PHARMACOGENETICS AND PHARMACOGENOMICS IN LATIN AMERICA: ETHNIC VARIABILITY, NEW INSIGHTS IN ADVANCES AND PERSPECTIVES: A RELIVAF-CYTED INITIATIVE

EDITED BY: Luis Abel Quiñones, Nelson Miguel Varela, Matias F. Martinez, Maria Ana Redal, Alberto Lazarowski, Andrés López-Cortés and Patricia Esperon

PUBLISHED IN: Frontiers in Pharmacology and Frontiers in Genetics

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88974-415-2
DOI 10.3389/978-2-88974-415-2

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

PHARMACOGENETICS AND PHARMACOGENOMICS IN LATIN AMERICA: ETHNIC VARIABILITY, NEW INSIGHTS IN ADVANCES AND PERSPECTIVES: A RELIVAF-CYTED INITIATIVE

Topic Editors:
Luis Abel Quiñones, University of Chile, Chile
Nelson Miguel Varela, University of Chile, Chile
Matias F. Martinez, University of Chile, Chile
Maria Ana Redal, University of Buenos Aires, Argentina
Alberto Lazarowski, University of Buenos Aires, Argentina
Andrés López-Cortés, Paradise One Health Genomics, Ecuador
Patricia Esperon, Universidad de la República, Uruguay

Citation: Quiñones, L. A., Varela, N. M., Martinez, M. F., Redal, M. A., Lazarowski, A., López-Cortés, A., Esperon, P., eds. (2022). Pharmacogenetics and Pharmacogenomics in Latin America: Ethnic Variability, New Insights in Advances and Perspectives: A RELIVAF-CYTED Initiative. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-415-2

Table of Contents

frontiers in Pharmacology

EDITORIAL
published: 11 January 2022
doi: 10.3389/fphar.2021.833000

Editorial: Pharmacogenetics and Pharmacogenomics in Latin America: Ethnic Variability, New Insights in Advances and Perspectives: A RELIVAF-CYTED Initiative

*Patricia Esperón[1,2], Matías F. Martínez[2,3], María A. Redal[2,4], Alberto Lazarowski[2,5], Andrés López-Cortés[2], Nelson M. Varela[2,3] and Luis A. Quiñones[2,3]**

[1]Molecular Genetics Laboratory, Clinical Biochemistry Department, School of Chemistry, Universidad de la República, Montevideo, Uruguay, [2]Latin American Network for the Implementation and Validation of Clinical Pharmacogenomics Guidelines (RELIVAF-CYTED), Madrid, Spain, [3]Laboratory of Chemical Carcinogenesis and Pharmacogenetics, Department of Basic-Clinical Oncology (DOBC), Faculty of Medicine, University of Chile, Santiago, Chile, [4]Genetic Division, Department of Medicine, INFIBIOC, Hospital de Clínicas José de San Martín, Buenos Aires University, Buenos Aires, Argentina, [5]Clinical Biochemistry Department, Hospital de Clínicas José de San Martín School of Pharmacy and Biochemistry, Institute for Research in Physiopathology and Clinical Biochemistry (INFIBIOC), University of Buenos Aires, Buenos Aires, Argentina

Keywords: pharmacogenetics, pharmacogenomics, CYTED, Latin America, precision medicine

Editorial on the Research Topic

Pharmacogenetics and Pharmacogenomics in Latin America: Ethnic Variability, New Insights in Advances and Perspectives: A RELIVAF-CYTED Initiative

OPEN ACCESS

Edited and reviewed by:
José A. G. Agúndez, University of Extremadura, Spain

****Correspondence:***
Luis A. Quiñones
lquinone@uchile.cl

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *10 December 2021*
Accepted: *23 December 2021*
Published: *11 January 2022*

Citation:
Esperón P, Martínez MF, Redal MA, Lazarowski A, López-Cortés A, Varela NM and Quiñones LA (2022) Editorial: Pharmacogenetics and Pharmacogenomics in Latin America: Ethnic Variability, New Insights in Advances and Perspectives: A RELIVAF-CYTED Initiative. Front. Pharmacol. 12:833000. doi: 10.3389/fphar.2021.833000

INTRODUCTION

Since the pharmacogenetics and pharmacogenomics (PGx) field started to rise, the information about the relationship between actionable genes, genotypes, and response to drugs has increased exponentially (Nicholson et al., 2021). There is evidence of the utility and impact of genetics in the choice of therapeutic regimens improving their effectiveness and safety (Arbitrio et al., 2021). Even some international efforts have created clinical guidelines that allow to implementation of pharmacogenomics in daily clinical practice. In addition to clinical outcomes, Economic benefits have been associated with the translation from "the bench to the bedside".

Moreover, several major PGx expert organizations such as the Clinical Pharmacogenetics Implementation Consortium (CPIC, 2021) and the Dutch Pharmacogenetics Working Group (DPWG), provide gene-drug guidelines for actionable variants. In addition, Ubiquitous Pharmacogenomics (U-PGx, 2021), the Latin American Network for the Implementation and Validation of Pharmacogenomics Guidelines (RELIVAF-CYTED, 2021), and the Southeast Asian Pharmacogenomics Research Network (SEAPharm, Chumnumwat et al., 2019) have investigated pharmacotherapeutic recommendations guided by pharmacogenetics. In this respect, based on scientific evidence the Food and Drug Administration (FDA) has published a list of PGx biomarkers for drug labelling (FDA, 2021).

Even though high-quality research addresses the utility of implementing pharmacogenetics programs in clinical practice, most of this evidence comes from the United States or Europe. Moreover, commonly, it does not include the Latin American population, or when the guidelines do, it is considering as one big group. Some recently regional formed scientific societies (RELAGH,

2014; SOLFAGEM, 2021) and international efforts (RELIVAF-CYTED) are looking to shorten the region's gap of evidence and information. In this respect, Latin America is a vast region with some characteristics that do not allow easy implementation of research made in other settings (Quiñones et al., 2014). It is one of the most genetically diverse areas having frequencies or polymorphisms not found in other regions. There is a lack of high-quality Latin American population-focused research about the relationship between specific genes and drug response, and, also, there is a lack of knowledge of frequencies. Altogether, there are many disadvantages to implementing pharmacogenetics in clinical practice in Latin America.

Sixteenth articles are included in this issue, eleven original/experimental research, two brief research reports, one review article, one case report, and one opinion, covering different and complementary aspects of the pharmacogenomic research in this region.

Workflows of data-driven modeling and model-driven experimentation have led to the development of *in silico* algorithms including pharmacogenomics data of disease risk at the patient-population level (Wolkenhauer et al., 2014). In this Research Topic four predictive models based on pharmacogenomics have been developed to identify patients who were suitable for preventive genotyping. Although the models must be validated with a larger number of patients and do not necessarily apply to all populations, they are a very good first approximation to predict the incidence of adverse effects among patients undergoing different therapies in Latin America. Miranda et al. proposed a model that included genetic polymorphisms in addition to clinical records, to predict the tamoxifen response in a population of 162 ER + breast cancer patients (Serral et al.) through an *in silico* approach, explored the druggable genes of two bacterial pathogens with a relevant impact in Latin America. The model proposed by Varela et al. was based on an integrated *in silico* analysis of breast and prostate cancer data genes. Finally, the algorithm created by Martinez et al. intended to predict the incidence of infections among patients under chemotherapy treatment.

The PGx of the immunosuppressive drug Tacrolimus (TAC) has been extensively studied, and according to the CPIC guidelines (Birdwell et al., 2015) an increase of starting dose for CYP3A5 expressers is recommended, followed by a therapeutic drug monitoring to guide dose adjustments. Thus, two manuscripts address the issue in Chilean kidney transplantation patients, one in children (Krall et al.), the other in an adult population (Contreras-Castillo et al.), for immunosuppressive treatment (cyclosporine and tacrolimus) after transplantation.

The antiretroviral treatment (ART) is generally not well tolerated and most patients present important adverse effects (ADR) that potentially limit treatment adherence or lead to this interruption (Saag et al., 2020). Poblete et al. retrospectively evaluated the UGT1A1*28 and CYP2B6 c.516G > T frequency and their influence on major ADRs in 67 adult HIV patients from Chile, as a starting point to validate in the nearest future CPIC guidelines in Latin America.

Two investigations referred to children with acute lymphoblastic leukemia from different angles. From Mexico, Gándara-Mireles et al. analyzed the frequency distribution and the association between the illness and the most common polymorphisms in ABCC1, NCF4, and CBR3 genes. The influence of TPMT-VNTR polymorphism on 6-MP related hematological toxicity was confirmed by Burgueño-Rodríguez et al. in 130 Uruguayan pediatric patients.

The studies performed in Duchenne Muscular Dystrophy (Luce et al.), cardiovascular disease (Gálvez et al.), and severe encephalopathy patients (Kravetz et al.) emphasized the importance of identifying both already known and novel variants to differential diagnosis and patient management. Luce et al. described the mutational spectrum of DMD gene in 400 Argentinian patients with a clinical diagnosis of dystrophinopathy. Gálvez et al. reported a significant association between APOB, APOE, and MTHFR polymorphisms and lipid levels, especially in women, among 193 healthy subjects from Chile. Identifying a genetic variant in KCNT1 channel in an Argentinian pediatric patient with a severe encephalopathy was crucial to include quinidine in the treatment regimen as an antiepileptic drug (Kravetz et al.).

Since discovering the non-coding RNA, its clinical relevance has become increasingly important. In particular, inter-individual variability in drug response, both in efficacy and toxicity, could be due to both, variation in miRNA gene sequences and circulating miRNA levels (Latini et al., 2019; Ubilla et al.) found an increase in miRNA-33b-5p levels in hypercholesterolemic patients under atorvastatin therapy and proposed this microRNA as a biomarker to follow the response to statins. Ruiz et al., worried about BCR-ABL1 tyrosine kinase inhibitor resistance in chronic myeloid leukemia patients, observed a global decrease of microRNA levels in resistant cells, founding a promiser field for future studies.

On other hand, the biobanks allow access to many well-classified, high-quality samples and establish the indispensable conditions for achieving reproducible research results (Coppola et al., 2019). Vargas and Cobar, express their opinion about creating biobanks and believe that the same requirements will be necessary to obtain pharmacogenetics information and efficient therapeutic responses in Latin America.

Barriers to PGx implementation include a lack of knowledge, training, and confidence among physicians to apply pharmacogenomic tests (Rigter et al., 2020). As such, Undurraga et al. reported on an anonymous online survey addressed to psychiatrists from Chile, observing a low acceptance of PGx tests, but a clear interest from psychiatrists in their potential incorporation into their clinical practice.

We proudly present this research topic which aims to address high-quality pharmacogenetic and pharmacogenomic research with a particular focus on the Latin American population and its needs. The main goal is to increase the information on the clinical implementation and the impact of pharmacogenetics in Latin American patients. Also, collecting experience and project the field in the region, looking for strategies and new perspectives. Furthermore, to potentiate the research among countries in the region.

AUTHOR CONTRIBUTIONS

PE: The conception of the Research Topic idea, writing the manuscript. MM: The conception of the Research Topic idea,

writing the manuscript. MR: The conception of the Research Topic idea, writing the manuscript. AL: The conception of the Research Topic idea, writing the manuscript. AL-C: The conception of the Research Topic idea, writing the manuscript. NV: The conception of the Research Topic idea, writing the manuscript. LQ: The conception of the Research Topic idea, writing the manuscript.

ACKNOWLEDGMENTS

We would like to thank all the authors for their contributions to this topic, to the Latin American Society for Pharmacogenomics and Personalized Medicine (SOLFAGEM) and the Latin American Network for Human Genetics (RELAGH) by their contribution and CYTED by sponsoring this article.

REFERENCES

Arbitrio, M., Scionti, F., Di Martino, M. T., Caracciolo, D., Pensabene, L., Tassone, P., et al. (2021). Pharmacogenomics Biomarker Discovery and Validation for Translation in Clinical Practice. *Clin. Transl Sci.* 14 (1), 113–119. Epub 2020 Oct 22. PMID: 33089968; PMCID: PMC7877857. doi:10.1111/cts.12869

Birdwell, K. A., Decker, B., Barbarino, J. M., Peterson, J. F., Stein, C. M., Sadee, W., et al. (2015). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for CYP3A5 Genotype and Tacrolimus Dosing. *Clin. Pharmacol. Ther.* 98 (1), 19–24. doi:10.1002/cpt.113

Chumnumwat, S., Lu, Z. H., Sukasem, C., Winther, M. D., Capule, F. R., Abdul Hamid, A. A. A. T., et al. (2019). Southeast Asian Pharmacogenomics Research Network (SEAPharm): Current Status and Perspectives. *Public Health Genomics.* 22 (3-4), 132–139. doi:10.1159/000502916

Coppola, L., Cianflone, A., Grimaldi, A. M., Incoronato, M., Bevilacqua, P., Messina, F., et al. (2019). Biobanking in Health Care: Evolution and Future Directions. *J. Transl Med.* 17 (1), 172. doi:10.1186/s12967-019-1922-3

CPIC (2021). Available at: https://cpicpgx.org/genes-drugs/(Accessed December 06, 2021).

FDA (2021). Available at: https://www.fda.gov/drugs/science-and-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling (Accessed December 06, 2021).

Gálvez, A. S., Ramírez, H., Placencia, P., Rojas, C., Urzúa, X., Kalergis, A. M., et al. (2021). Single Nucleotide Polymorphisms in Apolipoprotein B, Apolipoprotein E, and Methylenetetrahydrofolate Reductase Are Associated With Serum Lipid Levels in Northern Chilean Subjects. A Pilot Study. *Front. Genet.* 12, 640956. doi:10.3389/fgene.2021.640956

Latini, A., Borgiani, P., Novelli, G., and Ciccacci, C. (2019). miRNAs in Drug Response Variability: Potential Utility as Biomarkers for Personalized Medicine. *Pharmacogenomics.* 20, 1049–1059. doi:10.2217/pgs-2019-0089

Nicholson, W. T., Formea, C. M., Matey, E. T., Wright, J. A., Giri, J., and Moyer, A. M. (2021). Considerations When Applying Pharmacogenomics to Your Practice. *Mayo Clin. Proc.* 96 (1), 218–230. doi:10.1016/j.mayocp.2020.03.011

Quiñones, L. A., Lavanderos, M. A., Cayún, J. P., García-Martin, E., Agúndez, J. A., Cáceres, D. D., et al. (2014). Perception of the Usefulness of Drug/gene Pairs and Barriers for Pharmacogenomics in Latin America. *Curr. Drug Metab.* 15 (2), 202–208. doi:10.2174/1389200215666140202220753

RELAGH (2014). Latin American Network for Human Genetics: Rojas-Martínez A, Giraldo-Ríos A, Jiménez-Arce G, de Vargas AF, Giugliani R. RELAGH - The challenge of having a scientific network in Latin America: An account from the presidents. *Genet. Mol. Biol.* 37 (1 Suppl. l), 305–309. doi:10.1590/s1415-47572014000200017

RELIVAF-CYTED (2021). Available at: http://www.cyted.org/es/relivaf (Accessed December 06, 2021).

Rigter, T., Jansen, M. E., de Groot, J. M., Janssen, S. W. J., Rodenburg, W., and Cornel, M. C. (2020). Implementation of Pharmacogenetics in Primary Care: A Multi-Stakeholder Perspective. *Front. Genet.* 11 (10), 10. doi:10.3389/fgene.2020.00010

Saag, M. S., Gandhi, R. T., Hoy, J. F., Landovitz, R. J., Thompson, M. A., Sax, P. E., et al. (2020). Antiretroviral Drugs for Treatment and Prevention of HIV Infection in Adults. *JAMA.* 324 (16), 1651–1669. doi:10.1001/jama.2020.17025

SOLFAGEM (2021). Latin American Society for Pharmacogenomics and Personalized Medicine. Available at: www.solfagem.org (Accessed December 06, 2021).

U-PGx (2021). Available at: http://upgx.eu/(Accessed December 06, 2021).

Wolkenhauer, O., Auffray, C., Brass, O., Clairambault, J., Deutsch, A., Drasdo, D., et al. (2014). Enabling Multiscale Modeling in Systems Medicine. *Genome Med.* 6 (3), 21. doi:10.1186/gm538

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

BRIEF RESEARCH REPORT
published: 09 November 2020
doi: 10.3389/fphar.2020.594262

Ancestry and TPMT-VNTR Polymorphism: Relationship with Hematological Toxicity in Uruguayan Patients with Acute Lymphoblastic Leukemia

Gabriela Burgueño-Rodríguez[1], Yessika Méndez[2], Natalia Olano[2], Agustín Dabezies[2], Bernardo Bertoni[3], Jorge Souto[3], Luis Castillo[2], Julio da Luz[1] and Ana María Soler[1]**

[1]Laboratorio de Genética Molecular Humana, Centro Universitario Regional (CENUR) Litoral Norte-Sede Salto, Universidad de la República (UdelaR), Salto, Uruguay, [2]Servicio Hemato Oncológico Pediátrico (SHOP), Centro Hospitalario Pereira Rossell (CHPR), Montevideo, Uruguay, [3]Departamento de Genética, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay

OPEN ACCESS

Edited by:
Andrés López-Cortés,
Universidad UTE, Ecuador

Reviewed by:
Yang Shu,
Sichuan University, China
Lucia Taja-Chayeb,
National Institute of Cancerology (INCAN), Mexico

***Correspondence:**
Ana María Soler
anamasoler@gmail.com
Julio da Luz
jdal@fmed.edu.uy

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: 12 August 2020
Accepted: 24 September 2020
Published: 09 November 2020

Citation:
Burgueño-Rodríguez G, Méndez Y, Olano N, Dabezies A, Bertoni B, Souto J, Castillo L, da Luz J and Soler AM (2020) Ancestry and TPMT-VNTR Polymorphism: Relationship with Hematological Toxicity in Uruguayan Patients with Acute Lymphoblastic Leukemia. Front. Pharmacol. 11:594262. doi: 10.3389/fphar.2020.594262

6-Mercaptopurine (6-MP) is a thiopurine drug widely used in childhood acute lymphoblastic leukemia (ALL) therapy. Genes such as *TPMT* and *NUDT15* have an outstanding role in 6-MP metabolism. Mutations in both genes explain a significant portion of hematological toxicities suffered by ALL Uruguayan pediatric patients. A variable number tandem repeat in the *TPMT* promoter (*TPMT*-VNTR) has been associated with *TPMT* expression. This VNTR has a conservative architecture (AnBmC). To explore new causes of hematological toxicities related to ALL therapy, we genotyped the *TPMT*-VNTR of 130 Uruguayan pediatric patients. Additionally, individual genetic ancestry was estimated by 45 ancestry-informative markers (AIMs). Hematological toxicity was measured as the number of leukopenia events and 6-MP dose along the maintenance phase. As previously reported, we found *TPMT*2* and *TPMT*3C* alleles were associated to *TPMT*-VNTR A2BC and AB2C, respectively. However, contrasting with other reports, *TPMT*3A* allele was found in a heterogeneous genetic background in linkage equilibrium. Patients carrying more than 5 A repeats present a significant higher number of leukopenia events among patients without *TPMT* and/or *NUDT15* variants. Native American ancestry and the number of A repeats were significantly correlated with the number of leukopenia events. However, the correlation between Native American ancestry and the number of leukopenia events was lost when the number of A repeats was considered as covariate. This suggests that *TPMT*-VNTR alleles are more relevant than Native American ancestry in the hematological toxicity. Our results emphasize that *TPMT*-VNTR may be used as a pharmacogenetic biomarker to predict 6-MP-related hematological toxicity in ALL childhood therapy.

Keywords: Acute Lymphoblastic Leukemia, TPMT, TPMT-VNTR, NUDT15, 6-MP, Pharmacogenomics, Ancestry, Hematological toxicity

INTRODUCTION

Childhood acute lymphoblastic leukemia (ALL) is the most frequent children cancer worldwide. Uruguay is not an exception. The 5-year disease-free survival (DFS) in this country is similar to that observed in developed countries (Castillos et al., 2012). Although ALL treatments have been improved, approximately 20% of the patients suffer a diversity of adverse side effects due to the nonspecific action and the narrow therapeutic range of the drugs.

In Uruguay, Berlin-Frankfurt-Münster (IC-BFM) protocol is used to treat ALL pediatric patients. In this protocol, patients are classified as standard, intermediate, and high risk according to age of diagnosis, white blood cell (WBC) count, blast number in peripheral blood at day 8, blast percentage in the bone marrow at days 15 and 33, and the presence of translocations. The overall chemotherapy consists in a two-year (104 weeks) treatment encompassing many phases. According to the risk group, the maintenance phase varies between 63 and 74 weeks. In this phase, patients received 50 mg/m^2/day of 6-mercaptopurine (6-MP) and 20 mg/m^2/week of methotrexate (MTX). During the maintenance phase, complete blood count (CBC) is performed weekly along the first two months and then every two or three weeks.

Thiopurine S-methyltransferase (TPMT) is a cytosolic enzyme, in which physiological function remains uncertain. TPMT catalyzes the *S*-methylation of thiopurine drugs such as azathioprine, 6-MP, and 6-thioguanine (6-TG), which are widely used in autoimmune, inflammatory, and cancer diseases (Spire-Vayron de la Moureyre et al., 1998; Appell et al., 2013; Chouchana et al., 2014).

Normally, 6-MP and 6-TG are considered as purines and are metabolized to thioguanine nucleotides (TGNs) which will be incorporated into DNA, causing damage and triggering further cellular apoptosis (Pan and Nelson, 1990; Kotur et al., 2012). In the hematopoietic tissue, TPMT has an outstanding role in the downregulation of TGN formation due to its S-methyltransferase activity. There is therefore an inverse relationship between the activity of TPMT and active TGN metabolites' concentration (Relling et al., 2011).

TPMT illustrates one of the most characteristic examples of pharmacogenetics utility, attempting to a personalized drug therapy. To date, more than 44 *TPMT* mutant alleles have been described (https://www.hmv.liu.se/tpmtalleles), but *TPMT*2*, *TPMT*3A*, and *TPMT*3C* alleles represent more than 95% of inherited TPMT deficiency (Ameyaw et al., 1999; Chrzanowska et al., 2012). Furthermore, within the *TPMT* gene promoter, there is a variable number tandem repeat (VNTR) region (Spire-Vayron de la Moureyre et al., 1998; Spire-Vayron de la Moureyre et al., 1999), whose composition and number of repeats have been associated with *TPMT* expression (Zukic et al., 2010; Kotur et al., 2012; Kotur et al., 2015). This GC-rich VNTR region (*TPMT*-VNTR) has a three-element architecture, two of them being variable in their copy number (AmBnC). According to Vergnaud and Denoeud (2000) criteria, this VNTR is considered a minisatellite, and, hitherto, 19 different alleles have been described (Urbančič et al., 2018).

It has already been reported that, in Uruguayan pediatric patients under ALL therapy, approximately 30% of the hematological toxicities are explained by SNPs at *TPMT* and *NUDT15* genes (Soler et al., 2018). This result demonstrates once again the importance of specific studies for each population, especially in admixed ones such as Uruguayan population. The aim of this work is to analyze the *TPMT*-VNTR in ALL pediatric patients from Uruguay and determine its association with 6-MP hematological toxicity (measured as the number of leukopenia events and 6-MP cumulative dose) in the maintenance phase; taking individual genetic ancestry into account.

METHODOLOGY

Patients and Clinical Data

DNA samples of 130 ALL Uruguayan pediatric patients aged between 1 and 15 years were analyzed. All patients were followed up at the Servicio Hemato Oncológico Pediátrico—Centro Hospitalario Pereira Rossell, Montevideo, Uruguay, and treated with the IC-BFM protocol (58 according to ALL-IC-BFM 2002 and 72 according to ALL-IC-BFM 2009). The collected clinical data were 6-MP dose and the number of leukopenia events along the first 32 weeks of the maintenance phase. The cumulative 6-MP dose was calculated at weeks 8, 16, 24, and 32. Additionally, we calculated the mean weekly 6-MP dose (mg/m^2). On the contrary, the hematological toxicity (leukopenia events' grades 3 and 4) was analyzed as a whole and within eight-week intervals (1–8, 9–16, 17–24, and 25–32 weeks). The data analyzed in this study were collected blinded to genotypes from the patients' medical files by the two corresponding researchers.

The protocol and procedure employed in this study was approved by the CENUR Litoral Norte, Universidad de la República, institutional ethics committee, and informed consent was obtained from parents, guardians, and/or patients, as required (Exp. N° 311170-000332-17, www.expe.edu.uy).

TPMT, *TPMT*-VNTR, and *NUDT15* Genotyping

Genomic DNA was isolated from peripheral blood leukocytes by the saline extraction method (Miller et al., 1988). Patients were previously genotyped for *TPMT* and *NUDT15* genes which are the most common variants (Soler et al., 2018). The number and type of tandem repeats were determined by *TPMT* first-exon PCR amplification, including the promoter region, followed by Sanger sequencing. Furthermore, in order to discard other possible mutations, the *TPMT* gene was completely characterized by the amplification and sequencing of its nine exons. In addition, since *NUDT15* variants analyzed by Soler et al. (2018) were in exons 1 and 3, we also sequenced *NUDT15* exon 2. Amplification primers and PCR conditions are shown in **Supplementary Table S1**.

Linkage Disequilibrium

TPMT-VNTR genotypes and *TPMT* genotypic data previously reported by Soler et al. (2018) were used to estimate the gametic phase, applying Arlequin software v3.5.2 (Excoffier and Lischer, 2010). For each haplotype, we estimated the normalized linkage disequilibrium (D').

TABLE 1 | *TPMT*-VNTR genotype frequencies and *TPMT* genotype.

***TPMT*-VNTR**		***n***	**Freq. (%)**	***TPMT* (*n*)**	**Group***
3/3	ABC/ABC	1	0.77	*1*/*3A*	1
3/4a	ABC/A2BC	1	0.77	**1/*1*	1
3/5a	ABC/A2B2C	1	0.77	**1/*1*	1
4a/4a	A2BC/A2BC	33	25.38	**1/*1* (29) **1/*2* (1) **1/*3A* (3)	1
4a/4b	A2BC/AB2C	4	3.08	**1/*1*	1
4a/5a	A2BC/A2B2C	44	33.85	**1/*1* (43) **1/*3A* (1)	1
4a/5b	A2BC/A3BC	1	0.77	**1/*1*	2
4a/6a	A2BC/A2B3C	11	8.46	**1/*1*	1
4a/6b	A2BC/AB4C	2	1.54	**1/*1*	1
4a/6c	A2BC/A4BC	2	1.54	**1/*1*	2
4a/7a	A2BC/A5BC	3	2.31	**1/*1*	2
4b/5a	AB2C/A2B2C	2	1.54	**1/*1* (1) **1/*3C* (1)	1
5a/5a	A2B2C/A2B2C	15	11.54	**1/*1* (13) **1/*3A* (2)	1
5a/5b	A2B2C/A3BC	1	0.77	**1/*1*	2
5a/6a	A2B2C/A2B3C	5	3.85	**1/*1*	1
5a/7a	A2B2C/A5BC	2	1.54	**1/*1*	2
6b/8a	AB4C/A6BC	1	0.77	**1/*3A*	2
7a/7a	A5BC/A5BC	1	0.77	**1/*3A*	2
Total	—	**130**	—	—	—

Freq.: TPMT-*VNTR genotype frequency. *1): number of A repeats < 5; 2): number of A repeats ≥ 5.*

Ancestry

A total of 45 ancestry-informative markers (AIMs) located along all autosomes were genotyped (**Supplementary Table S2**). These AIMs were selected from the SNP panel published by Yaeger et al. (2008). Nineteen of them were performed by the SNaPshot multiplex system (Thermo Fisher Scientific, Waltham, Massachusetts, United States) and 26 by MassARRAY SNP genotyping (Agena Bioscience Inc., San Diego, United States). In order to estimate the European, Amerindian, and African ancestry, Structure software (Pritchard et al., 2000) was used with 100,000 iterations for the burn-in period and 1,000,000 additional iterations. The parental populations included 42 Europeans (Coriell's North American panel), 37 West Africans (nonadmixed Africans living in London, United Kingdom, and South Carolina, United States), and 30 Native Americans (15 Mayans and 15 Nahuas), who were genotyped on an Affymetrix 100 K SNP chip (data were kindly provided by Dr. Laura Fejerman (University of California, San Francisco)).

Statistical Analysis

Allele and genotype frequencies were estimated by gene counting. Correlation between clinical data and the number of repeats (A, B, and A + B) was analyzed by Spearman correlation test. Additionally, patients were classified into two groups according to the number of A repeats (A < 5 and A ≥ 5). The association between these two groups with 6-MP cumulative dose and the number of leukopenia events was analyzed by variance analyses (ANOVA) and Mann–Whitney test, respectively. Furthermore, association between individual ancestry, clinical data, and genetic information was analyzed by ANOVA, Spearman, and partial correlation. Classification tree was built to cluster the patients according to the number of leukopenia events, using the chi-squared automatic interaction detection (CHAID) algorithm. The categorical variables considered were the presence or absence of *TPMT* and/or *NUDT15* variants, A repeats' groups, and Native American ancestry (cutoff 15%). All analyzes were carried out in the total sample and in the sample subdivided by the presence or absence of *TPMT* and *NUDT15* variants. Statistical analysis was performed in R statistical package with a CI of 95%, with exception of the classification tree which was carried on SPSS 22.0.

RESULTS

TPMT-VNTR Genotypic and Allelic Frequencies

TPMT and *NUDT15* exon sequencing did not reveal any other mutation to those previously reported by Soler et al. (2018). Regarding *TPMT*-VNTR, we identified 18 different genotypes and 10 different alleles. The total number of VNTR repeats (A + B + C) ranged from 6 (ABC/ABC) to 14 (A5BC/A5BC) (**Table 1**).

The most frequent *TPMT*-VNTR genotypes were *4a/*5a (33.85%), *4a/*4a (25.38%), *5a/*5a (11.54%), and *4a/*6a (8.46%). Therefore, *TPMT*-VNTR alleles *4a (51.54%), *5a (32.59%), and *6a (6.15%) explain more than 90% of variability (**Table 1** and **2**).

Six different *TPMT*-VNTR genotypes (*3/*3, *4a/*4a, *4a/*5a, *5a/*5a, *6b/*8a, and *7a/*7a) were found in the nine heterozygous *TPMT*3A/*1* patients. The *TPMT*2/*1* and *TPMT*3C/*1* patients have the *4a/*4a and *4b/*5a *TPMT*-VNTR genotypes, respectively (**Table 1**).

Estimation of the gametic phase showed the *TPMT*3A* allele associated to five *TPMT*-VNTR alleles (*3, *4a, *5a, *7a, and *8a), whereas *TPMT*2* and *TPMT*3C* alleles were associated to *TPMT*-VNTR alleles *4a and *4b, respectively. With exception

TABLE 2 | Linkage disequilibrium between *TPMT*-VNTR alleles and *TPMT* alleles obtained from estimated haplotype combinations.

***TPMT*-VNTR alleles**		**Freq. (%)**	***TPMT* allele**	***n***	**D′**	r^2	X^2	***P***
*3	ABC	1,54	**1*	3	-0.217	0.017	4,325	0.038
			**3A*	1	0.223	0.022	5,640	0.018
*4a	A2BC	51,54	**1*	130	0.118	0.001	0.170	0.680
			**2*	1	1,000	0.004	0.944	0.331
			**3A*	3	-0.138	0.001	0.188	0.665
*4b	AB2C	2,31	**1*	5	-0.130	0.009	2,344	0.126
			**3C*	1	1,000	0.163	42,497	**<0.0001**
*5a	A2B2C	32,69	**1*	82	0.444	0.004	1,099	0.295
			**3A*	3	-0.320	0.002	0.464	0.496
*5b	A3BC	0,77	**1*	2	1,000	0.000	0.089	0.765
*6a	A2B3C	6,15	**1*	16	1,000	0.003	0.753	0.386
*6b	AB4C	1,15	**1*	3	1,000	0.001	0.134	0.714
*6c	A4BC	0,77	**1*	2	1,000	0.000	0.089	0.765
*7a	A5BC	2,69	**1*	6	-0.105	0.007	1,795	0.180
			**3A*	1	0.112	0.010	2,522	0.112
*8a	A6BC	0,38	**3A*	1	-1,000	0.087	22,724	**<0.0001**

Freq.: TPMT-*VNTR allele frequency. D′: normalized linkage disequilibrium coefficient.* r^2*: correlation coefficient.* X^2*: chi-square.* p = p *value.*

TABLE 3 | Association between the number of leukopenia events and 1) the number of A, B, and A + B repeats and 2) risk groups.

Interval (weeks)		Spearman correlation (1)						Mann–Whitney (2)				
		Number of A repeats		Number of B repeats		A + B repeats		Number of leukopenia events				*p*
		R	*p*	*R*	*p*	*r*	*P*	Risk group 1		Risk group 2		
								N	Mean	*N*	Mean	
1–8	Total	0.245	***0.008***	-0.122	0.116	-0.007	0.474	88	0.61	8	1.63	***0.010***
	wt							72	0.54	6	1.50	***0.044***
	mut							16	0.94	2	2	0.105
9–16	Total	0.112	0.140	0.056	0.294	0.106	0.151	88	0.85	8	1.62	0.105
	wt							72	0.71	6	1.17	0.328
	mut							16	1.50	2	3	0.163
17–24	Total	0.166	0.054	-0.170	0.050	-0.070	0.250	87	0.48	8	1.25	**0.025**
	wt							71	0.41	6	1.17	0.068
	mut							16	0.81	2	1.50	0.216
25–32	Total	0.137	0.093	0.099	0.172	0.126	0.114	86	0.34	8	0.63	0.154
	wt							71	0.28	6	0.67	0.148
	mut							15	0.60	2	0.50	0.669
Total (32 weeks)	Total	0.240	***0.009***	-0.075	0.233	0.044	0.336	88	2.27	8	5.13	***0.002***
	wt							72	1.93	6	4.50	***0.012***
	mut							16	3.81	2	7	0.078

1) Association between the number of leukopenia events and the number of A, B, and A + B repeats was done by Spearman test. 2) The number of leukopenia events between risk groups was compared by Mann–Whitney test. Total: total sample; wt: *patients without TPMT or NUDT15 variants; mut: patients with TPMT and/or NUDT15 variants. r: Spearman correlation coefficient.* p: p *value.* N*: number of individuals. Risk group 1: A < 5. Risk group 2: A ≥ 5. Mean: mean number of leukopenia events' grade 3 or 4.*

of haplotypes *3-*TPMT*1*, *3-*TPMT*3A*, *4a-*TPMT*3C*, and *8a-*TPMT*3A*, the remaining haplotypes were found in linkage equilibrium (**Table 2**).

Association of *TPMT*-VNTR with Hematological Toxicity

We found a significant correlation between the number of A repeats and the number of leukopenia events in the interval 1–8 ($p = 0.008$) and when the 32 weeks were analyzed as a whole ($p = 0.009$) (**Table 3**). Graphical representation of total leukopenia events at week 32 showed a trend toward a greater number of leukopenia events when the number of A repeats was equal or higher than five (**Supplementary Figure S1**). Therefore, patients were classified into two groups according to the number of A repeats: risk group 1 with less than five A repeats and risk group 2 with five or more A repeats. When the number of leukopenia events was compared among the risk groups by the Mann–Whitney test, the number of leukopenia events was significantly greater in risk group 2 than that in risk group 1 at intervals 1–8 and 17–24 and considering the 32 weeks as a whole (**Table 3**).

As the number of leukopenia events is strongly associated with *TPMT* and *NUDT15* variants (Relling et al., 2011; Cargnin et al., 2018; Soler et al., 2018), we also subdivided the sample according to the presence or absence of *TPMT* and/or *NUDT15* variants. The association between risk groups and the number of leukopenia events remains significant only in the subsample without *TPMT* and/or *NUDT15* variants. This association was observed in the interval 1–8 and in the total number of leukopenia events at week 32 (**Table 3**).

We did not detect significant correlations between 6-MP cumulative dose and the *TPMT*-VNTR length (A + B) as well as with the number of A or B repeats separately (**Supplementary Table S3**). Moreover, 6-MP cumulative dose did not show an association with risk groups by ANOVA neither in the total sample nor in the sample subdivided by the presence and absence of *TPMT* and/or *NUDT15* variants (data not shown).

Ancestry, Hematological Toxicity, and *TPMT*-VNTR

Ancestry was determined in 111 of 130 patients. European, Native American, and African ancestral proportions were 73.9 ± 14.8%, 16.9 ± 11.7%, and 9.12 ± 7.4% (mean ± SD), respectively (**Supplementary Figure S2**).

There is a significant negative correlation between the number of leukopenia events and Native American ancestry in patients without *TPMT* and/or *NUDT15* variants (**Supplementary Table S4**). Moreover, risk group 2 patients showed a significantly higher European and lower Native American ancestry than patients from risk group 1 among patients without *TPMT* and/or *NUDT15* variants (**Supplementary Figures 2C,D**). In order to analyze the relevance of the abovementioned variables, we additionally performed a partial correlation analysis between the number of leukopenia events and patients' Native American ancestry using risk groups as covariate. This analysis did not show a significant correlation ($p = 0.120$). Furthermore, classification tree first divided the sample by the presence or absence of *TPMT* and/or *NUDT15* variants and then by risk groups resulting in three groups (two nodes). Native American ancestry was not a significant variable to explain the number of leukopenia events (**Figure 1**).

6-MP cumulative dose does not show a significant correlation with European, Native American, or African ancestry in the total sample.

DISCUSSION

As it has been widely reported, *TPMT* and *NUDT15* genes have an outstanding role in the 6-MP metabolic pathway. Variants in both genes induce loss of their enzymatic activity and the subsequent accumulation of TGNs, resulting in the presence of adverse drug reaction due to 6-MP (Fabre et al., 2004; Yang et al., 2014; Moriyama et al., 2016; Soler et al., 2018; Cao et al., 2020). However, patients without *TPMT* and/or *NUDT15* variants also suffer 6-MP adverse effects. A VNTR region in the *TPMT* promoter has been postulated to modulate TPMT activity (Spire-Vayron de la Moureyre et al., 1998) and hence may explain part of the hematological toxicity due to 6-MP.

In our study population, an admixed one, overall *TPMT*-VNTR allele frequencies did not differ significantly from the previously reported for Portuguese and Mozambique (Alves et al., 2002), Asian British (Marinaki et al., 2003), and Serbian populations (Zukic et al., 2010; Kotur et al., 2012). The most frequent alleles found were *4a and *5a followed by *6a, which were similar to those observed in other populations (Alves et al., 2002; Marinaki et al., 2003; Zukic et al., 2010; Urbančič et al., 2018). Interestingly, alleles *3 and *4b were found at higher frequencies than those observed in European populations and similar to those reported for Mozambique and Asian British populations, respectively (Alves et al., 2002; Marinaki et al., 2003; Urbančič et al., 2018). This fact could be explained by the contribution of Native American and African populations to the Uruguayan genetic pool (Sans et al., 1997; Buchelli and Cabella., 2010; Bonilla et al., 2015).

Genotypic and linkage disequilibrium (LD) analyses showed that, with exception to the *8a allele, the functional *TPMT*1* allele is linked to the rest of the *TPMT*-VNTR alleles. In concordance with the previous reports, *TPMT*2* and *TPMT**3C alleles were found linked to *TPMT*-VNTR *4a and *4b alleles, respectively (Alves et al., 2001; Marinaki et al., 2003; Urbančič et al., 2018).

On the contrary, *TPMT*3A* (the most common deficient allele) was linked to six different *TPMT*-VNTR alleles. Six out of nine *TPMT*3A* alleles were linked in equilibrium to the most frequent *TPMT*-VNTR alleles (*4a and *5a), similar to what Spire-Vayron de la Moureyre et al. (1998) reported for a European population. Moreover, the *7a allele (the fourth most frequent *TPMT*-VNTR allele) was also found in linkage equilibrium with the *TPMT*3A* allele. Different to previous studies, we did not observe linkage between *TPMT*3A* and *6b (Alves et al., 2001; Marinaki et al., 2003). Additionally, *TPMT*3A* allele was in LD only with *TPMT*-VNTR alleles found in low frequencies (*3 and *8a). This result contrasts with previous studies, where *TPMT*3A* was found in LD with *5a, *6a, and ABnC (n > 2) alleles (Yan et al., 2000; Alves et al., 2001; Marinaki et al., 2003; Urbančič et al., 2018) (**Table 2**).

The linkage between *TPMT*-VNTR with the variable number of A repeats and *TPMT*3A* allele, added to the absence of LD when all alleles were considered (p = 0.198), does not adjust to the evolutionary model proposed by Urbančič et al. (2018). These differences may be explained by several reasons: 1) the loss of LD by recombination between the *TPMT*3A* allele and the common *TPMT*-VNTR alleles as suggested by Marinaki et al. (2003); 2) the heterogeneity in the Caucasic population that contributes to the Uruguayan genetic pool; and 3) microevolutionary events as genetic drift or founder effects occurring in the Uruguayan population.

Several studies had reported that *TPMT*-VNTR modulates the expression of the *TPMT* gene, possibly at the transcriptional level. Some authors had reported an inverse correlation between the overall repeat number of *TPMT*-VNTR and TPMT activity (Spire-Vayron de la Moureyre et al., 1998; Yan et al., 2000; Fabre et al., 2004), whereas other authors found a correlation between the number of A or B repeats with TPMT activity and/or with hematological toxicities (Alves et al., 2001; Zukic et al., 2010; Kotur et al., 2012; Kotur et al., 2015). Most of these studies had analyzed the relationship between *TPMT*-VNTR alleles and *TPMT* expression/activity at the molecular or biochemical level. In our study, we focus on the leukopenia events and on 6-MP dose, two indirect clinical measures of TPMT expression/activity.

We found a negative correlation between the number of A repeats and the number of leukopenia events (**Table 3**). These results agree with *TPMT* gene expression data previously reported by Zukic et al. (2010) and Kotur et al. (2012), Kotur et al. (2015). Interestingly, we did not find any significant correlation between either the total overall repeat number or the number of B repeats with the two clinical measures analyzed, as described by Spire-Vayron de la Moureyre et al. (1998) and Alves et al. (2001). Other study, with kidney transplant recipients receiving azathioprine, reported that patients with *TPMT*-VNTR genotypes composed of more than ten repeats showed a significant reduction in azathioprine dose compared to those with ten or less repeats (Fabre et al., 2004). Although this study does not discriminate between the number of A and B repeats, those individuals (>10 repeats) must have at least one *TPMT*-VNTR allele carrying more than six repeats. These alleles may show a bias toward a higher number of A than B repeats according to reported *TPMT*-VNTR allele frequencies (Urbančič et al., 2018).

According to Kotur et al. (2015), carriers of the *7a allele had the lowest average of *TPMT* gene expression level and the least increase in its expression level during chemotherapy. Interestingly, six of our 11 patients, belonging to risk group 2 (A ≥ 5), carry the *7a allele. Unfortunately, there are no other studies about *TPMT* expression in patients carrying the other *TPMT*-VNTR alleles belonging to risk group 2 (*5b, *6c, and *8a).

The relevance of genetic ancestry in 6-MP therapy has been demonstrated by several examples such as hematological toxicity observed in patients carrying NUDT15 variants, which is more frequent in Asian and South American populations (Yang et al., 2014; Soler et al., 2018). Global ancestry proportions were similar to those observed in previous studies and confirmed the trihybrid structure of the Uruguayan population (Sans et al., 1997; Sans et al., 2006; Bonilla et al., 2015). At the individual level, the observed wide variation was also similar to the previous report (Bonilla et al., 2015) (**Supplementary Figure S2A**).

Even though we found a significant negative correlation between the total number of leukopenia events and Native American ancestry in patients without *TPMT* and/or *NUDT15* variants (**Supplementary Table S4**), this correlation was no longer significant when risk groups were taken into consideration (p =

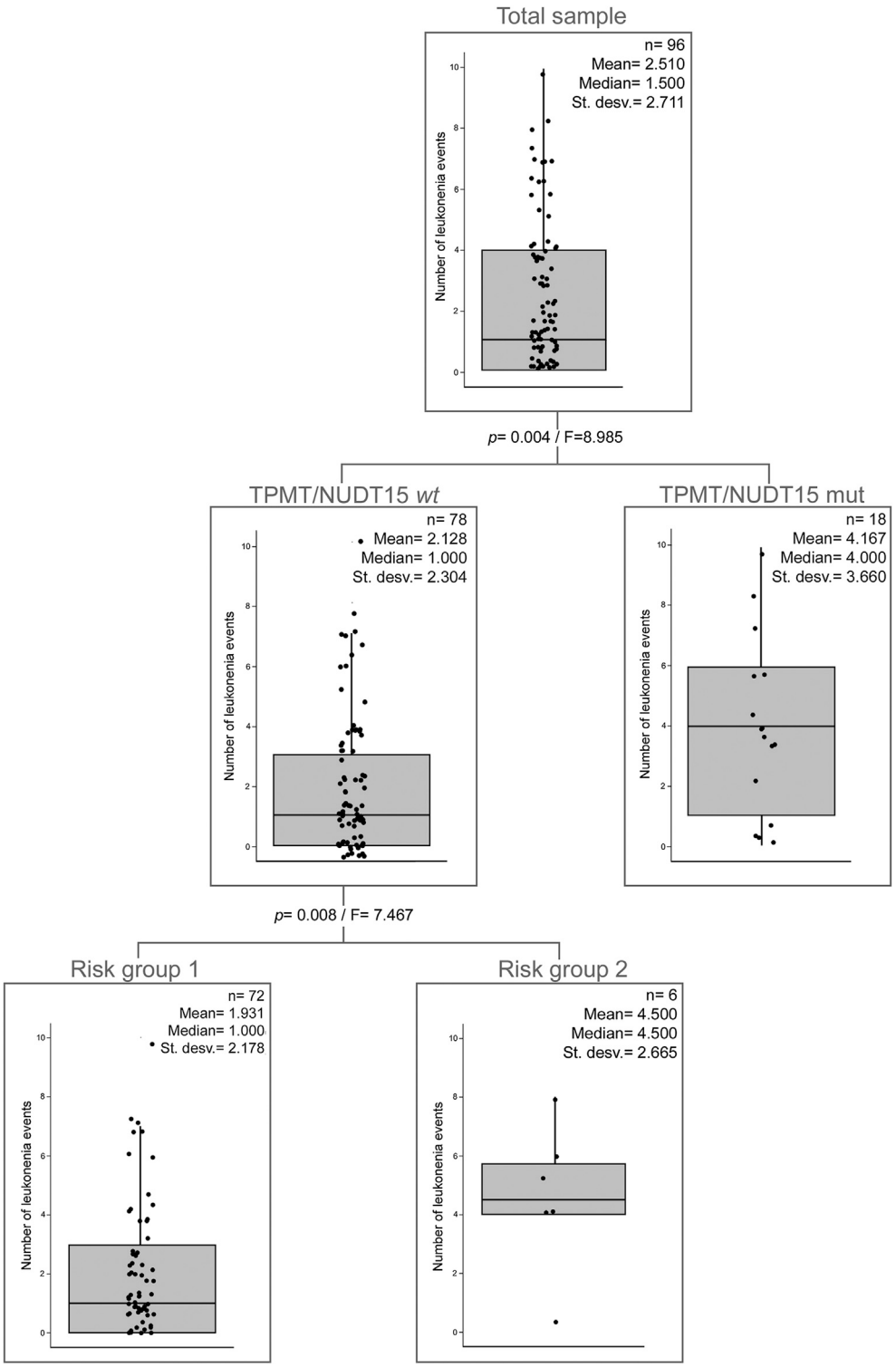


FIGURE 1 | Classification tree of the number of leukopenia events. Presence/absence of *TPMT* and/or *NUDT15* mutations, risk groups, and Native American ancestry (cutoff = 15%) were taken as categorical variables. *TPMT*/NUDT15wt: *TPMT*1* and *NUDT15*1*. *TPMT*/NUDT15mut: *TPMT* and/or *NUDT15* mutated. Risk group 1: A < 5. Risk group 2: A ≥ 5.

0.120). This may show that *TPMT*-VNTR alleles are more relevant than Native American ancestry. A heterogeneous distribution of *TPMT*-VNTR alleles among ancestral populations may explain the loss of the correlation between Native American ancestry and the total number of leukopenia events. However, although *TPMT*-VNTR frequencies in Asian and Native American populations are scarce, these data do not support this hypothesis (Marinaki et al., 2003; Urbančič et al., 2018). Another possible explanation is microevolutionary factors as genetic drift or founder effects, causing different levels of ancestry between risk groups.

Despite the significant association between the number of leukopenia events and risk group 2, as visualized at the classification tree and consistent with previous reports (Zukic et al., 2010; Kotur et al., 2012; Kotur et al., 2015), we must be cautious with this result due to the small sample size and the low number of patients with five or more A repeats analyzed in this study. Nevertheless, considering that Uruguay has only 3.5 million inhabitants and presents approximately 25 new ALL pediatric patients per year, a 130-individual sample represents more than 5 years of ALL patients in our country.

Furthermore, additionally to *NUDT15* and *TPMT*, there are other genes involved in 6-MP metabolism, like *ITPA* or *ABCC4,* which may influence in the hematological toxicity (Hareedy et al., 2015; Tanaka et al., 2015; Tanaka et al., 2018).

Although we found a clear association between risk group 2 and the number of leukopenia events, this association was not reflected in 6-MP dose. This could be due to a weaker effect of *TPMT*-VNTR compared with other genetic factors such as *TPMT* and *NUDT15* variants. Additionally, the decision to modify 6-MP dose is based on empirical information, and thus relative to the patient general state and to medical staff.

To our knowledge, this is the first report that studied the relationship between *TPMT*-VNTR alleles and clinical data, stratifying the sample by *TPMT* and/or *NUDT15* mutations.

In summary, we confirm the previous reported association between *4a and *4b *TPMT*-VNTR alleles with *TPMT**2* and *TPMT**3C,* respectively. However, we did not find the reported association between *TPMT*-VNTR pattern ABnC ($n \geq 2$) and *TPMT**3A*, suggesting a greater heterogeneity in the genetic background of patients carrying this variant. Moreover, a higher number of A repeats might be a risk factor for suffering leukopenia events. Finally, our results support the hypothesis that transcriptional genetic modifiers, such as the VNTR region in the *TPMT* promoter, may be used as a pharmacogenetic biomarker and could contribute to the design of personalized ALL childhood therapy.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité de Ética en Investigación Institucional, Centro Universitario Regional (CENUR) Litoral Norte, Universidad de la República (UdelaR). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

AS, JL, and GB-R designed the study, analyzed clinical and genetic data, and drafted the manuscript. BB and JS participated in the ancestry analyses. LC, AD, YM, and NO were responsible for patients' recruitment and helped in the analyses of clinical data.

FUNDING

Financial support from Centro Universitario Regional (CENUR) Litoral Norte, Universidad de la República (UdelaR) (2017-2019), and Comisión Sectorial de Investigación Científica-Vinculación Universidad Sociedad y Producción Modalidad 2, Universidad de la República (CSIC-VUSP-Mod. 2 2016, UdelaR).

ACKNOWLEDGMENTS

We are grateful to all the patients and parents who kindly participated in this study, as well as the Servicio Hemato Oncológico Pediátrico (SHOP)-Centro Hospitalario Pereira Rossell (CHPR) staff.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2020.594262/full#supplementary-material

REFERENCES

Alves, S., Amorim, A., Ferreira, F., and Prata, M. J. (2001). Influence of the variable number of tandem repeats located in the promoter region of the thiopurine methyltransferase gene on enzymatic activity. *Clin. Pharmacol. Ther.* 70 (2), 165–174. doi:10.1067/mcp.2001.117284

Alves, S., Amorim, A., and Prata, M. (2002). Evolution of a VNTR located within the promoter region of the thiopurine methyltransferase gene: inferences from population and sequence data. *Hum. Genet.* 111 (2), 172–178. doi:10.1007/s00439-002-0784-5

Ameyaw, M. M., Collie-Duguid, E. S., Powrie, R. H., Ofori-Adjei, D., and McLeod, H. L. (1999). Thiopurine methyltransferase alleles in British and Ghanaian populations. *Hum. Mol. Genet.* 8 (2), 367–370. doi:10.1093/hmg/8.2.367

Appell, M. L., Berg, J., Duley, J., Evans, W. E., Kennedy, M. A., Lennard, L., et al. (2013). Nomenclature for alleles of the thiopurine methyltransferase gene. *Pharmacogenetics Genom.* 23 (4), 242. doi:10.1097/FPC.0b013e32835f1cc0

Bonilla, C., Bertoni, B., Hidalgo, P. C., Artagaveytia, N., Ackermann, E., Barreto, I., et al. (2015). Breast cancer risk and genetic ancestry: a case–control study in Uruguay. *BMC Women's Health* 15 (1), 11. doi:10.1186/s12905-015-0171-8.

Bucheli, M., and Cabella, W. (2010). El perfil demográfico y socioeconómico de la población uruguaya según su ascendencia racial. Notas de población, Año XXXVII, N° 91, Naciones Unidas, Santiago de Chile.

Cao, M., Yin, D., Qin, Y., Liao, F., Su, Y., Xia, X., et al. (2020) Screening of novel pharmacogenetic candidates for mercaptopurine-induced toxicity in patients with acute lymphoblastic leukemia. *Front. Pharmacol.* 11, 267. doi:10.3389/fphar.2020.00267

Cargnin, S., Genazzani, A. A., Canonico, P. L., and Terrazzino, S. (2018). Diagnostic accuracy of NUDT15 gene variants for thiopurine-induced leukopenia: a systematic review and meta-analysis. *Pharmacol. Res.* 135, 102–111. doi:10.1016/j.phrs.2018.07.021

Castillo, L., Dabezies, A., Dufort, G., Carracedo, M., Castiglioni, M., Simón, E., et al. (2012). Evolucion del cáncer pediátrico en Uruguay (1992–2011). *Arch. Pediatría del Urug.* 83 (1), 26–30.

Chouchana, L., Narjoz, C., Roche, D., Golmard, J. L., Pineau, B., Chatellier, G., et al. (2014). Interindividual variability in TPMT enzyme activity: 10 years of experience with thiopurine pharmacogenetics and therapeutic drug monitoring. *Pharmacogenomics* 15 (6), 745–757. doi:10.2217/pgs.14.32

Chrzanowska, M., Kuehn, M., Januszkiewicz-Lewandowska, D., Kurzawski, M., and Droździk, M. (2012). Thiopurine S-methyltransferase phenotype-genotype correlation in children with acute lymphoblastic leukemia. *Acta Pol. Pharm.* 69 (3), 405–410.

Excoffier, L., and Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under linux and windows. *Mol. Ecol. Resour.* 10 (3), 564–567. doi:10.1111/j.1755-0998.2010.02847.x

Fabre, M. A., Jones, D. C., Bunce, M., Morris, P. J., Friend, P. J., Welsh, K. I., et al. (2004). The impact of thiopurine S-methyltransferase polymorphisms on azathioprine dose 1 year after renal transplantation. *Transpl. Int.* 17 (9), 531–539. doi:10.1111/j.1432-2277.2004.tb00483.x

Hareedy, M. S., El Desoky, E. S., Woillard, J. B., Thabet, R. H., Ali, A. M., Marquet, P., et al. (2015). Genetic variants in 6-mercaptopurine pathway as potential factors of hematological toxicity in acute lymphoblastic leukemia patients. *Pharmacogenomics* 16 (10), 1119–1134. doi:10.2217/pgs.15.62

Kotur, N., Dokmanovic, L., Janic, D., Stankovic, B., Krstovski, N., Tosic, N., et al. (2015). TPMT gene expression is increased during maintenance therapy in childhood acute lymphoblastic leukemia patients in a TPMT gene promoter variable number of tandem repeat-dependent manner. *Pharmacogenomics* 16 (15), 1701–1712. doi:10.2217/pgs.15.109

Kotur, N., Stankovic, B., Kassela, K., Georgitsi, M., Vicha, A., Leontari, I., et al. (2012). 6-mercaptopurine influences TPMT gene transcription in a TPMT gene promoter variable number of tandem repeats-dependent manner. *Pharmacogenomics* 13 (3), 283–295. doi:10.2217/pgs.11.153

Marinaki, A. M., Arenas, M., Khan, Z. H., Lewis, C. M., Shobowale-Bakre, E. M., Escuredo, E., et al. (2003). Genetic determinants of the thiopurine methyltransferase intermediate activity phenotype in British Asians and Caucasians. *Pharmacogenetics Genom.* 13 (2), 97–105. doi:10.1097/00008571-200302000-00006

Miller, S. A., Dykes, D. D., and Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16 (3), 1215. doi:10.1093/nar/16.3.1215

Moriyama, T., Nishii, R., Perez-Andreu, V., Yang, W., Klussmann, F. A., Zhao, X., et al. (2016). NUDT15 polymorphisms alter thiopurine metabolism and hematopoietic toxicity. *Nat. Genet.* 48 (4), 367–373. doi:10.1038/ng.3508

Pan, B. F., and Nelson, J. A. (1990). Characterization of the DNA damage in 6-thioguanine-treated cells. *Biochem. Pharmacol.* 40 (5), 1063–1069. doi:10.1016/0006-2952(90)90494-6

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155 (2), 945–959.

Relling, M. V., Gardner, E. E., Sandborn, W. J., Schmiegelow, K., Pui, C. H., Yee, S. W., et al. (2011). Clinical pharmacogenetics implementation consortium guidelines for thiopurine methyltransferase genotype and thiopurine dosing. *Clin. Pharmacol. Ther.* 89 (3), 387–391. doi:10.1038/clpt.2010.320

Sans, M., Merriwether, D. A., Hidalgo, P. C., Bentancor, N., Weimer, T. A., Franco, M. H. L., et al. (2006). Population structure and admixture in Cerro Largo, Uruguay, based on blood markers and mitochondrial DNA polymorphisms. *Am. J. Hum. Biol.* 18 (4), 513–524. doi:10.1002/ajhb.20520

Sans, M., Salzano, F. M., and Chakraborty, R. (1997). Historical genetics in Uruguay: estimates of biological origins and their problems. *Hum. Biol.* 69 (2), 161–170.

Soler, A. M., Olano, N., Méndez, Y., Lopes, A., Silveira, A., Dabezies, A., et al. (2018). TPMT and NUDT 15 genes are both related to mercaptopurine intolerance in acute lymphoblastic leukaemia patients from Uruguay. *Br. J. Haematol.* 181 (2), 252–255. doi:10.1111/bjh.14532

Spire-Vayron de la Moureyre, C., Debuysère, H., Fazio, F., Sergent, E., Bernard, C., Sabbagh, N., et al. (1999). Characterization of a variable number tandem repeat region in the thiopurine S-methyltransferase gene promoter. *Pharmacogenetics* 9 (2), 189–198.

Spire Vayron de la Moureyre, C., Debuysere, H., Mastain, B., Vinner, E., Marez, D., Lo Guidice, J. M., et al. (1998). Genotypic and phenotypic analysis of the polymorphic thiopurine S-methyltransferase gene (TPMT) in a European population. *Br. J. Pharmacol.* 125 (4), 879–887. doi:10.1038/sj.bjp.0702152

Tanaka, Y., Manabe, A., Fukushima, H., Suzuki, R., Nakadate, H., Kondoh, K., et al. (2015). Multidrug resistance protein 4 (MRP4) polymorphisms impact the 6-mercaptopurine dose tolerance during maintenance therapy in Japanese childhood acute lymphoblastic leukemia. *Pharmacogenomics J.* 15 (4), 380–384. doi:10.1038/tpj.2014.74

Tanaka, Y., Nakadate, H., Kondoh, K., Nakamura, K., Koh, K., and Manabe, A. (2018). Interaction between NUDT15 and ABCC4 variants enhances intolerability of 6-mercaptopurine in Japanese patients with childhood acute lymphoblastic leukemia. *Pharmacogenomics J.* 18 (2), 275–280. doi:10.1038/tpj.2017.12

Urbančič, D., Šmid, A., Stocco, G., Decorti, G., Mlinarič-Raščan, I., and Karas Kuželički, N. (2018). Novel motif of variable number of tandem repeats in TPMT promoter region and evolutionary association of variable number of tandem repeats with TPMT* 3 alleles. *Pharmacogenomics* 19 (17), 1311–1322. doi:10.2217/pgs-2018-0123

Vergnaud, G., and Denoeud, F. (2000). Minisatellites: mutability and genome architecture. *Genome Res.* 10 (7), 899–907. doi:10.1101/gr.10.7.899

Yaeger, R., Avila-Bront, A., Abdul, K., Nolan, P. C., Grann, V. R., Birchette, M. G., et al. (2008). Comparing genetic ancestry and self-described race in African Americans born in the United States and in Africa. *Cancer Epidemiol.* 17 (6), 1329–1338. doi:10.1158/1055-9965.EPI-07-2505

Yan, L., Zhang, S., Eiff, B., Szumlanski, C. L., Powers, M., O'Brien, J. F., et al. (2000). Thiopurine methyltransferase polymorphic tandem repeat: genotype-phenotype correlation analysis. *Clin. Pharmacol. Ther.* 68 (2), 210–219. doi:10.1067/mcp.2000.108674

Yang, S. K., Hong, M., Baek, J., Choi, H., Zhao, W., Jung, Y., et al. (2014). A common missense variant in NUDT15 confers susceptibility to thiopurine-induced leukopenia. *Nat. Genet.* 46 (9), 1017. doi:10.1038/ng.3060

Zukic, B., Radmilovic, M., Stojiljkovic, M., Tosic, N., Pourfarzad, F., and Dokmanovic, L., , et al. (2010). Functional analysis of the role of the TPMT gene promoter VNTR polymorphism in TPMT gene transcription. *Pharmacogenomics* 11 (4), 547–557. doi:10.2217/pgs.10.7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

frontiers in Pharmacology

ORIGINAL REASEARCH
published: 11 January 2021
doi: 10.3389/fphar.2020.612573

miRNome profiling of LSC-enriched CD34$^+$CD38$^-$CD26$^+$ fraction in Ph$^+$ CML-CP samples from Argentinean patients: a potential new pharmacogenomic tool

María Sol Ruiz[1], María Belén Sánchez[1], Simone Bonecker[2], Carolina Furtado[3], Daniel Koile[4], Patricio Yankilevich[4], Santiago Cranco[5], María del Rosario Custidiano[5], Josefina Freitas[6], Beatriz Moiraghi[7], Mariel Ana Pérez[8], Carolina Pavlovsky[9], Ana Inés Varela[7], Verónica Ventriglia[6], Julio César Sánchez Ávalos[5], Irene Larripa[10], Ilana Zalcberg[2], José Mordoh[11,12,13], Peter Valent[14,15] and Michele Bianchini[1]**

[1]Centro de Investigaciones Oncológicas—Fundación Cáncer, Ciudad Autónoma de Buenos Aires, Argentina, [2]Centro de Transplante de Medula Ôssea, Instituto Nacional de Câncer, Rio de Janeiro, Brazil, [3]Programa de Genética, Instituto Nacional de Câncer, Rio de Janeiro, Brazil, [4]Instituto de Investigación en Biomedicina de Buenos Aires (IBioBA)-CONICET-Partner Institute of the Max Planck Society, Ciudad Autónoma de Buenos Aires, Argentina, [5]Instituto Alexander Fleming, Ciudad Autónoma de Buenos Aires, Argentina, [6]Hospital Nacional Posadas, Buenos Aires, Argentina, [7]Hospital J. M. Ramos Mejía, Ciudad Autónoma de Buenos Aires, Argentina, [8]Hospital Interzonal General de Agudos, Buenos Aires, Argentina, [9]Fundaleu, Ciudad Autónoma de Buenos Aires, Argentina, [10]Instituto de Medicina Experimental, CONICET/Academia Nacional de Medicina, Ciudad Autónoma de Buenos Aires, Argentina, [11]Centro de Investigaciones Oncológicas, Fundación Cáncer, Buenos, Aires, Argentina, [12]IIBBA-CONICET, Fundación Instituto Leloir, Buenos, Aires, Argentina, [13]Instituto Alexander Fleming, Buenos, Aires, Argentina, [14]Division of Hematology and Hemostaseology, Department of Internal Medicine I, Medical University of Vienna, Vienna, Austria, [15]Ludwig Boltzmann Institute for Hematology and Oncology, Medical University of Vienna, Vienna, Austria

OPEN ACCESS

Edited by:
Alberto Lazarowski, University of Buenos Aires, Argentina

Reviewed by:
Danilo Perrotti, University of Maryland, Baltimore, United States
Jaume Folch Lopez, University of Rovira i Virgili, Spain

***Correspondence:**
María Sol Ruiz
ma.sol.ruiz@gmail.com
Michele Bianchini
mbianchini@conicet.gov.ar

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *30 September 2020*
Accepted: *30 November 2020*
Published: *11 January 2021*

Citation:
Ruiz MS, Sánchez MB, Bonecker S, Furtado C, Koile D, Yankilevich P, Cranco S, Custidiano MdelR, Freitas J, Moiraghi B, Pérez MA, Pavlovsky C, Varela AI, Ventriglia V, Sánchez Ávalos JC, Larripa I, Zalcberg I, Mordoh J, Valent P and Bianchini M (2021) miRNome profiling of LSC-enriched CD34$^+$CD38$^-$CD26$^+$ fraction in Ph$^+$ CML-CP samples from Argentinean patients: a potential new pharmacogenomic tool. Front. Pharmacol. 11:612573. doi: 10.3389/fphar.2020.612573

Chronic myeloid leukemia (CML) is a myeloid stem cell neoplasm characterized by an expansion of myeloid progenitor cells and the presence of BCR-ABL1 oncoprotein. Since the introduction of specific BCR-ABL1 tyrosine kinase inhibitors (TKI), overall survival has improved significantly. However, under long-term therapy patients may have residual disease that originates from TKI-resistant leukemic stem cells (LSC). In this work, we analyzed the miRNome of LSC-enriched CD34$^+$CD38$^-$CD26$^+$ and normal hematopoietic stem cells (HSC) fractions obtained from the same chronic phase (CP) CML patients, and stem and progenitor cells obtained from healthy donors (HD) by next-generation sequencing. We detected a global decrease of microRNA levels in LSC-enriched CD34$^+$CD38$^-$CD26$^+$ and HSC fractions from CML-CP patients, and decreased levels of microRNAs and snoRNAs from a genomic cluster in chromosome 14, suggesting a mechanism of silencing of multiple non-coding RNAs. Surprisingly, HSC from CML-CP patients, despite the absence of *BCR-ABL1* expression, showed an altered miRNome. We

Abbreviations: BFU-E, burst-forming unit-erythroid; BM, bone marrow; CFU-GEMM, colony-forming unit-granulocyte erythroid macrophage megakaryocyte; CFU-GM, colony-forming unit-granulocyte macrophage; CML, chronic myeloid leukemia; CP, chronic phase; FACS, fluorescent-activated cell sorting; FDA, U.S. Food and Drug Administration; GO, gene ontology; HD, healthy donor; HSC, hematopoietic stem cells; KEGG, kyoto encyclopedia of genes and genomes; LSC, leukemic stem cells; LT-HSC, (murine) long-term repopulating hematopoietic stem cell; MNC, mononuclear cells; NGS, next-generation sequencing; PB, peripheral blood; Ph, Philadelphia chromosome; qPCR, quantitative PCR; RT, reverse transcription; snoRNAs, small nucleolar RNAs; snRNAs, small nuclear RNAs; TKI, tyrosine kinase inhibitors; tRNAs, transfer RNAs.

confirmed by RT-qPCR that the levels of miR-196a-5p were increased more than nine-fold in $CD26^+$ (*BCR-ABL1*$^+$) vs. $CD26^-$ (*BCR-ABL1*$^-$) $CD34^+CD38^-$ fractions from CML-CP patients at diagnosis, and *in silico* analysis revealed a significant association to lipid metabolism and hematopoiesis functions. In the light of recent descriptions of increased oxidative metabolism in CML LSC-enriched fractions, these results serve as a guide for future functional studies that evaluate the role of microRNAs in this process. Metabolic vulnerabilities in LSCs open the road for new therapeutic strategies. This is the first report of the miRNome of CML-CP $CD34^+CD38^-$ fractions that distinguishes between $CD26^+$ (*BCR-ABL1*$^+$) and their $CD26^-$ (*BCR-ABL1*$^-$) counterparts, providing valuable data for future studies.

Keywords: microRNAs, metabolism, leukemic stem cell, leukemia, CD26

INTRODUCTION

Chronic myeloid leukemia (CML) originates from a hematopoietic stem cell (HSC) that acquires the reciprocal translocation t(9;22)(q34;q11) and thus the Philadelphia chromosome (Ph) (Rowley, 1973). The resulting fusion gene, *BCR-ABL1*, encodes an oncogenic protein with constitutive tyrosine kinase activity. Treatment of CML patients was revolutionized by the introduction of specific tyrosine kinase inhibitors (TKI), like imatinib, nilotinib or dasatinib. These TKIs effectively induce apoptosis in leukemic cells in patients with CML in chronic phase (CP) (Druker et al., 1996). However, the response of patients to TKI treatment is heterogeneous, and about 40% of imatinib-treated patients require a switch of TKI due to intolerance or resistance to treatment (Holyoake and Vetrie, 2017). Other patients with optimal response to TKI show persistence of the leukemic clone, even after several years of treatment (Chomel et al., 2011). A subset of TKI-treated CML patients can achieve a deep molecular response during therapy (Holyoake and Vetrie, 2017). However, only half of them or even less can sustain a treatment-free remission (Mahon et al., 2010; Etienne et al., 2017; Ross et al., 2018; Saussele et al., 2018).

Leukemic stem cells (LSC) are defined as a population of cells that gives rise and maintains the leukemic clone (Bonnet and Dick, 1997; Valent et al., 2012). The classical view of CML considers that LSC derive from the acquisition of BCR-ABL1 in a HSC (Nguyen et al., 2012). However, BCR-ABL1 alone is unable to induce a leukemia (Foley et al., 2013). Rather, additional molecular lesions and hits are required for full transformation of clonal pre-leukemic (stem) cells into fully malignant leukemic (stem) cells (Valent et al., 2012; Valent et al., 2013). Correspondingly, single-cell gene expression analysis revealed great heterogeneity within LSC populations (Giustacchini et al., 2017; Warfvinge et al., 2017). Moreover, the most primitive LSC population has been described as quiescent $CD34^+CFSE^{max}$, $CD34^+CD38^-CD90^+CD93^+$ or $Lin^-CD34^+CD38^{-/low}CD45RA^-cKIT^-CD26^+$ cells (Holyoake et al., 1999; Neviani et al., 2013; Warfvinge et al., 2017; Kinstrie et al., 2020). Their normal counterparts, HSC, also constitute a heterogeneous population, and individual HSC exhibit differences in properties related to their stem cell nature: self-renewal, quiescence, repopulation capacity, and differentiation potential (Mckenzie et al., 2006; Valent et al., 2013). The mechanisms underlying the regulation of such properties are not completely understood; however, they depend on both intrinsic (such as the levels of specific transcription factors) and extrinsic (such as signals coming from the bone marrow niche) factors (Mckenzie et al., 2006; Nakamura-Ishizu et al., 2014). In CML, most LSC and their subclones may be sensitive to TKI therapy. However, certain stem cell classes, especially pre-leukemic neoplastic stem cells may be resistant because they are slowly cycling cells and exhibit multiple forms of stem cell resistance (Valent et al., 2012; Valent et al., 2013). TKI-resistance of CML LSC has been associated to both cell-autonomous (Corbin et al., 2011; Kumari et al., 2012) and extrinsic factors (Arrigoni et al., 2018; Zhang et al., 2018; Silvestri et al., 2020). Sometimes even LSC may survive TKI therapy and thus persist in CML patients. The persistence of LSC in patients under TKI therapy has fueled intensive research on this topic, in order to identify novel therapeutic targets that enable the complete eradication of the leukemic clone in all patients (Vetrie et al., 2020). On the other hand, it is not clear whether the heterogeneity observed in the LSC population is related to different responses to TKI treatment.

In CML, recent reports have characterized the transcriptome (Bruns et al., 2009; Giustacchini et al., 2017; Kinstrie et al., 2020), protein networks (Abraham et al., 2016), and the metabolome of stem/progenitor fractions in CML (Kuntz et al., 2017). Gene expression profiling of the $CD34^+CD38^-$ fraction (which includes quiescent but mostly non-quiescent LSC) in CML patients revealed a transcriptional profile resembling normal $CD34^+$ myeloid progenitor cells, with decreased levels of transcription factors involved in maintenance of stem-cell fate, suggesting loss of quiescence (Bruns et al., 2009). Single-cell RNA sequencing revealed an enrichment of gene sets related to mechanistic target of rapamycin kinase (MTOR), targets of E2F transcription factors, G2/M checkpoints, oxidative phosphorylation, and glycolysis-associated gene expression in *BCR-ABL1*$^+$ stem cells (Giustacchini et al., 2017). However, little is known about microRNA-mediated regulation of gene expression in this population. MicroRNAs are small (19–25 nt), non-coding RNAs that can regulate multiple targets, mainly by mRNA destabilization or inhibition of protein translation. They are evolutionary conserved and have shown to be relevant for multiple physiological and pathological processes (Calin and Croce, 2006). One report has shown the involvement of

microRNAs in TKI sensitivity in CML LSC (Salati et al., 2017). Recent evidence has shown that miR-300 is a tumor suppressor microRNA able to induce quiescence in CML LSCs (Silvestri et al., 2020), and that miR-126-3p regulates quiescence, self-renewal and engraftment capacity of CML LSCs (Zhang et al., 2018). Advances in the characterization of aberrant expression of surface markers have allowed the prospective isolation of LSC and HSC from CML patients (Herrmann et al., 2014; Warfvinge et al., 2017). In this work, we characterized the miRNome of LSC-enriched $CD34^+CD38^-CD26^+$ fraction (*BCR-ABL1*$^+$) and its *BCR-ABL1*$^-$ counterpart ($CD34^+CD38^-CD26^-$ fraction, defined as "CML-CP HSC" in this article) isolated by fluorescence-activated cell sorting (FACS) from CML-CP patients at diagnosis by small RNA-Next-Generation Sequencing (NGS), in order to identify differential molecular mechanisms that contribute to unravel LSC biology, and the possible therapeutic implications of such differences. We observed a global decrease in microRNA levels in LSC-enriched and HSC fractions from CML-CP patients in comparison with HSC obtained from healthy donors (HD). Surprisingly, compared to HSC from HD, we detected decreased levels in the LSC-enriched fraction of microRNAs and snoRNAs belonging to a genomic cluster located in chromosome 14 (14q32) that contains imprinted genes, suggesting an epigenetic mechanism of silencing of multiple non-coding RNAs. Bioinformatic analysis of microRNAs with altered levels in LSC revealed a significant association with lipid metabolism and hematopoiesis. Finally, we confirmed an increase in the levels of miR-196a-5p in LSC-enriched $CD34^+CD38^-CD26^+$ fraction by reverse transcription followed by quantitative PCR (RT-qPCR) in additional CML-CP patients. Our results suggest that some microRNAs may act as mediators of the dysregulated metabolism observed in CML stem/progenitor fractions, and opens an exciting pathway for future research.

MATERIALS AND METHODS

Patient Samples

The project was approved by the Institutional Review Board, at Instituto Alexander Fleming (Buenos Aires, Argentina). All procedures involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments. All patients and HD gave written informed consent. Bone marrow (BM) or peripheral blood (PB) samples were obtained from newly diagnosed, untreated CML-CP patients. Patient samples used for library preparation for small RNA-NGS and validation by RT-qPCR are listed in **Table 1**. Mononuclear cells (MNC) were isolated by density-gradient centrifugation (Ficoll-Paque PLUS, GE Healthcare Life Sciences) for 30 min at 400 × *g*, followed by one wash in phosphate buffered saline (PBS, GIBCO), a red cells lysis step, and a low-speed centrifugation step (12–15 min at 200 × *g*) for removal of the platelet-rich fraction. Up to 2×10^8 MNC were used for isolation of $CD34^+$ cells.

TABLE 1 | CML-CP and HD samples used for small RNA-NGS and validation by RT-qPCR.

Code	Type	Sex	Age	No. events in HSC fraction	No. events in LSC-enriched fraction	No. events in progenitor fraction
CML-CP samples used for small RNA-NGS (pooled)						
N26	BM	M	24	6,896	1,857	ND
N33	PB	M	54	1,859	0	ND
IM	BM	M	22	NE*	3,046	ND
HD samples used for small RNA-NGS (pooled)						
2891	Buffy coat	M	36	1,578	N/A	7,624
2890	Buffy coat	F	31	1,316	N/A	13,724
2810	PB	F	50	156	N/A	ND
3060	PB	F	54	181	N/A	ND
3308	Buffy coat	F	37	1,700	N/A	13,000
CML-CP samples used for validation by RT-qPCR						
N36	PB	M	56	2,857	2,200	37,272
N47	PB	M	58	1,128	4,504	96,333
N55	PB	M	25	NE*	9,276	19,861
N38	PB	M	31	592	0	2,300
N46	PB	F	38	NE*	3,360	61,428
N56	PB	M	55	4,267	2,910	291,653
HD samples used for validation by RT-qPCR						
3984	Buffy coat	F	36	1,117	N/A	1,779
3308	Buffy coat	F	37	1,682	N/A	2,027
2771	Buffy coat	M	53	4,460	N/A	9,539
4532	Buffy coat	F	57	2,314	N/A	13,256

NE, fractions not evaluated (pattern 3); ND, not done; N/A, not applicable.*
The number of events refers to the number of sorted cells obtained from each fraction.

Isolation of $CD34^+$ Cells

In order to enrich for stem and progenitor cells, we performed a positive selection using CD34 MicroBeads (Miltenyi Biotech), according to manufacturer's instructions. The $CD34^+$ fraction was immediately used or cryopreserved in 1 ml of freezing medium (**Supplementary Material**).

CFU Assay for Assessment of Purity in Sorted Fractions

Between 250 and 500 $CD34^+$ cells were directly sorted into 250 µl of enriched methylcellulose (Methocult H4435 Medium, Stem Cell Technologies, Vancouver, Canada), and then plated into p35 culture dishes containing a final volume of 1.1 ml of enriched methylcellulose. Cells were incubated at 37°C in a humid chamber. After 14–18 days, pools of four to six colonies (CFU-GM, BFU-E and mixed CFU-GEMM) were plucked from methylcellulose, resuspended in 500 µl of Roswell Park Memorial Institute—1640 medium (RPMI-1640, GIBCO), and centrifuged. Cells were resuspended in 100 µl of lysis solution (RNAqueous-Micro Kit, Ambion), and kept at −20°C until RNA extraction was performed. Total RNA was extracted following manufacturer's instructions, and *BCR-ABL1* mRNA was measured by RT-qPCR (**Supplementary Material**).

Isolation of LSC and HSC by FACS

Total number of cells used for FACS varied according to the yield of each sample. $CD34^+$ cells or MNC from CML-CP patients or

HD were incubated with the following antibodies: 5 µl CD45-PerCP (2D1, BD Biosciences), 2.5 µl CD34-FITC (AC136, Miltenyi Biotech), 2.5 µl CD38-PE (IB6, Miltenyi Biotech), and 15 µl CD26-APC (FR10-11G9, Miltenyi Biotech), in a final volume of 100 µl of MACS buffer, for 15 min at room temperature. Cells were washed once with 1 ml of PBS (GIBCO) and resuspended in 300 µl of PBS. Sorting was performed in a FACS Aria II cytometer (BD Biosciences), located at Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires (Buenos Aires, Argentina). Setting of positive and negative gates for CD38, CD34, and CD26 was performed on the $CD45^{low}$ population; therefore, isotype control tube included 2.5 µl Mouse IgG2a-FITC (Miltenyi Biotech), 2.5 µl Mouse IgG2a-PE (Miltenyi Biotech), 15 µl Mouse IgG2a-FITC (Miltenyi Biotech), and 5 µl CD45-PerCP. In order to avoid electronic aborts that could affect the purity of sorted fractions, the parameter "window extension" was set to zero. Other parameters included 70 µm nozzle, and "purity." Cells were collected in aseptic conditions, directly into 100 µl of lysis buffer for RNA extraction (RNAqueous-Micro Kit, Ambion), in RNAse-free 200 µl tubes, or in enriched methylcellulose for assessment of purity. Flow-cytometry data analysis was performed with BD FACSDiva (version 6.1.3) and FlowJo (version 7.6.2) software.

Total RNA containing small RNAs (<200 nt) was extracted following the protocol from RNAaqueous-micro kit (Ambion) with a slight modification: 125 µl of EtOH 100% were added to the lysate and vortexed; the rest of the protocol was performed according to manufacturer's instructions. RNA elution was performed twice (9 µl each) using pre-warmed distilled water (75°C). RNA was kept at −80°C. Quality and conservation of the small RNA fraction were assessed with Agilent 2100 Bioanalyzer (total RNA Nano kit), at Fundación Instituto Leloir (Buenos Aires, Argentina).

Concentration of Pooled Samples for Small RNA-NGS

RNAs extracted from different samples were combined in order to increase RNA yield before NGS library preparation. After mixing, RNA was freezed at −80°C, and transported from Argentina to Brazil in dry ice. Samples were concentrated using a vacuum centrifuge and resuspended in 7 µl of distilled water (5 min at 50°C). Quantification of RNA was performed using Qubit 2.0. The entire content was used for library preparation (<140 ng for LSC-enriched, CML-CP HSC and HD HSC fractions, and 210 ng for HD progenitor fraction).

Preparation of Libraries for Small RNA-NGS in HiSeq 2500 (Illumina)

Libraries from each pool of samples (CML-CP LSC-enriched $CD34^+CD38^-CD26^+$, CML-CP HSC $CD34^+CD38^-CD26^-$, HD HSC $CD34^+CD38^{-/dim}$, HD progenitors $CD34^+CD38^+$) were prepared using Truseq Small RNA kit (Illumina), following manufacturer's instructions (15 PCR cycles). The protocol is based on the selective ligation of RNAs with free 3′OH and 5′phosphate ends, resulting from precursor cleavage during small RNA biogenesis (Bartel, 2018). Therefore, other small RNAs besides microRNAs are included in the library: fragments of tRNAs, small nucleolar RNAs (snoRNAs), small nuclear RNAs (snRNAs), and piwiRNAs. Estimated size of the libraries was 147–157 bp, which were purified by band excision after polyacrylamide denaturing gel electrophoresis (Novex 6% TBE, Invitrogen). Quantification of libraries was performed by qPCR (KAPA SYBR, Roche Life Sciences). Libraries were concentrated before sequencing by vacuum centrifugation. Single-end sequencing was performed on HiSeq 2500 (Illumina) at Instituto Nacional de Câncer (Rio de Janeiro, Brazil).

Bioinformatic Analysis of Small RNA-NGS

Low quality lectures were filtered (fastq_quality_filter; >80% reads with Q > 20), and contaminant sequences were removed (3′ and 5′ adaptors, indexes). Identification of known microRNAs was performed with Chimira (Vitsios and Enright, 2015); raw microRNA counts from each pool of samples is available (**Supplementary Data Sheet S2**). Differential expression analysis was performed using GFOLD algorithm (c = 0.01), which is especially suited for experiments without biological replicates, after mapping against a database of snoRNA/miRNA (HISAT2) (Feng et al., 2012). Complete results from GFOLD analysis are available (**Supplementary Data Sheet S1**). Traditional analysis of potential targets and related pathways was performed using miRPath (*Diana tools*) (Vlachos et al., 2015), and ChemiRs (Su et al., 2016). The parameters used for miRPath analysis were: "*KEGG analysis*," *Tarbase* (database of experimentally validated interactions), or *microT-CDS* in those cases with no experimental evidence, "*Pathway union*," "*p-value threshold:* 0.001," "*MicroT threshold*: 0.8," "*Enrichment analysis method = Fisher's exact test (hypergeometric distribution)*," "*FDR correction* (*Benjamini & Hochberg*)", "*Conservative stats*." Network analysis were performed with miRNet 2.0 (Chang et al., 2020) (query by lists of microRNAs; targets databases: Genes miRTarBase v8.0/ lncRNAs, degree cutoff: 1.1). Overrepresentation analysis were performed with TAM 2.0 (Li et al., 2018). Venn diagrams were created with BioVenn (Hulsen et al., 2008). Intersections between lists of microRNAs or targets were performed with R software (v.3.4.0).

Detection of microRNAs by RT-qPCR

We evaluated the following fractions from CML-CP patients or HD samples: CML-CP LSC-enriched $CD34^+CD38^-CD26^+$, CML-CP HSC $CD34^+CD38^-CD26^-$, CML progenitors $CD34^+CD38^+$, HD HSC $CD34^+CD38^{-/dim}$, HD progenitors $CD34^+CD38^+$. We applied a modification of the protocol reported by Chen et al., based on a reverse transcription (RT) using gene-specific stem-loop primers (Chen et al., 2005) (incubation times were modified as detailed below), followed by individual qPCR reactions for each microRNA using an intercalating agent. qPCR used a specific forward primer and a universal reverse primer designed to hybridize with the constant region included in the stem-loop primer. Two multiplex RT

reactions for microRNAs (M1 and M2) (**Supplementary Table S1**) were performed for each sample. Additionally, each sample was reverse transcribed with random primers in a separate reaction, in order to measure snRNA U6 as a reference gene for qPCR. Final concentrations of components of RT reaction were: dNTPs 0.25 mM (Invitrogen or INBIO Highway); DTT 10 mM (Invitrogen); Superscript II 2.5 U/μl (Invitrogen); RNAse inhibitor 0.2 U/μl (RNAseOUT, Invitrogen); *stem-loop* primer 0.05 μM (each) or random primers 0.01 μg/μl (Invitrogen). Incubation times were: 5 min of RNA, H_2O, and dNTPs at 65°C; tubes were immediately placed on ice; the remaining components were added to the tube and incubated for 30 min at 16°C, 40 cycles (30 s at 30°C + 30 s at 42°C + 1 s at 50°C), followed by a final step of 5 min at 85°C. cDNA was diluted (1/2) with distilled water and stored at −20°C. Two microliters of diluted cDNA was used for each qPCR reaction, using the following conditions: forward primer 0.3 μM; universal reverse primer 0.3 μM, and SYBR Green (PowerUp SYBR Green MasterMix, Applied Biosystems; according to the information provided by the manufacturer, Mg^{2+} concentration can vary between 4.76 and 6.44 mM); incubated for 2 min at 50°C, 2 min at 95°C, 50 cycles (15 s 95°C + 1 min at 60°C), in a Rotor-Gene Q qPCR equipment (Qiagen). Melting curves were evaluated in order to assess specificity of the reaction. Quantifications were performed in duplicate. In cases were duplicate measurements differed (ΔCt > 2), a triplicate measurement was performed. RT-qPCR efficiency was estimated by performing curves of RNA; formula used for efficiency estimation was $E = [10^{(-1/m)}]^{-1}$, m being the slope of the curve. Sequences of all primers used for quantification of microRNAs are available (**Supplementary Material**).

Statistical Analysis and Graphical Tools

GraphPad Prism 6, Microsoft Excel 2007, and Inkscape 0.92 software were used for graphics. Infostat v.2018e software (Córdoba, Argentina) was used for statistical analysis. Data from quantification of microRNAs by RT-qPCR were analyzed using the variable dCt = (Ct microRNA X − Ct snRNA U6), with a linear mixed-effects model (ANAVA): fraction (LSC-enriched, CML-CP HSC, CML-CP progenitors, HD HSC, HD progenitors) was set as a fixed effect, and sample (each patient or HD) was set as a random effect (correlation factor: compound symmetry). Variance was modeled using "VarIdent" function (using the variable "fraction"). False discovery rate was considered by multiplying p-values by the number of total microRNAs evaluated. *A posteriori* comparisons were performed using DCG formula (di Rienzo et al., 2011).

RESULTS

Global Patterns in the miRNome of LSC-Enriched and CML-CP HSC Fractions

We isolated highly pure LSC-enriched and HSC fractions from CML-CP patients at diagnosis or from HD, based on a combination of cell surface markers (CD34, CD38, CD45, CD26) and flow cytometry parameters (FSC, SSC) (**Supplementary Figure S1**). Some patients showed no clear separation of $CD26^-$ and $CD26^+$ populations (**Supplementary Figure S2**): in those cases, both fractions included leukemic *BCR-ABL1*$^+$ cells, therefore, we only used $CD26^+$ fraction. Purity was assessed by *BCR-ABL1* mRNA detection in CFU-derived colonies (**Supplementary Figure S2**). We extracted total RNA containing the small RNA fraction (<200 nt) from sorted cells; given that individual patient-derived fractions had low yields of RNA, samples from different patients or HD were pooled before preparation of libraries for small RNA-NGS (CML-CP LSC-enriched $CD34^+CD38^-CD26^+$, CML-CP HSC $CD34^+CD38^-CD26^-$, HD HSC $CD34^+CD38^{-/dim}$, HD progenitors $CD34^+CD38^+$). More than 1,000 (≥1 count) or 600 (≥10 counts) different microRNAs were detected in each fraction, with high abundance of a few specific microRNAs: top-10 most abundant microRNAs represented 57–69% of total microRNAs in CML-CP and HD (**Figure 1A**). Surprisingly, the pattern of most-abundant microRNAs showed more differences between fractions from CML-CP patients (LSC-enriched and CML-CP-HSC) than between LSC-enriched and HD-HSC fractions. Most (>80%) microRNAs dysregulated (GFOLD ≥ |1|) in LSC-enriched and HSC fractions from CML-CP patients had decreased levels compared to primitive ($CD34^+CD38^{-/dim}$) cells from HD, suggesting a global pattern of microRNA downregulation (**Figure 1B**).

microRNAs With Altered Levels in LSC-Enriched vs. HSC Fractions From CML-CP Patients and HD

Differential expression of microRNAs was assessed by calculation of a GFOLD value, which is a robust fold-change parameter that considers both the absolute number and the relative difference in microRNA levels between samples. With a cut-off value of GFOLD ≥ |1|, we found 120 microRNAs dysregulated between LSC-enriched and putative HSC fractions from CML-CP patients; and 46 microRNAs between LSC-enriched and HD-HSC fractions. The intersection of both lists resulted in 16 microRNAs (**Figure 2A**; **Supplementary Table S2**). Traditional *in silico* enrichment analysis of both predicted and experimentally validated targets showed a noteworthy proportion of false positives (**Supplementary Results, Supplementary Figures S3 and S4, Supplementary Table S3**).

Given the observed *bias*, and the difficulty of analyzing microRNA-related pathways through their mRNA targets due to the existence of "multiple-to-multiple relationships" (a given microRNA can regulate multiple genes, and a given gene can be regulated by multiple microRNAs), we performed a network analysis with the recently updated tool miRNet (v.2.0) (Chang et al., 2020). This tool allows the simultaneous inclusion of microRNA interactions with mRNAs, transcription factors (TF), lncRNAs, small non-coding RNAs and circular RNAs. We also applied an

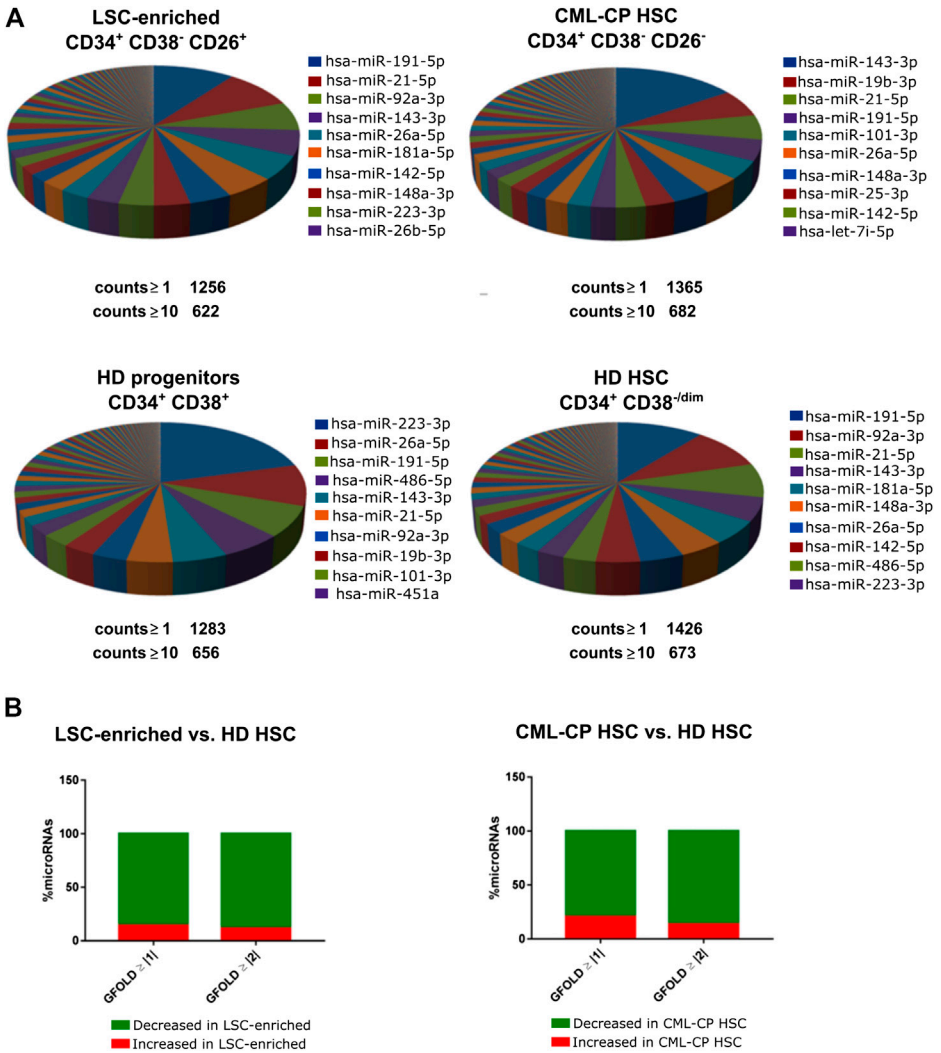


FIGURE 1 | Global patterns in the miRNome of LSC-enriched CD34$^+$CD38$^-$CD26$^+$ and CML-CP HSC fractions. **(A)** Pie chart representing the relative abundance of each microRNA in each fraction assessed by small RNA-NGS. The total number of different microRNAs (with at least 1 or 10 counts) is detailed below the pie chart. The names of the top-10 most abundant microRNAs in each fraction are detailed. **(B)** Global decrease in microRNA levels in LSC-enriched CD34$^+$CD38$^-$CD26$^+$ or CML-CP HSC fractions compared to HD HSC. Most microRNAs dysregulated in both fractions from CML-CP samples had decreased levels (GFOLD ≥ |1| or GFOLD ≥ |2|) compared to HD HSC.

overrepresentation analysis based on microRNA-sets using TAM 2.0, a manually curated database of functional and disease associations of microRNAs (Li et al., 2018). *Bias* assessment using the same random lists resulted in none statistically significant associations (**Supplementary Data Sheet S3**). Using as input microRNAs upregulated in CML-CP LSC-enriched fraction, this analysis detected a significant enrichment in "lipid metabolism" (False Discovery Rate, FDR: 0.0141), "hematopoiesis" (FDR: 0.0217), TF *Early growth response 1* (*EGR1*) (FDR: 0.007), and miR-99b cluster (FDR: 0.0138) (**Table 2**). Network analysis allowed the extraction of statistically significant modules (**Figures 2C,D**), suggesting the existence of mechanisms that regulate microRNA levels in a coordinated fashion.

Surprisingly, most (seven out of eight) microRNAs with decreased levels in LSC-enriched fraction belong to a genomic cluster located in region 14q32 (*DKL1/DIO3* locus). Moreover, inspection of microRNAs and snoRNAs from this locus revealed that 18 additional microRNAs and five snoRNAs had decreased levels in LSC-enriched vs. HD-HSC fractions (**Figure 3**; **Supplementary Table S4**). Accordingly, overrepresentation analysis by TAM 2.0 detected a significant enrichment in miR-379 cluster (located in the *DKL1/DIO3* locus) in microRNAs downregulated in the LSC-enriched fraction (FDR: 0.002, **Table 2**). Given that this region contains imprinted genes (Benetatos et al., 2013), this result suggests a mechanism of epigenetic silencing in the LSC-enriched CD34$^+$CD38$^-$CD26$^+$ fraction.

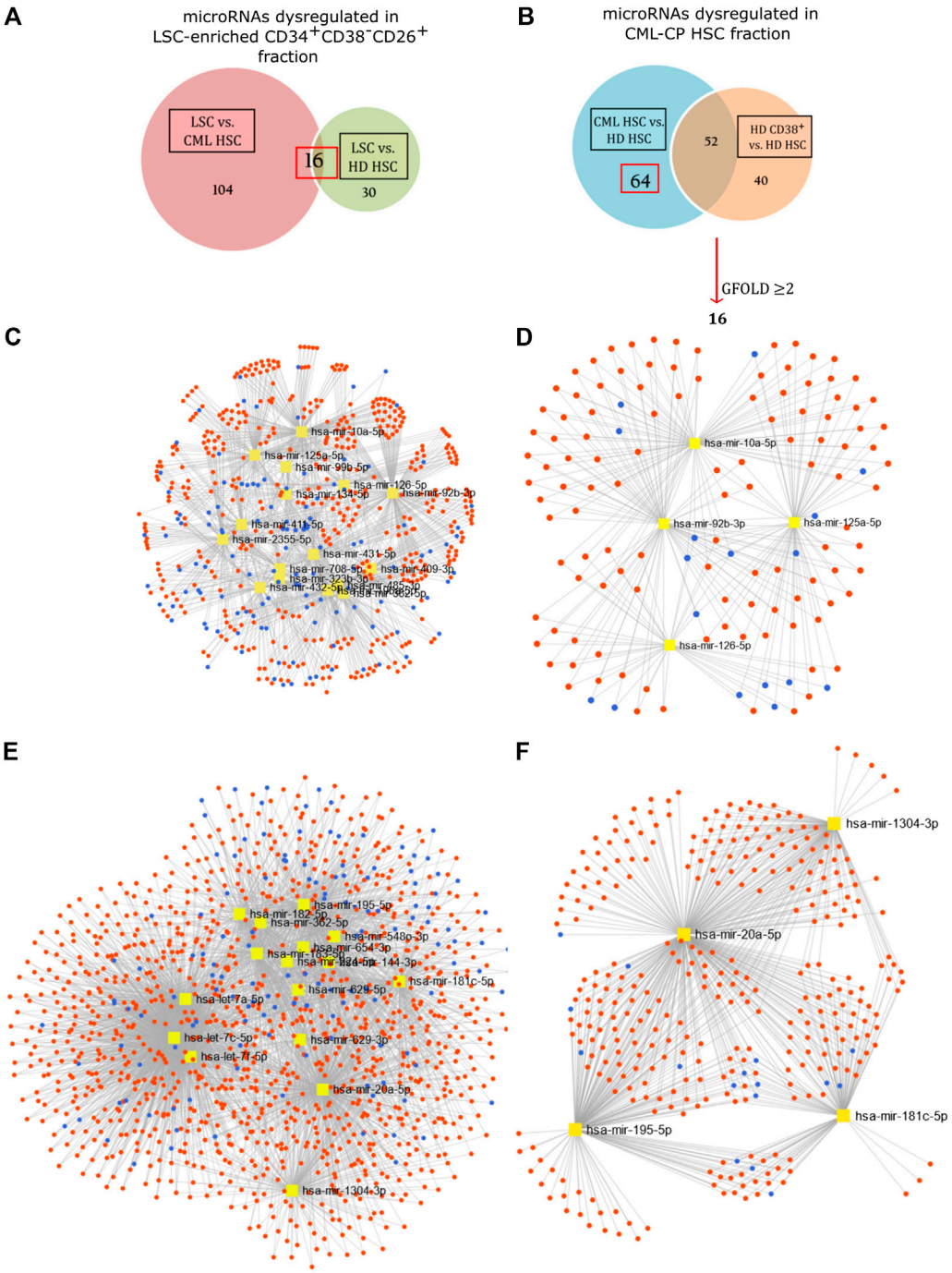


FIGURE 2 | Network analysis. Differentially expressed microRNAs in LSC-enriched CD34^{+}CD38^{-}CD26^{+} **(A,C)** and CML-CP HSC **(B,E)** fractions were analyzed as part of a network of microRNAs (yellow), target mRNAs (red) and lncRNAs (blue) using the bioinformatic tool miRNet (v2.0). Statistically significant modules were extracted for microRNAs dysregulated in LSC-enriched CD34^{+}CD38^{-}CD26^{+} **(D)** and CML-CP HSC **(F)** fractions.

microRNAs in HSC From CML-CP Patients Show a Dysregulated miRNome Despite the Absence of BCR-ABL1

Based on the hypothesis that HSC present in CML-CP patients are not equivalent to HSC in HD, we compared microRNAs between both fractions. We found 64 microRNAs significantly dysregulated (**Figure 2B**); further selection (GFOLD ≥ |2|) resulted in a list of 16 microRNAs (**Supplementary Table S5**). It is important to clarify that HSC from HD were sorted using a less strict gating of CD38-negative cells (resulting in a CD38$^{-/\mathrm{dim}}$ population), because we obtained very low yields

TABLE 2 | Overrepresented associations in differentially expressed microRNAs (TAM 2.0).

Category	Term	FDR	microRNA
Input: Upregulated microRNAs in LSC-enriched CD34$^+$CD38$^-$CD26$^+$ vs. CML-CP HSC fractions (miR-92b, miR-196a, miR-126, miR-125a, miR-2355, miR-99b, miR-411, miR-10a)			
Function	Lipid metabolism	0.0141	hsa-mir-125a, hsa-mir-126, hsa-mir-196a-2, hsa-mir-196a-1
Function	Hematopoiesis	0.0217	hsa-mir-125a, hsa-mir-126, hsa-mir-196a-2, hsa-mir-196a-1
TF	EGR1	0.00704	hsa-mir-125a, hsa-mir-99b, hsa-mir-10a, hsa-mir-92b
Cluster	hsa-mir-99b cluster	0.0138	hsa-mir-99b, hsa-mir-125a
Input: downregulated microRNAs in LSC-enriched CD34$^+$CD38$^-$CD26$^+$ vs. CML-CP HSC fractions (miR-708, miR-431, miR-134, miR-485, miR-409, miR-323b, miR-432, miR-382)			
Cluster	hsa-mir-379 cluster	0.00248	hsa-mir-382, hsa-mir-134, hsa-mir-485, hsa-mir-323b, hsa-mir-409
Input: downregulated microRNAs in CML-CP HSC vs. HD HSC (excluding let-7 family) (miR-20a, miR-195, miR-654, miR-224, miR-144, miR-181c, miR-548o, miR-182, miR-183)			
Function	Glucose metabolism	0.0514	hsa-mir-20a, hsa-mir-195, hsa-mir-144
Transcription factor	E2F1	0.0457	hsa-mir-20a, hsa-mir-224, hsa-mir-195
Transcription factor	STAT5	0.0557	hsa-mir-195, hsa-mir-20a

FDR, false discovery rate.
Only statistically significant results are shown.

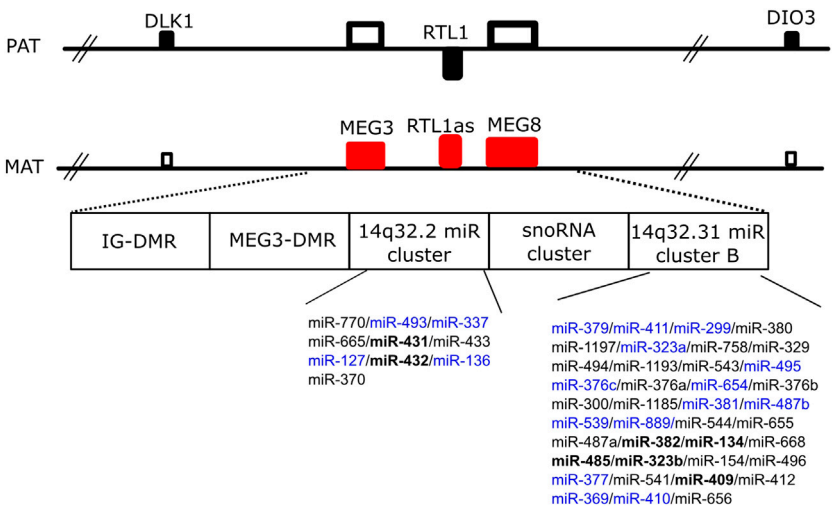


FIGURE 3 | Schematic representation of the genomic 14q32 region. This region includes coding genes with paternal (PAT) imprinting (*DLK1*, *RTL1*, *DIO3*), non-coding genes with maternal (MAT) imprinting (*MEG3*, *RTL1as*, *MEG8*), two clusters of microRNAs, and one cluster of snoRNAs. microRNAs in bold showed decreased levels in LSC-enriched CD34$^+$CD38$^-$CD26$^+$ vs. both HD HSC and CML-CP HSC; while microRNAs in blue showed decreased levels in LSC-enriched CD34$^+$CD38$^-$CD26$^+$ vs. HD HSC but did not vary compared to CML-CP HSC. IG, intergenic; DMR, differentially methylated regions. Adapted from (Nadal et al., 2014).

from individual samples. Therefore, in order to exclude differentially expressed microRNAs related to the inclusion of a CD38^{dim} population in HD, we excluded microRNAs that were differentially expressed between CD38$^{-/dim}$ and CD38$^+$ fractions from HD, under the assumption that some of these microRNAs would be related to the process of hematopoietic differentiation. Network (miRNet) and overrepresentation analysis by TAM 2.0 were performed. MicroRNAs belonging to Let-7 family were not included in the analysis by TAM 2.0 due to their association to multiple functions. Overrepresentation analysis resulted in "glucose metabolism" (FDR: 0.0514), TFs *E2F transcription factor 1* (*E2F1*, FDR: 0.0457) and *Signal transducer and activator of transcription 5* (*STAT5*, FDR: 0.0557) in those microRNAs downregulated in CML-CP HSC (**Table 2**). Network analysis allowed extraction of statistically significant modules (**Figures 2E,F**). These results suggest that putative HSC from CML-CP patients display an altered repertoire of microRNAs, as a consequence of either extrinsic factors (i.e., a niche altered by coexistence with leukemic cells), and/or that they are pre-leukemic neoplastic stem cells and thus harbor early, BCR-ABL1-independent genetic or epigenetic alterations that affect microRNA levels (i.e., mutations in microRNA-processing machinery).

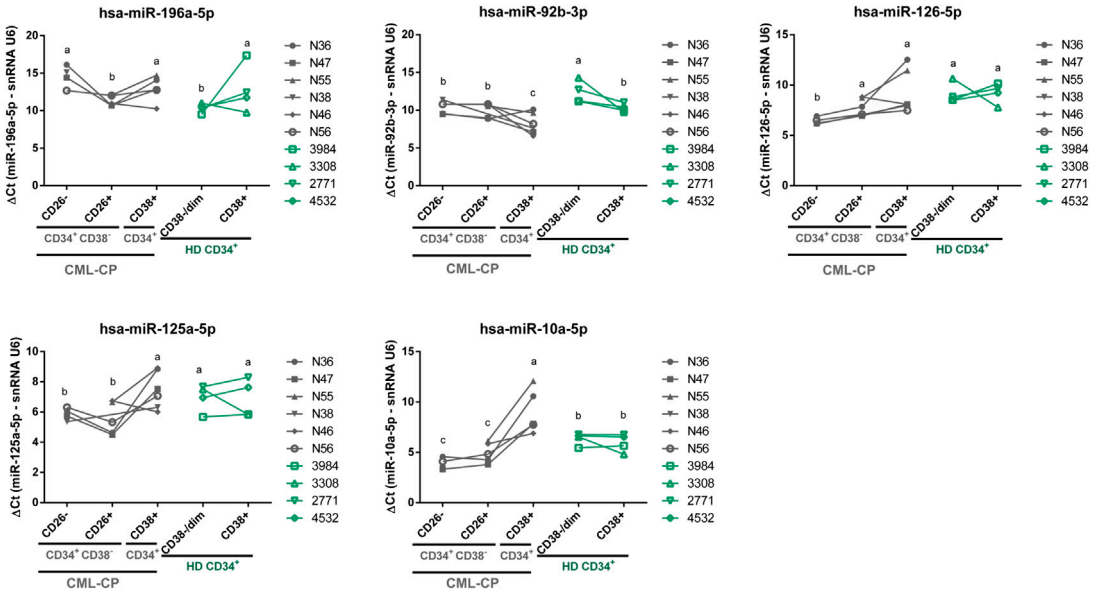


FIGURE 4 | Validation of microRNAs by RT-q PCR in a new cohort of CML-CP and HD samples. Results are expressed as ΔCt = Ct (microRNA) – Ct (snRNA U6). Each dot is the mean of technical duplicates from each patient or HD. CML-CP samples are represented in grey symbols, and HD samples in green symbols. Lines connect different fractions from the same patient or HD. Different letters indicate statistically significant differences (linear mixed-effects model, *a posteriori* comparison, global $\alpha = 0.05$).

Validation by RT-qPCR in a New Cohort of CML-CP Patients and HD

As other techniques, NGS is not free of intrinsic bias, mainly related to library preparation, platform used for sequencing, and data analysis (Baran-gale et al., 2015). In addition to perform a technical validation, we aimed to validate NGS results in a new cohort of patients and HD using RT-qPCR. We performed a multiplex RT step that allowed us to measure individual microRNA levels using very low inputs of RNA; therefore, pooling of samples from different patients or HD was not necessary, and we could assess intra-group variability. We evaluated the following fractions from CML-CP patients or HD samples: CML-CP LSC-enriched ($CD34^+CD38^-CD26^+$), CML-CP HSC ($CD34^+CD38^-CD26^-$), CML-CP progenitors ($CD34^+CD38^+$), HD HSC ($CD34^+CD38^{-/dim}$), HD progenitors ($CD34^+CD38^+$).

We selected six microRNAs upregulated in LSC-enriched vs. HSC fractions from CML-CP patients (miR-125a-5p, miR-10a-5p, miR-126-5p, miR-92b-3p, miR-196a-5p, miR-2355-5p), because of their high GFOLD values and clustering in pathway analysis, and measured their levels in six CML-CP patients, and four HD. We also included four additional small RNAs: one potential "novel" microRNA that emerged from NGS data ("novel-3"), and three additional microRNAs which were of interest in this population according to previous references (let-7a-5p, miR-132-3p, miR-182-5p). Purity of fractions was assessed by RT-qPCR of *BCR-ABL1* in RNA isolated from sorted cells (**Supplementary Figure S5**). miR-2355-5p and miR-182-5p were not detected in most fractions, therefore we excluded them from posterior analysis. We detected significant differences among fractions for miR-125a-5p, miR-10a-5p, miR-126-5p, miR-92b-3p, and miR-196a-5p (global p-value < 0.05; linear mixed-effects model) (**Figure 4**). We did not detect global differences in the levels of "novel-3," miR-let-7a-5p, and miR-132-3p (global p-value > 0.05; linear mixed-effects model) (**Supplementary Figure S6**). Only for miR-196a-5p the trend of change between CML-CP LSC-enriched and CML-CP HSC fractions was the same in NGS (fold-change LSC-enriched/CML-CP HSC = 9.6) and RT-qPCR (mean = 10.7, SD = 4.6) (**Supplementary Figure S7**). In contrast, differences in Ct values between CML-CP LSC-enriched and CML-CP HSC fractions were not statistically significant for miR-125a-5p and miR-92b-3p, but both were significantly increased compared to HD HSC (adjusted p-value < 0.05, *a posteriori comparison*, **Figure 4**) On the contrary, miR-126-5p and miR-10a-5p displayed opposite trends in NGS and RT-qPCR (**Supplementary Figure S7**).

CML-CP progenitors were only measured by RT-qPCR; interestingly miR-92b-3p and miR-10a-5p levels were significantly increased or decreased, respectively, in CML-CP progenitors in comparison with all the other CML-CP and HD fractions (**Figure 4**; **Supplementary Figure S8** for log2-fold change values).

According to significant functions detected by TAM 2.0, we searched for possible targets of miR-196a-5p included in gene lists classified in Reactome as "Metabolism of lipids" and Gene Ontology "hematopoietic stem cell differentiation". The search of potential targets included experimentally validated microRNA-mRNA interactions (assessed by ChemiRs) (**Supplementary Figure S9**). This analysis reduced the number of potential targets from 307 to 14 genes.

DISCUSSION

To our knowledge, this is the first description of the miRNome of CML-CP LSC-enriched $CD34^+CD38^-CD26^+$ fraction and its $CD26^-$ counterpart. In these analyses, major differences were found when analyzing *BCR-ABL1*$^+$ and *BCR-ABL1*$^-$ fractions in newly diagnosed patients. In addition, we detected several major differences in the miRNome pattern when comparing with HSC of HD. First, we observed a global downregulation of microRNAs in CML-CP LSC-enriched $CD34^+CD38^-CD26^+$ fraction and its $CD26^-$ counterpart in comparison with HSC of HD. Second, clusters and TF-associated networks of microRNAs were detected among differentially expressed microRNAs, suggesting that mature levels of functionally related microRNAs are (dys) regulated by common mechanisms. Third, compared to HSC from HD, we detected decreased levels in the LSC-enriched fraction of microRNAs and snoRNAs belonging to a genomic cluster located in chromosome 14 (14q32). Fourth, a high number of microRNAs were differentially expressed between putative HSC from CML-CP patients and HSC from HD, suggesting an altered phenotype of the "normal" HSC fraction in CML-CP patients. Finally, we confirmed by RT-qPCR that the levels of miR-196a-5p were increased more than nine-fold in LSC-enriched $CD34^+CD38^-CD26^+$ (*BCR-ABL1*$^+$) vs. $CD26^-$(*BCR-ABL1*$^-$) fractions from CML-CP patients at diagnosis, and *in silico* analysis revealed a significant association to lipid metabolism and hematopoiesis functions.

It is important to consider the limitations of this study in the interpretation of the results. In particular, the possible effects of pooling different samples (e.g., low statistical power or dilution of differences), and the transport between countries for small RNA-NGS (e.g., possible degradation of some microRNAs). In addition, future studies using a higher number of samples are desirable to confirm these observations. The measurement of a new cohort, non-pooled patient and HD samples in the validation step by RT-qPCR partially overcame these limitations.

Global downregulation of microRNAs in cancer has been reported in different tumors (Lu et al., 2005). Multiple mechanisms have been described to explain microRNA dysregulation in cancer, including genomic structural variations, altered regulation of microRNA transcription, epigenetic changes, defective microRNA processing machinery, and dysregulation of the complex that mediates pre-microRNA export from the nucleus (Croce, 2009). In CML, global microRNA depletion in patient samples has not been reported so far. In the work of Zhang et al., they showed, in K562 cells and CML $CD34^+$ cells, that BCR-ABL1 can affect the export of miR-126 precursors from the nucleus to the cytoplasm, through phosphorylation of SPRED1, a negative regulator of RAS superfamily of proteins, interfering with Ran-exportin 5-RCC1 complex (Zhang et al., 2018). Interestingly, this effect was reversible by treatment with Nilotinib. However, in our work, we observed decreased levels of mature microRNAs in both *BCR-ABL1*$^+$ and *BCR-ABL1*$^-$ primitive ($CD34^+CD38^-$) cells compared to HSC from HD, suggesting a BCR-ABL1-independent mechanism.

Clustering of microRNAs dysregulated in the LSC-enriched fraction suggests the existence of mechanisms of coordinated regulation. MicroRNAs can belong to families in which members are evolutionary related, therefore they share regions of common sequences, and can regulate similar or related targets. In this context, miR-125a and miR-10a belong to the miR-10/miR-100 family, and we found a significant positive correlation of both microRNAs in samples evaluated by RT-qPCR [r (Pearson) = 0.85; $p = 5.7 \times 10^{-7}$]. This suggests that future studies aimed at evaluating the functional relevance of microRNAs dysregulated in this system should take into consideration possible functional redundancy between related microRNAs. In fact, knockout experiments of microRNAs belonging to the same family have shown partially redundant effects on mice (Wong et al., 2015). Therefore, the combination of individual and simultaneous knockdown of correlated microRNAs would be an ideal approach.

In humans, the *DKL1/DIO3* locus at the 14q32 region, contains the paternally expressed genes *Delta-like 1 homolog* (*DLK1*), *Retrotransposon-like 1* (*RTL1*), and *Iodothyronine deiodinase 3* (*DIO3*), and the maternally expressed genes *MEG3*, *MEG8*, and anti-sense *RTL1*. *MEG3* and *MEG8* are long intergenic RNAs; *MEG3* has been found dysregulated in several types of tumors, and it is believed to function as a tumor-suppressor gene through interactions with p53 (Zhou et al., 2007). *MEG3* was shown to be downregulated in CML-CP samples, and patients in advanced phase and blast crisis showed further decreased levels of *MEG3* (Zhou et al., 2017). The largest mammalian cluster of microRNAs, together with a cluster of snoRNAs, are included in the maternally expressed strand of this locus. MicroRNAs in this locus were reported as mediators of ground-state pluripotency in mouse embryonic stem cells via inhibition of multi-lineage differentiation and promotion of self-renewal (Moradi et al., 2017), and their expression levels were correlated with pluripotency in mouse induced pluripotent stem cells (Liu et al., 2010). Non-coding RNAs (including microRNAs) from this locus maintained mouse fetal liver and adult long-term repopulating HSCs (LT-HSCs) through the suppression of the PI3K-mTOR pathway, which resulted in inhibition of mitochondrial biogenesis and metabolic activity (Qian et al., 2016). On the other hand, miR-300 belongs to the *DLK1/DIO3* locus and was reported to be highly expressed in quiescent CP and blast crisis CML LSC ($CD34^+CFSE^{max}$) (Silvestri et al., 2020). We detected a downregulation of microRNAs and snoRNAs from the 14q32 cluster, and the absence of expression of miR-300 in all fractions evaluated by NGS. Both results suggest that the fractions used in our study included mostly non-quiescent $CD34^+CD38^-$ cells with a possible decrease in multipotency. This could also explain the similarity observed between LSC-enriched $CD34^+CD38^-CD26^+$ and HD-HSC ($CD34^+CD38^{-/dim}$, expected to include more differentiated cells) fractions (**Figures 1A**, **2A**).

The role of microenvironmental factors in the development of hematological malignancies is an exciting field. In our study, a great number of microRNAs were dysregulated between HSC from CML-CP patients and HSC from HD. This observation could be attributed to cell-autonomous (i.e., genetic or epigenetic

alterations) and/or extrinsic factors. Evidence regarding a role of the BM microenvironment in CML include the alteration of the BM niche by LSC by secretion of costimulatory molecules and suppressive cytokines that target metabolic pathways (Zhang et al., 2012; Johnson et al., 2016; Agarwal et al., 2019), the presence of tumor-derived exosomes that can modulate immune responses (Boyiadzis and Whiteside, 2017), and hypoxic conditions that can influence LSC quiescence, differentiation, metabolism and therapy resistance (Ng et al., 2014; Silvestri et al., 2020). It would be of interest to assess whether the miRNome of CML-CP HSC is restored upon TKI treatment.

Bioinformatic tools for microRNA analysis are part of a growing field. Traditional enrichment analysis of pathways based on predicted targets of selected microRNAs resulted in a high rate of false positives. This type of *bias* has been reported as a result of higher redundancy of pathway information at the microRNA level than at gene level (Godard and van Eyll, 2015). Through an overrepresentation analysis that bypasses targets using manually curated lists of microRNA-associations, we detected an enrichment in microRNAs regulated by EGR1 TF, and also in the miR-99b cluster. Interestingly, EGR1 has been reported as regulator of homeostasis of HSC, where it is highly expressed, and downregulated after induction of cell division and migration (Min et al., 2008), whereas miR-99b cluster has been reported as a conserved cluster preferentially expressed in long-term HSC (Guo et al., 2010). By this strategy we also detected an enrichment in lipid metabolism and hematopoiesis (microRNAs upregulated in LSC-enriched vs. CML-CP HSC fractions) and glucose metabolism (microRNAs downregulated in CML-CP HSC vs. HD HSC). In the light of recent descriptions of increased oxidative metabolism in CML LSC-enriched fractions (Kuntz et al., 2017), the results obtained by us serve as a guide for future functional studies that evaluate the role of microRNAs in this process. Kuntz et al. described an increase in fatty acid oxidation and lipolysis, increased glucose oxidation and anaplerosis in stem cell enriched CML fractions, and showed that restriction of mitochondrial functions by treatment with tigecycline had *in vivo* cytotoxic effects on stem/progenitor CML cells, in combination with imatinib (Kuntz et al., 2017). Tigecycline is an FDA-approved antibiotic that inhibits bacterial protein synthesis, but also inhibits the synthesis of mitochondria-encoded proteins. Metabolic vulnerabilities in LSC open the road for new therapeutic strategies, and a thorough understanding of the differential mechanisms involved in CML vs. normal primitive cells is necessary in order to predict possible unfavorable side effects (Vetrie et al., 2020), which have been repeatedly observed in clinical trials that target this population.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: European Nucleotide Archive, PRJEB41369.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research ethics committee from Instituto Alexander Fleming, Ciudad Autónoma de Buenos Aires, Argentina. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MR and MB designed the experiments. MR performed most of the experiments and data analysis. MS contributed to the validation of microRNAs by RT-qPCR and processing of biological samples. SB and CF contributed to small RNA NGS-library construction and sequencing. DK and PY contributed to bioinformatic analysis of NGS-derived data. SC, MC, JF, BM, CP, MP, AV, and VV contributed with patient samples and analysis of clinical data. JS, IZ, IL, and JM participated with helpful discussion and ideas. PV contributed with conceptual input and technical support in CML stem cell experiments, reviewed the data and drafted parts of the manuscript. MR wrote the manuscript. MB supervised the entire work. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from CONICET-FAPERJ (2013), Agencia Nacional de Promoción Científica y Tecnológica of Argentina (PICT-2013 1710), Fundación Mosoteguy, and Fundación SALES. PV and his team are supported by the Austrian Science Fund FWF, grant F4704-B20. MB, IL, PY, and JM are researchers from the Consejo Nacional de Investigaciones Científicas y Tecnológicas of Argentina (CONICET). MR, MS, and DK received CONICET fellowships.

ACKNOWLEDGMENTS

We thank CML patients, healthy donors, and the service of Hemotherapy from Instituto Alexander Fleming (Buenos Aires, Argentina) for their invaluable help by providing samples for this project, and G. Cueto for his advice on statistical analysis of the data. A preprint version of this manuscript is available at bioRxiv (https://doi.org/10.1101/2020.03.16.989194) (Ruiz et al., 2020).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2020.612573/full#supplementary-material.

REFERENCES

Abraham, S. A., Hopcroft, L. E. M., Carrick, E., Drotar, M. E., Dunn, K., Williamson, A. J. K., et al. (2016). Dual targeting of p53 and c-MYC selectively eliminates leukaemic stem cells. *Nature* 534, 341–346. doi:10.1038/nature18288

Agarwal, P., Isringhausen, S., Li, H., Paterson, A. J., He, J., Gomariz, Á., et al. (2019). Mesenchymal niche-specific expression of Cxcl12 controls quiescence of treatment-resistant leukemia stem cells. *Cell Stem Cell* 24, 769–784. doi:10.1016/j.stem.2019.02.018

Arrigoni, E., del Re, M., Galimberti, S., Restante, G., Rofi, E., Crucitta, S., et al. (2018). Concise review: chronic myeloid leukemia: stem cell niche and response to pharmacologic treatment. *Stem Cells Transl. Med.* 7, 305–314. doi:10.1002/sctm.17-0175

Baran-gale, J., Kurtz, C. L., Erdos, M. R., Sison, C., Kalathur, R. C., and Baran-gale, J. (2015). Addressing bias in small RNA library preparation for sequencing: a new protocol recovers microRNAs that evade capture by current methods. *Front. Genet.* 6, 352. doi:10.3389/fgene.2015.00352

Bartel, D. P. (2018). Metazoan microRNAs. *Cell* 173, 20–51. doi:10.1016/j.cell.2018.03.006

Benetatos, L., Hatzimichael, E., Londin, E., Vartholomatos, G., Loher, P., Rigoutsos, I., et al. (2013). The microRNAs within the DLK1-DIO3 genomic region: involvement in disease pathogenesis. *Cell. Mol. Life Sci.* 70, 795–814. doi:10.1007/s00018-012-1080-8

Bonnet, D., and Dick, J. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* 3, 730–737. doi:10.1038/nm0797-730

Boyiadzis, M., and Whiteside, T. L. (2017). The emerging roles of tumor-derived exosomes in hematological malignancies. *Leukemia* 31, 1259–1268. doi:10.1038/leu.2017.91

Bruns, I., Czibere, A., Fischer, J. C., Roels, F., Cadeddu, R., Buest, S., et al. (2009). The hematopoietic stem cell in chronic phase CML is characterized by a transcriptional profile resembling normal myeloid progenitor cells and reflecting loss of quiescence. *Leukemia* 23, 892–899. doi:10.1038/leu.2008.392

Calin, G. A., and Croce, C. M. (2006). MicroRNA signatures in human cancers. *Nat. Rev. Cancer* 6, 857–866. doi:10.1038/nrc1997

Chang, L., Zhou, G., Soufan, O., and Xia, J. (2020). miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology. *Nucl. Acids Res.* 48, W244–W251. doi:10.1093/nar/gkaa467

Chen, C., Ridzon, D. A., Broomer, A. J., Zhou, Z., Lee, D. H., Nguyen, J. T., et al. (2005). Real-time quantification of microRNAs by stem–loop RT-PCR. *Nucl. Acids Res* 33, e179. doi:10.1093/nar/gni178

Chomel, J., Bonnet, M., Sorel, N., Bertrand, A., Meunier, M., Fichelson, S., et al. (2011). Leukemic stem cell persistence in chronic myeloid leukemia patients with sustained undetectable molecular residual disease. *Blood* 118, 3657–3660. doi:10.1182/blood-2011-02-335497

Corbin, A. S., Agarwal, A., Loriaux, M., Cortes, J., Deininger, M. W., and Druker, B. J. (2011). Human chronic myeloid leukemia stem cells are insensitive to imatinib despite inhibition of BCR-ABL activity. *J. Clin. Invest.* 121, 396–409. doi:10.1172/JCI35721DS1

Croce, C. M. (2009). Causes and consequences of microRNA dysregulation in cancer. *Nat. Rev. Genet.* 10, 704–714. doi:10.1038/nrg2634

di Rienzo, J. A., Casanoves, F., Balzarini, M. G., Gonzalez, L., Tablada, M., and Robledo, C. W. (2011). InfoStat version 2011. Available from: http://www.infostat.com.ar/

Druker, B., Tamura, S., Buchdunger, E., Ohno, S., Segal, G., Fanning, S., et al. (1996). Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat. Med.* 2, 561–566. doi:10.1038/nm0596-561

Etienne, G., Guilhot, J., Rea, D., Rigal-Huguet, F., Nicolini, F., Charbonnier, A., et al. (2017). Long-term follow-up of the French Stop Imatinib (STIM1) study in patients with chronic myeloid leukemia. *J. Clin. Oncol.* 35, 298–305. doi:10.1200/JCO.2016.68.2914

Feng, J., Meyer, C. A., Wang, Q., Liu, J. S., Liu, X. S., and Zhang, Y. (2012). GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. *Bioinformatics* 28, 2782–2788. doi:10.1093/bioinformatics/bts515

Foley, S. B., Hildenbrand, Z. L., Soyombo, A. A., Magee, J. A., Wu, Y., Oravecz-wilson, K. I., et al. (2013). Expression of BCR/ABL p210 from a knockin allele enhances bone marrow engraftment without inducing neoplasia. *Cell Rep.* 5, 51–60. doi:10.1016/j.celrep.2013.08.037

Giustacchini, A., Thongjuea, S., Barkas, N., Woll, P. S., Povinelli, B. J., Booth, C. A. G., et al. (2017). Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nat. Med.* 23, 692–702. doi:10.1038/nm.4336

Godard, P., and van Eyll, J. (2015). Pathway analysis from lists of microRNAs: common pitfalls and alternative strategy. *Nucl. Acids Res.* 43, 3490–3497. doi:10.1093/nar/gkv249

Guo, S., Lu, J., Schlanger, R., Zhang, H., Wang, J. Y., Fox, M. C., et al. (2010). MicroRNA miR-125a controls hematopoietic stem cell number. *Proc. Natl. Acad. Sci. U S A.* 107, 14229–14234. doi:10.1073/pnas.0913574107

Herrmann, H., Sadovnik, I., Cerny-Reiterer, S., Rülicke, T., Stefanzl, G., Willmann, M., et al. (2014). Dipeptidylpeptidase IV (CD26) defines leukemic stem cells (LSC) in chronic myeloid leukemia. *Blood* 123, 3951–3962. doi:10.1182/blood-2013-10-536078

Holyoake, T., Jiang, X., Eaves, C., and Eaves, A. (1999). Isolation of a highly quiescent subpopulation of primitive leukemic cells in chronic myeloid leukemia. *Blood* 94, 2056–2064.

Holyoake, T. L., and Vetrie, D. (2017). The chronic myeloid leukemia stem cell: stemming the tide of persistence. *Blood* 129, 1595–1606. doi:10.1182/blood-2016-09-696013

Hsu, C. L., Chang, H. Y., Chang, J. Y., Hsu, W. M., Huang, H. C., and Juan, H. F. (2016). Unveiling MYCN regulatory networks in neuroblastoma via integrative analysis of heterogeneous genomics data. *Oncotarget* 7, 36293–36310. doi:10.18632/oncotarget.9202

Hulsen, T., de Vlieg, J., and Alkema, W. (2008). BioVenn–a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genom.* 9, 488. doi:10.1186/1471-2164-9-488

Johnson, S. M., Dempsey, C., Chadwick, A., Harrison, S., Liu, J., Di, Y., et al. (2016). To the editor: metabolic reprogramming of bone marrow stromal cells by leukemic extracellular vesicles in acute lymphoblastic leukemia. *Blood* 128, 453–456. doi:10.1182/blood-2015-12-688051

Kinstrie, R., Horne, G. A., Morrison, H., Irvine, D., Munje, C., Castañeda, E. G., et al. (2020). CD93 is expressed on chronic myeloid leukemia stem cells and identifies a quiescent population which persists after tyrosine kinase inhibitor therapy. *Leukemia* 34, 1613–1625. doi:10.1038/s41375-019-0684-5

Kumari, A., Brendel, C., Hochhaus, A., Neubauer, A., and Burchert, A. (2012). Low BCR-ABL expression levels in hematopoietic precursor cells enable persistence of chronic myeloid leukemia under imatinib. *Blood* 119, 530–539. doi:10.1182/blood-2010-08-303495

Kuntz, E. M., Baquero, P., Michie, A. M., Dunn, K., Tardito, S., Holyoake, T. L., et al. (2017). Targeting mitochondrial oxidative phosphorylation eradicates therapy-resistant chronic myeloid leukemia stem cells. *Nat. Med.* 23, 1234–1240. doi:10.1038/nm.4399

Li, J., Han, X., Wan, Y., Zhang, S., Zhao, Y., Fan, R., et al. (2018). TAM 2.0: tool for MicroRNA set analysis. *Nucl. Acids Res.* 46, W180–W185. doi:10.1093/nar/gky509

Liu, L., Luo, G. Z., Yang, W., Zhao, X., Zheng, Q., Lv, Z., et al. (2010). Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. *J. Biol. Chem.* 285, 19483–19490. doi:10.1074/jbc.M110.131995

Lu, J., Getz, G., Miska, E. A., Alvarez-saavedra, E., Lamb, J., Peck, D., et al. (2005). MicroRNA expression profiles classify human cancers. *Nature* 435, 834–838. doi:10.1038/nature03702

Mahon, F.-X., Réa, D., Guilhot, J., Guilhot, F., Huguet, F., Nicolini, F., et al. (2010). Discontinuation of imatinib in patients with chronic myeloid leukaemia who have maintained complete molecular remission for at least 2 years: the prospective, multicentre Stop Imatinib (STIM) trial. *Lancet Oncol.* 11, 1029–1035. doi:10.1016/S1470-2045(10)70233-3

Mckenzie, J. L., Gan, O. I., Doedens, M., Wang, J. C. Y., and Dick, J. E. (2006). Individual stem cells with highly variable proliferation and self-renewal properties comprise the human hematopoietic stem cell compartment. *Nat. Immunol.* 7, 1225–1233. doi:10.1038/ni1393

Min, I. M., Pietramaggiori, G., Kim, F. S., Passegué, E., Stevenson, K. E., and Wagers, A. J. (2008). The transcription factor EGR1 controls both the proliferation and localization of hematopoietic stem cells. *Cell Stem Cell* 2, 380–391. doi:10.1016/j.stem.2008.01.015

Moradi, S., Sharifi-Zarchi, A., Ahmadi, A., Mollamohammadi, S., Stubenvoll, A., Günther, S., et al. (2017). Small RNA sequencing reveals Dlk1-Dio3 locus-embedded MicroRNAs as major drivers of ground-state pluripotency. *Stem Cell Rep.* 9, 2081–2096. doi:10.1016/j.stemcr.2017.10.009

Nadal, E., Zhong, J., Lin, J., Reddy, R. M., Ramnath, N., Orringer, M. B., et al. (2014). A microRNA cluster at 14q32 drives aggressive lung adenocarcinoma. *Clin. Cancer Res.* 20, 3107–3117. doi:10.1158/1078-0432.CCR-13-3348

Nakamura-Ishizu, A., Takizawa, H., and Suda, T. (2014). The analysis, roles and regulation of quiescence in hematopoietic stem cells. *Development* 141, 4656–4666. doi:10.1242/dev.106575

Neviani, P., Harb, J. G., Oaks, J. J., Santhanam, R., Walker, C. J., Ellis, J. J., et al. (2013). PP2A-activating drugs selectively eradicate TKI-resistant chronic myeloid leukemic stem cells. *J. Clin. Invest.* 123, 4144–4157. doi:10.1172/JCI68951DS1

Ng, K. P., Manjeri, A., Lee, K. L., Huang, W., Tan, S. Y., Chuah, C. T. H., et al. (2014). Physiologic hypoxia promotes maintenance of CML stem cells despite effective BCR-ABL1 inhibition. *Blood* 123, 3316–3326. doi:10.1182/blood-2013-07-511907

Nguyen, L. V., Vanner, R., Dirks, P., and Eaves, C. J. (2012). Cancer stem cells: an evolving concept. *Nat. Rev. Cancer* 12, 133–143. doi:10.1038/nrc3184

Qian, P., He, X., Paulson, A., Li, Z., Tao, F., Perry, J., et al. (2016). The Dlk1-Gtl2 locus preserves LT-HSC function by inhibiting the PI3K-mTOR pathway to restrict mitochondrial metabolism. *Cell Stem Cell* 18, 214–228. doi:10.1016/j.stem.2015.11.001

Ross, D. M., Pagani, I. S., Shanmuganathan, N., Kok, C. H., Seymour, J. F., Mills, A. K., et al. (2018). Long-term treatment-free remission of chronic myeloid leukemia with falling levels of residual leukemic cells. *Leukemia* 32, 2572–2579. doi:10.1038/s41375-018-0264-0

Rowley, J. (1973). A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and giemsa staining. *Nature* 243, 290–293. doi:10.1038/243290a0

Ruiz, M. S., Sánchez, M. B., Bonecker, S., Furtado, C., Koile, D., Yankilevich, P., et al. (2020). miRNome profiling of clonal stem cells in Ph + CML [Preprint]. *bioRxiv*. doi:10.1101/2020.03.16.989194

Salati, S., Salvestrini, V., Carretta, C., Genovese, E., Rontauroli, S., Zini, R., et al. (2017). Deregulated expression of miR-29a-3p, miR-494-3p and miR-660-5p affects sensitivity to tyrosine kinase inhibitors in CML leukemic stem cells. *Oncotarget* 8, 49451–49469. doi:10.18632/oncotarget.17706

Saussele, S., Richter, J., Guilhot, J., Gruber, F. X., Hjorth-Hansen, H., Almeida, A., et al. (2018). Discontinuation of tyrosine kinase inhibitor therapy in chronic myeloid leukaemia (EURO-SKI): a prespecified interim analysis of a prospective, multicentre, non-randomised, trial. *Lancet Oncol.* 19, 747–757. doi:10.1016/S1470-2045(18)30192-X

Silvestri, G., Trotta, R., Stramucci, L., Ellis, J. J., Harb, J. G., Neviani, P., et al. (2020). Persistence of drug-resistant leukemic stem cells and impaired NK cell immunity in CML patients depend on MIR300 antiproliferative and PP2A-activating functions. *Blood Cancer Discov.* 1, 48–67. doi:10.1158/0008-5472.bcd-19-0039

Su, E. C., Chen, Y., Tien, Y., Liu, J., Ho, B., Yu, S., et al. (2016). ChemiRs: a web application for microRNAs and chemicals. *BMC Bioinf.* 17. doi:10.1186/s12859-016-1002-0

Valent, P., Bonnet, D., de Maria, R., Lapidot, T., Copland, M., Melo, J. V., et al. (2012). Cancer stem cell definitions and terminology: the devil is in the details. *Nat. Rev. Cancer* 12, 767–775. doi:10.1038/nrc3368

Valent, P., Bonnet, D., Wöhrer, S., Andreeff, M., Copland, M., Chomienne, C., et al. (2013). Heterogeneity of neoplastic stem cells: theoretical, functional, and clinical implications. *Cancer Res.* 73, 1037–1045. doi:10.1158/0008-5472.CAN-12-3678

Vetrie, D., Helgason, G. V., and Copland, M. (2020). The leukaemia stem cell: similarities, differences and clinical prospects in CML and AML. *Nat. Rev. Cancer* 20, 158–173. doi:10.1038/s41568-019-0230-9

Vitsios, D. M., and Enright, A. J. (2015). Chimira: analysis of small RNA sequencing data and microRNA modifications. *Bioinformatics* 31, 3365–3367. doi:10.1093/bioinformatics/btv380

Vlachos, I. S., Zagganas, K., Paraskevopoulou, M. D., Georgakilas, G., Karagkouni, D., Vergoulis, T., et al. (2015). DIANA-miRPath v3.0: deciphering microRNA function with experimental support. *Nucl. Acids Res.* 43, W460–W466. doi:10.1093/nar/gkv403

Warfvinge, R., Geironson, L., Sommarin, M. N. E., Lang, S., Karlsson, C., Roschupkina, T., et al. (2017). Single-cell molecular analysis defines therapy response and immunophenotype of stem cell subpopulations in CML. *Blood* 129, 2384–2395. doi:10.1182/blood-2016-07-728873

Wong, S. F. L., Agarwal, V., Mansfield, J. H., Denans, N., Schwartz, M. G., Prosser, H. M., et al. (2015). Independent regulation of vertebral number and vertebral identity by microRNA-196 paralogs. *Proc. Natl. Acad Sci. U S A.* 112, E4884–E4893. doi:10.1073/pnas.1512655112

Zhang, B., Ho, Y. W., Huang, Q., Maeda, T., Lin, A., Lee, S. U.-K., et al. (2012). Altered microenvironmental regulation of leukemic and normal stem cells in chronic myelogenous leukemia. *Cancer Cell* 21, 577–592. doi:10.1016/j.ccr.2012.02.018

Zhang, B., Nguyen, L. X. T., Li, L., Zhao, D., Kumar, B., Wu, H., et al. (2018). Bone marrow niche trafficking of miR-126 controls the self-renewal of leukemia stem cells in chronic myelogenous leukemia. *Nat. Med.* 24, 450–462. doi:10.1038/nm.4499

Zhou, X., Yuan, P., Liu, Q., and Liu, Z. (2017). LncRNA MEG3 regulates imatinib resistance in chronic myeloid leukemia via suppressing microRNA-21. *Biomol. Ther.* 25, 490–496. doi:10.4062/biomolther.2016.162

Zhou, Y., Zhong, Y., Wang, Y., Zhang, X., Batista, D. L., Gejman, R., et al. (2007). Activation of p53 by MEG3 non-coding RNA. *J. Biol. Chem.* 282, 24731–24742. doi:10.1074/jbc.M702029200

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 02 February 2021
doi: 10.3389/fphar.2020.616630

Genotype Analysis of *ABCC1*, *NCF4* and *CBR3* Polymorphism and the Association With Childhood Acute Lymphoblastic Leukemia in Mexican Childhood Population

Jesús Alonso Gándara-Mireles[1], *Ismael Lares-Asseff*[1,2]*, *Elio Aarón Reyes Espinoza*[3], *Javier G. Blanco*[4], *Isaias Chairez Hernández*[1], *Lourdes Patricia Córdova Hurtado*[3], *Verónica Loera Castañeda*[1,2], *Leslie Patrón Romero*[5], *Cristina Venzor Sánchez*[3], *Hugo Payan Gándara*[3], *Dinora Arechiga Gurrola*[3] and *Horacio Almanza Reyes*[2,5]

[1]Instituto Politécnico Nacional, CIIDIR-Unidad Durango, Durango, México, [2]Red Latinoamericana de Implementación y Validación de Guías Clínicas Farmacogenómicas (RELIVAF-CYTED), Madrid, Spain, [3]Centro Estatal de Cancerología, Durango, México, [4]School of Pharmacy and Pharmaceutical Sciences, University of Buffalo, The State University of New York, Albany, NY, United States, [5]Facultad de Medicina y Psicología de la Universidad Autónoma de Baja California, Tijuana, México

OPEN ACCESS

Edited by:
Maria Ana Redal,
University of Buenos Aires, Argentina

Reviewed by:
Maoxiang Qian,
Fudan University, China
Lucia Taja-Chayeb,
National Institute of Cancerology (INCAN), Mexico

***Correspondence:**
Ismael Lares-Asseff
ismaelares@yahoo.com

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: 12 October 2020
Accepted: 27 November 2020
Published: 02 February 2021

Citation:
Gándara-Mireles JA, Lares-Asseff I, Reyes Espinoza EA, Blanco JG, Chairez Hernández I, Córdova Hurtado LP, Loera Castañeda V, Patrón Romero L, Venzor Sánchez C, Payan Gándara H, Arechiga Gurrola D and Almanza Reyes H (2021) Genotype Analysis of ABCC1, NCF4 and CBR3 Polymorphism and the Association With Childhood Acute Lymphoblastic Leukemia in Mexican Childhood Population. Front. Pharmacol. 11:616630. doi: 10.3389/fphar.2020.616630

Background: The identification of genetic risk factors for Acute Lymphoblastic Leukemia (ALL), are increasingly urgent and necessary.

Objective: The purpose of this study is to determine the association of the genetic polymorphisms *ABCC1* rs3743527, *NCF4* rs1883112 and *CBR3* rs1056892 with ALL.

Methods: DNA samples were obtained in 71 children with ALL (from 2 to 18 years) and in 71 controls without ALL, to determine the polymorphisms by real-time polymerase chain reaction (qPCR), using specific TaqMan probes in a StepOne® thermal cycler (Applied Biosystems, United States).

Results: The results of the Odds Ratio analysis show that in the rs1883112 polymorphism of the NCF4 gene, the heterozygous allele has a risk effect for ALL (OR = 3.1870, CI = 1.8880–7.9383 and p = 0.0002), in turn the mutated genotype (AA) is associated with a protective effect (OR = 0.26, 0.1248 to 0.5434 and p = 0.0003). On the other hand, the CBR3 rs1056892 polymorphism shows a significant association of risk to ALL, in the presence of the HT genotype (OR = 2.77, IC = 1.3837 to 5.5651 and p = 0.004) and the mutated genotype of this polymorphism has a significant association with protection to ALL in the HM genotype (OR = 0.52, IC = 0.2639 to 1.0304 and p = 0.05). While the inheritance models of the polymorphisms let us see that of the rs1883112 polymorphism of the NCF4 polymorphism; the HT genotype of the codominant model shows a protective effect against ALL (OR = 0.4117, IC = 0.1718 to 0.9866 and p = 0.04), the recessive model shows us and confirms what we already saw in table number 3, being that there is an association with protective effect in the HM genotype (OR = 0.2604, IC = 0.1248 to 0.5434 and p = 0.0003). In the polymorphism rs1056892 of the CBR3 gene, a protection association was found in the heterozygous allele of the codominant model (OR = 0.3448, IC = 0.1375 to 0.8896 and p =

0.0274). In addition, the recessive inheritance model for the HM genotype shows a protective effect to ALL, (OR = 0.52, CI = 0.9919 to 3.8638 and p = 0.05).

Conclusion: There is an evident impact of the NCF4 rs1883112 and CBR3 rs1056892 polymorphisms with an increased risk of susceptibility to ALL; Likewise, through the codominant inheritance model, the effect of the variation of the CBR3 rs1056892 gene as a protective factor against ALL was evaluated.

Keywords: genetic polymorphisms, ABCC1, NCF4, CBR3, association, leukemia

INTRODUCTION

Childhood Acute lymphoblastic leukemia (ALL) is the most common pediatric cancer in developed countries and also it is a malignant disease of the white blood cells with a multifactorial etiology that likely involve an interplay of environmental and genetic variables. Is estimated that more than 60% of patients diagnosed with ALL are children below the age of 15 years, with a peak of incidence at 2–5 years of age. With the actual treatment scheme, the rate of cure is 90% (Rivera-Luna et al., 2005; Pérez-Saldívar et al., 2011; Secretaria de Salud et al., 2011). The implementation of risk-stratified therapy has been successful, as survival rates for ALL have improved significantly, with overall survival rates of 5 years currently (Rodriguez et al., 2010; Muwakkit et al., 2012; Ceppi et al., 2015; Frost et al., 2018). A wide and deep understanding of genetic factors, as well as chromosomal abnormalities, gene expression, response to therapeutic treatment, and host pharmacogenomics, offer the potential to improve prognosis and therapeutic optimization in the treatment of childhood ALL (Vrooman and Silverman, 2009). The Genome Wide Association Study (GWAS) provide a candidate genes and cellular pathways to understand the etiology, also to determine candidates' genes to improve the treatment response and pharmacogenomics. These single nucleotide polymorphisms (SNP's) have been largely study over the last decades in ALL. Some of these pathways involve inflammation. Liu et al. (2020) in a case-control study found that polymorphisms of IL-6 and IL-10 are significantly correlated with the susceptibility and pathogenesis of ALL in childhood (Liu et al., 2020). Some other environmental factors associated with leukemiogenesis are maternal vitamins intake during pregnancy, particularity folic acid and its fetal metabolism. Folic acid (or vitamin B9) and its derivatives, collectively known as folates, are chemoprotective micronutrients of great interest belonging to the B vitamin group. Folic acid is involved in the correct production and maturation of blood cells from hematopoietic stem cells, DNA methylations, DNA synthesis and repair, proper gene expression, chromosomal and nuclear integrity. In particular, folic acid supplementation during pregnancy has a protective effect towards ALL, with a diminished incidence of 15–21% (Cantarella et al., 2017). Inside the cell, folic acid is transformed into tetrahydrofolate (THF), a reaction carried out by the enzyme NADPH oxidase-dependent hydrofolate reductase, this being a membrane-bound enzyme complex that faces the extracellular space (Navarro-Pérez et al., 2016). On the other hand, the neutrophil cytosolic factor 4 (NCF4) is a protein encoded in humans by the *NCF4* gene, which encodes the p40-phox (NADPH) oxidase subunit of the NOX2; this protein is part of the NADPH oxidase complex (Zhan et al., 1996). NOX2 complex has been related to at least ovarian cancer and renal cell carcinoma (Meitzler et al., 2014). Also, there is evidence that the NADPH-dependent oxidase NOX2 is an important effector of immune cell function, and its activity has been linked to oncogenic signaling (Adane et al., 2019). Gutiérrez-Salinas J. and collaborators (Gutiérrez-Salinas et al., 2016) in 2016 reported that oxidative stress promote the development of cancer, and carbonylated proteins may have a role in the carcinogenesis process. Carbonyl reductase three encoded by the *CBR3* gene catalyzes the reduction of a large number of biologically and pharmacologically active carbonyl compounds; a polymorphism present in this gene, such as rs1056892, may have implications for the risk of developing ALL. Other pathway involved in the overcome, prognostic and surveillance is the drug response to the treatment. One of the most important gene in this field is the *ABCC1* gene. Multidrug resistance-associated protein 1 (MRP1), encoded by the *ABCC1* gene, is an ATP-binding cassette transporter mediating efflux of organic anions and xenobiotics; its overexpression leads to multidrug resistance. Many drugs are good substrates for MRP1, so its overexpression leads to multidrug resistance (MDR), especially during cancer chemotherapy. This effect can be observed in many types of cancer cells including solid tumors (lung cancer, breast cancer, gastric and colon carcinomas, melanoma, prostate cancer, neuroblastoma) as well as in various types of leukemias. There exist ethnic differences in the frequency of *ABCC1* polymorphic variants and least 95 SNPs have been described in different population with also wide diversity (Słomka et al., 2016). Due to the importance of identifying risk factors for developing ALL, the objective of the present study is to determine if the *ABCC1* rs3743527, *NCF4* rs1883112 and *CBR3* rs1056892 genetic polymorphisms are associated with susceptibility to ALL in pediatric population. Therefore, we investigated the association between these three SNPs involved in the metabolism of antineoplastic drugs and ALL susceptibility.

METHODS

The Research Ethics Committee and the Research Committee of the General Hospital of Durango Torre Materno-Infantil, México, approved and validated the study according to the declaration of Helsinki and the General Health Law of Mexico. 71 pediatric patients of both sexes treated at the Pediatric Hemato-

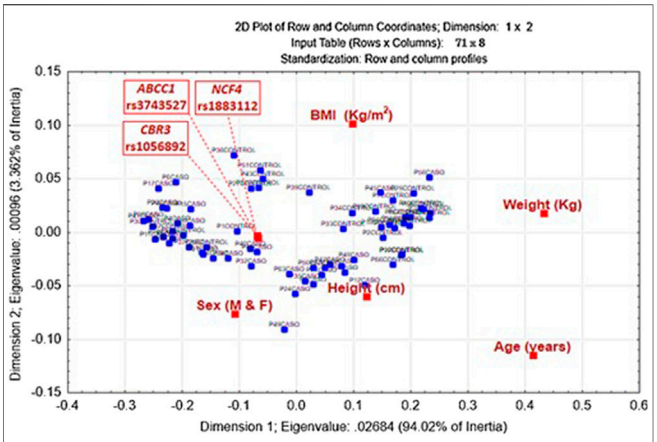


FIGURE 1 | Association of the eight variables (red box) with patients with All (blue box). The relative inertia of each variable shows that; Height (Relative Inertia: 0.610400), *NCF4* rs1883112 polymorphism (Relative inertia: 0.067791), and Sex (Relative Inertia: 0.058756) had a greater association with the development of ALL.

TABLE 1 | Demographic aspects of patients with ALL (cases) and volunteers without ALL (controls).

Variables	Cases (n = 71)	Controls =71	p
Age (years)	13.2 ± 4.8	14.1 ± 3.1	0.09
Gender (♂/♀)	45/22	28/39	
Weight (Kg)	32.2 ± 15.99	64.7 ± 13.2	<0.0005
Height (cm)	123.7 ± 26.4	164.0 ± 22.3	<0.0005
BMI (Kg/m^2)	18.5 ± 3.2	23.1 ± 3.5	<0.0005

[a]*T Student test.*

oncology Service of the State Cancer Center (CECAN) of the Secretary of Health of Durango, Mexico were studied. All patients were diagnosed with ALL according to the criteria of the Franco-American-British Hematology Association (Bennett et al., 1976). Each patient was undergoing chemotherapy treatment according to the St Jude TOTAL XV (Pui et al., 2010) protocol. Patients were also included in surveillance, that is, children who had completed their pharmacological treatment. In addition, a group of 71 adolescents without ALL was evaluated as a control group. All parents of patients were asked to sign the informed consent, in addition to children over 8 years of age, they were asked to accept the informed consent.

Genotypification

DNA was obtained from whole blood using the extraction procedure "DTAB-CTAB" (Gustincich et al., 1991), its integrity and purity was determined by horizontal electrophoresis in 1% agarose gel, stained with Texas Red and the quantification was carried out by spectrophotometry in Nanodrop® (Thermo Scientific, United States). The SNPs of the *ABCC1* rs3743527 gene (Probe number: C___8934057_30), the *NCF4* rs1883112 gene (Probe number: C__11521119_1_) and the *CBR3* rs1056892 gene (Probe number: C___9483603_10) were determined by qPCR using specific TaqMan probes for each polymorphism (by Thermo Fisher Scientific®), in a StepOne® thermocycler.

TABLE 2 | Genotypic and allelic frequencies of *NCF4* rs1883112, *CBR3* rs1056892 and *ABCC1*. rs3743527 polymorphisms in cases and controls in children with ALL.

Polymorphisms		Cases (n = 71)	Controls (=71)	p[a]	HWE[b]
NCF4 rs1883112	Genotype WT (GG	17 (24%)	18(25%)		
	HT (AG)	39(55%)	17(24%)	0.02	No
	HM (AA)	15 (21%)	36 (51%)		
	Allelos				
	Major allele (G)	73(52%)	53(37%)		
	Minor allele (A)	69(48%)	89(63%)		
CBR3 rs1056892	Genotype WT (AA)	11(15%)	17(24%)		
	HT (AG)	37(52%)	20(28%)	0.15	Yes
	HM(GG)	23 (33%)	34 (48%)		
	Allele				
	Major allele (G)	83(59%)	88(62%)		
	Minor allele (A)	59 (41%)	54 (38%)		
ABCC1 rs3743527	Genotype WT (CC)	31(44%)	26(37%)		
	HT (CT)	32(45%)	41(58%)	0.25	Yes
	HM (TT)	8 (11%)	4 (5%)		
	Allele				
	Major allele (C)	92(66%)	93(66%)		
	Minor allele (T)	48(34%)	49(34%)		

[a]*Chi square test.*
[b]*HWE (Hardy Weinberg Equilibrium).*

Statistic Analysis

The genotypic and allelic frequencies of the two groups were obtained; with ALL (cases) and without ALL (control). In addition, the Hardy-Weinberg equilibrium analysis (HWE) was performed. The software used was the SNPStats (Catalan Institute of Oncology, Barcelona, Spain) (Solé et al., 2006). Subsequently, the frequency of the alleles between the cases and the controls of each polymorphism was compared using the Chi square test (x^2), a value of $p < 0.05$ was considered statistically significant. Finally, an Odds Ratio (OR) (McHugh, 2009) risk association analysis was performed for each allele of each of the polymorphisms between the cases (with ALL) and the controls (without ALL), as well as the inheritance models in which a value p <0.05 was considered statistically significant.

RESULTS

Table 1 shows the demographic data for the two groups, patients with ALL (cases) and without ALL (controls), with a statistically

TABLE 3 | Estimation of risks by Odds Ratio between cases and controls of polymorphisms *NCF4* rs1883112, *CBR3* rs1056892 and *ABCC1* rs3743527.

Polymorphisms	Allele	Cases	Controls	OR	*IC*	*p*
	WT	17	18			
				0.92	0.4320 to 1.9892	0.84
	HM + HT	54	53			
NCF4rs1883112	HT	39	17	3.87	1.8880 to 7.9383	0.0002
	HM + WT	32	54			
	HM	15	36			
				0.26	0.1248 to 0.5434	0.0003
	WT + HT	56	35			
	WT	11	17	0.58	0.2507 to 1.3528	0.20
	HM + HT	60	54			
	HT	37	20			
CBR3rs1056892				2.77	1.3837 to 5.5651	0.004
	HM + WT	34	51			
	HM	23	34			
				0.52	0.2639 to 1.0304	0.05
	WT + HT	49	37			
	WT	31	26			
				1.34	0,6842 to 2,6295	0.39
	HM + HT	40	45			
	HT	32	41			
ABCC1rs3743527				0.60	0.3092 to 1.1657	0.13
	HM + WT	39	30			
	HM	08	04			
				2.12	0.6103 to 7.4133	0.23
	WT + HT	63	67			

significant difference between the biological variables considered, except for age ($p = 0.09$).

Table 2 shows the genotypic and allelic frequencies of patients with ALL for each polymorphism. Also as is shown, the NCF4 rs1883112 polymorphism does not find Hardy Weingber Equilibrium (EHW), while the CBR3 rs1056892 and ABCC1 rs374352 polymorphisms do.

Table 3 shows the association of WT, HT and HM genotypes with ALL through Odds Ratio values for each polymorphism. In the rs1883112 polymorphism of the *NCF4* gene, it is observed that the heterozygous allele has a risk effect for ALL (OR = 3.1870, CI = 1.8880 to 7.9383 and $p = 0.0002$), in turn the mutated genotype is associated with a protective effect (OR = 0.26, 0.1248 to 0.5434 and $p = 0.0003$). On the other hand, the *CBR3* rs1056892 polymorphism shows a significant association of risk to ALL, in the presence of the HT genotype (OR = 2.77, IC = 1.3837 and $p = 0.004$) and the mutated genotype of this polymorphism has a significant association with protection to ALL in the HM genotype (OR = 0.52, IC = 0.2639 to 1.0304 and $p = 0.05$). A significant association for any genotype in the *ABCC1* rs3743527 polymorphism was not found. Finally, the inheritance models of the studied polymorphisms are shown (**Table 4**) where it is observed that in the case of the rs1883112 polymorphism of the *NCF4* gene; the HT genotype of the codominant model shows a protective effect to ALL (OR = 0.4117, IC = 0.1718 to 0.9866 and $p = 0.04$), the recessive model shows us and confirms what we already saw in table number 3, since there is an association with protective effect on the HM genotype (OR = 0.2604, IC = 0.1248 to 0.5434 and $p = 0.0003$). In the rs1056892 polymorphism of the *CBR3* gene, a protective association was found in the heterozygous allele of the codominant model (OR = 0.3448, CI = 0.1375–0.8896 and $p = 0.0274$), in addition, the recessive inheritance model for the HM genotype shows a protective effect to ALL (OR = 0.52, IC = 0.9919 to 3.8638 and $p = 0.05$). Regarding *ABCC1* rs3743527, no statistically significant associations were found for ALL.

Figure 1 shows the multivariate analysis using the correspondence test, in this analysis it can be observed that the variables that had the greatest association with the development of ALL were 1)- height, 2)- *NCF4* rs1883112 polymorphism, and 3)- sex. The relative inertia values that indicate the level of association of each variable with the patients with ALL are shown in **Table 5**.

DISCUSSION

There are several factors that are involved in the development of ALL in pediatric patients, such factors can be chemical, environmental and genetic (Steinberg et al., 2007; Cangerana-Pereira et al., 2017), these factors can modify gene expression. Many of these genes are involved in the metabolism and transport of certain anti-leukemic drugs and may also be directly or indirectly involved in the development of cancer, as in the case of NADPH-dependent enzymes (Minotti et al., 2000; Blanco et al., 2012) and/or cellular transporters. Other cellular transporters are the ABCC1 superfamily (Zaruma-Torres et al., 2015) constituted by 48 structurally similar membrane transporters. Variants in *ABC* genes that affect gene function have clinically important effects on drug disposition and can be predictors of the risk of adverse drug reactions and efficacy of chemotherapeutics, calcium channel blockers, and protease

TABLE 4 | Association by models of inheritance of polymorphisms *NCF4* rs1883112, *CBR3*. rs1056892 and *ABCC1* rs3743527 at risk of developing LLA.

Polymorphisms		Cases(n = 71)	Controls(n = 71)	OR	CI	p
	Codominant					
	WT (GG)	17	18	1		
	HT (GA)	39	17	0.4117	0.1718 to 0.9866	0.04
	HM (AA)	15	36	2.2667	0.9256 to 5.5510	0.07
	ALLELE (G)	73	53			
NCF4 rs1883112	ALLELE (A)	69	89	1.7766	1.1068 to 2.8517	0.01
	Dominant					
	WT	17	18			
	HT + HT	54	53	0.9285	0.4320 to 1.9892	0.84
	Recessive					
	HM	15	36			
	WT + HT	56	35	0.2604	0.1248 to 0.5434	0.0003
	Codominant					
	WT (AA)	11	17	1		
	HT (AG)	37	20	0.3448	0.1375 to 0.8896	0.0274
	HM (GG)	23	34	0.4565	0.1611 to 2.2062	0.9249
CBR3	ALLELE (A)	59	54			
rs1056892	ALLELE (G)	83	88	0.3686	0.2094 to 0.6488	0.5445
	Dominant					
	WT	11	17	0.5824	0.2507 to 1.3528	0.2086
	HT + HM	60	54			
	Recessive					
	HM	23	34	0.5200	0.9919 to 3.8638	0.05
	WT + HT	49	37			
	Codominant					
	WT (CC)	31	26	1		
	HT (CT)	32	41	1.5276	0.7611 to 3.0662	0.2333
	HM (CT)	08	04	0.5962	0.1611 to 2.2062	0.4385
ABCC1	ALLELE (C)	94	93			
rs3743527	ALLELE (T)	48	49	1.0318	0.6318 to 1.6851	0.9004
	Dominant					
	WT	31	26	1.3413	0.6842 to 2.6295	0.3925
	HT + HM	40	45			
	Recessive					
	HM	08	04	2.1213	0.6103 to 7.4133	0.2361
	WT + HT	63	67			

inhibitors (Xiao et al., 2020). Kunická T. and Souček P., (Kunická and Souček 2014) mentioned that resistance to multiple drugs has one of the most important causes of cancer treatment failure, ABC transporters are membrane-bound proteins that participate in cell defense mechanisms, hence their function is to prevent toxicity as carcinogenesis on the one hand, but may contribute to resistance of tumor cells to a number of drugs, including chemotherapeutics. Zaruma-Torres F et al. (Zaruma-Torres et al., 2016) in 2016 found that the *ABCB1* rs1045642 and *ABCC5* rs3792585 polymorphisms were associated with an increased risk of ALL in Mexican children. In the current study, we determinate the impact of three gene polymorphisms involved in the metabolism of antineoplastic drugs, such as the *NCF4* rs1883112, *CBR*3 rs1056892 and *ABCC1* rs3743527 polymorphisms, we also provided frequencies of this variants and identify any association with the development of ALL in children. In 2016 Gutierrez-Salinas Torres et al. (Gutiérrez-Salinas et al., 2016), determined the concentrations of carbonylated proteins and the carbonyl reductase enzyme in Mexican women with breast cancer, finding that oxidative stress could promote cancer development and involve carbonylated proteins in the process carcinogenic, with a 3.76-fold increase in plasma carbonylated proteins in the patient group, compared to the healthy control group (5 ± 3.27 vs. 1.33 ± 2.31 nmol carbonyls/mg protein; $p < 0.05$); In addition, they found a 60% increase in the carbonyl reductase enzyme in the patients vs the control group (3.27 ± 0.124 vs. 2.04 ± 0.11 ng/mg protein; $p < 0.05$), a positive correlation was also found ($r = 0.95$; $p < 0.001$) between both measurements, which may suggest the presence of tissue damage caused by cancer. On the other hand, Osawa Y et al. (Osawa et al., 2015) had previously reported the importance of the expression levels of an enzyme of the "CBR" family, finding that the expression level of carbonyl reductase (CBR1) is related to tumor progression. Decreased CBR1 expression is associated with a poor prognosis in ovarian cancer; they also investigated the relationship between the level of CBR1 expression and the malignant potential of ovarian cancer. These results showed that the decrease in CBR1 promoted tumor proliferation and growth, as well as invasion and metastasis, suggesting that CBR1 has the potential to become a new candidate for molecular targeting therapy. Unlike the previous results, in our study, the importance of a polymorphism that modifies the sequence of the "CBR" family gene is reflected, since it was found that for the *CBR3* rs1056892 polymorphism there is a significant association of risk to ALL in presence of the HT genotype (OR = 2.7750 (IC = 1.3857–5.5651), $p = 0.004$), however, when

TABLE 5 | Relative inertias of each variable associated with the development of ALL.

	Relative inertia	*p*
Age	0.013539	0.0001[a]
Sex	0.058756	
Height	0.610400	
Weight	0.002540	
BMI	0.037701	
NCF4 rs1883112	0.067791	
CBR3 rs1056892	0.057491	
ABCC1 rs3743527	0.055791	

[a]*Chi square test*

analyzing the inheritance models we found that in the case of the codominant inheritance model, the heterozygous allele of the *CBR3* gene constitutes a protection factor according to the obtained values (OR = 0.3448 (IC = 1.375 to 0.8896), $p = 0.02$). It is evident that, although there are many factors that are involved in the development of leukemia, genetic factors play a very important role since within the genes that are involved in different cellular metabolic pathways, some of them seem to have more importance. Apparently, the folate route constitutes a very important pathway, such is the case of the work reported by Zaruma-Torres F et al. in 2016 (Zaruma-Torres et al., 2016) who determined the associations between six SNPs in four genes related to the folate transporter pathway, to determine a significant relationship with the appearance of ALL in Mexican children, concluding that certain genetic polymorphisms related to the folate transport route, particularly *COL18A1* rs2274808, *SLC19A1* rs2838956, *ABCB1* rs1045642 and *ABCC5* rs3792585, were associated with an increased risk of ALL in Mexican children. These results coincide, although to a lesser extent with what was found in our work, in which other polymorphisms are related to an increased susceptibility to ALL, such as the case of *NCF4* rs1883112, became evident since this factor NCF4 could influence the susceptibility of presenting ALL through this polymorphism, as it hinder the action of folic acid and promote the risk of developing ALL, possibly by making it difficult to transform the folic acid to THF, through the de-hydro folate reductase dDHFR which is dependent on NADPH oxidase. There is not experimental data to demonstrate its functional significance, however the rs1883112 polymorphism is found in the NCF4 promoter (p40phox), which participates in the negative regulation of NADPH oxidase (Zhang et al., 1996). In the study by Lopes LR (Lopez et al., 2004), they demonstrated that the addition of phosphorylated p40 (PHOX) to the cell-free system inhibits the NADPH oxidase activated by the C-phosphorylated protein kinase p47 (PHOX), an effect not observed with non-phosphorylated p40 (PHOX). Furthermore, phosphorylated p40 (PHOX) inhibits oxidase if added before or after full activation of the enzyme, thus they have postulated that phosphorylation of p40 (PHOX) in threonine 154 leads to an inhibitory conformation that changes the balance towards an inhibitory role and blocks oxidase activation in such a way that in our study we found that for the rs1883112 polymorphism of the *NCF4* gene, the heterozygous allele shows a risk effect for ALL (OR = 2.1267 (CI = 1.0020–4.5138), $p = 0.04$. The Mexican population has a great ethnic diversity with 68 indigenous groups (International Work Group for Indigenous Affairs, 2020). The presence of genetic bias is evidenced or ruled out through the analysis of the Hardy-Weinberg equilibrium test (EHW), in the present study it is observed as two of the three Polymorphisms are in equilibrium, however one is not in equilibrium with a predominance of the mutated allele in the control group, which was carefully selected not to include genetically related people, which suggests that perhaps there is a predominance of this mutated allele in the population studied for a possible founder effect.

CONCLUSION

In clinical practice, we can point out that a patient with an HT genotype for the polymorphisms of the CBR3 rs1056892 and NCF4 rs1883112 genes will have a higher risk of developing ALL. For the NCF4 rs1883112 gene polymorphism, the mutated allele will itself have an ALL risk effect, but two copies of it will not be additive; therefore, it is necessary for said mutated allele to bind to a copy of the wild allele (this being a heterozygous genotype) to consider it as a risk factor for ALL.

For the polymorphisms of the NCF4 rs1883112 and CBR3 rs1056892 genes, the codominant model shows us that both alleles (wild and mutated) have an association with a separate protective effect, even having two copies of these alleles still gives rise to a protective association against the ALL.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité de Ética en Investigación del Hospital General Materno Infantil de Durango, México. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

Conception: IL-A. Interpretation or analysis of data: JG-M, IL-A, JB, and VL. Preparation of the manuscript: JG-M, IL-A, JB, EE, LH, HR, LR, CS, HG, and DG. Revision for important intellectual content: IL-A, JB, and HR. Supervision: IL-A, EE, and JB.

FUNDING

To the Instituto Politécnico Nacional for project support: SIP 20200634.

REFERENCES

Adane, B., Haobin, Y., Nabilah, K., Shanshan, P., Mohammad, M., Brett, M., et al. (2019). The hematopoietic oxidase NOX2 regulates self-renewal of leukemic stem cells. *Cell Rep.* 27, 238–254. doi:10.1016/j.celrep.2019.03.009

Bennett, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D. A., Gralnick, H. R., et al. (1976). Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br. J. Haematol.* 33, 451–458. doi:10.1111/j.1365-2141.1976.tb03563.x

Blanco, J., Sun, C., Landier, W., Chen, L., Esparza-Duran, D., Leisenring, W., et al. (2012). Anthracycline-related cardiomyopathy after childhood cancer: role of polymorphisms in carbonyl reductase genes—a report from the children's oncology group. *J. Clin. Oncol.* 30, 1415–1421. doi:10.1200/JCO.2011.34.8987

Cangerana-Pereira, F. A., Mirra, A. P., Do Rosário, D. M., De Oliveira, L., and De Assunção, J. V. (2017). Environmental risk factors and acute lymphoblastic leukaemia. *Revista Ciencias de Salud.* 15, 129–144. doi:10.12804/revistas.urosario.edu.co/revsalud7a.5386

Cantarella, C. D., Ragusa, D., Giammanco, M., and Tosi, S. (2017). Folate deficiency as predisposing factor for childhood leukaemia: a review of the literature. *Genes Nutr.* 12, 1–15. doi:10.1186/s12263-017-0560-8

Ceppi, F., Cazzaniga, G., Colombini, A., Biondi, A., and Conter, V., (2015). Risk factors for relapse in childhood acute lymphoblastic leukemia: prediction and prevention. *Expert Rev. Hematol.* 8 (1), 57–70. doi:10.1586/17474086.2015.978281

Frost, A. E., Hoeper, M. M., Barberá, J. A., Jean-Luc, V., Blair, C., Langley, J., et al. (2018). Risk-stratified outcomes with initial combination therapy in pulmonary arterial hypertension: application of the REVEAL risk score. *J. Heart Lung Transplant.* 37 (12), 1410–1417. doi:10.1016/j.healun.2018.07.001

Gustincich, G., Manfidetti, G., Del Sal, C., Schneider, D., and Carninci, P. (1991). A fast method for high-quality genomic DNA extraction from whole human blood. *Biotechniques.* 11, 298–302.

Gutiérrez-Salinas, J., García-Ortiz, L., Mondragón-Terán, P., Hernández-Rodríguez, S., Ramírez-García, S., and Núñez-Ramos, N. (2016). Determinación de proteínas carboniladas y enzima carbonil reductasa en mujeres mexicanas con cáncer de mama: estudio piloto. *Gac. Med. Mex.* 152, 13–18.

International Work Group for Indigenous Affairs (2020). *El Mundo Indígena 2020: México - IWGIA*. Editor K. Wessendorf (Bolivia).

Kunická, T., and Souček, P. (2014). Importance of *ABCC1* for cancer therapy and prognosis. *Drug Metab. Rev.* 46, 325–342. doi:10.3109/03602532.2014.901348

Liu, R. T., Wang, X. X., Sun, J. R., Gao, N., and Yu, W.-Z. (2020). Correlations of IL-6 and IL-10 gene polymorphisms with childhood acute lymphoblastic leukemia. *Eur. Rev. Med. Pharmacol. Sci.* 24, 8048–8056. doi:10.26355/eurrev_202008_22488

Lopes, L. R., Dagher, M. C., Gutierrez, A., Young, B., Bouin, A.-P., Fuchs, A., et al. (2004). Phosphorylated p40PHOX as a negative regulator of NADPH oxidase. *Biochemistry.* 43, 3723–3730. doi:10.1021/bi035636s

McHugh, L. M. (2009). The odds ratio calculation, usage, and interpretation. *Biochem. Med.* 19, 120–126. doi:10.11613/BM.2009.011

Meitzler, J., Yongzhong, W., Smitha, A., Agnes, J., Han, L., Guojian, J., et al. (2014). NADPH oxidases: a perspective on reactive oxygen species production in tumor biology. *Antioxid. Redox Sign.* 20, 2873–2889. doi:10.1089/ars.2013.5603

Minotti, G., Licata, S., Saponiero, A., Menna, P., Calafiore, A. M., Di Giammarco, G., et al. (2000). Anthracycline Metabolism and toxicity in human myocardium: comparisons between doxorubicin, epirubicin, and a novel disaccharide analogue with a reduced level of formation and [4Fe-4S] reactivity of its secondary alcohol metabolite. *Chem. Res. Toxicol.* 13, 1336–2134. doi:10.1021/tx000143z

Muwakkit, S., Al-Aridi, C., Samra, A., Saab, R., Mahfouz, R. A., Farra, C., et al. (2012). Implementation of an intensive risk-stratified treatment protocol for children and adolescents with acute lymphoblastic leukemia in Lebanon. *Am. J. Hematol.* 87 (7), 678–683. doi:10.1002/ajh.23222

Navarro-Pérez, S., Mayorquín-Galván, E., Petarra-Del Río, S., Casas-Castañeda, M., Romero-Robles, B., Torres-Bugarín, O., et al. (2016). El ácido fólico como citoprotector después de una revisión. *El Resid.* 11 (2), 51–59.

Osawa, Y., Yokoyama, Y., Shigeto, T., Futagami, M., and Mizunuma, H. (2015). Decreased expression of carbonyl reductase 1 promotes ovarian cancer growth and proliferation. *Int. J. Oncol.* 46, 1252–1258. doi:10.3892/ijo.2014.2810

Pérez-Saldívar, M. L., Fajardo-Gutiérrez, A., Bernáldez-Ríos, R. M., Martínez-Avalos, A., Medina-Sanson, A., Espinosa-Hernández, L., et al. (2011). Childhood acute leukemias are frequent in Mexico City: descriptive epidemiology. *BMC Canc.* 11, 355. doi:10.1186/1471-2407-11-355

Pui, C. H., Pei, D., Sandlund, J. T., Ribeiro, R. C., Rubnitz, J. E., Raimondi, S. C., et al. (2010). Long-term results of St Jude total therapy studies 11, 12, 13A, 13B, and 14 for childhood acute lymphoblastic leukemia. *Leukemia.* 24, 371–382. doi:10.1038/leu.2009.252

Rivera-Luna, R., Cárdenas-Cardos, R., Olaya-Vargas, A., and Shalkow-Klincovstein, J. (2005). El niño de población abierta con cáncer en México. Consideraciones epidemiológicas. *An. Med. (Mex).* 60 (2), 91–97.

Rodriguez, L., Gonzalez-Llano, O., Mancias, C., Pompa, T., Gonzalez, G., Sandoval, A., et al. (2010). Observaciones sobre la incidencia de leucemias agudas en el noreste de México. *Rev. Hematol. Mex.* 11 (2), 78–81.

Secretaría de Salud (2011). *Subsecretaría de Prevención y Promoción de la Salud. Dirección General de Epidemiología 2011. Perfil Epidemiológico de Cáncer en Niños y Adolescentes*. Ciudad de México, 155.

Słomka, M., Sobalska-Kwapis, M., and Korycka-Machała, M. (2015). Genetic variation of the ABC transporter gene *ABCC1* (Multidrug resistance protein 1 – MRP1) in the Polish population. *BMC Genet.* 16, 114. doi:10.1186/s12863-015-0271-3

Solé, X., Guinó, E., Valls, J., Iniesta, R., and Moreno, V. (2006). SNPStats: a web tool for the analysis of association studies. *Bioinformatics.* 22, 1928–1929. doi:10.1093/bioinformatics/btl268

Steinberg, K. K., Relling, M., Gallagher, M., Greene, C., Rubin, C. S., and French, D. (2007). Genetic studies of a cluster of acute lymphoblastic leukemia cases in churchill county, Nevada. *Environ. Health Perspect.* 115, 158–164. doi:10.1289/ehp.9025

Vrooman, L. M., and Silverman, L. B. (2009). Childhood acute lymphoblastic leukemia: update on prognostic factors. *Curr. Opin. Pediatr.* 21 (1), 1–8. doi:10.1097/MOP.0b013e32831f1f24

Xiao, Q., Zhou, Y., and Lauschke, V. M. (2020). Ethnogeographic and inter-individual variability of human ABC transporters. *Hum. Genet.* 139, 1–24. doi:10.1007/s00439-020-02150-6

Zaruma-Torres, F., Lares-Asseff, I., Lima, A., Reyes-Espinoza, A., Loera-Castañeda, V., and Sosa-Macías, M. (2016). Genetic polymorphisms associated to folate transport as predictors of increased risk for acute lymphoblastic leukemia in Mexican children. *Front. Pharmacol.* 7, 238. doi:10.3389/fphar.2016.00238

Zaruma-Torres, F., Lares-Asseff, I., Reyes-Espinoza, A., Loera-Castañeda, V., Chairez-Hernández, I., Sosa-Macías, M., et al. (2015). Association of ABCB1, ABCC5 and Xanthine oxidase genetic polymorphisms with methotrexate adverse reactions in Mexican pediatric patients with ALL. *Drug Metab. Pers. Ther.* 30, 195–201. doi:10.1515/dmpt-2015-0011

Zhan, S., Vazquez, N., Zhan, S., Wientjes, F. B., Budarf, M. L., Schrock, E., et al. (1996). Genomic structure, chromosomal localization, start of transcription, and tissue expression of the human p40-phox, a new component of the nicotinamide adenine dinucleotide phosphate-oxidase complex. *Blood.* 88, 2714–2721. doi:10.1182/blood.V88.7.2714.bloodjournal8872714

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 10 March 2021
doi: 10.3389/fphar.2021.602676

Pharmacogenetics–Based Preliminary Algorithm to Predict the Incidence of Infection in Patients Receiving Cytotoxic Chemotherapy for Hematological Malignancies: A Discovery Cohort

Matias F. Martinez[1,2,3]*, *Enzo Alveal*[1], *Tomas G. Soto*[1,4], *Eva I. Bustamante*[5], *Fernanda Ávila*[6], *Shrikant I. Bangdiwala*[7,8], *Ivonne Flores*[5], *Claudia Monterrosa*[5], *Ricardo Morales*[5], *Nelson M. Varela*[1,3], *Alison E. Fohner*[9] *and Luis A. Quiñones*[1,3]*

[1]*Laboratory of Chemical Carcinogenesis and Pharmacogenetics (CQF), Department of Basic and Clinical Oncology (DOBC), Faculty of Medicine, University of Chile, Santiago, Chile,* [2]*Departamento de Ciencias y Tecnología Farmacéuticas, Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile, Santiago de Chile, Chile,* [3]*Latin American Network for the Implementation and Validation of Pharmacogenomic Clinical Guidelines (RELIVAF-CYTED), Madrid, Spain,* [4]*Departamento De Ciencias Básicas Santiago, Facultad De Ciencias, Universidad Santo Tomás, Santiago, Chile,* [5]*Cancer Institute Arturo López Pérez Foundation, Santiago, Chile,* [6]*Clinical Hospital of the University of Chile, Santiago, Chile,* [7]*Population Health Research Institute, McMaster University, Hamilton, ON, Canada,* [8]*Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada,* [9]*Department of Epidemiology and Institute of Public Health Genetics, University of Washington, Seattle, WA, United States*

Introduction: Infections in hematological cancer patients are common and usually life-threatening; avoiding them could decrease morbidity, mortality, and cost. Genes associated with antineoplastics' pharmacokinetics or with the immune/inflammatory response could explain variability in infection occurrence.

Objective: To build a pharmacogenetic-based algorithm to predict the incidence of infections in patients undergoing cytotoxic chemotherapy.

Methods: Prospective cohort study in adult patients receiving cytotoxic chemotherapy to treat leukemia, lymphoma, or myeloma in two hospitals in Santiago, Chile. We constructed the predictive model using logistic regression. We assessed thirteen genetic polymorphisms (including nine pharmacokinetic—related genes and four inflammatory response-related genes) and sociodemographic/clinical variables to be incorporated into the model. The model's calibration and discrimination were used to compare models; they were assessed by the Hosmer-Lemeshow goodness-of-fit test and area under the ROC curve, respectively, in association with Pseudo-R^2.

Results: We analyzed 203 chemotherapy cycles in 50 patients (47.8 ± 16.1 years; 56% women), including 13 (26%) with acute lymphoblastic and 12 (24%) with myeloblastic leukemia.

Pharmacokinetics-related polymorphisms incorporated into the model were *CYP3A4* rs2242480C>T and *OAT4* rs11231809T>A. Immune/inflammatory response-related

Edited by:
José A. G. Agúndez, Institute of Molecular Pathology Biomarkers, University of Extremadura, Spain

Reviewed by:
Jatinder K. Lamba, College of Pharmacy, University of Florida, United States
Pedro Ayuso Parejo, Hospital Universitario Infanta Leonor, Spain

***Correspondence:**
Matías F. Martínez
matiasmartinez@ug.uchile.cl
Luis A. Quiñones
lquinone@uchile.cl

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *04 September 2020*
Accepted: *20 January 2021*
Published: *10 March 2021*

Citation:
Martinez MF, Alveal E, Soto TG, Bustamante EI, Ávila F, Bangdiwala SI, Flores I, Monterrosa C, Morales R, Varela NM, Fohner AE and Quiñones LA (2021) Pharmacogenetics–Based Preliminary Algorithm to Predict the Incidence of Infection in Patients Receiving Cytotoxic Chemotherapy for Hematological Malignancies: A Discovery Cohort. Front. Pharmacol. 12:602676. doi: 10.3389/fphar.2021.602676

polymorphisms were *TLR2* rs4696480T>A and *IL-6* rs1800796C>G. Clinical/demographic variables incorporated into the model were chemotherapy type and cycle, diagnosis, days in neutropenia, age, and sex. The Pseudo-R^2 was 0.56, the *p*-value of the Hosmer-Lemeshow test was 0.98, showing good goodness-of-fit, and the area under the ROC curve was 0.93, showing good diagnostic accuracy.

Conclusions: Genetics can help to predict infections in patients undergoing chemotherapy. This algorithm should be validated and could be used to save lives, decrease economic costs, and optimize limited health resources.

Keywords: pharmacogenetics, hematological malignancies, infections, Prediction, Algorithm, pharmacogenomics, CYP3A4, OAT4

INTRODUCTION

Infections in patients diagnosed with hematological malignancies are common and usually life-threatening (Green 2017); 30% of patients undergoing chemotherapy for the treatment of these illnesses experience infection, and 11% dying as a result (Castagnola et al., 2007; Yilmaz et al., 2008; Dutronc et al., 2009). Several variables have been described as risk or protector factors. In addition to clinical variables associated with a high risk of infection, neutropenia, defined as an absolute neutrophil count lower than 500 cells/mm^3 (Klastersky et al., 2016), is common in patients undergoing cytotoxic chemotherapy (Li et al., 2016). Genetics can modify the occurrence of chemotherapy-related neutropenia. A polymorphic variant can alter the metabolism or elimination of the cytotoxic agent; this could increase the plasmatic level of the antineoplastic, and therefore increase the risk of dose-related toxicity (Lyman et al., 2014; Buaboonnam et al., 2019). Among these pharmacokinetics-related genetic factors, CYP3A4 and CYP3A5 are two enzymes that participate in the metabolism of most drugs used in the treatment of hematological malignancies (Lee et al., 2013; Daly 2015), and polymorphic variants in the genes that encode those proteins are known to decrease their expression or functionality, decreasing drug metabolism (Lamba et al., 2012; Park et al., 2014; Zhu et al., 2014). Moreover, polymorphic variants in genes that codify for drug transporters could affect the elimination of antineoplastic medications (Choi et al., 2015). Some of the transporters associated with cancer medications are ATP-Binding Cassette (ABC), specifically ABCB1, ABCC2, and ABCG2 (Jedlitschky et al., 2006; Vasiliou et al., 2009), and Solute Linking Carrier (SLC), specifically *SLC22A11* that codifies for the Organic Anion Transporter 4 (OAT4) (Burckhardt 2012). Genetic variants in these genes could increase the risk of adverse events due to decreased elimination leading to toxic drug concentrations (Vormfelde et al., 2006; Lewis et al., 2013).

Furthermore, some immune/inflammatory response-related proteins can improve the aggregation and survival of neutrophils. Interleukin 6 and 1β (IL6 and IL-1β) participate in the maturation and apoptosis inhibition of white blood cells (WBC) (Rose-John et al., 2017; Chiba et al., 2018). A lower expression of the genes that encode these interleukins could enhance the risk of neutropenia, make it more severe or prolong the length of neutropenia (Wright et al., 2014; Loft et al., 2018; Badawy et al., 2019). The caspase recruitment domain 8 (CARD8) participates in the activation of IL-1β, and a polymorphism that affects CARD8 functionality could have the same consequences of having less interleukin signal (Paramel et al., 2015). Toll-like receptor 2 (TLR2) is a protein that senses pathogen molecules and develops the intracellular signaling in response to a possible infection (Beutler et al., 2006); a genetic variant that affects the functionality of this receptor could increase the risk of infections (Hawn et al., 2009; Bielinski et al., 2011; Esposito et al., 2014).

No tool can predict the incidence of infections in hematological cancer patients in chemotherapy that use clinical and genetic variables. By accurately predicting which patients are likely to develop an infection while undergoing chemotherapy, dose adjustments and enhanced monitoring could be targeted to prevent infection, thereby reducing morbidity, mortality, and healthcare costs. Our study aims to create an algorithm to predict the incidence of infections among patients undergoing chemotherapy, using pharmacokinetics and immune response-related genetic polymorphisms in addition to clinical variables.

MATERIALS AND METHODS

Study Design

We carried out a prospective cohort study from November 2017 to October 2018 at the Oncologic Hospital "Fundación Arturo Lopez Pérez" (FALP) and the Clinical Hospital of the University of Chile (HCUCH) in Santiago, Chile. Patients were enrolled at those clinical centers before the first chemotherapy cycle, and they were followed prospectively through every cycle of chemotherapy. Infection was established by clinical criteria, including fever not explained by chemotherapy, positive bacterial cultures, or compatible imaging. A multidisciplinary team compound by specialist physicians and clinical pharmacists reviewed every febrile episode and decided if it was due to infection or not according to local practice. The occurrence of any infections was recorded.

Patients and Data Source

The study included patients 18 years or older diagnosed with leukemia or non-Hodgkin's lymphoma and undergoing cytotoxic chemotherapy. We excluded patients regularly taking immunosuppressive medication, pregnant women, and patients with a diagnosis of immunodeficiency. All clinical data were obtained from the medical record. We collect the information that allows clinic, pharmacotherapeutic, morbid, and demographic characterization of the sample.

Ethics Statement

All patients signed a written informed consent and an agreement to participate in this study. The study was carried out following the strict ethical procedures recommended by the Ethics Committee of the Clinical Hospital of the University of Chile (approval received on July 18, 2017) and the Eastern Metropolitan Health Service (approval received on July 4, 2017), following the procedures suggested in the Declaration of Helsinki, with Chilean Laws 20.120, 20.584, and 19.628 and with the guidelines of the Good Clinical Practices from the World Health Organization.

Genotyping Analysis

Genomic DNA was isolated from the subjects' peripheral blood samples using the High Pure PCR Template Preparation Kit (Catalog Number, 11796828001; Roche Diagnostics GmbH, Mannheim, Germany). The blood sample was collected after the first remission.

The nine pharmacokinetics-related polymorphisms were *SLC22A11* rs11231809; *ABCB1* rs2032582, rs1045642y rs1128503; *CYP3A4* rs2740574, rs2242480, *CYP3A5* rs15524, *ABCC2* rs12762549, and *ABCG2* rs2231142. The immune response-related polymorphisms were *IL6* rs1800796, *IL1*β rs1143627, *CARD8* rs2043211, and *TLR2* rs4696480. The potential effect in protein was the main criteria to pick the polymorphic variants in addition to the relationship with drug toxicity reported in the literature. Another factor considered was the minor allele frequency of more than 5% in the Latinamerican population when available.

All polymorphisms were analyzed using *TaqMan*® SNP Genotyping Assay (Catalog number, 4362691; Thermo Fisher Scientific, Waltham, MA, United States) in a Stratagene Mx3000p real-time PCR system (Agilent Technologies, Santa Clara, CA, United States). Every sample was analyzed in triplicate to ensure reliability. The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: [10.6084/m9.figshare.13444211].

Statistical Analysis

To determine which genetic variants were associated with infection incidence, it was a binary variable, and it was assessed in every chemotherapy cycle. We carried out multivariate logistic regression models to establish the relationship between every polymorphism by itself and the outcome. We assessed the inheritance in a dominant, recessive, and co-dominant model, and they were added depending on their statistical significance. These models included one genetic variable and the following control variables: sex; age (in groups: from 18 to 40 years; from 41 to 60 years and older than 60 years); the number of chemotherapy cycle; kind of chemotherapy scheme (induction or consolidation), diagnosis (acute lymphoblastic leukemia, acute myeloblastic leukemia, other leukemias or lymphoma) and the number of days in profound neutropenia, defined as an absolute neutrophil count of 0 cells/mm^3.

TABLE 1 | Sample characteristics.

Characteristics	n = 50 (%)
Age, years (median ± IQR)	40 (31–52)
18–40 years	17 (34)
41–60 years	20 (40)
>60 years	13 (26)
Female	27 (54)
Male	23 (46)
Diagnosis	
Acute lymphoblastic leukemia	13 (26)
Acute myeloblastic leukemia	12 (24)
Other leukemias	5 (10)
Lymphoma	20 (40)

Adjusted Odds Ratios (OR) were obtained from exponentiating the coefficient given by the regression model. The ORs were used to determine which variables were risk or protective factors of having an infection.

To develop the final model, we added all the control variables and those genetic factors that had a statically significant association with infection in the previous step. The best model was chosen according to the value of Pseudo-R^2, the calibration, and the discrimination of the model. The control variables were chosen based on the univariate relationship or previous reports of the association's association with the event.

We use a mixed-effect model to account for intraindividual variability regarding chemotherapy scheme and cycle and use correlated outcome data (also known as hierarchical models). Here we used a three-level model, so we had a random intercept for three characteristics. The first one was the patient level, the second chemotherapy scheme, and finally cycle level, where infections were assessed.

Calibration is the degree of similarity between the probability given by the model and the observed incidence. The goodness-of-fit test of Hosmer—Lemeshow compares frequencies of cases and controls using a chi2 test. In this case, a higher p-value indicates fewer differences between the predicted frequencies and the frequencies observed in the sample.

On the other hand, discrimination is the degree to which the proposed model can distinguish between patients who experience the event from those who do not; that is, the ability to indicate that a patient will experience an infection, and, on the other hand, the ability to predict when the patient will not experience it. Discrimination between cases and non-cases of infection was tested using the area under the Receiver Operating Characteristic curve (AUC-ROC).

TABLE 2 | Allele and genotypic frequencies of the studied polymorphism.

Genetic polymorphism	Allele frequency		Genotypic frequency, n (%)		
CARD8 rs2043211	A	T	A/A	A/T	T/T
	0,64	0,36	19 (39)	25 (51)	5 (10)
TLR2 rs4696480	T	A	T/T	T/A	A/A
	0,63	0,38	20 (42)	20 (42)	8 (17)
IL-6 rs1800796	G	C	G/G	G/C	C/C
	0,50	0,50	11 (22)	28 (56)	11 (22)
IL-1β rs1143627	G	A	G/G	G/A	A/A
	0,54	0,46	16 (33)	21 (43)	12 (24)
OAT4 rs11231809	T	A	T/T	T/A	A/A
	0,38	0,62	11 (24)	13 (28)	22 (48)
ABCB1 rs2032582	C	A	C/C	C/A	A/A
	0,68	0,32	21 (53)	12 (30)	7 (18)
ABCB1 rs1045642	A	G	AA	AG	GG
	0,33	0,67	3 (6)	27 (54)	20 (40)
ABCB1 rs1128503	A	G	AA	AG	GG
	0,38	0,62	7 (14)	24 (48)	19 (38)
CYP3A4 rs2740574	T	C	T/T	T/C	C/C
	0,83	0,17	31 (74)	8 (19)	3 (7)
CYP3A4 rs2242480	C	T	CC	CT	TT
	0,71	0,29	29 (59)	12 (24)	8 (16)
CYP3A5 rs15524	A	G	AA	AG	GG
	0,84	0,16	37 (74)	10 (20)	3 (6)
ABCC2 rs12762549	G	C	G/G	G/C	C/C
	0,62	0,38	19 (42)	18 (40)	8 (18)
ABCG2 rs2231142	G	T	G/G	G/T	T/T
	0,90	0,10	37 (80)	9 (20)	(0)

Some samples could not be fully genotyped.

The sample characteristics were mainly presented as proportions. The Shapiro–Wilk test was used to assess the normality of the distribution of continuous variables, and when the distribution was skewed, the variable was presented as median and interquartile range. All the analyses and figures were performed using STATA 15.0 software ®.

RESULTS

Sample Characteristics

Participants' median age was 40 years (IQR 31–52), and 27 (54.0%) were women. Half of the patients had acute leukemia [13 (26%) lymphoblastic and 12 (24%) myeloblastic] (**Table 1**). We detected infection in 82 (40.8%) of 203 total chemotherapy cycles, including 31(62.0%) of the 50 patients recruited. No patients died during the follow-up period. The genotype distribution of the thirteen polymorphisms assessed in the study is given in **Table 2**.

CONSTRUCTION OF THE PREDICTIVE MODEL

Model With Non-genetic Variables

To assess how genetics can improve the predictive model's performance, we carried out primarily a non-genetic algorithm. For this, we incorporated chemotherapy type and cycle, diagnosis, days in neutropenia, age, and sex. The *p*-value of the Hosmer – Lemeshow test was 0.29, the AUC-ROC was 0.82, and the Pseudo-R^2 was 0.23.

Model With Genetic Variables

We found that Toll-Like Receptor 2 (*TLR2*), Interleukin 6 (*IL6*), CYP3A4, Solute Linking Carrier family 22 member 11 (*SLC22A11*) or Organic Anion Transporter 4 (OAT4), and ATP-Binding Cassette Subfamily C member 2 (*ABCC2*) polymorphisms were associated with the incidence of infection by themselves (**Supplementary Table S1**), and they were added to the final model.

In **Figure 1**, we summarize the OR obtained from the adjusted final model. There we found that TLR2, IL6, OAT4, and CYP3A4 were significantly associated with the occurrence of infection. The Pseudo R2 of the model was 0.5327; i.e., the model explained 53% of the variability of the incidence of infections.

The multilevel model with a random intercept for patient, chemotherapy scheme, and cycle variables gave the same calibration and discrimination performance. As we had a predictive objective and allowing a more straightforward application in clinical practice, we construct the model only with fixed intercepts, i.e., a "normal" logistic regression.

Calibration of the Model

The *p*-value of the test of Hosmer—Lemeshow was 0.9516; this means that the model had an excellent goodness-of-fit or, in other words, there is no statistical difference between the frequency of cases detected by the model and that observed in the sample.

Discrimination of the Model

The AUC-ROC indicates the probability of assigning a higher probability of incidence to patients who develop an infection than those who do not. The ROC curve compares sensitivity, that is, the ability to assign the event correctly when it occurred, and one- specificity, that is, the proportion of patients who did not have an infection, but the model indicated that they would present it.

The AUC-ROC curve can range from 0 to 1, with 0.5 indicating that the instrument cannot discriminate outcomes better than chance. A value of one would be for a model that perfectly predicts the occurrence of the event. For our proposed model, the AUC-ROC curve was 0.93. **Figure 2** shows the ROC curve for the model to predict the incidence of infections in the sample. Concerning non-genetic variables, we found that the number of chemotherapy cycles was a protective factor. This result could be interpreted as follows: with each cycle of chemotherapy that passes, the risk of infection decreases, perhaps because if the patient had an infection in an initial cycle, in the following cycles, the prophylaxis would be optimized to avoid futures events. Besides, the male sex was protective compared to the female sex; induction was associated with more risk than consolidation, and the longer neutropenia episodes were associated with more risk of infection.

The coefficients for every variable obtained from the final regression model allow us to build an equation to predict the

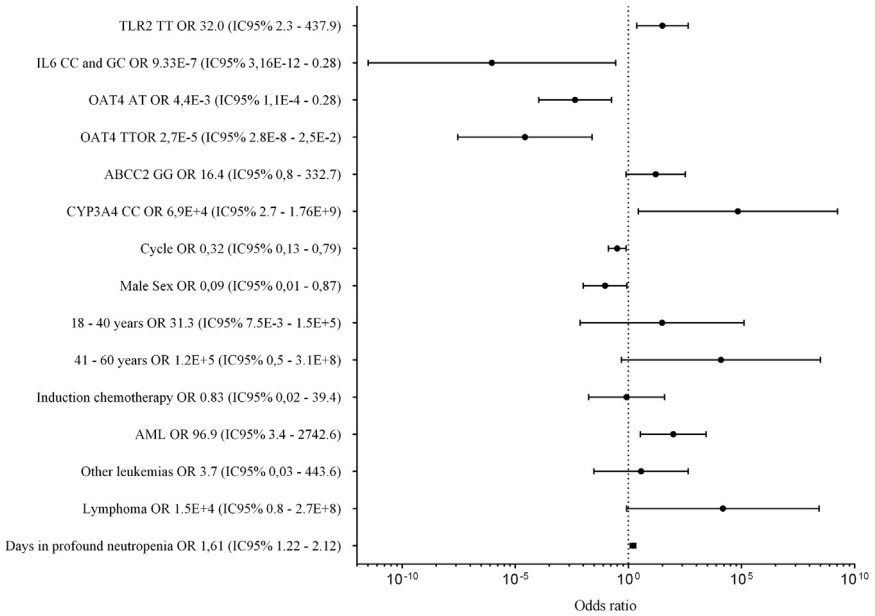


FIGURE 1 | Genetic and non-genetic factors associated with infection in patients treated with cytotoxic chemotherapy. The figure presents the point estimate of the odds ratio for infection, and the bars around each point show the 95% confidence interval.

probability of infection occurrence. Thus, the model, including genetic and non-genetic variables, is:

$$\ln\left(\frac{p}{1-p}\right) = 3.47 \times (TLR2\ TT) - 13.9 \times (IL6\ GG) 5.4 \times (OAT4\ AT) - 10.5 \times (OAT4\ TT) + 2.8 \times (ABCC2\ GG) + 11.1 \times (CYP3A4\ CC) - 1.1 \times (Cycle) - 2.4 \times (Male\ sex) + 3.4 \times (From\ 18\ to\ 40\ years) + 9.4 \times (From\ 41\ to\ 60\ years) + 0.2 \times (Induction) + 4.6 \times (AML) + 1.3 \times (Other\ leukemias) + 9.6 \times (Lymphoma) + 0.5 \times (Days\ in\ deep\ neutropenia) - 1.9$$

where p is the probability of occurrence of infection, with values ranging from 0 to 1 (or 0%–100% of occurrence probability). Some factors increased the odds of having an infection (positive factors), and others decreased the odds of the event (negative factor), according to the OR.

For genetics, the model works with dichotomous variables (i.e., 0 and 1). Thus when the variant showed in the model is present, we should use a 1, and when it is another genotype, we should use a 0. For example, in the OAT variant, if the genotype TT were determined, the coefficient for OAT AT would be multiplied by 0 (5.4 × 0) because the patient presents another genotype, and the coefficient for TT genotype should be multiplied by 1 (10.5 × 1). Alternatively, if the genotype AA were determined, the coefficients for TT and AT should be multiplied by 0. Days in neutropenia should be used as a continuous variable. Meanwhile, the days in neutropenia increase, also the risk of infection does.

DISCUSSION

Infections are frequent and potentially lethal events in patients treated with cytotoxic chemotherapy for hematological malignancies. Neutropenia due to the medication and the illness by itself is common in these patients, and it is one of the more critical factors in the incidence of infections. This study

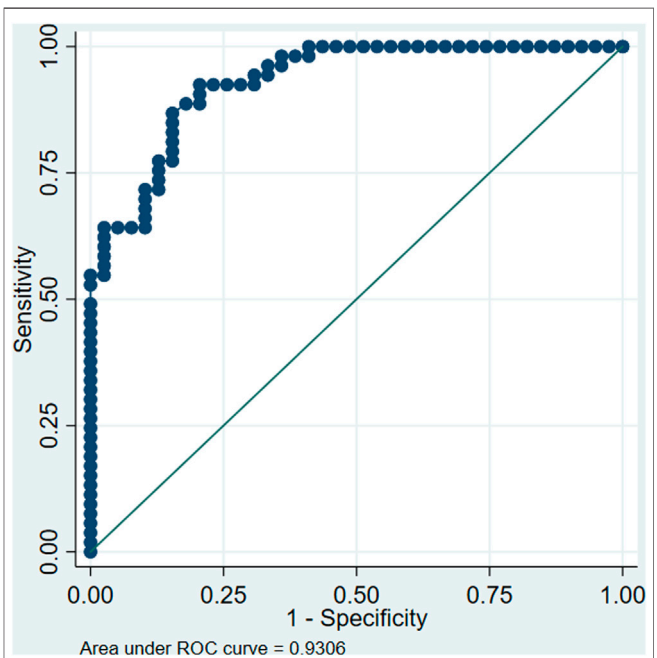


FIGURE 2 | Receiver Operating Characteristic (ROC) curve for the model proposed to predict infections in patients undergoing cytotoxic chemotherapy. (References TLR2 TA/AA, IL6 GG, OAT4 AA, ABCC2 CC/CG, CYP3A4 TT and CT, Female sex, 60 years or older, Consolidation chemotherapy, Acute Lymphoblastic Leukemia, respectively for each category).

is the first that aims to predict their occurrence using genes associated with both antineoplastic pharmacokinetics and immune response.

Relationship Between Genetics and the Incidence of Infections

Concerning immune response-related genes, we found two polymorphisms significantly associated with infections. One of them was *IL6* rs1800796 (−572C>G), which is a variant located in the promoter of the gene and where the G allele is related to a lower expression of the protein (Sharma et al., 2018). We found that the GG genotype was a risk factor for having an infection, agreeing with previous reports (Tang et al., 2014; Badawy et al., 2019), probably due to a lower neutrophil mobilization and activation (Wright et al., 2014; Rose-John et al., 2017).

The other immune response-related polymorphism associated with infections was *TLR2* rs4696480 (c.-373+1614T>A). Although the variant's effect at the protein level is not well established, the T allele has been associated with lowering receptor functionality (Loft et al., 2018), a decreased function of *TLR2* prevents the pathogen recognition and delay the immune response (Hawn et al., 2009). We found the TT genotype was a risk factor of having an infection compared to TA and AA genotypes as we expected because the *TLR2* receptor also promotes neutrophil recruitment and survival (Malcolm et al., 2003; Sabroe et al., 2005).

Concerning pharmacokinetics–related polymorphisms, *SLC22A11* rs11231809 polymorphism was associated with the risk of infection. We found that the genotype TT and AT are protector factors compared to the AA genotype. This finding is consistent with previous reports describing that the A allele is associated with a lower functionality of the protein OAT4. This variant has been associated with a decreased clearance of some drugs (Vormfelde et al., 2006; Lima et al., 2015), so perhaps a lower elimination leads to a higher plasmatic level of antineoplastic that causes dose-related toxicities.

We found the polymorphic variant rs2242480 in the gene *CYP3A4* that encodes the biotransformation enzyme CYP3A4 was related to the incidence of infection. The allele CC genotype has been related to higher levels of some drugs (Zhu et al., 2014), and we found that the CC genotype was associated with an increased risk of infections, so this could be due to higher plasmatic levels of antineoplastic and, therefore, a higher probability of adverse events.

The Usefulness of the Model

We can calculate the probability of having an infection by combining all factors and coefficients in the model we developed. Nevertheless, the clinical action taken due to the calculated value should be consistent with the clinical objective. An important issue is determining the output probability over which the patient will be considered as a probable case or not. Depending on the probability cutoff selected, the model's sensibility and specificity would change, and it is relevant to ponder which indicator will be better for clinical outcomes, workflows, and resources.

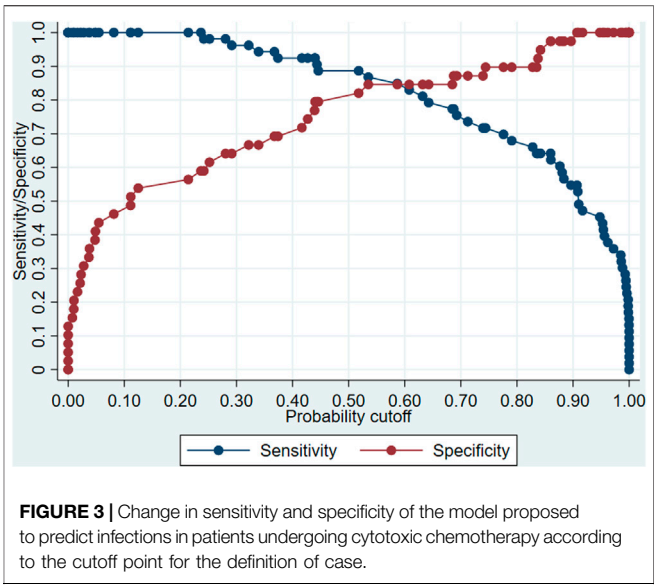


FIGURE 3 | Change in sensitivity and specificity of the model proposed to predict infections in patients undergoing cytotoxic chemotherapy according to the cutoff point for the definition of case.

Figure 3 shows the change of sensitivity (blue line) and specificity (red line) of the proposed model according to different cutoffs. If a balance between sensitivity and specificity is sought, a better cutoff point would be 0.6 (that is, above 0.6, the patient would be classified as having an infection and under the value, as without infection). This cutoff is the intersection point of curves, where both specificity and sensitivity are in values around 85%.

If it is more important that all or a large part of the possible cases are correctly classified, sensitivity should be privileged, and a lower probability cutoff should be chosen. For example, with a cutoff point of 0.3, the sensitivity is 100%, which ensures that most cases (true positives) are detected, with the disadvantage that some patients who will not suffer an infection will be misclassified (false positives). In this example, the model could be used to give medical discharge to those patients who have a low risk or probability of infection.

On the other hand, if it is sought to correctly classify all patients (or the vast majority) who will not suffer an infection (true negatives), a higher cutoff point should be set. For example, by setting a cutoff point of 0.9, we ensure that no patient (or very few) who will not suffer an infection are classified as a case. The disadvantage of this approach is that the rate of those classified as non-cases would increase, even though they will present an infection (false negatives). In this example, the model could help patients at high risk of infection be referred to an isolation room or receive optimized antimicrobial prophylaxis from the start of chemotherapy since the model indicates the patient has a high probability of suffering an infection.

Concerning the performance of the proposed model, the value of the AUC-ROC is high. With an AUC-ROC value over 0.8, the model makes a proper classification of people who will have the event and who will not (Harrell, 2015). The Hosmer–Lemeshow test *p*-value indicates that the frequency of cases predicted by the model was similar to our sample's observed values.

Several studies have tried to predict the incidence of infections in patients with hematological malignancies, but no one has combined genetic and clinical factors in a single model. Webb et al. created an algorithm to predict the incidence of bloodstream infections due to vancomycin-resistant *enterococcus* in patients undergoing leukemia induction and included severe neutropenia as one of the factors adding to the use of some antimicrobials previously; they found an AUC-ROC curve of 0.84 (Webb et al., 2017).

Schalk et al. used a modified Infection Probability Score (mIPS) to predict the incidence of central venous catheter-related bloodstream infections in patients with hematological malignancies. The mIPS includes clinical variables such as the heart and respiratory rates and also the WBC count. They compared the patients' score at the moment of catheter insertion and removal; the AUC-ROC curve was 0.77 (Schalk et al., 2015). Apostolopoulou et al. included in their model the presence and length of neutropenia, and similarly to what we found with longer neutropenia associated with more risk of infection, they also identified chemotherapy as a risk factor (Apostolopoulou et al., 2010).

Non-Genetic Factors Associated With the Risk of Infections

Other risk factors we identified as associated with the risk of infection were acute lymphoblastic leukemia diagnosis compared to acute myeloblastic leukemia and the number of days in profound neutropenia. The first factor could be due to differences in treatment and even the illness's pathophysiology (Sung et al., 2007; Chandran et al., 2012; Inaba et al., 2017). An absolute neutrophil count lower than 500 cells/mm3 is one of the more significant factors associated with infection, so the longer neutropenia lasts, the higher the risk (Gil et al., 2007). Male sex seems to be a protective factor. It could be because women are more susceptible to hematological toxicities during chemotherapy, increasing the risk of infection (Singh et al., 2005), making them more likely to have a bloodstream infection with hematology-oncology illness (Apostolopoulou et al., 2010). Additionally, women often have smaller body size than men and may be more likely to experience chemotoxicity.

It is necessary to consider non-genetic variables in the analysis because they can explain the intraindividual variability of infection, having genetics unchanged. Differences could be attributed to changes in the chemotherapy scheme or doses, different use of G-CSF or prophylaxis, and other unmeasured variables as food, visits, or bed availability. With this model, we can better understand some of the factors (genetic or not) involved in chemotherapy response.

Limitations

This study aimed to create a preliminary algorithm and identify genetic and clinical variables associated with infection risk in patients undergoing chemotherapy. Nevertheless, these results could be used as a base for new studies. Because of the small sample size, we could not split our cohort into a training and test cohort for the model development. As a result, the performance metrics of our prediction model likely reflect a degree of overfitting. The prediction algorithm should not be directly applied to clinical practice. Future studies are needed to validate the model and set cutoff decision parameters to improve clinical care and outcomes.

One of the proposed model's main limitations is that it uses the number of days in profound neutropenia as an explanatory variable and other data obtained directly from the clinical record. However, if we want to use this prediction model at the beginning of the treatment cycle, we should use the expected number of days, causing the model to lose accuracy because it would be an approximate number, not the actual length of neutropenia. A solution to this problem is to generate a second model to predict neutropenia duration and feed its results into the previous model to predict infection risk.

Due to the sample size, we decided not to incorporate some pharmacotherapeutic variables, mainly the specific chemotherapy scheme or the use of granulocyte colony-stimulating factors. In a larger sample, these variables should be included to improve the precision of the model. Besides, other possible not measured explanatory variables, such as the severity of the illness or the gut microbiome, may influence infection risk (Hakim et al., 2018).

The sample size should be analyzed in the context of a developing way country, with a population of 18 million and an incidence of hematological malignancies of 15 cases per 100,000 habitants/year. Also, the prospective character of the study makes the rate of recruitment similar to the incidence rate. This sample size means almost all the new cases in a year in two hospitals in Chile. Although we did not reach a power to discard other polymorphisms, we got the confidence of four polymorphism were related to the event.

CONCLUSION

CYP3A4 rs2242480C>T, *SLC22A11* (OAT4) rs11231809T>A, *TLR2* rs4696480T>A, and *IL6* rs1800796C>G genetic polymorphisms are associated with the incidence of infections among patients undergoing cytotoxic chemotherapy. Including these genetic variables with clinical variables leads to a useful prediction tool. This study is the first to use genetic variables in addition to clinical variables to predict the incidence of infections in patients with hematological malignancies undergoing cytotoxic chemotherapy and could lead to improved clinical outcomes for patients.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found on doi.org/10.6084/m9.figshare.13444211. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Clinical Hospital of the University of Chile (approval received on July 18, 2017) and the Eastern Metropolitan Health Service (approval received on July 4, 2017). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial and direct contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

We acknowledge the Hospital Clínico de la Universidad de Chile's oncologic hematology department, mainly to Rebeca Aguayo. Moreover, to Javier Morales and Felipe Oyarzún for critically reviewing this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.602676/full#supplementary-material.

REFERENCES

Apostolopoulou, E., Raftopoulos, V., Terzis, K., and Elefsiniotis, I. (2010). Infection Probability Score, Apache II and KARNOFSKY scoring systems as predictors of bloodstream infection onset in hematology-oncology patients. *BMC Infect. Dis.* 10 (1), 135. doi:10.1186/1471-2334-10-135

Badawy, A. A., Othman, G., Elabbasy, L., Abd Elsalam, M., Shrief, R., Fahmy, E., et al. (2019). IL-6− 572G/C and− 174G/C polymorphisms association with hepatitis C virus-induced hepatocellular carcinoma. *Br. J. Biomed. Sci.* 76 (4), 201–204. doi:10.1080/09674845.2019.1642562

Beutler, B., Jiang, Z., Georgel, P., Crozat, K., Croker, B., Rutschmann, S., et al. (2006). Genetic analysis of host resistance: Toll-like receptor signaling and immunity at large. *Annu. Rev. Immunol.* 24, 353–389. doi:10.1146/annurev.immunol.24.021605.090552

Bielinski, S. J., Hall, J. L., Pankow, J. S., Boerwinkle, E., Matijevic-Aleksic, N., He, M., et al. (2011). Genetic variants in TLR2 and TLR4 are associated with markers of monocyte activation: the atherosclerosis risk in communities MRI study. *Hum. Genet.* 129 (6), 655–662. doi:10.1007/s00439-011-0962-4

Buaboonnam, J., Sripatanatadasakul, P., Treesucon, A., Glomglao, W., Siraprapapat, P., Narkbunnam, N., et al. (2019). Effect of NUDT15 on incidence of neutropenia in children with acute lymphoblastic leukemia. *Pediatr. Int.* 61 (8), 754–758. doi:10.1111/ped.13905

Burckhardt, G. (2012). Drug transport by organic anion transporters (OATs). *Pharmacol. Ther.* 136 (1), 106–130. doi:10.1016/j.pharmthera.2012.07.010

Castagnola, E., Fontana, V., Caviglia, I., Caruso, S., Faraci, M., Fioredda, F., et al. (2007). A prospective study on the epidemiology of febrile episodes during chemotherapy-induced neutropenia in children with cancer or after hemopoietic stem cell transplantation. *Clin. Infect. Dis.* 45 (10), 1296–1304. doi:10.1086/522533

Chandran, R., Hakki, M., and Spurgeon, S. (2012). "Infections in leukemia," in *Sepsis-an ongoing and significant challenge*. Editors L. Azevedo (London, United Kingdom: IntechOpen).

Chiba, Y., Mizoguchi, I., Hasegawa, H., Ohashi, M., Orii, N., Nagai, T., et al. (2018). Regulation of myelopoiesis by proinflammatory cytokines in infectious diseases. *Cell. Mol. Life Sci.* 75 (8), 1363–1376. doi:10.1007/s00018-017-2724-5

Choi, J. R., Kim, J. O., Kang, D. R., Shin, J. Y., Zhang, X. H., Oh, J. E., et al. (2015). Genetic variations of drug transporters can influence on drug response in patients treated with docetaxel chemotherapy. *Cancer Res. Treat.* 47 (3), 509. doi:10.4143/crt.2014.012

Daly, A. K. (2015). Polymorphic variants of cytochrome P450: relevance to cancer and other diseases. *Adv. Pharmacol.* 74, 85–111. doi:10.1016/bs.apha.2015.03.001

Dutronc, H., Billhot, M., Dupon, M., Eghbali, H., Donamaria, C., Dauchy, F.-A., et al. (2009). Prise en charge de 315épisodes neutropéniques fébriles dans un centre anticancéreux. *Med. Mal. Infect.* 39 (6), 388–393. doi:10.1016/j.medmal.2008.10.014

Esposito, S., Zampiero, A., Pugni, L., Tabano, S., Pelucchi, C., Ghirardi, B., et al. (2014). Genetic polymorphisms and sepsis in premature neonates. *PLoS One* 9 (7), e101248. doi:10.1371/journal.pone.0101248

Gil, L., Styczynski, J., and Komarnicki, M. (2007). Infectious complication in 314 patients after high-dose therapy and autologous hematopoietic stem cell transplantation: risk factors analysis and outcome. *Infection* 35 (6), 421–427. doi:10.1007/s15010-007-6350-2

Green, M. L. (2017). Viral pneumonia in patients with hematopoietic cell transplantation and hematologic malignancies. *Clin. Chest Med.* 38 (2), 295–305. doi:10.1016/j.ccm.2016.12.009

Hakim, H., Dallas, R., Wolf, J., Tang, L., Schultz-Cherry, S., Darling, V., et al. (2018). Gut microbiome composition predicts infection risk during chemotherapy in children with acute lymphoblastic leukemia. *Clin. Infect. Dis.* 67 (4), 541–548. doi:10.1093/cid/ciy153

Harrell, F. E., Jr (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Berlin, Germany: Springer.

Hawn, T. R., Scholes, D., Li, S. S., Wang, H., Yang, Y., Roberts, P. L., et al. (2009). Toll-like receptor polymorphisms and susceptibility to urinary tract infections in adult women. *PLoS One* 4 (6), e5990. doi:10.1371/journal.pone.0005990

Inaba, H., Pei, D., Wolf, J., Howard, S., Hayden, R., Go, M., et al. (2017). Infection-related complications during treatment for childhood acute lymphoblastic leukemia. *Ann. Oncol.* 28 (2), 386–392. doi:10.1093/annonc/mdw557

Jedlitschky, G., Hoffmann, U., and Kroemer, H. K. (2006). Structure and function of the MRP2 (ABCC2) protein and its role in drug disposition. *Expert Opin. Drug Metab. Toxicol.* 2 (3), 351–366. doi:10.1517/17425255.2.3.351

Klastersky, J., De Naurois, J., Rolston, K., Rapoport, B., Maschmeyer, G., Aapro, M., et al. (2016). Management of febrile neutropaenia: ESMO clinical practice guidelines. *Ann. Oncol.* 27 (Suppl 5), v111–v118. doi:10.1093/annonc/mdw325

Lamba, J., Hebert, J. M., Schuetz, E. G., Klein, T. E., and Altman, R. B. (2012). PharmGKB summary: very important pharmacogene information for CYP3A5. *Pharmacogenet. Genomics* 22 (7), 555. doi:10.1097/FPC.0b013e328351d47f

Lee, J. S., Cheong, H. S., Kim, L. H., Kim, J. O., Seo, D. W., Kim, Y. H., et al. (2013). Screening of genetic polymorphisms of CYP3A4 and CYP3A5 genes. *Korean J. Physiol. Pharmacol.* 17 (6), 479–484. doi:10.4196/kjpp.2013.17.6.479

Lewis, L. D., Miller, A. A., Owzar, K., Bies, R. R., Markova, S., Jiang, C., et al. (2013). The relationship of polymorphisms in ABCC2 and SLCO1B3 with docetaxel pharmacokinetics and neutropenia: CALGB 60805 (Alliance). *Pharmacogenet. Genomics* 23 (1), 29–33. doi:10.1097/FPC.0b013e32835b16d8

Li, Y., Klippel, Z., Shih, X., Reiner, M., Wang, H., and Page, J. H. (2016). Relationship between severity and duration of chemotherapy-induced neutropenia and risk of infection among patients with nonmyeloid malignancies. *Support. Care Cancer* 24 (10), 4377–4383. doi:10.1007/s00520-016-3277-0

Lima, A., Bernardes, M., Azevedo, R., Medeiros, R., and Seabra, V. (2015). Pharmacogenomics of methotrexate membrane transport pathway: can clinical response to methotrexate in rheumatoid arthritis be predicted? *Int. J. Mol. Sci.* 16 (6), 13760–13780. doi:10.3390/ijms160613760

Loft, N., Skov, L., Iversen, L., Gniadecki, R., Dam, T., Brandslund, I., et al. (2018). Associations between functional polymorphisms and response to biological treatment in Danish patients with psoriasis. *Pharmacogenomics J.* 18 (3), 494. doi:10.1038/tpj.2017.31

Lyman, G. H., Abella, E., and Pettengell, R. (2014). Risk factors for febrile neutropenia among patients with cancer receiving chemotherapy: a

systematic review. *Crit. Rev. Oncol. Hematol.* 90 (3), 190–199. doi:10.1016/j.critrevonc.2013.12.006

Malcolm, K. C., Arndt, P. G., Manos, E. J., Jones, D. A., and Worthen, G. S. (2003). Microarray analysis of lipopolysaccharide-treated human neutrophils. *Am. J. Physiol. Lung Cel Mol Physiol* 284 (4), L663–L670. doi:10.1152/ajplung.00094.2002

Paramel, G., Sirsjö, A., and Fransén, K. (2015). Role of genetic alterations in the NLRP3 and CARD8 genes in health and disease. *Mediators Inflamm.* 2015, 846782. doi:10.1155/2015/846782

Park, J.-Y., Cha, Y.-J., and Kim, K.-A. (2014). CYP3A5*3Polymorphism and its clinical implications and pharmacokinetic role. *Transl Clin. Pharmacol.* 22 (1), 3–7. doi:10.12793/tcp.2014.22.1.3

Rose-John, S., Winthrop, K., and Calabrese, L. (2017). The role of IL-6 in host defence against infections: immunobiology and clinical implications. *Nat. Rev. Rheumatol.* 13 (7), 399. doi:10.1038/nrrheum.2017.83

Sabroe, I., Dower, S. K., and Whyte, M. K. (2005). The role of Toll-like receptors in the regulation of neutrophil migration, activation, and apoptosis. *Clin. Infect. Dis.* 41 Suppl 7 (Suppl 7), S421–S426. doi:10.1086/431992

Schalk, E., Hanus, L., Färber, J., Fischer, T., and Heidel, F. H. (2015). Prediction of central venous catheter-related bloodstream infections (CRBSIs) in patients with haematologic malignancies using a modified Infection Probability Score (mIPS). *Ann. Hematol.* 94 (9), 1451–1456. doi:10.1007/s00277-015-2387-y

Sharma, A., Singh, K., Biswas, A., Ranjan, R., Kishor, K., Pandey, H., et al. (2018). Impact of interleukin 6 promoter polymorphisms (-174 G > C, -572 G > C and -597 G > A) on plasma IL-6 levels and their influence on the development of DVT: a study from India. *Hematology* 23 (10), 833–838. doi:10.1080/10245332.2018.1483546

Singh, S., Parulekar, W., Murray, N., Feld, R., Evans, W. K., Tu, D., et al. (2005). Influence of sex on toxicity and treatment outcome in small-cell lung cancer. *J. Clin. Oncol.* 23 (4), 850–856. doi:10.1200/JCO.2005.03.171

Sung, L., Lange, B. J., Gerbing, R. B., Alonzo, T. A., and Feusner, J. (2007). Microbiologically documented infections and infection-related mortality in children with acute myeloid leukemia. *Blood* 110 (10), 3532–3539. doi:10.1182/blood-2007-05-091942

Tang, S., Yuan, Y., He, Y., Pan, D., Zhang, Y., Liu, Y., et al. (2014). Genetic polymorphism of interleukin-6 influences susceptibility to HBV-related hepatocellular carcinoma in a male Chinese Han population. *Hum. Immunol.* 75 (4), 297–301. doi:10.1016/j.humimm.2014.02.006

Vasiliou, V., Vasiliou, K., and Nebert, D. W. (2009). Human ATP-binding cassette (ABC) transporter family. *Hum. Genomics* 3 (3), 281. doi:10.1186/1479-7364-3-3-281

Vormfelde, S. V., Schirmer, M., Hagos, Y., Toliat, M. R., Engelhardt, S., Meineke, I., et al. (2006). Torsemide renal clearance and genetic variation in luminal and basolateral organic anion transporters. *Br. J. Clin. Pharmacol.* 62 (3), 323–335. doi:10.1111/j.1365-2125.2006.02655.x

Webb, B. J., Healy, R., Majers, J., Burr, Z., Gazdik, M., Lopansri, B., et al. (2017). Prediction of bloodstream infection due to vancomycin-resistant enterococcus in patients undergoing leukemia induction or hematopoietic stem-cell transplantation. *Clin. Infect. Dis.* 64 (12), 1753–1759. doi:10.1093/cid/cix232

Wright, H. L., Cross, A. L., Edwards, S. W., and Moots, R. J. (2014). Effects of IL-6 and IL-6 blockade on neutrophil function *in vitro* and *in vivo*. *Rheumatology (Oxford)* 53 (7), 1321–1331. doi:10.1093/rheumatology/keu035

Yilmaz, Ş., Ören, H., Demircioğlu, F., and Irken, G. (2008). Assessment of febrile neutropenia episodes in children with acute leukemia treated with BFM protocols. *Pediatr. Hematol. Oncol.* 25 (3), 195–204. doi:10.1080/08880010801938231

Zhu, X., Yun, W., Sun, X., Qiu, F., Zhao, L., and Guo, Y. (2014). Effects of major transporter and metabolizing enzyme gene polymorphisms on carbamazepine metabolism in Chinese patients with epilepsy. *Pharmacogenomics* 15 (15), 1867–1879. doi:10.2217/pgs.14.142

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

OPINION
published: 19 March 2021
doi: 10.3389/fphar.2021.620043

The Urgent Need for Management of Biological Samples and Data Accessibility in Latin America

Rodrigo José Vargas[1,2,3]* *and Oscar M. Cobar*[1,3,4]

[1]Escuela de Estudios de Postgrado, Facultad de Ciencias Médicas, Universidad de San Carlos de Guatemala, Guatemala City, Guatemala, [2]Universidad Galileo, Guatemala City, Guatemala, [3]Latin American Network for the Implementation and Validation of Pharmacogenomic Clinical Guidelines (RELIVAF-CYTED), Madrid, Spain, [4]Universidad del Istmo, Fraijanes, Guatemala

Keywords: biobank, Amerindian, COVID-19, mestizo, diversity

INTRODUCTION

Elucidation of the human genome has increased understanding of human body responses to drug administration (Oprea et al., 2018; Nabirotchkin et al., 2020). Likewise, recent studies on human genetic diversity have shown that it is still necessary to delve into individual genetic differences, the adverse effects associated with drug metabolism, drug response variability with the diet, and even the human microbiome (Sharma et al., 2019). Access to biological samples in Latin America is essential to further understanding the presence of genes that may be associated with adverse effects and pharmacological interactions due to their diversity of populations and in anticipating the effects of new treatments.

The Coronavirus disease 2019 (COVID-19) pandemic has reinforced the urgent need to study the genetic differences among people with mild symptoms and those with complex responses to the disease (Ovsyannikova et al., 2020). During the pandemic, different drugs have been studied in the search for therapeutic alternatives to combat it, mainly due to the "Solidarity Trial" and "Repurposing Drugs" initiatives of the World Health Organization (OMS) (Harrison, 2020). Similarly, new drugs for actual and future diseases can be designed by using pharmacogenetic information. As an example, due to the diversity of therapeutic action mechanisms of a drug, it is necessary to in parallel study human cell susceptibility to the entry of SARS-CoV-2 in genetic world populations. Pharmacogenetics studies of their allele variants are essential. This approach is imperative in enabling national health systems to make informed decisions about the therapeutic strategies used, especially in countries with multiple ethnic groups.

Examples that illustrate the importance of understanding therapeutic effectiveness responses in target groups include the premature administration of hydroxychloroquine based on affect glycosylation of angiotensin converting enzyme-2, without information on genetic variability (Ferner and Aronson, 2020) and remdesivir administration (Beigel et al., 2020). Research on ethnic group pharmacogenetics in Latin America is still scarce (Sosa-Macías et al., 2016; Leitão et al., 2020), but there is evidence of pharmacogenetic differences (Suarez-Kurtz and Parra, 2018). For example, Mestizo populations are not considered a unique group for these types of studies, since miscegenation varies in each Latin American country (Botton et al., 2019).The COVID-19 pandemic has shown us that data and access to pharmacogenetic information are required for making clinical decisions in a prioritized and sometimes urgent ways.

The reservoirs of genetic material in biobanks in the United Kingdom (McInnes and Altman, 2020) and even in Africa (Matimba et al., 2008), allow researchers to quickly access information on the allelic distributions of important pharmacogenetic markers. However, in Latin America, biobanks for pharmacogenetic research purposes have had only slight development. The bioethical and legal aspects must be considered, and access to biomedical sample information is

OPEN ACCESS

Edited by:
Matias F. Martinez, University of Chile, Chile

Reviewed by:
M. Isabel Lucena, University of Malaga, Spain
Alfonso Dueñas-González, National Autonomous University of Mexico, Mexico

***Correspondence:**
Rodrigo José Vargas
rodrigovargas@galileo.edu

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *21 October 2020*
Accepted: *15 February 2021*
Published: *19 March 2021*

Citation:
Vargas RJ and Cobar OM (2021) The Urgent Need for Management of Biological Samples and Data Accessibility in Latin America. Front. Pharmacol. 12:620043. doi: 10.3389/fphar.2021.620043

necessary in a region with high ethnical diversity and little-studied ancestral Amerindian populations.

BIOBANKS

Biobanks can hold genetic data for a significant percentage of an entire population (Malsagova et al., 2020). In Estonia, e.g., genetic information on about 5% of the adult population can be found in a DNA bank (Reisberg et al., 2019). Additionally, biobanks contribute to reducing the cost of pharmacogenetic studies (Huttin and Liebman, 2013) avoiding the sampling of the population of interest in each study, also allowing a collaborative approach. In Europe, biobanks like one in Poland, have used different management models, preserving essential fundamental information for the benefit of public health (Sak et al., 2012). Moreover, biobanks are very useful for establishing pharmacogenetic relationships for research drug interactions with genes identified in a population (Muller et al., 2020; Rollinson et al., 2020).

BIOETHICS AND LEGAL ASPECTS

There are ethical dilemmas involved with asking a donor to provide unique informed consent. This has, however, been improved with the development of the model of dynamic consent (Steinsbekk et al., 2013). Ethics committees play a key role in the ethical debates concerning approval of donations to biobanks and access to the stored genetic material (Hansson, 2009). A legal framework is also required for the use of biobank samples (Helgesson et al., 2007). Recently, legal and ethical issues have focused on sample ownership and access, donor protection, and long-term storage of biological samples (Paskal et al., 2018; Facca et al., 2020).

CONCLUSION

Due to high population diversity, Latin America faces the challenge of addressing genetic variability in studies to improve pharmacological responses to therapeutics for diseases. The creation of biobanks, their strengthening, and collaboration among them, would be a fundamental contribution to obtain pharmacogenetic information and efficient therapeutic responses in Latin America.

AUTHOR CONTRIBUTIONS

RV and OC designed and drafted the manuscript. Both authors contributed to the article and approved the submitted version.

REFERENCES

Beigel, J. H., Tomashek, K. M., Dodd, L. E., Mehta, A. K., Zingman, B. S., Kalil, A. C., et al. (2020). Remdesivir for the treatment of covid-19—final report. *New Engl. J. Med.* 0 (0). doi:10.1056/NEJMoa2007764

Botton, M. R., Campos, K. C., Duconge, J., and Scott, S. A. (2019). "Chapter 12-pharmacogenomics in Latin American populations," in *Pharmacogenomics*. 2nd Edn, Editors E. Y. W. F. Lam and S. A. Scott (Academic Press), 329–368.

Facca, D., Smith, M., Shelley, J., Lizotte, D., and Donelle, L. (2020). Exploring the ethical issues in research using digital data collection strategies with minors: a scoping review. *PLoS One* 15 (8), e0237875. doi:10.1371/journal.pone.0237875

Ferner, R. E., and Aronson, J. K. (2020). Chloroquine and hydroxychloroquine in covid-19. *BMJ* 369, m1432. doi:10.1136/bmj.m1432

Hansson, M. G. (2009). Ethics and biobanks. *Br. J. Cancer* 100 (1), 8–12. doi:10.1038/sj.bjc.6604795

Harrison, Ch. (2020). Coronavirus puts drug repurposing on the fast track. *Nat. Biotechnol.* 38 (4), 379–391. doi:10.1038/d41587-020-00003-1

Helgesson, G., Dillner, J., Carlson, J., Bartram, C. R., and Hansson, M. G. (2007). Ethical framework for previously collected biobank samples. *Nat. Biotechnol.* 25 (9), 973–976. doi:10.1038/nbt0907-973b

Huttin, C. C., and Liebman, M. N. (2013). The economics of biobanking and pharmacogenetics databasing. *Technol. Health Care* 21 (2), 183–190. doi:10.3233/THC-130729

Leitão, L. P., Souza, T. P., Rodrigues, J. C., Fernandes, M. R., Santos, S., and Santos, N. P. (2020). The metabolization profile of the CYP2D6 gene in amerindian populations: a review. *Genes* 11 (3), 262. doi:10.3390/genes11030262

Malsagova, K., Kopylov, A., Stepanov, A., Butlova, T., Sinitsyna, A., Izotov, A., et al. (2020). Biobanks-A platform for scientific and biomedical research. *Diagnostics* 10, 485. doi:10.3390/diagnostics10070485

Matimba, A., Oluka, M. N., Ebeshi, B. U., Sayi, J., Bolaji, O. O., Guantai, A. N., et al. (2008). Establishment of a biobank and pharmacogenetics database of African populations. *Eur. J. Hum. Genet.* 16 (7), 780–783. doi:10.1038/ejhg.2008.49

McInnes, G., and Altman, R. B. (2020). Drug response pharmacogenetics for 200,000 UK biobank participants. bioRxiv. [Preprint]. doi:10.1101/2020.08.09.243311v1

Muller, H., Dagher, G., Loibner, M., Stumptner, C., Kungl, P., and Zatloukal, K. (2020). Biobanks for life sciences and personalized medicine. importance of standarization, biosafety, biosecurity and data management. *Curr. Opin. Biotechnol.* 65, 45–51. doi:10.1016/j.copbio.2019.12.004

Nabirotchkin, S., Pelufo, A., Rinaudo, Ph., Yu, J., Hajj, R., and Cohen, D. (2020). Next-generation drug repurposing using human genetics and network biology. *Curr. Opin. Pharmacol.* 51, 78–92. doi:10.1016/j.coph.2019.12.004

Oprea, T. I., Bologa, C. G., Brunak, S., Campbell, A., Gan, G. N., Gaulton, A., et al. (2018). Unexplored therapeutic opportunities in the human genome. *Nat. Rev. Drug Discov.* 17 (5), 317–332. doi:10.1038/nrd.2018.14

Ovsyannikova, I. G., Haralambieva, I. H., Crooke, S. N., Poland, G. A., and Kennedy, R. B. (2020). The role of host genetics in the immune response to SARS-CoV-2 and COVID-19 susceptibility and severity. *Immunological Rev.* 296 (1), 205–219. doi:10.1111/imr.12897

Paskal, W., Paskal, A. M., Dębski, T., Gryziak, M., and Jaworowski, J. (2018). Aspects of modern biobank activity—comprehensive review. *Pathol. Oncol. Res.* 24 (4), 771–785. doi:10.1007/s12253-018-0418-4

Reisberg, S., Krebs, K., Lepamets, M., Kals, M., Mägi, R., Metsalu, K., et al. (2019). Translating genotype data of 44,000 biobank participants into clinical pharmacogenetic recommendations: challenges and solutions. *Genet. Med.* 21 (6), 1345–1354. doi:10.1038/s41436-018-0337-5

Rollinson, V., Turner, R., and Pirmohamed, M. (2020). Pharmacogenomics for primary care: an overview. *Genes* 11 (11), 1337. doi:10.3390/genes11111337

Sak, J., Pawlikowski, J., Goniewicz, M., and Witt, M. (2012). Population biobanking in selected European countries and proposed model for a polish national DNA bank. *J. Appl. Genet.* 53 (2), 159–165. doi:10.1007/s13353-012-0082-4

Sharma, A., Buschmann, M. M., and Gilbert, J. A. (2019). Pharmacomicrobiomics: the holy grail to variability in drug response? *Clin. Pharmacol. Ther.* 106 (2), 317–328. doi:10.1002/cpt.1437

Sosa-Macías, M., Teran, E., Waters, W., Fors, M. M., Altamirano, C., Jung-Cook, H., et al. (2016). Pharmacogenetics and ethnicity: relevance for clinical implementation, clinical trials, pharmacovigilance and drug regulation in

Latin America. *Pharmacogenomics* 17 (16), 1741–1747. doi:10.2217/pgs-2016-0153

Steinsbekk, K. S., Kåre Myskja, B., and Solberg, B. (2013). Broad consent versus dynamic consent in biobank research: is passive participation an ethical problem? *Eur. J. Hum. Genet.* 21 (9), 897–902. doi:10.1038/ejhg.2012.282

Suarez-Kurtz, G., and Parra, E. J. (2018). "Chapter six—population diversity in pharmacogenetics: a Latin American perspective," in *Advances in pharmacology*. Editors En. K. Brøsen and P. Damkier (Academic Press), Vol. 83, 133–154.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 12 April 2021
doi: 10.3389/fphar.2021.630658

A New Insight for the Identification of Oncogenic Variants in Breast and Prostate Cancers in Diverse Human Populations, With a Focus on Latinos

Nelson M. Varela[1,2], *Patricia Guevara-Ramírez*[2], *Cristian Acevedo*[1,3], *Tomás Zambrano*[4], *Isaac Armendáriz-Castillo*[5], *Santiago Guerrero*[5], *Luis A. Quiñones*[1,2]* and *Andrés López-Cortés*[2,5,6]*

[1]Laboratory of Chemical Carcinogenesis and Pharmacogenetics, Department of Basic and Clinical Oncology, Faculty of Medicine, University of Chile, Santiago, Chile, [2]Latin American Network for the Implementation and Validation of Clinical Pharmacogenomics Guidelines (RELIVAF-CYTED), Madrid, Spain, [3]Department of Basic and Clinical Oncology, Clinical Hospital University of Chile, Santiago, Chile, [4]Department of Medical Technology, Faculty of Medicine, University of Chile, Santiago, Chile, [5]Centro de Investigación Genética y Genómica, Facultad de Ciencias de la Salud Eugenio Espejo, Universidad UTE, Quito, Ecuador, [6]Department of Computer Science and Information Technologies, Faculty of Computer Science, University of A Coruna, A Coruña, Spain

OPEN ACCESS

Edited by:
Marcelo Rizzatti Luizon,
Federal University of Minas Gerais, Brazil

Reviewed by:
Jatinder K Lamba,
University of Florida, United States
Yong Xu,
First Hospital of Shanxi Medical University, China

***Correspondence:**
Luis Abel Quiñones
lquinone@uchile.cl
Andrés López-Cortés
aalc84@gmail.com

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: 18 November 2020
Accepted: 25 January 2021
Published: 12 April 2021

Citation:
Varela NM, Guevara-Ramírez P, Acevedo C, Zambrano T, Armendáriz-Castillo I, Guerrero S, Quiñones LA and López-Cortés A (2021) A New Insight for the Identification of Oncogenic Variants in Breast and Prostate Cancers in Diverse Human Populations, With a Focus on Latinos. Front. Pharmacol. 12:630658. doi: 10.3389/fphar.2021.630658

Background: Breast cancer (BRCA) and prostate cancer (PRCA) are the most commonly diagnosed cancer types in Latin American women and men, respectively. Although in recent years large-scale efforts from international consortia have focused on improving precision oncology, a better understanding of genomic features of BRCA and PRCA in developing regions and racial/ethnic minority populations is still required.

Methods: To fill in this gap, we performed integrated *in silico* analyses to elucidate oncogenic variants from BRCA and PRCA driver genes; to calculate their deleteriousness scores and allele frequencies from seven human populations worldwide, including Latinos; and to propose the most effective therapeutic strategies based on precision oncology.

Results: We analyzed 339,100 variants belonging to 99 BRCA and 82 PRCA driver genes and identified 18,512 and 15,648 known/predicted oncogenic variants, respectively. Regarding known oncogenic variants, we prioritized the most frequent and deleterious variants of BRCA (n = 230) and PRCA (n = 167) from Latino, African, Ashkenazi Jewish, East Asian, South Asian, European Finnish, and European non-Finnish populations, to incorporate them into pharmacogenomics testing. Lastly, we identified which oncogenic variants may shape the response to anti-cancer therapies, detailing the current status of pharmacogenomics guidelines and clinical trials involved in BRCA and PRCA cancer driver proteins.

Conclusion: It is imperative to unify efforts where developing countries might invest in obtaining databases of genomic profiles of their populations, and developed countries might incorporate racial/ethnic minority populations in future clinical trials and cancer researches with the overall objective of fomenting pharmacogenomics in clinical practice and public health policies.

Keywords: breast, prostate, oncogenic variants, latino population, precision oncology

INTRODUCTION

Cancer is the second leading cause of death globally (Bray et al., 2018); meanwhile, breast cancer (BRCA) and prostate cancer (PRCA) are the most commonly diagnosed cancer types in Latin American women and men, respectively, (López-Cortés et al., 2017). BRCA and PRCA are complex and heterogeneous diseases characterized by an intricate interplay between different biological features, such as driver mutations, ethnicity, protein expression deregulation, signaling pathway alterations, and environmental determinants (López-Cortés et al., 2013; López-Cortés et al.,2018; López-Cortés et al.,2020b; López-Cortés et al.,2020a).

Starting with the Human Genome Project in 1990, genomics has progressively become an essential tool in basic and translational research (Green et al., 2020). The development of high-throughput technologies focused on large-scale DNA sequencing has allowed us to better understand the molecular landscape of oncogenesis. Thus, considerable progress has been made in discovering cancer driver genes (Kandoth et al., 2013; Lawrence et al., 2014), coding and non-coding cancer driver mutations (Sjöblom et al., 2006; Tamborero et al., 2013; Porta-Pardo et al., 2017; Rheinbay et al., 2020), germline variants (Lu et al., 2015), druggable enzymes (Rubio-Perez et al., 2015), and drug resistance (Vasan et al., 2019).

A main goal in oncology research is to understand the mechanisms of malignant cell transformation to develop efficient therapeutic approaches. One milestone towards this objective is the identification of cancer driver genes carrying mutations capable of driving BRCA and PRCA tumorigenesis. Nowadays, it is known that cancer driver genes are under positive selection in tumorigenesis, and the development of carefully designed bioinformatics pipelines such as the Integrative OncoGenomics (IntOGen) and the Cancer Genome Interpreter (CGI) is fundamental to identify oncogenic variants across tumors (Tamborero et al., 2018; Martínez-Jiménez et al., 2020). Similarly, The Cancer Genome Atlas (TCGA) and the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) projects have established molecular tumor classification based on DNA, RNA and protein alterations (ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, 2020).

These genomic signatures are allowing the development of personalized cancer treatments. Over the past years it has become clear that oncological patients, diagnosed with the same cancer type, may have different responses to generic treatments such as radiation or chemotherapy. To overcome these variable responses, cancer precision medicine aims to provide the right dose of the right drug for the right patient at the right time (Quinones et al., 2014). Thus, precision medicine has become an important tool in cancer treatment; it allows the identification of specific mutations in driver genes responsible for tumor progression (ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, 2020). Based on each human multi-omics profile, drug development can be tailored for each individual improving not only efficiency of the drug but minimizing the possibility of acquiring adverse reactions (López-Cortés et al., 2020c).

Despite these efforts, fundamental and applied cancer researchers have failed to include ethnically diverse populations (Guerrero et al., 2018). In that respect, several studies have shown that race/ethnicity has a great impact on cancer incidence, survival, drug response, molecular pathways, and epigenetics (Ma et al., 2010; Patel, 2015; Kader and Ghai, 2017). Astonishingly, relevant cancer genomic databases, such as TCGA and TARGET, are overrepresented by Caucasian individuals (91.1%) (Guerrero et al., 2018). Consequently, this bias greatly limits the development of pharmacogenomics (PGx) and precision medicine in developing regions, such as Latin America. To fill in this gap, we performed integrated *in silico* analyses to elucidate oncogenic variants from BRCA and PRCA driver genes, and to calculate their deleteriousness scores and allele frequencies from seven human populations worldwide, with a focus on the Latino population.

METHODS

Incidence and Mortality of BRCA and PRCA

The Global Cancer Observatory (GLOBOCAN) (https://gco.iarc.fr/) enables a comprehensive assessment of the cancer burden worldwide (Bray et al., 2018). From the latest version of GLOBOCAN 2020, we have retrieved the estimated crude incidence and mortality rates related to the top cancer types worldwide, and the estimated crude incidence and mortality rates of BRCA and PRCA from Latin American and the Caribbean countries.

BRCA and PRCA Driver Genes

The intOGen (https://www.intogen.org) framework identifies cancer genes and pinpoints their mechanism of action across tumor types (Martínez-Jiménez et al., 2020). The current version of the intOGen pipeline uses seven methods to identify cancer driver genes from somatic point mutations: dNdScv (Martincorena et al., 2017), CBaSE (Weghorn and Sunyaev, 2017), MutPanning (Dietlein et al., 2020), OncodriveCLUSTL (Arnedo-Pac et al., 2019), HotMAPS (Tokheim et al., 2016), smRegions (Martínez-Jiménez et al., 2020), and OncodriveFML (Mularoni et al., 2016). Therefore, we have retrieved 99 BRCA driver genes and 82 PRCA driver genes from intOGen, and have identified its involvement as oncogenes (Sondka et al., 2018), tumor suppressor genes (Sondka et al., 2018), kinase genes (Manning et al., 2002), DNA-repair genes (Chae et al., 2016), RNA-binding proteins (Hentze et al., 2018), cell cycle genes (Bar-Joseph et al., 2008), and cancer immunotherapy genes (Patel et al., 2017). Lastly, BRCA and PRCA driver gene sets are fully detailed in **Supplementary Table S1**.

Identification of Oncogenic Variants

The identification of oncogenic variants was divided in two steps. In the first step, we extracted 339,100 single nucleotide variants and insertion/deletion variants belonging to 99 BRCA diver genes (n = 183,616) and 82 PRCA driver genes (n = 155,484) from the Genome Aggregation database (gnomAD v2.2.1), using GRCh37/hg19 as the human genome reference (Collins et al., 2020; Karczewski et al., 2020). In the second step, we performed the OncodriveMUT method integrated into the Cancer Genome Interpreter platform (https://www.cancergenomeinterpreter.org) to assess the tumorigenic potential of the 339,100 aforementioned genomic variants (Tamborero et al., 2018). OncodriveMUT is a developed rule-based approach that combines genomic features such as gene signals of positive selection, clusters of somatic mutations, gene mechanism of action, and regions depleted by germline variants to classify driver mutations into known, predicted tier 1, predicted tier 2, and passenger mutations using the Catalog of Validated Oncogenic Mutations (Tamborero et al., 2018).

Deleteriousness Score of Oncogenic Variants

Combined Annotation-Dependent Depletion (CADD) (https://cadd.gs.washington.edu/) is an integrative annotation built from more than 60 genomic features, and measures the deleteriousness of single nucleotide variants as well as insertion/deletion variants in the human genome (Kircher et al., 2014; Rentzsch et al., 2019). In this study, we calculated the CADD phred score for ranking the deleteriousness of known and predicted oncogenic variants located in BRCA and PRCA driver genes. The deleteriousness of oncogenic variants was categorized according to its CADD phred score in very high (30–50), high (25–30), medium (15–25), low (10–15), and very low (0–10).

Allele Frequencies in Human Populations

The gnomAD (https://gnomad.broadinstitute.org/) is a database that harmonize exome and genome sequencing data from a variety of large-scale sequencing projects worldwide (Karczewski et al., 2020). The gnomAD database version 2.1.1 is integrated by 15,708 exomes and 125,748 genomes (total = 141,456). In this study, we calculated the allele frequencies of BRCA and PRCA oncogenic variants belonging to seven human populations, such as Latino (424 exomes and 17,296 genome), African (4,359 exomes and 8,128 genomes), Ashkenazi Jewish (145 exomes and 5,040 genomes), East Asian (780 exomes and 9,197 genomes), South Asian (15,308 genomes), European Finnish (1,738 exomes and 10,824 genomes), and European non-Finnish (7,718 exomes and 56,885 genomes) (Collins et al., 2020; Karczewski et al., 2020; Wang et al., 2020).

Validation of Known Oncogenic Variants Through the Pan-Cancer Atlas

The Pan-Cancer Atlas project, which belongs to TCGA consortium, provides a comprehensive, in-depth, and interconnected understanding of human cancer, and is an essential resource for the development of new treatments in the pursuit of precision medicine (Hoadley et al., 2018; Huang et al., 2018). Therefore, the previously obtained known oncogenic variants were identified and the allele frequencies were calculated in a cohort of 850 TCGA-BRCA patients encompassing 162 black/African individuals, 600 white individuals (not Hispanic or Latino), 29 white individuals (Hispanic or Latino), and 59 Asian individuals; and in a cohort of 150 TCGA-PRAD patients encompassing seven black/African individuals, and 143 white individuals (not Hispanic or Latino). Lastly, mutational data was taken from the Genomics Data Commons of the National Cancer Institute (https://portal.gdc.cancer.gov/), and the cBioPortal database (http://www.cbioportal.org/) (Cerami et al., 2012; Gao et al., 2013).

Current Pharmacogenomics Guidelines in Clinical Practice

PharmGKB (https://www.pharmgkb.org/) is a pharmacogenomics knowledge resource that encompasses potentially clinically actionable gene-drug associations and precise guidelines for the application of pharmacogenomics in clinical practice (Whirl-Carrillo et al., 2012; Barbarino et al., 2018). This database collects information from the Clinical Pharmacogenetics Implementation Consortium (Saito et al., 2016; Relling et al., 2020), the Canadian Pharmacogenomics Network for Drug Safety (Ross et al., 2010), the Royal Dutch Association for the Advancement of Pharmacy (Swen et al., 2011), the National Comprehensive Cancer Network (NCCN), and the European Society for Medical Oncology (ESMO). Consequently, we have retrieved clinical annotations, gene-drug pairs, and genomic variants associated to BRCA and PRCA pharmacogenomics guidelines.

In silico Drug Prescription

Another CGI approach is the *in silico* drug prescription that contains the putative biomarker of drug response found in the tumor organized according to distinct levels of clinical relevance. The CGI employs two resources to explore the association between genomic variants and drug response: the Cancer Biomarker database (Dienstmann et al., 2015), and the Cancer Bioactivities database (Tamborero et al., 2018). Therefore, we performed an in silico analysis to determine the druggability of known and predicted oncogenic variants located in BRCA and PRCA driver genes, and consequently the most relevant precision oncology treatments.

Clinical Trials

The Open Targets Platform (https://www.targetvalidation.org) is comprehensive and robust data integration for access to and visualization of drugs involved in clinical trials associated with BRCA and PRCA proteins, detailing its phase, type, action type, and target class (Carvalho-Silva et al., 2019). Additionally, we created Sankey plots to better understand which drugs are involved in the most advanced phases (3 and 4) of clinical trials in both cancer types.

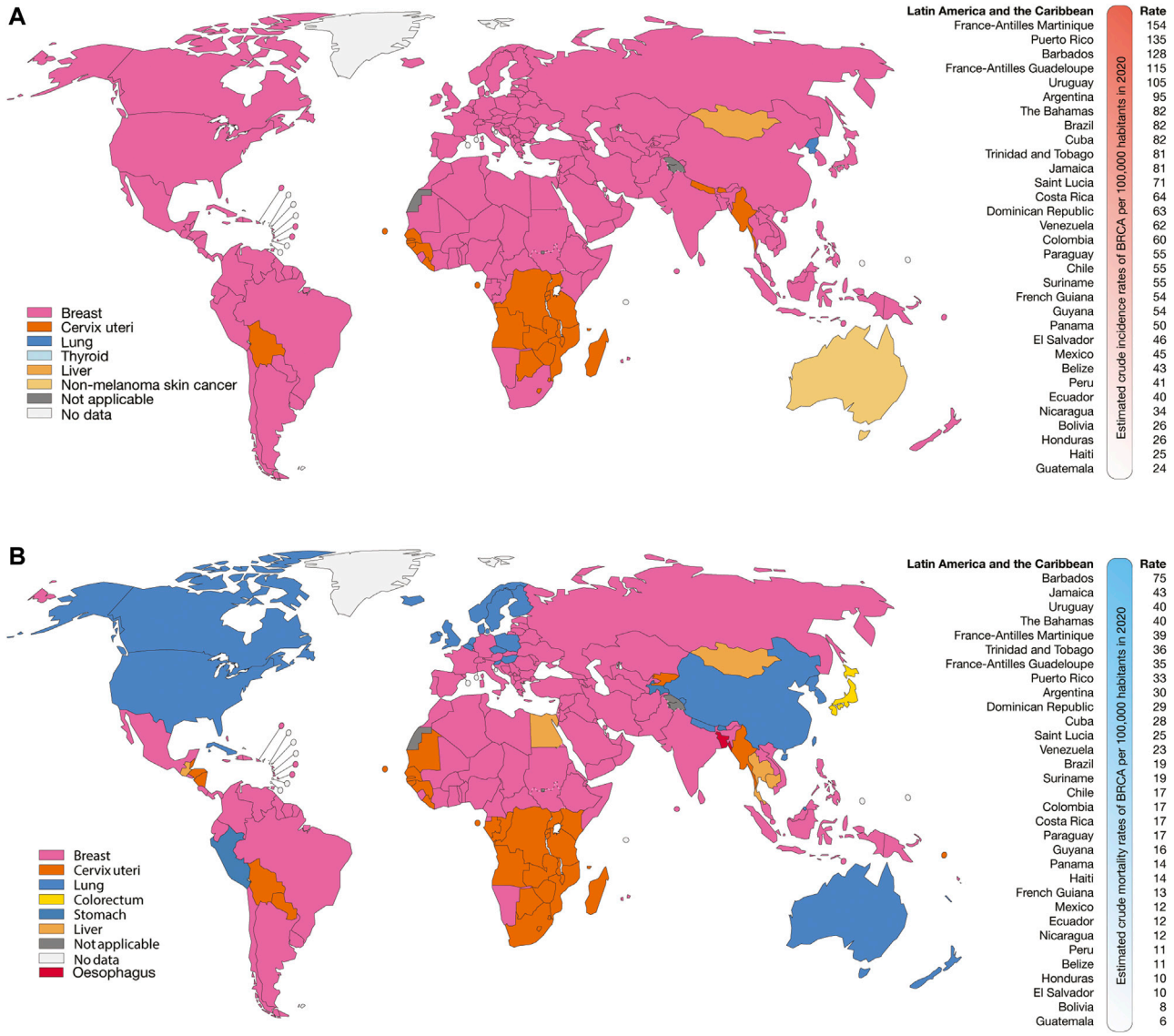


FIGURE 1 | Epidemiology of BRCA. **(A)** Top cancer crude incidence per country in females of all ages **(left)**, and ranking of estimated crude incidence rates of BRCA per 100,000 inhabitants in Latin American and the Caribbean countries **(right)**. **(B)** Top cancer mortality per country in females of all ages **(left)**, and ranking of mortality crude rates of BRCA per 100,000 inhabitants in Latin American and the Caribbean countries **(right)**.

RESULTS

Incidence and Mortality of BRCA and PRCA

According to GLOBOCAN 2020, the Latin American and the Caribbean countries with the highest estimated crude incidence rates of BRCA per 100,000 inhabitants were France-Antilles Martinique (154), Puerto Rico (135), Barbados (128), France-Antilles Guadeloupe (115), Uruguay (105), Argentina (95), The Bahamas (82), and Brazil (82) (**Figure 1A**); the countries with the highest estimated crude mortality rates of BRCA were Barbados (75), Jamaica (43), Uruguay (40), The Bahamas (40), France-Antilles Martinique (39), Trinidad and Tobago (36), and France-Antilles Guadeloupe (35) (**Figure 1B**); the countries with the highest estimated crude incidence rates of PRCA were France-Antilles Guadeloupe (391), France-Antilles Martinique (382), Puerto Rico (202), Barbados (201), Saint Lucia (149), Trinidad and Tobago (128), Cuba (113), and Jamaica (106) (**Figure 2A**); lastly, the countries with the highest estimated crude mortality rates of PRCA were Barbados (99), France-Antilles Martinique (62), Cuba (61), Saint Lucia (60), Trinidad and Tobago (58), Jamaica (57), France-Antilles Guadeloupe (53), and Dominican Republic (41) (**Figure 2B**) (Bray et al., 2018).

BRCA and PRCA Driver Genes

We have retrieved 99 BRCA driver genes and 82 PRCA driver genes from the intOGen framework (Martínez-Jiménez et al.,

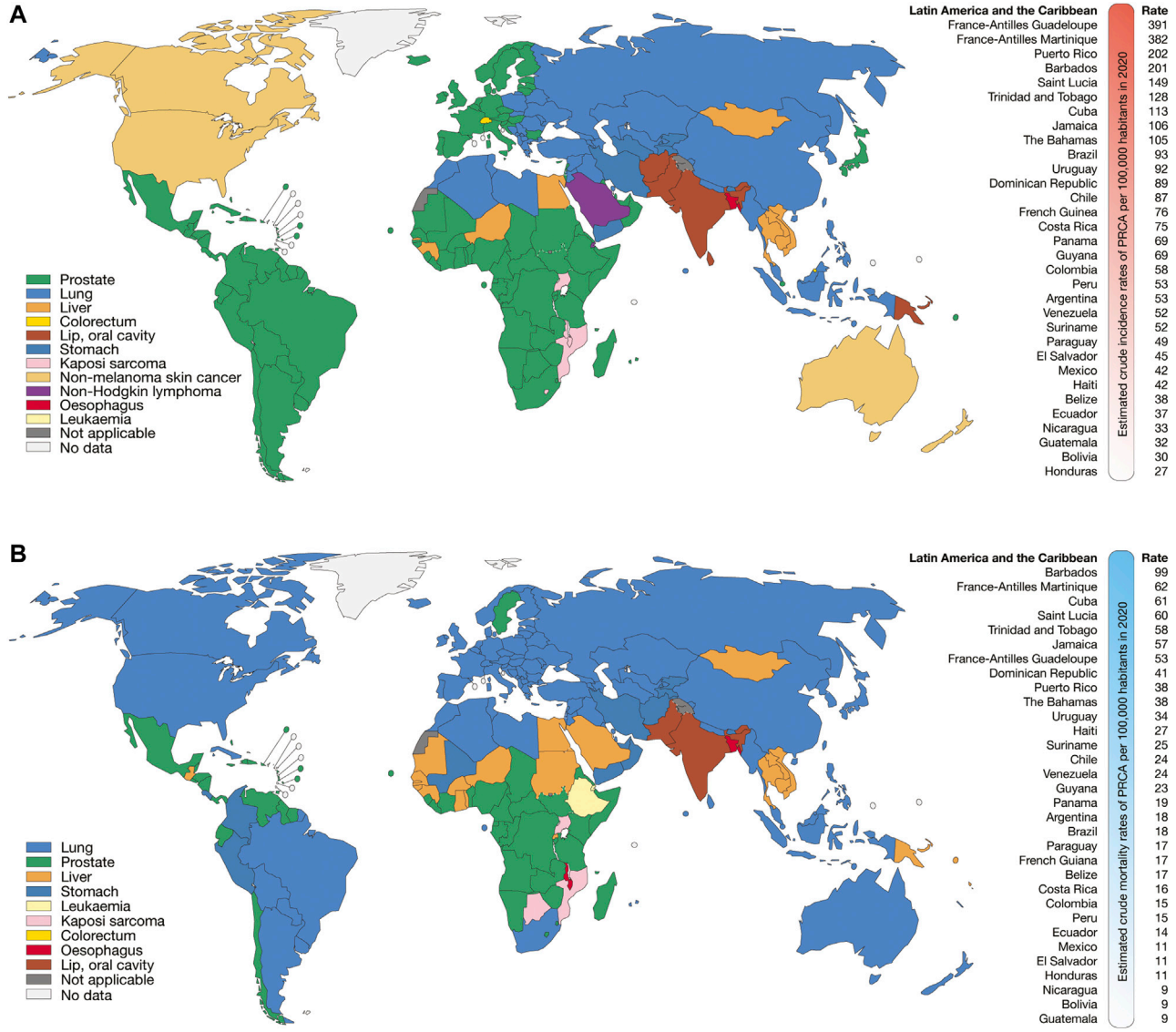


FIGURE 2 | Epidemiology of PRCA. **(A)** Top cancer crude incidence per country in males of all ages **(left)**, and ranking of estimated crude incidence rates of PRCA per 100,000 inhabitants in Latin American and the Caribbean countries **(right)**. **(B)** Top cancer mortality per country in males of all ages **(left)**, and ranking of mortality crude rates of PRCA per 100,000 inhabitants in Latin American and the Caribbean countries **(right)**.

2020). Regarding BRCA driver genes, 41.4% were tumor suppressor genes (Sondka et al., 2018), 28.3% were oncogenes (Sondka et al., 2018), 15.2% were kinase genes (Manning et al., 2002), 10.1% encode RNA-binding proteins (Hentze et al., 2018), 9.1% were cancer immunotherapy genes (Patel et al., 2017), 4% were cell cycle genes (Bar-Joseph et al., 2008), and 4% were DNA repair genes (Chae et al., 2016). Regarding PRCA driver genes, 37.8% were tumor suppressor genes (Sondka et al., 2018), 31.7% were oncogenes (Sondka et al., 2018), 13.4% were cancer immunotherapy genes (Patel et al., 2017), 12.2% were kinase genes (Manning et al., 2002), 8.5% encode RNA-binding proteins (Hentze et al., 2018), 3.7% were cell cycle genes (Bar-Joseph et al., 2008), and 3.7% were DNA repair genes (Chae et al., 2016) (**Supplementary Table 1**).

Identification of Oncogenic Variants

Figure 3 shows the results of the OncodriveMUT analysis to identify oncogenic variants in 99 BRCA driver genes. After the analysis of 183,616 variants, we identified 18,512 oncogenic variants. Of them, 240 (1%) were known, 10,437 (56%) were predicted tier 1, and 7,835 (42%) were predicted tier 2. Regarding gene role, 9,766 (53%) variants produced protein loss of function and 4,541 (25%) produced protein activation. In addition, 6,246 (34%) oncogenic variants had very high CADD phred score, and 11,557 (62%) had high score (**Figure 3A**).

Figure 3B shows violin plots and ranking of CADD phred score of known, predicted tier 1 and tier 2 oncogenic variants in BRCA. The known oncogenic variant with the highest score was *ATM* rs371638537 (score = 54). The predicted tier 1 variant with

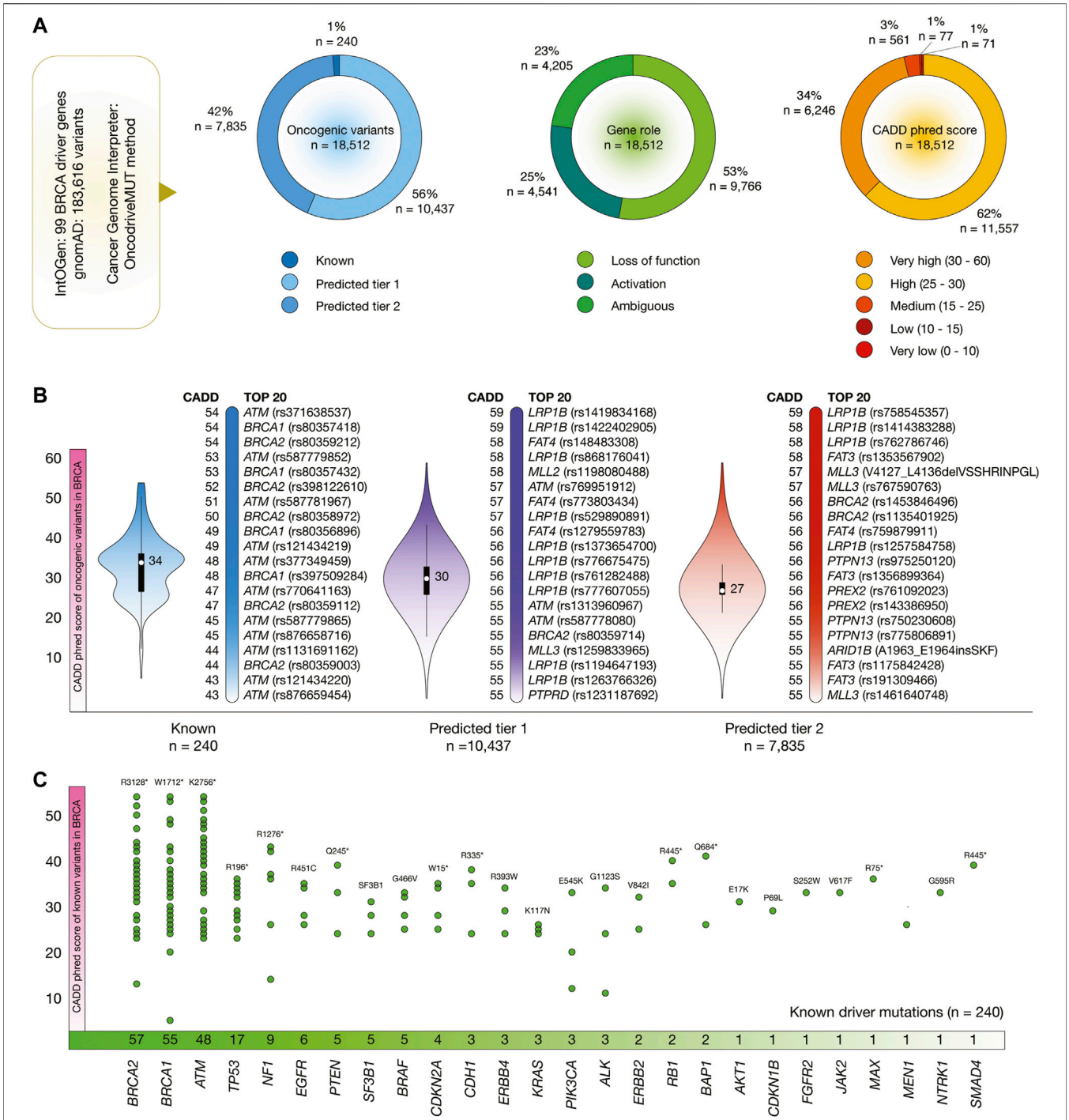


FIGURE 3 | BRCA driver genes, oncogenic variants, and deleteriousness scores. **(A)** Features of BRCA driver genes and oncogenic variants. **(B)** Bean plots of CADD phred scores of BRCA oncogenic variants, and ranking of known oncogenic variants with the highest CADD phred scores. **(C)** BRCA driver genes with the highest number of oncogenic variants.

the highest score was *LRP1B* rs1419834168 (59), and the predicted tier 2 variant with the highest score was *LRP1B* rs758545357 (54). The ranking of the 18,512 BRCA oncogenic variants is fully detailed in the **Supplementary Table S2**.

Figure 3C details the number of known oncogenic variants per BRCA driver gene. Genes with the highest number of known variants were *BRCA2* (*n* = 57), *BRCA1* (55), *ATM* (48), *TP53* (17), *NF1* (9), *EGFR* (6), *PTEN* (5), *SF3B1* (5), *BRAF* (5), and *CDKN2A* (4).

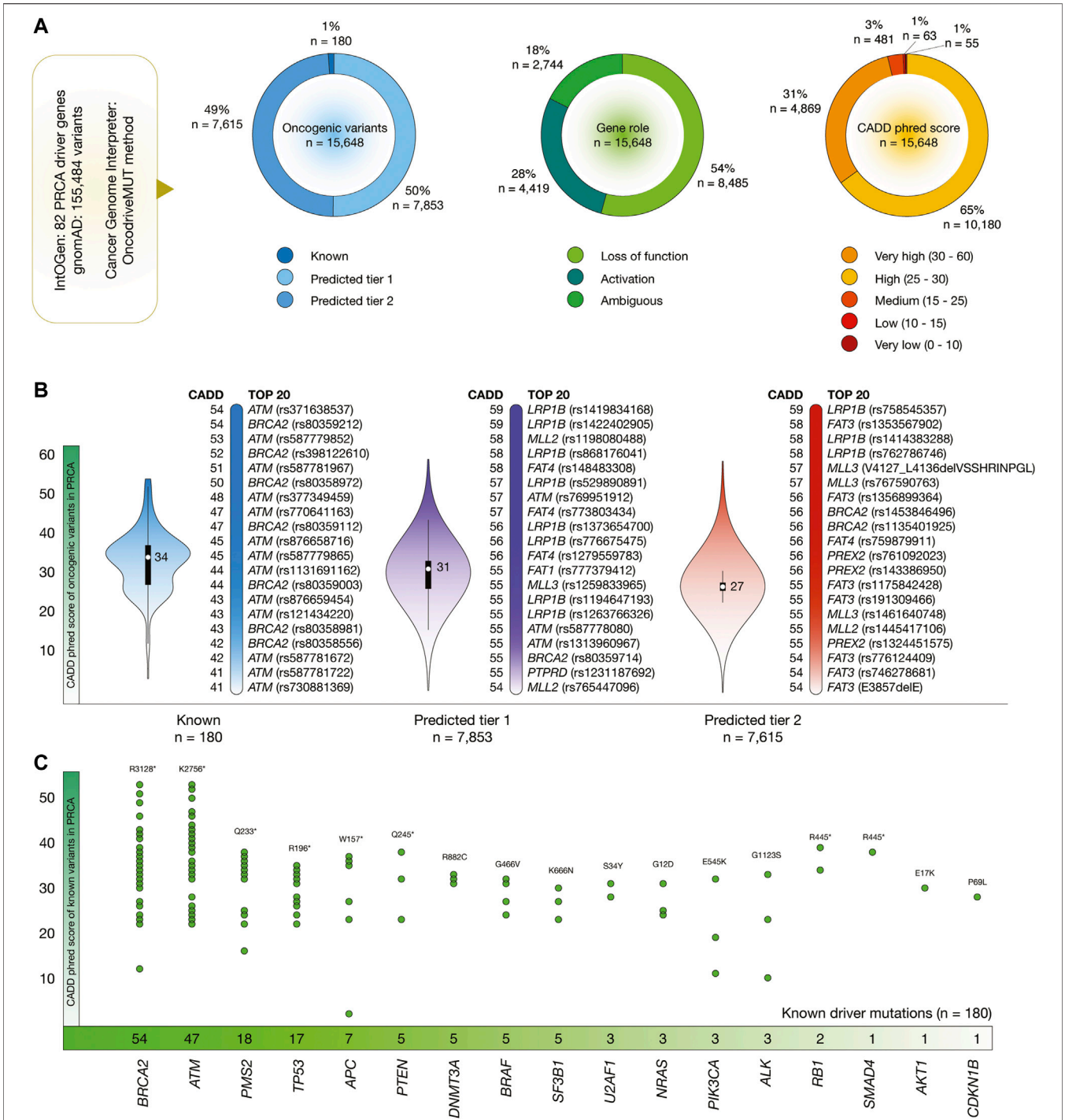


FIGURE 4 | PRCA driver genes, oncogenic variants, and deleteriousness scores. **(A)** Features of PRCA driver genes and oncogenic variants. **(B)** Bean plots of CADD phred scores of PRCA oncogenic variants, and ranking of known oncogenic variants with the highest CADD phred scores. **(C)** PRCA driver genes with the highest number of oncogenic variants.

Figure 4 shows the results of the OncodriveMUT analysis to identify oncogenic variants in 82 PRCA driver genes. After the analysis of 155,484 variants, we identified 15,648 oncogenic variants. Of them, 180 (1%) were known, 7,853 (50%) were predicted tier 1, and 7,615 (49%) were predicted tier 2 oncogenic variants. Regarding gene role, 8,485 (54%) variants produced protein loss of function and 4,419 (28%) variants produced protein activation. Additionally, 4,869 (31%) oncogenic variants had very high CADD phred score, and 10,180 (65%) variants had high (**Figure 4A**).

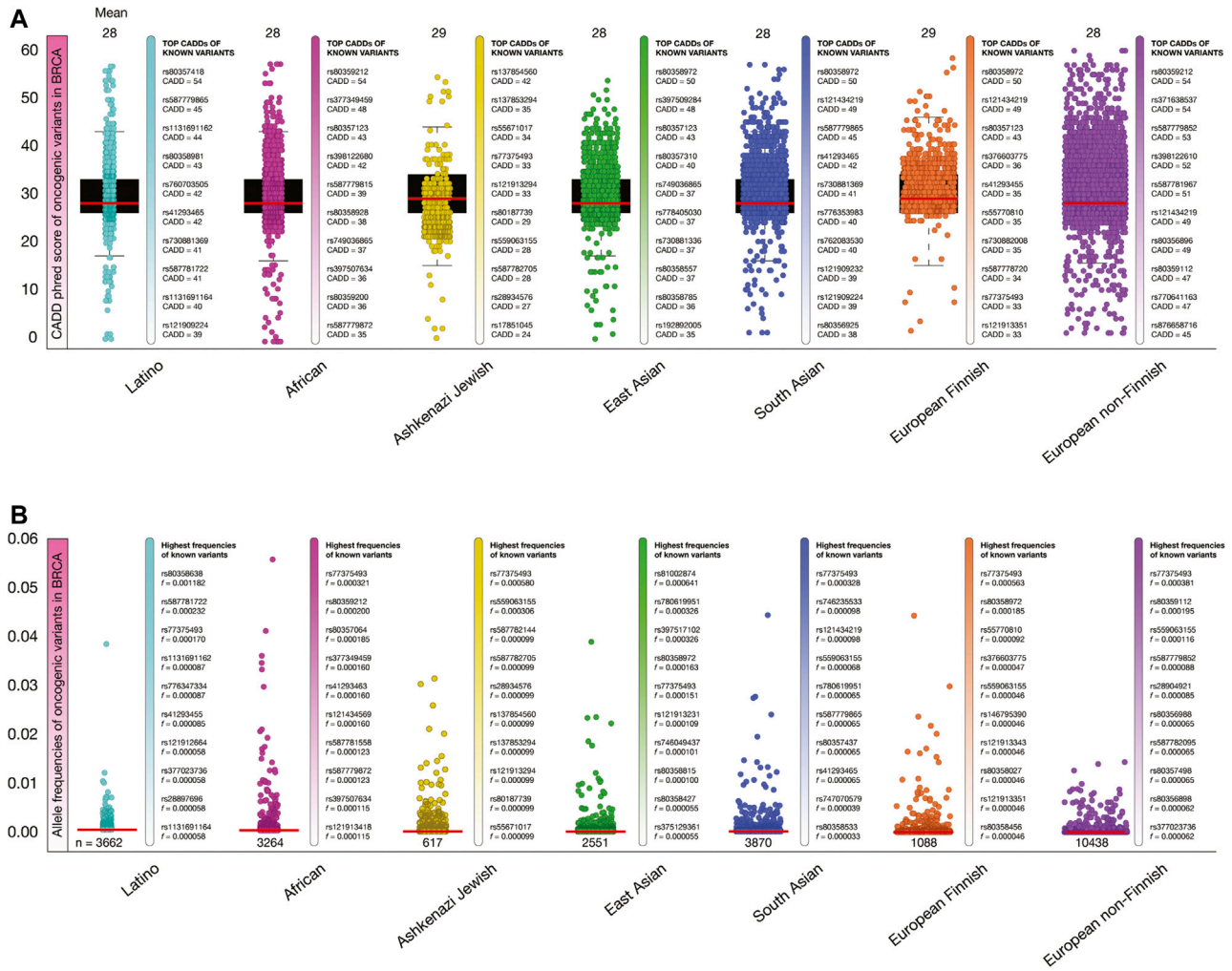


FIGURE 5 | Deleteriousness scores and allele frequencies of BRCA oncogenic variants per human population. **(A)** Box plots show CADD phred scores of BRCA known and predicted oncogenic variants per human population. Vertical bars show ranking of known oncogenic variants with the highest CADD phred scores per human population. Red lines show the mean of deleteriousness scores per human population. **(B)** Box plots show allele frequencies of known and predicted BRCA oncogenic variants per human population. Vertical bars show ranking of known oncogenic variants with the highest allele frequencies per human population. Red lines show the mean of deleteriousness scores per human population. **(C)** Ranking of known oncogenic variants with the highest allele frequencies identified in breast cancer patients from TCGA-BRCA project.

Figure 4B shows violin plots and ranking of CADD phred score of known, predicted tier 1 and tier 2 oncogenic variants in PRCA. The known oncogenic variant with the highest score was *ATM* rs371638537 (score = 54), the predicted tier 1 variant with the highest score was *LRP1B* rs1419834168 (59), and the predicted tier 2 variants with the highest score was *LRP1B* rs758545357 (59). The ranking of the 15,648 PRCA oncogenic variants is fully detailed in the **Supplementary Table S3**.

Finally, **Figure 4C** details the number of known oncogenic variants per PRCA driver gene. Genes with the highest number of known variants were *BRCA2* (*n* = 54), *ATM* (47), *PMS2* (18), *TP53* (17), *APC* (7), *PTEN* (5), *DNMT3A* (5), *BRAF* (5), *SF3B1* (5), and *U2AF1* (3).

Deleteriousness Scores, Allele Frequencies, and Validation of Oncogenic Variants per Human Population

Figure 5A shows box plots of CADD phred score of BRCA oncogenic variants per human population. The mean of deleteriousness score was 29 in Ashkenazi Jewish and South Asian populations, and 28 in Latino, African, East Asian, European Finnish and European non-Finish populations. The known oncogenic variant with the highest CADD phred score in Latinos was *BRCA1* rs80357418 (score = 54), in Africans was *BRCA2* rs80359212 (54), in Ashkenazi Jewish was *NF1* rs137854560 (42), in East Asians was *BRCA2* rs80358972 (50), in South Asians was *BRCA2* rs80358972 (50), in European

FIGURE 6 | Deleteriousness scores and allele frequencies of PRCA oncogenic variants per human population. **(A)** Box plots show CADD phred scores of PRCA known and predicted oncogenic variants per human population. Vertical bars show ranking of known oncogenic variants with the highest CADD phred scores per human population. Red lines show the mean of deleteriousness scores per human population. **(B)** Box plots show allele frequencies of known and predicted PRCA oncogenic variants per human population. Vertical bars show ranking of known oncogenic variants with the highest allele frequencies per human population. Red lines show the mean of deleteriousness scores per human population. **(C)** Ranking of known oncogenic variants with the highest allele frequencies identified in prostate cancer patients from TCGA-PRCA project.

Finnish was *BRCA2* 80358972 (50), and in European non-Finnish was *BRCA2* rs803559212 (54).

Figure 5B shows box plots of allele frequencies of oncogenic variants in BRCA driver genes per human population. The known oncogenic variant with the highest allele frequency in Latinos was *BRCA2* rs80358638 (f = 0.001182), in Africans was *JAK2* rs77375493 (0.000321), in Ashkenazi Jewish was *JAK2* rs77375493 (0.000580), in East Asians was *ERBB4* rs535202189 (0.000870), in South Asians was *JAK2* rs77375493 (0.000328), in European Finnish was *JAK2* rs77375493 (0.000563), and in European non-Finnish was *JAK2* rs77375493 (0.000381). The allele frequencies of the 18,512 BRCA oncogenic variants are fully detailed in the **Supplementary Table S4**.

Additionally, we identified some of the previously obtained allele frequencies of known oncogenic variants using the TCGA-BRCA project. From the 240 known oncogenic variants, we identified 32 variants and calculated its allele frequencies in 850 TCGA-BRCA patients with ethnicity data. The known oncogenic variants with the highest allele frequencies in the 162 black/African individuals were *PIK3CA* H1047R (f = 0.068), *AKT1* E17K (0.037), *PIK3CA* E542K (0.037), *TP53* R175H (0.037), and *TP53* Y220C (0.019); in the 600 white individuals (not Hispanic or Latino) were *PIK3CA* H1047R (0.419), *PIK3CA* E542 (0.154), *AKT1* E17K (0.093), *TP53* R175H (0.056), and *PIK3CA* H1047L (0.056); in the 29 white individuals (Hispanic or Latino) were *PIK3CA* H1047R (0.019), *PIK3CA* E542K (0.006), *PIK3CA* H1047L (0.006), *TP53* R273H (0.006), and *CDH1* Q23* (0.006); and in the 59 Asian individuals were *PIK3CA* H1047R (0.068), *PIK3CA* E542K (0.037), *TP53* R273H (0.012),

AKT1 E17K (0.037), and *TP53* R175H (0.037) (**Supplementary Table S5**).

Figure 6A shows box plots of CADD phred score of PRCA oncogenic variants per human population. The mean of deleteriousness score was 28 in the seven human populations. The known oncogenic variant with the highest CADD phred score in Latinos was *ATM* rs587779865 (score = 45), in Africans was *BRCA2* rs80359212 (54), in Ashkenazi Jewish was *PMS2* rs200640585 (37), in East Asians was *BRCA2* rs80358972 (50), in South Asians was *BRCA2* rs80358972 (50), in European Finnish was *BRCA2* rs80358972 (50), and in European non-Finnish was *ATM* rs371638537 (54).

Figure 6B shows box plots of allele frequencies of oncogenic variants in PRCA driver genes per human population. The known oncogenic variant with the highest allele frequency in Latinos was *BRCA2* rs80358638 (f = 0.001182), in Africans was *DNMT3A* rs147001633 (0.000616), in Ashkenazi Jewish was *SF2B1* rs559063155 (0.000306), in East Asians was *PMS2* rs63751228 (0.000643), in South Asians was *DNMT3A* rs147001633 (0.000229), in European Finnish was *BRCA2* rs81002862 (0.000370), and in European non-Finnish was *PMS2* rs121434629 (0.000295). The allele frequencies of the 15,648 PRCA oncogenic variants are fully detailed in the **Supplementary Table S6**.

Lastly, we identified some of the previously obtained allele frequencies of known oncogenic variants using the TCGA-PRAD project. From the 180 known oncogenic variants, we identified five variants and calculated its allele frequencies in 150 TCGA-PRAD patients. The known oncogenic variant with the highest allele frequencies in the seven black/African individuals was *TP53* R158H (f = 0.142); and in the 143 white individuals (not Hispanic or Latino) were AKT1 E17K (0.007), TP53 R175H (0.007), TP53 R282W (0.007), TP53 R248Q (0.007), and TP53 R158H (0.007) (**Supplementary Table S7**).

Current Pharmacogenomics Guidelines in Clinical Practice

PharmGKB details the current status of pharmacogenomics guidelines applied in clinical practice of patients with BRCA and PRCA. Clinical annotations provide information about variant-drug pairs based on variant annotations (Whirl-Carrillo et al., 2012; Barbarino et al., 2018). Regarding BRCA, there are currently 160 clinical annotations with responsive and resistant effects involving 73 human genes. Of them, 47 clinical annotations have responsive drug effects on 30 human proteins as shown in **Figure 7A**, and 12 clinical annotations have responsive and resistant drug effects on BRCA driver genes. For instance, carboplatin, docetaxel, and trastuzumab have efficacy on patients with *ERBB3* rs773123, *ERBB3* rs2229046, and *ERBB2* rs1136201; docetaxel and epirubicin have efficacy on patients with *MDM4* rs1563828; exemestane generates toxicity on patients with *ESR1* rs2813543; everolimus produces toxicity on patients with *PIK3R1* rs10515074; tamoxifen generates toxicity on patients with *NCOA1* rs1804645; cyclophosphamide, doxorubicin, and fluorouracil produces toxicity on patients with *ATM* rs1801516; examestane and letrozole generates toxicity on patients with *ESR1* rs9322335; letrozole produces toxicity on patients with *ERS1* rs4870061; cyclophosphamide, epirubicin, and fluorouracil generates toxicity on patients with *TP53* rs4968187; and trastuzumab produces toxicity on patients with *ERBB2* rs1136201 (**Supplementary Table S8**). Regarding PRCA, there are currently 33 clinical annotations with responsive and resistant drug effects involving 15 human genes. Of them, 14 clinical annotations have responsive drug effects on six human proteins as shown in **Figure 7B**, but no clinical association is related to PRCA driver genes (**Supplementary Table S9**). However, the identification of numerous oncogenic variants with high deleteriousness scores in BRCA and PRCA driver genes provides the ability to improve drug discovery on potential therapeutic targets.

In silico Drug Prescription

One impressive resource that the CGI employs is the Cancer Biomarker database, an extension of a previous collection of genomic biomarkers of anti-cancer drug response, which contains 310 drugs across 130 cancer types (Dienstmann et al., 2015). **Figure 8A** shows a circos plot of putative biomarkers of drug response involved in BRCA treatments. Individuals with *AKT* oncogenic mutations have responsive treatments with non-allosteric and allosteric AKT inhibitors; *BRCA1* and *BRCA2* oncogenic mutations with PARP inhibitor (veliparib) and chemotherapy (cisplatin); *CDKN2A* oncogenic mutations with AURKA-VEGF inhibitor (ilorasertib); *ERBB2* oncogenic mutations with ERBB inhibitor (neratinib); *ESR1* oncogenic mutations with hormonal therapy (fluvestrant); *HRAS* oncogenic mutations with farnesyltransferase inhibitor (tipifarnib); *NOTCH2* oncogenic mutations with gamma secretase inhibitor (mk-0752); *PIK3CA* oncogenic mutations with MTOR inhibitor (everolimus) plus ERBB2 mAb inhibitor (trastuzumab), PIK3CA inhibitors, PI3K pathway inhibitors and AKT inhibitors; and *PTEN* oncogenic mutations with MTOR inhibitor (sirolimus). On the other hand, individuals with *TP53* oncogenic mutations have resistant treatment with CD4/6 inhibitor (abemaciclib); and *ESR1* oncogenic mutations with hormonal therapy (exemestane). All data is fully detailed in **Supplementary Table S10**.

Figure 9A shows a circos plot of putative biomarkers of drug response involved in PRCA treatments. Individuals with the *AKT1* E17K oncogenic mutation have responsive treatment with non-allosteric and allosteric AKT inhibitors; *ATM* oncogenic mutations with PARP inhibitor (olaparib); *BRCA2* oncogenic mutations with PARP inhibitor (olaparib); *HRAS* oncogenic mutations with farnesyltransferase inhibitor (tipifarnib); and *PTEN* oncogenic mutations with MTOR inhibitors (sirolimus and everolimus) (**Figure 8A**). All data is fully detailed in **Supplementary Table S11**.

Clinical Trials

The Open Targets Platform shows the current status of clinical trials involved in BRCA and PRCA driver proteins (Carvalho-Silva et al., 2019). In regards to BRCA, there were 369 clinical trials in phase 3 (94%) and phase 4 (6%). Small molecules were the most analyzed type of drugs (61%), followed by antibodies

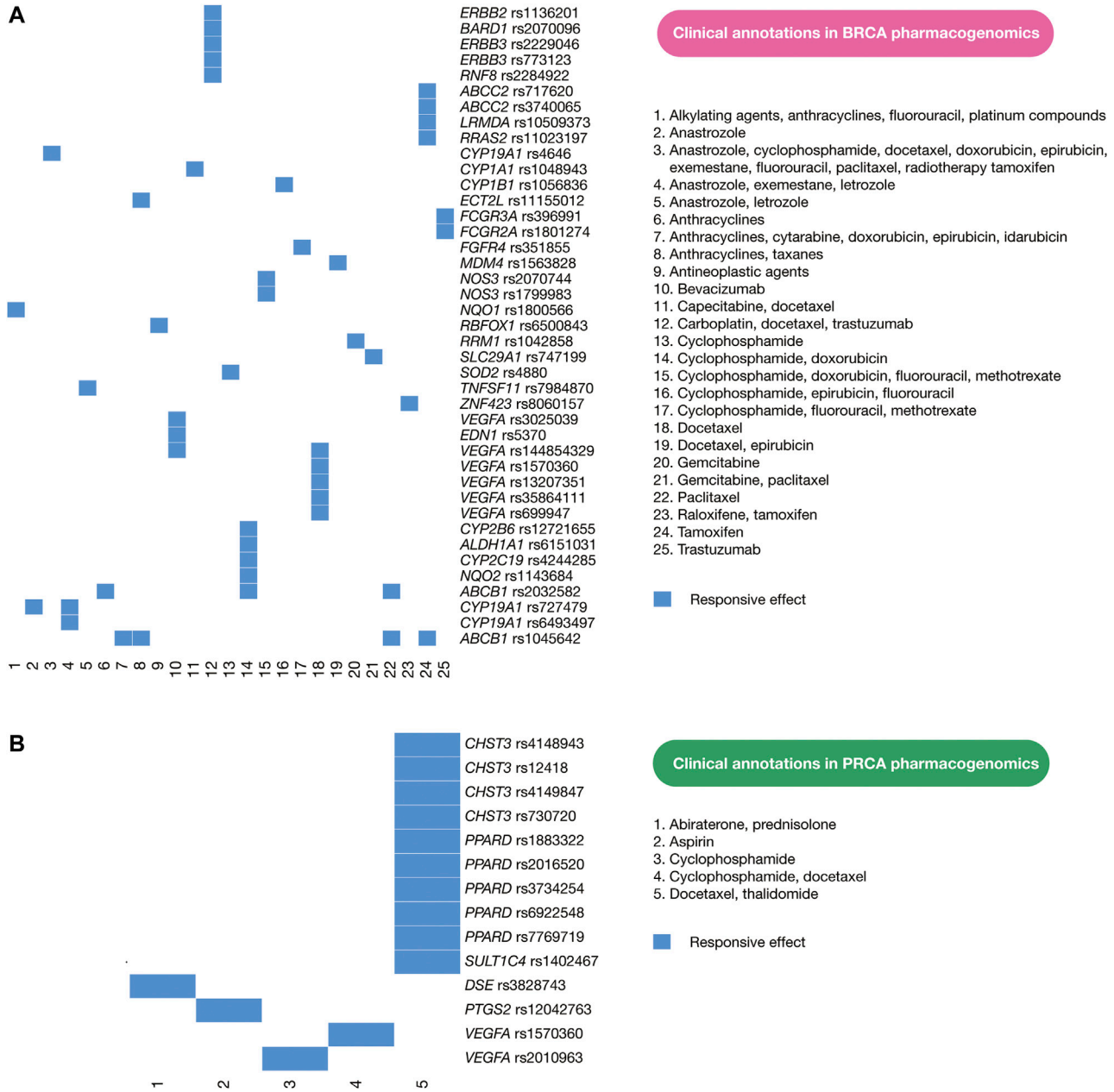


FIGURE 7 | Current pharmacogenomics guidelines. **(A)** Clinical annotations of BRCA pharmacogenomics focused on responsive drug effects on human proteins. **(B)** Clinical annotations of PRCA pharmacogenomics focused on responsive drug effects on human proteins. These clinical annotations encompass cancer driver proteins and non-cancer driver proteins.

(39%). Inhibitors were the most predominantly action type of drugs (68%), followed by modulators (21%), antagonists (9%), binding agents (1%), and non-allosteric modulators (1%). The target classes with the highest number of clinical trials was AGC kinase AKT family (60%), followed by enzymes (31%), NMDA receptors (7%), nuclear hormone receptors (0.5%), TKL kinase RAF family (0.5%), transferases (0.5%), and tyrosine protein kinases (0.5%) (**Figure 8B**). On the other hand, the Sankey plot showed 25 drugs currently analyzed in 369 clinical trials in 11 BRCA driver proteins. Druggable proteins with the highest number of clinical trials were ERBB2 (n = 178), ESR1 (n = 114), EGFR (n = 35), and POLD1 (n = 21). Lastly, drugs with the highest number of clinical trials in advanced stages were trastuzumab (n = 100), a recombinant humanized IgG1 monoclonal antibody against the ERBB2 receptor (Bange et al., 2001); tamoxifen (n = 73) that inhibits estrogen binding to its receptor (Jordan, 1993); lapatinib (n = 50) that is a 4-anilinoquinazoline kinase inhibitor of the intracellular tyrosine

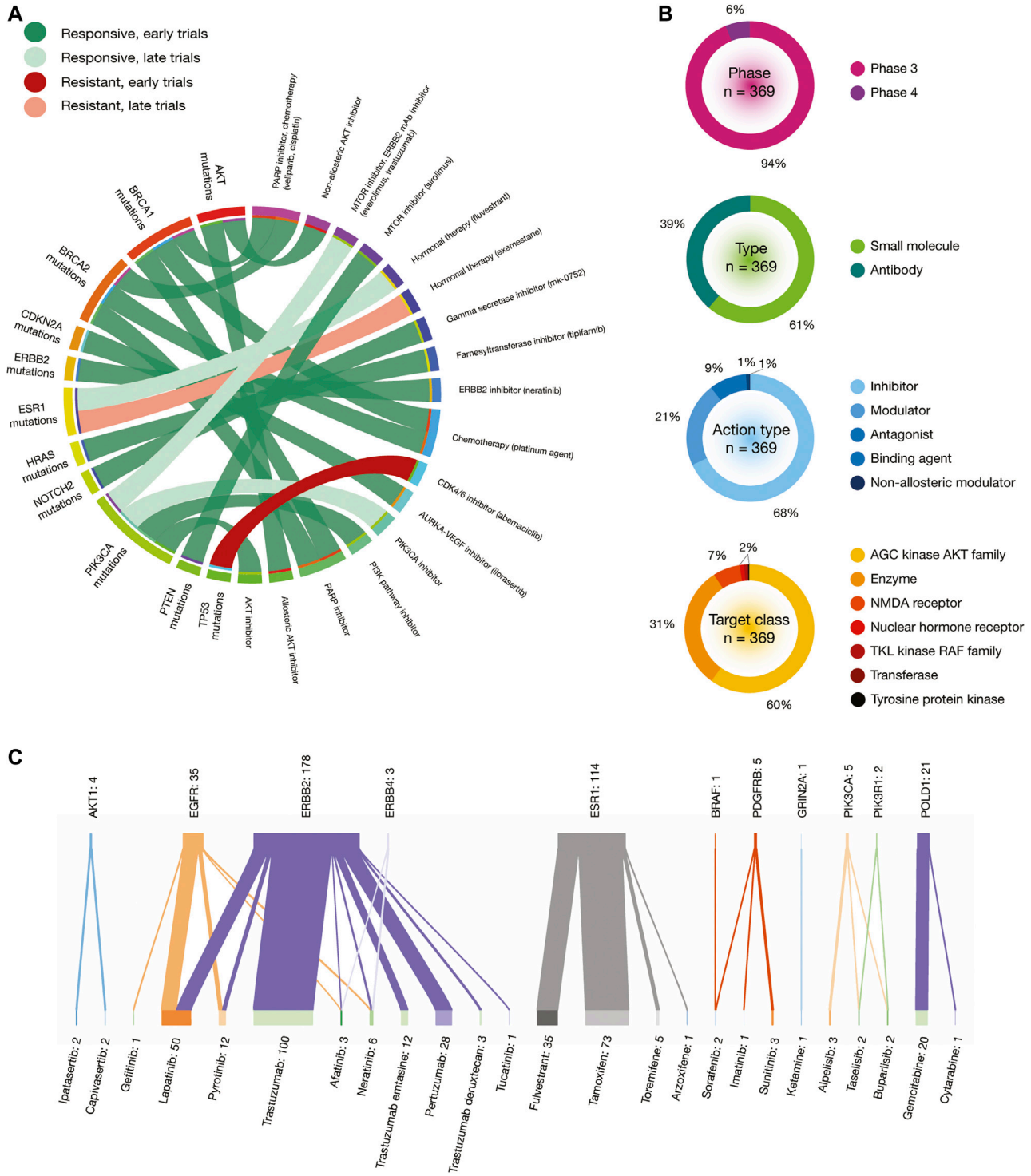


FIGURE 8 | *In silico* drug prescriptions and clinical trials involved in BRCA. **(A)** Circos plot showing precision medicine application between drugs and BRCA driver mutations. **(B)** Clinical trial features on BRCA. **(C)** Sankey plot of drugs with the highest number of clinical trials on BRCA driver proteins.

kinase domain of EGFR and ERBB2 (Xia et al., 2002); fulvestrant (n = 35) that achieves its anti-estrogen effects through downregulation and degradation of estrogen receptors (Chen et al., 2002); and pertuzumab (n = 28), a monoclonal antibody that targets the extracellular dimerization domain of ERBB2, thereby inhibiting intracellular signaling via the PI3K and MAP kinase pathways (Adams et al., 2006) (**Figure 8C**). All data is fully detailed in **Supplementary Table S12**.

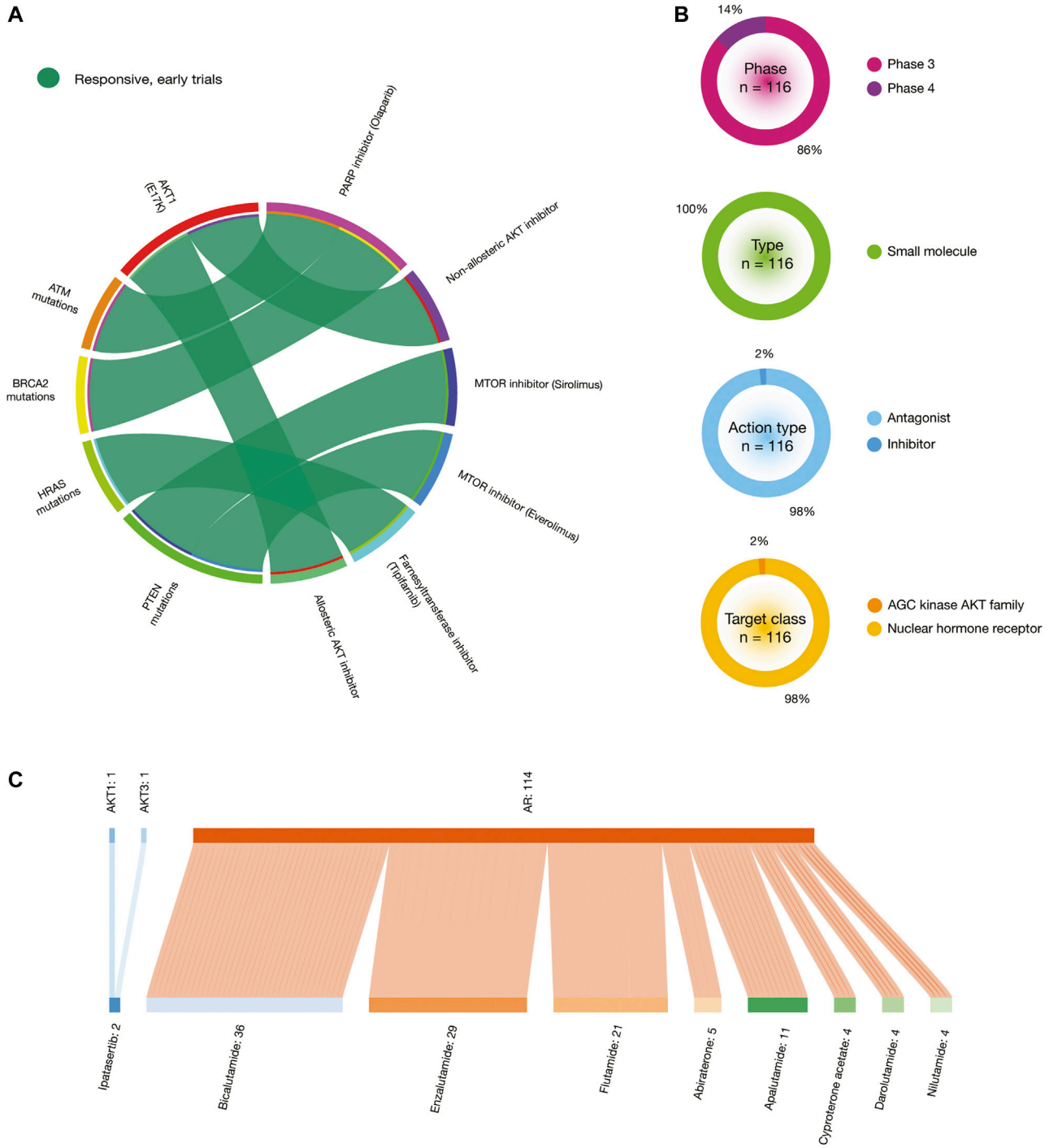


FIGURE 9 | *In silico* drug prescriptions and clinical trials involved in PRCA. **(A)** Circos plot showing precision medicine application between drugs and PRCA driver mutations. **(B)** Clinical trial features on PRCA. **(C)** Sankey plots of drugs with the highest number of clinical trials on PRCA driver proteins.

Regarding PRCA, there are 116 clinical trials in phase 3 (86%) and phase 4 (14%). Small molecules were the only type of drugs analyzed (100%). Antagonists were the most predominantly action type of drugs (98%), followed by inhibitors (2%). The target classes with the highest number of clinical trials were AGC kinase AKT family (98%) and nuclear hormone receptors (2%) (**Figure 9B**). On the other hand, the Sankey plot showed nine drugs that are being analyzed in 116 clinical trials in 3 PRCA driver proteins. Druggable proteins with the highest number of clinical trials were AR ($n = 114$), AKT1 ($n = 1$), AKT3 ($n = 1$), and

POLD1. Lastly, drugs with the highest number of clinical trials in advanced stages were bicalutamide ($n = 36$), a small molecule that blocks the action of androgens of adrenal and testicular origin (Chang et al., 1999); enzalutamide ($n = 29$), an androgen receptor inhibitor for the treatment of castration-resistant prostate cancer (Nadiminty et al., 2013); flutamide ($n = 21$), a nonsteroidal antiandrogen that blocks the action testosterone by binding to the androgen receptor (Balk, 2002); apalutamide ($n = 11$) that impairs the translocation of AR from the cytoplasm to the nucleus; and abiraterone (n = 5), a small molecule that is a derivative of steroidal progesterone and is an orally active inhibitor of CYP17A1 (de Bono et al., 2011) (**Figure 9C**). All data is fully detailed in **Supplementary Table S13**.

DISCUSSION

Precision oncology is a treatment paradigm that takes into account the molecular and cellular features of a tumor as well as its environment and additional traits of the individual, such as genetics and lifestyle, to create a tailor-made treatment (Le Tourneau et al., 2019). Most molecular alterations in tumors exist in multiple tumor types, and it has been hypothesized that anticancer therapy should be tailored to each patient according to their tumor molecular profile. Hence, the interpretation of molecular profiles through bioinformatics tools is imperative to analyze omics data and provide the most effective therapy to patients (Valencia and Hidalgo, 2012).

The most important aim in the interpretation of cancer genomes is to identify the variants responsible for tumorigenic traits. In this context, OncodriveMUT is a machine-learning approach integrated into the CGI platform to assess oncogenic variant's tumorigenic potential. OncodriveMUT combines genomic features such as gene signals of positive selection, clusters of somatic mutations, gene mechanism of action, and regions depleted by germline variants (Tamborero et al., 2018). In this study, we analysed 183,616 variants located into 99 BRCA driver genes, and identified 18,512 known and predicted oncogenic variants. Of them, 240 were known oncogenic variants, and 9,766 were loss-of-function variants. Additionally, we analysed 155,484 variants located into 82 PRCA driver genes, and identified 15,648 known and predicted oncogenic variants. Of them, 180 were known oncogenic variants, and 8,485 were loss-of-function variants. Consequently, we calculated the CADD phred scores that represents the deleteriousness of single nucleotide variants as well as insertion/deletion variants involved in the molecular landscape of oncogenesis (Kircher et al., 2014; Rentzsch et al., 2019). The known BRCA oncogenic variants with the highest deleteriousness scores were *ATM* rs371638537 (CADD = 54), *BRCA1* rs80357418 (CADD = 54), and *BRCA2* rs80359212 (CADD = 54) (**Figure 3B**); and the known PRCA oncogenic variants with the highest deleteriousness scores were *ATM* rs371638537 (CADD = 54), and *BRCA2* rs80359212 (CADD = 54) (**Figure 4B**).

The ability to identify oncogenic variants and their deleteriousness scores in BRCA and PRCA tumors is an important step to apply PGx in clinical practice. Nevertheless, there are two main barriers for implementing PGx in developing regions. On the one hand, the most relevant cancer genome projects worldwide, such as TCGA(The Cancer Genome Atlas Research Network, 2013), TARGET or PCAWGC (ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, 2020), are overrepresented by Caucasian individuals (91.1%), and do not include enough individuals from minority populations (Guerrero et al., 2018). On the other hand, developing regions lack of investment in cancer genomics tests, have fragmented healthcare systems, and have insufficient characterization of pharmacogenetics variability in their populations (Quinones et al., 2014). Therefore, in this study we proposed a new insight for identification of the most frequent oncogenic variants in the Latino, African, Ashkenazi Jewish, East Asian, South Asian, European Finnish, and European non-Finnish populations in order to focus economic resources on analyzing the most frequent and relevant molecular targets.

The gnomAD database harmonize exome and genome sequencing data from a variety of large-scale sequencing projects worldwide (Karczewski et al., 2020). We calculated allele frequencies of the previously identified known and predicted BRCA and PRCA oncogenic variants from Latinos, Africans, Ashkenazi Jewish, East Asians, South Asians, European Finnish, and European non-Finnish. Regarding BRCA, there are 42 known oncogenic variants with allele frequencies >0 in Latinos, 32 in Africans, 11 in Ashkenazi Jewish, 36 in East Asians, 35 in South Asians, 19 in European Finnish, and 156 in European non-Finnish (**Supplementary Table S4**). Regarding PRCA, there are 33 known oncogenic variants with allele frequencies >0 in Latinos, 27 in Africans, 12 in Ashkenazi Jewish, 25 in East Asians, 28 in South Asians, 15 in European Finnish, and 117 in European non-Finnish (**Supplementary Table S5**). Nevertheless, not all proteins carrying these oncogenic variants are actionable therapeutic targets or have clinical annotations in PGx guidelines.

Consequently, the second major aim of the effort to interpret cancer genomes is to identify which oncogenic variants may shape the response to anti-cancer therapies. After identifying the most frequent oncogenic variants per human population, we integrated these results with the current clinical annotations of the PGx guidelines from PharmGKB (Whirl-Carrillo et al., 2012; Barbarino et al., 2018), with the *in silico* drug prescriptions from the Cancer Genome Interpreter (Tamborero et al., 2018), and with the current clinical trials in advanced stages from the Open Targets Platform (Carvalho-Silva et al., 2019). The main idea of the integration of precision oncology per human population is to prioritize the possible oncogenic variants found in cancer patients, focusing economic resources for PGx testing in a most effective way.

Figure 10 summarizes the integration of precision oncology of BRCA per human population. There are 11 druggable driver proteins carrying 138 known oncogenic variants with the highest deleteriousness scores, and the highest allele frequencies per human population. Latinos have 16 variants in five actionable therapeutic targets (BRCA1, BRCA2, CDKN2A, PTEN, and EGFR); Africans have 13 variants in four druggable proteins

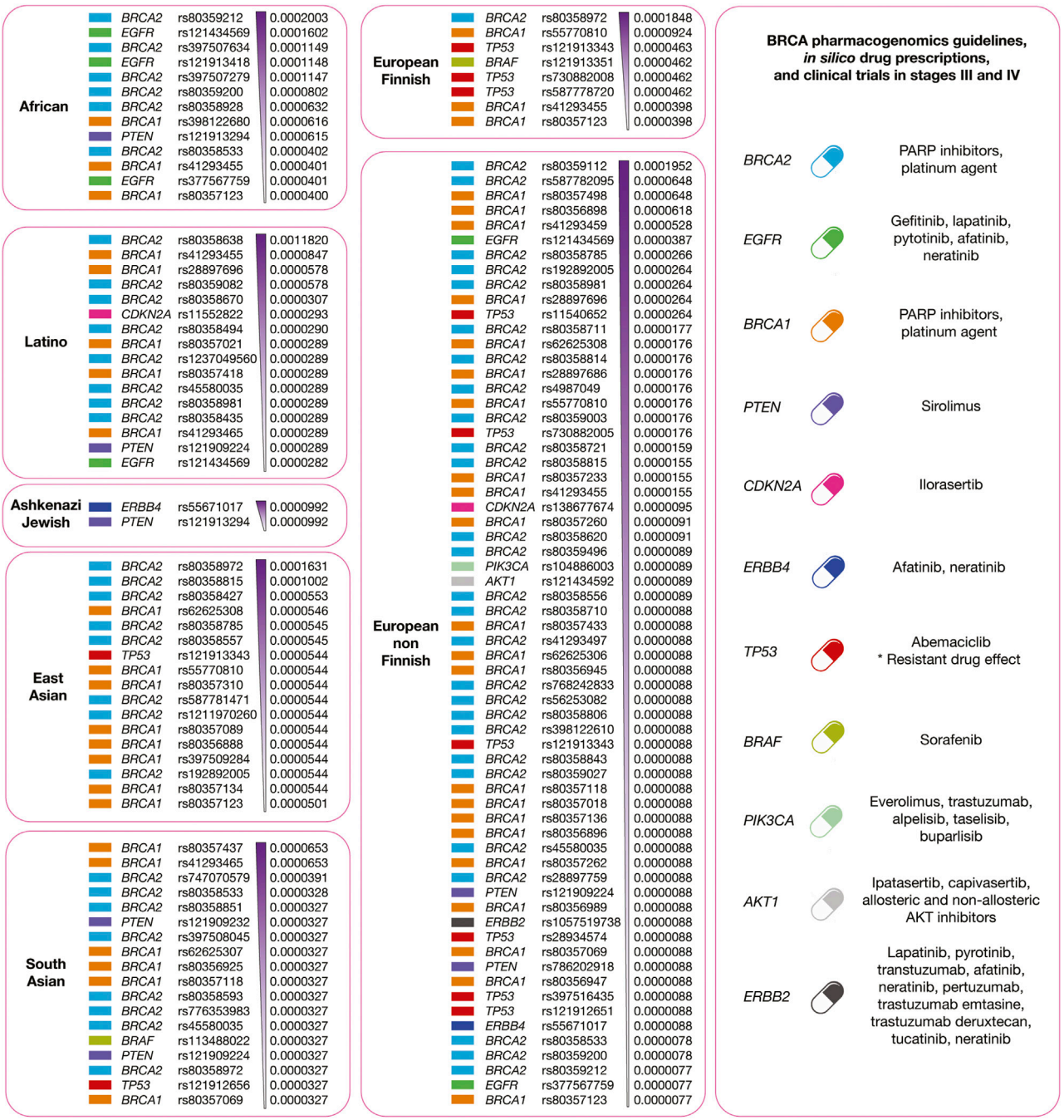


FIGURE 10 | Precision medicine of BRCA per human population. On the left, druggable BRCA driver proteins carrying known oncogenic variants with the highest allele frequencies and the highest deleteriousness scores per human population. On the right, integration of pharmacogenomics guidelines, drug prescription, and clinical trials according to druggable BRCA driver proteins.

(BRCA1, BRCA2, EGFR, and PTEN); Ashkenazi Jewish has 2 variants in 2 actionable therapeutic targets (ERBB4 and PTEN); East Asians have 17 variants in 3 druggable proteins (BRCA1, BRCA2, and TP53); European Finnish have eight variants in four actionable therapeutic targets (BRCA1, BRCA2, TP53, and BRAF); European non-Finnish have 64 variants in 10 druggable proteins (BRCA1, BRCA2, EGFR, TP53, CDKN2A, PIK3CA, AKT1, PTEN, ERBB2, and ERBB4); and South Asians have 18 variants in five actionable therapeutic targets (BRCA1, BRCA2, PTEN, BRAF, and TP53) (**Supplementary Table S14**). Regarding BRCA responsive treatments, ipatasertib, capivasertib, allosteric AKT inhibitors, and non-allosteric AKT inhibitors act on AKT (Kostaras et al., 2020); veliparib and cisplatin respond on BRCA1 and BRCA2 (Diéras et al., 2020); ilorasertib reacts on

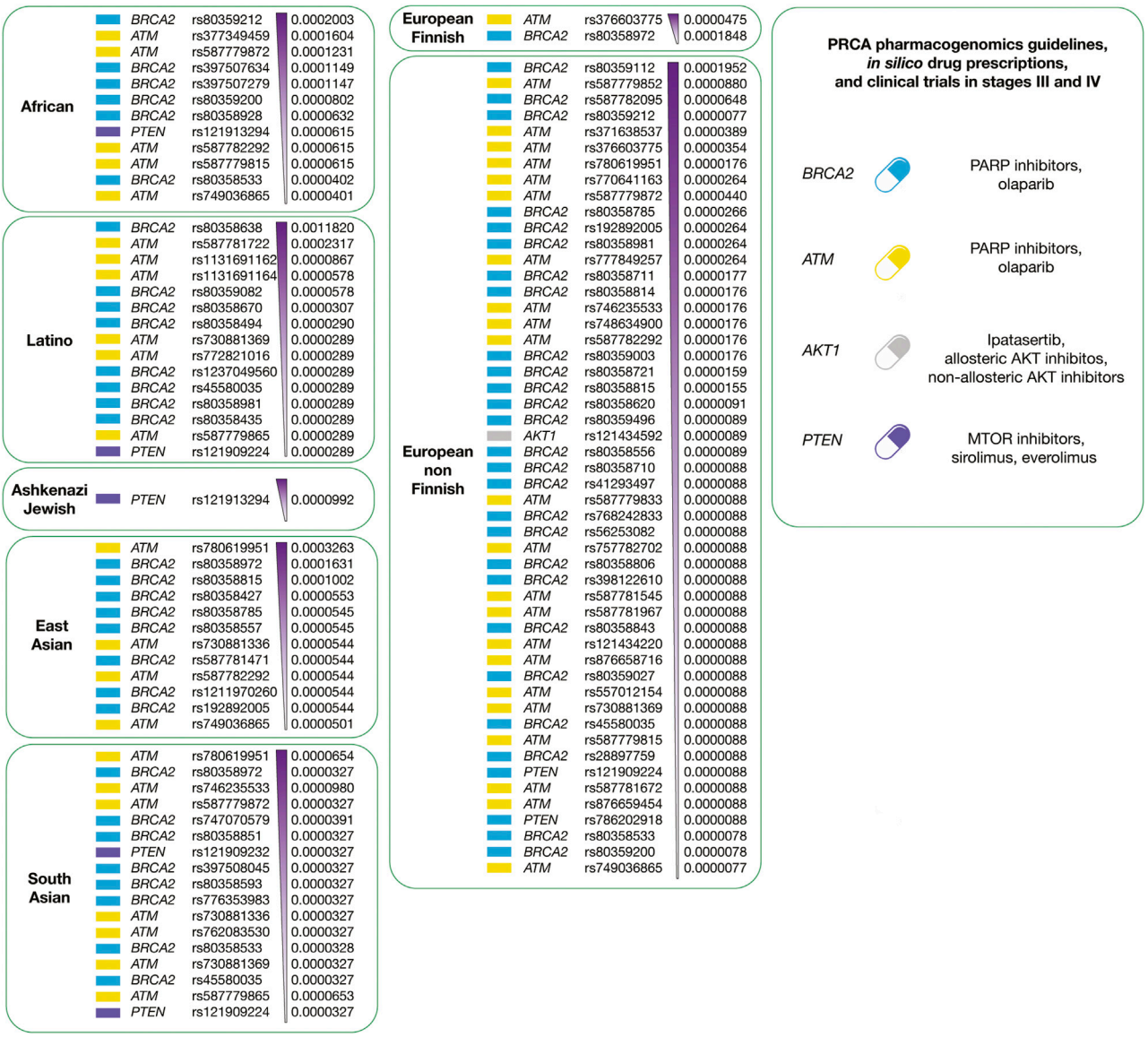


FIGURE 11 | Precision medicine of PRCA per human population. On the left, druggable PRCA driver proteins carrying known oncogenic variants with the highest allele frequencies and the highest deleteriousness scores per human population. On the right, integration of pharmacogenomics guidelines, drug prescription, and clinical trials according to druggable PRCA driver proteins.

CDKN2A (Aftab et al., 2019); sirolimus acts on PTEN (Schmid et al., 2014); gefitinib, lapatinib, pytotinib, afatinib, and neratinib respond on EGFR; afatinib and neratinib react on ERBB4; sorafenib acts on BRAF; everolimus, trastuzumab, alpelisib, taselisib, and buparlisib work on PIK3CA (Chen et al., 2019); lastly, lapatinib, pyrotinib, transtuzumab, afatinib, neratinib, trastuzumab emtasine, pertuzumab, trastuzumab deruxtecan, tucatinib, and neratinib act on ERBB2 (Ben-Baruch et al., 2015).

Figure 11 summarizes the integration of precision oncology of PRCA per human population. There are four druggable driver proteins carrying 110 known oncogenic variants with the highest deleteriousness scores, and the highest allele frequencies per human population. Latinos have 15 variants in 3 actionable therapeutic targets (BRCA2, ATM, and PTEN); Africans have 12 variants in 3 druggable proteins (BRCA2, ATM, and PTEN); Ashkenazi Jewish has 1 variant in 1 actionable therapeutic target (PTEN); East Asians have 12 variants in 2 druggable proteins (BRCA2 and ATM); European Finnish have 2 variants in 2 actionable therapeutic targets (BRCA2 and ATM); European non-Finnish have 51 variants in four druggable proteins (BRCA2, AKT1, ATM, and PTEN); and South Asians have 17 variants in 3 actionable therapeutic targets (BRCA2, ATM, and PTEN) (**Supplementary Table S15**). Regarding PRCA responsive treatments, ipatasertib, allosteric and non-allosteric AKT inhibitors act on AKT1 (Mundi et al., 2016);

olaparib (PARP inhibitor) reacts on ATM and BRCA2 (de Bono et al., 2020); and lastly, sirolimus and everolimus (MTOR inhibitors) act on PTEN (Morgan et al., 2009).

In the era of precision oncology, PGx testing will make it possible to improve the efficiency on the use of resources, patient safety, and drug dosage in BRCA and PRCA treatments. Hence, it is imperative to unify efforts where developing countries might invest in obtaining databases of their population's genomic profiles, and developed countries might incorporate racial/ethnic minority populations in future clinical trials and cancer researches with the main aim of fomenting PGx in public health policies and clinical practice.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.The datasets generated for this study are included in this published article (and its Supplementary Information files).

AUTHOR CONTRIBUTIONS

AL-C conceived the subject, the conceptualization of the study, and wrote the manuscript. NV, PG-R, IA-C, SG, CA, TZ, and LQ edited the manuscript, did data curation, and gave valuable scientific input. NV and LQ did funding acquisition. Lastly, all authors reviewed and approved the manuscript.

FUNDING

Publication of this article was funded by University of Chile. This work was supported by the Latin American Society of Pharmacogenomics and Personalized Medicine (SOLFAGEM).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.630658/full#supplementary-material.

REFERENCES

Adams, C. W., Allison, D. E., Flagella, K., Presta, L., Clarke, J., Dybdal, N., et al. (2006). Humanization of a recombinant monoclonal antibody to produce a therapeutic HER dimerization inhibitor, pertuzumab. *Cancer Immunol. Immunother.* 55, 717–727. doi:10.1007/s00262-005-0058-x

Aftab, A., Shahzad, S., Hussain, H. M. J., Khan, R., Irum, S., and Tabassum, S. (2019). CDKN2A/P16INK4A variants association with breast cancer and their in-silico analysis. *Breast Cancer* 26, 11–28. doi:10.1007/s12282-018-0894-0

Arnedo-Pac, C., Mularoni, L., Muiños, F., Gonzalez-Perez, A., Lopez-Bigas, N., and Schwartz, R. (2019). OncodriveCLUSTL: a sequence-based clustering method to identify cancer drivers. *Bioinformatics* 35, 4788–4790. doi:10.1093/bioinformatics/btz501

Balk, S. P. (2002). Androgen receptor as a target in androgen-independent prostate cancer. *Urology* 60, 132–138. doi:10.1016/S0090-4295(02)01593-5

Bange, J., Zwick, E., and Ullrich, A. (2001). Molecular targets for breast cancer therapy and prevention. *Nat. Med.* 7, 548–552. doi:10.1038/87872

Bar-Joseph, Z., Siegfried, Z., Brandeis, M., Brors, B., Lu, Y., Eils, R., et al. (2008). Genome-wide transcriptional analysis of the human cell cycle identifies genes differentially regulated in normal and cancer cells. *Proc. Natl. Acad. Sci.* 105, 955–960. doi:10.1073/pnas.0704723105

Barbarino, J. M., Whirl-Carrillo, M., Altman, R. B., and Klein, T. E. (2018). PharmGKB: a worldwide resource for pharmacogenomic information. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 10, e1417. doi:10.1002/wsbm.1417

Ben-Baruch, N. E., Bose, R., Kavuri, S. M., Ma, C. X., and Ellis, M. J. (2015). HER2-mutated breast cancer responds to treatment with single-agent neratinib, a second-generation HER2/EGFR tyrosine kinase inhibitor. *JNCCN J. Natl. Compr. Cancer Netw.* 13, 1061–1064. doi:10.6004/jnccn.2015.0131

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* 68, 394–424. doi:10.3322/caac.21492

Carvalho-Silva, D., Pierleoni, A., Pignatelli, M., Ong, C. K., Fumis, L., Karamanis, N., et al. (2019). Open Targets Platform: new developments and updates two years on. *Nucleic Acids Res.* 47, D1056–D1065. doi:10.1093/nar/gky1133

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2, 401–404. doi:10.1158/2159-8290.CD-12-0095

Chae, Y. K., Anker, J. F., Carneiro, B. A., Chandra, S., Kaplan, J., Kalyan, A., et al. (2016). Genomic landscape of DNA repair genes in cancer. *Oncotarget* 7, 23312–23321. doi:10.18632/oncotarget.8196

Chang, H. C., Miyamoto, H., Marwah, P., Lardy, H., Yeh, S., Huang, K. E., et al. (1999). Suppression of Δ5-androstenediol-induced androgen receptor transactivation by selective steroids in human prostate cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* 96, 11173–11177. doi:10.1073/pnas.96.20.11173

Chen, X., Ji, Z. L., and Chen, Y. Z. (2002). TTD: therapeutic target database. *Nucleic Acids Res.* 30, 412–415. doi:10.1093/nar/30.1.412

Chen, Z., Zheng, Y., Cao, W., Zhang, Y., Zhao, Z., Wang, G., et al. (2019). Everolimus-containing therapy vs conventional therapy in the treatment of refractory breast cancer patients with PI3K/AKT/mTOR mutations: a retrospective study. *Cancer Med.* 8, 5544–5553. doi:10.1002/cam4.2460

Collins, R. L., Brand, H., Karczewski, K. J., Zhao, X., Alföldi, J., Francioli, L. C., et al. (2020). A structural variation reference for medical and population genetics. *Nature* 581, 444–451. doi:10.1038/s41586-020-2287-8

de Bono, J., Mateo, J., Fizazi, K., Saad, F., Shore, N., Sandhu, S., et al. (2020). Olaparib for metastatic castration-resistant prostate cancer. *N. Engl. J. Med.* 382, 2091–2102. doi:10.1056/nejmoa1911440

de Bono, J. S., Logothetis, C. J., Molina, A., Fizazi, K., North, S., Chu, L., et al. (2011). Abiraterone and increased survival in metastatic prostate cancer. *N. Engl. J. Med.* 364, 1995–2005. doi:10.1056/nejmoa1014618

Dienstmann, R., Jang, I. S., Bot, B., Friend, S., and Guinney, J. (2015). Database of genomic biomarkers for cancer drugs and clinical targetability in solid tumors. *Cancer Discov.* 5, 118–123. doi:10.1158/2159-8290.CD-14-1118

Diéras, V., Han, H. S., Kaufman, B., Wildiers, H., Friedlander, M., Ayoub, J. P., et al. (2020). Veliparib with carboplatin and paclitaxel in BRCA-mutated advanced breast cancer (BROCADE3): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol.* 21, 1269–1282. doi:10.1016/S1470-2045(20)30447-2

Dietlein, F., Weghorn, D., Taylor-Weiner, A., Richters, A., Reardon, B., Liu, D., et al. (2020). Identification of cancer driver genes based on nucleotide context. *Nat. Genet.* 52, 208–218. doi:10.1038/s41588-019-0572-y

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* 6, 1. doi:10.1126/scisignal.2004088

Green, E. D., Gunter, C., Biesecker, L. G., Di Francesco, V., Easter, C. L., Feingold, E. A., et al. (2020). Strategic vision for improving human health at the Forefront of Genomics. *Nature* 586, 683–692. doi:10.1038/s41586-020-2817-4

Guerrero, S., López-Cortés, A., Indacochea, A., García-Cárdenas, J. M., Zambrano, A. K., Cabrera-Andrade, A., et al. (2018). Analysis of racial/ethnic

representation in select basic and applied cancer research studies. *Sci. Rep.* 8, 13978. doi:10.1038/s41598-018-32264-x

Hentze, M. W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* 19, 327–341. doi:10.1038/nrm.2017.130

Hoadley, K. A., Yau, C., Hinoue, T., Wolf, D. M., Lazar, A. J., Drill, E., et al. (2018). Cell-of-Origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell* 173, 291–304. doi:10.1016/j.cell.2018.03.022

Huang, K. lin., Mashl, R. J., Wu, Y., Ritter, D. I., Wang, J., Oh, C., et al. (2018). Pathogenic germline variants in 10,389 adult cancers. *Cell* 173, 355–370. doi:10.1016/j.cell.2018.03.039

ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium (2020). Pan-cancer analysis of whole genomes. *Nature* 578, 82–93. doi:10.1038/s41586-020-1969-6

Jordan, V. C. (1993). A current view of tamoxifen for the treatment and prevention of breast cancer. *Br. J. Pharmacol.* 110, 507–517. doi:10.1111/j.1476-5381.1993.tb13840.x

Kader, F., and Ghai, M. (2017). DNA methylation-based variation between human populations. *Mol. Genet. Genomics.* 292, 5–35. doi:10.1007/s00438-016-1264-2

Kandoth, C., McLellan, M. D., Vandin, F., Ye, K., Niu, B., Lu, C., et al. (2013). Mutational landscape and significance across 12 major cancer types. *Nature* 502, 333–339. doi:10.1038/nature12634

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581, 434–443. doi:10.1038/s41586-020-2308-710.1530/ey.17.14.310.1530/ey.17.14.3

Kircher, M., Witten, D. M., Jain, P., O'roak, B. J., Cooper, G. M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* 46, 310–315. doi:10.1038/ng.2892

Kostaras, E., Kaserer, T., Lazaro, G., Heuss, S. F., Hussain, A., Casado, P., et al. (2020). A systematic molecular and pharmacologic evaluation of AKT inhibitors reveals new insight into their biological activity. *Br. J. Cancer* 123, 542–555. doi:10.1038/s41416-020-0889-4

Lawrence, M. S., Stojanov, P., Mermel, C. H., Robinson, J. T., Garraway, L. A., Golub, T. R., et al. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 505, 495–501. doi:10.1038/nature12912

Le Tourneau, C., Borcoman, E., and Kamal, M. (2019). Molecular profiling in precision medicine oncology. *Nat. Med.* 25 (5), 711–712. doi:10.1038/s41591-019-0442-2

López-Cortés, A., Jaramillo-Koupermann, G., Muñoz, M. J., Cabrera, A., Echeverría, C., Rosales, F., et al. (2013). Genetic polymorphisms in MTHFR (C677T, A1298C), MTR (A2756G) and MTRR (A66G) genes associated with pathological characteristics of prostate cancer in the ecuadorian population. *Am. J. Med. Sci.* 346, 447–454. doi:10.1097/MAJ.0b013e3182882578

López-Cortés, A., Guerrero, S., Redal, M. A., Alvarado, A. T., and Quiñones, L. A. (2017). State of art of cancer pharmacogenomics in Latin American populations. *Int. J. Mol. Sci.* 18 (6), 639. doi:10.3390/ijms18060639

López-Cortés, A., Paz-y-Miño, C., Cabrera-Andrade, A., Barigye, S. J., Munteanu, C. R., González-Díaz, H., et al. (2018). Gene prioritization, communality analysis, networking and metabolic integrated pathway to better understand breast cancer pathogenesis. *Sci. Rep.* 8, 16679. doi:10.1038/s41598-018-35149-1

López-Cortés, A., Cabrera-Andrade, A., Vázquez-Naya, J. M., Pazos, A., González-Díaz, H., Paz-y-Miño, C., et al. (2020a). Prediction of breast cancer proteins involved in immunotherapy, metastasis, and RNA-binding using molecular descriptors and artificial neural networks. *Sci. Rep.* 10, 8515. doi:10.1038/s41598-020-65584-y

López-Cortés, A., Paz-y-Miño, C., Guerrero, S., Cabrera-Andrade, A., Barigye, S. J., Munteanu, C. R., et al. (2020b). OncoOmics approaches to reveal essential genes in breast cancer: a panoramic view from pathogenesis to precision medicine. *Sci. Rep.* 10, 5285. doi:10.1038/s41598-020-62279-2

López-Cortés, A., Paz-y-Miño, C., Guerrero, S., Jaramillo-Koupermann, G., León Cáceres, Á., Intriago-Baldeón, D. P., et al. (2020c). Pharmacogenomics, biomarker network, and allele frequencies in colorectal cancer. *Pharmacogenomics J.* 20, 136–158. doi:10.1038/s41397-019-0102-4

Lu, C., Xie, M., Wendl, M. C., Wang, J., McLellan, M. D., Leiserson, M. D. M., et al. (2015). Patterns and functional implications of rare germline variants across 12 cancer types. *Nat. Commun.* 16, 10086. doi:10.1038/ncomms10086

Ma, B. B., Hui, E. P., and Mok, T. S. (2010). Population-based differences in treatment outcome following anticancer drug therapies. *Lancet Oncol.* 11, 75–84. doi:10.1016/S1470-2045(09)70160-3

Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science*, 80, 1912–1934. doi:10.1126/science.1075762

Martincorena, I., Raine, K. M., Gerstung, M., Dawson, K. J., Haase, K., Van Loo, P., et al. (2017). Universal patterns of selection in cancer and somatic tissues. *Cell* 171, 1029–1041. doi:10.1016/j.cell.2017.09.042

Martínez-Jiménez, F., Muiños, F., Sentís, I., Deu-Pons, J., Reyes-Salazar, I., Arnedo-Pac, C., et al. (2020). A compendium of mutational cancer driver genes. *Nat. Rev. Cancer* 20, 555–572. doi:10.1038/s41568-020-0290-x

Morgan, T., Koreckij, T., and Corey, E. (2009). Targeted therapy for advanced prostate cancer: inhibition of the PI3K/Akt/mTOR pathway. *Curr. Cancer Drug Targets* 9, 237–249. doi:10.2174/156800909787580999

Mularoni, L., Sabarinathan, R., Deu-Pons, J., Gonzalez-Perez, A., and López-Bigas, N. (2016). OncodriveFML: a general framework to identify coding and non-coding regions with cancer driver mutations. *Genome Biol.* 17, 128. doi:10.1186/s13059-016-0994-0

Mundi, P. S., Sachdev, J., McCourt, C., and Kalinsky, K. (2016). AKT in cancer: new molecular insights and advances in drug development. *Br. J. Clin. Pharmacol.* 82, 943–956. doi:10.1111/bcp.13021

Nadiminty, N., Tummala, R., Liu, C., Yang, J., Lou, W., Evans, C. P., et al. (2013). NF-kB2/p52 induces resistance to enzalutamide in prostate cancer: role of androgen receptor and its variants. *Mol. Cancer Ther.* 12, 1629–1637. doi:10.1158/1535-7163.MCT-13-0027

Patel, J. N. (2015). Cancer pharmacogenomics: implications on ethnic diversity and drug response. *Pharmacogenet. Genomics* 25, 223–230. doi:10.1097/FPC.0000000000000134

Patel, S. J., Sanjana, N. E., Kishton, R. J., Eidizadeh, A., Vodnala, S. K., Cam, M., et al. (2017). Identification of essential genes for cancer immunotherapy. *Nature* 548, 537–542. doi:10.1038/nature23477

Porta-Pardo, E., Kamburov, A., Tamborero, D., Pons, T., Grases, D., Valencia, A., et al. (2017). Comparison of algorithms for the detection of cancer drivers at subgene resolution. *Nat. Methods* 14 (8), 782–788. doi:10.1038/nmeth.4364

Quinones, L., Lavanderos, M., Cayun, J., Garcia-Martin, E., Agundez, J., Caceres, D., et al. (2014). Perception of the usefulness of drug/gene pairs and barriers for pharmacogenomics in Latin America. *Curr. Drug Metab.* 15, 202–208. doi:10.2174/1389200215666140202220753

Relling, M. V., Klein, T. E., Gammal, R. S., Whirl-Carrillo, M., Hoffman, J. M., and Caudle, K. E. (2020). The clinical pharmacogenetics implementation consortium: 10 Years later. *Clin. Pharmacol. Ther.* 107, 171–175. doi:10.1002/cpt.1651

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., and Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* 47, D886–D894. doi:10.1093/nar/gky1016

Rheinbay, E., Nielsen, M. M., Abascal, F., and Wala, J. A. (2020). Analyses of non-coding somatic drivers in 2, 658 cancer whole genomes. *Nature* 578, 102–111. doi:10.1038/s41586-020-1965-x

Ross, C. J. D., Visscher, H., Sistonen, J., Brunham, L. R., Pussegoda, K., Loo, T. T., et al. (2010). The Canadian pharmacogenomics network for drug safety: a model for safety pharmacology. *Thyroid* 20, 681–687. doi:10.1089/thy.2010.1642

Rubio-Perez, C., Tamborero, D., Schroeder, M. P., Antolín, A. A., Deu-Pons, J., Perez-Llamas, C., et al. (2015). In silico prescription of anticancer drugs to cohorts of 28 tumor types reveals targeting opportunities. *Cancer Cell* 27, 382–396. doi:10.1016/j.ccell.2015.02.007

Saito, Y., Stamp, L. K., Caudle, K. E., Hershfield, M. S., McDonagh, E. M., Callaghan, J. T., et al. (2016). CPIC: clinical pharmacogenetics implementation consortium of the pharmacogenomics research network. *Clin. Pharmacol. Ther.* 99, 36–37. doi:10.1002/cpt.161

Schmid, G. L., Kässner, F., Uhlig, H. H., Körner, A., Kratzsch, J., Händel, N., et al. (2014). Sirolimus treatment of severe PTEN hamartoma tumor syndrome: case report and *in vitro* studies. *Pediatr. Res.* 75 (4), 527–534. doi:10.1038/pr.2013.246

Sjöblom, T., Jones, S., Wood, L. D., Parsons, D. W., Lin, J., Barber, T. D., et al. (2006). The consensus coding sequences of human breast and colorectal cancers. *Science* 314, 268–274. doi:10.1126/science.1133427

Sondka, Z., Bamford, S., Cole, C. G., Ward, S. A., Dunham, I., and Forbes, S. A. (2018). The COSMIC cancer gene census: describing genetic dysfunction across all human cancers. *Nat. Rev. Cancer* 18 (11), 696–705. doi:10.1038/s41568-018-0060-1

Swen, J. J., Nijenhuis, M., De Boer, A., Grandia, L., Maitland-Van Der Zee, A. H., Mulder, H., et al. (2011). Pharmacogenetics: from bench to byte an update of guidelines. *Clin. Pharmacol. Ther.* 89, 662–673. doi:10.1038/clpt.2011.34

Tamborero, D., Gonzalez-Perez, A., Perez-llamas, C., Deu-Pons, J., Kandoth, C., Reimand, J., et al. (2013). Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Sci. Rep.* 3, 2650. doi:10.1038/srep02650

Tamborero, D., Rubio-Perez, C., Deu-Pons, J., Schroeder, M. P., Vivancos, A., Rovira, A., et al. (2018). Cancer genome interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med.* 10 (1), 25. doi:10.1186/s13073-018-0531-8

The Cancer Genome Atlas Research Network (2013). The cancer genome atlas pan-cancer analysis project. *Nat. Genet.* 45, 1113–1118. doi:10.1038/ng.2764

Tokheim, C., Bhattacharya, R., Niknafs, N., Gygax, D. M., Kim, R., Ryan, M., et al. (2016). Exome-scale discovery of hotspot mutation regions in human cancer using 3D protein structure. *Cancer Res.* 76 (13), 3719–3731. doi:10.1158/0008-5472.CAN-15-3190

Valencia, A., and Hidalgo, M. (2012). Getting personalized cancer genome analysis into the clinic: the challenges in bioinformatics. *Genome Med.* 4 (7), 61. doi:10.1186/gm362

Vasan, N., Baselga, J., and Hyman, D. M. (2019). A view on drug resistance in cancer. *Nature* 575, 299–309. doi:10.1038/s41586-019-1730-1

Wang, Q., Pierce-Hoffman, E., Cummings, B. B., Alföldi, J., Francioli, L. C., Gauthier, L. D., et al. (2020). Landscape of multi-nucleotide variants in 125,748 human exomes and 15,708 genomes. *Nat. Commun.* 11 (1), 2539. doi:10.1038/s41467-019-12438-5

Weghorn, D., and Sunyaev, S. (2017). Bayesian inference of negative and positive selection in human cancers. *Nat. Genet.* 49 (12), 1785–1788. doi:10.1038/ng.3987

Whirl-Carrillo, M., McDonagh, E. M., Hebert, J. M., Gong, L., Sangkuhl, K., Thorn, C. F., et al. (2012). Pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* 92 (4), 414–417. doi:10.1038/clpt.2012.96

Xia, W., Mullin, R. J., Keith, B. R., Liu, L. H., Ma, H., Rusnak, D. W., et al. (2002). Anti-tumor activity of GW572016: a dual tyrosine kinase inhibitor blocks EGF activation of EGFR/erbB2 and downstream Erk1/2 and AKT pathways. *Oncogene* 21 (41), 6255–6263. doi:10.1038/sj.onc.1205794

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

BRIEF RESEARCH REPORT
published: 15 April 2021
doi: 10.3389/fphar.2021.657985

Pharmacogenetics in Psychiatry: Perceived Value and Opinions in a Chilean Sample of Practitioners

Juan Undurraga[1,2], *Ignacio Bórquez-Infante*[3], *Nicolás A. Crossley*[4,5,6], *Miguel L. Prieto*[7,8,9,10] *and Gabriela M. Repetto*[11,12]*

[1]Department of Neurology and Psychiatry, Clínica Alemana Universidad Del Desarrollo, Santiago, Chile, [2]Early Intervention Program, Instituto Psiquiátrico Dr J. Horwitz Barak, Santiago, Chile, [3]Institute of Sociology, Pontificia Universidad Católica de Chile, Santiago, Chile, [4]Department of Psychiatry, Pontificia Universidad Católica de Chile, Santiago, Chile, [5]Biomedical Imaging Center, Pontificia Universidad Católica de Chile, Santiago, Chile, [6]Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom, [7]Department of Psychiatry, Faculty of Medicine, Universidad de Los Andes, Santiago, Chile, [8]Mental Health Service, Clínica Universidad de Los Andes, Santiago, Chile, [9]Center for Biomedical Research and Innovation, Universidad de Los Andes, Santiago, Chile, [10]Department of Psychiatry and Psychology, Mayo Clinic College of Medicine, Rochester, MN, United States, [11]Center for Genetics and Genomics, Instituto de Ciencias e Innovación en Medicina, Facultad de Medicina, Clínica Alemana Universidad Del Desarrollo, Santiago, Chile, [12]Department of Pediatrics, Clínica Alemana, Santiago, Chile

OPEN ACCESS

Edited by:
Patricia Esperon, Universidad de la República, Uruguay

Reviewed by:
Chad A. Bousman, University of Calgary, Canada
Dante Caceres, University of Chile, Chile

***Correspondence:**
Gabriela M. Repetto
grepetto@udd.cl

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *24 January 2021*
Accepted: *08 March 2021*
Published: *15 April 2021*

Citation:
Undurraga J, Bórquez-Infante I, Crossley NA, Prieto ML and Repetto GM (2021) Pharmacogenetics in Psychiatry: Perceived Value and Opinions in a Chilean Sample of Practitioners. Front. Pharmacol. 12:657985. doi: 10.3389/fphar.2021.657985

Use of pharmacogenetics (PGx) testing to guide clinical decisions is growing in developed countries. Published guidelines for gene–drug pair analysis are available for prescriptions in psychiatry, but information on their utilization, barriers, and health outcomes in Latin America is limited. As a result, this work aimed at exploring current use, opinions, and perceived obstacles on PGx testing among psychiatrists in Chile, via an online, anonymous survey. Among 123 respondents (5.9% of registered psychiatrists in the country), 16.3% reported ever requesting a PGx test. The vast majority (95%) of tests were ordered by clinicians practicing in the Metropolitan Region of Santiago. Having more than 20 years in practice was positively associated with prior use of PGx (p 0.02, OR 3.74 (1.19–11.80)), while working in the public health system was negatively associated (OR 0.30 (0.10–0.83)). Perceived barriers to local implementation included insufficient evidence of clinical utility, limited clinicians' knowledge on PGx and on test availability, and health systems' issues, such as costs and reimbursement. Despite the recognition of these barriers, 80% of respondents asserted that it is likely that they will incorporate PGx tests in their practice in the next five years. Given these results, we propose next steps to facilitate implementation such as further research in health outcomes and clinical utility of known and novel clinically actionable variants, growth in local sequencing capabilities, education of clinicians, incorporation of clinical decision support tools, and economic evaluations, all in local context.

Keywords: pharmacogenetics, psychiatry, implementation, Chile, Latin America

INTRODUCTION

Use of pharmacogenetics (PGx) tests to guide clinical decisions and care is appealing. From a broad perspective, this approach is intended to predict drug efficacy and adverse effects with the goal of improving clinical outcomes (Roden et al., 2019).

Genotype-guided prescribing or PGx testing considers interindividual genetic or genomic variation to inform about the likelihood of response to a specific treatment and/or of presenting adverse drug effects. It identifies genetic or genomic variants influencing drug effects, commonly via the identification of common and rare variations in genes related to pharmacokinetics or pharmacodynamics (reviewed in (Relling and Evans, 2015)). Allelic variation in genes encoding cytochrome P450 (CYP) monooxygenase system enzymes *CYP2D6* and *CYP2C19* plays an important role in commonly used psychiatric drugs (Zanger and Schwab, 2013). Also, variants in *HLA* alleles (*HLA-B**15:02 and *HLA-A**31:01), more commonly occurring in people of Asian ancestry, increase the risk of severe adverse reactions such as Stevens–Johnson or toxic epidermal necrolysis syndromes when using carbamazepine or oxcarbazepine (Phillips et al., 2018).

Deciding to use PGx in clinical care depends on many factors, including analytic validity, clinical validity, and clinical utility (Relling and Evans, 2015). The latter refers to whether using the test leads to improved health outcomes. In psychiatry, there is controversy regarding the clinical validity of pharmacogenetics testing for most gene-psychotropic drug pairs. Comprehensive resources such as the Pharmacogenomic Knowledge Base (PharmGKB) (PharmGKB, 2000), the Clinical Pharmacogenetics Implementation Consortium (CPIC) (CPIC, 2021), and the Dutch Pharmacogenetics Working Group (Swen et al., 2011; DPWG, 2019) aid in the implementation of PGx in clinical practice. Today, about 15% of United States FDA-approved medications contain PGx information in their label, with 38 drugs in psychiatry (Superintendencia De Salud, 2020). They include commonly used antidepressants (selective serotonin reuptake inhibitors, tricyclic antidepressants, venlafaxine, and vortioxetine), antipsychotics (aripiprazole, brexpiprazole, clozapine, haloperidol, iloperidone, risperidone, and zuclopenthixol), and mood stabilizers (carbamazepine and oxcarbazepine), among others (ISPG, 2020).

Some authors consider the available evidence insufficient to support clinical use. PGx testing is not considered, in general, as a standard of care in the field of mental health, since psychiatric disorders, such as major depressive disorder, are "determined by a large number of genes, and, except in rare cases, no single gene or limited gene set, even those for drug metabolism and drug targets, determines more than a few percent of the risk of illness or course of treatment," with environmental factors and comorbid states and medications are important determinants (Zubenko et al., 2018). Acknowledging this, the International Society for Psychiatric Genetics (ISPG) recommended in 2019 that PGx "should be viewed as a decision-support tool to assist in thoughtful implementation of good clinical care," and that evidence was still inconclusive to support widespread use (ISPG, 2020), but anticipated changes in the field. A more recent (2021) consensus by experts convened by the ISPG considers that current evidence does support the use of PGx testing for cytochrome P450 genes CYP2D6 and CYP2C19 to inform selection and dosing of commonly used antidepressant and antipsychotic medication, as well as testing for HLA-A and HLA-B when using carbamazepine, HLA-B for oxcarbazepine, PYP2C9, and HLA-B for phenytoin, and when a mitochondrial disorder or a urea cycle disorder is suspected, POLG, OTC, and CSP1 for valproate use (Bousman et al., 2021). Consequently, the authors predict that PGx testing will become a relevant tool in psychiatry.

In addition to information on validity and utility, other factors also impact implementation, such as physician awareness, availability and accessibility of testing, associated costs and reimbursement, and cost-effectiveness analyses in local contexts among others (Relling and Evans, 2015; Luzum et al., 2017). Some studies have assessed physicians' attitudes toward PGx. A recent survey on primary care and mental health providers enrolled in the Veterans Affairs' PRecision Medicine In MEntal Health Care (PRIME Care) study (Hull et al., 2019), revealed that less than 15% of 342 participants had ordered a PGx test to guide psychotropic medication in the previous year. Earlier also performed in the United States showed similar results: 12.9% of more than 10,000 responders of a nationwide survey of physicians in different specialties (Stanek et al., 2012) and 20% of 597 United States internists and family medicine practitioners (Haga et al., 2012), although these studies did not describe the specific area of ordered tests. Other studies show that a substantial proportion of psychiatrists consider that PGx testing will become a useful aid in drug prescribing (Thompson et al., 2015; Walden et al., 2015).

Most work on discovery, applications, and implementation of PGx has been carried out in developed countries, and while the global information is based on larger samples than in Latin America, there is limited information on use and views among practitioners in this region. For this reason, we conducted a survey on opinions about current practices, perceived value and barriers to clinical use of PGx testing, in a sample of clinical psychiatrists from Chile.

METHODS

Sample and Survey

We conducted an anonymous, cross-sectional survey via e-mail among clinical psychiatrists working in Chile. Contact information was obtained from directories of medical societies and professional associations, and smaller databases from the authors. The survey was approved by Research Ethics Committee at Facultad de Medicina Clinica Alemana Universidad del Desarrollo, and participants gave informed consent.

The survey was sent in November and December 2020, with an explanatory cover letter, an informed consent form, and a 10-question survey exploring opinions and barriers toward use of PGx in their clinical practice and future intention to use. Prior to submission, the questions were reviewed by a psychiatrist and a nonpsychiatrist physician, for comprehension and readability.

TABLE 1 | Sample characteristics and associations with prior use of PGx tests.

	Total		Ever used pharmacogenetic tests				p-value
			No N = 103 (83.7%)		Yes N = 20 (16.3%)		
	N	%	N	%	N	%	
Metropolitan region**	92	74.8	79	70.9	19	95.0	0.02
Practice at (% yes)[a]							
Public system**	67	54.5	61	59.2	6	30	0.02
University or research center	27	22	21	20.4	6	30	0.34
Private medical center*	90	73.2	72	69.9	18	90	0.06
Nonprofit foundation	5	4.1	4	3.9	1	5	0.81
In training	3	2.4	3	2.9	0	0	0.44
Other	16	13	13	12.6	3	15	0.77
Clinical practice (% yes)[a]							
Child psychiatry	19	15.4	15	14.6	4	20	0.53
Adolescent psychiatry	31	25.2	25	24.3	6	30	0.59
Adult psychiatry	104	84.6	87	84.5	17	85	0.95
Years in practice**							
0 to 10	56	45.5	51	49.5	5	25	
11 to 20	26	21.1	22	21.4	4	20	
21 to 30	22	17.9	19	18.4	3	15	0.02
31 to 40	16	13	9	8.7	7	35	
41 or more	3	2.4	2	1.9	1	5	
Participated in research in past 5 years							
Yes	75	61	62	60.2	13	65	0.68

*Note: N = 123. ***p < 0.01, **p < 0.05, *p < 0.1.*

[a]*More than one answer permitted.*

The survey was sent in Spanish. The translated questions can be found in **Supplementary Table S1**.

Statistical Analysis

A descriptive analysis was performed to report the participant characteristics and their responses, using frequencies and percentages. Association analysis was made using the chi-square test for categorical variables. The included variables predicting usage of PGx were 1) practice in the Metropolitan Region of Santiago; 2) practice in the public system; 3) practice at a university or research center; 4) practice at private medical center; 5) practice in adult psychiatry; and 6) years in practice. All analyses were conducted using Stata v16 (StataCorp. 2019, Stata Statistical Software: release 16. College Station, TX: StataCorp LLC).

RESULTS

At the time of the survey submission, there were 2093 registered psychiatrists in Chile, 1,022 of them (49%) in the Metropolitan Region of Santiago (FDA, 2021). The survey was sent to 546 psychiatrists registered in medical societies and professional associations. Responses were received from 123 (5.9% of registered psychiatrists and 22.5% response rate). Answers for some questions were not mutually exclusive; for example, a psychiatrist could have more than one practice setting or see patients of different age groups. Demographic characteristics are described in **Table 1**. Briefly, ninety-two participants (74.8%) worked in Santiago and the remainder in nine of the other fifteen regions of the country. The majority worked in combinations of public, private, and academic health settings, and 84% were adult psychiatrists. Close to half of respondents, 45.5%, had less than ten years in practice, and 61% had participated in research in the previous five years.

Twenty participants (16.3%) had ever requested PGx tests for psychotropic medication prescribing, most of them (80%), more than once. The median number of times they had requested a test was 3, with a range of 1–10. Most of the requests (90%) were made by physicians who declared working in the private healthcare system. The reported reasons for ordering a PGx test were to predict drug response (85%) or side effects (40%), or to explain observed side effects (25%). No tests by these psychiatrists were ordered per patient's request or other reasons not specified in choices in the survey.

Factors significantly associated with prior use of PGx testing included working in the Metropolitan Region (p 0.02, OR 7.81, 95% CI 1.00–60.98) and having 21 or more years in practice (vs. 10 or less years) (p 0.02, OR 3.74, 95% CI 1.19–11.80). In fact, only one clinician who reported use of PGx testing did not work in the Metropolitan Region. Practicing in the public health system was negatively associated with PGx testing (p 0.02 OR 0.3 [95% CI, 0.10–0.83]).

The main reason reported for not having ordered PGx testing was lack of awareness of tests that could be relevant for one's own clinical practice (62.1%), followed by costs (44.7%) and by perception of lack of utility (23.3%).

Perceived barriers to use of PGx included lack of personal knowledge about pharmacogenetics (51.2%), high cost of available exams (48.8%), and lack of availability of pharmacogenetic tests in Chile (46.3%). We grouped surveyed barriers in three main areas: 1) available evidence, 2) provider-related barriers, and 3) system-related barriers.

TABLE 2 | Perceived barriers to pharmacogenetic tests use.

	Ever used pharmacogenetic tests				Total		p-value
	No N = 103 83.7%		Yes N = 20 16.3%				
	N	%	N	%	N	%	
What are the main barriers to the use of pharmacogenetic tests? (% yes)							
Provider-related barriers ***	91	88.4	13	65	104	84.6	0.00
Lack of personal knowledge about pharmacogenetics	56	54.4	7	35	63	51.2	0.11
Not knowing where they are available **	51	49.5	4	20	55	44.7	0.02
Not knowing how to request them **	38	36.9	2	10	40	32.5	0.02
Complexity to request tests	13	12.6	2	10	15	12.2	0.74
Complexity to interpret the results **	7	6.8	5	25	12	9.8	0.01
Healthcare system-related barriers	81	78.6	18	90	99	80.5	0.24
High cost of available tests **	46	44.7	14	70	60	48.8	0.04
Lack of availability of pharmacogenetic tests in Chile	49	47.6	8	40	57	46.3	0.53
Lack of coverage by health systems	41	39.8	11	55	52	42.3	0.2
Lack of evidence of clinical utility ***	26	25.2	11	55	37	30.1	0.00
Other	5	4.9	0	0	5	4.1	0.31
Will incorporate the use of pharmacogenetic tests in its clinical practice in the next 5 years							
Yes	73	80.2	14	77.8	87	79.8	0.81

*** p < *0.01*, ** p < *0.05*, * p < *0.1*.

Provider-related barriers were more common among those who have not used PGx (88.4 vs. 65.0% in past users, p < 0.01), whereas healthcare system barriers were mentioned more frequently by those that have used tests in the past (90.0 vs. 78.6% in not users), although this difference was not statistically significant (p = 0.24) (**Table 2**). Perception of lack of evidence of clinical utility was mentioned more frequently by prior users (55%) than nonusers (25.2%) (p < 0.01). Despite recognition of these gaps, the majority (80%) considered likely that they will use PGx tests in their practice in the next 5 years, independent of having used them or not in the past (p 0.81).

DISCUSSION

This work represents the first survey of clinicians in Chile regarding the use and perceived barriers of PGx testing in psychiatric practice. Most respondents worked as adult psychiatrists in the Metropolitan Region of Santiago. However, the full sample included a broad spectrum of geographic locations, practice settings, and years of clinical experience. Prior use of PGx testing was reported by 16%, similar in proportion to studies done in the United States (Haga et al., 2012; Stanek et al., 2012; Hull et al., 2019) but each clinician in our survey had ordered tests 10 times or less, evidencing that, overall, its use was quite limited.

Factors associated with utilization of PGx testing included practice in the Metropolitan Region of Santiago and in private healthcare settings. These results could be related to the fact that tests are not currently available in certified clinical laboratories in Chile; therefore, samples need to be sent to laboratories abroad. This involves logistical challenges to order and to arrange for shipment and payments, processes that must be organized by the physician since these centers do not have local offices or contracts. In addition, testing performed out of the country is not covered by private or public insurance and it poses a significant out-of-pocket expense for patients (Villalobos Dintrans, 2018). Also, given the added shipment days, turnaround times for results are long, limiting the ability to guide prescription decisions, especially in acute care settings. Finally, and although not explored in this study, the lack of standards for PGx panels results in differences in the available ones; this heterogeneity may also add to challenges in ordering decisions (Fan and Bousman, 2020; Bousman et al., 2021).

Being in mid to late career stages (more than 20 years in practice) was positively associated with prior use, which suggests that knowledge in the field may have been acquired through continuing education or exposure to international research. PGx is not formally included the curriculum in medical schools or residency training programs in Chile. The reasons reported to use tests were mostly reactive to specific prescribing situations, that is, prediction of effects or explanation of side effects of particular drugs. No mention was made of preemptive testing; however, this was not explicitly explored in this study.

Among barriers to PGx testing was the perception of insufficient evidence of clinical utility. This was mentioned more frequently by prior users than nonusers, who may be more aware of the current limitations. As stated in the introduction, the clinical value of PGx testing in psychiatry is still a matter of debate and an evolving issue. The recently published review and updated consensus on PGx testing in psychiatry by a group of international experts assembled by the ISPG might help change this perception and advance in expanding the use of PGx testing in clinical settings for specific validates genes and gene-pair drugs (Bousman et al., 2021).

Most participants predicted that they will include PGx testing in their clinical practice in the next five years. This finding mirrors

opinions of psychiatrists in the United States and Canada (Thompson et al., 2015; Walden et al., 2015) where more than 80% of respondents in these studies consider that PGx will become a common standard in practice. Similar favorable opinions toward PGx testing were reported in a study of physicians in Jordan (Muflih, 2017). Nevertheless, the respondents of our study also recognized the presence of barriers to clinical use. In addition to insufficient evidence for clinical utility, other perceived difficulties were related to clinicians'personal knowledge about PGx, the logistics involved in requesting tests, and health system barriers, such as cost and reimbursement. These are also similar to the barriers identified by the largest published nationwide survey in the United States (Stanek et al., 2012). Another important issue is that, although CPIC and FDA guidelines are available for specific drug pairs, knowledge of their application and results in terms of health outcomes in Latin-American populations is practically nonexistent. In addition, there are differences in allelic composition and frequencies in several relevant genes, compared to European and other populations (Naranjo et al., 2018; Rodrigues-Soares et al., 2020), which may have an impact on clinical use in the region.

To address these limitations, several next steps are needed. Although most of the described challenges are similar to those described by practitioners in North America, there is a need to generate local information. Research on validation of known variant drug pairs in Latin-American populations and discovery of potential novel actionable associations are crucial; since due to the admixed genetic nature of local populations, findings from participants with other ancestries cannot be immediately transferred to the region. For instance, the RIBEF-CEIBA Network Consortium published a study analyzing CYP2D6, CYP2C9, and CYP2C19 genetic polymorphisms on 6,060 healthy individuals from Ibero-America, classified according to their self-reported ancestry, and found that Amerindians, that contribute to the admixed populations in the region, had significant differences in allelic frequencies with other ethnic groups, and also between native populations from the north and the south (Naranjo et al., 2018). These differences show the need for further local research in clinical applications and outcomes.

Once validity and utility are established in local context and training of clinicians is necessary. This requires the incorporation of PGx content in curricula in medical schools, graduate, and continuing medical education since this field is rapidly changing. Ease of access to pertinent tests is another relevant factor, and efforts are needed to increase local sequencing and interpretation capacities for accurate and rapid turnaround genotyping. Development and inclusion of clinical decision support tools that also consider nongenetic information, and their integration into health records are also facilitators, particularly, since as mentioned above, prediction of beneficial and side effects is of multifactorial nature. Economic analysis is also needed to inform coverage and reimbursement decisions for health services. Differences in allelic frequencies across global populations can impact on cost-effectiveness predictions in different countries (Zhou et al., 2021).

Limitations of this study include that it is a convenience sample, with a relatively small sample size and low response rate, representing only 6% of psychiatrists working in Chile. Nevertheless, the number of respondents were similar to previous similar surveys in more populated countries such the United States and Canada (Thompson et al., 2015; Walden et al., 2015), and to a study of physicians' response rate to web-based surveys, in the range of 27–46%, with the lower figure for psychiatrists (Cunningham et al., 2015). Most psychiatrists were affiliated both to private and public sectors and this study did not explore in which specific setting the tests were requested. It is possible that responses were mostly from physicians that may be more aware of PGx and therefore more inclined to participate. These results limit generalizability of the conclusions of the study. In addition, the survey did not explore whether the results of PGx testing had modified therapeutic decisions and/or patient outcomes in those cases in which it had been used.

In summary, this exploratory study shows that the current uptake of PGx testing by psychiatrists in Chile is low, but there is interest in the area and in potential incorporation in clinical practice. To bridge the gap between current and future appropriate use, local research and implementation strategies are needed. Understanding the scope and problems of PGx implementation can also benefit from collaboration throughout Latin America since the similar challenges are shared by other countries in the region (Abou Diwan et al., 2019).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité Etico-Científico, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

JU, NC, MP, and GR conceived the study and survey, collected the data, analyzed and interpreted the results, and wrote and edited the manuscript for scientific content. IB-I performed the statistical analysis and contributed to interpretation and writing and edition of the manuscript. All authors approved the submitted version.

FUNDING

Supported in part by ANID-PIA-ACT192064 (JU, NC, GMR), FONDECYT 1180358 (JU), 1200601 (NAC), 1171014 (GMR), 1181365 (MLP), FONDEF ID19I10116 (MLP) and Clínica

Alemana de Santiago ID 863 (JU). Funding for researchers' time. and publication fees The agencies had no role in study design, analysis, interpretation of results, or manuscript writing.

ACKNOWLEDGMENTS

The authors thank George Patrinos, PhD and Paula B. Repetto, PhD for critical review of the manuscript; Anne Bliss, PhD, for English language editing, and the participants for their time in responding the survey questions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.657985/full#supplementary-material.

REFERENCES

Abou Diwan, E., Zeitoun, R. I., Abou Haidar, L., Cascorbi, I., and Khoueiry Zgheib, N. (2019). Implementation and obstacles of pharmacogenetics in clinical practice: an international survey. *Br. J. Clin. Pharmacol.* 85, 2076–2088. doi:10.1111/bcp.13999

Bousman, C. A., Bengesser, S. A., Aitchison, K. J., Amare, A. T., Aschauer, H., Baune, B. T., et al. (2021). Review and consensus on pharmacogenomic testing in psychiatry. *Pharmacopsychiatry* 54, 5–17. doi:10.1055/a-1288-1061

CPIC (2021). Clinical pharmacogenetics implementation Consortium (CPIC). Available from: https://cpicpgx.org/ (Accessed Jan 17, 2021).

Cunningham, C. T., Quan, H., Hemmelgarn, B., Noseworthy, T., Beck, C. A., Dixon, E., et al. (2015). Exploring physician specialist response rates to web-based surveys. *BMC Med. Res. Methodol.* 15, 32. doi:10.1186/s12874-015-0016-z

DPWG (2019). Dutch pharmacogenetics working group [internet]. Available from: https://www.pharmgkb.org/page/dpwg (Accessed Feb 25, 2021).

Fan, M., and Bousman, C. A. (2020). Commercial pharmacogenetic tests in psychiatry: do they facilitate the implementation of pharmacogenetic dosing guidelines? *Pharmacopsychiatry* 53, 174–178. doi:10.1055/a-0863-4692

Haga, S., Burke, W., Ginsburg, G., Mills, R., and Agans, R. (2012). Primary care physicians' knowledge of and experience with pharmacogenetic testing. *Clin. Genet.* 82, 388–394. doi:10.1111/j.1399-0004.2012.01908.x

Hull, L. E., Lynch, K. G., and Oslin, D. W. (2019). VA primary care and mental health providers' comfort with genetic testing: survey results from the PRIME care study. *J. Gen. Intern. Med.* 34, 799–801. doi:10.1007/s11606-018-4776-0

ISPG (2020) International society of psychiatric genetics-genetic testing statement. Available from: https://ispg.net/genetic-testing-statement/ (Accessed Dec 29, 2020).

Luzum, J., Pakyz, R., Elsey, A., Haidar, C., Peterson, J., Whirl-Carrillo, M., et al. (2017). The Pharmacogenomics research Network translational pharmacogenetics program: outcomes and metrics of pharmacogenetic implementations across diverse healthcare systems. *Clin. Pharmacol. Ther.* 102, 502–510. doi:10.1002/cpt.630

Muflih, S. (2017). Measuring knowledge and attitudes regarding the use of pharmacogenetic testing among patients and prescribers: diffusion of innovation theory. PhD dissertation. Fort Lauderdale (FL): Nova Southeastern University.

Naranjo, M. E. G., Rodrigues-Soares, F., Peñas-Lledó, E. M., Tarazona-Santos, E., Fariñas, H., Rodeiro, I., et al. (2018). Interethnic variability in CYP2D6, CYP2C9, and CYP2C19 genes and predicted drug metabolism phenotypes among 6060 Ibero- and native Americans: RIBEF-CEIBA Consortium report on population Pharmacogenomics. *OMICS: A J. Integr. Biol.* 22, 575–588. doi:10.1089/omi.2018.0114

PharmGKB (2000) Pharmacogenomics knowledge Base (PharmGKB). Available from: https://www.pharmgkb.org/ (Accessed Jan 28, 2019).

Phillips, E. J., Sukasem, C., Whirl-Carrillo, M., Müller, D. J., Dunnenberger, H. M., Chantratita, W., et al. (2018). Clinical pharmacogenetics implementation Consortium guideline for *HLA* genotype and use of carbamazepine and oxcarbazepine: 2017 update. *Clin. Pharmacol. Ther.* 103, 574–581. doi:10.1002/cpt.1004

Superintendencia De Salud (2020) Registro Nacional de Prestadores Individuales de Salud. Available at: https://rnpi.superdesalud.gob.cl/ (Accessed Dec 31, 2020).

Relling, M. V., and Evans, W. E. (2015). Pharmacogenomics in the clinic. *Nature* 526, 343–350. doi:10.1038/nature15817

Roden, D. M., McLeod, H. L., Relling, M. V., Williams, M. S., Mensah, G. A., Peterson, J. F., et al. (2019). Pharmacogenomics. *The Lancet* 394, 521–532. doi:10.1016/s0140-6736(19)31276-0

Rodrigues-Soares, F., Peñas-Lledó, E. M., Tarazona-Santos, E., Sosa-Macías, M., Terán, E., López-López, M., et al. (2020). Genomic ancestry, CYP2D6, CYP2C9, and CYP2C19 among Latin Americans. *Clin. Pharmacol. Ther.* 107 (1), 257–268. doi:10.1002/cpt.1598

Stanek, E. J., Sanders, C. L., Taber, K. A. J., Khalid, M., Patel, A., Verbrugge, R. R., et al. (2012). Adoption of pharmacogenomic testing by US physicians: results of a nationwide survey. *Clin. Pharmacol. Ther.* 91, 450–458. doi:10.1038/clpt.2011.306

Swen, J. J., Nijenhuis, M., de Boer, A., Grandia, L., Maitland-van der Zee, A. H., Mulder, H., et al. (2011). Pharmacogenetics: from bench to byte-an update of guidelines. *Clin. Pharmacol. Ther.* 89, 662–673. doi:10.1038/clpt.2011.34

FDA (2021). Table of Pharmacogenomic Biomarkers in Drug Labeling. Available at: https://www.fda.gov/drugs/science-and-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling (Accessed Jan 11, 2021).

Thompson, C., Steven P. Hamilton, P. H., and Hippman, C. (2015). Psychiatrist attitudes towards pharmacogenetic testing, direct-to-consumer genetic testing, and integrating genetic counseling into psychiatric patient care. *Psychiatry Res.* 226, 68–72. doi:10.1016/j.psychres.2014.11.044

Villalobos Dintrans, P. (2018). Out-of-pocket health expenditure differences in Chile: insurance performance or selection? *Health Policy* 122, 184–191. doi:10.1016/j.healthpol.2017.11.007

Walden, L. M., Brandl, E. J., Changasi, A., Sturgess, J. E., Soibel, A., Notario, J. F. D., et al. (2015). Physicians' opinions following pharmacogenetic testing for psychotropic medication. *Psychiatry Res.* 229, 913–918. doi:10.1016/j.psychres.2015.07.032

Zanger, U. M., and Schwab, M. (2013). Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol. Ther.* 138, 103–141. doi:10.1016/j.pharmthera.2012.12.007

Zhou, Y., Krebs, K., Milani, L., and Lauschke, V. M. (2021). Global frequencies of clinically important HLA alleles and their implications for the cost-effectiveness of preemptive pharmacogenetic testing. *Clin. Pharmacol. Ther.* 109, 160–174. doi:10.1002/cpt.1944

Zubenko, G. S., Sommer, B. R., and Cohen, B. M. (2018). On the marketing and use of pharmacogenetic tests for psychiatric treatment. *JAMA Psychiatry* 75, 769–770. doi:10.1001/jamapsychiatry.2018.0834

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 22 April 2021
doi: 10.3389/fphar.2021.653525

CYP3A5 and UGT1A9 Polymorphisms Influence Immunosuppressive Therapy in Pediatric Kidney Transplant Recipients

Paola Krall[1†], Dominique Yañez[2†], Angélica Rojo[1], Ángela Delucchi[1], Miguel Córdova[2], Jorge Morales[3], Pía Boza[2], Alonso de la Rivera[2], Natalie Espinoza[2], Natalia Armijo[4], Luis E. Castañeda[5], Mauricio J. Farfán[1,2] and Carolina Salas[2]**

[1]Departamento de Pediatría y Cirugía Infantil Oriente, Facultad de Medicina, Universidad de Chile, Santiago de Chile, Chile, [2]Laboratorio Clínico, Hospital Luis Calvo Mackenna, Santiago de Chile, Chile, [3]Servicio de Farmacia, Hospital Luis Calvo Mackenna, Santiago de Chile, Chile, [4]Unidad de Nefrología, Hospital Luis Calvo Mackenna, Santiago de Chile, Chile, [5]Programa de Genética Humana, Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, Santiago de Chile, Chile

OPEN ACCESS

Edited by:
Patricia Esperon,
Universidad de la República, Uruguay

Reviewed by:
Su-Jun Lee,
Inje University, South Korea
Chakradhara Rao Satyanarayana Uppugunduri,
Université de Genève, Switzerland

***Correspondence:**
Mauricio J. Farfán
mfarfan@uchile.cl
Carolina Salas
csalas@calvomackenna.cl

†These authors have contributed equally to this work.

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: 14 January 2021
Accepted: 23 March 2021
Published: 22 April 2021

Citation:
Krall P, Yañez D, Rojo A, Delucchi Á, Córdova M, Morales J, Boza P, de la Rivera A, Espinoza N, Armijo N, Castañeda LE, Farfán MJ and Salas C (2021) CYP3A5 and UGT1A9 Polymorphisms Influence Immunosuppressive Therapy in Pediatric Kidney Transplant Recipients. Front. Pharmacol. 12:653525. doi: 10.3389/fphar.2021.653525

Background: Tacrolimus (TAC) and mycophenolic acid (MPA) are the main immunosuppressive drugs used in pediatric kidney transplantation. Single nucleotide polymorphisms (SNPs) in metabolizing enzymes and transporters might influence plasma levels of these drugs. Herein, we sought to determine the influence of SNPs on *CYP3A5*, *MRP2* and *UGT1A9* genes in Chilean pediatric kidney recipients using TAC and MPA.

Patients and Methods: A prospective study was performed on 104 pediatric kidney recipients that used TAC and MPA for immunosuppression. The median age at the time of transplantation was 8.1 years [Q1–Q3 4.5–11.6 years] and the main clinical diagnosis was a structural anomaly. In a subgroup of patients, a complete steroid withdrawal was made at day 7. The *CYP3A5* polymorphism (ancestral allele *1; variant allele *3) was determined in the entire cohort, while *MRP2* -24G > A, *UGT1A9* -275T > A, and *UGT1A9* -2152C > T polymorphisms were determined in 53 patients. Genotypes were associated with trough drug concentrations (C_0), dose requirements normalized by weight (TAC-D mg/kg) or body surface (MPA-D mg/m2), trough levels normalized by dose requirements (C_0/D), and area under the curve in 12 h normalized by dose requirements ($AUC_{0–12h}$/D).

Results: The frequencies of the variant alleles *CYP3A5*3*, *MRP2-24A*, *UGT1A9-275A*, and *UGT1A9-2152T* were 76.9, 22.1, 6.6, and 2.9%, respectively. $AUC_{0–12h}$/TAC-D were 1.6-fold higher in *CYP3A5*3/*3* patients than in *CYP3A5*1* carriers (*CYP3A5*1/*3* and *CYP3A5*1/*1*). When analyzing patients with steroid withdrawal, *CYP3A5*3/*3* patients had 1.7-fold higher $AUC_{0–12h}$/TAC-D than the other genotypes. Patients carrying the *CYP3A5*3/*3* genotype had higher TAC-C_0, lower TAC-D and higher TAC-C_0/D, consistently in a 6-months follow-up. Creatinine clearance was stable during the follow-up, regardless of the genotype. No significant differences between *MRP2* and *UGT1A9* genotypes were observed in MPA-C_0, MPA-D or MPA-C_0/D. However, patients carrying the *UGT1A9-275A* allele had lower $AUC_{0–12h}$/MPA-D than those carrying the *UGT1A9-275T* ancestral allele.

Conclusions: These results support that *CYP3A5* and *UGT1A9* genotyping in pediatric recipients might be useful and advisable to guide TAC and MPA dosing and monitoring in children that undergo kidney transplantation.

Keywords: pediatric kidney transplantation, pharmacogenetics, tacrolimus, mycophenolic acid, pharmacokinetics

INTRODUCTION

Kidney transplantation is the renal replacement therapy of choice in the pediatric population, improving survival, growth and quality of life (McDonald and Craig, 2004). Immunosuppressive therapy to avoid rejection is the cornerstone of graft survival (Winterberg and Garro, 2019). Tacrolimus (TAC) and mycophenolic acid (MPA) are the main immunosuppressive drugs used in pediatric kidney transplantation in combination with steroids. However, an early tapering of steroids until complete withdrawal maintaining the TAC and MPA dosage in pediatric recipients with low immunological risk is associated with a better growth pattern (Delucchi et al., 2007).

TAC and MPA need continuous monitoring to verify that the plasma concentrations are within the therapeutic ranges, to maintain the balance between efficacy and toxicity and, consequently, to ensure graft and patient survival (Zwart et al., 2018). The drug bioavailability might be influenced by genetic variations, such as single nucleotide polymorphisms (SNPs) in metabolizing enzymes or transport proteins, causing increases or decreases in plasmatic drug concentrations (Golubovic et al., 2016).

TAC is a calcineurin inhibitor that is metabolized by the cytochrome P450 enzymes, mainly CYP3A4 and CYP3A5, involved in the first step of drug metabolism. CYP3A5 plays the most important role when a dose of this drug is prescribed, particularly in pediatric patients (Zheng et al., 2003; De Wildt et al., 2011). Several studies have shown that the CYP3A5 expression is correlated with reduced TAC plasma levels and a delay in achieving adequate plasma concentrations with the conventional exposure dose, which increases the risk of acute rejection incidence (Kuehl et al., 2001; Staatz, Taylor, and Tett 2001; Macphee et al., 2002; MacPhee et al., 2004; Min et al., 2010). By contrast, recipients treated with TAC who do not express this enzyme are more exposed to nephrotoxicity that might be explained by plasma TAC accumulation (Thervet et al., 2003).

The variability in *CYP3A5* enzyme expression is explained mainly by the *CYP3A5* gene that contains a SNP described as a nucleotide change located in intron 3 (rs776746; *CYP3A5* c. A6986G). This genetic variation results in two alleles: the ancestral allele called *CYP3A5*1* and the variant allele called *CYP3A5*3*. The presence of two *CYP3A5*3* alleles (*CYP3A5*3/*3*) results in a truncated protein as a consequence of alternative splicing causing the absence of enzyme activity, also known as a non-expressor genotype. On the other hand, the presence of at least a single *CYP3A5*1* allele (*CYP3A5*1/*3* or *CYP3A5*1/*1*) is related to a functional enzyme, and these genotypes are cataloged as expressors (Kuehl et al., 2001). Based on genetic studies, the frequency of the *CYP3A5*3* allele displays variability according to the geographic and ethnic origin, ranging from 32% in African individuals to 93% in European individuals. This variability reinforces the need to analyze the frequency of the *CYP3A5* genotype and the association with plasma TAC levels in different human geographic populations or ethnic groups (Xie et al., 2004).

MPA is an inhibitor of T-cell proliferation strongly influenced by the metabolizing diphosphate glucuronyl transferase (UGT) enzymes and the multidrug resistance-associated proteins (MRP) that act as drug transporters on the membrane of hepatocytes (Barraclough et al., 2010). Similar to *CYP3A5*, there is genetic variability and several reports have demonstrated that the specific variants *MRP2* -24C (rs717620), *UGT1A9* -275A (rs6714486) and *UGT1A9* -2152T (rs17868320) polymorphisms might have a role in the inter-individual variability of MPA plasma levels in pediatric organ recipients (Bernard and Guillemette, 2004; Naesens et al., 2006; Fukuda et al., 2012).

Considering that polymorphisms involved in the metabolism of TAC and MPA have a geographic and ethnic variability, in this study we aimed to determine the frequency of SNPs in *CYP3A5*, *MRP2*, and *UGT1A9* genes in Chilean pediatric kidney recipients. Additionally, we evaluated the influence of these polymorphisms on TAC and MPA plasma levels. Our data support that the *CYP3A5* and *UGT1A9* genotyping in pediatric recipients might be useful and advisable to guide TAC and MPA dosing and monitoring in children that undergo kidney transplantation.

PATIENTS AND METHODS

Patient Data

Pediatric patients that underwent kidney transplantation with a deceased or living-related donor between 2001 and 2019 at the referral transplant center Hospital Luis Calvo Mackenna (Santiago, Chile) were invited to participate in this descriptive, prospective, longitudinal study. In our center, 85–90% of the transplantations were performed with deceased donors each year. We studied 104 recipients who received TAC and MPA as immunosuppressive therapy combined with steroids. Patients were induced with basiliximab on day 0 during surgery before reperfusion and on day 4 post-transplantation. TAC (starting dose 0.15 mg/kg twice per day), MPA (starting dose 800 mg/m2 twice per day), and steroids were administered according to the institutional protocol. Steroid use was individualized according to immunological risk and graft function in each patient. Starting steroid therapy was methylprednisolone 2 mg/kg/day for 2 days and switched to prednisone at day 3. The patients undergoing steroid withdrawal were treated successively with 2, 1, 0.5, and 0.25 mg/kg/day, each dose for one day. The patients that continued steroid therapy switched to prednisone 2 mg/kg/day at day 3 and reduced gradually until reaching 0.12 mg/kg/day at

month 6, and this was kept constant from that month onwards. Therapeutic drug monitoring of TAC was applied regularly to all recipients and dose adjustment was performed when necessary to achieve target trough levels (10–15 ng/ml in the first 3 months and 5–7 ng/ml after 3 months post-transplantation). At the end of a 6-months follow-up of TAC therapy, associations of the *CYP3A5* SNP with drug concentrations (TAC-C_0), daily dose normalized by weight (TAC-D mg/kg) and drug concentrations normalized by dose (TAC- C_0/D) were analyzed. MPA levels were regularly monitored and associations of the *MRP2* and *UGT1A9* SNPs with drug concentrations (MPA-C_0), daily dose normalized by body surface (MPA-D mg/m2) and drug concentrations normalized by dose (MPA-C_0/D) were analyzed at one time point, once stable immunosuppression had been achieved. The analysis of the area under the curve in 12 h normalized by dose (AUC_{0-12h}/D) was conducted on 98 and 44 patients with stable immunosuppression based on TAC or MPA therapy, respectively. Relative frequencies of ABO and Rh blood groups demonstrated that the study population was representative of the Chilean population (data not shown).

The study was conducted according to the declaration of Helsinki and was approved by the Ethics Committees of the Universidad de Chile and Hospital Luis Calvo Mackenna. Patients with evidence of graft dysfunction during the first year post-transplantation, with change of immunosuppressive medication to azathioprine or cyclosporine A within the first 6 months after transplantation, or parents that denied participation were excluded.

DNA Extraction and SNP Genotyping

Nucleic acids were extracted from whole peripheral blood samples using QIAmp DNA Blood Mini Kit (QIAGEN) or MagNA Pure Compact Nucleic Acid Isolation Kit I (Roche) following the manufacturer's instructions. *CYP3A5*, *MRP2*, and *UGT1A9* were selected as genes with variants of interest for this study, because they have been proposed as the major factors to explain the variability related to plasma drug levels (TAC and MPA) and have shown similar results in different populations. The presence of the *CYP3A5* c. A6986G polymorphism (rs776746) was determined using the PCR-RFLP technique with specific primers (forward 5′-CAT GAC TTA GTA GAC AGA TGA-3 ′and reverse 5′-GGT CCA AAC AGG GAA GAA ATA-3′) and the restriction enzyme *SspI*, according to previously described protocols (Rong et al., 2010). The assessment of the polymorphisms *UGT1A9* 2152C > T (rs17868320), *UGT1A9* -275T > A (rs6714486) and *MRP2* -24C > T (rs717620) was performed with the TaqMan™ Drug Metabolism Genotyping Assay (ThermoFisher). All SNPs were processed on LightCycler® 480II and analyzed with the LightCycler® 480II software (Roche). Each assay included controls confirmed previously by Sanger sequencing for the three genotypes (homozygous reference allele, heterozygous, homozygous variant allele) if available and negative controls with an equal volume of nuclease-free pure water.

Therapeutic Drug Monitoring: TAC and MPA

To measure plasma drug concentration, four blood samples of 2 ml were taken in EDTA tubes, corresponding to C_0 (trough level before dose), C_1, C_2, and C_4 time points. Whole blood levels of TAC were measured using the Abbott Architect i1000 immunoassay (Abbott Laboratories). Plasma levels of MPA were measured using a liquid-liquid extraction and high-performance liquid chromatography equipment with a diode array detector (Agilent 1260, Agilent Technologies). The value of TAC AUC_{0-12h} was calculated in 98 patients using the abbreviated equation: $AUC_{0-12h} = 10 + 1.4*C_0 + 0.8*C_1 + 1.6*C_2 + 5.5*C_4$) (Wong et al., 2000). The value of MPA AUC_{0-12h} was calculated in 44 patients using the abbreviated equation: $AUC_{0-12h} = 8.217 + 3.163*C_0 + 0.994*C_1 + 1.334*C_2 + 4.183*C_4$ (Filler, 2004). The AUC_{0-12h} values were normalized by dose (TAC or MPA) for the following reasons: 1) AUC_{0-12h} was performed in each patient according to the clinical need to determine empirically drug exposure and was required at different time points, before and after 3 months of transplantation, that differ in AUC_{0-12h} and C_0 therapeutic targets, requiring higher doses in the first 3 months period; 2) the study cohort is a pediatric population with different etiologies/co-morbidities and their ages ranged from 1.5 to 15.3 years at the time of transplantation, causing differences in drug metabolism that confer, in consequence, variability in dose requirements to reach therapeutic targets.

Statistical Analysis

Pearson chi-square (χ^2) goodness-of-fit test for the Hardy Weinberg equilibrium was applied to assess deviation of allele and genotype frequencies. Normality and homoscedasticity assumptions were checked for all variables or transformed with log to achieve normal distribution. AUC_{0-12h}/TAC-D was analyzed by a one-way ANOVA to compare CYP3A5 genotypes. The analysis of AUC_{0-12h} without TAC dose normalization to compare *CYP3A5* genotypes are shown as supplemental material (**Supplementary Figures S1, S2**). TAC blood concentrations (TAC-C_0), TAC daily dose requirement (TAC-D), TAC blood concentrations normalized by daily dose requirement (TAC- C_0/D), and creatinine levels throughout the 6-months follow-up were analyzed by a repeated-measures ANOVA with genotype and time as main effects using the *ez* package (Lawrence, 2013). *Post hoc* tests were applied in the case of significant differences between groups by Tukey HSD test. MPA blood concentrations (MPA-C_0), MPA dose requirement (MPA-D), and MPA blood concentrations normalized by dose requirement (MPA- C_0/D) were analyzed by Wilcoxon-Mann Whitney test, grouping the heterozygous and homozygous carriers of the variant allele. The AUC_{0-12h}/MPA-D values between carriers and non-carriers of the variant allele were compared with an unpaired one-tailed Wilcoxon-Mann Whitney test. The analysis of AUC_{0-12h} without MPA dose normalization to compare *UGT1A9 -275* variant carriers are shown as supplemental material (**Supplementary Figure S3**). Outlier values (mean ± 3*SD) suggestive of errors in sampling procedure, technical measurements or data manipulation were excluded from the analysis. Box plots show median and interquartile range (IQR) and whiskers represent the 1.5*IQR. All statistical analyses were

TABLE 1 | Basic clinical characteristics and *CYP3A5* genotype of the 104 pediatric kidney recipients.

Characteristics	
Male/Female	52/52
Cause of ESRD	
Structural anomaly	70 (67.3%)
Glomerulopathy	14 (13.5%)
Monogenic cause (confirmed)	8 (7.7%)
Vascular cause	4 (3.8%)
Other	6 (5.8%)
Undetermined or unknown	2 (1.9%)
Age median at time of transplantation (years) [IQR]	8.1 [4.5–11.6]
Weight median at time of transplantation (kg) [IQR]	25.5 [15.9–35.2]
CYP3A5 genotype	
*CYP3A5 *1/*1*	9 (8.7%)
*CYP3A5 *3/*1*	30 (28.8%)
*CYP3A5 *3/*3*	65 (62.5%)

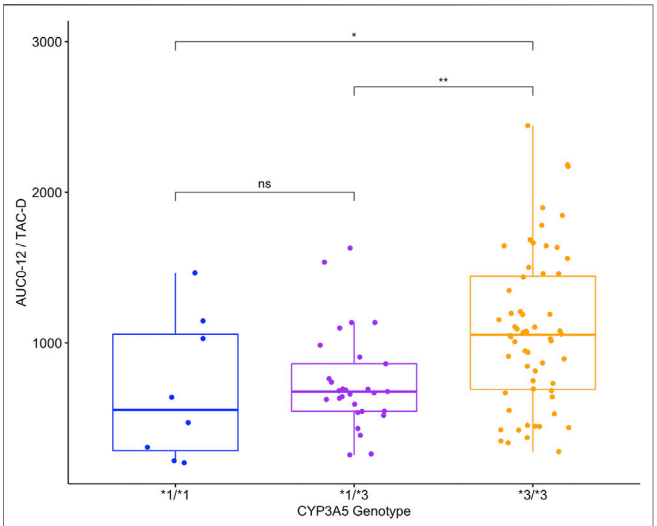


FIGURE 1 | Area under the curve in 12 h normalized by TAC dose requirements ($AUC_{0\text{-}12h}$/TAC-D ng*hr/ml/mg/kg) in 97 patients with different *CYP3A5* genotypes.

performed with R (Wilson and Norden, 2015) and plots were made using *ggpubr* (Kassambara, 2020a) and *rstatix* (Kassambara, 2020b) packages for R. P-values < 0.05 were considered to be statistically significant.

RESULTS

Characteristics of the Patients and Frequency of the *CYP3A5* Genotype

Altogether 104 pediatric kidney recipients from a single referral transplant center were included in this study. Clinical characteristics and *CYP3A5* genotype data are shown in **Table 1**. The median age at the time of transplantation was 8.1 years, and the leading cause of the end stage renal disease to require transplantation was a structural anomaly.

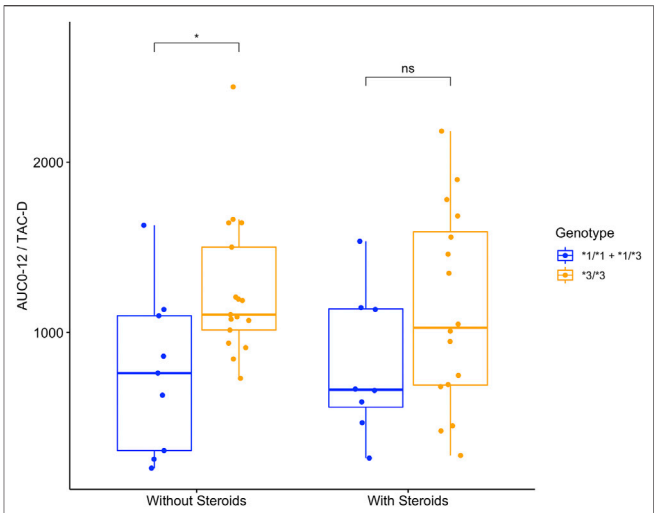


FIGURE 2 | Area under the curve in 12 h normalized by dose requirements ($AUC_{0\text{-}12h}$/TAC-D ng*hr/ml/mg/kg) in 52 patients with different *CYP3A5* genotypes with or without steroids from day 7 post-transplantation onwards.

The SNP rs776746 in *CYP3A5* was analyzed in the entire cohort and we found that 8.7% presented the *CYP3A5*1/*1* genotype, 28.8% the heterozygous *CYP3A5*1/*3* genotype and 62.5% the *CYP3A5*3/*3* genotype (**Table 1**). The frequency of the *CYP3A5*3* variant allele resulted in 76.9%. Genotype and allele frequencies were not significantly different than expected if the population was in Hardy-Weinberg equilibrium ($\chi^2 = 3.66$, $p = 0.06$). Compared to the Latino/Admixed American population reported in the gnomAd database v2.1.1 in December 2020 (https://gnomad.broadinstitute.org/), no statistical differences were found in our population in terms of genotype ($\chi^2 = 3.09$, $p = 0.21$) or allele ($\chi^2 = 0.49$, $p = 0.48$) frequencies of this SNP.

Associations of the CYP3A5 Genotypes with $AUC_{0–12h}$/D, Trough Levels (C_0), Dose Requirements (D) and C_0/D in TAC Therapy.

The analysis of $AUC_{0–12h}$/TAC-D in 97 patients revealed significant differences ($p < 0.01$) among the *CYP3A5* genotypes (**Figure 1**). The *CYP3A5*3/*3* patients showed the highest values (1058 [IQR 688.6–1460] ng*hr/ml/mg/kg), which was 1.56-fold higher than *CYP3A5*1/*3* patients (676.9 [IQR 546.0–883.6] ng*hr/ml/mg/kg; $p < 0.01$) and 1.91-fold higher than *CYP3A5*1/*1* patients (554.7 [IQR 239.7–1117]ng/h*ml/mg/kg; $p < 0.05$), respectively. $AUC_{0–12h}$/TAC-D values did not differ between *CYP3A5*1/*3* and *CYP3A5*1/*1* patients ($p = 0.79$).

Regarding the steroid therapy, the analysis showed some differences between *CYP3A5* genotypes (**Figure 2**). When analyzing patients with early steroid withdrawal, the carriers of the *CYP3A5*3/*3* genotype showed $AUC_{0–12h}$/TAC-D values (1099 [IQR 930.7–1537] ng*hr/ml/mg/kg), 1.58-fold significantly higher than the carriers of the *CYP3A5*1* ancestral allele (*CYP3A5*1/*3* $n = 7$ and *CYP*1/*1*; $n = 2$)

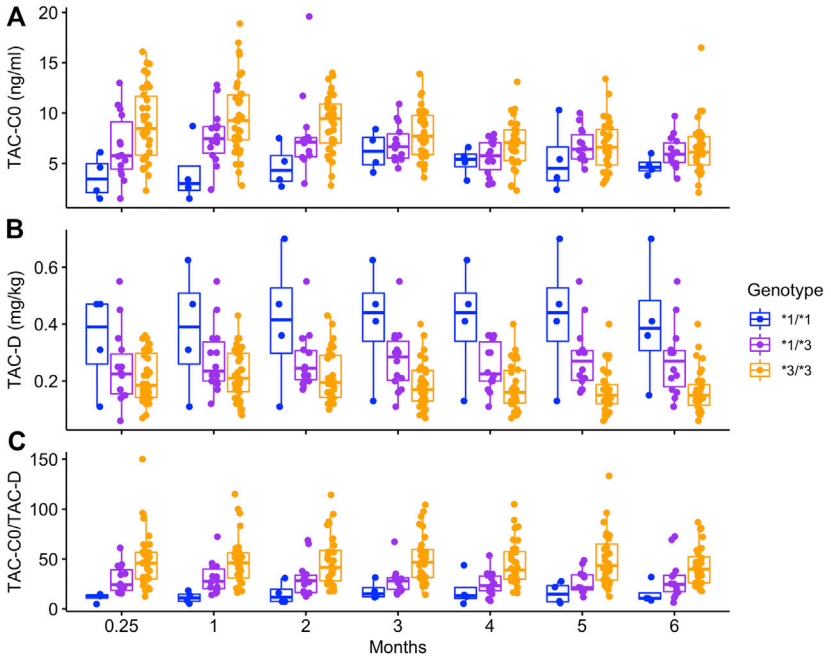


FIGURE 3 | Trough levels (TAC-C_0 ng/ml), dose requirements (TAC-D mg/kg) and trough levels normalized by dose requirements (TAC-C_0/D ng*kg/ml/mg) in 52 patients with different *CYP3A5* genotypes in TAC therapy that underwent a 6-month follow-up.

(697 [IQR 404.5–1108] ng*hr/ml/mg/kg) ($p < 0.05$). No statistical differences were observed between *CYP3A5* genotypes that continued steroid therapy ($p = 0.34$).

In the analysis performed in 52 patients with 6 months of follow-up, the TAC blood concentrations (TAC-C_0), TAC daily dose requirement (TAC-D) and TAC blood concentrations normalized by dose requirement (TAC- C_0/D) differed significantly between the *CYP3A5* genotypes (TAC-C_0 $p < 0.001$, TAC-D $p < 0.01$, TAC- C_0/D $p < 0.001$) (**Figures 3A–C**). Specifically, patients with the *CYP3A5*3/*3* genotype showed higher TAC-C_0 values than patients with the *CYP3A5*1/*3* and *CYP3A5*1/*1* genotypes ($p < 0.01$ and $p < 0.001$, respectively). Additionally, patients with the *CYP3A5*1/*3* genotype also exhibited higher TAC-C_0 values than patients with the *CYP3A5*1/*1* genotype ($p < 0.001$). Monitoring within the first three months, showed that 5–15% of the patients carrying the *CYP3A5*1* allele had achieved target levels, lower than the 22–43% of the patients with the *CYP3A5*3/*3* genotype. At the sixth month, 44.6% of the patients had achieved TAC-C_0 in the range between 5 and 7 ng/mL, regardless of their *CYP3A5* genotype (**Figure 3A**). According to the analysis of TAC-D, patients with *CYP3A5*1/*1* genotype showed higher TAC-D values than patients with the *CYP3A5*1/*3* genotype and patients with the *CYP3A5*3/*3* genotype ($p < 0.01$ and $p < 0.001$, respectively). Additionally, patients with the *CYP3A5*1/*3* genotype also displayed higher TAC-D values than patients with the *CYP3A5*3/*3* genotype ($p < 0.001$). Interestingly, significant differences between genotypes were not detected until the third month of follow-up, which is corroborated by an interaction between genotype and time ($p < 0.01$) (**Figure 3B**). Finally, patients with the *CYP3A5*3/*3* genotype showed higher TAC-C_0/D values than patients with the *CYP3A5*1/*3* and *CYP3A5*1/*1* genotypes ($p < 0.001$ in both cases). Similar to the observations in the TAC- C_0 analysis, patients with the *CYP3A5*1/*3* genotype also exhibited higher TAC-C_0/D values than patients with the *CYP3A5*1/*1* genotype ($p < 0.001$) (**Figure 3C**). No significant temporal variation (TAC-C_0 $p = 0.32$; TAC-D $p = 0.43$; and TAC- C_0/D $p = 0.44$) or genotype-time interaction were observed (TAC-C_0 $p = 0.18$; and TAC- C_0/D $p = 0.56$), except for the genotype-time interaction for TAC-D (see above).

The change in dose requirements comparing the initial prescribed TAC dose with the required TAC dose at the end of the 6-months follow-up was different between *CYP3A5* genotypes, experiencing a 209% increase in carriers of the *CYP3A5*1* ancestral allele (*CYP3A5*1/*3* and *CYP*1/*1* patients; 0.314 ± 0.151 mg/kg, $p < 0.01$). By contrast, the TAC dose requirement of *CYP3A5*3/*3* patients remained unaltered at the end of the follow-up in comparison to the initial prescribed TAC dose for these patients (0.164 ± 0.073 mg/kg, $p = 0.99$). Creatinine clearance was stable during the 6 months, regardless of the genotype ($p = 0.76$) (data not shown).

Association of *MRP2* and *UGT1A9* Alleles with AUC_{0-12h}/D, Trough Levels (C_0), Dose Requirements (D) and C_0/D in MPA Therapy.

When analyzing the SNPs in *MRP2* and *UGT1A9* that might influence MPA pharmacokinetics in a subgroup of 53 patients,

TABLE 2 | Basic clinical characteristics and *MRP2/UGT1A9* genotypes of 53 pediatric kidney recipients.

Characteristics	
Male/Female	28/25
Cause of ESRD	
Structural anomaly	35 (66.0%)
Glomerulopathy	5 (9.4%)
Monogenic cause (confirmed)	5 (9.4%)
Vascular cause	3 (5.7%)
Other	2 (3.8%)
Undetermined or unknown	3 (5.7%)
Age median at time of transplantation (years) [IQR]	8.8 [4.1–11.2]
Body surface median at time of transplantation (m2) [IQR]	1.01 [0.66–1.24]
MRP2 -24G > A genotype[a]	
-24G/-24G	30 (57.7%)
-24G/-24A	21 (40.4%)
-24A/-24A	1 (1.9%)
UGT1A9 -275T > A genotype	
-275T/-275T	46 (86.8%)
-275T/-275A	7 (13.2%)
-275A/-275A	0 (0%)
UGT1A9 -2152C > T genotype[a]	
-2152C/-2152C	49 (94.2%)
-2152C/-2152T	3 (5.8%)
-2152T/-2152T	0 (0%)

[a]*In one sample the genotype was not determined.*

genotype and allele frequencies were not significantly different than expected if the population was in Hardy-Weinberg equilibrium (*MRP2* -24C > T χ^2 = 1.54, p = 0.21; *UGT1A9* -275T > A χ^2 = 0.26, p = 0.61; *UGT1A9* -2152C > T χ^2 = 0.046, p = 0.83), although a low frequency of some variant alleles was observed (**Table 2**). In particular, the frequencies of the *MRP2* -24A, *UGT1A9* -275A and *UGT1A9* -2152T alleles were 22.1, 6.6, and 2.9%, respectively. No significant differences between genotypes were observed in MPA-C_0, MPA-D or MPA-C_0/D (*MRP2* -24C > T: MPA-C_0 p = 0.93, MPA-D p = 0.66, MPA-C_0/D p = 0.86; *UGT1A9* -275T > A: MPA-C_0 p = 0.27, MPA-D p = 0.50, MPA- C_0/D p = 0.54; *UGT1A9* 2152C > T: MPA-C_0 p = 0.30, MPA-D p = 0.99, MPA- C_0/D p = 0.36).

Since the *UGT1A9* and *MRP2* alleles have been reported as variants associated with higher enzyme and transporter activities and higher MPA dose requirements, AUC_{0-12h}/MPA-D was compared between carriers and non-carriers of the variant alleles. Interestingly, patients carrying the *UGT1A9*-275A variant allele had lower AUC_{0-12h}/MPA-D (0.053 [IQR 0.040–0.100] ug*hr/ml/mg/m2) than patients carrying only the *UGT1A9*-275T ancestral allele genotype (0.117 [IQR 0.058–0.150]ug*hr/ml/mg/m2) with a difference of marginal significance (p = 0.05) (**Figure 4**). This difference was not observed in the other *UGT1A9* (p = 0.33) or *MRP2* (p = 0.29) alleles explored in this study.

DISCUSSION

An individualized immunosuppressive therapy is essential to minimize risks after transplantation. TAC is used as the first-line calcineurin inhibitor in many centers performing kidney transplantation in pediatric recipients given the evidence of its clinical effectiveness in preventing acute rejection and optimizing the function of the transplanted kidney (KDIGO Transplant Work, 2009). The main enzyme responsible for the TAC metabolism is *CYP3A5* and contains a well-studied genetic variant, *CYP3A5*3*, which represents the most frequent allele associated with a non-functional CYP3A5 enzyme worldwide (Kuehl et al., 2001).

According to our results, the *CYP3A5*3* allele seems to be predominant in the Chilean population, which could be due in part to the genetic structure of Chileans as a consequence of the admixture between Native Amerindians, the historical migration of individuals from Europe and, to a lesser extent, the arrival of African individuals (Eyheramendy et al., 2015). Similar results have been observed in other admixed Latin American populations in terms of the presence of the three *CYP3A5* genotypes, the association with different TAC pharmacokinetic patterns and the influence of ancestry (Ferraris et al., 2011; Cusinato et al., 2014; Galaviz-Hernández et al., 2020). GWAS studies in North American, Caucasian, African and Asiatic populations have validated the role of *CYP3A5* as the main predictor of TAC metabolism, although a few other common variants in cytochrome P450 enzymes and drug transporters have been suggested as contributing to the variability in plasma drug levels (Oetting et al., 2018; Bruckmueller et al., 2015; Oetting et al., 2016). Latin American populations constitute a challenge with a hidden genetic complexity and are highly underrepresented in clinical-pharmacological research studies, but deserve to be analyzed to propose individualized therapies based on local experience for the pediatric as well as the adult population. The disparities in studies for some geographic and ethnic groups such as Latin American countries limit our

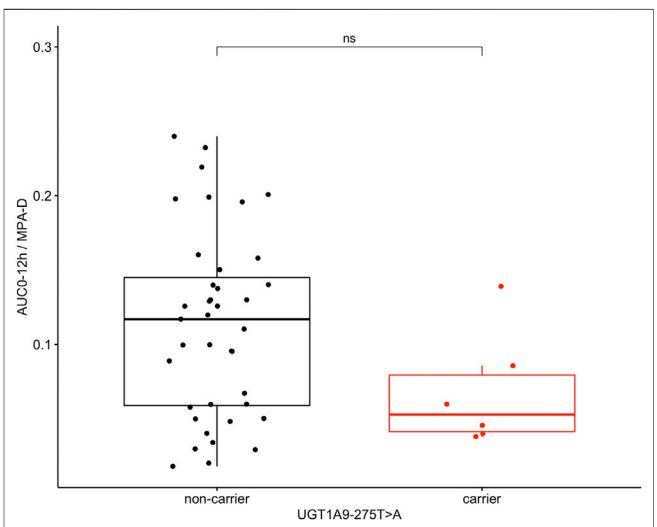


FIGURE 4 | Area under the curve in 12 h normalized by MPA dose requirements (AUC_{0-12h}/MPA-D ug*hr/ml/mg/m2) in 45 patients as non-carriers or carriers of the *UGT1A9*-275T>A variant allele.

understanding of their clinical phenotypes, and reduce access to personalized medicine to improve health care. In addition, association studies between polymorphisms and drug levels are more frequently performed on adult populations, with less evidence for pediatric populations as we present in this study.

In this study we decided to normalize AUC_{0-12h} by dose to reduce differences between patients and their characteristics at the time of transplantation, and the fact that AUC calculations were performed according to the clinical requirement to determine drug exposure that occurred at different days in each patient (before or after 3 months post-transplantation). The AUC_{0-12h}/TAC-D values were higher for *CYP3A5*3/*3* carriers and remained high in patients with steroid withdrawal. This is particularly important since the AUC calculation is considered an adequate marker for drug exposure and high values are associated with chronic nephrotoxicity. However, steroid withdrawal in patients using TAC and MPA as immunosuppressive therapy might be valuable to achieve better long-term graft function and patient survival by preventing hypertension, impaired glucose metabolism, growth retardation, obesity and infections, among other adverse effects (Sarwal et al., 2001; Gülhan et al., 2014). Although some patients presented high AUC_{0-12h}/TAC-D values, we saw no evidence of supratherapeutic immunosuppression during the 6-months follow-up. On the other hand, the *CYP3A5*1* allele was less predominant (23.1%), but strongly associated with lower AUC_{0-12h}/TAC-D values without evidence of subtherapeutic immunosuppression during the follow-up and lower trough TAC levels. In fact, within the first three months fewer than 15% of patients carrying the *CYP3A5*1* allele had reached target levels, suggesting that these patients require fine monitoring and that they may steroid therapy during the first post-transplantation period to prevent graft rejection.

TAC pharmacokinetics were strongly associated with the *CYP3A5* genotype in the study population. Of note, empirical dose adjustment during the follow-up showed substantial differences from the third month onwards between genotypes, with the daily dose requirement being 1.6–1.9 fold higher in CYP3A5 expressors (*CYP3A5*3/*1* and *CYP3A5*1/*1*) compared to CYP3A5 non-expressors (*CYP3A5*3/*3*). This observation is consistent with the CPIC (Clinical Pharmacogenetics Implementation Consortium) guideline for TAC dosing, which recommends increasing the starting dose by 1.5–2 times in children and adolescents CYP3A5 expressors followed by therapeutic drug monitoring as recommended for adults (Birdwell et al., 2015).

The influence of genetic variants on MPA pharmacokinetics has been less validated than those related to TAC pharmacokinetics, which might be explained in part by the low frequency of variants in metabolizing enzymes and transporters. According to our data, the SNP in the *UGT1A9* promoter region, *UGT1A9* -275T > A, presents low allele frequency (6.6%) and was associated with lower MPA exposure due to higher glucouronidating activity in the liver as proposed previously in the literature (Girard et al., 2004). A larger cohort is needed to continue validation of *MRP2* and *UGT1A9* variants or explore non-genetic determinants to better understand MPA pharmacokinetics in pediatric kidney recipients (Hesselink and Van Gelder, 2005).

Our study had certain limitations, such as not having considered the haplotypes or the co-administration of TAC and MPA in the analysis. Given the limited number of participants and the low frequency of some alleles, the haplotype analysis was not included in this study, because the statistical analysis might not meet the targets. In addition, the combined use of TAC, MPA and steroids is recommended in transplantation of solid organs to prevent graft rejection, but several studies have suggested that drug interaction may occur when TAC and MPA are combined. In fact, a lower TAC clearance was observed in healthy volunteers when this drug was co-administered with MPA, indicating that enzymes and transporters are involved in the metabolism of both drugs (Kim et al., 2018). However, this drug-drug interaction is still controversial and a larger dataset is required to test this effect to prevent under- or over-immunosuppression in patients of different ethnic and geographic origin (Kagaya et al., 2008; Park et al., 2009).

According to the evidence, the expression of hepatic enzymes varies with age. During the fetal period, CYP3A7 predominates being replaced gradually after birth by CYP3A4 and CYP3A5. As a consequence, it may be that CYP3A5 protein has different effect on drug metabolism in pediatric patients at different ages. Although the analysis of ontogeny was not presented as part of this study, we compared patients according to the Tanner score (prepubescent vs pubescent patients) and found no significant differences between them that could be a consequence of the limited dataset of pubescent patients (data not shown). However, it might be hypothesized that variants in *CYP3A5* cause multiple clinical phenotypes, generating a wide spectrum that ranges from low expressor-poor metabolizer in younger patients to high expressor-rapid metabolizer phenotypes in older patients.

On the other hand, we are aware of the need to perform a long-term follow-up in association with genetic determinants to describe the presence of adverse events and/or acute graft rejection, since these situations might evolve from sub- or supra-therapeutic immunosuppression with no evident clinical manifestations in months. The use of biomarkers in the pre-transplantation and post-transplantation periods related to pharmacodynamic effects has been proposed as a complementary tool to predict the effects of immunosuppression (Wieland et al., 2010).

Finally, we could not rule out the nonadherence that is known to be prevalent among pediatric transplant recipients and that places them at risk of rejection and graft loss, especially those of older age and specific sociodemographic characteristics (Dew et al., 2009). Medication adherence can be measured by evaluating the fluctuation of plasma drugs levels with a median that results usually below the target levels (Fredericks et al., 2007). Nonadherence was not analyzed in this study population, but the *CYP3A5* and *UGT1A9* genotype assessment might give a better understanding of the patients' metabolism and particularly in those of high enzymatic activity with low trough levels and high dose requirements to enhance credibility of their adherence to immunosuppressive drugs.

In conclusion, our results reinforce the need to consider the *CYP3A5* and *UGT1A9* genotype analysis in local transplantation

guidelines for pediatric recipients that could be assessed while on the waiting list. Additionally, the routine parameter C_0/D as part of therapeutic drug monitoring and drug adherence could be better interpreted according to the genotype to guide TAC and MPA dosing safely and perform a fine follow-up of patients at risk of adverse events (viral infections, malignancies, graft rejection). In the future, further analysis will be needed to evaluate the impact of early TAC and MPA dose adjustment according to the *CYP3A5* and *UGT1A9* genotypes.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher. The dataset Krall et al. 2021 TAC-MPA for this study can be found in the FigShare with DOI: 10.6084/m9.figshare.13574555.

ETHICS STATEMENT

The study was conducted according to the declaration of Helsinki and was approved by the Ethics Committees of the Universidad de Chile and Hospital Luis Calvo Mackenna. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

AR, AD, MF, and CS conceived the study. DY, AR, AD, MC, JM, PB, AD, NE, and NA provided samples and clinical data. PK and LEC performed the analysis. PK, AR, MF, and CS compiled the draft manuscript and CS provided funding. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the "Sociedad Chilena de Trasplante-Concurso de investigación Científica 2017–2019" awarded to CS.

ACKNOWLEDGMENTS

The authors acknowledge the contributions of Iris Delgado from the Instituto de Ciencias e Innovación en Medicina, Universidad del Desarrollo in Chile that aided the efforts of the authors at the beginning of this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.653525/full#supplementary-material.

REFERENCES

Barraclough, K. A., Lee, K. J., and Staatz, C. E. (2010). Pharmacogenetic influences on mycophenolate therapy. *Pharmacogenomics* 11, 369. doi:10.2217/pgs.10.9

Bernard, O., and Guillemette, C. (2004). The main role of UGT1A9 in the hepatic metabolism of mycophenolic acid and the effects of naturally occurring variants. *Drug Metab. Dispos.* 32 (8), 775–778. doi:10.1124/dmd.32.8.775

Bruckmueller, H., Werk, A. N., Renders, L., Feldkamp, T., Tepel, M., Borst, C., et al. (2015). Which Genetic Determinants Should be Considered for Tacrolimus Dose Optimization in Kidney Transplantation? A Combined Analysis of Genes Affecting the CYP3A Locus. *Ther Drug Monit.* 37, 288–295. doi:10.1097/FTD.0000000000000142

Birdwell, K., Decker, B., Barbarino, J., Peterson, J., Stein, C., Sadee, W., et al. (2015). Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines for CYP3A5 genotype and tacrolimus dosing. *Clin. Pharmacol. Ther.* 98 (1), 19–24. doi:10.1002/cpt.113

Cusinato, D. A. C., Lacchini, R., Romao, E. A., Moysés-Neto, M., and Coelho, E. B. (2014). Relationship of CYP3A5 genotype and ABCB1 diplotype to tacrolimus disposition in Brazilian kidney transplant patients. *Br. J. Clin. Pharmacol.* 78 (2), 364–372. doi:10.1111/bcp.12345

De Wildt, S. N. D., Van Schaik, R. H. N., Soldin, O. P., Soldin, S. J., Brojeni, P. Y., Van Der Heiden, I. P., et al. (2011). The interactions of age, genetics, and disease severity on tacrolimus dosing requirements after pediatric kidney and liver transplantation. *Eur. J. Clin. Pharmacol.* 67 (12), 1231–1241. doi:10.1007/s00228-011-1083-7

Delucchi, A., Valenzuela, M., Ferrario, M., Lillo, A. M., Lillo, J. L., Rodriguez, E., et al. (2007). Early steroid withdrawal in pediatric renal transplant on newer immunosuppressive drugs. *Pediatr. Transplant.* 11 (7), 743–748. doi:10.1111/j.1399-3046.2007.00735.x

Dew, M. A., Dabbs, A. D. V., Myaskovsky, L., Shyu, S., Shellmer, D. A., DiMartini, A. F., et al. (2009). Meta-analysis of medical regimen adherence outcomes in pediatric solid organ transplantation. *Transplantation* 88 (5), 736–746. doi:10.1097/tp.0b013e3181b2a0e0

Eyheramendy, S., Martinez, F. I., Manevy, F., Vial, C., and Repetto, G. M. (2015). Genetic structure characterization of Chileans reflects historical immigration patterns. *Nat. Commun.* 6, 6472. doi:10.1038/ncomms7472

Ferraris, J. R., Argibay, P. F., Costa, L., Jimenez, G., Coccia, P. A., Ghezzi, L. F. R., et al. (2011). Influence of CYP3A5 polymorphism on tacrolimus maintenance doses and serum levels after renal transplantation: age dependency and pharmacological interaction with steroids. *Pediatr. Transplant.* 15 (5), 525–532. doi:10.1111/j.1399-3046.2011.01513.x

Filler, G. (2004). Abbreviated mycophenolic acid AUC from C0, C1, C2, and C4 is preferable in children after renal transplantation on mycophenolate mofetil and tacrolimus therapy. *Transpl. Int.* 17 (3), 120–125. doi:10.1007/s00147-003-0678-z

Fredericks, E. M., Lopez, M. J., Magee, J. C., Shieck, V., and Opipari-Arrigan, L. (2007). Psychological functioning, nonadherence and health outcomes after pediatric liver transplantation. *Am. J. Transpl.* 7 (8), 1974–1983. doi:10.1111/j.1600-6143.2007.01878.x

Fukuda, T., Goebel, J., Cox, S., Maseck, D., Zhang, K., Sherbotie, J. R., et al. (2012). UGT1A9, UGT2B7, and MRP2 genotypes can predict mycophenolic acid pharmacokinetic variability in pediatric kidney transplant recipients. *Ther. Drug Monit.* 34 (6), 671–679. doi:10.1097/ftd.0b013e3182708f84

Galaviz-Hernández, C., Lazalde-Ramos, B. P., Lares-Assef, I., Macías-Salas, A., Ortega-Chavez, M. A., Rangel-Villalobos, H., et al. (2020). Influence of genetic admixture components on CYP3A5*3 allele-associated hypertension in amerindian populations from Northwest Mexico. *Front. Pharmacol.* 11, 638. doi:10.3389/fphar.2020.00638

Girard, H., Court, M. H., Bernard, O., Fortier, L.-C., Villeneuve, L., Hao, Q., et al. (2004). Identification of common polymorphisms in the promoter of the UGT1A9 gene. *Pharmacogenetics* 14 (8), 501–515. doi:10.1097/01.fpc.0000114754.08559.27

Golubovic, B., Prostran, M., Miljkovic, B., Vucicevic, K., Radivojevic, D., and Grabnar, I. (2016). Population pharmacokinetic approach of immunosuppressive therapy in kidney transplant patients. *Cmc* 23 (19), 1998–2011. doi:10.2174/0929867323666151221150214

Gülhan, B., Topaloğlu, R., Karabulut, E., Özaltın, F., Aki, F. T., Bilginer, Y., et al. (2014). Post-transplant hypertension in pediatric kidney transplant recipients. *Pediatr. Nephrol.* 29 (6), 1075–1080. doi:10.1007/s00467-013-2721-3

Hesselink, D. A., and Van Gelder, T. (2005). Genetic and nongenetic determinants of between-patient variability in the pharmacokinetics of mycophenolic acid. *Clin. Pharmacol. Ther.* 78 (4), 317–321. doi:10.1016/j.clpt.2005.06.008

Kagaya, H., Miura, M., Satoh, S., Inoue, K., Saito, M., Inoue, T., et al. (2008). No pharmacokinetic interactions between mycophenolic acid and tacrolimus in renal transplant recipients. *J. Clin. Pharm. Ther.* 33 (2), 193–201. doi:10.1111/j.1365-2710.2008.00906.x

Kassambara, A. (2020a). "Ggpubr: 'ggplot2' based publication ready plots". R Package Version 0.4.0. Available at: https://CRAN.R-Project.Org/Package=ggpubr.

Kassambara, A. (2020b). "Rstatix: pipe-friendly framework for basic statistical tests". R Package Version 0.6.0. Available at: https://CRAN.R-Project.Org/Package=rstatix.

KDIGO Transplant Work (2009). Special issue: KDIGO clinical practice guideline for the care of kidney transplant recipients. *Am. J. Transplant.* 9, S1–S155. doi:10.1111/j.1600-6143.2009.02834.x

Kim, J. H., Han, N., Kim, M. G., Yun, H.- Y., Lee, E., Bae, E., et al. (2018). Increased exposure of tacrolimus by co-administered mycophenolate mofetil: population pharmacokinetic analysis in healthy volunteers. *Sci. Rep.* 8 (1), 1687. doi:10.1038/s41598-018-20071-3

Kuehl, P., Zhang, J., Lin, Y., Lamba, J., Assem, M., Schuetz, J., et al. (2001). Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. *Nat. Genet.* 27 (4), 383–391. doi:10.1038/86882

Lawrence, M. A. (2013). Ez: easy analysis and visualization of factorial experiments. https://rdrr.io/cran/ez/.

Macphee, I. A. M., Fredericks, S., Tai, T., Syrris, P., Carter, N. D., Johnston, A., et al. (2002). Tacrolimus pharmacogenetics: polymorphisms associated with expression of cytochrome P4503a5 and P-glycoprotein correlate with dose requirement. *Transplantation* 74 (11), 1486–1489. doi:10.1097/00007890-200212150-00002

MacPhee, I. A. M., Fredericks, S., Tai, T., Syrris, P., Carter, N. D., Johnston, A., et al. (2004). The influence of pharmacogenetics on the time to achieve target tacrolimus concentrations after kidney transplantation. *Am. J. Transpl.* 4 (6), 914–919. doi:10.1111/j.1600-6143.2004.00435.x

McDonald, S. P., and Craig, J. C. (2004). Long-term survival of children with end-stage renal disease. *N. Engl. J. Med.* 350 (26), 2654–2662. doi:10.1056/nejmoa031643

Min, S.-I., Kim, S. Y., Ahn, S. H., Min, S.-K., Kim, S. H., Kim, Y. S., et al. (2010). CYP3A5 *1 allele: impacts on early acute rejection and graft function in tacrolimus-based renal transplant recipients. *Transplantation* 90 (12), 1394–1400. doi:10.1097/tp.0b013e3181fa93a4

Naesens, M., Kuypers, D. R. J., Verbeke, K., and Vanrenterghem, Y. (2006). Multidrug resistance protein 2 genetic polymorphisms influence mycophenolic acid exposure in renal allograft recipients. *Transplantation* 82 (8), 1074–1084. doi:10.1097/01.tp.0000235533.29300.e7

Oetting, W. S., Schladt, D. P, Guan, W., Miller, M. B., Remmel, R. P., and Dorr, C. (2016). Genomewide Association Study of Tacrolimus Concentrations in African American Kidney Transplant Recipients Identifies Multiple CYP3A5 Alleles. *Am J Transplant.* 16 (8), 574–582. doi:10.1111/ajt.13495

Oetting, W. S., Wu, B., Schladt, D. P., Guan, W., Remmel, R. P., and Dorr, C. (2018). Attempted validation of 44 reported SNPs associated with tacrolimus troughs in a cohort of kidney allograft recipients. *Pharmacogenomics* 19, 175–184. doi:10.2217/pgs-2017-0187

Park, S.-I., Felipe, C. R., Pinheiro-Machado, P. G., Garcia, R., Fernandes, F. B., Casarini, D. E., et al. (2009). Tacrolimus pharmacokinetic drug interactions: effect of prednisone, mycophenolic acid or sirolimus. *Fundam. Clin. Pharmacol.* 23 (1), 137–145. doi:10.1111/j.1472-8206.2008.00644.x

Rong, G., Jing, L., Deng-Qing, L., Hong-Shan, Z., Shai-Hong, Z., and Xin-Min, N. (2010). Influence of CYP3A5 and MDR1(ABCB1) polymorphisms on the pharmacokinetics of tacrolimus in Chinese renal transplant recipients. *Transplant. Proc.* 42 (9), 3455–3458. doi:10.1016/j.transproceed.2010.08.063

Sarwal, M. M., Yorgin, P. D., Alexander, S., Millan, M. T., Belson, A., Belanger, N., et al. (2001). Promising early outcomes with a novel, complete steroid avoidance immunosuppression protocol in pediatric renal transplantation1. *Transplantation* 72 (1), 13–21. doi:10.1097/00007890-200107150-00006

Staatz, C., Taylor, P., and Tett, S. (2001). Low tacrolimus concentrations and increased risk of early acute rejection in adult renal transplantation. *Nephrol. Dial. Transplant.* 16 (9), 1905–1909. doi:10.1093/ndt/16.9.1905

Thervet, E., Anglicheau, D., King, B., Schlageter, M. H., Schlageter, M.-H., Cassinat, B., et al. (2003). Impact of cytochrome P450 3A5 genetic polymorphism on tacrolimus doses and concentration-to-dose ratio in renal transplant recipients1 2. *Transplantation* 76 (8), 1233–1235. doi:10.1097/01.tp.0000090753.99170.89

Wieland, E., Olbricht, C. J., Süsal, C., Gurragchaa, P., Böhler, T., Israeli, M., et al. (2010). Biomarkers as a tool for management of immunosuppression in transplant patients. *Ther. Drug Monit.* 32 (5), 560–572. doi:10.1097/ftd.0b013e3181efb3d2

Wilson, A., and Norden, N. (2015). The R project for statistical computing the R project for statistical computing. Available at: https://www.r-project.org/254.

Winterberg, P. D., and Garro, R. (2019). Long-term outcomes of kidney transplantation in children pediatric clinics of North America. 66 (1), 269–280. doi:10.1016/j.pcl.2018.09.008

Wong, K. M., Shek, C. C., Chau, K. F., and Li, C. S. (2000). Abbreviated tacrolimus area-under-the-curve monitoring for renal transplant recipients. *Am. J. Kidney Dis.* 35 (4), 660–666. doi:10.1016/s0272-6386(00)70013-8

Xie, H. G., Wood, A. J., Kim, R. B, Stein, C. M., Wilkinson, G. R., et al. (2004). Genetic variability in CYP3A5 and its possible consequences. *Pharmacogenomics* 5, 243–272. doi:10.1517/phgs.5.3.243.29833

Zheng, H., Webber, S., Zeevi, A., Schuetz, E., Zhang, J., Bowman, P., et al. (2003). Tacrolimus dosing in pediatric heart transplant patients is related to CYP3A5 and MDR1 gene polymorphisms. *Am. J. Transplant.* 3 (4), 477–483. doi:10.1034/j.1600-6143.2003.00077.x

Zwart, T. C., Gokoel, S. R. M., van der Boog, P. J. M., de Fijter, J. W., Kweekel, D. M., Swen, J. J., et al. (2018). Therapeutic drug monitoring of tacrolimus and mycophenolic acid in outpatient renal transplant recipients using a volumetric dried blood spot sampling device. *Br. J. Clin. Pharmacol.* 84 (12), 2889–2902. doi:10.1111/bcp.13755

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 19 May 2021
doi: 10.3389/fphar.2021.660965

Pharmacogenetic Associations Between Atazanavir/*UGT1A1*28* and Efavirenz/rs3745274 (*CYP2B6*) Account for Specific Adverse Reactions in Chilean Patients Undergoing Antiretroviral Therapy

Daniela Poblete[1], *Fernando Bernal*[2], *Gabriel Llull*[3], *Sebastian Archiles*[3], *Patricia Vasquez*[2], *Leonardo Chanqueo*[2], *Nicole Soto*[1], *María A. Lavanderos*[1,4], *Luis A. Quiñones*[1,4] *and Nelson M. Varela*[1,4]*

[1]*Laboratory of Chemical Carcinogenesis and Pharmacogenetics, Department of Basic and Clinical Oncology, Faculty of Medicine, University of Chile, Santiago, Chile,* [2]*Department of Infectious Diseases, Hospital San Juan de Dios, Santiago, Chile,* [3]*Clinical Laboratory, Hospital San Juan de Dios, Santiago, Chile,* [4]*Latin American Network for the Implementation and Validation of Clinical Pharmacogenomics Guidelines (RELIVAF-CYTED), Madrid, Spain*

OPEN ACCESS

Edited by:
Chonlaphat Sukasem, Mahidol University, Thailand

Reviewed by:
Chalirmporn Atasilp, Thammasat University, Thailand
Salvador F. Aliño, University of Valencia, Spain

***Correspondence:**
Nelson M. Varela
nvarela@med.uchile.cl

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *30 January 2021*
Accepted: *27 April 2021*
Published: *19 May 2021*

Citation:
*Poblete D, Bernal F, Llull G, Archiles S, Vasquez P, Chanqueo L, Soto N, Lavanderos MA, Quiñones LA and Varela NM (2021) Pharmacogenetic Associations Between Atazanavir/*UGT1A1*28 and Efavirenz/rs3745274 (*CYP2B6*) Account for Specific Adverse Reactions in Chilean Patients Undergoing Antiretroviral Therapy. Front. Pharmacol. 12:660965. doi: 10.3389/fphar.2021.660965*

Background: Efavirenz (EFV), a non-nucleoside reverse transcriptase inhibitor, and atazanavir (ATV), a protease inhibitor, are drugs widely used in antiretroviral therapy (ART) for people living with HIV. These drugs have shown high interindividual variability in adverse drug reactions (ADRs). *UGT1A1*28* and *CYP2B6* c.516G>T have been proposed to be related with higher toxicity by ATV and EFV, respectively.

Objective: To study the association between genetic polymorphisms and ADRs related to EFV or ATV in patients living with HIV treated at a public hospital in Chile.

Methods: Epidemiologic, case–control, retrospective, observational study in 67 adult patients under EFV or ATV treatment was conducted, in the San Juan de Dios Hospital. Data were obtained from patients' medical records. Genotype analyses were performed using rtPCR for rs887829 (indirect identification of *UGT1A1*28* allele) and rs3745274 (*CYP2B6* c.516G>T), with TaqMan® probes. The association analyses were performed with univariate logistic regression between genetic variants using three inheritance models (codominant, recessive, and dominant).

Results: In ATV-treated patients, hyperbilirubinemia (total bilirubin >1.2 mg/dl) had the main incidence (61.11%), and moderate and severe hyperbilirubinemia (total bilirubin >1.9 mg/dl) were statistically associated with *UGT1A1*28* in recessive and codominant inheritance models (OR = 16.33, $p = 0.028$ and OR = 10.82, $p = 0.036$, respectively). On the other hand, in EFV-treated patients adverse reactions associated with CNS toxicity reached 34.21%. In this respect, nightmares showed significant association with *CYP2B6* c.516G>T, in codominant and recessive inheritance models (OR = 12.00, $p = 0.031$ and OR = 7.14, $p = 0.042$, respectively). Grouped CNS ADRs (nightmares, insomnia, anxiety, and suicide attempt) also showed a statistically significant association with *CYP2B6*

c.516G > T in the codominant and recessive models (OR = 30.00, p = 0.011 and OR = 14.99, p = 0.021, respectively).

Conclusion: Our findings suggest that after treatment with ATV or EFV, *UGT1A1*28* and *CYP2B6* c.516G>T influence the appearance of moderate-to-severe hyperbilirubinemia and CNS toxicity, respectively. However, larger prospective studies will be necessary to validate these associations in our population. Without a doubt, improving adherence in patients living with HIV is a critical issue to the success of therapy. Hence, validating and applying international pharmacogenetic recommendations in Latin American countries would improve the precision of ART: a fundamental aspect to achieve the 95–95–95 treatment target proposed by UNAIDS.

Keywords: UGT1A1, CYP2B6, ADRs, atazanavir, efavirenz, HIV, pharmacogenetic, antiretroviral

INTRODUCTION

Human immunodeficiency virus (HIV) affects up to 38 million people worldwide, and it is considered a pandemic. In 2010, HIV/acquired immunodeficiency syndrome (AIDS) was the eighth cause of death, but in 2019, it was the 19th. It means that prevention, detection, and treatment initiatives have been successful in recent decades. But in recent years, the World Health Organization (WHO) reported a general slowdown in progress against infectious diseases such as HIV, tuberculosis, and malaria. In that context, the Joint United Nations Programme on HIV/AIDS (UNAIDS) proposed an ambitious goal: the 95–95–95 treatment target. It means that in 2030, 95% of all people living with HIV will know their HIV status, 95% of all people with diagnosed HIV infection will receive sustained antiretroviral therapy, and 95% of all people receiving antiretroviral therapy will have viral suppression (ONUSIDA, 2014). Currently, in Chile, according UNAIDS, we have reached 90-68-62 (ONUSIDA, 2019).

Antiretroviral therapy (ART) with associations of three drugs has a great impact. It is highly effective in suppressing viral replication (Moore et al., 2005), allows the qualitative and quantitative recovery of the immune response (Gras et al., 2007), prevents clinical progression, and significantly reduces associated mortality, despite there being no eradication of the virus (FrankPalella et al., 2003). The generally accepted formula for a potent antiretroviral regimen includes two nucleoside reverse transcriptase inhibitors (NRTIs) added to a third from another family: a non-nucleoside reverse transcriptase inhibitor such as efavirenz (EFV), a protease inhibitor such as atazanavir (ATV), or integrase inhibitor such as raltegravir (MINSAL, 2013). Recently, in Chile, there has been an increase in the prescription of integrase inhibitors, thus reducing the prescription of EFV and ATV. This is mainly due to the safety criteria, since it is expected that patients will have fewer adverse drug reactions (ADRs) associated with ART.

To improve the current HIV situation in our country, therapeutic adherence to ART becomes relevant, considering that ADRs are one of the main causes of treatment abandonment. For example, 5–10% of patients receiving ATV discontinue treatment due to ADRs, whereas at least 50% of patients receiving EFV present ADRs (Gaida et al., 2016). This variability is based on the characteristics of the drug and the patient. At the last point, the patient's own genetics helps to explain the interindividual variability in the presence of adverse events (Roden and George, 2002; Quiñones et al., 2017).

The Clinical Pharmacogenetics Implementation Consortium (CPIC; https://cpicpgx.org/) is an international consortium of individual volunteers and staff who are interested in facilitating the use of pharmacogenetic test for patient care. Its goal is to address this barrier to clinical implementation of pharmacogenetic test by creating, curating, and posting freely available, peer-reviewed, evidence-based, updatable, and detailed gene/drug clinical practice guidelines (https://cpicpgx.org/guidelines/). Each clinical guide provides information to allow the interpretation of clinical genotype test so that the results can be used to inform the prescribing of drugs (Caudle et al., 2014).

EFV and ATV, drugs widely used in ART in people living with HIV, have shown to possess high interindividual variability in ADRs. EFV is predominantly metabolized by CYP2B6, a highly polymorphic enzyme. Patients with certain genetic variants may be at increased risk of developing ADRs at the central nervous system (CNS) level. The single-nucleotide polymorphism (SNP) *CYP2B6* c.516G>T (NM_000767.5; NG_007929.1:g.20638G > T; rs3745274; *p*. Q172H) has been extensively studied. All alleles carrying c.516G>T [e.g., *CYP2B6*6*, in haplotype with c. A785G (Manosuthi et al., 2014)] are considered to lead to a reduced enzyme activity (Hofmann, et al., 2008). There is substantial evidence that links this SNP (rs3745274) with increase in plasma EFV concentrations (Sukasem et al., 2013; Desta et al., 2019). In this case, for these patients, there is a moderate recommendation from the CPIC guideline to consider starting efavirenz with a reduced dose of 400 or 200 mg/day. This recommendation reflects the fact that most CYP2B6 poor metabolizers do not discontinue efavirenz 600 mg/day due to ADRs (Desta et al., 2019).

On the other hand, ATV inhibits hepatic uridine diphosphate glucuronosyltransferase (UGT)1A1, the major UGT1A subfamily enzyme, expressed primarily in the liver and gastrointestinal tract (Court et al., 2012), and is essential for the metabolization (conjugation) of bilirubin and some drugs (Bosma et al., 1994; Hongkaew et al., 2018). Reduced UGT1A1 activity either through

genetic variation (Kadakol et al., 2000; Strassburg, 2008) or catalytic inhibition by drugs (Zhang et al., 2005) results in the accumulation of unconjugated (indirect) bilirubin in blood and tissues. Mechanistic studies using promoter–reporter constructs have shown that the presence of a TATAA consensus element in *UGT1A1* promotor containing seven TA dinucleotide repeats (TA_7 or *UGT1A1*28* allele) causes reduction in gene transcription as compared with the reference TA_6 (*UGT1A1*1*) allele (Bosma et al., 1995), possibly due to reduced binding affinity for transcription factors including TATA-binding protein (Hsieh et al., 2007). Thereby, the presence of ATV prevents the glucuronidation and elimination of bilirubin; resultant indirect hyperbilirubinemia can cause jaundice and a potential discontinuation of atazanavir treatment. Risk for bilirubin-related discontinuation is highest among individuals who carry *UGT1A1*28*, decreased function alleles. The CPIC guideline suggests a *UGT1A1* genotype is most helpful if available before atazanavir is prescribed (Gammal et al., 2016). Due to the methodological complexity involved in determining the presence of *UGT1A1*28*, the majority of population studies apply an indirect identification strategy, identifying a SNP located less than 300 base pairs from the TA repeat, rs887829 (c. −364C>T; *UGT1A1*80*), where the T allele is in a very strong linkage disequilibrium (LD) with the TA_7 allele (*UGT1A1*28*) ($r^2 \cong 0.99$) (Gammal et al., 2016). It is important to highlight that UGT1A1*6 is also associated to a reduced enzymatic function and could be relevant (Sukasem et al., 2016); however, the frequency in populations of non-Asian descent is extremely low (1,000 Genome) to be included in this study.

Since *UGT1A1*28* (rs8175347, TA_7) and *CYP2B6 c.516G>T* (rs3745274) have been proposed to be related with higher toxicity by ATV and EFV, respectively, the main goal of this research was to study the association between EFV or ATV ADRs and the presence of these genetic polymorphisms in patients from a public hospital in Chile.

PATIENTS AND METHODS

Patients

We carried out an epidemiologic, case–control, retrospective, observational study in eighty patients. The inclusion criteria were applicable for the following patients: 1) ≥18 years old; 2) under EFV (600 mg/day) or ATV (300 mg/day with 100 mg/day of ritonavir as a booster) treatment; 3) with A1, A2, B1, or B2 (patients without AIDS) clinical category (outpatients) according to the Chilean HIV Clinical Guide (MINSAL, 2013) in "San Juan de Dios Hospital" (SJDH) (Santiago, Chile); 4) adherent to treatment; and 5) with the clinical record available for reviewing. Exclusion criteria were the following: 1) patients with pre-ART diagnoses of kidney or liver insufficiency, 2) patients under voriconazole treatment, 3) existence of untreated active opportunistic disease susceptible to severe Immune Reconstitution Inflammatory Syndrome (IRIS) with the initiation of ART and/or *tuberculosis*, and 4) patients who do not respect the treatment and control protocol established by the hospital unit. Data were obtained from patients' medical records and organized in a coded database which included patient identification, clinical record number, sex, age (years), weight (kilograms), lymphocyte count (cells/mm^3), viral loads (copies/mL), liver transaminases (U/L), total bilirubin level (mg/dl), serum creatinine levels (mg/dl), and lipid levels (total cholesterol, HDL, LDL, and triglycerides).

Before starting treatment with EFV, patients are evaluated by a psychologist to rule out preexisting CNS pathologies. Patients are a part of the Adherence Program of SJDH; it consists of a basal control before starting treatment, and then after two weeks, one month, and every three months. All controls include a session with a pharmacist and a physician, and allow the generation of records of ADRs to ART, which are classified according to their presence or absence, and results of clinical and hematological laboratory analyses. ADRs registered were hyperbilirubinemia (total bilirubin level >1.2 mg/dl), gastrointestinal upset, rash, dizziness, nightmares, insomnia, headache, fever, anxiety, lipodystrophy, dyslipidemia, and suicide attempt. Hyperbilirubinemia was classified for total bilirubin levels as follows: grade 1 (mild), 1.3–1.9 mg/dl; grade 2 (moderate), 1.9–3.1 mg/dl; grade 3 (severe), 3.1–6.1 mg/dl; and grade 4 (serious), >6.1 mg/dl (Fellay et al., 2001; MINSAL, 2013). Adherence to treatment was assessed by applying an adherence questionnaire (Knobel et al., 2002) and correlated with pharmacy withdrawals and viral load levels. The need for volunteer patients to agree to answer the questionnaire was established in the document associated with the informed consent procedure.

Genotyping Analyses

Whole venous EDTA blood samples were obtained at the SJDH and analyzed in Laboratory of Chemical Carcinogenesis and Pharmacogenetics of the Faculty of Medicine, the University of Chile. The sample was immediately refrigerated at 4–6°C until centrifugation. Then the sample was centrifuged at 3,000 rpm at 4°C for 15 min. Subsequently, 250 µL of buffy coat was carefully removed from the interface and deposited in a properly labeled tube. The buffy coat sample was frozen at −80°C until DNA extraction with E.Z.N.A.® Blood DNA Mini Kit, OMEGA Bio-tek, Inc. (Norcross, Georgia, United States). Genomic DNA (gDNA) samples were quantified using a nanodrop Denovix® DS-11 FX Series Spectrophotometer (Wilmington, Delaware, United States). The absorbance ratio of 260/280 nm was used to evaluate the purity of the gDNA, considering a value higher than 1.7 as acceptable. The samples were genotyped using a Stratagene® Mx3000P Real-Time PCR system (Agilent Technologies, Waldbronn, Germany). The reaction mixture consisted of 30 ng of gDNA, 5.0 µL of *TaqMan*® Genotyping Master Mix 2X (Applied Biosystems™; catalog no. 4371355), 0.50 µL of *TaqMan*® SNP genotyping assay (Thermo Fischer Scientific Inc.®), and nuclease-free water (sufficient quantity for 10 uL). The thermal cycling conditions were as follows: an initial denaturation and enzyme activation at 95°C for 10 min, followed by 50 cycles, each consisting of two phases: denaturation at 95°C for 15 s and annealing/extension at 60°C for 90 s. For the determination of *CYP2B6* c.516G>T (rs3745274), the *TaqMan*® SNP genotyping assay catalog no.

TABLE 1 | Baseline characteristics of the patients (n = 67).

Age	**Years**
Mean ± *SD*	35.7 ± 10.4
Median	34
Age-group (years)	**n (%)**
20–29	23 (34.33)
30–39	24 (35.82)
40–49	15 (22.39)
≥50	5 (7.46)
Sex	**n (%)**
Female	13 (19.40)
Male	54 (80.60)
ART	**n (%)[a]**
Patients treated with ATV	36
+ abacavir/lamivudine	24 (66.67)
+ tenofovir/emtricitabine	10 (27.78)
+ zidovudine/lamivudine	2 (5.56)
Patients treated with EFV	38
+ abacavir/lamivudine	21 (55.26)
+ tenofovir/emtricitabine	13 (34.21)
+ zidovudine/lamivudine	4 (10.53)

SD, standard deviation; n, number of patients; ART, antiretroviral therapy; EFV, efavirenz; ATV, atazanavir.

[a]*Seven patients received EFV and ATV in different treatment regimens.*

4362691 (assay ID: C_7,817,765_60) was used. For the indirect determination of the *UGT1A1*28* allele, the identification of rs887829 is used as a strategy, through the *TaqMan*® SNP genotyping assay catalog no. 4351379 (assay ID: C_2,669,357_10).

Statistical Analyses

The association analyses were performed with univariate logistic regression between genetic variants using three inheritance models (codominant, recessive, and dominant) on STATA 12.0 (Copyright 1985–2011 StataCorp, LP, TX, United States), and a *p-value* < 0.05 was considered statistically significant. All association studies were assayed by testing three genetic models of inheritance, that is, dominant, codominant, and recessive models, and choosing parameters with better statistical association for each analysis. χ^2 and Fisher`s exact test were applied to compare proportions between independent groups.

Ethical Issues

The study was carried out under strict ethical procedures approved by the Ethics Committee of the Faculty of Medicine of the University of Chile (September 5, 2016) and authorized by the director of the SJDH (September 23, 2016) resolution No. 4855, in accordance with the procedures suggested in the Declaration of Helsinki (Declaration of Helsinki, 1964), and according to Chilean Laws 20,120, 20,584, and 19,628, and the guidelines of the Good Clinical Practices (GCP). All patients signed informed consent.

RESULTS

Table 1 shows baseline characteristics of the 67 Chilean patients who live with HIV. The mean age was 35.7 ± 10.4 years, median age was 34 years, and the age group from 30 to 39 years (35.82%) was the most prevalent. 54 patients were males (80.6%) and 13 females (19.4%). Our patients were treated with different pharmacological schemes from their diagnosis; 31 patients were treated with EFV, 29 used ATV, and 7 patients first received EFV or ATV in a specific scheme and afterward changed to ATV or EFV, respectively, or *vice versa*, so they were included individually on ATV and EFV statistical analysis. Antiretroviral drugs (NRTIs) that were administered concurrently with EFV included abacavir/lamivudine (n = 21; 55.26%), tenofovir/emtricitabine (n = 13; 34.21%), and zidovudine/lamivudine (n = 4; 10.53%); and the ATV-treated patients completed their tritherapy with abacavir/lamivudine (n = 24; 66.67%), tenofovir/emtricitabine (n = 10; 27.78%), and zidovudine/lamivudine (n = 2; 5.56%).

Twelve types of ADRs were registered from clinical records, and frequency was classified by the drug (**Table 2**). Hyperbilirubinemia had the main incidence (26.04% of all ADRs) and was observed in 61.11% of the patients treated with ATV. Gastrointestinal upset was observed in 18 patients (18.75%), and central nervous system (CNS) ADRs were present in 15 patients (15.63% of all ADRs).

Although there were common adverse events in both groups of patients (treated with EFV or ATV), the presence of ADRs at the CNS level was significantly more frequent in the group of patients who received EFV (χ2: 9.3924; p: 0.002), while hyperbilirubinemia (total bilirubin level >1.2 mg/dl) was more frequent in patients who received atazanavir (χ2: 23.4030; p: <0.0001), considering that the distribution of antiretroviral drug regimens that accompanied ATV and EFV (**Table 1**) did not present significant differences between both groups (p: 0.592, Fisher's exact test).

Table 3 shows genotypic and allelic frequency of polymorphisms in patients included in this study. Minor allele frequencies obtained for both polymorphisms were relatively high (*CYP2B6* c.516G>T, T: 0.38 and *UGT1A1*28*, TA_7: 0.36), which allowed a good statistical potency for associations.

Thirty-six patients were treated with ATV (**Table 1**); of the latter, only seven suffered severe hyperbilirubinemia (total bilirubin level >3.1 mg/dl). Univariate logistic regression analysis showed that the *UGT1A1*28* allele was significantly associated with severe hyperbilirubinemia, in a recessive inheritance model (OR = 8.33, p = 0.023, 95% CI: 1.33–52.03). The other inheritance models could not be evaluated due to the absence of cases. None of the patients had serious hyperbilirubinemia (total bilirubin level >6.1 mg/dl).

Table 4 shows results for the univariate logistic regression analysis between the *UGT1A1*28* allele and total bilirubin levels >1.9 mg/dl in patients treated with ATV; a statistically significant association was observed in the codominant and recessive inheritance models (OR = 16.33, p = 0.028 and OR = 10.82, p = 0.036, respectively). Headache, rash, and gastrointestinal upset did not show a statistically significant difference (data not shown).

TABLE 2 | Type of adverse drugs reactions in patients undergoing antiretroviral therapy with EFV or ATV.

ADR	n	%	EFV-treated patients	ATV-treated patients
Hyperbilirubinemia[a]	25	26.04	3	22
Gastrointestinal upset	18	18.75	6	12
Rash	10	10.42	7	3
Dizziness	13	13.54	12	1
Nightmares	12	12.50	11	1
Insomnia	5	5.21	4	1
Headache	6	6.25	3	3
Fever	2	2.08	1	1
Anxiety	1	1.04	1	0
Lipodystrophy	1	1.04	1	0
Dyslipidemia	1	1.04	1	0
Suicide attempt	2	2.08	1	1
CNS[b]	15	15.63	13	2

ADR, adverse drugs reaction; CNS, central nervous system; EFV, efavirenz; ATV, atazanavir.
[a]*Total bilirubin level >1.2 mg/dl.*
[b]*Nightmare, insomnia, anxiety, and suicide attempt were also grouped as CNS ADRs.*

TABLE 3 | Genotype and allele frequencies in patients recruited for this study.

Genetic variant	Genotypic frequency n (%)			Allelic frequency	
***CYP2B6* c.516G>T**	G/G	G/T	T/T	*f* G	*f* T
All patients (*n* = 67)	26 (38.81)	31 (46.27)	10 (14.92)	0.62	0.38
EFV-treated patients (*n* = 38)	14 (36.84)	18 (43.37)	6 (15.79)	0.61	0.39
UGT1A1*28[a]	TA_6/TA_6	TA_6/TA_7	TA_7/TA_7	*f* TA_6	*f* TA_7
All patients (*n* = 67)	31 (46.27)	24 (35.82)	12 (17.91)	0.64	0.36
ATV-treated patients (*n* = 36)	10 (27.78)	18 (50.00)	8 (22.22)	0.53	0.47

n, number of patients; f, frequency; EFV, efavirenz; ATV, atazanavir.
[a]*Indirect identification of the UGT1A1*28 allele using rs887829.*

TABLE 4 | Univariate logistic regression between *UGT1A1*28* (TA_7) and total bilirubin level >1.9 mg/dl (moderate and severe hyperbilirubinemia) in patients treated with atazanavir.

Inheritance model	Genotypes[a]	OR	*p*-value	95% CI
Codominant model	TA_6/TA_6	1.00	Ref	-
	TA_6/TA_7	1.87	0.456	0.36–9.63
	TA_7/TA_7	16.33	0.028*	1.35–197.77
Recessive model	TA_6/TA_6 + TA_6/TA_7	1.00	Ref	-
	TA_7/TA_7	10.82	0.036*	1.17–100.44
Dominant model	TA_6/TA_6	1.00	Ref	-
	TA_6/TA_7 + TA_7/TA_7	3.18	0.146	0.67–15.15

*OR: odds ratio; CI: confidence interval; ref: reference; *p-value < 0.05.*
[a]*Indirect identification of the UGT1A1*28 (TA_7) allele using rs887829.*

TABLE 5 | Univariate logistic regression between *CYP2B6* c.516G>T (rs3745274) and nightmares in EFV-treated patients.

Inheritance model	Genotypes	OR	*p*-value	95% CI
Codominant model	G/G	1.00	Ref	-
	G/T	2.31	0.367	0.37–14.21
	T/T	12.00	0.031*	1.25–115.36
Recessive model	G/G + G/T	1.00	Ref	-
	T/T	7.14	0.042*	1.08–47.42
Dominant model	G/G	1.00	Ref	-
	G/T + T/T	3.6	0.142	0.65–19.90

*OR, odds ratio; CI, confidence interval; ref, reference; *p-value < 0.05.*

Same analysis was made on 38 patients treated with EFV. We did not find statistical significance among hyperbilirubinemia, gastrointestinal upset, rash, headache, insomnia, and dizziness, and polymorphisms included in this study (data not shown). However, nightmares had statistical significance on the codominant and recessive heritage model for *CYP2B6* c.516G>T (OR = 12.00, $p = 0.031$ and OR = 7.14, $p = 0.042$, respectively). No significance was observed in the dominant model (**Table 5**).

On the other hand, nightmare, insomnia, anxiety, and suicide attempt were grouped as CNS ADRs. The analysis of univariate logistic regression showed statistical significance on codominant and recessive heritage models with *CYP2B6* c.516G>T (OR = 30.00, $p = 0.011$ and OR = 14.99, $p = 0.021$, respectively). No significance was observed in the dominant model (**Table 6**).

DISCUSSION

We report here that the most frequently observed ADRs were hyperbilirubinemia, and gastrointestinal and CNS alterations (26.04, 18.75, and 15.63%, respectively). This pattern of ADRs is similar to that observed in a study carried out in 2011 in the same hospital; precisely, ADRs such as hyperbilirubinemia, and gastrointestinal and CNS alterations were the most

TABLE 6 | Univariate logistic regression between *CYP2B6* c.516G>T (rs3745274) and CNS toxicity (nightmares, insomnia, anxiety, and suicide attempt grouped) in EFV-treated patients.

Inheritance model	Genotypes	OR	*p*-value	95% CI
Codominant model	G/G	1.00	Ref	-
	G/T	3.00	0.229	0.50–17.95
	T/T	30.00	0.011*	2.19–410.99
Recessive model	G/G + G/T	1.00	Ref	-
	T/T	14.99	0.021*	1.52–148.31
Dominant model	G/G	1.00	Ref	-
	G/T + T/T	5.08	0.061	0.93–27.75

*OR, odds ratio; CI, confidence interval; ref, reference; *p-value < 0.05.*

frequent (19.7, 21.1, and 13.2%, respectively) (Bernal Ortiz et al., 2013).

Interestingly, in our study, we observed that some ADRs occur in a cross-sectional manner in patients on antiretroviral therapy such as gastrointestinal upset and headache, and other ADRs are specific for a determinate group of patients; it is very relevant that certain ADRs are significantly associated with a specific therapy such as CNS ADR-EFV and hyperbilirubinemia-ATV. As statistically it is demonstrated that CNS ADRs are practically specific in patients treated with EFV, pharmacogenetic analysis was performed, and it was demonstrated that the presence of *CYP2B6* c.516G>T (rs3745274) is significantly associated with the development of toxicity at the CNS level. The same situation was observed with ATV where we demonstrated a statistical significant association between moderate and severe hyperbilirubinemia (total bilirubin level >1.9 mg/dl) and *UGT1A1*28*. Conversely, despite gastrointestinal ADRs being common within antiretroviral therapies (Bernal Ortiz et al., 2013), when we performed univariate logistic regression analysis in the dominant, codominant, and recessive model for *CYP2B6* c.516G>T (rs3745274) and *UGT1A1*28*, we were not able to visualize a statistically significant relationship. Similarly, no association was observed among headache and fever, and a given genotype of variants studied.

On the other hand, adverse reactions at the CNS level, analyzed as a whole or separately, as in the case of nightmares, are widely documented, although there is controversy regarding their appearance (10–74%) and the discontinuation of therapies (2–11%) due to these events (Kenedi and Goforth, 2011; Rosenblatt et al., 2017). In the univariate model of codominant inheritance for the nightmare event, the association with T/T genotype of *CYP2B6* c.516G>T (rs3745274) was observed with efavirenz (OR = 12, $p = 0.031$). Similarly, when we performed the univariate analysis for the recessive inheritance model, an association with the T/T genotype was observed (OR = 7.14, $p = 0.042$) but not with the dominant inheritance model (OR = 3.6, $p = 0.142$), possibly attributable to the low number of recruited patients.

Despite the fact that some ADRs such as insomnia, anxiety, and suicide attempt failed to have a statistically significant association in the univariate statistical models, when grouping them, including nightmares, as CNS ADRs, associations with the *CYP2B6* c.516G>T (rs3745274) T/T genotype and the toxicity, both in the codominant and recessive models, were observed (OR = 30.00, $p = 0.011$; OR = 14.99, $p = 0.021$, respectively), but not in the dominant model (OR = 5.08, $p = 0.061$).

In the present research, we observe a frequency of 0.36 for *UGT1A1*28*, which is comparable to the 0.33 reported in a previous study on the prevalence of Gilbert's syndrome in the Chilean population, and where the presence of the *UGT1A1*28* allele was determined indirectly by analyzing the rs6742078 (G > T) variant located in intron 1 of the *UGT1A1* gene, which is also in strong linkage disequilibrium with *UGT1A1*28* (Méndez et al., 2013). These frequencies are very similar to those described in other Latin American (Colombia: 0.34; Peru: 0.45; Mexico: 0.37) and European (0.30; Project 1000 genomes) populations, but different from Asian populations, where allele frequencies of 0.16 have been reported (Atasilp et al., 2020). The genotype frequencies observed in this study for *CYP2B6* c.516G>T (rs3745274) (G/G: 38.81%; G/T: 46.27%; T/T: 14.92%, **Table 3**) were very similar to those of a previous study carried out in a Chilean population under therapy with EFV (G/G: 43%; G/T: 42%; T/T: 15%) (Carr et al., 2010), where the frequency of allele T (0.38) was very similar to that of other Latin American populations such as Colombian (0.37), Mexicans (0.31), and Puerto Rican (0.35), and higher than that described in Europeans (0.24; Project 1000 genomes). Given the frequencies observed for *CYP2B6* c.516G>T and *UGT1A1*28* in Chile and Latin America, it is necessary to consider promoting studies that allow validating the recommendations of the CPIC guidelines for ATV (Gammal et al., 2016) and EFV (Desta et al., 2019) in our countries. These recommendations not only suggest the use of other alternative drugs, to avoid toxicity reactions in patients with genotypes associated with poor metabolizer phenotypes (e.g., homozygous genotype for *CYP2B6* c.516G>T), but also suggest lower initial doses of the same drug (e.g., starting efavirenz with a reduced dose of 400 or 200 mg/day; Desta et al., 2019).

Although this study reached a small sample size, statistically significant associations were observed between specific RAMs for ATV and EFV, with the variants *CYP2B6* c.516G>T and *UGT1A1*28*, respectively, demonstrating that these studies are necessary: a need that has been highlighted in previous national studies where the association of genetic variants in *CYP2B6* with elevated plasma levels of EFV was evidenced (Carr et al., 2010).

The present study has some shortcomings. The relatively small sample size of examined patients could mask potential associations. Moreover, some other potentially candidate genes/polymorphisms with a low level of evidence (e.g., *CYP3A4/5*, *CYP2A6*, and *SLCO1B1*) were not evaluated, but could be still relevant. Besides, some missing clinical values and possible differential misclassification bias could be relevant, affecting estimated associations between potentially relevant combinations of risk factors and adverse reactions.

To our knowledge, this is the first Latin American study performed to evaluate association between these genetic variants and ADRs in antiretroviral therapy, which allows us to initially characterize the therapeutic behavior in Chilean patients living with HIV.

CONCLUSION

Our findings suggest that after treatment with ATV or EFV, *UGT1A1*28* and *CYP2B6* c.516G>T (rs3745274) genetic polymorphisms influence the appearance of moderate-to-severe hyperbilirubinemia and CNS toxicity, respectively. Larger prospective studies will be necessary to validate these associations in our population; however, this preliminary study validates the recommendations of the CPIC guidelines to use the pharmacogenetic criteria in prescription of EFV or ATV to patients.

Without a doubt, improving adherence in patients living with HIV is a critical issue for the success of therapy. We must also rescue and optimize the use of drugs classified as "unsafe" whose prescription has decreased, given the type and frequency of adverse reactions they can generate. Many times, their application is strongly questioned despite being very effective. However, pharmacogenetics studies show that the generation of ADRs is not only an effect caused by the drug but rather the relationship with the patient. In this way, clinically, high effective drugs could be kept in force by following the recommendations of pharmacogenetic clinical guidelines by, for example, reducing the prescribed standard doses according to patients' genotype.

Hence, validating and applying international pharmacogenetics recommendations in Latin American countries would improve the precision of ART: a fundamental aspect to achieve the 90–90–90 treatment target proposed by UNAIDS.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study can be requested from the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Faculty of Medicine of the University of Chile. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Study conception: NV, PV, and FB; patient recruitment and informed consent procedure: FB, PV, and LC; obtaining clinical record data, preparation of CRFs, and tabulation of clinical data: GL, SA, NS, DP, and FB; obtaining gDNA and genotypic analysis: GL, SA, NS, and DP; statistical analysis: ML and DP; final analysis and generation of the manuscript: NV, LQ, DP, and FB.

FUNDING

This research was funded by Health Project 2016, Faculty of Medicine, the University of Chile (PI: NV) and National Research and Development Agency Scholarship No. 21160655 (DP).

ACKNOWLEDGMENTS

The authors wish to thank patients living with HIV from the "San Juan de Dios" Hospital for their altruistic collaboration in pursuit of the common welfare. We also thank the Latin American Society of Pharmacogenomics and Personalized Medicine (SOLFAGEM) by sponsoring this article.

REFERENCES

Atasilp, C., Chansriwong, P., Sirachainan, E., Reungwetwattana, T., Sirilerttrakul, S., Chamnanphon, M., et al. (2020). Effect of Drug Metabolizing Enzymes and Transporters in Thai Colorectal Cancer Patients Treated with Irinotecan-Based Chemotherapy. *Scientific Rep.* 10 (1), 1–10. doi:10.1038/s41598-020-70351-0

Bernal Ortiz, F., Vásquez, P., Giadalah, C., Rodríguez, L., and Villagrán, A. (2013). Incidencia de reacciones adversas a medicamentos en pacientes que inician o cambian terapia anti-retroviral. *Revista Chilena de Infectologia* 30 (5), 507–512. doi:10.4067/s0716-10182013000500007

Bosma, P. J., Chowdhury, J. R., Bakker, C., Gantla, S., de Boer, A., Oostra, B. A., et al. (1995). The Genetic Basis of the Reduced Expression of Bilirubin UDP-Glucuronosyltransferase 1 in Gilbert's Syndrome. *New Engl. J. Med.* 333 (18), 1171–1175. doi:10.1056/nejm199511023331802

Bosma, P. J., Seppen, J., Goldhoorn, B., Bakker, C., Oude Elferink, R. P. J., Chowdhury, J. R., et al. (1994). Bilirubin UDP-Glucuronosyltransferase 1 Is the Only Relevant Bilirubin Glucuronidating Isoform in Man. *J. Biol. Chem.* 269 (27), 17960–17964. doi:10.1016/s0021-9258(17)32403-1

Carr, D. F., la Porte, C., Pirmohamed, M., Owen, A., and Cortes, C. P. (2010). Haplotype Structure of CYP2B6 and Association with Plasma Efavirenz Concentrations in a Chilean HIV Cohort. *J. Antimicrob. Chemother.* 65 (9), 1889–1893. doi:10.1093/jac/dkq260

Caudle, K., Klein, T., Hoffman, J., Muller, D., Whirl-Carrillo, M., Gong, L., et al. (2014). Incorporation of Pharmacogenomics into Routine Clinical Practice: the Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline Development Process. *Curr. Drug Metab.* 15 (2), 209–217. doi:10.2174/1389200215666140130124910

Court, M. H., Zhang, X., Ding, X., Yee, K. K., Hesse, L. M., and Finel, M. (2012). Quantitative Distribution of mRNAs Encoding the 19 Human UDP-Glucuronosyltransferase Enzymes in 26 Adult and 3 Fetal Tissues. *Xenobiotica* 42 (3), 266–277. doi:10.3109/00498254.2011.618954

Desta, Z., Gammal, R. S., Gong, L., Whirl-Carrillo, M., Gaur, A. H., Sukasem, C., et al. (2019). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2B6 and Efavirenz-Containing Antiretroviral Therapy. *Clin. Pharmacol. Ther.* 106 (4), 726–733. doi:10.1002/cpt.1477

Fellay, J., Boubaker, K., and Ledergerber, B., (2001). Prevalence of Adverse Events Associated with Potent Antiretroviral Treatment: Swiss HIV Cohort Study. *The Lancet* 358 (9290), 1322–1327. doi:10.1016/s0140-6736(01)06413-3

FrankPalella, J. J. M., Maria Deloria-Knoll, P., JoanChmiel, S. P., AnneMoorman, Bsn, C. M., KathleenWood, C. B., Alan, E., et al. (2003). Article Survival Benefit of Initiating Antiretroviral Therapy in HIV-Infected. *Ann. Intern. Med.* 128, 620–626. doi:10.7326/0003-4819-138-8-200304150-00007

Gaida, R., Truter, I., and Grobler, C. (2016). Incidence of Neuropsychiatric Side Effects of Efavirenz in HIV-Positive Treatment-Naïve Patients in Public-Sector Clinics in the Eastern Cape. *South. Afr. J. HIV Med.* 17 (1), 1–6. doi:10.4102/hivmed.v17i1.452

Gammal, R. S., Court, M. H., Haidar, C. E., Iwuchukwu, O. F., Gaur, A. H., Alvarellos, M., et al. (2016). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for UGT1A1 and Atazanavir Prescribing. *Clin. Pharmacol. Ther.* 99 (4), 363–369. doi:10.1002/cpt.269

Gras, L., Kesselring, A. M., Griffin, J. T., Van Sighem, A. I., Fraser, C., Ghani, A. C., et al. (2007). CD4 Cell Counts of 800 Cells/mm3 or Greater after 7 Years of

Highly Active Antiretroviral Therapy Are Feasible in Most Patients Starting with 350 Cells/mm3 or Greater. *J. Acquired Immune Deficiency Syndromes* 45 (2), 183–192. doi:10.1097/qai.0b013e31804d685b

Hofmann, M. H., Blievernich, J. K., Klein, K., Saussele, T., Schaeffeler, E., Schwab, M., et al. (2008). Aberrant Splicing Caused by Single Nucleotide Polymorphism c.516G>T [Q172H], a Marker of CYP2B6*6, Is Responsible for Decreased Expression and Activity of CYP2B6 in Liver. *J. Pharmacol. Exp. Ther.* 325 (1), 284–292. doi:10.1124/jpet.107.133306

Hongkaew, Y., Medhasi, S., Pasomsub, E., Ngamsamut, N., Puangpetch, A., Vanwong, N., et al. (2018). UGT1A1 Polymorphisms Associated with Prolactin Response in Risperidone-Treated Children and Adolescents with Autism Spectrum Disorder. *Pharmacogenomics J.* 18 (6), 740–748. doi:10.1038/s41397-018-0031-7

Hsieh, T. Y., Shiu, T. Y., Huang, S. M., Lin, H. H., Lee, T. C., Chen, P. J., et al. (2007). Molecular Pathogenesis of Gilbert's Syndrome: Decreased TATA-Binding Protein Binding Affinity of UGT1A1 Gene Promoter. *Pharmacogenetics and Genomics* 17 (4), 229–236. doi:10.1097/fpc.0b013e328012d0da

Kadakol, A., Ghosh, S. S., Sappal, B. S., Sharma, G., Chowdhury, J. R., and Chowdhury, N. R. (2000). Genetic Lesions of Bilirubin Uridine-Diphosphoglucuronate Glucuronosyltransferase (UGT1A1) Causing Crigler-Najjar and Gilbert Syndromes: Correlation of Genotype to Phenotype. *Hum. Mutat.* 16 (4), 297–306. doi:10.1002/1098-1004(200010)16:4<297::aid-humu2>3.0.co;2-z

Kenedi, C. A., and Goforth, H. W. (2011). A Systematic Review of the Psychiatric Side-Effects of Efavirenz. *AIDS Behav.* 15 (8), 1803–1818. doi:10.1007/s10461-011-9939-5

Knobel, H., Alonso, J., Casado, J. L., Collazos, J., González, J., Ruiz, I., et al. (2002). Validation of a Simplified Medication Adherence Questionnaire in a Large Cohort of HIV-Infected Patients: The GEEMA Study. *Aids* 16 (4), 605–613. doi:10.1097/00002030-200203080-00012

Manosuthi, W., Sukasem, C., Lueangniyomkul, A., Mankatitham, W., Thongyen, S., Nilkamhang, S, et al. (2014). CYP2B6 Haplotype and Biological Factors Responsible for Hepatotoxicity in HIV-Infected Patients Receiving Efavirenz-Based Antiretroviral Therapy. *Int. J. Antimicrob. Agents* 43 (3), 292–296. doi:10.1016/j.ijantimicag.2013.10.022

Méndez, L., Lagos, M., Quiroga, T., Margozzini, P., Azócar, L., Molina, H. R., et al. (2013). Prevalencia de síndrome de Gilbert y sus determinantes genéticas en población chilena. *Revista Med. de Chile* 141 (10), 1266–1274. doi:10.4067/s0034-98872013001000005

MINSAL (2013). Guía Clínica AUGE Síndrome de la Inmunodeficiencia adquirida VIH/SIDA. del Gobierno de Chile: Ministerio de Salud (MINSAL) 1, 1–214. Available at: https://www.minsal.cl/sites/default/files/files/GPCVIH.pdf

Moore, R. D., Keruly, J. C., Gebo, K. A., and Lucas, G. M. (2005). An Improvement in Virologic Response to Highly Active Antiretroviral Therapy in Clinical Practice from 1996 through 2002. *J. Acquirimmune Defic Syndr.* 39 (39), 195–198.

Onusida (2019). Country Factsheets Chile. Available at https://www.unaids.org/es/regionscountries/countries/chile.19.

Onusida (2014). Fast Track. Ending The AIDS by 2030. 1–40. Available at https://www.unaids.org/sites/default/files/media_asset/JC2686_WAD2014report_es.pdf.

Quiñones, L., Roco, Á., Cayún, J. P., Escalante, P., Miranda, C., Varela, N., et al. (2017). *Farmacogenómica como herramienta fundamental para la medicina personalizada: aplicaciones en la práctica clínica*, 483–500. doi:10.4067/s0034-98872017000400009

Roden, D. M., and George, A. L. (2002). The Genetic Basis of Variability in Drug Responses. *Nat. Rev. Drug Discov.* 1 (1), 37–44. doi:10.1038/nrd705

Rosenblatt, L., Broder, M. S., Bentley, T. G. K., Chang, E., Reddy, S. R., Papoyan, E., et al. (2017). Tolerability of Central Nervous System Symptoms Among HIV-1 Infected Efavirenz Users: Analysis of Patient Electronic Medical Record Data. *AIDS Care - Psychol. Socio-Medical Aspects AIDS/HIV* 29 (8), 1067–1073. doi:10.1080/09540121.2017.1280123

Strassburg, C. P. (2008). Pharmacogenetics of Gilbert's Syndrome. *Pharmacogenomics* 9, 703–715. doi:10.2217/14622416.9.6.703

Sukasem, C., Atasilp, C., Chansriwong, P., Chamnanphon, M., Puangpetch, A., and Sirachainan, E. (2016). Development of Pyrosequencing Method for Detection of UGT1A1 Polymorphisms in Thai Colorectal Cancers. *J. Clin. Lab. Anal.* 30 (1), 84–89. doi:10.1002/jcla.21820

Sukasem, C., Chamnanphon, M., Koomdee, N., Puangpetch, A., Santon, S., Jantararoungtong, T., et al. (2013). High Plasma Efavirenz Concentration and CYP2B6 Polymorphisms in Thai HIV-1 Infections. *Drug Metab. Pharmacokinet.* 28 (5), 391–397. doi:10.2133/dmpk.dmpk-12-rg-120

Zhang, D., Chando, T. J., Everett, D. W., Patten, C. J., Dehal, S. S., and Griffith Humphreys, W. (2005). *In vitro* inhibition of Udp Glucuronosyltransferases by Atazanavir and Other HIV Protease Inhibitors and the Relationship of This Property to In Vivo Bilirubin Glucuronidation. *Drug Metab. Disposition* 33 (11), 1729–1739. doi:10.1124/dmd.105.005447

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 21 May 2021
doi: 10.3389/fphar.2021.674252

MicroRNA-33b is a Potential Non-Invasive Biomarker for Response to Atorvastatin Treatment in Chilean Subjects With Hypercholesterolemia: A Pilot Study

*Carmen Gloria Ubilla[1], Yalena Prado[1], Jeremy Angulo[1], Ignacio Obreque[1], Isis Paez[1], Nicolás Saavedra[1], Kathleen Saavedra[1], Tomás Zambrano[2] and Luis A. Salazar[1]**

[1]Center of Molecular Biology and Pharmacogenetics, Department of Basic Sciences, Faculty of Medicine, Universidad de La Frontera, Temuco, Chile, [2]Department of Medical Technology, Faculty of Medicine, Universidad de Chile, Santiago, Chile

OPEN ACCESS

Edited by:
Maria Ana Redal, University of Buenos Aires, Argentina

Reviewed by:
Claudia Pisanu, University of Cagliari, Italy
Peter Igaz, Semmelweis University, Hungary

***Correspondence:**
Luis A. Salazar
luis.salazar@ufrontera.cl

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *01 March 2021*
Accepted: *06 May 2021*
Published: *21 May 2021*

Citation:
Ubilla CG, Prado Y, Angulo J, Obreque I, Paez I, Saavedra N, Saavedra K, Zambrano T and Salazar LA (2021) MicroRNA-33b is a Potential Non-Invasive Biomarker for Response to Atorvastatin Treatment in Chilean Subjects With Hypercholesterolemia: A Pilot Study. Front. Pharmacol. 12:674252. doi: 10.3389/fphar.2021.674252

Evidence accumulated so far indicates that circulating levels of microRNAs (miRNAs) are associated with several pathologies. Therefore, differential expression of extracellular miRNAs exhibits promising potential for screening and diagnosis purposes. We evaluated plasma miRNAs in response to the lipid-lowering drug atorvastatin in patients with hypercholesterolemia (HC) and controls. Methods: We selected miRNAs based on previous data reported by our group and also by employing bioinformatics tools to identify 10 miRNAs related to cholesterol metabolism and statin response genes. Following miRNA identification, we determined plasma levels of miRNA-17-5p, miRNA-30c-5p, miRNA-24-3p, miRNA-33a-5p, miRNA-33b-5p, miRNA-29a-3p, miRNA-29b-3p, miRNA-454-3p, miRNA-590-3p and miRNA-27a-3p in 20 HC patients before and after 1 month of 20 mg/day atorvastatin treatment, evaluating the same miRNA set in a group of 20 healthy subjects, and employing qRT-PCR to determine differential miRNAs expression. Results: HC individuals showed significant overexpression of miRNA-30c-5p and miRNA-29b-3p vs. NL ($p = 0.0008$ and $p = 0.0001$, respectively). Once cholesterol-lowering treatment was concluded, HC individuals showed a substantial increase of three extracellular miRNAs (miRNA-24-3p, miRNA-590, and miRNA-33b-5p), the latter elevated more than 37-fold ($p = 0.0082$). Conclusion: Data suggest that circulating miRNA-30c-5p and miRNA-29b-3p are associated with hypercholesterolemia. Also, atorvastatin induces a strong elevation of miRNA-33b-5p levels in HC individuals, which could indicate an important function that this miRNA may exert upon atorvastatin therapy. Additional studies are needed to clarify the role of this particular miRNA in statin treatment.

Keywords: hypercholesterolemia, circulating microRNAs, atorvastatin, epidrugs, statins

INTRODUCTION

Ever since the discovery that miRNAs can circulate in the extracellular medium with extraordinary stability, there has been substantial interest in identifying these molecules as non-invasive biomarkers for several problematic pathologies to public health, such as cardiovascular disease. Advancements made so far show that miRNAs are not only expressed differently between a varied number of conditions, but their ability to be measured in diverse bodily fluids such as serum or plasma, urine, tears, breast milk, amniotic fluid, cerebrospinal fluid, saliva, and semen show promise for screening, diagnosis, prognosis, and follow-up purposes.

It is well-established that abnormal lipid accumulation sets an atherogenic milieu (Summerhill and et al., 2019), and to date, numerous miRNAs have been reported to be involved in lipid metabolism (Fernández-Hernando et al., 2011; Aryal et al., 2017). We previously showed that miRNAs expression is altered *in vitro* (Zambrano et al., 2015) and *in vivo* (Zambrano et al., 2018) following statin treatment, a lipid-lowering drug of widespread use that also reduces the cardiovascular risk (Garcia-Gil et al., 2018). However, the improvement in lipid profiles is usually underachieved, as almost half of the patients undergoing statins will show a suboptimal response to these drugs (Gitt et al., 2012; Akyea et al., 2019). Moreover, very little is known about the potential use of miRNAs as biomarkers of poor response to statin therapy, which would ultimately help detect patients not meeting appropriate LDL-C reduction goals. Therefore, we carried out bioinformatics and relative expression studies to determine circulating miRNAs in healthy subjects and patients with hypercholesterolemia undergoing treatment with 20 mg/day atorvastatin for one month, to identify candidate miRNAs as potential therapeutic targets.

MATERIALS AND METHODS

Subjects

A total of forty individuals were selected for this study. Twenty subjects were normolipidemic and comprised the control group, whereas 20 were hypercholesterolemic patients treated with 20 mg/day atorvastatin for 1 month. Biochemical profiles were performed on all patients, and the treated group was sampled before and after completion of the lipid-lowering treatment. The patients were recruited from public health centers. Blood samples were then processed to separate the plasma, which was frozen for further analysis. The Ethics Committee of University of La Frontera (Protocol #045_17) approved the study protocol. All subjects gave their written informed consent to participate from this investigation.

Biochemical Analyses

To establish plasma lipid levels before and after atorvastatin treatment, blood collection was performed by direct venous puncture following an overnight fast using EDTA tubes. Total cholesterol, HDL cholesterol (HDL-C) and triglycerides (TG) were measured by routine enzymatic-colorimetric methods. LDL cholesterol (LDL-C) was calculated using the Friedewald equation when TG levels were not above 400 mg/dl.

Selection of miRNAs

As aforementioned, we based the current selection of ten circulating miRNAs on previous data obtained by our group (Zambrano et al., 2015; Zambrano et al., 2018). Briefly, in the first report, we identified deregulated miRNAs by analyzing their expression profile in HepG2 cells treated during 24 h with different statins. In the second study, we employed multiple bioinformatic tools such as TargetScan, miRanda, DianaLab, MicroCosm, and PicTar to narrow down the previous *in vitro* miRNA data, evaluating a subset of 84 miRNAs potentially associated with 28 key genes involved in cholesterol metabolism and statin response in subjects undergoing 1 month of different low-dose statins. In the present study, we crossed our previous results and updated the selection criteria by running Pharmaco miR, a miRNA pharmacogenomics database that identifies associations of miRNAs and both the genes they regulate and the drugs annotated, to finally select the ten most probably dysregulated miRNAs in patients with hypercholesterolemia following a 20 mg/day atorvastatin dose for 4 weeks and controls.

RNA Extraction

Circulating miRNAs were extracted using the miRNeasy serum/plasma kit (Quiagen, Hilden, Germany) according to the manufacturer's instruction, using 200 µl of plasma. miR-39 of *C. elegans* was incorporated as a spike-in control to normalize quantification (1.6×10^8 copies/µl). RNA concentration and purity were determined by spectrophotometry (Infinite® 200 PRO NanoQuant) using the 260 nm/280 nm absorbance ratio. The samples were diluted to a concentration of 5 ng/µL and stored at −80°C for ulterior use.

Reverse Transcription

The cDNA synthesis was performed using the TaqMan™ Advanced miRNA cDNA synthesis kit. The protocol carried out consisted of four steps: 1) polyadenylation, 2) ligation of adapters, 3) retrotransciption and 4) miRNAs obtention. Afterward, miRNAs obtained were stored at −20°C for later use.

qPCR Expression Analyses

Each sample was diluted to 1:10 with ultrapure water. The PCR Mix was prepared in sufficient quantity for each assay, for the identification of miRNA-17-5p, miRNA-30c-5p, miRNA-24p-3p, miRNA-33a-5p, miRNA-33b-5p, miRNA-29a-3p, miRNA-29b-3p, miRNA-454-3p, miRNA-590-3p and miRNA-27-3p following the manufacturer's protocol.

Interaction Network of Differentially Expressed miRNAs

To generate the interaction network of selected miRNAs, we employed miRTargetLink 2.0 (Kern and et al., 2021), a tool containing experimentally validated interactions on human microRNA-mRNA pairs. Data shown correspond to miRNA-

TABLE 1 | Clinical, demographics and plasma lipid levels before and after treatment with atorvastatin (20 mg/day/4 weeks) of normolipidemic (NL) and hypercholesterolemic (HC) individuals.

Parameter	NL	HC	HC post-treatment	% Change	*p*-value
Female/Male	15/5	15/5	—	—	-—
Age (years)	31.2 ± 7.3	47.3 ± 11.3	—	—	—
BMI (kg/m^2)	26.4 ± 4.4	27.0 ± 3.0	—	—	—
Total cholesterol (mg/dl)	151.3 ± 28.4	239.3 ± 28.2	158.1 ± 33.4	34.1 ± 10.7	<0.001
LDL-cholesterol (mg/dl)	87.5 ± 17.4	176.0 ± 15.5	96.1 ± 29.9	44.6 ± 14.0	<0.001
HDL-cholesterol (mg/dl)	50.6 ± 19.7	44.4 ± 10.0	41.2 ± 9.4	6.4 ± 14.9	0.056
Triglycerides (mg/dl)	80.2 ± 30.5	150.4 ± 66.2	121.6 ± 55.4	10.9 ± 37.8	<0.05
VLDL-cholesterol (mg/dl)	17.0 ± 6.5	27.0 ± 7.0	21.6 ± 6.2	15.0 ± 36.7	<0.05

Values expressed as mean ± standard deviation, n, number of subjects; BMI: body mass index, LDL-C: low-density lipoprotein cholesterol, HDL-C: High-density lipoprotein cholesterol, VLDL: Very low-density lipoprotein cholesterol.

target interactions with strong support, i.e., validated experimentally by reporter assay, western blot, qPCR, microarray, and/or next-generation sequencing experiments. The software obtains miRNAs annotations from the latest version of miRBase (v.22.1), while the experimentally validated targets are retrieved from miRTarBase (v.8) (Huang et al., 2020) and miRATBase (Kern et al., 2021). miRTargetLink 2.0 can be freely accessible from the following link (https://ccb-compute.cs.uni-saarland.de/mirtargetlink2/).

Statistical Analysis

Data were analyzed using GraphPad Prism version 5.00 for Windows (GraphPad Software, San Diego, CA, United States). Normal distribution was assessed using D'Agostino-Pearson testing. As data came from a limited set of ten miRNAs, correction for multiple testing was not employed, and differential miRNA expression was calculated using $2^{-\Delta CT}$ to compare normolipidemic and hypercholesterolemic subjects. The Mann-Whitney test was applied for unpaired data and the Wilcoxon test for paired data. A p-value <0.05 was considered statistically significant.

RESULTS

Clinical Data, Plasma Lipids and Statin Response

Table 1 shows anthropometric data and lipid profiles of normolipidemic (NL) and hypercholesterolemic (HC) patients before and after 20 mg/day of atorvastatin during a 4-weeks period. Significant differences were observed for all lipid levels except for HDL-C.

Expression of Circulating miRNAs Associated With Cholesterol Metabolism and Response to Statins

We found two extracellular miRNAs differentially expressed between normolipidemic (NL) and hypercholesterolemic (HC) individuals. **Figure 1** shows the fold change of miRNA-30c and miRNA-29b in these subjects. The rest of the miRNAs evaluated did not show significant differences between the study groups.

Expression of Circulating miRNAs in Hypercholesterolemic (HC) Patients and Response to Atorvastatin Treatment

According to the expression of extracellular miRNAs in patients at the beginning and post-treatment, significant differences were observed for miRNA-33-5p, miRNA-24-3p and miRNA-590. **Table 2** illustrates the relative expression of miRNAs before and after 1 month with atorvastatin (20 mg/day).

Correlation of miRNAs and Lipid Reduction

Table 3 shows the correlation between basal circulating miRNAs and the reduction percentage (%) of LDL-C following 1 month of treatment with a 20 mg/day atorvastatin dose. All miRNAs showed positive correlation with LDL-C reductions, but only five (miRNA-30c, miRNA-29a, miRNA-454, miRNA-24-3p and miRNA-590) were statistically significant. **Supplementary Figure S1** displays the correlation of circulating miRNAs and LDL-C reduction.

Predicted Interactions of Deregulated miRNAs

We found five deregulated miRNAs, two of them were overexpressed in hypercholesterolemic vs. normal subjects (miRNA-30c-5p and miRNA-29b-3p), while three were significantly upregulated in HC patients after concluding atorvastatin treatment (miRNA-24-3p, miRNA-590, and miRNA-33b-5p). Importantly, miRNA-33b-5p showed a substantial 37-fold increase since treatment initiation. **Figure 2** shows a network of validated mRNA targets for miRNA-33b-5p. Additional miRNA-mRNA pairs interaction networks from the remaining deregulated miRNAs are presented as supplementary material (**Supplementary Figures S2–S5**).

DISCUSSION

The LDL-C lowering efficacy for a 20 mg atorvastatin dose is reported to be around 42% (Adams et al., 2015). In our study, half of the patients did not achieve this goal, and a quarter showed

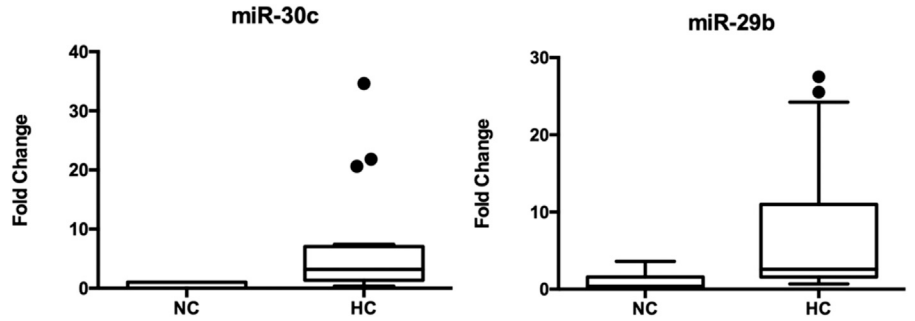


FIGURE 1 | Expression of circulating miRNA-30c and miRNA-29b in hypercholesterolemic patients vs. controls. *Mann-Whitney test. *C. elegans* miRNA-39 was used as a normalizing control. The values of the threshold cycle (Ct) of each miRNA, normalized against this sequence (ΔCt = Ct target miRNA—Ct endogenous control), the plasma expression levels of these miRNAs were calculated according to $2^{-\Delta Ct}$ method. The total number of samples analyzed was 20 for each study group.

TABLE 2 | Circulating miRNA levels in hypercholesterolemic (HC) subjects before and following 20 mg/day/4 weeks atorvastatin treatment.

miRNA	HC basal	HC post-treatment	p-value*
miR-17-5p	1 ± 1.583	4.727 ± 8.082	0.1297
miR-30c	1 ± 1.314	0.789 ± 1.140	0.7983
miR-33a-5p	1 ± 1.408	4.230 ± 9.728	0.2413
miR-33b-5p	1 ± 1.644	37.479 ± 111.932	0.0082
miR-29a	1 ± 1.466	3.057 ± 7.289	0.4900
miR-29b	1 ± 1.189	3.031 ± 4.002	0.0799
miR-454	1 ± 1.389	2.750 ± 3.968	0.2253
miR-24-3p	1 ± 1.271	4.815 ± 6.527	0.0494
miR-590	1 ± 1.791	9.207 ± 12.656	0.0047
miR-27a-3p	1 ± 2.239	2.161 ± 3.597	0.1336

**Wilcoxon matched-pairs signed rank test. The plasma expression levels of these miRNAs were calculated according to 2 - $^{\Delta\Delta Ct}$ [ΔΔCt = Ct Basal (target miRNA—Ct endogenous control)—Ct treatment (target miRNA—Ct endogenous control)].*

LDL-C reductions lower than 27%. This highly heterogeneous response to statins is well documented (Karlson et al., 2016) and is consistent with previous reports made by our group (Rosales et al., 2012). As for miRNAs, we observed an up-regulation of two miRNAs, i.e., miRNA-30c and miRNA-29b, between patients having abnormally high cholesterol levels vs. healthy subjects. Previously, Soh and colleagues reported that miRNA-30c reduce lipid synthesis and LDL-precursors due to its interaction with the 3-UTR of the microsomal triglyceride transfer protein (MTP) (Soh et al., 2013), a protein central for lipoprotein assembly and production of LDL precursors. Hence, miRNA-30c contributes to control hepatic and plasma lipids by decreasing hyperlipidemia, which is a likely explanation of why this miRNA is elevated in hypercholesterolemic individuals. Moreover, Sodi et al. reported a significant and positive correlation between miRNA-30c with total cholesterol (TC) and LDL-C (Sodi et al., 2017), which further demonstrates an important regulatory role for this miRNA in lipid homeostasis. Authors also showed that pravastatin but not rosuvastatin increased serum miRNA-30c, which points toward a dissimilar effect of statins on miRNAs expression, even though both statins assessed share a hydrophilic nature. One reason supporting this differential regulation could be the differing treatment periods (1-year pravastatin vs. 8-weeks rosuvastatin). Likewise, we have previously reported that different statins produce variable effects on miRNA expression (Zambrano et al., 2015; Zambrano et al., 2018). This distinctive effect has also been documented at different levels of gene regulation (Leszczynska et al., 2011). As far as we know, this is the first study disclosing the relationship between miRNA-30c and atorvastatin treatment.

The most considerable finding observed is a 37-fold increase of miRNA-33b, together with a significant elevation of circulating miRNA-590 and miRNA-24-3p levels in HC patients upon completion of atorvastatin treatment. The miRNA-33b is part of the miRNA-33 family, formed by miRNAs 33a and 33b. This pair represent one of the most widely characterized miRNA mediators involved in lipid homeostasis. Both are intronic miRNAs differing by 2-nucleotides, but unlike miRNA-33a, miRNA-33b is encoded within the sterol regulatory element-binding protein-1 (*SREBP1*), a transcription factor implicated in fatty acid metabolism. Existing data connecting miRNA-33b and statins are abundant. Reports show that miRNA-33b is co-transcribed along *SREBP1*, working coordinately to control lipid levels (Marquart et al., 2010; Najafi-Shoushtari et al., 2010; Rayner et al., 2010; Davalos et al., 2011). Regarding statins, investigations display mixt outcomes. In 2012, Takwi et al. showed an upregulation of miRNA-33b expression following lovastatin treatment of medulloblastoma cells (Takwi et al., 2012). 4 years later, Zhang and colleagues demonstrated that pitavastatin reverted the oxLDL-mediated suppression of miRNA-33b in human THP-1 cells (Zhang et al., 2016). Additional evidence supporting a relation between statins and miRNA-33 came from studying statin-naïve subjects with metabolic syndrome (MetS) (Chen et al., 2016). The authors showed that MetS subjects had significantly higher plasma values of miRNA-33 than their healthy controls counterparts. Afterward, MetS subjects treated with atorvastatin or pitavastatin experimented an additional increase in circulating miRNA-33 levels. Moreover, when studying the murine macrophage cell line RAW264.7 and bone marrow-derived macrophages (BMDM), authors found a dose-dependent up-regulation of miRNA-33 in both cell lysates and medium,

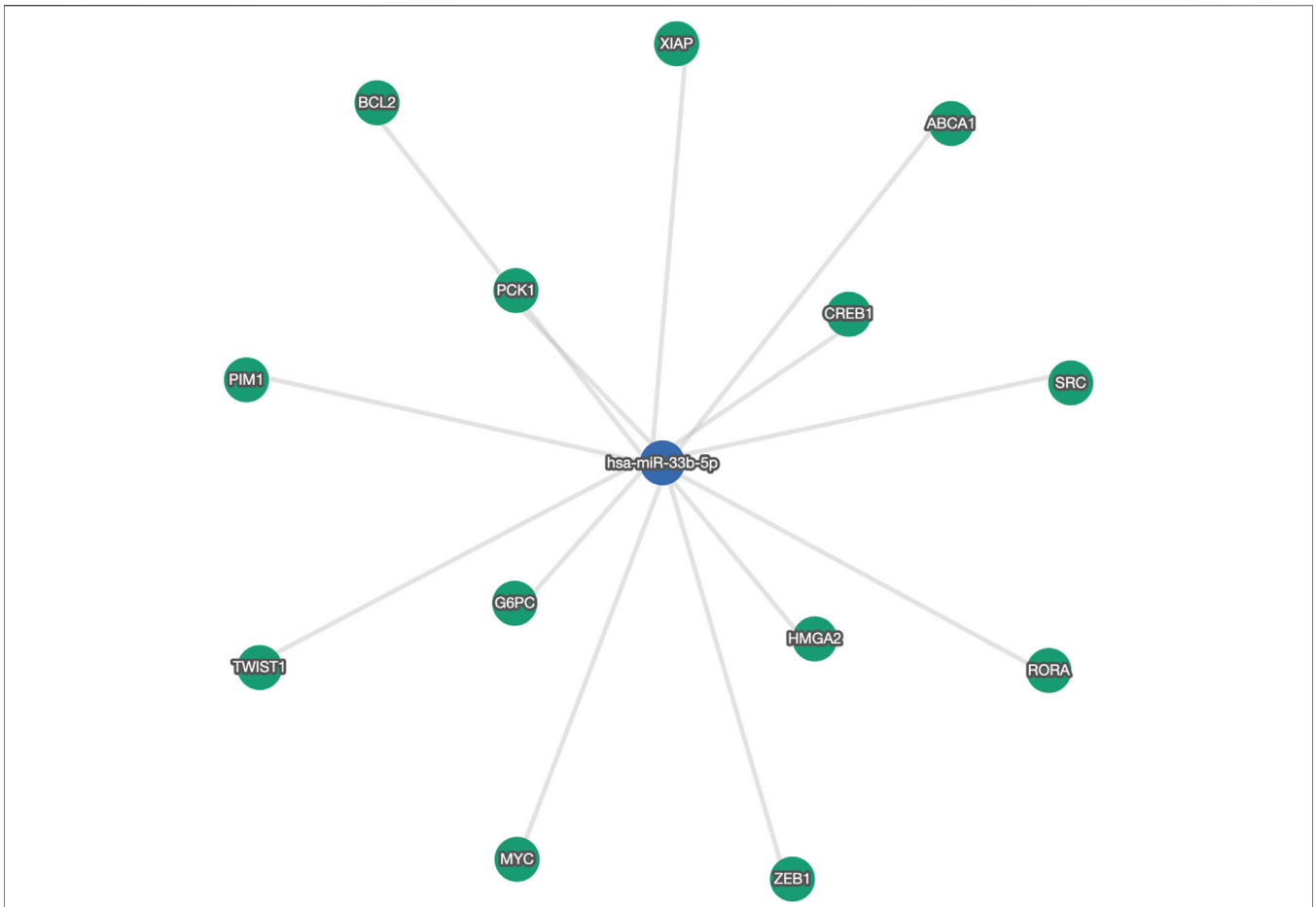


FIGURE 2 | Network of validated gene targets for miRNA-33b-5p. ***ABCA1***: ATP binding cassette subfamily A member one; ***BCL2***: B-cell lymphoma two Apoptosis Regulator; ***CREB1***: cAMP responsive element binding protein one; ***G6PC***: glucose-6-phosphatase, catalytic; ***HMGA2***: High Mobility Group AT-Hook two; ***MYC***: MYC Proto-Oncogene, BHLH Transcription Factor; ***PCK1***: Phosphoenolpyruvate carboxykinase one; ***PIM1***: Proto-Oncogene, Serine/Threonine Kinase; ***RORA***: RAR Related Orphan Receptor A; ***SRC***: SRC Proto-Oncogene, Non-Receptor Tyrosine Kinase; ***TWIST1***: Twist Family BHLH Transcription Factor 1; ***XIAP***: X-linked inhibitor of apoptosis; ***ZEB1***: zinc finger E-box binding homeobox 1.

regardless of the statin used. Our previous assessment of statin treatment on miRNAs from hepatocellular carcinoma cells (HepG2) revealed that neither miRNA-33a nor miRNA-33b was affected by low-dose atorvastatin or simvastatin (Zambrano et al., 2015). However, in peripheral blood mononuclear cells (PBMC) (Zambrano et al., 2018), we observed that atorvastatin but not simvastatin repressed the cellular miRNA-33b content in humans, representing a contrasting outcome from the MetS cohort. Nonetheless, several differences must be underlined, especially regarding the subjects investigated. First, in the study from (Chen et al., 2016), included individuals were defined by traditional MetS criteria, none of which considers high LDL-C levels, while we specifically included individuals having elevated LDL-C concentrations according to clinical characteristics from both studies (MetS: 155 mg/dl; PBMC: 182 mg/dl). Second, our cohort was treated during 1 month with a 10 mg/day statin dose after a thirty-day wash-out period, vs. the 3-months statin therapy for statin-naïve MetS individuals. In addition to the dissimilar ethnic background between Taiwanese and Brazilians, and the fact that we did not evaluate circulating miRNAs, another critical difference is represented by the cellular model employed. While we studied a variety of peripheral cells containing a single, rounded nucleus such as lymphocytes and monocytes i.e., PBMC, (Chen et al., 2016) specifically assessed macrophages. Even though macrophages are differentiated from monocytes, macrophages represent a very specialized and highly heterogeneous cellular type with a distinct phenotype and functionality than monocytes due to their particularly active role as one of the first lines of immune defense. Lastly, our study evaluated the miRNA-33 family separately, i.e., miRNA-33a and miRNA-33b, in contrast to what was displayed by (Chen et al., 2016), where they showed that the upregulation induced by atorvastatin affected miRNA-33 generally, making its specific impact on miRNA-33a or 33-b indistinguishable from one another.

On the other hand, the evidence surrounding the role that miRNA-590 and miRNA-24-3p portray in statin therapy is scarce. Studies confer miRNA-590 a function in lipid

TABLE 3 | Correlation analysis of miRNAs and the percentage of lipid reduction after statin therapy.

miRNA	Slope	R^2	*p*-value
miRNA-17-5p	0.393 ± 0.336	0.0703	0.2583
miRNA-30c	0.495 ± 0.172	0.313	0.0103*
miRNA-33a-5p	2.845 ± 1.484	0.169	0.0712
miRNA-33b-5p	9.434 ± 6.099	0.117	0.1393
miRNA-29a	1.765 ± 0.759	0.230	0.0321*
miRNA-29b	0.666 ± 0.786	0.038	0.4080
miRNA-454	1.763 ± 0.807	0.209	0.0425*
miRNA-24-3p	2.318 ± 1.047	0.205	0.0393*
miRNA-590	16.97 ± 7.239	0.244	0.0314*
miRNA-27a-3p	0.084 ± 0.075	0.065	0.2753

*Statistically significant.

homeostasis by inhibiting lipoprotein lipase (LPL), an enzyme that degrades circulating triglycerides, resulting in attenuated lipid accumulation in human THP-1 macrophages (He et al., 2014). Regarding statins, we previously showed that Brazilian patients not meeting their expected LDL-C reduction goal following 10 mg/day atorvastatin had a significant downregulation of this particular miRNA in PBMC (Zambrano et al., 2018). Conversely, the present study's data revealed a >9-fold increase in circulating levels of miRNA-590 following 20 mg/day atorvastatin. This conflicting miRNA behavior is not clear, but one key difference is that presently, we did not assess miRNA performance according to subgroups of LDL-C goals achievements. Besides, we evaluated extracellular rather than intracellular miRNAs. In the case of miRNA-24-3p, studies show that its expression is significantly increased in the livers of high-fat diet-treated mice (Ng et al., 2014). The same study revealed insulin-induced gene 1 (*Insig1*) -a lipogenesis inhibitor-as a validated target of miRNA-24. Therefore, elevated miRNA-24 levels decrease hepatic lipid accumulation *via Insig1* up-regulation. In the same way, Wang et al. (Wang et al., 2018) showed that obesity-induced miRNA-24 overexpression inhibited Scavenger Receptor B1 (*SR-B1*), a member of the CD36 family of scavenger receptors B that facilitates selective cholesterol uptake from high-density lipoproteins (HDL). Additionally, miRNA-24 increased the expression of important genes related to cholesterol synthesis, such as 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase (*HMGCR*) and sterol regulatory element-binding protein 2 (*SREBF2*). Therefore, miRNA-24 assists in regulating cholesterol homeostasis and steroidogenesis by repressing HDL uptake in HepG2 cells. However, its role regarding statin therapy remains to be clarified.

One of the main limitations of the present study is its small sample size, preventing us from reporting more robust conclusions. Even though it did not restrict the successful identification of extracellular miRNA deregulation, our preliminary data should be interpreted in light of the limited cohort evaluated. We emphasize the need for additional in-depth stratified analysis on larger population and functional studies in different models to clarify further the relationship between statins and miRNAs, as they represent promising candidates for therapeutic manipulation.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the **COMITÉ ÉTICO CIENTÍFICO** UNIVERSIDAD DE LA FRONTERA. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CU, JA, IO, IP and YP performed laboratory analysis. CU, NS and TZ contributed to data analysis. TZ, NS, KS, and LS contributed to the conception and design of the study. CU wrote the first draft of the manuscript. TZ revised the manuscript and wrote the final version. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This research was funded by Fondecyt—Chile (Grant number 1171765), Dirección de Investigación of the Universidad de La Frontera (Grant number DI19-2018), and ANID—Chile Doctoral Scholarship to IP.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.674252/full#supplementary-material

REFERENCES

Adams, S. P., Tsang, M., and Wright, J. M. (2015). Lipid-lowering Efficacy of Atorvastatin. *Cochrane Database Syst. Rev.* (3), CD008226. doi:10.1002/14651858.CD008226.pub3

Akyea, R. K., Kai, J., Qureshi, N., Iyen, B., and Weng, S. F. (2019). Sub-optimal Cholesterol Response to Initiation of Statins and Future Risk of Cardiovascular Disease. *Heart* 105 (13), 975–981. doi:10.1136/heartjnl-2018-314253

Aryal, B., Singh, A. K., Rotllan, N., Price, N., and Fernández-Hernando, C. (2017). MicroRNAs and Lipid Metabolism. *Curr. Opin. Lipidol.* 28 (3), 273–280. doi:10.1097/mol.0000000000000420

Chen, W. M., Sheu, W. H., Tseng, P. C., Lee, T. S., Lee, W. J., Chang, P. J., et al. (2016). Modulation of microRNA Expression in Subjects with Metabolic Syndrome and Decrease of Cholesterol Efflux from Macrophages via microRNA-33-Mediated Attenuation of ATP-Binding Cassette Transporter A1 Expression by Statins. *PLoS One* 11 (5), e0154672. doi:10.1371/journal.pone.0154672

Davalos, A., Goedeke, L., Smibert, P., Ramirez, C. M., Warrier, N. P., Andreo, U., et al. (2011). miR-33a/b Contribute to the Regulation of Fatty Acid Metabolism and Insulin Signaling. *Proc. Natl. Acad. Sci.* 108 (22), 9232–9237. doi:10.1073/pnas.1102281108

Fernández-Hernando, C., Suárez, Y., Rayner, K. J., and Moore, K. J. (2011). MicroRNAs in Lipid Metabolism. *Curr. Opin. Lipidol.* 22 (2), 86–92. doi:10.1097/mol.0b013e3283428d9d

Garcia-Gil, M., Comas-Cufí, M., Blanch, J., Martí, R., Ponjoan, A., Alves-Cabratosa, L., et al. (2018). Effectiveness of Statins as Primary Prevention in People with Different Cardiovascular Risk: A Population-Based Cohort Study. *Clin. Pharmacol. Ther.* 104 (4), 719–732. doi:10.1002/cpt.954

Gitt, A. K., Drexel, H., Feely, J., Ferrières, J., Gonzalez-Juanatey, J. R., Thomsen, K. K., et al. .(2012). Persistent Lipid Abnormalities in Statin-Treated Patients and Predictors of LDL-Cholesterol Goal Achievement in Clinical Practice in Europe and Canada. *Eur. J. Prev. Cardiolog.* 19 (2), 221–230. doi:10.1177/1741826711400545

He, P.-P., Ouyang, X.-P., Tang, Y.-Y., Liao, L., Wang, Z.-B., Lv, Y.-C., et al. (2014). MicroRNA-590 Attenuates Lipid Accumulation and Pro-inflammatory Cytokine Secretion by Targeting Lipoprotein Lipase Gene in Human THP-1 Macrophages. *Biochimie* 106, 81–90. doi:10.1016/j.biochi.2014.08.003

Huang, H. Y., Lin, Y. C. D., Li, Jing., Huang, K. Y., Shrestha, S., Hong, H. C., et al. 2020: miRTarBase 2020: Updates to the Experimentally Validated microRNA-Target Interaction Database. *Nucleic Acids Res.* 48(D1): p. D148–D154.doi:10.1093/nar/gkz896

Karlson, B. W., Wiklund, O., Palmer, M. K., Nicholls, S. J., Lundman, P., and Barter, P. J. (2016). Variability of Low-Density Lipoprotein Cholesterol Response with Different Doses of Atorvastatin, Rosuvastatin, and Simvastatin: Results from VOYAGER. *Eur. Heart J. Cardiovasc. Pharmacother.* 2 (4), 212–217. doi:10.1093/ehjcvp/pvw006

Kern, F., et al. (2021). miRTargetLink 2.0-interactive miRNA Target Gene and Target Pathways Networks. *Nucleic Acids Res.* 8.

Kern, F., Krammes, L., Danz, K., Diener, C., Kehl, T., Küchler, O., et al. (2021). Validation of Human microRNA Target Pathways Enables Evaluation of Target Prediction Tools. *Nucleic Acids Res.* 49 (1), 127–144. doi:10.1093/nar/gkaa1161

Leszczynska, A., Gora, M., Plochocka, D., Hoser, G., Szkopinska, A., Koblowska, M., et al. (2011). Different Statins Produce Highly Divergent Changes in Gene Expression Profiles of Human Hepatoma Cells: a Pilot Study. *Acta Biochim. Pol.* 58 (4), 635–639. doi:10.18388/abp.2011_2235

Marquart, T. J., Allen, R. M., Ory, D. S., and Baldan, A. (2010). miR-33 Links SREBP-2 Induction to Repression of Sterol Transporters. *Proc. Natl. Acad. Sci.* 107 (27), 12228–12232. doi:10.1073/pnas.1005191107

Najafi-Shoushtari, S. H., Kristo, F., Li, Y., Shioda, T., Cohen, D. E., Gerszten, R. E., et al. (2010). MicroRNA-33 and the SREBP Host Genes Cooperate to Control Cholesterol Homeostasis. *Science* 328 (5985), 1566–1569. doi:10.1126/science.1189123

Ng, R., Wu, H., Xiao, H., Chen, X., Willenbring, H., Steer, C. J., et al. (2014). Inhibition of microRNA-24 Expression in Liver Prevents Hepatic Lipid Accumulation and Hyperlipidemia. *Hepatology* 60 (2), 554–564. doi:10.1002/hep.27153

Rayner, K. J., Suarez, Y., Davalos, A., Parathath, S., Fitzgerald, M. L., Tamehiro, N., et al. (2010). MiR-33 Contributes to the Regulation of Cholesterol Homeostasis. *Science* 328 (5985), 1570–1573. doi:10.1126/science.1189862

Rosales, A., Alvear, M., Cuevas, A., Saavedra, N., Zambrano, T., and Salazar, L. A. (2012). Identification of Pharmacogenetic Predictors of Lipid-Lowering Response to Atorvastatin in Chilean Subjects with Hypercholesterolemia. *Clin. Chim. Acta* 413 (3-4), 495–501. doi:10.1016/j.cca.2011.11.003

Sodi, R., Eastwood, J., Caslake, M., Packard, C. J., and Denby, L. (2017). Relationship between Circulating microRNA-30c with Total- and LDL-Cholesterol, Their Circulatory Transportation and Effect of Statins. *Clinica Chim. Acta* 466, 13–19. doi:10.1016/j.cca.2016.12.031

Soh, J., Iqbal, J., Queiroz, J., Fernandez-Hernando, C., and Hussain, M. M. (2013). MicroRNA-30c Reduces Hyperlipidemia and Atherosclerosis in Mice by Decreasing Lipid Synthesis and Lipoprotein Secretion. *Nat. Med.* 19 (7), 892–900. doi:10.1038/nm.3200

Summerhill, V. I., Grechko, A. V., Yet, S. F., Sobenin, I. A., and Orekhov, A. N. (2019). The Atherogenic Role of Circulating Modified Lipids in Atherosclerosis. *Int. J. Mol. Sci.* 20 (14). doi:10.3390/ijms20143561

Takwi, A. A. L., Li, Y., Becker Buscaglia, L. E., Zhang, J., Choudhury, S., Park, A. K., et al. (2012). A Statin-regulated microRNA Represses Human c-Myc Expression and Function. *EMBO Mol. Med.* 4 (9), 896–909. doi:10.1002/emmm.201101045

Wang, M., Li, L., Liu, R., Song, Y., Zhang, X., Niu, W., et al. (2018). Obesity-induced Overexpression of miRNA-24 Regulates Cholesterol Uptake and Lipid Metabolism by Targeting SR-B1. *Gene* 668, 196–203. doi:10.1016/j.gene.2018.05.072

Zambrano, T., Hirata, R. D. C., Hirata, M. H., Cerda, Á., and Salazar, L. A. (2015). Altered microRNome Profiling in Statin-Induced HepG2 Cells: A Pilot Study Identifying Potential New Biomarkers Involved in Lipid-Lowering Treatment. *Cardiovasc. Drugs Ther.* 29 (6), 509–518. doi:10.1007/s10557-015-6627-0

Zambrano, T., Hirata, R. D. C., Hirata, M. H., Cerda, Á., and Salazar, L. A. (2018). Statins Differentially Modulate microRNAs Expression in Peripheral Cells of Hyperlipidemic Subjects: A Pilot Study. *Eur. J. Pharm. Sci.* 117, 55–61. doi:10.1016/j.ejps.2018.02.007

Zhang, H., Lamon, B. D., Moran, G., Sun, T., Gotto, A. M., and Hajjar, D. P. (2016). Pitavastatin Differentially Modulates MicroRNA-Associated Cholesterol Transport Proteins in Macrophages. *PLoS One* 11 (7), e0159130. doi:10.1371/journal.pone.0159130

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

CASE REPORT
published: 28 May 2021
doi: 10.3389/fphar.2021.648519

Case Report of Novel Genetic Variant in KCNT1 Channel and Pharmacological Treatment With Quinidine. Precision Medicine in Refractory Epilepsy

M. C. Kravetz[1], M. S. Viola[1], J. Prenz[2], M. Curi[2], G. F. Bramuglia[1,3] and S. Tenembaum[4]*

[1]Department of Pharmacology, Faculty of Farmacy and Biochemistry, University of Buenos Aires, Buenos Aires City, Argentina, [2]Department of Cardiology, Garrahan Hospital, Buenos Aires City, Argentina, [3]Fundacion Investigar, Buenos Aires City, Argentina, [4]Department of Neurology, Garrahan Hospital, Buenos Aires City, Argentina

Case introduction: In this work we present a female infant patient with epilepsy of infancy with migrating focal seizures (EIMFS). Although many pharmacological schemes were attempted, she developed an encephalopathy with poor response to antiepileptic drugs and progressive cerebral dysfunction.

Aim: To present the pharmacological response and therapeutic drug monitoring of a paediatric patient with a severe encephalopathy carrying a genetic variant in KCNT1 gene, whose identification led to include quinidine (QND) in the treatment regimen as an antiepileptic drug.

Case report: Patient showed slow rhythmic activity (theta range) over left occipital areas with temporal propagation and oculo-clonic focal seizures and without tonic spasms three months after birth. At the age of 18 months showed severe impairments of motor and intellectual function with poor eye contact. When the patient was 4 years old, a genetic variant in the exon 24 of the KCNT1 gene was found. This led to the diagnosis of EIMFS. Due to antiepileptic treatment failed to control seizures, QND a KCNT1 blocker, was introduced as a therapeutic alternative besides topiramate (200 mg/day) and nitrazepam (2 mg/day). Therapeutic drug monitoring (TDM) of QND plasma levels needed to be implemented to establish individual therapeutic range and avoid toxicity. TDM for dose adjustment was performed to establish the individual therapeutic range of the patient. Seizures were under control with QND levels above 1.5 mcg/ml (65–70 mg/kg q. i.d). In addition, QND levels higher than 4.0 mcg/ml, were related to higher risk of suffering arrhythmia due to prolongation of QT segment. Despite initial intention to withdrawal topiramate completely, QND was no longer effective by itself and failed to maintain seizures control. Due to this necessary interaction between quinidine and topiramate, topiramate was stablished in a maintenance dose of 40 mg/day.

Conclusion: The implementation of Precision Medicine by using tools such as Next Generation Sequencing and TDM led to diagnose and select a targeted therapy for the treatment of a KCNT1-related epilepsy in a patient presented with EIMFS in early infancy

OPEN ACCESS

Edited by:
Alberto Lazarowski, University of Buenos Aires, Argentina

Reviewed by:
Sandra Orozco-Suárez, Mexican Institute of Social Security, Mexico,
Marta Vázquez, Universidad de la República, Uruguay

****Correspondence:***
M. C. Kravetz
mariaceciliakravetz@gmail.com

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *31 December 2020*
Accepted: *26 March 2021*
Published: *28 May 2021*

Citation:
Kravetz MC, Viola MS, Prenz J, Curi M, Bramuglia GF and Tenembaum S (2021) Case Report of Novel Genetic Variant in KCNT1 Channel and Pharmacological Treatment With Quinidine. Precision Medicine in Refractory Epilepsy. Front. Pharmacol. 12:648519. doi: 10.3389/fphar.2021.648519

and poor response to antiepileptic drugs. QND an old antiarrhythmic drug, due to its activity as KCNT1 channel blocker, associated to topiramate resulted in seizures control. Due to high variability observed in QND levels, TDM and pharmacokinetic characterization allowed to optimize drug regimen to maintain QND concentration between the individual therapeutic range and diminish toxicity.

Keywords: epilepsy, pharmacogenomics, pharmacokinetics, precision medicine, TDM, quinidine, HPLC

GRAPHICAL ABSTRACT |

INTRODUCTION

Epilepsy of infancy with migrating focal seizures (EIMFS) is a severe encephalopathy affecting children at very early stages. Seizures can begin as early as six months of life, and even a few weeks after delivery. This rare epileptic syndrome is characterized for pharmacological treatment resistance and poor prognosis. Its low prevalence varies from 1/200,000 to 1/1,000,000 habitants depending on of the population studied (Epilepsy of infancy with migrating focal seizures (2020) Epilepsy of infancy with migrating focal seizures | Epilepsy Action).

About etiology, the origin of this complex syndromes of infancy with migrating partial onset is not certain, but it is described the association with pathologic variants of KCNT1, SCN1A, SCN2A, PLCB1, TBC1D24 and CHD2 (International League Against Epilepsy, 2020 Epilepsy of infancy with migrating focal seizures, 2020). Around 70% of patients carries mutations in these and others genes and KCNT1 variants are implied in almost 40% of EIMFS (Lim et al., 2016).

In vitro functional studies have shown that KCNT1 carrying pathogenic variants generate larger currents when compared to wild-type KCNT1 channels, leading to the concept that a gain-of-function mechanism is responsible for epileptogenesis associated with KCNT1 variants (Barcia et al., 2012). Based on the ability to inhibit KCNT1 mutant channels, KCNT1 blockers like quinidine (QND), and old antiarrhythmic drug, and bepridil, have been proposed for the treatment of EIMFS (Milligan et al., 2014). Targeted therapy with QND has been introduced for compassionate use in EIMFS patients, resulting in variable anticonvulsant effect ranging from high pharmacological responses (Fukuoka et al., 2017) to a lack of efficacy or excessive toxicity (Chong et al., 2016; Madaan et al., 2018).

Previous data showed QND produced a concentration-dependent decrease in KCNT1 outward currents (Rizzo et al., 2016; Dilena et al., 2018) and that QND was even more potent in blocking KCNT1 channel carrying previously identified genetic variants, such as R950Q or E893K, in comparison with KCNT1 wild type (Dilena et al., 2018).

In this work we present a female pediatric patient with a severe case of EIMFS. Although many pharmacological schemes were attempted, the patient developed an epileptic syndrome with poor response to antiepileptic drugs and progressive cerebral dysfunction.

The patient had two episodes of apnea during the first week of life and showed poor eye contact, a trend toward a conjugate deviation of the eyes and head to the right or left, with horizontal nystagmus. She was extremely irritable. Three month later the patient was admitted due to three epileptic seizures recorded by electroencephalogram, which indicated slow rhythmic activity (theta range) over left occipital areas with temporal propagation. She presented oculo-clonic focal seizures which she repeated several times during examinations without tonic spasms.

The first year of life the patient received a drug regimen that included phenobarbital, valproic acid 750 mg/day, levetiracetam 600 mg/day, and ketogenic diet. Immunoglobulin therapy was also included because of low serum levels of Ig A and IgG.

As time progressed, the patient demonstrated spastic asymmetric quadriplegia, more prominent on the right, truncal hypotonia with poor head control, and lower-limb

pyramidal tract signs. There were no laboratory abnormalities in blood nor CSF, and no sign of skeletal malformation on physical examination.

Electroencephalography during sleep (multiple recordings) demonstrated paroxysmal multifocal discharges with periods of electric attenuation or suppression which were not continuous. Electrographic seizures were also recorded, arising from the left frontal-temporal and right temporal-occipital regions. The clinical presentation and electroencephalographic features resembled an early myoclonic encephalopathy.

When the patient was 18 months old topiramate (TPM) was initiated that was related with a decrease in the frequency of seizures. At two years old (September 2011) seizures could be controlled receiving TPM 7 mg/kg/day, levetiracetam 48 mg/kg/day, and nitrazepam 2 mg/day (ketogenic diet was discontinuated). Nevertheless, electroencephalography during sleep in October 2011 revealed a typical sleep pattern with sleep spindles and vertex waves, in addition to infrequent sharp waves in left occipital region without generalized paroxysms or electric suppression.

A genomic study was performed, and no pathogenic variants were found in Glut1 gene, MCP2, CDKL5, FOXG1. Levetiracetam was finally withdrawn in April 2012 with no seizure recurrence. Due to a clear right-sided hemidystonia, levo-dopa with carbidopa in increasing dose was initiated in May 2012.

A brain MRI was performed at 3 years of age, and revealed subtle signs of cortical atrophy, hypoplastic corpus callosum, T2 and FLAIR high signal abnormalities involving the brain white matter, more prominent in the left hemisphere. Although neurological improvement could be observed, the patient showed a persistent right-sided hemidystonia, increased muscle tone in lower limbs with pyramidal signs and sleep pattern discharges.

AIM

To present the pharmacological response and therapeutic drug monitoring (TDM) of a paediatric patient with a severe encephalopathy refractory to pharmacological treatment, carrying a genetic variant in KCNT1 gene, whose identification by Next Generation Sequencing (NGS) led to include QND as antiepileptic drug.

CASE REPORT

To summarize, the initial years of the patient included epileptic encephalopathy development, cognitive and physical impairment which were part of an undiagnosed syndrome resistant to pharmacological treatment.

On 2013, when the patient was 4 years old, a Whole Exome Sequencing (WES) was performed. The entire exome was captured using reagent that targets coding exons from the consensus coding sequence. Capture enrichment was followed by sequencing on Illumina platform using previously described standard protocols. Illumina sequence analysis was performed using the Human Genome Sequencing Center's integrated Mercury pipeline (Baylor College of Medicine, United States of America). A pathogenic genetic variant associated to EIMFS (Barcia et al., 2012; The ILAE Genetics Commission Blog, 2020 KCNT1 – this is what you need to know, 2020) located in exon 24 (2795 T > C) of the KCNT1 gene (Potassium Sodium-Activated Channel Subfamily T Member 1) was identified.

Due to the remaining nocturnal activity despite many pharmacological schemes, a different approach was implemented.

At that moment new clinical studies showed benefits in using QND in epilepsy (Mikati et al., 2015), due to its activity as KCNT1 blocker. Due to its pharmacological effect and the variant described in the patient, at the age of six years old (in 2015), QND was included as "off label" drug regimen, besides TPM (175 mg/day) and nitrazepam (1.5 mg/day).

Because QND toxicity was well described it was needed to monitor the cardiac function and plasma concentrations. For this aim our laboratory developed and validated HPLC-UV method to quantify QND in human serum. Propranolol was added to the samples as an internal standard, followed by liquid extraction with cyclohexane and evaporation. The chromatography condition included a reversed phase column and a mobile phase consisting of acetonitrile and phosphate buffer (25:75).

Validation parameters studied were specificity (no interferences were found), lineality (0.5–8.0 mcg/ml), repeatability, accuracy (less than 10% CD), sensitivity and stability (in different conditions and time periods). Tests were evaluated according US FDA Guidance for Industry Bioanalytical Method Validation (US FDA Guidance for Industry Bioanalytical Method Validation, 2018).

During QND dose titration, TPM was decreasing the daily doses (initially 175 mg/day) to complete withdrawal, while nitrazepam dose regimen of 2 mg/day was maintained.

TDM was implemented, to establish the individual therapeutic range, as well as electrocardiogram (ECG) studies to assist cardiac follow up. When QND dose was above 50 mg/kg/day in four administrations (achieving a minimum serum level of 1.0 mcg/ml) the therapy started to modify the discharges profile. Only when doses reached 65–70 mg/kg q. i.d (and serum levels were above 1.5 mcg/ml) the patient was seizure free.

According to the original strategy, TPM was almost withdrawn. Surprisingly, QND doses of 65–70 mg/kg which had initially shown therapeutic effectiveness, failed to gain seizures control. Due to this observed synergic effect between TPM and QND, TPM was again indicated. TPM maintenance dose was 2.5 mg/kg/day and then increased to 4.5 mg/kg/day (which achieved levels of 3 mcg/ml and 6 mcg/ml of TPM respectively).

Table 1 shows the variability of QND levels at higher doses compared with lower doses. Initially, when doses increased from 240 to 640 mg/day, QND concentration augmented from 0.6 mcg/ml to 1.0 mcg/ml. But analyzing the same increment in dose (approximately 400 mg/day) with higher doses (from 1,160 to 1,600 mg/day), it resulted in a higher variation of serum levels obtained (1.5–4.1 mcg/ml).

TABLE 1 | Comparison of different variation in lower and higher doses, occurred during patient follow up with TDM. Levels are more disperse in association with doses increment.

Range of doses mg/day	Doses variation (Max–min)	Range of plasma concentration mcg/ml	Plasma concentration variation (Max-min)
240–640	400.0	0.6–1	0.4
1,160–1,600	440.0	1.5–4.1	2.6

TABLE 2 | Serum levels and doses related to QT segment prolongation (Reference <0.46 s). In bold is highlighted those situations where reducing doses was necessary to correct QT segment prolongation.

Date dd/mm/yy	Dose (mg/day)	Plasmatic level (mcg/ml)	QT segment (sec.)
22/03/16	560	0.5	0.48
08/05/17	**1,320**	**4.0**	**0.60**
06/09/17	1,360	3.0	0.46
22/09/17	1,320	2.8	0.52
16/01/18	1,400	1.9	0.46
25/09/18	1,440	1.8	0.52
21/01/19	1,560	3.4	0.50
20/08/19	**1,560**	**5.9**	**0.50**

Values in bold indicate toxicity events.

It is important to mention that the patient received different pharmacological treatment according particular necessities, such as: antimicrobial agents or laxatives, which might affect QND levels. Moreover, changes in TPM doses could contribute to the variability observed. At the beginning of the treatment the increase in QND doses did not correlate with a proportional increase of QND plasma levels. After a reduction in TPM dose regimen below 50 mg/day the increase in QND doses yielded a proportional dose-concentration relationship.

It is also known that QND cardiotoxicity could produce arrhythmia due to prolongation of QT segment. Our patient experienced the risk of suffering this adverse event on several occasions. Nevertheless, we identified two different related patterns (**Table 2**). The first one was observed during the titration period, when QND doses were incremented every 14 days. QT segment values corrected themselves without any medical intervention, once normalized titration could continue. On the other hand, when the levels reached were higher than 4.0 mcg/ml, the prolongation was persistent, and it was mandatory to reduce the doses. To sum up, Graphic 1 shows QND levels achieved during the follow-up, correlated with QND doses.

DISCUSSION

Although pathogenic variants on *KCNT1* are rare in the population, it has been proposed as an important cause of epilepsy with a wide phenotypic spectrum (Barcia et al., 2012). We characterize the genotype, as a pharmacological target, but also a therapeutic drug monitoring strategy to optimize the pharmacological treatment with the potassium channel blocker QND in association with other antiepileptic drugs in a pediatric patient.

KCNT1 encodes the largest potassium channel subunit and is thought to regulate the hyperpolarization that follows repetitive firing. Functional studies have shown that genetic variants in the cytoplasmic C-terminal domain caused constitutive activation of the potassium channel (Milligan et al., 2016), which modify normal neuronal firing and directly led to epileptogenesis.

A KCNT1 variant previously reported as pathogenic (Vanderver et al., 2014; Yang et al., 2014) was identified in the patient. This '*de novo*' variant is highly conserved, and the p.F932I change is, in principle, the most damaging non-synonymous change associated to this variant.

The KCNT1 variant identified in our work is located immediately adjacent to two previously described pathogenic variants (pArg928Cys and pAla934Thr), and therefore was proposed as the underlying disease-causing mechanism in this case. KCNT1 related epilepsies is included in a broader group of potassium channel related epilepsies that may be amendable to channel specific therapies, such as QND.

QND an old antiarrhythmic and antimalarial drug, has emerged as a potential precision therapy for KCNT1-related epilepsy. *In vitro* functional assays have demonstrated that gain of-function effects associated to KCNT-1 gene variants can be reversed by QND. Although previous case reports showed a significant reduction in seizure burden (Bearden et al., 2014; Milligan et al., 2014), subsequent reports demonstrated less favorable results, raising questions about QND effectiveness for treating KCNT1-related epilepsy (Chong et al., 2016; Numis et al., 2018).

In this manuscript we described the pharmacological effects and therapeutic drug monitoring strategy applied after including QND to a patient's drug regimen who carries a KCNT1 variant, receiving also TPM and nitrazepam as pharmacological therapy.

One interesting point to note is that in a previous study, a different pattern of pharmacological response was reported related to the KCNT-1 variant identified. In our work, the less frequent 2795 T > C variant resulted in sustained seizure freedom after quinidine treatment, in agreement with the data showed by Fitzgerald et al. (2019) in a patient treated with QND and receiving TPM and nitrazepam as antiepileptic drugs.

It has been also proposed that the limited efficacy of QND previously reported in KCNT1 patients could be related to the limited bioavailability of QND in the CNS. In this way, it has been shown in healthy volunteers that QND levels in CSF could be lower than in plasma where QND CSF concentration were 16% of unbound serum concentrations (range, 4–37%) (Ochs et al., 1980). Anyway, there is no data available of CSF levels in KCNT1 patients receiving QND to evaluate the CSF/plasma quinidine concentration relationship and antiepileptic response.

Moreover, the QND dose increment at higher doses resulted in greater variability in QND plasma levels. This phenomenon could in part be explained by the rise in QND free fraction in the range of concentration observed (1.5–4.0 ug/ml), that could be available to be eliminated or re-distributed according to different factors affecting the patient (Ochs et al., 1980). In this sense, the variability observed, justified the QND monitoring of plasma levels with the aim to optimize dose regimen.

In the largest cohort of patient with KCNT1 related epilepsy, that included 43 patients, QND treatment resulted in a >50% reduction in seizure frequency in only 20% of patients receiving this therapy, with sustained seizure freedom occurring rarely (Fitzgerald et al., 2019). In the study mentioned, therapeutic blood levels were not achieved in many patients, with 45% of patients failing to reach a blood level of 2 ug/ml.

There are several potential explanations for this finding, but we can hypothesize that QTc prolongation concern, infrequent use of this drug or lack of knowledge regarding the most appropriate strategies for therapeutic monitoring, could be some of the reasons that prevented reaching higher therapeutic levels.

TDM is a useful tool when allows us to select a goal and to reach a target by developing a pharmacokinetic/pharmacodynamics (PK/PD) model for a particular patient, but it is useless when there is a lack of skills related to drug monitoring practice, because blood monitoring of a particular drug is very infrequent, or if no pharmacokinetics recommendations are given.

In our study QND drug regimen optimization was performed and a high proportion of QND levels reached the therapeutic target. During the follow-up, intensive therapeutic drug monitoring was accompanied by pharmacokinetics recommendation of QND dose adjustment with a high rate of physician compliance. This strategy allowed to modify QND dose regimen to achieve plasma levels into the individual therapeutic range for this patient.

To highline from our work, the patient got free from seizures receiving QND 70–90 mg/kg q. i.d, TPM 4.5 mg/kg/day and nitrazepam 2.5 mg/day as pharmacological treatment. Due to the fact that QND effectiveness decreased when TPM doses were below 20–30 mg/day during TPM withdrawal, and that QND and TPM does not share the same mechanism of action, we can suggest a potentially synergic effect in this case. Initially, high doses of TPM showed to be ineffective to control seizures in this patient. However low doses of TPM were needed associated to QND to get free from seizures. Although the nature of this interaction needs further studies, QND as monotherapy may not be as effective acting at the KCNT1 receptor level making necessary the association of another antiepileptic drug to maintain the pharmacological response.

From a pharmacokinetic point of view QND is a very potent inhibitor of the efflux transporter P-glycoprotein (PGP), and is involved in the increased bioavailability at the CNS and toxicity of different drugs. It has been shown that TPM is a substrate of PGP, suggesting that up-regulation of MDR-1 gene could affect the TPM brain levels. QND could be acting as a "booster" and might be increasing TPM levels that reach CNS (Sadeque et al., 2000), allowing the use of lower doses of TPM that showed to be effective.

In our experience QND have shown to be affected by other conditions. The patient levels drop off on many occasions neither for changes in doses nor changes in the administration of QND. Levels that were below expected could be associated for example to other pharmacological treatment such as wide spectrum antibiotics which might have affected QND absorption. Laxatives also may have decreased QND absorption because of the increment in bowel motility.

Increasing TPM doses are also suspected of altering serum QND concentrations through several different mechanisms that need to be studied more carefully. According to bibliography, the inhibition of carbonic anhydrase enzyme by TPM (Ruiz Granados et al., 2015) might conduct to metabolic acidosis, producing an increase in urinary pH and QND plasma levels. By the contrary, in our work at higher TPM doses, QND titration was not accompanied by a proportional increase of QND plasma levels until TPM regimen was below 50 mg/day. Anyway, as we mentioned, the fluctuation of QND levels was more significant at higher QND doses, and TDM was successfully applied in order to reach the individual therapeutic range that was defined between 1.5–4.0 ug/ml for this patient.

Considering previous studies and the results presented in this report, we suggest that in patients with KCNT-1 related epilepsy, the inclusion of QND should be accompanied of the analysis of the KCNT-1 genetic variants identified, together with a TDM strategy that allows to characterize the pharmacokinetic of QND, with the aim to reach the individual therapeutic target in order to increase efficacy with minimal side effects.

CONCLUSION

The implementation of Precision Medicine by using tools such as NGS and TDM led to diagnose and select a targeted therapy for the treatment of a KCNT1-related epilepsy in a patient presented with EIMFS in early infancy and poor response to antiepileptic drugs. QND an old antiarrhythmic drug, due to its activity as KCNT1 channel blocker, associated to TPM resulted in seizures control. Due to high variability observed in QND levels, TDM and pharmacokinetic characterization allowed to optimize drug regimen to maintain QND concentration between the individual therapeutic range and diminish toxicity.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comite de Ética de la Facultad de Farmacia y Bioquímica-UBA. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

Pharm. Rodriguez Basso A. G. for assisting HPLC method validation, Bioch. Barrionuevo N.M. and Pharm. Fernández F.A. for technical assistance. Nuñez M and the Institute of Pharmacology at University of Medicine- UBA for equipment backup.

REFERENCES

Barcia, G., Fleming, M. R., Deligniere, A., Gazula, V.-R., Brown, M. R., Langouet, M., et al. (2012). De novo gain-of-function KCNT1 channel mutations cause malignant migrating partial seizures of infancy. *Nat. Genet.* 44, 1255–1259. doi:10.1212/NXG.0000000000000036310.1038/ng.2441

Bearden, D., Strong, A., Ehnot, J., DiGiovine, M., Dlugos, D., and Goldberg, E. M. (2014). Targeted treatment of migrating partial seizures of infancy with quinidine. *Ann. Neurol.* 76, 457–461. doi:10.1002/ana.24229

Chong, P. F., Nakamura, R., Saitsu, H., Matsumoto, N., and Kira, R. (2016). Ineffective quinidine therapy in early onset epileptic encephalopathy withKCNT1mutation. *Ann. Neurol.* 79, 502–503. doi:10.1002/ana.24598

Dilena, R., DiFrancesco, J. C., Soldovieri, M. V., Giacobbe, A., Ambrosino, P., Mosca, I., et al. (2018). Early treatment with quinidine in 2 patients with epilepsy of infancy with migrating focal seizures (EIMFS) due to gain-of-function KCNT1 mutations: functional studies, clinical responses, and critical issues for personalized therapy. *Neurotherapeutics* 15 (4), 1112–1126. doi:10.1007/s13311-018-0657-9

Epilepsy of infancy with migrating focal seizures (2020). Epilepsy of infancy with migrating focal seizures | Epilepsy Action. Available at: https://www.epilepsy.org.uk/info/syndromes/migrating-partial-epilepsy-in-infancy (Accessed November 30, 2020).

Fitzgerald, M. P., Fiannacca, M., Smith, D. M., Gertler, T. S., Gunning, B., Syrbe, S., et al. (2019) Treatment responsiveness in KCNT1-related epilepsy. *Neurotherapeutics* 16, 848–857. doi:10.1007/s13311-019-00739-y

Fukuoka, M., Kuki, I., Kawawaki, H., Okazaki, S., Kim, K., Hattori, Y., et al. (2017). Quinidine therapy for West syndrome with KCNTI mutation: a case report. *Brain Development* 39, 80–83. doi:10.1016/j.braindev.2016.08.002

International League Against Epilepsy (2020). *Epilepsy of infancy with migrating focal seizures* Available at: https://www.epilepsydiagnosis.org/syndrome/infancy-migrating-focal-genetics.html (Accessed December 20, 2020).

Lim, C. X., Ricos, M. G., Dibbens, L. M., and Heron, S. E. (2016). KCNT1mutations in seizure disorders: the phenotypic spectrum and functional effects. *J. Med. Genet.* 53, 217. doi:10.1136/jmedgenet-2015-103508

Madaan, P., Jauhari, P., Gupta, A., Chakrabarty, B., and Gulati, S. (2018). A quinidine non responsive novel KCNT1 mutation in an Indian infant with epilepsy of infancy with migrating focal seizures. *Brain Development* 40, 229–232. doi:10.1016/j.braindev.2017.09.008

Mikati, M. A., Jiang, Y.-h., Carboni, M., Shashi, V., Petrovski, S., Spillmann, R., et al. (2015). Quinidine in the treatment of KCNT1-positive epilepsies. *Ann. Neurol.* 78, 995–999. doi:10.1002/ana.24520

Milligan, C. J., Li, M., Gazina, E. V., Heron, S. E., Nair, U., Trager, C., et al. (2014). KCNT1gain of function in 2 epilepsy phenotypes is reversed by quinidine. *Ann. Neurol.* 75, 581–590. doi:10.1002/ana.24128

Milligan, C. X., Ricos, M. G., Dibbens, L. M., and Heron, S. E. (2016). KCNT1mutations in seizure disorders: the phenotypic spectrum and functional effects. *J. Med. Genet.* 53, 217–225. doi:10.1002/ana.2452010.1136/jmedgenet-2015-103508

Numis, A. L., Nair, U., Datta, A. N., Sands, T. T., Oldham, M. S., Patel, A., et al. (2018). Lack of response to quinidine in KCNT1 -related neonatal epilepsy. *Epilepsia* 59 (10), 1889–1898. doi:10.1111/epi.14551

Ochs, H. R., Greenblatt, D. J., Lloyd, B. L., Woo, E., Sonntag, M., and Smith, T. W. (1980). Entry of quinidine into cerebrospinal fluid. *Am. Heart J.* 100 (3), 341–346. doi:10.1016/0002-8703(80)90148-9

Rizzo, F., Ambrosino, P., Guacci, A., Chetta, M., Marchese, G., Rocco, T., et al. (2016). Characterization of two de novo KCNT1 mutations in children with malignant migrating partial seizures in infancy. *Mol. Cell Neurosci.* 72, 54–63. doi:10.1016/j.mcn.2016.01.004

Ruiz Granados, V. J., and Márquez Romero, J. M. (2015). Topiramato en monoterapia o en combinación como causa de acidosis metabólica en adultos con epilepsia. *Rev Neurol* 60, 159–163. doi:10.33588/rn.6004.2014451

Sadeque, A., Wandel, C., He, H., Shah, S., and Wood, A. J. J. (2000). Increased drug delivery to the brain by P-glycoprotein inhibition. *Clin. Pharmacol. Ther.* 68, 231–237. doi:10.1067/mcp.2000.109156

The ILAE Genetics Commission Blog (2020). *KCNT1 – this is what you need to know*. Available at: http://epilepsygenetics.net/the-epilepsiome/kcnt1-this-is-what-you-need-to-know/ (Accessed December 29, 2020).

US FDA Guidance for Industry Bioanalytical Method Validation (2018). Available at: https://www.fda.gov/files/drugs/published/Bioanalytical-Method-Validation-Guidance-for-Industry.pdf (Accessed on February 25, 2021).

Vanderver, A., Simons, C., Schmidt, J. L., Pearl, P. L., Bloom, M., Lavenstein, B., et al. (2014). Identification of a novel de novo p.Phe932Ile KCNT1 mutation in a patient with leukoencephalopathy and severe epilepsy. *Pediatr. Neurol.* 50 (1), 112–114. doi:10.1016/j.pediatrneurol.2013.06.024

Yang, Y., Muzny, D. M., Xia, F., Niu, Z., Person, R., Ding, Y., et al. (2014). Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* 312, 18708–18709. doi:10.1001/jama.2014.14601

Conflict of Interest: The handling Editor declared a shared affiliation, though no other collaboration, with several of the authors (MCK, MSV, GFB).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 03 June 2021
doi: 10.3389/fphar.2021.648390

Theragnosis for Duchenne Muscular Dystrophy

Leonela Luce[1,2], *Micaela Carcione*[1,2], *Chiara Mazzanti*[1,2], *Paula I. Buonfiglio*[3], *Viviana Dalamón*[3], *Lilia Mesa*[4], *Alberto Dubrovsky*[4], *José Corderí*[4] *and Florencia Giliberto*[1,2]*

[1]*Laboratorio de Distrofinopatías, Cátedra de Genética, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Buenos Aires, Argentina,* [2]*Instituto de Inmunología, Genética y Metabolismo (INIGEM), CONICET—Universidad de Buenos Aires, Buenos Aires, Argentina,* [3]*Instituto de Investigaciones en Ingeniería Genética y Biología Molecular (INGEBI) "Dr. Héctor N. Torres", CONICET, Buenos Aires, Argentina,* [4]*Instituto de Neurociencias, Fundación Favaloro, Buenos Aires, Argentina*

Dystrophinopathies cover a spectrum of rare progressive X-linked muscle diseases, arising from *DMD* mutations. They are among the most common pediatric muscular dystrophies, being Duchenne muscular dystrophy (DMD) the most severe form. Despite the fact that there is still no cure for these serious diseases, unprecedented advances are being made for the development of therapies for DMD. Some of which are already conditionally approved: exon skipping and premature stop codon read-through. The present work aimed to characterize the mutational spectrum of *DMD* in an Argentinian cohort, to identify candidates for available pharmacogenetic treatments and finally, to conduct a comparative analysis of the Latin American (LA) frequencies of mutations amenable for available DMD therapies. We studied 400 patients with clinical diagnosis of dystrophinopathy, implementing a diagnostic molecular algorithm including: MLPA/PCR/Sanger/Exome and bioinformatics. We also performed a meta-analysis of LA's metrics for DMD available therapies. The employed algorithm resulted effective for the achievement of differential diagnosis, reaching a detection rate of 97%. Because of this, corticosteroid treatment was correctly indicated and validated in 371 patients with genetic confirmation of dystrophinopathy. Also, 20 were eligible for exon skipping of exon 51, 21 for exon 53, 12 for exon 45 and another 70 for premature stop codon read-through therapy. We determined that 87.5% of DMD patients will restore the reading frame with the skipping of only one exon. Regarding nonsense variants, UGA turned out to be the most frequent premature stop codon observed (47%). According to the meta-analysis, only four LA countries (Argentina, Brazil, Colombia and Mexico) provide the complete molecular algorithm for dystrophinopathies. We observed different relations among the available targets for exon skipping in the analyzed populations, but a more even proportion of nonsense variants (~40%). In conclusion, this manuscript describes the theragnosis carried out in Argentinian dystrophinopathy patients. The implemented molecular algorithm proved to be efficient for the achievement of differential diagnosis, which plays a crucial role in patient management, determination of the standard of care and genetic counseling. Finally, this work contributes with the international efforts to characterize the frequencies and variants in LA, pillars of drug development and theragnosis.

OPEN ACCESS

Edited by:
Luis Abel Quiñones, University of Chile, Chile

Reviewed by:
Félix Javier Jiménez-Jiménez, Hospital Universitario del Sureste, Spain
Jorge Alfredo Bevilacqua, University of Chile, Chile

***Correspondence:**
Florencia Giliberto
gilibertoflor@gmail.com

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *31 December 2020*
Accepted: *17 May 2021*
Published: *03 June 2021*

Citation:
Luce L, Carcione M, Mazzanti C, Buonfiglio PI, Dalamón V, Mesa L, Dubrovsky A, Corderí J and Giliberto F (2021) Theragnosis for Duchenne Muscular Dystrophy. Front. Pharmacol. 12:648390. doi: 10.3389/fphar.2021.648390

Keywords: dystrophinopathies, duchenne muscular dystrophy (DMD), meta-analysis, theragnosis, mutagenic spectrum, nonsense, exon skipping, Latin America

INTRODUCTION

Muscular Dystrophies (MDs) are hereditary disorders that cause weakness and progressive degeneration of skeletal muscles. These diseases are caused by molecular alterations in a wide range of genes that encode proteins that participate in the stability, maintenance, repair, regeneration and proper functioning of muscle fibers (Wallace and McNally, 2009). Although DMD clinical features are quite typical for the trained physician, there are other less frequent forms of MDs with similar clinical characteristics such as sarcoglycanopathies, laminopathies and other forms of LGMD. Therefore, the clinical diagnosis can be misled by these overlapping features, turning the molecular diagnosis into a crucial tool for the achievement of a differential diagnosis.

Dystrophinopathies are the most frequent form of MDs among the pediatric population. These are X-linked recessive diseases caused by pathogenic variants in the *DMD* gene (OMIM ID: 300377) (Hoffman et al., 1987; Koenig et al., 1988). Although in theory dystrophinopathies can be subdivided into three distinctive clinical conditions, Duchenne muscular dystrophy (DMD), Becker muscular dystrophy (BMD), and DMD-associated dilated cardiomyopathy (DCM), they actually entail a continuous spectrum of muscle diseases (Darras et al., 2000; Brandsema and Darras, 2015). DMD is the most prevalent and severe pediatric form of MD, with an incidence of 1:3500–5000 male births (Mendell et al., 2012). It is characterized by progressive muscle-waste, which leads to disability and premature death (Aartsma-Rus et al., 2017). On the other hand, BMD affects 1:18.000 born males and has a milder symptomatology pattern and/or slower progression rate than DMD.

The genotype/phenotype correlation relies on the impact of the molecular alteration on dystrophin function. DMD is mainly associated with mutations leading to complete absence of functional dystrophin, such as frameshift or nonsense variants. Instead, BMD is caused by a decrease in the amount or function of dystrophin, as it would be the case of in-frame variants (Guiraud et al., 2015). Nonetheless, in some cases, the phenotype predicted on the basis of molecular alterations detected at genomic level do not correlate with the observed clinical picture. This would be the case of patients carrying out-of-frame or nonsense variants but showing a mild progression of the disease, which could be explained by an endogenous exon skipping restoring the reading frame or avoiding the premature stop codon, respectively.

Mutational spectrum of the *DMD* gene comprise mainly copy number variants (CNVs), such as deletions (~68%) or duplications (~11%) of one or more exons, and small molecular alterations in the remaining ~20% (Aartsma-Rus et al., 2016). In addition, around half of small sequence variants are nonsense substitutions.

Accurate molecular diagnosis, given by the identification and precise characterization of deleterious variants, is crucial for dystrophinopathy patients to confirm the clinical presumptive diagnosis, to access to the specific and optimal standard of care (Bushby et al., 2010) and determine eligibility for the available pharmacogenetic treatments. For example, molecular confirmation of dystrophinopathy determines applicability of corticosteroid therapy, as DMD is one of the MDs showing fruitful results from this treatment (Albuquerque et al., 2014; Bello et al., 2015). On the other hand, molecular diagnosis plays a key role in family planning and, therefore, prevention.

Despite the fact that there is still no cure for these serious diseases, unprecedented advances are being made for the development of therapies for DMD. Hitherto, three mutation specific treatments already have conditional approval: premature stop codon read-through (Ataluren) by the European Medicines Agency (EMA) and exon skipping for exon 51 (Eteplirsen) and exon 53 (Golodirsen) by the Food and Drug Administration (FDA) (Haas et al., 2015; Mah, 2018).

The rationale of "exon skipping" is to restore the *DMD* reading frame by the removal of one or several exons adjacent to any of the deletion's borders, which is accomplished by targeting regulatory splice sites in the pre-mRNA (Syed, 2016; Kinane et al., 2018). Therefore, the resulting spliced transcript might generate a partially functional dystrophin, albeit internally deleted and quantitatively reduced, capable of shifting the patient's severe phenotype into a milder one (Syed, 2016; Kinane et al., 2018). Apart from the previously mentioned exon skipping for exon 51 and exon 53, which apply to 10–15% and 8–10% of DMD patients respectively, antisense oligonucleotides to target exon 45 (Casimersen) are now pursuing FDA's approval.

On the other hand, the principle behind Ataluren is the endogenous process known as "stop codon suppression or readthrough", which entails the recognition of stop codons by a near-cognate aminoacyl-tRNA (Keeling and Bedwell, 2011). The efficacy of the suppression process depends on several conditions: the innate readthrough capacity of each stop codon (UGA > UAG > UAA), the sequence surrounding the termination codon and the functionality of the incorporated amino acid (Miller and Pearce, 2014). This therapy specifically applies to patients carrying *DMD* nonsense mutations (10–15%).

In addition, under the name of dystrophin restoration therapies are included the gene-transfer strategy, which incorporates short versions of the *DMD* gene but encoding functional mini/micro-dystrophins, and the *DMD* gene-editing approach, that applies CRISPR-Cas9 technology to correct the molecular alteration carried by each individual (Duchêne et al., 2018; Verhaart and Aartsma-Rus, 2019; Lim et al., 2020; Mendell et al., 2020). Furthermore, gene-transfer therapies for other types of muscular

dystrophies (*CAPN3*, *SGCB*, *SGCA*, *DYSF*, *SGCG* and *ANO5*) are burgeoning (Gene Therapy Engine; Chu and Moran, 2018)[1].

On the other hand, one of the pillars for drug development and theragnosis is the information regarding the frequency and types of molecular alterations that take place in a certain gene. However, this knowledge principally comes from Europe and the United States, as little is known about the Latin American frequencies, which is also true for *DMD*.

Therefore, the present work has three major aims. Firstly, the characterization of the mutational spectrum of the *DMD* gene in an Argentinian dystrophinopathy cohort. Secondly, the identification of candidate patients for the available pharmacogenetic treatments for DMD. Finally, the conduction of a comparative analysis of the Latin American frequencies of the mutations amenable for the available DMD therapies.

MATERIALS AND METHODS

Patients and Samples

A cohort of 400 boys with presumptive clinical diagnosis of dystrophinopathy was referred to our laboratory in pursuit of differential molecular diagnosis. The criteria followed for the clinical diagnosis was the one described in Birnkrant et al., 2018. The algorithm began with the clinical assessment. Clinical suspicion of DMD arose in cases with DMD family history or based on the observation of progressive muscular weakness, Gowers sign, calf muscle pseudohypertrophy, difficulty at climbing stairs, waddling gait and/or toe walking. The second step was the determination of the CK level, followed by molecular studies. If no pathogenic variant is found by genetic testing the guideline recommends a muscle biopsy.

Whole blood was drawn by venipuncture with 5% ethylene-diamine tetraacetic acid (EDTA) as anticoagulant for all study subjects. Genomic DNA was isolated using the cetyl-trimethyl-ammonium bromide (CTAB) method (Murray and Thompson, 1980). DNA concentration and quality were measured by absorbance at 260 nm and by the ratio of A260 nm/A280 nm, respectively. All samples were stored at −20°C.

The protocol was approved by the Institutional Review Board. Informed consent was obtained for all study subjects prior to the molecular studies.

Multiplex Ligation-dependent Probe Amplification (MLPA)

The commercially available MLPA kit for the *DMD* gene (Salsas PO34–PO35) was used to screen for gene deletions/duplications (Schwartz and Dunø, 2004; Gatta et al., 2005; Janssen et al., 2005). Reactions were carried out according to the manufacturer's recommendations [MRC-Holland, Amsterdam, Netherlands (www.mlpa.com)]. Products were analyzed using a fragment analyzer sequencer (ABI 3730XL; Applied Biosystems, Foster City, California) and 500Liz as internal size standard. Data analysis was performed using Coffalyser (MRC-Holland, Amsterdam, Netherlands) and GeneMarker V2.2.0 (Softgenetics, State College, Pennsylvania) software. Wild-type, deleted, and duplicated controls were included in all reactions. Following the best practice guidelines for genetic testing for dystrophinopathies, cases with single-exon deletion were confirmed by PCR and/or Sanger sequencing (Fratter et al., 2020).

Whole Exome Sequencing (WES)

WES was carried out by Macrogen Services (Republic of Korea). Exome libraries were constructed by hybridization capture with the Agilent SureSelect V4/V5/V6 Target Enrichment Kits (Agilent Technologies, Santa Clara, United States). WES was performed on the Illumina HiSeq4000/NovaSeq6000 platforms (Illumina, San Diego, United States), following the manufacturer's recommendations. FASTQ sequencing files were aligned to the Human Reference Genome hg19 from UCSC (original GRCh37 from NCBI, Feb. 2009) applying Burrows-Wheeler Alignment Tool (BWA−0.7.12). Analysis proceeded using Picard (picard-tools-1.130) and Genome Analysis Toolkit (GATK3.v4). Finally, variant annotation was carried out applying SnpEff (SnpEff_v4.1g), dbSNP database (version 142), 1000Genomes phase 3, ClinVar database (version 05/2015) and ESP database (ESP6500SI_V2). Furthermore, in order to determine the coverage, coverage depth and the quality of the reads, bam files were analyzed using the Integrative Genomics Viewer (IGV) software (Broad Institute, University of California, United States).

Selection of Disease-Associated Candidate Variants

The screening of disease-associated variants from the NGS results started with the analysis of the *DMD* gene. When no *DMD* disease associated variants were identified, we broadened the analysis to genes associated with other monogenic neuromuscular disorders (NMDs), beginning with muscular dystrophies, group 1 of "The Gene table of Neuromuscular disorders" (Benarroch et al., 2019). When no disease-associated variants were found, we extended the search to all the groups listed in the table previously mentioned (Benarroch et al., 2019).

The detected sequence variants were classified according to the standards and guidelines of the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (Richards et al., 2015). Nomenclature of the identified variants was achieved following the HGVS standards (Dunnen et al., 2016). The classification of variants was performed on the basis of the information gathered from: 1) Type and effect of the molecular alteration; 2) Population data from 1000Genomes and gnomAD (https://gnomad.broadinstitute.org/); 3) disease/gene specific databases, such as Leiden open variation database (LOVD) (http://www.lovd.nl/3.0/home/) and ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/); 4) *In silico* predictive analysis: PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (http://sift.jcvi.org/), Mutation Taster (http://www.mutationtaster.org/), Mutation Assessor (http://mutationassessor.org/r3/), CADD (http://cadd.gs.washington.edu/), UMD Predictor (http://umd-predictor.eu/analysis.php), Human Splicing Finder (https://www.genomnis.com/access-hsf), etc.; 5) Phenotypic features; 6) Familial segregation; and, 7) Bibliographic reports of functional assays.

[1]https://www.sarepta.com/science/gene-therapy-engine.

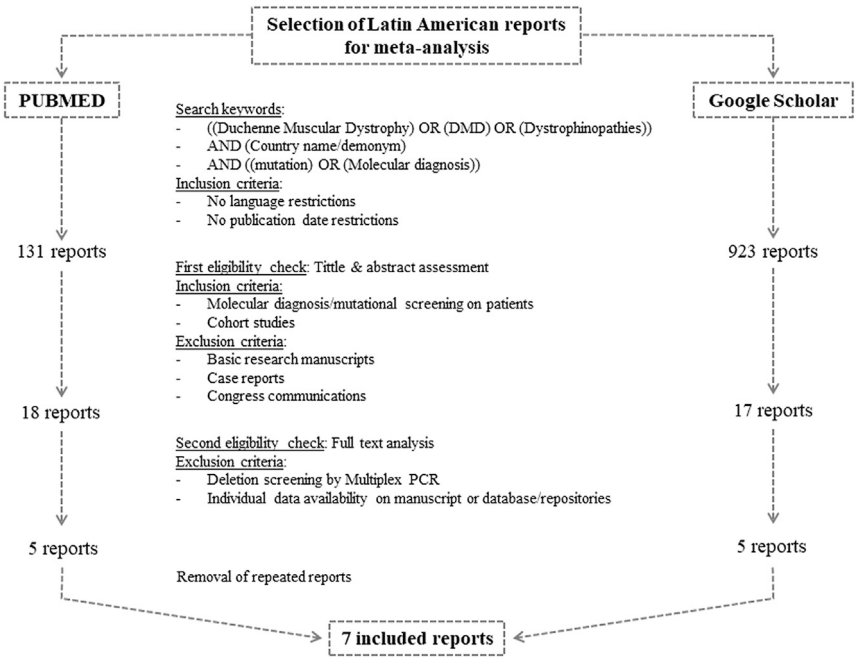


FIGURE 1 | Flow chart of search and selection algorithm of Latin American reports for the meta-analysis.

Polimerase Chain Reaction (PCR)-Sanger Sequencing

Every disease associated or likely pathogenic variant identified by WES and single exon deletion observed by MLPA was corroborated by PCR-Sanger sequencing. Also, this technique was employed for the analysis of patients with known familial causative small molecular alteration. Primer sequences and PCR conditions were obtained from the Leiden muscular dystrophy site [Leiden muscular dystrophy webpages (www.dmd.nl)]. All PCR reactions were performed in a thermal cycler (Veriti; Applied Biosystems, Foster City, California). PCR amplicons were analyzed by 2% agarose (Genbiotech SRL) gel electrophoresis in 1X TBE buffer and dyed with GelRed™ (Biotium). Positive controls (wild-type DNA) and negative controls (no DNA) were included in all reactions. The exons were sequenced using both PCR primers and the reaction products were analyzed using a DNA analyzer (ABI 3730 XL; Applied Biosystems, Foster City, California). The quality of the obtained sequence was determined using FinchTV software (Geospiza, Seattle, United States) and the results were analyzed by comparison with the GenBank sequence of the *DMD* muscular isoform (Dp427m, NM_004006.3).

Analysis of Exonic Targets for Exon Skipping

In order to establish the most frequent targets for exon skipping in our cohort, we selected a subset of 112 patients carrying out-of-frame deletions in the *DMD* gene. Duplications were excluded from this analysis as MLPA results do not provide information about the location of the duplicated exons nor the direction in which they were inserted. According to the GenBank sequence of the dystrophin muscular isoform (Dp427m, NM_004006.3), we determined for each deletion the minimum number of exons, both at the 5′ and 3’ borders of the molecular alteration, that could be skipped in order to restore the reading frame. This analysis was not restricted to the exons targeted by the available or underdevelopment therapies.

Meta-Analysis of Latin America’s Metrics for Duchenne Muscular Dystrophy Available Therapies

In order to determine the frequency of candidate patients for the available therapies for each Latin American country, we performed a systematic review of the literature regarding molecular diagnosis of dystrophinopathies.

The study was carried out following the “Preferred Reporting Items for Systematic reviews and Meta-Analyses” (PRISMA) guidelines (Page et al., 2021). **Figure 1** summarizes the search and selection process. The search was conducted in PubMed from the National Library of Medicine (National Center for Biotechnology Information—NCBI) and in Google Scholar from Google (White, 2020). The following keywords were used to browse in both search engines [(Duchenne muscular dystrophy) OR (DMD) OR (Dystrophinopathies)] AND (Country name/demonym) AND [(mutation) OR (molecular diagnosis)]. We applied no publication date nor language restrictions. The last search was performed on November 30, 2020.

Three authors (LL, MC, and CM) independently carried out the study selection from the retrieved manuscripts. Firstly,

potentially eligible reports were identified by assessing their title and abstract. At this stage, we included cohort studies and thesis conducting molecular diagnosis or mutational screening on patients with clinical presumptive diagnosis of dystrophinopathies. Manuscripts regarding basic research, case reports and congress communications were excluded.

The second eligibility step included the review of the full text. We excluded reports implementing multiplex-PCR, given that the exact deletion borders might not have been determined and that it resembles an underestimation of the amount of deletions. Moreover, so as to calculate the frequencies using a common criteria, we only considered manuscripts including the results of each individual or that had their results submitted on public repositories or databases such as the Leiden open variation database (LOVD) (https://www.lovd.nl/).

Once we had the selected reports, we manually extracted the following information: the employed molecular techniques, total amount of analyzed patients, the amount of individuals with genetic confirmation of dystrophinopathy, unrelated patients carrying deletions, unrelated individuals with deletions amenable by exon skipping of exon 45, 51 or 53, unrelated boys carrying small variants in *DMD* and unrelated patients with nonsense variants.

Additionally, so as to compare the calculated Latin American frequencies with the well-known and highly regarded frequencies from Europe and the United States, we conducted the screening described above with minor modifications for Spain, Italy, Portugal and the United States. We selected the above mentioned European countries on the basis of the most relevant migratory waves of the Latin American history. Granted that these four countries have been providing state of the art molecular diagnosis for dystrophinopathies for many years, the amount of available reports considerably exceeded the Latin American ones. Therefore, we decided to restrict the publication date (2005–2020) and select only one report per country, opting for the latest and/or the one with the largest cohort with available individual data.

RESULTS

Molecular Diagnosis and Selection of Candidate Patients

From the studied cohort, dystrophinopathy clinical diagnosis could be confirmed in 371 from 400 analyzed patients. The employed molecular algorithm, based on the best practice guidelines for genetic testing for dystrophinopathies and the characteristics of each case (familial/sporadic case, known/unknown causative mutation and type of molecular alteration), reached a detection rate of 92.8% (**Figure 2**). Granted that we already had the WES results of the 29 patients without identified mutation in *DMD,* we broadened the screening of pathogenic variants to genes associated with other muscular dystrophies (Group 1). This extended algorithm allowed us to provide a differential diagnosis to other 17 patients and, also, to increase the detection rate to 97%. These patients showed overlapping symptoms with DMD/BMD but turned out to be principally affected by limb-girdle muscular dystrophies, as we found disease causing variants in the following genes: *FKRP* (4), *SGCA* (2), *SGCG* (2), *SGCB* (1), *CAPN3* (1), *FKTN* (1), *POMT2* (2), *SYNE1* (1), *COL6A1* (1), *COL6A3* (1) and *PHKA1* (1). In the remaining 12 patients, we proceeded with the screening of all the other groups of NMD. However, we could not identify any disease associated variants.

Given the broadening of the analysis of WES results, we have identified two patients from our cohort carrying pathogenic or likely pathogenic molecular alterations in two MDs causing genes. Patient one presented an intronic variant in *DMD* (NM_004006.3:c.1332-9A > G), mainly reported as pathogenic in LOVD and ClinVar and probed to affect the splicing process. However, he also carried an heterozygous missense variant in *CAPN3* (NM_000070.2:c.1303G > A; NP_000061.1: p(Glu435Lys), classified as pathogenic in LOVD and as likely pathogenic/pathogenic in ClinVar. Patient two showed two variants in *DMD* and a heterozygous variant in *SYNE1*. In *DMD* gene, not only did he carry a splice site variant (NM_004006.3:c.9975-2A > T) but also a missense molecular alteration (NM_004006.3:c.10010G > A; NP_003997.2: p.Cys3337Tyr), both reported in LOVD database in a single occasion and classified as pathogenic and likely pathogenic, respectively. As for *SYNE1*, he presented a novel frameshift deletion [NM_182961.3:c.7310del; NP_892006.3: p(Gly2437Valfs*6)].

On the other hand, as it was mentioned above, the identification of the disease causative molecular alteration contributes to the selection of the suitable treatment for each individual (**Figure 2**). Firstly, the 371 patients with genetic confirmation of dystrophinopathy resulted candidates for corticosteroid treatment to ameliorate the inflammation and improve muscle strength and function. Moreover, regarding the available mutation-specific therapies for Dystrophinopathy, the precise characterization of the *DMD* mutation allowed us to determine that 20 patients were candidates for exon skipping of exon 51, 21 for exon 53 and 12 for exon 45, while another 70 were eligible for premature stop codon read-through therapy (**Table 1**). Alternatively, molecular diagnosis prevented the unnecessary and ineffective corticosteroid treatment of 17 patients diagnosed with other MDs. However, this differential diagnosis enabled us to determine that six patients were candidates for gene-transfer therapies for LGMD (2 LGMD2D—*SGCA*, 2 LGMD2C—*SGCG*, 1 LGMD2E—*SGCB* and 1 LGMD2A—*CAPN3*).

Argentinian Duchenne Muscular Dystrophy Mutagenic Spectrum and Analysis of Exonic Targets for Exon Skipping

So as to collaborate with the international efforts that aim to determine mutation frequencies from Latin America, we used our results to establish the *DMD* mutagenic spectrum for the Argentinian affected population. As expected, CNVs were the most frequent type of molecular alterations taking place in *DMD*, accounting for 71.5% of cases. Deletions of one or more exons were the major contributors of CNVs, being detected in 56.6% of

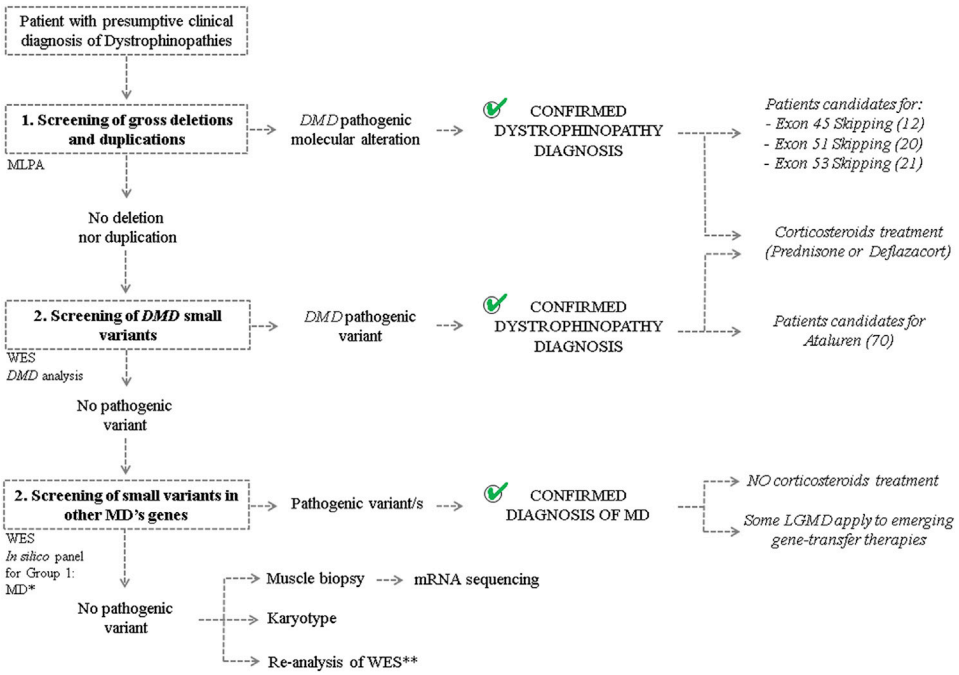


FIGURE 2 | Molecular algorithm The figure shows the workflow carried out to achieve differential diagnosis of patients with clinical suspicion of Dystrophinopathy and to determine candidates for mutation-specific therapies. "MLPA": multiplex ligation-dependent probe amplification. "WES": whole exon sequencing. "MD": Muscular Dystrophy. *In silico* panel was created considering genes associated with the development of MDs (Group 1) according to the "Gene table of neuromuscular disorders (nuclear genome)" and its annual updates. **Re-analysis of the WES results on the basis of the discovery of new genes involved with MDs and the revision of the ACMG classification of sequence variants.

cases, while duplications of one or more exons were found in 14.9%. CNVs were followed by small pathogenic sequence variants, which were identified in 25.4% of cases (**Figure 3A**). According to the classification by effect of the sequence variants, the three types most commonly found were nonsense (42.6%), followed by frameshift (32%) and splice site variants (20.5%) (**Figure 3B**). Furthermore, not only have we detected a small fraction of patients carrying a deletion and a duplication in the same allele, but also some non-contiguous duplications.

On the other hand, we wondered which were the most useful exonic targets for exon skipping in our cohort. To answer this query, we only took into account the 112 out-of-frame deletions identified. The putative single or multiple exonic targets that would restore the reading frame of the 112 deletions are depicted in **Figure 4**. The skipping of only one exon could restore the reading frame of 87.5% of patients, whereas the remaining 12.5% would require multiple-exon skipping. The deletions of 14 patients can even be corrected by two different single-exon skipping strategies, as removing the 5′ or 3' exons adjacent to the alteration can turn it into an in-frame mutation.

As it was expected, most of the identified molecular alterations (47.3%) would restore the reading frame by skipping exons 45, 51 or 53, that is to say that they are eligible for the already available therapies. They are followed by exon 44, which applies to 9.8% of the out-of-frame deletions. It is worth mentioning that exon skipping of exon 44 is currently under preclinical test. Exon skipping strategies targeting 50 and 52, which are also undergoing preclinical trials, only probed to be eligible for 4.5 and 3.4% of the identified mutations, respectively. Strikingly, we found that 5.4% of the mutations from our cohort are amenable for exon skipping of exon 2 or exon 8, as all these patients carried the same deletion of exons 3 to 7 (**Figure 4**).

Characterization of Nonsense Variants

Ataluren is the drug developed to enable ribosomal read-through of premature stop codons in nonsense mutations for Duchenne patients. From the total cohort analyzed we identified 70 patients with nonsense variants, who were candidates for Ataluren drug. We wonder if the subtypes of nonsense variants could affect Ataluren effectiveness, therefore we proposed to characterize the diversity of nonsense in an Argentine cohort. We described the different types of nonsense found, the number of times the same variant was observed in unrelated patients, and the exons and protein domains affected. Furthermore, we analyzed which was the wild-type amino acid that switched to a premature stop codon, and finally, the codon position of the transition/transversion (**Table 1**).

From the 70 patients with nonsense identified, 60 were unrelated. These 60 variants were distributed in 33 of the 79 *DMD* exons, and almost 66% were localized in the dystrophin rod domain. Moreover, in unrelated patients, nonsense mutations were more frequently found in exons 20, 23, 66, and 70. The UGA turned out to be the most frequent premature stop codon observed (47%) and it was in the first

TABLE 1 | Nonsense variants characterization.

Lab identification	Patients*	Nonsense variant (HGVS, c./p.)	Exon	Dys domains	WT aa	WT codón	Stop codon	DNA subst mut
#392	1	c.433C > T/p.(Arg145*)	6	Actin binding	Arg	CGA	**T**GA	Transition
#56	1	c.620T > G/p.(Leu207*)	7	Actin binding	Leu	TTA	T**G**A	Transversion
#598	1	c.701C > G/p.(Ser234*)	8	Actin binding	Ser	TCG	T**G**A	Transversion
#620	1	c.826C > T/p.(Gln276*)	8	Actin binding	Gln	CAA	**T**AA	Transition
#104	1	c.907C > T/p.(Gln303*)	9	Central rod	Gln	CAG	**T**AG	Transition
#586	1	c.1132C > T/p.(Gln378*)	10	Central rod	Gln	CAG	**T**AG	Transition
#246	1	c.1388G > A/p.(Trp463*)	12	Central rod	Trp	TGG	T**A**G	Transition
#362	1	c.1793C > G/p.(Ser598*)	15	Central rod	Ser	TCA	**T**GA	Transversion
#461	1	c.1777C > T/p.(Gln593*)	15	Central rod	Gln	CAA	**T**AA	Transition
#307	1	c.1928G > A/p.(Trp643*)	16	Central rod	Trp	TGG	T**A**G	Transition
#619	1	c.2032C > T/p.(Gln678*)	17	Central rod	Gln	CAG	**T**AG	Transition
#288	1	c.2270C > G/p.(Ser757*)	18	Central rod	Ser	TCA	T**G**A	Transversion
#132	1	c.2317A > T/p.(Lys773*)	19	Central rod	Lys	AAG	**T**AG	Transversion
#110/#725/#824	3	c.2407C > T/p.(Gln803*)	20	Central rod	Gln	CAA	**T**AA	Transition
#326	1	c.3151C > T/p.(Arg1051*)	20	Central rod	Arg	CGA	**T**GA	Transition
#303	1	c.2440G > T/p.(Glu814*)	20	Central rod	Glu	GAA	**T**AA	Transversion
#775	1	c.2566C > T/p.(Gln856*)	20	Central rod	Gln	CAA	**T**AA	Transition
#773	1	c.2414C > G/p.(Ser805*)	20	Central rod	Ser	TCA	T**G**A	Transversion
#723	1	c.2626G > T/p.(Glu876*)	21	Central rod	Glu	GAA	**T**AA	Transversion
#762/#695	2	c.2991C > G/p.(Tyr997*)	23	Central rod	Tyr	TAC	**T**AG	Transversion
#394/#460	2	c.3151C > T/p.(Arg1051*)	23	Central rod	Arg	CGA	**T**GA	Transition
#717	1	c.3136C > T/p.(Gln1046*)	23	Central rod	Gln	CAA	**T**AA	Transition
#125	1	c.3742C > T/p.(Gln1248*)	27	Central rod	Gln	CAG	**T**AG	Transition
#686	1	c.4108C > T/p.(Gln1370*)	30	Central rod	Gln	CAG	**T**AG	Transition
#258	1	c.4375C > T/p.(Arg1459*)	32	Central rod	Arg	CGA	**T**GA	Transition
DMD191	1	c.4499C > A/p.(Ser1500*)	32	Central rod	Ser	TCA	T**A**A	Transversion
#675/#677	2	c.4729C > T/p.(Arg1577*)	34	Central rod	Arg	CGA	**T**GA	Transition
#603	1	c.4820T > A/p.(Leu1607*)	34	Central rod	Leu	TTG	T**A**G	Transversion
#639	1	c.5530C > T/p.(Arg1844*)	39	Central rod	Arg	CGA	**T**GA	Transition
#649	1	c.6254G > A/p.(Trp2085*)	43	Central rod	Trp	TGG	T**A**G	Transition
#710	1	c.6715G > T/p.(Glu2239*)	46	Central rod	Glu	GAA	**T**AA	Transversion
#303/#338	2	c.6973C > T/p.(Gln2325*)	48	Central rod	Gln	CAG	**T**AG	Transition
#769	1	c.7010T > G/p.(Leu2337*)	48	Central rod	Leu	TTA	T**G**A	Transversion
#774	1	c.7657C > T/p.(Arg2553*)	52	Central rod	Arg	CGA	**T**GA	Transition
#465	1	c.7792C > T/p.(Gln2598*)	53	Central rod	Gln	CAG	**T**AG	Transition
#508	1	c.7750C > T/p.(Gln2584*)	53	Central rod	Gln	CAA	**T**AA	Transition
#689	1	c.8098A > T/p.(Lys2700*)	55	Central rod	Lys	AAG	**T**AG	Transversion
#285	1	c.8608C > T/p.(Arg2870*)	58	Central rod	Arg	CGA	**T**GA	Transition
#623	1	c.8774G > A/p.(Trp2925*)	59	Central rod	Trp	TGG	T**A**G	Transition
#483	1	c.8944C > T/p.(Arg2982*)	60	Central rod	Arg	CGA	**T**GA	Transition
#295	1	c.9337C > T/p.(Arg3113*)	64	Cysteine-rich	Arg	CGA	**T**GA	Transition
#194	1	c.9459T > A/p.(Cys3153*)	65	Cysteine-rich	Cys	TGT	TG**A**	Transversion
#673	1	c.9474T > G/p.(Tyr3158*)	65	Cysteine-rich	Tyr	TAT	TA**G**	Transversion
#542/#495/#700	3	c.9568C > T/p.(Arg3190*)	66	Cysteine-rich	Arg	CGA	**T**GA	Transition
#617	1	c.9802C > T/p.(Gln3268*)	67	Cysteine-rich	Gln	CAA	**T**AA	Transition
#196	1	c.9928C > T/p.(Gln3310*)	68	Cysteine-rich	Gln	CAG	**T**AG	Transition
#469/#437/#753/#846	4	c.10108C > T/p.(Arg3370*)	70	Cysteine-rich	Arg	CGA	**T**GA	Transition
#250	1	c.10141C > T/p.(Arg3381*)	70	Carboxy-terminal	Arg	CGA	**T**GA	Transition
#854	1	c.10171C > T/(p.Arg3391*)	70	Carboxy-terminal	Arg	CGA	**T**GA	Transition

Lab, laboratory; Patients, number of non-related patients with the same nonsense; Nonsense variant (HGVS, c./p.), HGVS-nomenclature (https://varnomen.hgvs.org/); p, (protein); c, (coding DNA) (Dp427m, NM_004006.3); WT aa, Wild type amino acid; WT Codón, Wild type codon, in bold the base implicated in the substitution; Dys Domains, dystrophin domains; DNA subst mut, DNA substitution mutations.

position of the codon that 73% of the substitutions occurred, being 2.2 times more frequent transitions than transversions. It is important to highlight that 35% of the 60 unrelated patients carried nonsense mutations involving Arginine and 30% involving Glutamine. Both of them were caused by transitions, and all mutations in the case of arginine, were CGA codons. Of the 20 amino acids encoded in our genetic code, only ten can convert to a stop codon by mutations. We observed nine of them present in our cohort, only glycine was not found (**Table 1**).

Meta-Analysis of Latin America's Metrics for Duchenne Muscular Dystrophy Available Therapies

A total of seven reports (six published manuscripts and a thesis) matched the inclusion/exclusion criteria used for the systematic review of the literature, representing only six Latin American countries (**Table 2**). Furthermore, in order to compare the information from Latin America with the highly regarded

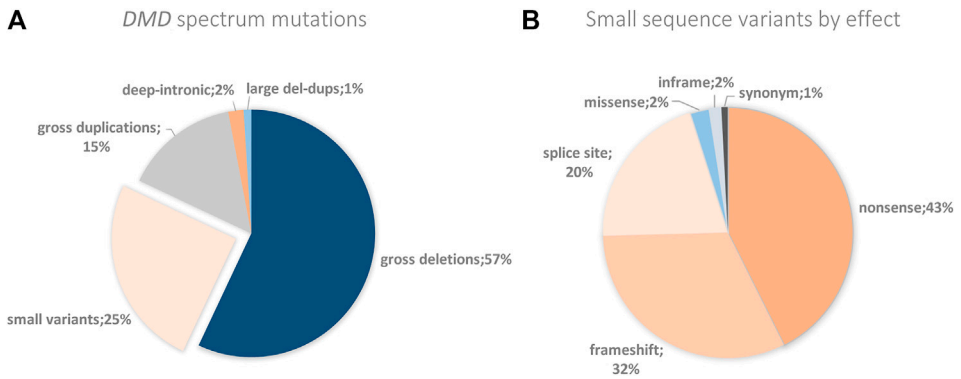


FIGURE 3 | DMD spectrum mutations and small variants by effect. **(A)** The figure shows the *DMD* percentages of the different genetic alterations found in the Argentine cohort. **(B)** *DMD* percentages of the small variants by their effect found in the Argentine cohort.

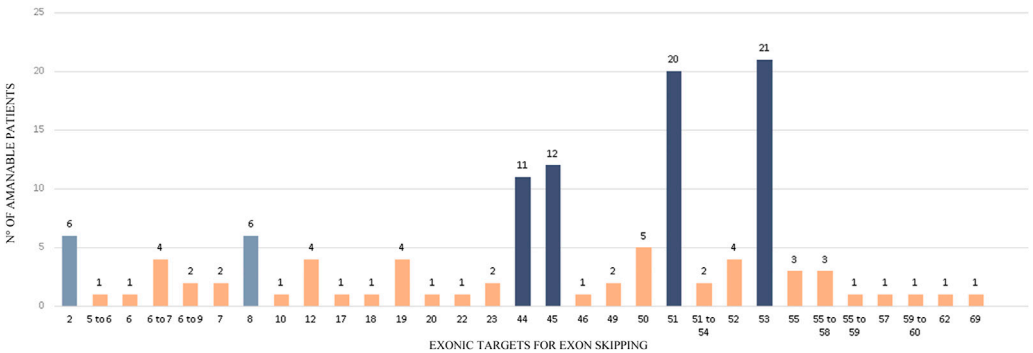


FIGURE 4 | Exonic targets for Exon skipping. The figure shows the targets for exon skipping that could restore the reading frame of a subset of 112 patients carrying out-of-frame deletions in *DMD*. 14 mutations can be corrected by two different Exon Skipping strategies.

knowledge from Europe and North America, manuscripts from Spain, Italy, Portugal and the United States were included in our meta-analysis (**Table 2**). Unfortunately, though the manuscript from Portugal included a large cohort and performed the complete diagnostic algorithm, the data submitted in LOVD was incomplete and showed a bias toward small sequence variants.

Strikingly, according to the gathered reports, only four Latin American countries (Argentina, Brazil, Colombia and Mexico) conduct the screening of *DMD* small sequence variants, that is to say, the complete molecular diagnostic algorithm for dystrophinopathies (CNV and SNV). Thus, most of the Latin American countries only provide studies to detect CNVs. As it was part of the inclusion criteria, only these seven manuscripts report the usage of MLPA, the rest of the available Latin American literature in PubMed and Google Scholar not only were they more than 15 years old but also entailed studies of multiplex-PCR to detect deletions in the *DMD* gene. On the other hand, information about the age at molecular diagnosis was only provided for three countries, being the earliest the Argentinian at the age of 6 years old and the latest the Colombian at the age of 9 years (**Table 2**).

Regarding the frequencies of exon skipping, as in the manuscripts they were estimated in alternative ways, we standardized the calculations as: total amount of unrelated patients amenable for each exon skipping/total amount of unrelated patients with *DMD* deletions. We observed different relations among the three targets for exon skipping in the analyzed countries. As it is generally reported in the literature, the pattern E51S > E53S > E45S was detected in Argentina and Puerto Rico, which also coincided with the Spanish results. Alternatively, Colombia and Costa Rica showed the pattern E51S > E45S > E53S, which was shared with the United States. In addition, exon 45 proved to be the most frequent target for exon skipping in Brazil (E45S > E53S > E51S) and Peru (E45S > E51S > E53S), but they presented different relations for exons 51 and 53. Notably, Mexico exhibited the same frequency for the three targets (E45S = E51S = E53S). Similarly, Italy depicted almost the same frequency for exon 51 and 53 skipping, yet a reduced proportion for exon 45 (E51S ≈ E53S > E45S).

Concerning the premature stop codon read-through therapy, we calculated the proportion of nonsense variants as follows: total amount of unrelated patients with nonsense

TABLE 2 | Meta-analysis of Latin America's metrics for DMD available therapies.

Country	Total/diagnosed cases[a]	Employed techniques	AD	Exon 45 skipping (E45S)[b]	Exon 51 skipping (E51S)[b]	Exon 53 skipping (E53S)[b]	Ataluren[c]	Reference
Argentina	371/359	MLPA and *DMD* seq. by NGS	6 years	9/155 (5.8%)	20/155 (12.9%)	18/155 (11.6%)	60/143 (42%)	Present work
Brazil	177/177	MLPA and *DMD* seq. by NGS	—	19/103 (18.4%)	11/103 (10.7%)	13/103 (12.6%)	22/52 (42.3%)	de Almeida et al. (2017)
Peru	40/21	PCRm and MLPA	—	4/17 (23.5%)	3/17 (17.6%)	2/17 (11.8%)	—	Huaman-Dianderas et al. (2019)
Colombia	52/52	MLPA and *DMD* seq. by NGS	9 years	1/28 (3.6%)	2/28 (7.1%)	0/28 (0%)	6/11 (54.5%)	García-Acero et al. (2018)
Costa Rica	74/53	PCRm and MLPA	7.5 years	7/45 (15.6%)	8/45 (17.8%)	1/45 (2.2%)	—	Thesis[3]
Puerto Rico	84/65	MLPA	—	1/56 (1.8%)	9/56 (16.1%)	5/56 (8.9%)	—	Ramos et al. (2016)
Mexico	170/116	MLPA, PM-MLPA, HRM and sanger seq	—	11/86 (12,8%)[d]	11/86 (12,8%)[d]	11/86 (12,8%)[d]	—	López-Hernández et al. (2015)
	63/52	MLPA, NGS and sanger seq	—	—	—	—	11/29 (37.9%)	Alcántara-Ortigoza et al. (2019)[§]
Spain	284/284	PCRm, MLPA and sanger seq	—	11/131 (8.4%)	16/131 (12.2%)	8/131 (6.1%)	49/97 (50.5%)	Vieitez et al. (2017)
Italy	1902/1902	PCRm, Log-PCR, MLPA, NGS and sanger seq	—	39/610 (6.4%)[d]	51/610 (8.4%)[d]	53/610 (8.7%)[d]	200/469 (42.6%)	Neri et al. (2020)
Portugal	312/284	Southern blot, PCRm, MLPA and sanger seq	—	0/11 (0%)[d]	1/11 (9.1%)[d]	0/11 (0%)[d]	5/25 (20%)	Santos et al. (2014)
EUA	933/933	SCAIP, MLPA, sanger seq and cDNA seq	—	53/426 (12.4%)	70/426 (16.4%)	50/426 (11.7%)	226/400 (56.5%)	Flanigan et al. (2009)

[a]*Tslrotal of molecularly analyzed/diagnosed unrelated male patients, members of a family were counted as 1 case; AD: Mean age at diagnosis*

[b]*Percentage of deletions amenable with exon skipping of exon 45, 51, and 53, respectively. Calculations were performed as follows: N° of unrelated patients amenable for each of the therapies/N° of unrelated patients carrying deletions*

[c]*Percentage of* DMD *small variants candidates for Ataluren or premature stop codon read through. Calculations were performed considering: N° of unrelated patients carrying nonsense variants/N° of unrelated patients with sequence variants. N/A: Data not available.*

[d]*Frequencies were determined on the basis of the patients reported on LOVD and linked to the manuscript.*

variants/total amount of unrelated individuals with *DMD* small molecular alterations. The proportions of nonsense variants for Latin America were mainly around 40%, which correlated with the results from the Spanish and Italian populations. Colombia exhibited the greater amount of nonsense mutations (54%) among the Latin American countries and was similar to the 56.5% observed in the United States.

DISCUSSION

Dystrophinopathies cover a spectrum of X-linked muscle diseases ranging from mild to severe phenotypes. Although rare, they are among the most common pediatric muscular dystrophies, being DMD the most prevalent and severe form. Luckily, in the last two decades, unprecedented advances have been made in the field of drug development for rare diseases and DMD is a great example. Nowadays, not only are corticosteroids available, but also mutation-dependent therapies aiming to generate a functional dystrophin mRNA and/or protein. These major advances have turned the molecular diagnosis of dystrophinopathies into a key element for the selection of the best standard of care. In other words, screening and precise characterization of the *DMD* causative mutation is now the basis of the theragnosis for these diseases.

In the present work, we pursued differential molecular diagnosis of 400 Argentinian patients with presumptive clinical diagnosis of dystrophinopathy, so as to determine eligibility for the available therapies. For this, as it was mentioned above, we set up a general diagnostic algorithm following the best practice guidelines for genetic testing for dystrophinopathies (Birnkrant et al., 2018; Fratter et al., 2020), which was tailored on the basis of the particular characteristics of each case. Moreover, this strategy was improved by taking into consideration the possibility of mistaken clinical diagnosis given the existence of overlapping phenotypic features with other types of MDs. Here we have shown that, at least for the Argentinian analyzed cohort, the employed algorithm is highly effective for the detection of the disease causing molecular alterations and for the achievement of differential diagnosis, as we reached a detection rate of 97%.

Dystrophinopathy clinical diagnosis was confirmed by genetic testing in 371/400 (92.8%) patients. Apart from being the foundations of familial genetic assessment, molecular diagnosis plays a key role in the selection of the suitable standard of care for each individual (Birnkrant et al., 2018). Corticosteroids are the recommended standard therapy for DMD, thus molecular confirmation of the diagnosis is essential for starting treatment. Therefore in these 371 DMD confirmed patients, corticosteroid treatment was correctly

indicated and validated. The addition of corticosteroids in the standards of care for dystrophinopathy was subject of great debate given their well-known side effects. However, they have extensively proved to ameliorate the inflammation and improve muscle strength, which translates in a delay at the age of loss of ambulation. This is the main reason why their benefits are thought to surpass their side effects. Nowadays, the discussion relies on which corticosteroid is the best for dystrophinopathy patients, Prednisone or Deflazacort (Griggs et al., 2016; Shieh et al., 2018).

On the one hand, the employed molecular algorithm was able to identify 17 clinically misdiagnosed patients. While most of these patients were affected by recessive or dominant forms of limb-girdle muscular dystrophies, one of them had a pathogenic variant in *PHKA1* associated with the X-linked form of muscle glycogenosis. Individuals suffering from the autosomal recessive LGMD2I (*FKRP*) were the most frequently misdiagnosed as DMD. This large proportion of mistaken clinical diagnosis relies on the overlapping signs and symptoms among MDs and the requirement of an experienced neurologist/physician to detect the characteristic features of each clinical picture. Even the highly regarded muscle biopsy is no longer considered as an unequivocal diagnostic test, as it has been shown that patients with absence or decreased levels of dystrophin in the immunohistochemistry can carry molecular alterations in genes encoding for dystrophin-related proteins (Yamamoto et al., 2008). Moreover, it must be highlighted that if we would have followed the recommended molecular algorithm for dystrophinopathies without modifications, the 29 individuals not carrying pathogenic variants in *DMD* would have undergone muscle biopsy. This, therefore, shows that when WES or NGS panels results are available, broadening the screening to genes linked to the development of NMDs is useful not only to reach the differential diagnosis without further tests but also to prevent patients from going through the invasive biopsy procedure.

On the other hand, providing differential diagnosis to the 17 individuals with other forms of MDs was also useful to determine the suitable standard of care and eligibility on gene-transfer therapies. They were prevented from a corticosteroid treatment, as for these diseases there is still no strong evidence for its efficacy (Walter et al., 2013; Albuquerque et al., 2014). Furthermore, six of them could already be determined as candidates for gene-transfer therapies that are under preclinical or clinical tests. Even though these are gene-dependent therapeutic approaches, counting with the precise characterization of the disease causative mutations will allow them to start with these treatments as soon as they are approved and become available, so as to prevent further muscle deterioration.

Given the usage of WES for the screening of small variants, we have identified two patients carrying likely pathogenic/pathogenic variants in *DMD* and in other MDs genes (*CAPN3* and *SYNE1*) associated with both dominant and recessive inheritance patterns. One of them even carried a pathogenic variant and a VUS in *DMD*, being the latter reported as likely pathogenic in LOVD. We are observing this type of cases more frequently as NGS tests are becoming more broadly used for genetic diagnosis. On the one side, these findings highlight the importance of performing a deep genetic analysis so as to provide an accurate theragnosis, given that the selected treatment might not be as effective as it should because of the existence of a second deleterious molecular alteration further affecting the skeletal muscle. On the other side, these observations could explain the phenotypic heterogeneity among dystrophinopathy patients.

Regarding *DMD* mutational spectrum, the observed proportion of CNVs and small sequence variants were in accordance with what was reported in literature for the European and North American populations (Flanigan et al., 2009; Falzarano et al., 2015; Aartsma-Rus et al., 2016; Vieitez et al., 2017). We even detected 0.7% of large allelic del-dups. Notably, ~7% of the identified gross duplications were non-contiguous alterations.

In particular, 5% of the patients with single exon deletions detected by MLPA, actually had small sequence variants affecting the hybridization of the hemiprobes. Despite the fact of being included in the best practice guidelines for genetic testing for dystrophinopathies, at least in our country, corroboration of single exon deletions with an alternative technique is not the norm for every laboratory. Yet, this has a huge impact for the patient, as a mistaken diagnosis can affect management and theragnosis.

As for the 12 individuals with clinical presumptive diagnosis of dystrophinopathy but without identified pathogenic mutation, we could determine that five of them had a biopsy with immunohistochemistry compatible with dystrophinopathy (dystrophin deficient or absent). Therefore, we presumed they could be dystrophinopathy cases with deep intronic alterations, chromosomal rearrangements or regulatory mutations not detected by the employed methodology (Gurvich et al., 2008; Tran et al., 2013; Aartsma-Rus et al., 2016). Another seven patients had highly increased CK levels (ranging from 1.600 to 18.000 UI/L), however they did not have a biopsy. In conclusion, to continue with the study of these undiagnosed patients complementary studies will be necessary. A muscle biopsy could be valuable to perform Immunohistochemistry analysis, western blot and/or mRNA sequencing. Other options are MRI, electromyography and other genetics tests that analyze genes that were not included in the NMD gene table (new NMD-associated genes) or genes carrying variants not detected by the employed NGS pipeline. Such are the cases of facioscapulohumeral muscular dystrophy (FSHD) or type three spinal muscular atrophy (SMA).

Regarding exon skipping, the selection of candidate patients is generally performed considering the exonic borders of the observed deletion at gDNA level and determining if the removal of one or several of the surrounding exons would restore the reading frame. This way we could establish that 53 of the dystrophinopathy patients apply for the available strategies targeting exons 45, 51, and 53. However, it should be noticed that as the cDNA is not evaluated for these patients, we do not know if the molecular alteration at the mRNA level resembles the one identified in the gDNA. That is to say, it cannot be excluded that, given the location of the intronic breakpoints of the deletion, non-

canonical regulatory splicing sites might be altering the mRNA processing. Hence, as this could modify the effect of the exon skipping therapy and make patients go through an ineffective treatment, taking this matter into consideration, would further improve the selection of candidate patients (Gualandi et al., 2006; Tuffery-Giraud et al., 2017).

As for the screening of putative exonic targets for the identified out-of-frame deletions, we observed that the majority of patients (87.5%) would require a mono-target exon skipping strategy to restore their reading frame. Luckily, this agrees with the knowledge obtained from the development of different exons skipping mechanisms, as the mono-target therapies have reached more fruitful results than the multi-target ones. This is related to the great difficulty with the delivery of the required chemically modified AONs, turning the task more laborious as the amount of targets increases (Aslesh et al., 2018; Echigoya et al., 2019). Furthermore, among the most useful exonic targets, we found exons 44, 50 and 52 (~17.8%) which are already undergoing preclinical trials[2].

Also, for 14 patients, we identified two different single-exon skipping strategies capable of restoring the reading frame. In such cases the question is which would be better or would be more effective for the patients. We will try to answer this query using as an example the deletion of exons 3 to 7 which is amenable for exon skipping of exon 2 and exon 8. As far as we are concerned, the decision should be made on the basis of the location of the molecular alteration, the role of the implicated area in the protein, information of gene/protein structure and reports of patients having the resulting deletion and their phenotype or clinical course of the disease. In this case, although the deletion affects the actin-binding domain, it is known that there is another actin anchorage site within the rod domain (Mias-Lucquin et al., 2020). Moreover, it has been reported the existence of three Internal Ribosome Entry Sites (IRES) or, in other words, three internal in-frame start codons of the translations in exon 8 (Malhotra et al., 1988; Winnard et al., 1995). In addition, LOVD counted with five reports of patients with deletion of exons 2 to 7 (three classified as BMD, one as DMD and 1 DMD/BMD), while there were only three reports of the deletion of exons 3 to 8 (1 BMD and 2 MD). Finally, according to the provided information, we think that in this case the exon skipping of exon 2 would show better results.

Concerning premature stop codon read-through therapy, given the fact that nonsense variants are generally considered as truncating alterations, candidate patients are selected under the simple consideration of presenting this type of small sequence variants. Following this criteria, we identified in our cohort 70 individuals eligible for the treatment with Ataluren. Nonetheless, this year it was reported that not every *DMD* nonsense variant should be rendered as truncating nor strictly associated with DMD, as due to their genetic/exonic location they could be actually having a milder effect on the phenotype (Neri et al., 2020; Torella et al., 2020).

Granted the existence of a therapeutic protocol for nonsense variants, it is of the utmost importance to deepen the knowledge of the effect of this type of alterations and their characterization. In our cohort, we have found 60 nonsense in unrelated patients, distributed in 33 of the 79 *DMD* exons and mainly affecting the rod domain (65.6%) of dystrophin. These substitutions principally took place in the first codon position (73%), followed by 19% affecting the second and 8% the third position, mostly disrupting codons coding for arginine and glutamine. Also, it was observed that transitions occurred 2.2 times more frequent than transversions, being C > T substitution the most prevalent. Finally, the stop codon generation rate was: UGA (46.9%), UAG (32.8%) and UAA (20.3%).

As expected, G:C > A:T transitions were the most prevalent stop mutation class (72%), we obtained results similar to Flanigan et al., 2009. We also found arginine as the most frequent amino acid converted to stop. From the total of the 60 substitutions, the 35% were transitions due to CpG from arginine (CGA) to Stop, presumably due to the spontaneous deamination of 5-methylcytosine to thymidine at methylated CpG dinucleotides (Cooper and Krawczak, 1989; Flanigan et al., 2009).

Although up to now nonsense mutations are treated with Ataluren, it is possible that in a short term, combined therapies will begin to be implemented. For example, 13 of the 60 nonsense (exons: 9, 10, 15, 16, 23, 27, 30, 32, 34, 39, 48, 60, and 64) are located in inframe exons, these exons could be used as targets for exon skipping and be combined with the premature stop codon read-through therapy.

Another important point to take into account in the trials is that the populations (controls vs. treated) to be compared should be as homogeneous as possible in terms of nonsense variants. Beyond the fact that most nonsense are considered null alleles, it must be taken into account that their location in the *DMD* gene will have different functional impact since it will depend on how many isoforms will be affected. In our case, patients who have nonsense after exon 60 will theoretically have all the isoforms of dystrophin affected. On the other hand, as more 5′ the nonsense would be located, less isoforms would be altered. Regarding the above mentioned nonsense characteristics, although 10 patients are currently being treated with Ataluren, unfortunately, we do not have enough clinical data to make genotype-phenotype statistical comparisons to determine the effectiveness of the treatment.

In the matter of the performed meta-analysis, it surprised us the lack of updated reports about the genetic and molecular characterization of dystrophinopathy patients from Latin America. Moreover, it is alarming the reduced number of countries providing the *DMD* full mutational analysis to the affected individuals, which, as it was highlighted throughout the manuscript, is now considered the foundations of the theragnosis for dystrophinopathies. However, we cannot discard the chance that these studies are in fact carried out but the results are not shared in publications. Another possibility is that the studies are performed abroad, as it is well-known that for developing countries sometimes it is cost-effective to send the samples to experienced and equipped laboratories.

In addition, it was difficult to obtain information regarding the age at molecular diagnosis, though manuscripts more frequently

[2]https://www.sareptatherapeutics.ch/en/our-pipeline.

detailed the age at onset. Nonetheless, we reckon that the age at molecular diagnosis is more objective and precise than the age at onset, as the latter depends on the ability of the family to detect the appearance of the first symptoms and the capacity of the clinician to get the information during anamnesis. On the other hand, the age at molecular diagnosis provides insights about the availability of genetic studies in a certain place. It must be noticed that an early confirmation of the presumptive clinical diagnosis has a major impact on the management and treatment of the patient, as this translates into a major amount of healthy muscle fibers to treat. From the gathered data, Argentina presented the earliest age at molecular diagnosis (6 years) while Colombia showed the latest one (9 years). So, efforts must be done in Latin America to achieve early diagnosis in dystrophinopathy patients.

As for exon skipping, not only has the meta-analysis depicted an ample variability in the frequencies for each target exon, but also has demonstrated the existence of different patterns among them. This was true for the Latin American and European countries, proving that there is no such thing as a general continental pattern. Furthermore, as it has been reported for the Italian population, there can even exist differences in the exon skipping frequencies within a country (Neri et al., 2020). Hence, these results suggest that the selection of target exons for the development of exon skipping therapies, based on frequencies rendered as "general" in the literature, might not be the best approach. As local/ethnic differences are not being considered, many dystrophinopathy patients carrying frequent molecular alterations might miss the opportunity to access a mutation-dependent therapy suitable for them.

On the other hand, the obtained proportions of nonsense variants were similar in the analyzed countries, not even displaying differences between continents. The majority of the countries presented frequencies around 40%. However, Colombia and the United States showed an increased proportion of nonsense mutations, reaching ~55%. Thus, this information suggests an even distribution of this type of small variants among different populations.

In conclusion, the present manuscript describes the theragnosis carried out in one of the reference centers for the molecular diagnosis of dystrophinopathies in Argentina. Firstly, the implemented diagnostic molecular algorithm proved to be efficient for the achievement of differential diagnosis, which nowadays plays a crucial role in patient management, the determination of the suitable standard of care and genetic counseling. Secondly, we have performed a thorough characterization of the *DMD* molecular alterations and, particularly, of the nonsense variants observed in an Argentinian dystrophinopathy cohort. Thirdly, we conducted a meta-analysis that allowed us to compare the frequencies of the amenable mutations for the available DMD therapies and the current situation of the dystrophinopathy molecular studies throughout Latin America. Finally, this work contributes with the international efforts to characterize the frequencies and variants of Latin America, pillars of drug development and theragnosis.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in the LOVD database, accession link: https://databases.lovd.nl/shared/variants/in_gene#object_id=Transcript%2CVariantOnTranscript%2CVariantOnGenome&id=0&order=geneid%2CASC&search_VariantOnGenome/Reference=luce%202021&page_size=100&page=1.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Comité de Ética en Investigación Clínica (CEIC), Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Argentina. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

FG designed the study, raised funding and assured the general supervision of the research group; LL, MC, and CM performed the genetic studies; PB and VD contributed with graphic design and variant classification; LM, AD, and CJ provided clinical assessment; FG and LL wrote the manuscript. All the authors have read, edited and approved the final version of the manuscript.

FUNDING

This study was supported by grants from the University of Buenos Aires (UBACyT 2016—N° 20020150100058BA) and PTC Therapeutics. The sponsor did not interfere with study design; collection, analysis and interpretation of data; writing of the manuscript and in the decision to submit the present work for publication.

ACKNOWLEDGMENTS

We want to thank the support and help provided by the neurologists, patients and their families from our country and the Asociación Distrofia Muscular para las Enfermedades Neuromusculares de Argentina (ADM).

REFERENCES

Aartsma-Rus, A., Ginjaar, I. B., and Bushby, K. (2016). The Importance of Genetic Diagnosis for Duchenne Muscular Dystrophy. *J. Med. Genet.* 53, 145–151. doi:10.1136/jmedgenet-2015-103387

Aartsma-Rus, A., Straub, V., Hemmings, R., Haas, M., Schlosser-Weber, G., Stoyanova-Beninska, V., et al. (2017). Development of Exon Skipping Therapies for Duchenne Muscular Dystrophy: A Critical Review and a Perspective on the Outstanding Issues. *Nucleic Acid Ther.* 27, 251–259. doi:10.1089/nat.2017.0682

Albuquerque, M. A. V., Abath-Neto, O., Maximino, J. R., Chadi, G., Zanoteli, E., and Reed, U. C. (2014). Clinical Aspects of Patients with Sarcoglycanopathies under Steroids Therapy. *Arq. Neuro-psiquiatr.* 72, 768–772. doi:10.1590/0004-282x20140126

Alcántara-Ortigoza, M. A., Reyna-Fabián, M. E., González-del Angel, A., Estandia-Ortega, B., Bermúdez-López, C., Cruz-Miranda, G. M., et al. (2019). Predominance of Dystrophinopathy Genotypes in Mexican Male Patients Presenting as Muscular Dystrophy with A Normal Multiplex Polymerase Chain Reaction DMD Gene Result: A Study Including Targeted Next-Generation Sequencing. *Genes* 10, 856. doi:10.3390/genes10110856

Aslesh, T., Maruyama, R., and Yokota, T. (2018). Skipping Multiple Exons to Treat DMD-Promises and Challenges. *Biomedicines* 6, 1. doi:10.3390/biomedicines6010001

Bello, L., Gordish-Dressman, H., Morgenroth, L. P., Henricson, E. K., Duong, T., Hoffman, E. P., et al. (2015). Prednisone/prednisolone and Deflazacort Regimens in the CINRG Duchenne Natural History Study. *Neurology* 85, 1048–1055. doi:10.1212/wnl.0000000000001950

Benarroch, L., Bonne, G., Rivier, F., and Hamroun, D. (2019). The 2020 Version of the Gene Table of Neuromuscular Disorders (Nuclear Genome). *Neuromuscul. Disord.* 29, 980–1018. doi:10.1016/j.nmd.2019.10.010

Birnkrant, D. J., Bushby, K., Bann, C. M., Apkon, S. D., Blackwell, A., Brumbaugh, D., et al. (2018). Diagnosis and Management of Duchenne Muscular Dystrophy, Part 1: Diagnosis, and Neuromuscular, Rehabilitation, Endocrine, and Gastrointestinal and Nutritional Management. *Lancet Neurol.* 17, 251–267. doi:10.1016/s1474-4422(18)30024-3

Brandsema, J., and Darras, B. (2015). Dystrophinopathies. *Semin. Neurol.* 35, 369–384. doi:10.1055/s-0035-1558982

Bushby, K., Finkel, R., Birnkrant, D. J., Case, L. E., Clemens, P. R., Cripe, L., et al. (2010). Diagnosis and Management of Duchenne Muscular Dystrophy, Part 1: Diagnosis, and Pharmacological and Psychosocial Management. *Lancet Neurol.* 9, 77–93. doi:10.1016/s1474-4422(09)70271-6

Chu, M. L., and Moran, E. (2018). The Limb-Girdle Muscular Dystrophies: Is Treatment on the Horizon?. *Neurotherapeutics* 15, 849–862. doi:10.1007/s13311-018-0648-x

Cooper, D. N., and Krawczak, M. (1989). Cytosine Methylation and the Fate of CpG Dinucleotides in Vertebrate Genomes. *Hum. Genet.* 83, 181–188. doi:10.1007/bf00286715

Darras, B. T., Urion, D. K., and Ghosh, P. S. (2000). "Dystrophinopathies" in *GeneReviews*. Editors M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, et al. (Seattle (WA): University of Washington, Seattle).

de Almeida, P. A. D., Machado-Costa, M. C., Manzoli, G. N., Ferreira, L. S., Rodrigues, M. C. S., Bueno, L. S. M., et al. (2017). Genetic Profile of Brazilian Patients with Dystrophinopathies. *Clin. Genet.* 92, 199–203. doi:10.1111/cge.12975

den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., et al. (2016). HGVS Recommendations for the Description of Sequence Variants: 2016 Update. *Hum. Mutat.* 37, 564–569. doi:10.1002/humu.22981

Duchêne, B. L., Cherif, K., Iyombe-Engembe, J.-P., Guyon, A., Rousseau, J., Ouellet, D. L., et al. (2018). CRISPR-induced Deletion with SaCas9 Restores Dystrophin Expression in Dystrophic Models *In Vitro* and *In Vivo*. *Mol. Ther.* 26, 2604–2616. doi:10.1016/j.ymthe.2018.08.010

Echigoya, Y., Lim, K. R. Q., Melo, D., Bao, B., Trieu, N., Mizobe, Y., et al. (2019). Exons 45-55 Skipping Using Mutation-Tailored Cocktails of Antisense Morpholinos in the DMD Gene. *Mol. Ther.* 27, 2005–2017. doi:10.1016/j.ymthe.2019.07.012

Falzarano, M., Scotton, C., Passarelli, C., and Ferlini, A. (2015). Duchenne Muscular Dystrophy: From Diagnosis to Therapy. *Molecules* 20, 18168–18184. doi:10.3390/molecules201018168

Flanigan, K. M., Dunn, D. M., von Niederhausern, A., Soltanzadeh, P., Gappmaier, E., Howard, M. T., et al. (2009). Mutational Spectrum of DMD Mutations in Dystrophinopathy Patients: Application of Modern Diagnostic Techniques to a Large Cohort. *Hum. Mutat.* 30, 1657–1666. doi:10.1002/humu.21114

Fratter, C., Dalgleish, R., Allen, S. K., Santos, R., Abbs, S., Tuffery-Giraud, S., et al. (2020). EMQN Best Practice Guidelines for Genetic Testing in Dystrophinopathies. *Eur. J. Hum. Genet.* 28, 1141–1159. doi:10.1038/s41431-020-0643-7

García-Acero, M., Pineda, T., Guerra-Torres, M., García-Robles, R., Ayala-Ramírez, P., Buitrago, T., et al. (2018). Análisis del espectro mutacional de la distrofia muscular de Duchenne en un grupo de pacientes colombianos. *Neurología Argentina* 10, 137–146. doi:10.1016/j.neuarg.2018.05.001

Gatta, V., Scarciolla, O., Gaspari, A. R., Palka, C., De Angelis, M. V., Di Muzio, A., et al. (2005). Identification of Deletions and Duplications of the DMD Gene in Affected Males and Carrier Females by Multiple Ligation Probe Amplification (MLPA). *Hum. Genet.* 117, 92–98. doi:10.1007/s00439-005-1270-7

Griggs, R. C., Miller, J. P., Greenberg, C. R., Fehlings, D. L., Pestronk, A., Mendell, J. R., et al. (2016). Efficacy and Safety of Deflazacort vs Prednisone and Placebo for Duchenne Muscular Dystrophy. *Neurology* 87, 2123–2131. doi:10.1212/wnl.0000000000003217

Gualandi, F., Rimessi, P., Trabanelli, C., Spitali, P., Neri, M., Patarnello, T., et al. (2006). Intronic Breakpoint Definition and Transcription Analysis in DMD/BMD Patients with Deletion/duplication at the 5′ Mutation Hot Spot of the Dystrophin Gene. *Gene* 370, 26–33. doi:10.1016/j.gene.2005.11.002

Guiraud, S., Aartsma-Rus, A., Vieira, N. M., Davies, K. E., van Ommen, G.-J. B., and Kunkel, L. M. (2015). The Pathogenesis and Therapy of Muscular Dystrophies. *Annu. Rev. Genom. Hum. Genet.* 16, 281–308. doi:10.1146/annurev-genom-090314-025003

Gurvich, O. L., Tuohy, T. M., Howard, M. T., Finkel, R. S., Medne, L., Anderson, C. B., et al. (2008). DMDpseudoexon Mutations: Splicing Efficiency, Phenotype, and Potential Therapy. *Ann. Neurol.* 63, 81–89. doi:10.1002/ana.21290

Haas, M., Vlcek, V., Balabanov, P., Salmonson, T., Bakchine, S., Markey, G., et al. (2015). European Medicines Agency Review of Ataluren for the Treatment of Ambulant Patients Aged 5 Years and Older with Duchenne Muscular Dystrophy Resulting from a Nonsense Mutation in the Dystrophin Gene. *Neuromuscul. Disord.* 25, 5–13. doi:10.1016/j.nmd.2014.11.011

Hoffman, E. P., Brown, R. H., Jr, and Kunkel, L. M. (1987). Dystrophin: the Protein Product of the Duchenne Muscular Dystrophy Locus. *Cell* 51, 919–928. doi:10.1016/0092-8674(87)90579-4

Huaman-Dianderas, F. D., Guevara-Fujita, M. L., Rojas Málaga, D., Estrada-Cuzcano, A., and Fujita, R. (2019). Detección de mutaciones causantes de distrofia muscular de Duchenne/Becker: reacción en cadena de la polimerasa multiplex vs. amplificación múltiple dependiente de ligación por sondas. *Rev. Peru Med. Exp. Salud Publica* 36, 475. doi:10.17843/rpmesp.2019.363.4085

Janssen, B., Hartmann, C., Scholz, V., Jauch, A., and Zschocke, J. (2005). MLPA Analysis for the Detection of Deletions, Duplications and Complex Rearrangements in the Dystrophin Gene: Potential and Pitfalls. *Neurogenetics* 6, 29–35. doi:10.1007/s10048-004-0204-1

Keeling, K. M., and Bedwell, D. M. (2011). Suppression of Nonsense Mutations as a Therapeutic Approach to Treat Genetic Diseases. *WIREs RNA* 2, 837–852. doi:10.1002/wrna.95

Kinane, T. B., Mayer, O. H., Duda, P. W., Lowes, L. P., Moody, S. L., and Mendell, J. R. (2018). Long-Term Pulmonary Function in Duchenne Muscular Dystrophy: Comparison of Eteplirsen-Treated Patients to Natural History. *J. Neuromuscul. Dis.* 5, 47–58. doi:10.3233/jnd-170272

Koenig, M., Monaco, A. P., and Kunkel, L. M. (1988). The Complete Sequence of Dystrophin Predicts a Rod-Shaped Cytoskeletal Protein. *Cell* 53, 219–228. doi:10.1016/0092-8674(88)90383-2

Lim, K. R. Q., Nguyen, Q., Dzierlega, K., Huang, Y., and Yokota, T. (2020). CRISPR-generated Animal Models of Duchenne Muscular Dystrophy. *Genes* 11, 342. doi:10.3390/genes11030342

López-Hernández, L., Gómez-Díaz, B., Luna-Angulo, A., Anaya-Segura, M., Bunyan, D., Zúñiga-Guzman, C., et al. (2015). Comparison of Mutation Profiles in the Duchenne Muscular Dystrophy Gene Among Populations:

Implications for Potential Molecular Therapies. *Ijms* 16, 5334–5346. doi:10.3390/ijms16035334

Mah, J. K. (2018). An Overview of Recent Therapeutics Advances for Duchenne Muscular Dystrophy. *Methods Mol. Biol.*, 3, 3–17. doi:10.1007/978-1-4939-7374-3_1

Malhotra, S., Hart, K., Klamut, H., Thomas, N., Bodrug, S., Burghes, A., et al. (1988). Frame-shift Deletions in Patients with Duchenne and Becker Muscular Dystrophy. *Science* 242, 755–759. doi:10.1126/science.3055295

Mendell, J. R., Sahenk, Z., Lehman, K., Nease, C., Lowes, L. P., Miller, N. F., et al. (2020). Assessment of Systemic Delivery of rAAVrh74.MHCK7.Micro-dystrophin in Children with Duchenne Muscular Dystrophy. *JAMA Neurol.* 77, 1122–1131. doi:10.1001/jamaneurol.2020.1484

Mendell, J. R., Shilling, C., Leslie, N. D., Flanigan, K. M., al-Dahhak, R., Gastier-Foster, J., et al. (2012). Evidence-based Path to Newborn Screening for Duchenne Muscular Dystrophy. *Ann. Neurol.* 71, 304–313. doi:10.1002/ana.23528

Mias-Lucquin, D., Dos Santos Morais, R., Chéron, A., Lagarrigue, M., Winder, S. J., Chenuel, T., et al. (2020). How the central Domain of Dystrophin Acts to Bridge F-Actin to Sarcolemmal Lipids. *J. Struct. Biol.* 209, 107411. doi:10.1016/j.jsb.2019.107411

Miller, J. N., and Pearce, D. A. (2014). Nonsense-mediated Decay in Genetic Disease: Friend or Foe?. *Mutat. Research/Reviews Mutat. Res.* 762, 52–64. doi:10.1016/j.mrrev.2014.05.001

Murray, M. G., and Thompson, W. F. (1980). Rapid Isolation of High Molecular Weight Plant DNA. *Nucl. Acids Res.* 8, 4321–4326. doi:10.1093/nar/8.19.4321

Neri, M., Rossi, R., Trabanelli, C., Mauro, A., Selvatici, R., Falzarano, M. S., et al. (2020). The Genetic Landscape of Dystrophin Mutations in Italy: A Nationwide Study. *Front. Genet.* 11, 131. doi:10.3389/fgene.2020.00131

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). PRISMA 2020 Explanation and Elaboration: Updated Guidance and Exemplars for Reporting Systematic Reviews. *BMJ* 372, n160. doi:10.1136/bmj.n160

Ramos, E., Conde, J. G., Berrios, R. A., Pardo, S., Gómez, O., and Mas Rodríguez, M. F. (2016). Prevalence and Genetic Profile of Duchene and Becker Muscular Dystrophy in Puerto Rico. *Jnd* 3, 261–266. doi:10.3233/jnd-160147

Richards, S., Aziz, N., Aziz, N., Bale, S., Bick, D., Das, S., et al. (2015). Standards and Guidelines for the Interpretation of Sequence Variants: a Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* 17, 405–423. doi:10.1038/gim.2015.30

Santos, R., Gonçalves, A., Oliveira, J., Vieira, E., Vieira, J. P., Evangelista, T., et al. (2014). New Variants, Challenges and Pitfalls in DMD Genotyping: Implications in Diagnosis, Prognosis and Therapy. *J. Hum. Genet.* 59, 454–464. doi:10.1038/jhg.2014.54

Schwartz, M., and Dunø, M. (2004). Multiplex Ligation-dependent Probe Amplification Is superior for Detecting Deletions/duplications in Duchenne Muscular Dystrophy. *Clin. Genet.* 67, 189–191. doi:10.1111/j.1399-0004.2004.00382.x

Shieh, P. B., Mcintosh, J., Jin, F., Souza, M., Elfring, G., Narayanan, S., et al. (2018). Deflazacort versus Prednisone/prednisolone for Maintaining Motor Function and Delaying Loss of Ambulation: A Post Hoc Analysis from the ACT DMD Trial. *Muscle & Nerve* 58, 639–645. doi:10.1002/mus.26191

Syed, Y. Y. (2016). Eteplirsen: First Global Approval. *Drugs* 76, 1699–1704. doi:10.1007/s40265-016-0657-1

Torella, A., Zanobio, M., Zeuli, R., Del Vecchio Blanco, F., Savarese, M., Giugliano, T., et al. (2020). The Position of Nonsense Mutations Can Predict the Phenotype Severity: A Survey on the DMD Gene. *PLoS One* 15, e0237803. doi:10.1371/journal.pone.0237803

Tran, T. H., Zhang, Z., Yagi, M., Lee, T., Awano, H., Nishida, A., et al. (2013). Molecular Characterization of an X(p21.2;q28) Chromosomal Inversion in a Duchenne Muscular Dystrophy Patient with Mental Retardation Reveals a Novel Long Non-coding Gene on Xq28. *J. Hum. Genet.* 58, 33–39. doi:10.1038/jhg.2012.131

Tuffery-Giraud, S., Miro, J., Koenig, M., and Claustres, M. (2017). Normal and Altered Pre-mRNA Processing in the DMD Gene. *Hum. Genet.* 136, 1155–1172. doi:10.1007/s00439-017-1820-9

Verhaart, I. E. C., and Aartsma-Rus, A. (2019). Therapeutic Developments for Duchenne Muscular Dystrophy. *Nat. Rev. Neurol.* 15, 373–386. doi:10.1038/s41582-019-0203-3

Vieitez, I., Gallano, P., González-Quereda, L., Borrego, S., Marcos, I., Millán, J. M., et al. (2017). Mutational Spectrum of Duchenne Muscular Dystrophy in Spain: Study of 284 Cases. *Neurología (English Edition)* 32, 377–385. doi:10.1016/j.nrleng.2015.12.004

Wallace, G. Q., and McNally, E. M. (2009). Mechanisms of Muscle Degeneration, Regeneration, and Repair in the Muscular Dystrophies. *Annu. Rev. Physiol.* 71, 37–57. doi:10.1146/annurev.physiol.010908.163216

Walter, M. C., Reilich, P., Thiele, S., Schessl, J., Schreiber, H., Reiners, K., et al. (2013). Treatment of Dysferlinopathy with Deflazacort: a Double-Blind, Placebo-Controlled Clinical Trial. *Orphanet J. Rare Dis.* 8, 26. doi:10.1186/1750-1172-8-26

White, J. (2020). PubMed 2.0. *Med. Reference Serv. Q.* 39, 382–387. doi:10.1080/02763869.2020.1826228

Winnard, A. V., Mendell, J. R., Prior, T. W., Florence, J., and Burghes, A. H. (1995). Frameshift Deletions of Exons 3-7 and Revertant Fibers in Duchenne Muscular Dystrophy: Mechanisms of Dystrophin Production. *Am. J. Hum. Genet.* 56, 158–166.

Yamamoto, L. U., Velloso, F. J., Lima, B. L., Fogaça, L. L. Q., de Paula, F., Vieira, N. M., et al. (2008). Muscle Protein Alterations in LGMD2I Patients with Different Mutations in the Fukutin-Related Protein Gene. *J. Histochem. Cytochem.* 56, 995–1001. doi:10.1369/jhc.2008.951772

Conflict of Interest: The authors GF, CJ, ML, and DA have received honoraria/grants for either: teaching, consultation, advisory board and speaker activities from: PTC Therapeutics, Sarepta Therapeutics, Sanofi Genzyme and Biomarin. GF received research grants from PTC pharmaceuticals to perform genetic testing of the patients.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Pharmacology

REVIEW
published: 09 June 2021
doi: 10.3389/fphar.2021.647060

From Genome to Drugs: New Approaches in Antimicrobial Discovery

Federico Serral[1†], Florencia A. Castello[1†], Ezequiel J. Sosa[2,3†], Agustín M. Pardo[3], Miranda Clara Palumbo[2], Carlos Modenutti[2,3], María Mercedes Palomino[2,3], Alberto Lazarowski[4], Jerónimo Auzmendi[4,5], Pablo Ivan P. Ramos[6], Marisa F. Nicolás[7], Adrián G. Turjanski[2,3], Marcelo A. Martí[2,3]* and Darío Fernández Do Porto[1,2]**

[1]Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina, [2]Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina, [3]Instituto de Química Biológica de la Facultad de Ciencias Exactas y Naturales (IQUIBICEN) CONICET, Ciudad Universitaria, Buenos Aires, Argentina, [4]Departamento de Bioquímica Clínica, Instituto de Investigaciones en Fisiopatología y Bioquímica Clínica (INFIBIOC), Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Buenos Aires, Argentina, [5]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina, [6]Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS), Instituto Gonçalo Moniz, Fundação Oswaldo Cruz (FIOCRUZ), Salvador, Brazil, [7]Laboratório Nacional de Computação Científica (LNCC), Petrópolis, Brazil

OPEN ACCESS

Edited by:
Alessio Squassina, University of Cagliari, Italy

Reviewed by:
Hui Zhang, Shanghai Children's Medical Center, China
Ernesto Satoshi Nakayasu, Pacific Northwest National Laboratory (DOE), United States

***Correspondence:**
Adrián G. Turjanski
aturjans@gmail.com
Marcelo A. Martí
marti.marcelo@gmail.com
Darío Fernández Do Porto
dariofd@gmail.com

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *28 December 2020*
Accepted: *17 May 2021*
Published: *09 June 2021*

Citation:
Serral F, Castello FA, Sosa EJ, Pardo AM, Palumbo MC, Modenutti C, Palomino MM, Lazarowski A, Auzmendi J, Ramos PIP, Nicolás MF, Turjanski AG, Martí MA and Fernández Do Porto D (2021) From Genome to Drugs: New Approaches in Antimicrobial Discovery. Front. Pharmacol. 12:647060. doi: 10.3389/fphar.2021.647060

Decades of successful use of antibiotics is currently challenged by the emergence of increasingly resistant bacterial strains. Novel drugs are urgently required but, in a scenario where private investment in the development of new antimicrobials is declining, efforts to combat drug-resistant infections become a worldwide public health problem. Reasons behind unsuccessful new antimicrobial development projects range from inadequate selection of the molecular targets to a lack of innovation. In this context, increasingly available omics data for multiple pathogens has created new drug discovery and development opportunities to fight infectious diseases. Identification of an appropriate molecular target is currently accepted as a critical step of the drug discovery process. Here, we review how diverse layers of multi-omics data in conjunction with structural/functional analysis and systems biology can be used to prioritize the best candidate proteins. Once the target is selected, virtual screening can be used as a robust methodology to explore molecular scaffolds that could act as inhibitors, guiding the development of new drug lead compounds. This review focuses on how the advent of omics and the development and application of bioinformatics strategies conduct a "big-data era" that improves target selection and lead compound identification in a cost-effective and shortened timeline.

Keywords: drug discovery, drug target, metabolic reconstruction, structural modeling, target prioritization, virtual screening

INTRODUCTION

Antibiotics have revolutionized medicine in many aspects, and countless lives have been saved since their discovery at the beginning of the 20th century. However, although antimicrobials have enabled the control of most bacterial diseases considered deadly in the pre-antibiotic era, the emergence of resistant or multiresistant strains, often called "superbugs", is now a huge source of concern for

human health. The extraordinary genetic capacities of bacteria have benefited from man's overuse of antibiotics, leading to multiple antibiotic resistance mechanisms for each antibiotic introduced in the clinical practice (Davies and Davies, 2010). In this context, novel drugs or therapies are urgently required. Still, in a scenario where private investment in the development of new antimicrobials is declining, efforts to combat drug-resistant infections is becoming a worldwide concern.

Generally, antibiotic discovery and development processes are ineffective and costly. It is predicted that around 90% of drugs entering phase 1 clinical trials will not reach approval and that the overall cost for each approved compound is about 1.4 billion dollars (Hay et al., 2014; DiMasi et al., 2016). The decision-making process in a drug discovery project requires a thorough understanding of as many variables as possible to maximize the chance of success. The reasons for the failure of many new antimicrobial development projects range from inadequate selection of the molecular targets to a lack of innovation and, quite significantly, the appearance of severe side effects. However, the availability of pathogen genomic-scale datasets has created new opportunities for drug discovery, including those against new resistant and multiresistant strains. Subtractive genomics, structural bioinformatics, and metabolic pathways analysis approaches are currently applied for the development of new drugs and fight antimicrobial resistance, acting as a complement to traditional wet-lab approaches. Although not enough time has elapsed to exploit all capabilities of in silico approaches in drug discovery, target-based drug discovery has been effective for many therapeutic targets, most notably for HIV/AIDS (Zhan et al., 2016), and was also successful in identifying potent antibacterial inhibitors of peptide deformylase (Hackbarth et al., 2002). Other examples of genomic approaches that resulted in promising compounds include AFN-1252, a selective inhibitor of the *Staphylococcus aureus* enzyme enoyl-acyl carrier protein reductase, FabI, which showed potent *in vitro* activity and *in vivo* efficacy (Kaplan et al., 2012). Early genome-wide studies pointed to the essentiality of proteins involved in fatty acid biosynthesis that, coupled to structural differences between enzymes from bacteria and mammals, made these a noteworthy target (Forsyth et al., 2002). BamA, a component of the β-barrel assembly machine of Gram-negative bacteria, has also been proposed as a target due to its essentiality and extensive conservation in these organisms. A monoclonal antibody that selectively inhibits this protein has been developed and demonstrated to have bactericidal activity (Storek et al., 2018). Other inhibitors targeting Gram-negative outer membrane proteins have also been proposed (MacNair et al., 2020). Combined, these examples reinforce the utility of target-based approaches that, informed by genome evidence, can result in the successful identification of novel drug candidates.

Targeted drug development projects consist of several steps that range from candidate selection and validation, the performance of *in vitro* and *in vivo* experiments to identify lead and candidate molecules, pre-clinical development in animal models, and finally, clinical trials in human subjects to establish safety and effectiveness. Along this long and winding road, several significant challenges must be met to avoid failure, and as in any race, an optimal start is of great advantage. The mentioned advent of omics approaches (e.g., genomics, transcriptomics, and proteomics) has fostered the development of bioinformatics tools guiding to a "big-data era" that allows improved identification of putative targets and lead compounds. Other informatics approaches to enhance antimicrobial discovery, such as Machine Learning (ML), are out of the scope of the present work and are reviewed elsewhere (Lau et al., 2021). Opportunities to apply ML occur in all stages of antimicrobial discovery (Vamathevan et al., 2019; Lau et al., 2021). Examples include target validation, identification of prognostic biomarkers, and analysis of digital pathology data in clinical trials. Halicin is one of the most notable discoveries of new antimicrobials using ML techniques (Stokes et al., 2020). This drug was effective against many multidrug resistant microbes *in vitro* and *in vivo*.

In this review, we will focus on the different bioinformatics strategies used for prioritizing drug targets in pathogens. Particularly, we include results of prioritized targets with their potential molecule inhibitor candidates for two bacteria that cause endemic diseases in Latin American countries, namely *Mycobacterium tuberculosis* (Mtb) and *Bartonella bacilliformis* (Bb).

HOW TO PRIORITIZE DRUG TARGETS IN PATHOGENIC BACTERIA?

Since experimental research of putative drug targets is time-consuming and expensive, it is worthwhile to conduct bioinformatic analysis to select proteins that are good candidates as molecular targets for antimicrobial discovery projects. These analyses consider features commonly thought to be desirable in a target, including druggability (whether drug-like compounds are likely to interact with the protein), essentiality (which suggest that inhibiting the target function will have the desired bactericidal effect), specificity/selectivity (potential for inhibiting the pathogen without harming the host and its microbiota), and relevance in metabolic stages of the pathogen during infection.

From a general point of view, druggability is a concept used to describe the ability of a given protein to bind a drug-like molecule, which in turn modulates its function in some "desired" way (Gashaw et al., 2012). From a structural point of view, it can be related to the likelihood that a small molecule binds a given protein target with high affinity (Sheridan et al., 2010), a property usually referred to bindability. Taking this into account, druggable proteins should have a well-defined pocket with suitable physicochemical properties to bind a drug. Our group has developed a fast whole genome approach for druggability prediction based on the open-source algorithm fpocket (http://fpocket.sourceforge.net/) (Guilloux et al., 2009), which combines several physicochemical descriptors to estimate the druggability of the pockets present in proteins. This approach was extensively tested, both on experimental structures and homology-based models, in the context of whole proteome target search studies in our previous works on the subject (Radusky et al., 2014;

Defelipe et al., 2016; Ramos et al., 2018; Sosa et al., 2018; Farfán-López et al., 2020). Based on previous analysis of the druggability score distribution, for all pockets that host a drug-like compound in the Protein Data Bank (Sussman et al., 1998), we have classified pockets into four categories: non-druggable (ND), poorly druggable (PD), druggable (D), and highly druggable (HD). Good candidate targets are, from a structural standpoint, proteins that fall either into D or HD classes. Most of the pockets that actually host a drug in the PDB (80%) could be classified as druggable or highly druggable by our methodology.

Moving from the structural to a more general druggability concept, the early steps of rational antimicrobial target identification usually involve integrating the structural druggability assessment with the information present in the host and pathogen whole genomes. This strategy, called subtractive genomics, allows to select those targets relevant for the pathogen and absent in the host. Identifying a group of proteins that are essential to pathogens but are not present in the host minimizes the chance of unwanted side effects during treatment (Barh et al., 2011). Three hierarchical levels (sequence → DNA/protein, structure → protein, and enzymatic/regulatory reactions → regulatory/metabolic network) have generally been used alone to select candidate targets (Radusky et al., 2015; Defelipe et al., 2016; Kaur et al., 2017; Wadood et al., 2017; Uddin and Jamil, 2018). We drive our focus to the analysis of these multiple omics layers under an integrative framework.

Metabolic Reconstruction Contextualizes Target Importance and Directs Selection in Early Phases

The first layer of information that can be used to direct target prioritization efforts is the evaluation of the metabolic importance of a given protein. Metabolism refers to the set of biochemical reactions and regulatory pathways leading to cellular homeostasis and functioning. Early studies on microbial metabolism elucidated the major pathways related to energy production, amino acid synthesis, and lipid formation, which, combined with the current availability of full genomes and proteomes, helped set the stage for the study of metabolism on a large-scale. Computational methods that rely on whole-genome sequences, gene annotations, or both, allow for rapid generation of an initial metabolic draft for any given organism, which must be followed by careful manual curation to achieve a high-quality metabolic reconstruction. Pathway Tools (Karp et al., 2015) is one of such software providing a module (PathoLogic) that takes as input the genome and associated annotations of an organism of interest and, by mapping these annotations onto enzymatic reactions within the MetaCyc database using an enzyme-name matching tool, infers the set of reactions (or the reactome) for the desired species. A pathway-scoring algorithm is employed to predict pathways within the expected taxonomic range. Among other capabilities, the tool allows manual curation to be performed and supports the gap-filling process of pathways that could not be determined entirely by name matching alone. This process relies on the gene sequences. ModelSEED (Devoid et al., 2013) is a web resource that facilitates the reconstruction, exploration, comparison, and analysis of organism-specific metabolic models. This tool relies on an initial genome annotation using RAST and the SEED ontology, clustering metabolic pathways into subsystems, which are further subclassified (Devoid et al., 2013). KEGG Mapper tools (Kanehisa and Sato, 2020) also allow automatic assignment of enzymatic roles and pathway contextualization using genomes or proteomes as input and relies on the KEGG ontology to perform annotations based on sequence similarity. Other tools, reviewed elsewhere, allow additional curation and pathway-specific analysis to be performed once a draft reconstruction is attained (Pitkänen et al., 2010; Abd Algfoor et al., 2017). Common to all described tools is their dependency upon a vocabulary of metabolic elements, or ontology, which depicts the complex, and often multi-level relationships among genes, proteins, enzymes, biochemical reactions, and regulators (Stobbe et al., 2012). Accordingly, metabolic reconstructions performed using different strategies may lead to differing outcomes for the same organism, as pathway representations and modeling varies among each developing group (Green and Karp, 2006).

Once a metabolic reconstruction is obtained for the studied pathogen, this compendium can be used during the early prioritization step that involves target identification, aiming to rank the proteins involved in critical metabolic roles or participate as key intermediaries of multiple pathways. This analysis can be facilitated using a graph representation of a metabolism (Ramos et al., 2018). Multiple topological criteria can then be employed to rank the proteins identified as belonging to one or more metabolic pathways (**Box 1**). The rationale behind this strategy is that drugs that inhibit such targets have higher chances of success than those that target non-essential cellular functions.

Target Selection Databases

The increased availability of pathogen genomes and genome-scale datasets are expected to guide target-based drug discovery projects. However, a major bottleneck has been the complexity of capturing and integrating relevant information available, making them accessible to experimental researchers, thus facilitating the identification and prioritization of potential antimicrobial targets. Nowadays, there are several freely available academic resources designed for antimicrobial target identification. Most of these tools focus on specific protein characteristics. For example, Drug Target Database is a useful resource to select potential targets based on a reverse docking approach. The Therapeutic Targets Database provides a large volume of data of already known therapeutic targets. Another database that includes data of known targets is TargetDB/TargetTrack (Chen et al., 2004), in spite of its focus on structural information.

There are also a few existing databases and resources aimed at a particular group of pathogens. TDR targets (Magariños et al., 2012) is an interesting tool focused on neglected tropical diseases. Regarding the prioritization of molecular targets in *M. tuberculosis*, two specialized databases are currently available, TuberQ and TargetTB. TuberQ provides a druggability analysis of the Mtb proteome contributing to a better selection of potential

BOX 1 | Key concepts used to assign metabolic importance during target prioritization.

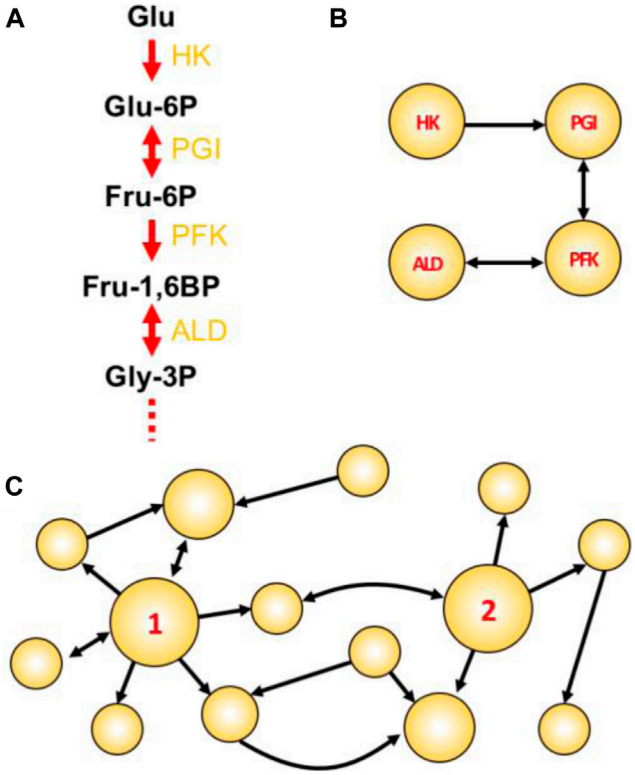


Figure Box Network concepts illustrated. **(A)** The initial four steps of glycolysis. In such traditional textbook representations, the emphasis is usually given to compounds. Enzyme names abbreviations are depicted in yellow. **(B)** A reaction-reaction directed graph constructed using the initial reactions shown in A. Here, the emphasis is given to reactions/enzymes. **(C)** A toy network with two particularly attractive nodes: node 1 having a high degree; and node 2 having high betweenness centrality. This example shows a directed graph, in which links are directed from one node to another. In undirected graphs, all links are bidirectional and represented by a single line connecting a node pair.

Choke-point reaction: A biochemical reaction that uniquely consumes (or synthesizes) a given substrate (or product) (Yeh et al., 2004). Enzymes that perform these reactions are termed choke-point enzymes, and their blocking could lead to the accumulation of the unique substrate (potentially toxic to the cell) or to the inability to produce an essential product (impairing the cellular homeostasis). For this reason, the identification of metabolic choke-points is integral to the prioritization of potential targets.

Reaction-reaction graph: In a graph-oriented study of metabolism (reviewed in (Cottret and Jourdan, 2010), biochemical reactions can be modeled as the network nodes, which consequently also model the enzyme(s) catalyzing the reaction. A link is placed between two reactions if one consumes a metabolite produced by the other reaction [Panel Box **(A,B)**].

Degree centrality (DC): The degree is one of many centrality measures (reviewed in (Jalili et al., 2016; Ashtiani et al., 2018) useful to define important metabolic nodes and represents the number of links connecting to a node. The higher the DC, the more shared metabolites a given reaction has with other immediate reactions. In directed graphs, the total DC is the sum of the in-degree (the number of incoming links) and out-degree (the number of outgoing links). An example of a high-degree node is shown in Panel Box **(C)**, where node 1 has in-degree = 2, out-degree = 6, and a total DC of 8.

Betweenness centrality (BC): Represents the frequency with which a given node appears as an intermediate between the paths of other possible node pairs. In the metabolic context, a reaction node with high BC would involve a metabolite that participates in many other reactions (not necessarily of the same direct pathway), thus having an important metabolic role. An example of a node having high BC is shown in Panel Box **(C)**, where node 2 is the only intermediate able to connect reactions on the left with the three reactions that appear on the right-hand part of the graph. Thus, node 2 is a key intermediate node in this graph.

drug targets for screening campaigns (Radusky et al., 2014). TargetTB integrates network analysis of the protein-protein interaction, metabolism, essentiality, sequence analyses, and structural data (Raman et al., 2008). Some databases allow the use of metabolic network data to target prioritization, such as FindTargetsWEB. This web server takes as input an extended Systems Biology Markup Language (SBML) file of a metabolic model of the pathogen under study. It performs both flux balance analysis (FBA) and flux variability analysis (FVA) to prioritize bacterial molecular targets (Merigueti et al., 2019).

Most of the available user-friendly web servers use few data sources to prioritize targets. However, continuing efforts to allow target prioritization by applying integrated multi-data approaches are in ongoing focus. In this context, by combining structural druggability, essentiality analysis, metabolic context, as well as genomic and expression data, our group has developed Target-Pathogen (TP) (Sosa et al., 2018) (**Figure 1**). TP is a web server that enables to select and prioritize drug targets of several clinical pathogens, including *M. tuberculosis*, *M. leprae*, *K. pneumoniae*, *S. aureus*, *Schistosoma*

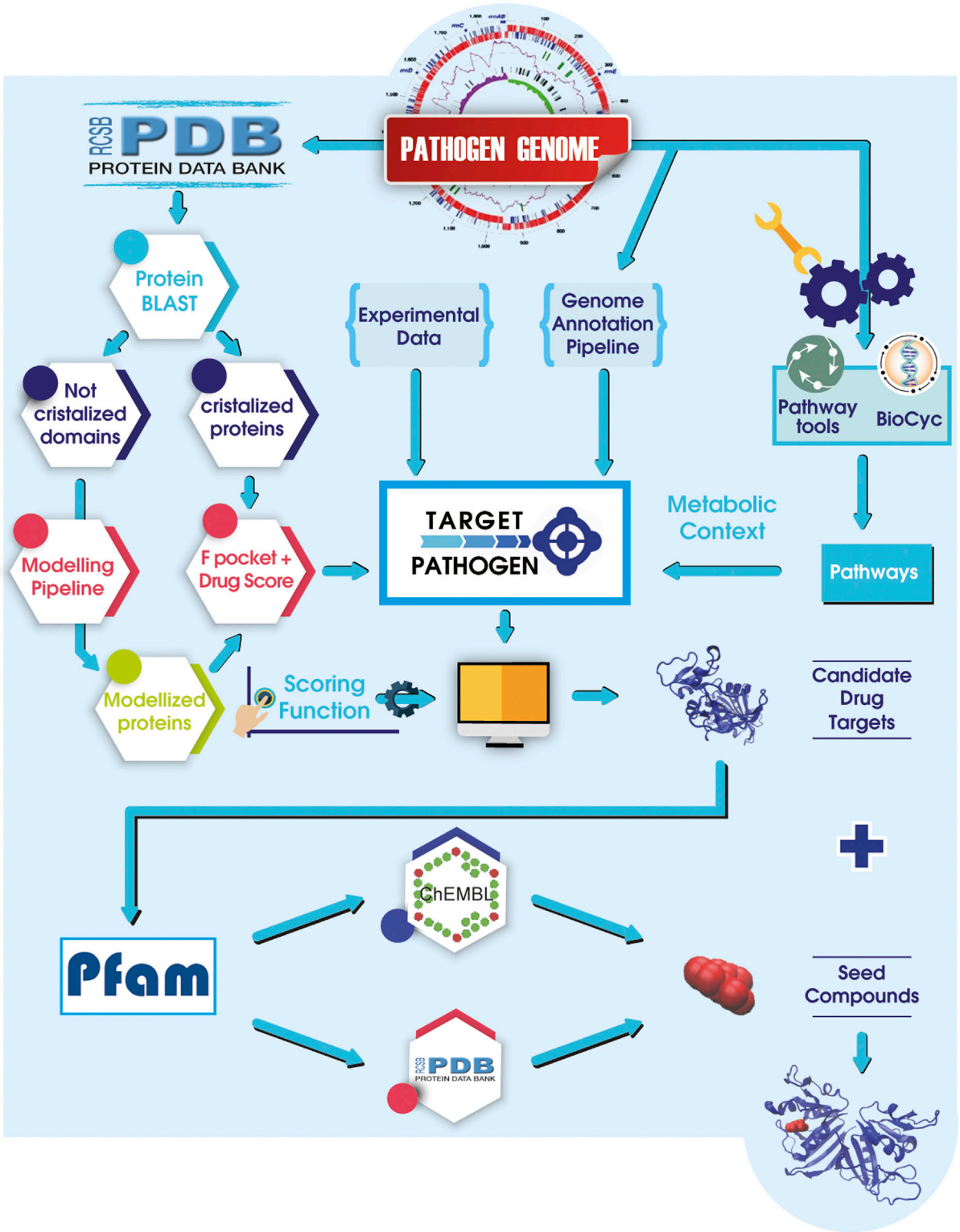


FIGURE 1 | A general sketch of Target-Pathogen integrated with LigQ pipeline. Structural druggability and metabolic analyses are integrated with available experimental data and *in silico* analysis data. After all, data is integrated into Target-Pathogen, a user-designed scoring function is used to weight different features to obtain a ranked list of candidate drug targets. Once the target is selected, LigQ pipeline finds all known binders of similar proteins in the PDB and ChEMBL. Target and putative binders can be used in further molecular docking assays.

mansoni, *Shigella dysenteriae*, *Toxoplasma gondii*, *Leishmania major*, *Trypanosoma cruzi*, *Acinetobacter baumannii*, and *Bartonella bacilliformis* among others. Under the TP framework, researchers can easily prioritize proteins of interest quickly and intuitively, running simple queries (such as searching for proteins with high druggability score or associated with metabolic reactions of high centrality), filtering by different data, assigning numerical weights for additional customized features and merge these results to obtain a ranked list of targets. A distinct advantage of the Target-Pathogen server is its capacity to rank, not solely proteins but entire pathways, thus allowing synergistically to attack several proteins of the same metabolic pathway. Another attractive feature of TP is that it will enable users to upload their data to be used in the prioritization pipeline. As of December 2020, there are 25 of the most relevant microorganisms from the human health perspective. Users can also request new genomes to be included in the platform by emailing target@biargentina.com.ar. By abiding to open-science practices, data associated with protein structures can be downloaded to perform further *in silico* analysis outside TP.

Identifying Lead Compounds to Treat Bacterial Infections

Once the protein target is selected, the challenge moves from biology to chemistry and consists of the identification of a small, usually drug-like, molecule that can inhibit the target's function, allows further pharmacological validation of the target, and ultimately paves the way for the development of a new antibiotic. To test a molecule's capacity to inhibit the desired target, *in vitro* protein activity (or binding) assays can be performed, as well as cell culture MIC determinations. However, the problem is that the universe of molecules that could act as inhibitors is vast. Conducting experimental high-throughput screening is beyond the capacity of most academic research labs in Latinamerican countries, where a typical research group can afford and test about 100 compounds in a typical one by one *in vitro* assay each year. Therefore, usually, only a moderate number of compounds are tested, and bioinformatics methods capable of screening for potential binders are highly appreciated.

The computational selection of potential inhibitors against a defined target is generally referred to as Virtual Screening (VS). VS methodologies can be further divided into two main techniques, which can be applied sequentially to obtain a best set of potential inhibitors. The first relies mainly on previous biological information and chemical similarity analysis of the compounds. It is usually referred to as compound filtering or pre-selection (as will be described below). The second, which is computationally demanding, involves molecular docking of each compound in the protein target, estimating its binding free energy, and finally performing a ranking. This technique is commonly and traditionally referred to as VS in the strictest sense.

Compound filtering traditionally involves selecting drug-like compounds using a set of driving principles, for instance, Lipinski's rules (Lipinski et al., 2001). However, the increasing amount of information available in public databases allows the derivation of improved filters, e.g., using the "guilt by association" principle, as described in our previous work LigQ (Radusky et al., 2017), and similar developments (O'Boyle et al., 2011; Volkamer et al., 2012). The idea is that similar proteins bind similar compounds. Therefore, for a given target, those compounds that are similar (in chemo-structural properties) to known binders of similar (homolog) proteins are good candidates. Starting from the selected target (protein name or UniprotId), LigQ first finds all known binders of similar proteins. Binders are classified in groups according to the degree of protein similarity [starting from high identity >60% homologs to binders to the same domain in PFAM (Mistry et al., 2020)] and available information (such as the structure of the protein-ligand complex) in different databases such as Protein Data Bank (PDB; http://rcsb.org), Pfam (http://pfam.xfam.org/), and ChEMBL (EMBL-EBI; http://www.ebi.ac.uk/chembl/).

This set of compounds is called the "seed set." It is used to retrieve from large datasets of commercially available compounds, those that are chemically similar to a specific -user-defined- degree. Chemical similarity can be defined based on the Tanimoto Index (Bajusz et al., 2015), and the similarity retrieval cut-off can be used to select the number of compounds to be retrieved, which are also organized in clusters according to their chemical similarity.

The information is extracted for each database, constituting four individual seed sets (Seed I–IV). Seed I and III are obtained through the direct search of the protein of interest by its corresponding identifier (ID) for each base (PDB (Sussman et al., 1998) and ChEMBL (Gaulton et al., 2012), respectively). On the other hand, the seeds II and IV are extracted by previously obtaining the functional domains (Pfam) that compose the protein of interest by using HMMER and later searching in PDB and ChEMBL for the compounds that interact with these domains.

In the following sections, we present and review prioritized targets and their potential binders, identified using the above-described methodology for two bacterial pathogens with an important impact in Latin America: *Bartonella bacilliformis* (causal agent of Carrion's disease) and *Mycobacterium tuberculosis*.

BARTONELLA BACILLIFORMIS AND CARRION'S DISEASE

Carrion's disease is an ancient vector-borne biphasic illness dating from the pre-Columbian era, restricted to the South American Andes, including Peru, Ecuador, and Colombia (Gomes and Ruiz, 2017). It is an endemic illness found in Andean valleys at an altitude of 600–3,200 m above sea level. *B. bacilliformis* (Bb) is transmitted to humans by female sandflies belonging to the *Lutzomyia* genus, which are commonly present in Andean valleys' high-altitude regions (Clemente et al., 2012; Minnick et al., 2014). However, since the end of the last century, an expansion of the illness into bare areas including jungle and coastal regions, such as the coastal areas of Guayas and Manabi in

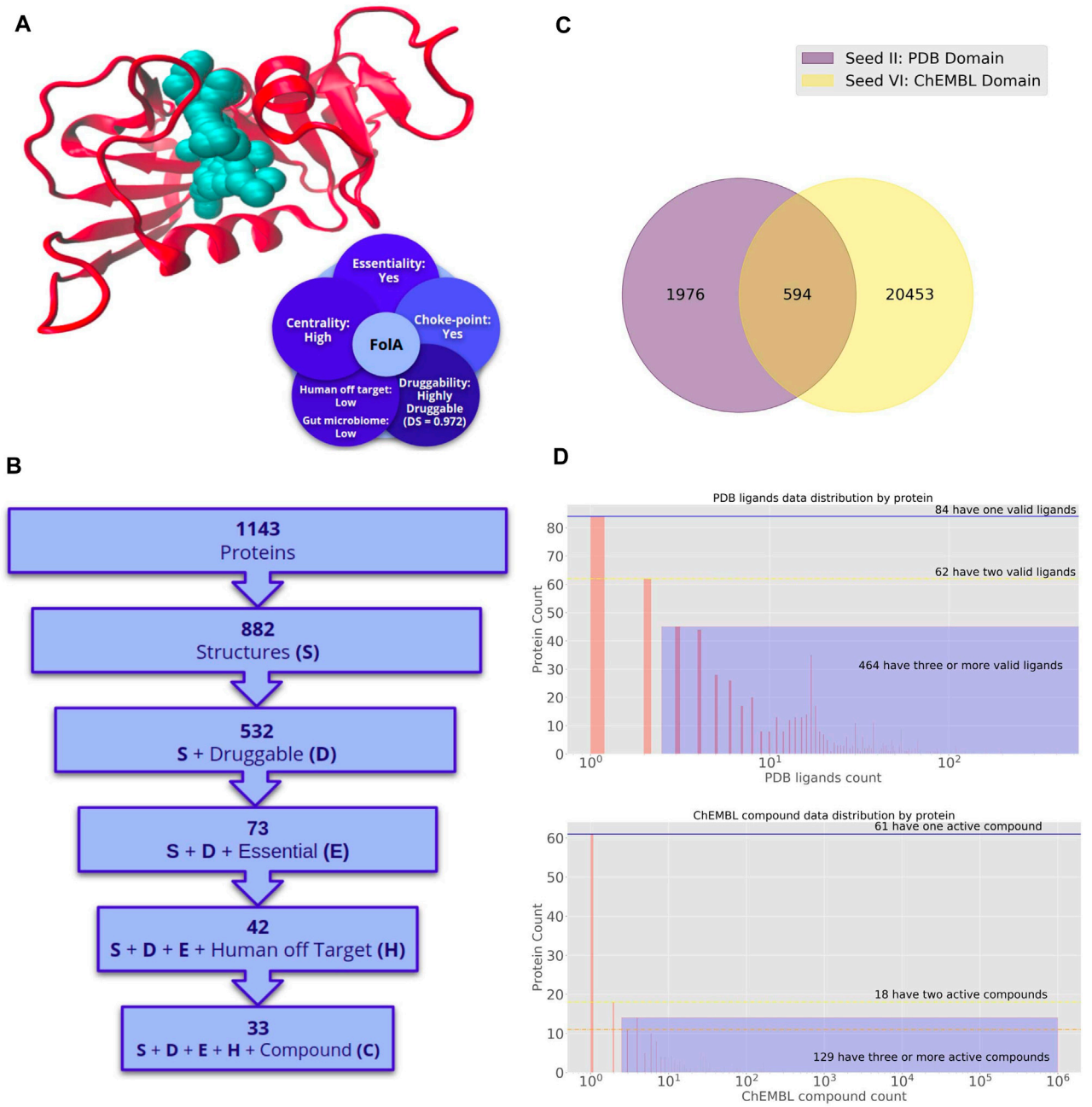


FIGURE 2 | Potential druggable targets and putative lead compounds to combat *Bartonella bacilliformis*. **(A)** Dihydrofolate reductase (FolA) structure. Crystallographic structure and a set of attractive features for target prioritization are shown. **(B)** Number of proteins in *Bartonella bacilliformis* genome with desirable properties for drug targets. Different filters are sequentially applied to obtain a shortlist of druggable, essential, and low identities with proteins in the human genome. The last filter is applied to get the list of proteins with putative binders. **(C)** Seed Compounds for *Bartonella bacilliformis*. Venn diagram of the number of seed compounds corresponding to different sets. Binders are classified into two seed groups. Seed II are those drugs that bind any protein that harbor the same Pfam domains with Bb proteins and have been co-crystallized with such proteins in the PDB. Seed IV is the set of drugs that bind any protein that shares Pfam domains with *Bartonella bacilliformis* proteins and was reported as active in Chembl. **(D)** Number of proteins in the *Bartonella bacilliformis* genome for which a ligand can be predicted. The top panel corresponds to drugs in Seed II. The bottom panel corresponds to drugs in Seed IV.

Ecuador, has been reported (Gomes and Ruiz, 2017; Garcia-Quintanilla et al., 2019). It is also thought that climate change could favor the expansion of Bb infections, presumably affecting the vector proliferation. In this sense, it is worthwhile to mention the El niño phenomenon, the unusual warming of surface waters in the eastern Pacific Ocean, which leads to a temperature and

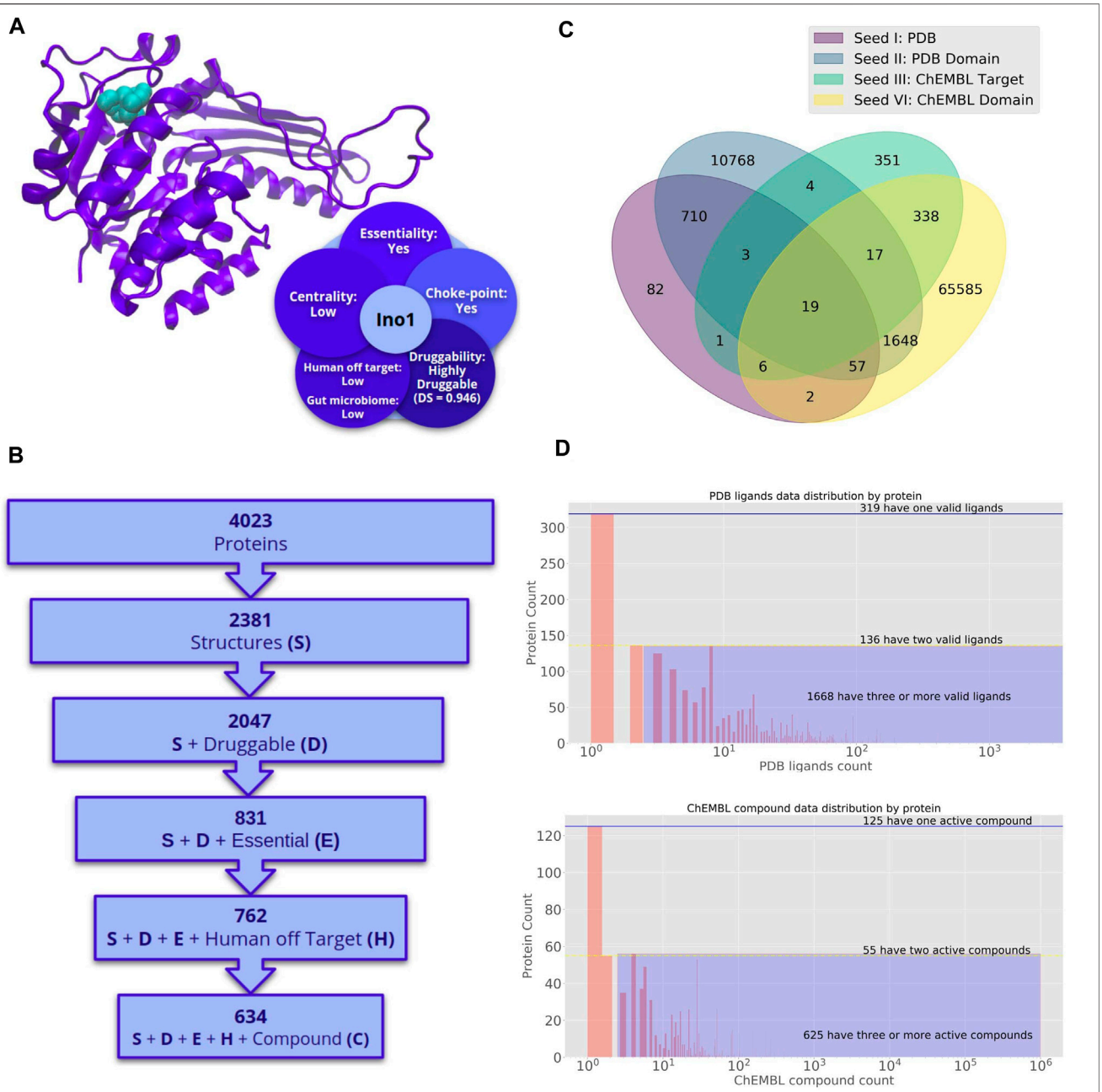


FIGURE 3 | Potential druggable targets and putative lead compounds to combat *Mycobacterium tuberculosis*. **(A)** Inositol-3-phosphate synthase (Ino1) structure. Crystallographic structure and a set of attractive features for target prioritization are shown. **(B)** Number of proteins in *Mycobacterium tuberculosis* genome with desirable properties for drug targets. Different filters are sequentially applied to obtain a shortlist of druggable, essential, and low identities with proteins in the human genome. The last filter is applied to get the list of proteins with putative binders **(C)** Seed Compounds for *Mycobacterium tuberculosis*. Venn diagram of the number of seed compounds corresponding to different sets. Binders are classified in groups according to the degree of protein similarity [starting from high identity >60% homologs to binders to the same domain in PFAM (Mistry et al., 2020)] and available information (such as the structure of the protein-ligand complex) in different databases such as Protein Data Bank (PDB), Pfam, ChEMBL (EMBL-EBI). Venn diagram of the number of seed compounds corresponding to different sets. Binders are classified into two seed groups. Seed I and III are obtained through the direct search of the protein of interest by its corresponding identifier (ID) for each base and ChEMBL. Seed II are those drugs that bind any protein that harbor the same Pfam domains with Bb proteins and have been co-crystallized with such proteins in the PDB. Seed IV is the set of drugs that bind any protein that shares Pfam domains with *Mycobacterium proteins* and was reported as active in Chembl. **(D)** Number of proteins in the *Mycobacterium tuberculosis* genome for which a ligand can be predicted. The top panel corresponds to drugs in Seed II. The bottom panel corresponds to drugs in Seed IV.

humidity increasing. These climate characteristics especially favor the sandfly spreading and promoting new Carrion's disease outbreaks (Pons et al., 2016).

The infection caused by Bb has two well defined clinical phases. The early stage, denominated Oroya fever, causes a severe acute hemolytic anemia. High case-fatality rates as 40–88% have been described in the Oroya fever phase in patients without any antibiotic treatment. Even with timely antibiotic treatment, the fatality rate is around 11% (Farfán-López et al., 2020). The chronic phase of Carrion's disease is characterized by the development of dermal eruptions known as Peruvian warts and commonly present on the head and extremities. Although this phase is seldom fatal, dermal eruptions can be accompanied by fever, headache, lymphadenopathy, and acute pains in joints and bones (Minnick et al., 2014).

Regarding antimicrobial therapy to treat Carrion's disease, different antibacterial agents have been used since the beginning of the antibiotic era, such as beta-lactams (including penicillins and cephalosporins), aminoglycosides, and quinolones (Battisti et al., 1998). Although most Bb strains are sensitive to a broad set of antimicrobials *in vitro*, there is still a potential risk of developing antibiotic-resistance during clinical treatment. Oroya fever has been traditionally treated with chloramphenicol, a successful drug due to the frequent coinfection with *Salmonella* spp. However, it is nowadays restricted for humans because of its potential to produce side effects in the bone marrow. Other drugs to treat Carrion's disease include beta-lactams such as ampicillin and penicillin G, tetracyclines (doxycycline), macrolides (erythromycin, roxithromycin), trimethoprim-sulfamethoxazole, and fluoroquinolones (norfloxacin, ciprofloxacin) (Rolain et al., 2004). Although the second-generation fluoroquinolone ciprofloxacin is the drug of choice for treating acute cases, it should be judiciously recommended because of the ability of Bb to become quinolone resistant. Several studies showed that quinolone resistance-determining regions (QRDR) are consequences of synonymous or non-synonymous mutations and responsible for the intrinsic resistance of *Bartonella* spp to this antimicrobial (Valle et al., 2010); (Espinoza-Culupú et al., 2014). Additionally, mutations conferring resistance to ciprofloxacin, erythromycin, rifampin, aminoglycosides, and folate inhibitor targets have been molecularly characterized in clinical isolates (Biswas et al., 2007). The current scenario is worse, considering the antibiotic resistance mediated by efflux pump overexpression (Gomes et al., 2016).

Mycobacterium Tuberculosis

Tuberculosis (TB) is an infectious disease that accounted for 1.2 million deaths in 2019 (Harding, 2020) being one of the top ten causes of death worldwide and the leading cause of death from a single infectious agent (ranking above HIV/AIDS). More than 95% of cases and deaths occur in developing countries (Ascenzi and Visca, 2008). Tuberculosis epidemiology varies markedly between Latin American countries (Woodman et al., 2019). The incidence of tuberculosis in Central America (including Mexico), the Caribbean, and South America were 25.9, 46.2, and 61.2 per 100,000 people. Drug resistance is an increasing problem throughout the Americas, particularly in Peru, where drug-resistant tuberculosis accounts for 9% of the cases (Woodman et al., 2019). In this framework, only 33% of patients received drug-susceptibility testing, resulting in an estimated 7,000 undiagnosed or untreated patients with drug-resistant tuberculosis (Woodman et al., 2019). About a quarter of the Latin American population is latently infected with Mtb.

Immune response to Mtb relies on phagocytosis of the bacteria by macrophages leading to granuloma formation. Inside the macrophages, bacilli face stressful conditions characterized by the presence of Reactive Nitrogen and Oxygen Species (RNOS). Based on this observation, we have hypothesized that identifying Mtb RNOS protein targets would permit us to select inhibitors against them and synergize with the macrophages attack in the latent phase of the infection (Defelipe et al., 2016).

Exploring the Druggable Genomes of *Bartonella bacilliformis* and *Mycobacterium tuberculosis*

We applied the previously described pipeline to the pathogens mentioned above, Bb and Mtb. Below we present the application of successive filters (drugability, essentiality, etc.) along their genomes. Bb genome codes for 1,143 different proteins (**Figure 2B**), from which we were able to build a total of 882 high-quality structural models (no experimental structures are available in the PDB for any Bb protein). Homology-based models are built for all proteome sequences using MODELLER (Webb and Sali, 2016) only when an adequate template is available (coverage 80%, E-value > 1×10^5). Only those models with GA341 score above 0.7, QMEAN between −2 and two are retained. It has been shown that the RMSD between the Modeller models and the native structures is <3 Å (Wallner and Elofsson, 2005), which shows the quality of the obtained models. 532 (~60%) of the models harbored a druggable pocket (DS > 0.5). From this subset of structurally relevant proteins, only 73 can be predicted as essential (i.e., close homologs in the Database of Essential Genes were found). After further filtering those proteins with close homologs in the human genome, 42 proteins remained (identity < 0.4). When an additional filter was applied in TP to disclose proteins that could potentially bind at least one compound, a final set of 33 candidate proteins was obtained.

Additionally, 18 of the 42 proteins mentioned above also have a low impact on the gut microbiome, and 17 have putative binders. Ten targets (all with predicted ligands) are also associated with choke-points reactions, and four with high-centrality reactions from the metabolic network point of view (concepts defined in **Box 1**). All this information is provided in the TP database, while a detailed list of mentioned targets is also presented in **Supplementary Table S1**.

The Mtb proteome comprises 4,023 proteins (**Figure 3B**), being 2,381 structurally defined (382 experimental structures and 1,999 models). From these proteins, 2,047 had DS > 0.5 (~85%), 831 were also essential, and putative drugs delivered for 762 are predicted to have a low impact in humans. We could predict

TABLE 1 | Proteins of Bb with desirable features to become a promising drug target.

Bartonella bacilliformis

Protein name	Druggability	Choke point	Centrality	Human off-target	Gut microbiome	Essentiality
Enoyl-[acyl-carrier-protein] Reductase (FabI)	0,992	Yes	High	Low	Low	Yes
Dihydrofolate reductase (FolA)	0,972	Yes	High	Low	Low	Yes
3-Phosphoshikimate carboxyvinyltransferase (AroA)	0,775	Yes	High	Low	Low	Yes
FADH(2)-oxidising methylenetetrahydrofolate –tRNA-(uracil(54)-C(5))- methyltransferase (TrmFO)	0,746	Yes	High	Low	Low	Yes
Undecaprenyl-diphosphatase (UppP)	0,738	Yes	High	Low	Low	Yes
UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase (MurE)	0,952	Yes	High	Low	Low	Yes

possible binders for 634 proteins out of this 762. From this subset, 635 satisfy microbiome off-target criteria (527 has possible binding compounds). If metabolic perspective is also considered, 140 catalyze choke-points reactions (130 with potential binders), and seven are associated with-high centrality reactions (all with predicted ligands). All this information is provided in the TP database, while a detailed list of mentioned targets is also presented in **Supplementary Table S2**.

The successive steps of the pipeline indicate how the application of sequential filtering steps narrows the universe of potential targets. We describe the most promising targets and their potential inhibitors for Bb and Mtb in the following sections.

Bb Prioritized Protein Targets and Their Potential Inhibitors

In Farfán-López et al. (2020), our group participated in a work that combined the efforts of scientific groups from Argentina, Brazil, and Peru to perform an integrative genomic-scale data analysis, which allowed us to shortlist a set of proteins that could serve as attractive targets for new antimicrobial discovery projects against Bb. This study was based on the genomic analysis of Bb USM-LMMB 07, firstly isolated in 2011 during an outbreak in Carmen de la Frontera district, Huancabamba Province, Piura (Guillen et al., 2016). The combination of genomic, structural, metabolic, and functional data integrated inside Target-Pathogen, finally led to shortlisting six proteins (FabI, FolA, AroA, TrmFO, UppP, and MurE) with unique characteristics (**Table 1**). FolA provides the main dihydrofolate reductase activity in the tetrahydrofolate or vitamin B9 pathway (**Figure 2A**). As is well known, tetrahydrofolate is a crucial intermediate in the biosynthesis of nucleic acids and proteins, which is biosynthesized *de novo* in bacteria. It participates in essential biosynthesis pathways, such as methionine, purines, and thymidylate. Since dihydrofolate reductase is essential for cell division and growth, it could become an attractive target for drug development. Another top-ranking protein is Enoyl- [acyl-carrier-protein] reductase (FabI), which is involved in fatty acid biosynthesis processes and was also described to be essential in many other bacteria, such as *E. coli* and Mtb (Heath et al., 1998; Kaplan et al., 2012). The gene product of *aroA* also meets the standard requirements to become a potential molecular target. Our ontology analysis results revealed that this protein is involved in aromatic amino acids and chorismate biosynthesis and showed an essential role in *Rhodopseudomonas palustris* CGA009 and *Caulobacter crescentus* (53.3 and 50.9% sequence identity with Bb *aroA,* respectively). Another attractive target found by our subtractive genomic approach is the tRNA modification enzyme, TrmFO. This protein showed a high identity against the *gid* essential gene of *Staphylococcus aureus* N315. Interestingly, several pathogens such as *E. coli*, *P. aeruginosa*, and *S. enterica* (Yim et al., 2006; Gupta et al., 2009; Shippy et al., 2011) show pleiotropic effects when carrying a mutant *gidA*; thus in this line, TrmFO becomes an appealing target in Bb. Finally, we prioritized UppP and MurE, enzymes involved in peptidoglycan (PG) biosynthesis, usually considered one of the principal antimicrobial targets. PG is a crucial component of the cell envelope of Eubacteria. It has an essential role in bacterial physiology due to its functions in maintaining the shape and integrity during growth and cell division, controlling the internal turgor pressure resistance, and serving as a structural scaffold to other cell envelope components. We now turn our attention to their potential inhibitors.

As described in the introduction, we implemented the LigQ pipeline in the context of Target-Pathogen to allow the identification of potential ligands that interact with desired protein targets. As mentioned above, there are no Bb protein structures in the PDB. There is also no information on experimental assays in ChEMBL for this pathogen; therefore, the set of possible inhibitors is based on seed sets II and IV, i.e., derived from ligands observed for proteins that share domains with the selected Bb targets. Seed set II consisted of 1,976 compounds, while seed IV was composed of 20,453 drugs. In contrast, 594 compounds are retrieved from both PDB and ChEMBL simultaneously (**Figure 2C**). This fact makes these drug-like compounds attractive to combat Bb infections. From a total of 1,143 Bb proteins, we could predict possible binders for 610 in the PDB and 201 in ChEMBL (**Figure 2D**). Potential inhibitors for the predicted targets are shown in **Supplementary Table S3**. As an example, Isoniazid (CHEMBL64—INH) was indicated as a potential inhibitor of Bb FabI. INH is one of the most important first-line drugs against tuberculosis. Although antimicrobial

TABLE 2 | Mtb proteins with worthy properties that make them good candidate targets.

Mycobacterium tuberculosis H37Rv						
Protein name	**Druggability**	**Choke point**	**Centrality**	**Human off-target**	**Gut microbiome**	**Essentiality**
Inositol-3-phosphate synthase (Ino1, Rv0046c)	0,946	Yes	Low	Low	Low	Yes
3-Phosphoshikimate 1-carboxyvinyltransferase (Rv3227)	0,696	Yes	High	Low	Low	Yes
O-Acetylhomoserine aminocarboxypropyltransferase (Rv3340)	0,679	Yes	Low	Low	High	Yes
3-Oxoacyl-[acyl-carrier-protein] synthase 2 (Rv2246)	0,709	Yes	Low	Low	Low	Yes
Octanoyltransferase (Rv2217	0,703	Yes	Low	Low	Low	Yes
Bifunctional protein GlmU (Rv1018c)	0,833	Yes	High	Low	Low	Yes
Rv1465	0,802	Yes	Low	Low	Low	Yes

activity of INH is thought to be selective for mycobacteria, likely due to its ability to inhibit mycolic acid synthesis, Bb FabI and Mtb InhA (the protein target of INH) share the same and domains and are structural homologs, except for the presence of a long loop of interaction with the substrate found in InhA (Andrade et al., 2008). Moreover, it was recently shown that isoniazid in conjugation with nanoparticles could prevent the growth of *Enterococcus faecalis, E. coli, Pseudomonas aeruginosa,* and *S. aureus* (Zargarnezhad et al., 2020). This prodrug is activated by the heme enzyme catalase-peroxidase (KatG) endogenous to *M. tuberculosis.* Given this information, it is possible to propose INH, or its Mtb endogenous product (after reaction with KatG), as a potential compound for future trials against Bb.

Another interesting compound found was Fosmidomycin (CHEMBL203125). This compound has recently completed the clinical phase III for *Plasmodium* infections, although its mechanism of action is not entirely understood. It is reported that this compound is active against UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA) in *E. coli*. MurA shares the same Pfam domain (PF00275) as Bb. In this way, we can think of Fosmidomycin as an attractive seed compound to be used in drug discovery projects against this bacteria.

Mycobacterium tuberculosis Prioritized Targets and Their Potential Inhibitors

To further analyze the potential of the 743 Mtb proteins, which are essential and druggable, an analysis of available expression data under different infection mimicking conditions was previously performed (Starvation, Hypoxia, RNOS stress, and mice infection) (Defelipe et al., 2016). We found that 24 of these proteins were also overexpressed in at least three conditions, including DevS protein, known to be involved in RNOS sensing and signal transduction, harboring a druggable kinase ATP binding pocket.

As a last step in the prioritization procedure, a comprehensive Mtb metabolic network was built. As mentioned above, Target-Pathogen allowed us to score not individual proteins but entire pathways, according to their potential to be used as targets in latent tuberculosis drug discovery projects. In this framework, all pathways that do not have at least one druggable protein were ruled out, and a scoring function was developed to combine each protein data into a global network score.

This analysis revealed several high-scoring "druggable" pathways, which include a set of targets with great potential for further drug discovery projects (**Table 2**). One of them was the mycothiol biosynthesis pathway. Mycothiol is crucial for the intracellular redox balance and plays a crucial role in Mtb survival within macrophages (Buchmeier et al., 2003). Inositol-3-phosphate synthase (Ino1, Rv0046c), an enzyme involved in the early steps of this pathway, is highly druggable and over-expressed in RNOS stress, hypoxia, and starvation, three of the four latent infection mimicking conditions (**Figure 3A**). Mycolate biosynthesis pathway is also at the top of the ranking. Mycolate is an integral cell wall component of Mtb that participates in the survival ability of the bacilli within infected hosts, virulence, and evasion of the immune system. This pathway is targeted by first-line tuberculosis drugs such as isoniazid and ethambutol (Barry et al., 2007) and harbors the promising target 3-oxoacyl-[acyl-carrier protein] synthase 2 (KasB, Rv2246) involved in meromycolate extension. The scoring function also reveals the relevance of sulfur metabolism, essential for the bacilli's survival and virulence. Moreover, most genes are absent in humans. Among these pathways, methionine degradation to homocysteine is performed by the druggable protein Rv3340 (O-acetylhomoserine amino carboxypropyl transferase), another interesting target for future developments. Chorismate biosynthesis was another prioritized pathway. Chorismate is a key biochemical intermediate, being a precursor for aromatic amino acids. Within this pathway, 3-phosphoshikimate 1-carboxyvinyltransferase (Rv3227) could be selected for further studies. We found that it is druggable, essential, and not present in humans, and appears overexpressed under different conditions that mimic infections.

Among other top-scoring pathways revealed by our analysis are those related to lipoate synthesis. The two key genes (*lipA*, Rv2218 and *lipB*, and Rv2217) are essential, and *lipB* was also found to be druggable and expressed under starvation conditions. Moreover, the druggable pocket of LipB has Cys 176, Tyr22, and Tyr 91, making the pocket potentially sensitive to RNOS. Although this process is not ubiquitous in Bacteria, lipoate has been implicated in microbial pathogenesis, including immune response-induced oxidative and nitrosative stress in mycobacteria. It has also been acknowledged that lipoylated proteins take part in crucial antioxidant processes (Spalding and Prigge, 2010), thus promoting this high-scoring pathway from the target-finding aspect. Moreover, LipB has been structurally characterized and shown to have promising therapeutic properties (Ma et al., 2006). Other worth mentioning pathways are the UDP-N-acetyl-D-glucosamine

biosynthesis I and iron-sulfur cluster biosynthesis, which harbors attractive targets, such as Rv1018c (GlmU) and Rv1465 that have a set of desirable characteristics to be considered as promising targets to combat latent tuberculosis (**Table 2**). Regarding possible binders for Mtb proteome, 82, 10,768, 351, and 65,585 compounds make up seed I, II, III, IV sets, respectively. Whereas 19 ligands are obtained in the four seeds in parallel, which make them attractive compounds to treat Mtb infections (**Figure 3C**). Concerning the distribution of ligands by proteins (**Figure 3D**), 2,123 Mtb proteins interact with at least one compound, in the case of PDB (**Figure 3D**, top), 319 interact with a single ligand, 136 with two, and 1,668 with at least three (Purple shading). In the case of ChEMBL (**Figure 3D**, bottom), 125 interact with a single ligand, 55 with two, and 625 with at least three (purple shading). We could predict ligands for 2,123 Mtb proteins by searching in the PDB and 805 by looking in ChEMBL. The putative inhibitors found for the targets mentioned above are summarized in **Supplementary Table S4**. Worth to mention is Disulfiram (CHEMBL964) that was found to target GlmU. This drug inhibits enzymatic oxidation and is widely used to support the treatment of chronic alcoholism, different types of cancer, and parasitic infections. Furthermore, it has recently been proposed as an antibacterial compound against methicillin-resistant *S. aureus* (MRSA) and vancomycin-resistant *Enterococcus* (VRE) (Frazier et al., 2019) and particularly Mtb (Horita et al., 2012; Chaudhary et al., 2020).

DISCUSSION

In the last decades, antimicrobial drug development has observed a shift from the traditional approaches based mostly on phenotypic screening of natural/synthetic compounds to a rational genome-based target-driven lead discovery approach. Since wet-lab investigations of candidate targets and lead compounds are time-consuming and expensive, it is worthwhile to conduct bioinformatic analyses to identify the proteins and ligands most worthy of experimental follow-up. *In silico* analyses are particularly important in developing countries (such as those from Latin America), where the research investment is usually limited.

Our developed bioinformatics pipeline and the underlying methodology, briefly presented here for Bb and Mtb, and freely available to the scientific community at http://target.sbg.qb.fcen.uba.ar/patho/, allows starting from a pathogen whole genome, the modeling and classification of the proteome. General results show that a large fraction of protein structures harbor a druggable pocket (60–85%). Interestingly, effects on the essential proteins yield substantial differences between both bacterial pathogens. While 20% of Mtb druggable proteins were considered essential, only 6% of the Bb proteome resulted in druggable and essential proteins. These differences could be explained by the different amounts of data available, particularly in terms of essentiality and knowledge on gene/protein function for both microorganisms, Mtb and Bb. Although there is vast information available for Mtb, Bb is a neglected disease with only a regional impact. Specifically, Mtb essentiality criteria were based on experimental mutagenesis studies; meanwhile, Bb essential genes were inferred by homology analysis with the Database of Essential Genes.

The results presented here (further expanded in the web Target-Pathogen) provide two crucial assets for those researchers in the field of Bb or Mtb antimicrobial development. In the first place, we provide a shortlist of attractive protein targets in each pathogen (**Tables 1**, **2**). We also provide a detailed analysis of those characteristics that make it a good target for each gene/protein (essentiality, druggability, biological relevant role, and lack of cross-reactivity with the host). We hope this analysis will allow wet-lab researchers to develop upon the targets disclosed herein, moving research forward. The second, and more important issue, is that we provide a list of potential inhibitors and their chemical scaffolds for several prioritized targets. (**Supplementary Tables S3, S4**). We expect that researchers working with those targets and familiar with whole-cell and protein-based *in vitro* will become interested and directly try some of these compounds for their antimicrobial activity. Furthermore, the current pipeline is also presented for other targets and pathogens in our freely accessible website, thus providing the community with a general platform to drive the development of antimicrobial compounds forward.

In silico approaches are rapid, efficient, and cost-effective techniques for screening drug targets and narrowing the search space of drug like-compounds for any given pathogen. The goal of these techniques is not to replace wet-lab strategies. Instead, it is to become a useful resource for researchers working in target identification and drug discovery to translate biological questions in a computationally tractable way by filtering and weighting the vast quantity of genome-scale data sets. High-throughput screening (HTS) campaigns against molecular targets *in vitro*, although extremely valuable, typically do not yield directly good antimicrobial compounds (Payne et al., 2007; Tommasi et al., 2015). The bacterial cell envelope has evolved to refract toxic compounds from entering into the cell and even those drugs that cross the barrier, can be extruded by efflux pumps in multidrug resistant bugs (Li et al., 2015), thus resulting in poor *in vivo* activity. Another limitation of HTS approaches, is that only a finite amount of chemicals, with limited diversity, are available in any given library. As this chemical space limitation will hardly be overcome, novel approaches are needed to tackle the ongoing problem of bacterial resistance to current treatments. Our work provides a framework for which such novel strategies can be developed and further adapted to use by mean sized research laboratories including those of developing countries. Our strategy looks first for potential best molecular targets, and subsequently applies *in silico* screening to find best drug candidates. Novel methods for drug delivery, in particular nanomaterials and molecular transporters have started to be investigated as alternative antibacterials or anti-infective carrier systems to improve the internalization of bactericidal drugs against bacterial infections, which are particularly problematic in the case of having to reach the cytoplasm, specially in Gram-negative pathogens. Some of these promising molecules that could help overcome the bacterial envelopes and are currently being tested are siderophores, cyclodextrins, and metal nanoparticles, antimicrobial/cell-penetrating peptides and fusogenic liposomes (Santos et al.,

2018). In this sense, we believe in a strategy that combines omics data and drug screening to discover lead antimicrobials, in which *in silico* and wet-lab approaches act synergically to maximize the success rate of drug discovery projects.

AUTHOR CONTRIBUTIONS

PR, MN, AT, MM, and DP conceived the study design. FS, FC, ES, AP, MC, CM, and PR, MN, AT, MM, DP contributed tools and performed data analysis. PR, MN, MM, DP, AT drafted the manuscript with input from the other authors. All authors read and approved the final version of the manuscript.

FUNDING

Fellowships supported MN from CNPq (process no. 306894/2019-0) and granted by CAPES (process no. 88887.368759/2019-00). Fellowship supported from CONICET to AP, CM, FS and FC. AP, FS, DP, MM, JA, MP, and AT are members of CONICET research career. This work was supported by Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT, PICT-2018-04663 to DP). DFD, JM, ES, NY, MM, BL, AT, and VR and Universidad de Buenos Aires (20020190200275BA to DP).

ACKNOWLEDGMENTS

We acknowledge Martín Palomino for graphic design.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.647060/full#supplementary-material

REFERENCES

Abd Algfoor, Z., Shahrizal Sunar, M., Abdullah, A., and Kolivand, H. (2017). Identification of Metabolic Pathways Using Pathfinding Approaches: a Systematic Review. *Brief. Funct. Genomics* 16, 87–98. doi:10.1093/bfgp/elw002

Sussman, J. L., Lin, D., Jiang, J., Manning, N. O., Prilusky, J., Ritter, O., et al. (1998). Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. *Acta Crystallogr. D Biol. Crystallogr.* 54, 1078–1084.

Andrade, C. H., Pasqualoto, K. F. M., Zaim, M. H., and Ferreira, E. I. (2008). Abordagem racional no planejamento de novos tuberculostáticos: inibidores da InhA, enoil-ACP redutase Do *M. tuberculosis*. *Rev. Bras. Cienc. Farm.* 44, 167–179. doi:10.1590/s1516-93322008000200002

Ascenzi, P., and Visca, P. (2008). Scavenging of Reactive Nitrogen Species by Mycobacterial Truncated Hemoglobins. *Methods Enzymol.* 436, 317–337. doi:10.1016/s0076-6879(08)36018-2

Ashtiani, M., Salehzadeh-Yazdi, A., Razaghi-Moghadam, Z., Hennig, H., Wolkenhauer, O., Mirzaie, M., et al. (2018). A Systematic Survey of Centrality Measures for Protein-Protein Interaction Networks. *BMC Syst. Biol.* 12, 80. doi:10.1186/s12918-018-0598-2

Bajusz, D., Rácz, A., and Héberger, K. (2015). Why Is Tanimoto index an Appropriate Choice for Fingerprint-Based Similarity Calculations? *J. Cheminform.* 7, 20. doi:10.1186/s13321-015-0069-3

Barh, D., Tiwari, S., Jain, N., Ali, A., Santos, A. R., Misra, A. N., et al. (2011). In Silico subtractive Genomics for Target Identification in Human Bacterial Pathogens. *Drug Dev. Res.* 72, 162–177. doi:10.1002/ddr.20413

Barry, C. E., Crick, D. C., and McNeil, M. R. (2007). Targeting the Formation of the Cell wall Core of *M. tuberculosis*. *Infect. Disord. Drug Targets* 7, 182–202. doi:10.2174/187152607781001808

Battisti, J. M., Smitherman, L. S., Samuels, D. S., and Minnick, M. F. (1998). Mutations in Bartonella Bacilliformis gyrB Confer Resistance to Coumermycin A1. *Antimicrob. Agents Chemother.* 42, 2906–2913. doi:10.1128/aac.42.11.2906

Biswas, S., Raoult, D., and Rolain, J.-M. (2007). Molecular Mechanisms of Resistance to Antibiotics in Bartonella Bacilliformis. *J. Antimicrob. Chemother.* 59, 1065–1070. doi:10.1093/jac/dkm105

Buchmeier, N. A., Newton, G. L., Koledin, T., and Fahey, R. C. (2003). Association of Mycothiol with protection of *Mycobacterium tuberculosis* from Toxic Oxidants and Antibiotics. *Mol. Microbiol.* 47, 1723–1732. doi:10.1046/j.1365-2958.2003.03416.x

Chaudhary, D., Marzuki, M., Lee, A., Bouzeyen, R., Singh, A., Gosain, T. P., et al. (2020). Disulfiram Inhibits *M. tuberculosis* Growth by Altering Methionine Pool, Redox Status and Host-Immune Response. *Cold Spring Harbor Lab.* doi:10.1101/2020.09.01.277368

Chen, L., Oughtred, R., Berman, H. M., and Westbrook, J. (2004). TargetDB: a Target Registration Database for Structural Genomics Projects. *Bioinformatics* 20, 2860–2862. doi:10.1093/bioinformatics/bth300

Clemente, N. S., Ugarte-Gil, C. A., Solórzano, N., Maguiña, C., Pachas, P., Blazes, D., et al. (2012). Bartonella Bacilliformis: A Systematic Review of the Literature to Guide the Research Agenda for Elimination. *PLoS Negl. Trop. Dis.* 6, e1819. doi:10.1371/journal.pntd.0001819

Cottret, L., and Jourdan, F. (2010). Graph Methods for the Investigation of Metabolic Networks in Parasitology. *Parasitology* 137, 1393–1407. doi:10.1017/s0031182010000363

Davies, J., and Davies, D. (2010). Origins and Evolution of Antibiotic Resistance. *Mmbr* 74, 417–433. doi:10.1128/mmbr.00016-10

Defelipe, L. A., Do Porto, D. F., Pereira Ramos, P. I., Nicolás, M. F., Sosa, E., Radusky, L., et al. (2016). A Whole Genome Bioinformatic Approach to Determine Potential Latent Phase Specific Targets in *Mycobacterium tuberculosis*. *Tuberculosis* 97, 181–192. doi:10.1016/j.tube.2015.11.009

del Valle, L. J., Flores, L., Vargas, M., García-de-la-Guarda, R., Quispe, R. L., Ibañez, Z. B., et al. (2010). Bartonella Bacilliformis, Endemic Pathogen of the Andean Region, Is Intrinsically Resistant to Quinolones. *Int. J. Infect. Dis.* 14, e506–10. doi:10.1016/j.ijid.2009.07.025

Devoid, S., Overbeek, R., DeJongh, M., Vonstein, V., Best, A. A., and Henry, C. (2013). *Automated Genome Annotation and Metabolic Model Reconstruction in the SEED and Model SEED," in Systems Metabolic Engineering*. Totowa, NJ: Humana Press, 17–45. doi:10.1007/978-1-62703-299-5_2

DiMasi, J. A., Grabowski, H. G., and Hansen, R. W. (2016). Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs. *J. Health Econ.* 47, 20–33. doi:10.1016/j.jhealeco.2016.01.012

Espinoza-Culupú, A., Quispe-Gaspar, R., Jaramillo, M., Icho, M., Eca, A., Ramírez, P., et al. (2014). Caracterización molecular de la región determinante de resistencia a quinolonas (QRDR) de la topoisomerasa IV de Bartonella bacilliformis en aislados clínicos. *Rev. Peru Biol.* 21, 89–98. doi:10.15381/rpb.v21i1.8251

Farfán-López, M., Espinoza-Culupú, A., García-de-la-Guarda, R., Serral, F., Sosa, E., Palomino, M. M., et al. (2020). Prioritisation of Potential Drug Targets against Bartonella Bacilliformis by an Integrative In-Silico Approach. *Mem. Inst. Oswaldo Cruz* 115. e200184. doi:10.1590/0074-02760200184

Forsyth, R. A., Haselbeck, R. J., Ohlsen, K. L., Yamamoto, R. T., Xu, H., Trawick, J. D., et al. (2002). A Genome-wide Strategy for the Identification of Essential Genes in *Staphylococcus aureus*. *Mol. Microbiol.* 43, 1387–1400. doi:10.1046/j.1365-2958.2002.02832.x

Frazier, K. R., Moore, J. A., and Long, T. E. (2019). Antibacterial Activity of Disulfiram and its Metabolites. *J. Appl. Microbiol.* 126, 79–86. doi:10.1111/jam.14094

Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., et al. (2012). ChEMBL: a Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res.* 40, D1100–D1107. doi:10.1093/nar/gkr777

Gomes, C., Martínez-Puchol, S., Ruiz-Roldán, L., Pons, M. J., del Valle Mendoza, J., and Ruiz, J. (2016). Development and Characterisation of Highly Antibiotic Resistant Bartonella Bacilliformis Mutants. *Sci. Rep.* 6. 33584. doi:10.1038/srep33584

Gomes, C., and Ruiz, J. (2017). Carrion's Disease: the Sound of Silence. *Clin. Microbiol. Rev.* 31. e00056. doi:10.1128/cmr.00056-17

Green, M. L., and Karp, P. D. (2006). The Outcomes of Pathway Database Computations Depend on Pathway Ontology. *Nucleic Acids Res.* 34, 3687–3697. doi:10.1093/nar/gkl438

Guilloux, V. L., Le Guilloux, V., Schmidtke, P., and Tuffery, P. (2009). Fpocket: An Open Source Platform for Ligand Pocket Detection. *BMC Bioinformatics* 10, 168. doi:10.1186/1471-2105-10-168

Hackbarth, C. J., Chen, D. Z., Lewis, J. G., Clark, K., Mangold, J. B., Cramer, J. A., et al. (2002). N-alkyl Urea Hydroxamic Acids as a New Class of Peptide Deformylase Inhibitors with Antibacterial Activity. *Aac* 46, 2752–2764. doi:10.1128/aac.46.9.2752-2764.2002

Harding, E. (2020). WHO Global Progress Report on Tuberculosis Elimination. *Lancet Respir. Med.* 8, 19. doi:10.1016/s2213-2600(19)30418-7

Hay, M., Thomas, D. W., Craighead, J. L., Economides, C., and Rosenthal, J. (2014). Clinical Development success Rates for Investigational Drugs. *Nat. Biotechnol.* 32, 40–51. doi:10.1038/nbt.2786

Heath, R. J., Yu, Y.-T., Shapiro, M. A., Olson, E., and Rock, C. O. (1998). Broad Spectrum Antimicrobial Biocides Target the FabI Component of Fatty Acid Synthesis. *J. Biol. Chem.* 273, 30316–30320. doi:10.1074/jbc.273.46.30316

Horita, Y., Takii, T., Yagi, T., Ogawa, K., Fujiwara, N., Inagaki, E., et al. (2012). Antitubercular Activity of Disulfiram, an Antialcoholism Drug, against Multidrug- and Extensively Drug-Resistant *Mycobacterium tuberculosis* Isolates. *Antimicrob. Agents Chemother.* 56, 4140–4145. doi:10.1128/aac.06445-11

Jalili, M., Salehzadeh-Yazdi, A., Gupta, S., Wolkenhauer, O., Yaghmaie, M., Resendis-Antonio, O., et al. (2016). Evolution of Centrality Measurements for the Detection of Essential Proteins in Biological Networks. *Front. Physiol.* 7, 375. doi:10.3389/fphys.2016.00375

Kanehisa, M., and Sato, Y. (2020). KEGG Mapper for Inferring Cellular Functions from Protein Sequences. *Protein Sci.* 29, 28–35. doi:10.1002/pro.3711

Kaplan, N., Albert, M., Awrey, D., Bardouniotis, E., Berman, J., Clarke, T., et al. (2012). Mode of Action,In VitroActivity, andIn VivoEfficacy of AFN-1252, a Selective Antistaphylococcal FabI Inhibitor. *Antimicrob. Agents Chemother.* 56, 5865–5874. doi:10.1128/aac.01411-12

Karp, P. D., Latendresse, M., Paley, S. M., Krummenacker, M., Ong, Q. D., Billington, R., et al. (2015). Pathway Tools Version 19.0 Update: Software for Pathway/genome Informatics and Systems Biology. *Brief. Bioinform.* 17, 877–890. doi:10.1093/bib/bbv079

Kaur, D., Kutum, R., Dash, D., and Brahmachari, S. K. (2017). Data Intensive Genome Level Analysis for Identifying Novel, Non-toxic Drug Targets for Multi Drug Resistant. *Mycobacterium tuberculosis. Sci. Rep.* 7, 46595. doi:10.1038/srep46825

Lau, H. J., Lim, C. H., Foo, S. C., and Tan, H. S. (2021). The Role of Artificial Intelligence in the Battle against Antimicrobial-Resistant Bacteria. *Curr. Genet.* 67, 421–429. doi:10.1007/s00294-021-01156-5

Li, X.-Z., Plésiat, P., and Nikaido, H. (2015). The challenge of Efflux-Mediated Antibiotic Resistance in Gram-Negative Bacteria. *Clin. Microbiol. Rev.* 28, 337–418. doi:10.1128/cmr.00117-14

Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001). Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings 1PII of Original Article: S0169-409X(96)00423-1. The Article Was Originally Published in Advanced Drug Delivery Reviews 23 (1997) 3-25. 1. *Adv. Drug Deliv. Rev.* 46, 3–26. doi:10.1016/s0169-409x(00)00129-0

Ma, Q., Zhao, X., Eddine, A. N., Geerlof, A., Li, X., Cronan, J. E., et al. (2006). The *Mycobacterium tuberculosis* LipB Enzyme Functions as a Cysteine/lysine Dyad Acyltransferase. *Proc. Natl. Acad. Sci.* 103, 8662–8667. doi:10.1073/pnas.0510436103

MacNair, C. R., Tsai, C. N., and Brown, E. D. (2020). Creative Targeting of the Gram-negative Outer Membrane in Antibiotic Discovery. *Ann. N.Y. Acad. Sci.* 1459, 69–85. doi:10.1111/nyas.14280

Magariños, M. P., Carmona, S. J., Crowther, G. J., Ralph, S. A., Roos, D. S., Shanmugam, D., et al. (2012). TDR Targets: a Chemogenomics Resource for Neglected Diseases. *Nucleic Acids Res.* 40, D1118–D1127. doi:10.1093/nar/gkr1053

Merigueti, T. C., Carneiro, M. W., Carvalho-Assef, A. P. D., Silva-, F. P., Jr, and da Silva, F. A. B. (2019). FindTargetsWEB: A User-Friendly Tool for Identification of Potential Therapeutic Targets in Metabolic Networks of Bacteria. *Front. Genet.* 10, 633. doi:10.3389/fgene.2019.00633

Minnick, M. F., Anderson, B. E., Lima, A., Battisti, J. M., Lawyer, P. G., and Birtles, R. J. (2014). Oroya Fever and Verruga Peruana: Bartonelloses Unique to South America. *Plos Negl. Trop. Dis.* 8, e2919. doi:10.1371/journal.pntd.0002919

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., et al. (2020). Pfam: The Protein Families Database in 2021. *Nucleic Acids Res.* 49, D412–D419. doi:10.1093/nar/gkaa913

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., and Hutchison, G. R. (2011). Open Babel: An Open Chemical Toolbox. *J. Cheminform.* 3, 33. doi:10.1186/1758-2946-3-33

Payne, D. J., Gwynn, M. N., Holmes, D. J., and Pompliano, D. L. (2007). Drugs for Bad Bugs: Confronting the Challenges of Antibacterial Discovery. *Nat. Rev. Drug Discov.* 6, 29–40. doi:10.1038/nrd2201

Pitkänen, E., Rousu, J., and Ukkonen, E. (2010). Computational Methods for Metabolic Reconstruction. *Curr. Opin. Biotechnol.* 21, 70–77. doi:10.1016/j.copbio.2010.01.010

Pons, M. J., Gomes, C., del Valle-Mendoza, J., and Ruiz, J. (2016). Carrion's Disease: More Than a Sand Fly-Vectored Illness. *Plos Pathog.* 12, e1005863. doi:10.1371/journal.ppat.1005863

Radusky, L., Defelipe, L. A., Lanzarotti, E., Luque, J., Barril, X., Marti, M. A., et al. (2014). TuberQ: a *Mycobacterium tuberculosis* Protein Druggability Database. *Database* 2014, bau035. doi:10.1093/database/bau035

Radusky, L. G., Hassan, S., Lanzarotti, E., Tiwari, S., Jamal, S., Ali, J., et al. (2015). An Integrated Structural Proteomics Approach along the Druggable Genome of Corynebacterium Pseudotuberculosis Species for Putative Druggable Targets. *BMC Genomics* 16 (Suppl. 5), S9. doi:10.1186/1471-2164-16-s5-s9

Radusky, L., Ruiz-Carmona, S., Modenutti, C., Barril, X., Turjanski, A. G., and Martí, M. A. (2017). LigQ: A Webserver to Select and Prepare Ligands for Virtual Screening. *J. Chem. Inf. Model.* 57, 1741–1746. doi:10.1021/acs.jcim.7b00241

Raman, K., Yeturu, K., and Chandra, N. (2008). targetTB: a Target Identification Pipeline for *Mycobacterium tuberculosis* through an Interactome, Reactome and Genome-Scale Structural Analysis. *BMC Syst. Biol.* 2, 109. doi:10.1186/1752-0509-2-109

Ramos, P. I. P., Fernández Do Porto, D., Lanzarotti, E., Sosa, E. J., Burguener, G., Pardo, A. M., et al. (2018). An Integrative, Multi-Omics Approach towards the Prioritization of *Klebsiella pneumoniae* Drug Targets. *Sci. Rep.* 8, 10755. doi:10.1038/s41598-018-28916-7

Rolain, J. M., Brouqui, P., Koehler, J. E., Maguina, C., Dolan, M. J., and Raoult, D. (2004). Recommendations for Treatment of Human Infections Caused by Bartonella Species. *Antimicrob. Agents. Chemother.* 48, 1921–1933. doi:10.1128/aac.48.6.1921-1933.2004

Santos, R. S., Figueiredo, C., Azevedo, N. F., Braeckmans, K., and De Smedt, S. C. (2018). Nanomaterials and Molecular Transporters to Overcome the Bacterial Envelope Barrier: Towards Advanced Delivery of Antibiotics. *Adv. Drug Deliv. Rev.* 136-137, 28–48. doi:10.1016/j.addr.2017.12.010

Sheridan, R. P., Maiorov, V. N., Holloway, M. K., Cornell, W. D., and Gao, Y.-D. (2010). Drug-like Density: A Method of Quantifying the "Bindability" of a Protein Target Based on a Very Large Set of Pockets and Drug-like Ligands from the Protein Data Bank. *J. Chem. Inf. Model.* 50, 2029–2040. doi:10.1021/ci100312t

Sosa, E., Burguener, G., Pardo, A., Marti, M., Turjanski, A., and Porto, D. A. F. D. (2018). Target-Pathogen: A Structural Bioinformatic Approach to Prioritize Drug Targets in Pathogens. *Int. J. Infect. Dis.* 73, 84. doi:10.1016/j.ijid.2018.04.3616

Spalding, M. D., and Prigge, S. T. (2010). Lipoic Acid Metabolism in Microbial Pathogens. *Microbiol. Mol. Biol. Rev.* 74, 200–228. doi:10.1128/mmbr.00008-10

Stobbe, M. D., Jansen, G. A., Moerland, P. D., and van Kampen, A. H. C. (2012). Knowledge Representation in Metabolic Pathway Databases. *Brief. Bioinform.* 15, 455–470. doi:10.1093/bib/bbs060

Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., et al. (2020). A Deep Learning Approach to Antibiotic Discovery. *Cell* 180, 688–702. doi:10.1016/j.cell.2020.01.021

Storek, K. M., Auerbach, M. R., Shi, H., Garcia, N. K., Sun, D., Nickerson, N. N., et al. (2018). Monoclonal Antibody Targeting the β-barrel Assembly Machine ofEscherichia Coliis Bactericidal. *Proc. Natl. Acad. Sci. USA* 115, 3692–3697. doi:10.1073/pnas.1800043115

Tommasi, R., Brown, D. G., Walkup, G. K., Manchester, J. I., and Miller, A. A. (2015). ESKAPEing the Labyrinth of Antibacterial Discovery. *Nat. Rev. Drug Discov.* 14, 529–542. doi:10.1038/nrd4572

Uddin, R., and Jamil, F. (2018). Prioritization of Potential Drug Targets against *P. aeruginosa* by Core Proteomic Analysis Using Computational Subtractive Genomics and Protein-Protein Interaction Network. *Comput. Biol. Chem.* 74, 115–122. doi:10.1016/j.compbiolchem.2018.02.017

Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., et al. (2019). Applications of Machine Learning in Drug Discovery and Development. *Nat. Rev. Drug Discov.* 18, 463–477. doi:10.1038/s41573-019-0024-5

Volkamer, A., Kuhn, D., Rippmann, F., and Rarey, M. (2012). DoGSiteScorer: a Web Server for Automatic Binding Site Prediction, Analysis and Druggability Assessment. *Bioinformatics* 28, 2074–2075. doi:10.1093/bioinformatics/bts310

Wadood, A., Ghufran, M., Khan, A., Azam, S. S., Uddin, R., Waqas, M., et al. (2017). The Methicillin-Resistant S. Epidermidis Strain RP62A Genome Mining for Potential Novel Drug Targets Identification. *Gene Rep.* 8, 88–93. doi:10.1016/j.genrep.2017.06.002

Woodman, M., Haeusler, I., and Grandjean, L. (2019). Tuberculosis Genetic Epidemiology: A Latin American Perspective. *Genes* 10, 53. doi:10.3390/genes10010053

Yeh, I., Hanekamp, T., Tsoka, S., Karp, P. D., and Altman, R. B. (2004). Computational Analysis of Plasmodium Falciparum Metabolism: Organizing Genomic Information to Facilitate Drug Discovery. *Genome Res.* 14, 917–924. doi:10.1101/gr.2050304

Zargarnezhad, S., Gholami, A., Khoshneviszadeh, M., Abootalebi, S. N., and Ghasemi, Y. (2020). Antimicrobial Activity of Isoniazid in Conjugation with Surface-Modified Magnetic Nanoparticles against *Mycobacterium tuberculosis* and Nonmycobacterial Microorganisms. *J. Nanomater.*, 2020 , 1–. doi:10.1155/2020/7372531

Zhan, P., Pannecouque, C., De Clercq, E., and Liu, X. (2016). Anti-HIV Drug Discovery and Development: Current Innovations and Future Trends. *J. Med. Chem.* 59, 2849–2878. doi:10.1021/acs.jmedchem.5b00497

Gashaw, I., Ellinghaus, P., Sommer, A., and Asadullah, K. (2012). What makes a good drug target? *Drug Discovery Today* 17, S24–S30. doi:10.1016/j.drudis.2011.12.008

Garcia-Quintanilla, M., Dichter, A., Guerra, H., and Kempf, V. A. J. (2019). Carrion's disease: more than a neglected disease. *Parasites Vectors* 12. doi:10.1186/s13071-019-3390-2

Webb, B., and Sali, A. (2016). Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Protein Sci.* 86, 1–37.

Wallner, B., and Elofsson, A. (2005). All are not equal: A benchmark of different homology modeling programs. *Protein Sci.* 14, 1315–1327. doi:10.1110/ps.041253405

Guillen, Y., Casadellà, M., García-de-la-Guarda, R., Espinoza-Culupú, A., Paredes, R., Ruiz, J., et al. (2016). Whole-Genome Sequencing of Two Bartonella bacilliformis Strains. *Genome Announc.* 4. doi:10.1128/genomeA.00659-16

Yim, L., Moukadiri, I., Björk, G. R., and Armengod, M.-E. (2006). Further insights into the tRNA modification process controlled by proteins MnmE and GidA of Escherichia coli. *Nucleic Acids Res.* 34, 5892–5905.

Gupta, R., Gobble, T. R., and Schuster, M. (2009). GidA posttranscriptionally regulates rhl quorum sensing in Pseudomonas aeruginosa. *J. Bacteriol.* 191, 5785–5792.

Shippy, D. C., Eakley, N. M., Bochsler, P. N., and Fadl, A. A. (2011). Biological and virulence characteristics of Salmonella enterica serovar Typhimurium following deletion of glucose-inhibited division (gidA) gene. *Microbial. Pathog.* 50, 303–313.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 20 September 2021
doi: 10.3389/fgene.2021.640956

Single Nucleotide Polymorphisms in *Apolipoprotein B, Apolipoprotein E*, and *Methylenetetrahydrofolate Reductase* Are Associated With Serum Lipid Levels in Northern Chilean Subjects. A Pilot Study

*Anita S. Gálvez[1], Hugo Ramírez[1], Pablo Placencia[1], Claudio Rojas[1], Ximena Urzúa[1], Alexis M. Kalergis[2,3], Luis A. Salazar[4] and Jorge Escobar-Vera[1]**

[1] Laboratorio de Genética, Departamento Biomédico, Facultad de Ciencias de la Salud, Universidad de Antofagasta, Antofagasta, Chile, [2] Departamento de Genética Molecular y Microbiología, Facultad de Ciencias Biológicas, Millennium Institute on Immunology and Immunotherapy (MIII), Pontificia Universidad Católica de Chile, Santiago, Chile, [3] Departamento de Endocrinología, Facultad de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile, [4] Departamento de Ciencias Básicas, Facultad de Medicina, Centro de Biología Molecular & Farmacogenética, Universidad de La Frontera, Temuco, Chile

OPEN ACCESS

Edited by:
Matias F. Martinez,
University of Chile, Chile

Reviewed by:
Cheryl D. Cropp,
Samford University, United States
Yogita Ghodke-Puranik,
Icahn School of Medicine at Mount Sinai, United States

***Correspondence:**
Jorge Escobar-Vera
jorge.escobar@uantof.cl

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Genetics

Received: 12 December 2020
Accepted: 24 August 2021
Published: 20 September 2021

Citation:
Gálvez AS, Ramírez H, Placencia P, Rojas C, Urzúa X, Kalergis AM, Salazar LA and Escobar-Vera J (2021) Single Nucleotide Polymorphisms in Apolipoprotein B, Apolipoprotein E, and Methylenetetrahydrofolate Reductase Are Associated With Serum Lipid Levels in Northern Chilean Subjects. A Pilot Study. Front. Genet. 12:640956. doi: 10.3389/fgene.2021.640956

Characterization of allelic variants is relevant to demonstrate associations among genetic background and susceptibility to develop cardiovascular diseases, which are the main cause of death in Chile. Association of APOB, APOE, and MTHFR polymorphisms with higher lipid levels and the risk of developing hypertension and cardiovascular diseases have been described. Thus, the aim of this study was to assess genotype distribution and relative allelic frequency of ApoB rs693, ApoE rs7412, ApoE rs429358, MTHFR rs1801131, and MTHFR rs1801133 allelic variants and their effects on lipid profile in young healthy men and women from Northern Chile. A group of 193 healthy subjects were enrolled for this study. Genotyping of rs693 (*APOB*), rs7412 and rs429358 (*APOE*), and rs1801131 and rs1801133 (*MTHFR*) polymorphisms were performed by real time PCR. In addition, lipid profiles were determined and associated to genetic data. The genotype distribution was *APOB* rs693 (CC = 37%, CT = 41%, and TT = 22%), *APOE* rs7412/rs429358 (E4 = 0.06, E3 = 0.91, and E2 = 0.03), *MTHFR* rs1801131 (AA = 57%, AC = 30%, and CC = 13%), and *MTHFR* rs1801133 (CC = 20%, CT = 47%, and TT = 33%). The association of the genetic variants with plasma lipid levels showed that women, but not men, carrying *APOB* mutated allele (T) and Apo E4 allele presented lower values of total cholesterol when compared with C/C homozygous genotype or E3 allele, respectively ($p < 0.05$). In addition, a subgroup analysis revealed that ApoB C/C homozygous women exhibited higher values of HDL-C when compared with men carrying identical genotype ($p < 0.01$). On the other hand, women carrying E4 allele exhibited lower values of triglycerides when compared with male carrying identical genotype ($p < 0.05$). Finally, women carrying mutate allele (C) for *MTHFR* rs1801131

showed lower levels of triglycerides when compared with A/A homozygous genotype ($p < 0.05$) and lower levels of LDL-C for *MTHFR* rs1801133 in females carrying (T) allele when compared with males carrying identical genotype ($p < 0.05$). In summary, the present data showed that *APOB, APOE*, *and* MTHFR single nucleotide polymorphisms are associated to lipid levels in a gender-dependent manner among healthy subjects from Northern Chile, especially in women.

Keywords: Lipid levels, genetic variants, Apo B, Apo E, MTHFR, polymorphism

INTRODUCTION

Studies based on genetic markers are frequently used to evaluate the relationships between the presence of an allelic variant and the susceptibility to develop chronic diseases. However, the precise magnitude of inheritance is modified depending on the polygenic model, including other factors such as disease type and age onset (Smolková et al., 2015); therefore, one of the main goals of biomedical research is to correlate genotype with biochemical or molecular abnormalities (Liggett et al., 2007). On this issue, lipoproteins, playing a main function in lipid transport and metabolism (Defesche et al., 2017), along with hyperlipidemia are considerably important in the development of cardiovascular diseases (CVDs) (Homsma et al., 2008), which are the main cause of death in Chile during the past years (DEIS, 2021). In this respect, the National Health Survey showed that around 40% of individuals have a high prevalence of two or more major risk factors, including low high-density lipoprotein cholesterol (HDL-C) concentrations (39%) and hypercholesterolemia (35%), among others (ENS, 2003). Although CVD remains the leading cause of death in both men and women in Chile, the latest report revealed that CVD-related mortality is higher in women (28.9%) than in men (26.2%) (Varleta et al., 2020). In addition, in-hospital mortality due to acute myocardial infarction was significantly greater in women than in men at any age (Nazzal and Alonso, 2013); however, according to some reports, women are not fully aware of this situation (Varleta et al., 2020). Therefore, a better knowledge about lipoprotein allelic variants and their effects to regulate metabolism and transport decreasing cholesterol and lipid blood levels in young healthy individuals, especially in women, may result in improving therapeutic treatments along with preventive public health strategies.

Despite the fact that a genome-wide association study focuses on association of numerous single nucleotide polymorphisms (SNPs) as molecular markers of a disease, analyses of specific genetic variants are helpful in the early detection of individuals carrying genetic susceptibility for hyperlipidemia (Defesche et al., 2017). Apolipoprotein (Apo) B gene, located on the short arm of chromosome 2 with a span of 43 kilobases and 29 exons (Innerarity et al., 1990), is the most important of the atherogenic lipoproteins, having a critical function in the dynamic equilibrium of cholesterol being relevant for the assembly and secretion of very low-density lipoproteins (VLDL) from the liver and uptake of low-density lipoproteins (LDL) mediated by LDL receptor (LDLR). Therefore, elevated Apo B and LDL-cholesterol (LDL-C) levels are risk factors for atherosclerosis; in contrast, low levels of Apo B may contribute to protect against atherosclerosis (Benn, 2009). Specifically, Apo B rs693 genetic variant (*Xba*I, Thr2515Thr) consisted of a change from cytocine (C) to thymine (T) in exon 26, codon 2488 creating a synonymous mutation from ACC to ACT with no change of amino acid residue Thr (Niu et al., 2017). On the other hand, Apo E gene, located on chromosome 19, contains four exons encoding a 34,200-Da protein of 299 amino acids. Apo E is a multifunctional protein synthesized mainly in the liver, taking part in the reverse cholesterol transport and triglyceride metabolism. Apo E binds to hepatic Apo E or LDLR to mediate chylomicrons and VLDL clearances from the plasma (Eichner et al., 2002). Two single-nucleotide polymorphism variants in the Apo E gene (rs429358 and rs7412) will generate different protein isoforms with opposite effects on lipid metabolism because they differ in their receptor-binding activity. Distinction based on cysteine (Cys) or arginine (Arg) at positions 112 and 158 amino acid residue determines three major isoforms commonly referred to as E2 (Cys112/Cys158), E3 (Cys112/Arg158), and E4 (Arg112/Arg158). These isoforms produce six genotypes, including three homozygous (E2/E2, E3/E3 and E4/E4) and three heterozygous (E3/E2, E4/E3 and E4/E2), which produce three isoforms of Apo E, referred to as Apo E2, Apo E3, and Apo E4 (Eichner et al., 2002; Khan et al., 2013). In addition, the enzyme methylenetetrahydrofolate reductase (MTHFR), which catalyzes the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, presents several allelic variants associated to modifications in its activity, resulting in hyperhomocysteinemia as an emerging risk factor for various cardiovascular, cerebrovascular, and neurological diseases along with several cancers (Bailey and Gregory, 1999; Liu et al., 2020). MTHFR gene, located in chromosome 1, encoded two enzyme isoforms of 70 and 77 kDa due to multiple transcription start sites (Tran et al., 2002). The two most common MTHFR genetic variants are rs1801133, C677T resulting in alanine (C) to valine (T) substitution with a reduction of enzymatic activity and rs1801131, A1298C resulting in glutamic acid (A) to alanine (C) substitution, showing a significant enzymatic activity decrease

Abbreviations: ANOVA, Analysis of variance; Apo B, Apolipoprotein B; Apo E, Apolipoprotein E; BMI, Body Mass Index; CVD, cardiovascular diseases; DNA, Deoxyribonucleic acid; dNTPs, deoxyribonucleotides; DBP, Diastolic blood pressure; HBP, high blood pressure; HDL-C, high-density lipoprotein-cholesterol; HWE, HardyŰWeinberg equilibrium; IAM, acute myocardial infarction; LDL-C, low-density lipoprotein-cholesterol; MTHFR, methylenetetrahydrofolate reductase; MgCl2, Magnesium chloride; PCR, Polymerase chain reaction; SNP, Single nucleotide polymorphism; SBP, Systolic blood pressure; VLDL, very low-density lipoproteins.

in individuals carrying C/C genotype (Abu-Hassan et al., 2019; Mahesh et al., 2019). *MTHFR* rs1801133 polymorphism was associated to elevated risk of myocardial infarction in young and middle-aged Caucasian individuals (Liew and Gupta, 2015); however, this issue remains controversial showing both association between MTHFR rs1801133 and hypertension and the lack of association (Amrani-Midoun et al., 2016; Ghogomu et al., 2016).

Consequently, the current study aims to determine the frequencies of rs693 (*APOB)*, rs7412 and rs429358 (*APOE),* and rs1801131 and rs1801133 (*MTHFR*) genetic polymorphisms, commonly associated to cardiovascular risk, in healthy men and women from Northern Chile and their possible relationships with plasma lipid levels. Previous studies conducted in our country have investigated some of these genetic polymorphisms in subjects from the Central and Southern regions of Chile. To our knowledge, this is the first study in healthy subjects from Northern Chile.

MATERIALS AND METHODS

Subjects

A group of 193 unrelated subjects (138 females and 55 males) volunteered to be enrolled in the current study. All participants were university students born in the region of Antofagasta (Northern Chile), having good health condition and no diagnosis of CVD, and were informed regarding the study design and goals. All the participants signed a written informed consent before enrolling in the study and blood sample extraction.

Individuals participating in the study required to answer a standardized questionnaire regarding basic health information; the collected data included age, gender, and body weight besides smoking, drinking, recreational drug consumption, exercise habits, and history of CVD based on self-reports. The type of recreational drugs used was not specified. They were all required to have normal fasting glucose levels (lower than 100 mg/dl), systolic blood pressure (SBP) lower than 120 mm Hg, diastolic blood pressure (DBP) lower than 80 mm Hg, and body mass index (BMI) lower than 25. Individuals taking anti-hypercholesterolemic, anti-hypertensive, or hypoglycemic medicines or those who declared to have any personal or family history of CVD were excluded from the study.

Anthropometric and clinical parameters as well as blood sample analyses were processed following standard procedures. Briefly, BMI was defined as the body weight in kilograms divided by the square of the body height in meters and expressed in units of kilograms per square meter. Blood samples were withdrawn and divided in two fractions; a plasma fraction was separated by centrifugation for biochemical analyses and stored at −20°C, and another fraction for analyses of total genomic DNA was conserved at 4°C until processing.

This study has been approved by the Ethics Committee of Universidad de Antofagasta (Chile). Every subject agreed to voluntarily take part in the study by signing a written informed consent. The investigation was performed accordingly to the ethical principles of the World Medical Association Inc (2008).

Biochemical Measurements

Venous blood samples were withdrawn from antecubital vein after overnight fasting. The plasma levels of glucose, total cholesterol, triglyceride, and HDL-C were quantitated using enzymatic-colorimetric kits commercially available from Human Diagnostics Worldwide, Germany. LDL-C value was calculated using Friedewald's formula if triglyceride levels did not exceed 400 mg/dl.

DNA Genotyping

Genomic DNA samples were extracted from peripheral blood leukocytes by using salting out method depicted by Salazar et al. (2001). Genotyping analysis for ApoB, Apo E, and MTHFR polymorphisms was performed by real-time PCR using 4351379 TaqMan SNP Genotyping Assays, Human, SM (ID assays were C_7615420_20 for rs693, C_904973_10 for rs7412, C_3084793_20 for rs429358, C_850486_20 for rs1801131, and C_1202883_20 for rs1801133; Applied Biosystems; Foster City, CA, United States). The PCR reactions were carried out accordingly to standard conditions provided by the manufacturer. In summary, assays contained 1 μl DNA (20 ng), 6.25 μl 2 × Taqman PCR master mix, 0.625 μl 20 Taqman genotyping assay, and 5.125 μl nuclease-free water for 13 μl total volume. Thermal cycling conditions for real-time system were initial denaturation at 95°C for 10 min, followed by 50 cycles of denaturation at 95°C for 15 s, annealing and extension at 60°C for 1 min, and a final extension at 60°C for 30 s.

For analytical purposes, the six APOE genotype groups (E2/E2, E3/E2, E3/E3, E4/E3, E4/E4, and E4/E2) were classified into three groups. The E3/E3 genotype occurs at high frequency in the population and it is considered the normal variant (Liehn et al., 2018). The E2/E2 and E3/E2 genotypes were combined and presented as E2 carriers. Since E4/E4 genotype individuals were absent in the study, the E4/E3 genotype is presented as E4 carriers (Wu et al., 2007). Previous studies have shown that the impact of the E2 allele on serum lipids is greater than that of the E4 allele (Wilson et al., 1994); therefore, individuals carrying E4/E2 genotype were excluded from the analysis.

Statistical Analyses

Genotype and allele frequencies were acquired by direct gene counting. The genotype distribution for Hardy–Weinberg equilibrium (HWE) was evaluated by χ^2 analysis. The Shapiro–Wilk test was performed to assess for normal distribution of continuous variables. Data were presented as mean ± SD. Comparison of continuous variables was accomplished using Student's *t*-test or ANOVA one-way test. The multiple comparisons were accomplished by the Bonferroni method. The Sigma Stat statistical software (Systat Software Inc., San Rafael, CA, United States) was used for statistical analyses. A p value < 0.05 was considered as statistically significant.

TABLE 1 | Anthropometric features and clinical parameters of the studied individuals.

Parameter	Total (193)	Male (55)	Female (138)
Age (years)	22.5 ± 8.8	22.2 ± 7.6	22.6 ± 9.2
BMI (kg/m^2)	23.0 ± 2.5	23.2 ± 2.2	22.4 ± 2.2
SBP (mm Hg)	116.2 ± 14.6	118.7 ± 12.9	115.4 ± 15.2
DBP (mm Hg)	70.1 ± 11.2	71.2 ± 13.3	69.9 ± 10.3
Glucose (mg/dl)	86.1 ± 28.0	84.7 ± 23.6	86.7 ± 28.9
Cholesterol (mg/dl)	139.7 ± 37.4	137.1 ± 31.7	141.1 ± 39.9
Triglycerides (mg/dl)	83.7 ± 55.6	77.7 ± 33.8	87.4 ± 65.2
HDL-cholesterol (mg/dl)	56.4 ± 23.8	49.2 ± 11.8	59.2 ± 26.5
LDL-cholesterol (mg/dl)	67.7 ± 39.0	72.4 ± 32.0	65.5 ± 42.2
Cigarette smoking (%)	22	20	24
Alcohol consumption (%)	69	74	66
Drugs consumption (%)	14	12	15
Physical activity (%)	55	67	50

Number of subjects in parentheses. Values are expressed as mean ± SD. Physical activity performed at least once a week. BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL-C, high-density lipoprotein cholesterol; LDL, low-density lipoprotein cholesterol.

RESULTS

Anthropometric features and clinical parameters of 193 subjects (28% male and 72% female) enlisted in the study are summarized in **Table 1**. The average age was 22.5 ± 8.8 years (22.2 ± 7.6 and 22.6 ± 9.2 years for male and female, respectively). Every volunteer had normal value for anthropometric BMI 23.0 ± 2.5 kg/m^2 (23.2 ± 2.2 kg/m^2 and 22.4 ± 2.2 kg/m^2 for male and female, respectively). Clinical parameters were normal, showing SBP of 116.2 ± 14.6 mm Hg (118.7 ± 12.9 mm Hg and 115.4 ± 15.2 mm Hg for male and female, respectively) and DBP of 70.1 ± 11.2 mm Hg (71.2 ± 13.3 mm Hg and 69.9 ± 10.3 mm Hg for male and female, respectively). Plasma glucose and lipid levels were all within the reference range (measured as mg/dl) (glucose 86.1 ± 28.0; total cholesterol 139.7 ± 37.4; triglycerides 83.7 ± 55.6; HDL-cholesterol 56.4 ± 23.8; LDL-cholesterol 67.7 ± 39.0). Analysis by gender is shown in **Table 1**. No significant differences were found among groups.

Genotype and allele frequencies for the studied polymorphisms are shown in **Table 2**. The relative allele frequency for *APOB* rs693 polymorphism was 0.42 for mutated allele (T). The relative allele frequencies for APOE rs7412/rs429358 polymorphisms were 0.06, 0.91, and 0.03 for E4, E3, and E2 alleles, respectively. The relative allele frequencies for *MTHFR* rs1801131 and rs1801133 polymorphisms were 0.28 and 0.56 for mutated alleles (C) and (T), respectively. Overall, *MTHFR* rs1801133 was the only genetic variant to show the distribution pattern predicted by HWE ($p = 0.6$), whereas the other three genotypes differed from the Hardy–Weinberg prediction ($p < 0.05$).

Based on previous reports (Utermann, 1987; Lagos et al., 2015; Han et al., 2016), we studied the potential association of ApoB rs693, ApoE rs7412, Apo E rs429358, MTHFR rs1801131, and MTHFR rs1801133 allelic variants with plasma lipid profile. No changes in lipid levels were observed for the entire group of subjects studied. The results are shown in **Table 3**. However, when the analysis was performed by gender, the female group carrying ApoB C/T heterozygous genotype exhibited significant lower levels of total cholesterol ($p < 0.05$) compared with those female carrying C/C genotype. In addition, we observed higher levels of HDL-C in females carrying C/C homozygous genotype compared with males carrying identical genotype ($p < 0.01$). On the other hand, analysis of ApoE rs7412 and ApoE rs429358 genetic variants exhibited significant lower levels of total cholesterol along with LDL-C in females carrying E4 compared with E3 allele ($p < 0.05$). Higher levels of triglycerides were found in males carrying E4 compared with E3 allele ($p < 0.01$); on the contrary, higher level of triglycerides, although not significant, were observed in females carrying E2 allele. In addition, lower levels of triglycerides were found in females carrying E4 allele compared with males carrying identical genotype. Analysis of *MTHFR* rs1801131 genetic variant showed significant lower value of triglycerides in females carrying A/C heterozygous compared with A/A homozygous genotype ($p < 0.05$). On the other hand, females carrying MTHFR rs1801133 T/T homozygous genotype showed significant lower values of LDL-C compared with males carrying identical genotype ($p < 0.05$). A comparative analysis of allele frequencies of the four genetic variants found in the population worldwide along with people living in different regions of Chile is shown in **Table 4**.

DISCUSSION

This is the first study highlighting the frequency of ApoB, ApoE, and MTHFR genetic variants and their relationship with plasma lipid levels in young healthy individuals from Antofagasta city (Northern Chile). In the studied group, the relative frequency for mutated allele (T) of rs693 (ApoB) was 0.42, similar to that previously described in Chile (Eyheramendy et al., 2015) and not very different of that found in individuals from Brazil (Nakazone et al., 2009; Tamburus et al., 2018). However, this rate was different from the allele frequency found in Asian individuals as shown in **Table 4** and depicted by Niu et al. (2017). For ApoE alleles, the most frequent genotype was E3/3 (84%), followed by E4/3 (10.2%), and E3/2 (3.9%). These data were similar to those described by Roco et al. (2015) in healthy individuals from the central region of Chile; however, we did not detect any subject with genotype E4/4. Similar frequencies were observed in Japanese, Chinese, and Mexican-American populations (Eichner et al., 2002). On the other hand, our data were different from those previously described in Caucasians from Eastern Europe and Amerindian hypercholesterolemic patients from Southern Chile (Eichner et al., 2002; Lagos et al., 2015). The analysis of MTHFR rs1801131 minor allele (C) frequency was 0.28, comparable with those previously described by Báez et al. (2010) and Eyheramendy et al. (2015) in Chile and also similar to those found in American (Mahesh et al., 2019), European, and Asian individuals as shown in **Table 4**. In addition, MTHFR rs1801133 minor allele (T) frequency was 0.56, different to the one depicted in healthy individuals from the central region of

TABLE 2 | Genotype distribution and relative allele frequency for ApoB, ApoE, and MTHFR polymorphisms.

		Genotype frequency (%)			P-value	Allele frequency	
		CC	CT	TT		C	T
ApoB (rs693) *n* = 189	Total	70 (37)	78 (41)	41 (22)	0.03	0.58	0.42
	Male	18 (37)	24 (49)	7 (14)			
	Female	52 (37)	54 (39)	34 (24)			

		Genotype frequency (%)					P-value	Allele frequency		
		E4/3	E4/2	E3/3	E3/2	E2/2		E4	E3	E2
ApoE (rs7412, rs429358) *n* = 127	Total	13 (10.2)	1 (0.8)	107 (84)	5 (3.9)	1 (0.8)	<0.05	0.06	0.91	0.03
	Male	6 (17.6)	1 (2.9)	27 (79.5)	0	0				
	Female	7 (7.5)	0	80 (86)	5 (5.4)	1 (1.1)				

		Genotype frequency (%)			P-value	Allele frequency	
		AA	AC	CC		A	C
MTHFR (rs1801131) *n* = 186	Total	106 (57)	56 (30)	24 (13)	0.001	0.72	0.28
	Male	29 (60)	12 (25)	7 (15)			
	Female	77 (56)	44 (32)	17 (12)			

		Genotype frequency (%)			P-value	Allele frequency	
		CC	CT	TT		C	T
MTHFR (rs1801133) *n* = 193	Total	39 (20)	91 (47)	63 (33)	0.6	0.44	0.56
	Male	12 (22)	28 (51)	15 (27)			
	Female	27 (20)	63 (46)	48 (35)			

χ^2 analysis was used to test Hardy–Weinberg equilibrium. Genotype E4/E4 was not found and genotype E4/E2 was excluded from the study; therefore, the E4/E3 is presented as E4 carriers. The E3/E2 and E2/E2 genotypes were combined and presented as E2 carriers.

Chile (Nitsche et al., 2003; Roco et al., 2015). Overall, those frequencies are very different from those described by Liew and Gupta (2015) in other regions of the world as shown in **Table 4**. The differences observed in ApoE and MTHFR rs1801133 allelic frequencies in individuals from the northern region of Chile compared with other regions might be explained based on the genetic background of Chilean people as a consequence of some miscegenation events where larger European and Native-American and smaller African ancestry contributed to admixing, showing that African ancestry contribution is higher in the northern region (3.9%) and diminished in southern regions (1.6%), while European ancestry contribution is higher in the central region and Native-American ancestry prevails in the southern region (Fuentes et al., 2014; Ruiz-Linares et al., 2014; Eyheramendy et al., 2015). Altogether, these observations indicate that the genetic background of Chilean subjects is unevenly spread throughout the country, and this could explain why the studied allelic variants exhibited a geographic dependent pattern as was previously demonstrated for *LDLR* and *PCSK9* polymorphisms (Rojas et al., 2019).

Regarding lipid profiles, no association of any of the genetic variants with total cholesterol, triglycerides, HDL-C, and LDL-C profile was found among the complete studied group of individuals, similar to those previously described (Nakazone et al., 2009; Roco et al., 2015; Mahesh et al., 2019); however, some associations appeared when the analyses were performed for each gender separately. For Apo B genetic variant, C/C genotype was associated with higher total cholesterol along with HDL-C levels in females and therefore may be a risk factor that predisposes to CVD. A similar association was described by Nakazone et al. (2009) in a group of control individuals that consisted of 64% females; however, other authors depicted a similar association in males but not in females (Niu et al., 2017). To this respect, association of ApoB rs693 and lipid levels is controversial since this genetic variant is a silent mutation that cannot directly affect lipid metabolism; however, association with higher levels of total cholesterol, triglyceride, and LDL-C and lower levels of HDL-C has been reported, suggesting that gender, ethnicity, and health status might modulate its association to lipid levels (Peacock et al., 1992; Benn et al., 2008; Niu et al., 2017; Tamburus et al., 2018), but some of those studies presented association of high lipid levels with males, which are opposite to our findings.

In addition, for ApoE genetic variants, E3 allele carriers showed a total cholesterol level similar to that previously described by Roco et al. (2015); however, E4 allele is associated with lower levels of total cholesterol, triglycerides, and LDL-C exclusively in women, whereas lipid level profiles in E2 allele carriers were not different from those observed in E3 allele carriers. These data were not consistent with those found in the

TABLE 3 | Lipid profile for ApoB, ApoE, and MTHFR polymorphisms in healthy individuals from Northern Chile.

		TC			TG			HDL-C			LDL-C		
		C/C	**C/T**	**T/T**	**C/C**	**C/T**	**T/T**	**C/C**	**C/T**	**T/T**	**C/C**	**C/T**	**T/T**
ApoB (rs693)	Total	147.8 ± 39.5	140.5 ± 41.4	138.5 ± 31.4	78.9 ± 30.7	76.2 ± 38.9	80.9 ± 38.4	53.7 ± 15.5	55.2 ± 18.0	52.6 ± 14.6	77.4 ± 40.2	67.2 ± 34.4	69.7 ± 31.6
	Male	139.2 ± 27.9	142.5 ± 45.6	130.3 ± 25.5	87.9 ± 37.1	72.7 ± 37.4	79.7 ± 35.9	45.5 ± 11.7	52.8 ± 11.8	47.1 ± 11.9	74.9 ± 32.8	70.0 ± 28.9	67.3 ± 22.2
	Female	151.9 ± 43.9	135.6 ± 36.9*	140.9 ± 32.9	78.4 ± 41.4	75.3 ± 38.5	77.9 ± 36.8	57.2 ± 14.9‡‡	56.7 ± 19.9	53.9 ± 15.2	73.0 ± 31.3	66.8 ± 38.5	71.5 ± 33.9
		E4	**E3**	**E2**	**E4**	**E3**	**E2**	**E4**	**E3**	**E2**	**E4**	**E3**	**E2**
ApoE (rs7412, rs429358)	Total	124.2 ± 35.2	134.8 ± 29.4	137.8 ± 18.4	86.7 ± 46.4	74.9 ± 35.5	82.8 ± 30.8	53.2 ± 13.6	58.6 ± 42.7	44.7 ± 13.9	61.5 ± 33.4	68.4 ± 33.9	71.9 ± 18.4
	Male	142.8 ± 31.1	134.5 ± 33.3	N.D.	114.1 ± 14.2**	68.2 ± 30.1	N.D.	45.8 ± 8.4	48.2 ± 13.6	N.D.	74.2 ± 29.9	72.6 ± 32.2	N.D.
	Female	107.6 ± 28.3*	135.1 ± 28.9	137.8 ± 18.4	70.4 ± 55.9‡	74.6 ± 29.5	82.8 ± 30.8	60.7 ± 15.5	57.0 ± 17.6	44.6 ± 13.9	48.9 ± 34.4*	73.1 ± 64.8	71.9 ± 18.4
		A/A	**A/C**	**C/C**	**A/A**	**A/C**	**C/C**	**A/A**	**A/C**	**C/C**	**A/A**	**A/C**	**C/C**
MTHFR (rs1801131)	Total	136.6 ± 37.9	141.9 ± 40.3	138.8 ± 28.7	88.5 ± 62.4	73.3 ± 36.9	86.1 ± 28.9	65.8 ± 38.7	66.9 ± 38.2	55.9 ± 22.3	67.6 ± 40.8	72.2 ± 38.3	62.3 ± 27.6
	Male	130.9 ± 33.2	143.6 ± 33.5	136.2 ± 29.7	76.7 ± 39.5	83.5 ± 27.6	96.7 ± 18.6	67.9 ± 44.8	63.4 ± 31.2	48.9 ± 15.1	71.8 ± 39.3	77.1 ± 30.3	70.3 ± 31.1
	Female	138.7 ± 39.6	141.4 ± 42.4	140.0 ± 29.2	92.8 ± 68.7	70.3 ± 39.0*	81.2 ± 32.1	66.8 ± 35.7	67.9 ± 40.3	62.5 ± 19.3	67.9 ± 40.5	70.9 ± 40.5	63.3 ± 22.1
		C/C	**C/T**	**T/T**	**C/C**	**C/T**	**T/T**	**C/C**	**C/T**	**T/T**	**C/C**	**C/T**	**T/T**
MTHFR (rs1801133)	Total	146.7 ± 30.7	136.7 ± 42.3	138.2 ± 34.0	92.6 ± 70.9	77.4 ± 49.6	88.1 ± 45.7	65.2 ± 31.3	63.1 ± 35.9	62.3 ± 36.2	72.2 ± 31.7	67.4 ± 46.1	65.2 ± 39.5
	Male	134.6 ± 28.4	127.7 ± 31.6	146.9 ± 39.8	74.7 ± 42.4	77.9 ± 32.7	96.5 ± 30.3	57.7 ± 30.9	57.4 ± 40.4	63.8 ± 39.9	74.4 ± 31.6	61.3 ± 39.8	85.3 ± 36.9
	Female	151.4 ± 31.1	140.9 ± 46.0	136.4 ± 32.2	99.7 ± 73.8	77.8 ± 55.7	85.4 ± 49.5	69.2 ± 31.2	67.5 ± 35.5	61.9 ± 35.5	71.9 ± 32.7	70.2 ± 48.5	59.5 ± 38.6‡

*Value are expressed in mg/dl and represented as mean ± SD. The E2/E2 and E2/E3 genotypes were combined and presented as E2 carriers. Since E4/E4 individuals were absent, the E3/E4 genotype was presented as E4 carriers. The E2/E4 genotype was excluded from the analysis. Data were analyzed by one-way ANOVA, followed by a post hoc Bonferroni or Student t-test. *$p < 0.05$ when compared with C/C or A/A alleles; *$p < 0.05$ or **$p < 0.01$ when compared with E3 allele; ‡$p < 0.05$ or ‡‡$p < 0.01$ when compared with the same genotype by gender. N.D.: not determined.*

TABLE 4 | Allele frequencies for *APOB*, *APOE*, and *MTHFR* polymorphisms in population worldwide and Chilean subjects from different regions.

Population	Apo B rs693			Apo E rs7412/rs429358				MTHFR rs1301131			MTHFR rs1801133		
	Sample size	C	T	Sample size	E4	E3	E2	Sample size	A	C	Sample size	C	T
African	11,472	0.77	0.23					11,466	0.83	0.17	11,720	0.88	0.12
Asian	842	0.94	0.06					884	0.75	0.25	3,984	0.66	0.34
European	207,254	0.51	0.49					211,780	0.69	0.31	322,326	0.65	0.35
Latin American	9,062	0.63	0.37					5,020	0.81	0.19	7,238	0.55	0.45
*Chileans	626	0.62	0.38					626	0.77	0.23			
Northern Chile (Antofagasta)	189	0.58	0.42	127	0.06	0.91	0.03	186	0.72	0.28	193	0.44	0.56
‡Central Chile				146	0.09	0.88	0.03				146	0.52	0.48
‡‡Central Chile								105	0.79	0.21			
‡‡‡Central Chile											184	0.60	0.40
+Southern Chile				116	0.18	0.80	0.02						

*All data were retrieved from the ALFA project. Release version: 20201027095038, except those indicated. *Eyheramendy et al. (2015). ‡Roco et al. (2015). ‡‡Báez et al. (2010). ‡‡‡Nitsche et al. (2003). +Lagos et al. (2015).*

general population, suggesting an association of E4 allele with high total cholesterol and LDL-C concentrations (Utermann, 1987) and therefore with high risk of CVD (Eichner et al., 2002; Larifla et al., 2017) or no association at all (Roco et al., 2015). On the other hand, gender-specific studies exhibited high triglyceride levels in men carrying both E2 and E4 alleles (Wilson et al., 1994) similar to our observations for E4 allele, although, as we previously stated, the limitation of a small sample size did not produce any males with E2 allele; therefore, comparisons were done with E4 allele exclusively. On the contrary, Kolovou et al. (2009) depicted increased total cholesterol in both women and men having E4 allele, which is opposite to our findings.

Finally, MTHFR genetic variants showed a gender-specific association with triglyceride (rs1801131) and LDL-C (rs1801133). MTHFR rs1801133 polymorphism has been thoughtfully studied because of its health implications. In general, (T) allele carriers showed higher total cholesterol and LDL-C in a study including Caucasian, Asian, African, and other ethnicities; however, no association with lipid profile was described in Mexican and American individuals carrying (T) allele (Leal-Ugarte et al., 2019; Mahesh et al., 2019), although authors proposed statins' effect on lowering lipid levels, and therefore no association. Analyses of different subgroups showed significant associations with higher levels of total cholesterol along with LDL-C detected exclusively in females (Luo et al., 2018); however, once again these data are not consistent with our findings showing lower levels of LDL-C in (T) allele carriers. On the other hand, MTHFR rs1801131 showed a significant association with lower levels of HDL-C (Mahesh et al., 2019) and higher levels of total cholesterol and LDL-C in individuals having C/C genotype; in this case, those results were dependent of hypermethylated promoter (Santana Bezerra et al., 2019). However, once again our data were opposite, showing only lower triglyceride levels in females carrying (C) allele.

Overall, these apparent differences in the association of genetic polymorphisms and lipid profile with previous reports may be attributable to the peculiar genetic background admixing existing in the studied individuals. Moreover, the underlying results showed that healthy age-matched subjects from central and southern regions of Chile exhibited lower level of lipids (Arteaga et al., 2010; Caamaño Navarrete et al., 2015), probably due to seasonal variations that produce higher values in winter and lower values in summer (Kreindl et al., 2014; Zhou et al., 2016). Since the northern region of Chile has a hot and dry desertic weather with temperature range 3 to 5°C higher than other regions and no seasons (McKay et al., 2003), it could account for lower lipid values found in our study. Considering these geographical differences on lipid values, the effects of the studied genetic variants on lipid-lowering therapy in subjects with hypercholesterolemia need to be investigated. Several authors have demonstrated the effect of ApoB (Guzmán et al., 2000; Ye et al., 2003), ApoE (Ye et al., 2003; Tavintharan et al., 2007; Lagos et al., 2015), and MTHFR (Maitland-van der Zee et al., 2008; Jiang et al., 2013) polymorphisms on response to statins (HMG-CoA reductase inhibitors), specifically on the pharmacodynamics of these drugs. For example, Ye et al. (2003) observed that the relative frequency of ApoB rs693 genetic variant is high in individuals with hyperlipidemia, in whom the cholesterol-lowering efficacy is diminished after treatment of simvastatin. In relation to APOE variants, Lagos et al. (2015) showed that E4/3 genotype carriers exhibited lower cholesterol reduction compared with E3/3 genotype (LDL-C: −18% vs. −29%, $p<0.001$), when the hypecholesterolemic subjects were treated with atorvastatin. Finally, in relation to MTHFR gene polymorphisms, Maitland-van der Zee et al. (2008) observed that subjects carrying C/C genotype for MTHFR rs1801133 genetic variant exhibited a significantly protective effect against CHD [0.71 (95% CI 0.58–0.87)] when the patients were treated with pravastatin. Similarly, Jiang et al. (2013) showed that MTHFR rs1801133 genetic variant contributes to the effects of simvastatin in Chinese subjects with primary hyperlipidemia.

In summary, our data showed associations between ApoB (rs693), ApoE (rs7412 and rs429358), and MTHFR (1801131 and 1801133) single nucleotide polymorphisms with lipid levels in healthy individuals from Northern Chile, especially in women. However, to confirm the influence of any of the studied genetic variants on lipid profile, enrolling a greater number of individuals in a case–control study is required. Finally, we propose that understanding new candidate genes that predispose to CVD might allow to elaborate risk score tables to improve therapeutic and preventive strategies. The clinical utility of these genetic polymorphisms in the prediction of cardiovascular risk or the influence on lipid-lowering therapy needs to be investigated in a future study including a larger number of subjects, considering a similar distribution of males and females. Moreover, future directions for this study need to include functional studies.

DATA AVAILABILITY STATEMENT

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité de Ética en Investigación Científica, Universidad de Antofagasta. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JE-V conceived and designed the study. HR, CR, PP, and XU performed the experiments. HR, PP, CR, and ASG analyzed the data. LAS participated in the design of the study, contributed reagents/materials, and analysis tools. LAS and AMK reviewed and edited the manuscript. ASG and JE-V wrote the paper. All authors have read the article and agreed with content.

FUNDING

This work was supported by Grant Number 5311—Programa Semilleros de Investigación, DGI, VRIIP, Universidad de Antofagasta, Chile, and also by the Regional Government of Antofagasta through the Innovation Fund for Competitiveness FIC-R 2017 (BIP Code: 30488811-0).

ACKNOWLEDGMENTS

We are grateful to all participants who donated their blood to perform this study.

REFERENCES

Abu-Hassan, D. W., Alhouri, A. N., Altork, N. A., Shkoukani, Z. W., Altamimi, T. S., Alqaisi, O. M., et al. (2019). MTHFR gene polymorphisms in hypothyroidism and hyperthyroidism among Jordanian females. *Arch. Endocrinol. Metab.* 63, 280–287.

Amrani-Midoun, A., Kiando, S. R., Treard, C., Jeunemaitre, X., and Bouatia-Naji, N. (2016). The relationship between MTHFR C677T gene polymorphism and essential hypertension in a sample of an Algerian population of Oran city. *Int. J. Cardiol.* 225, 408–411. doi: 10.1016/j.ijcard.2016.10.027

Arteaga, A., Bustos, P., Soto, R., Velasco, N., and Amigo, H. (2010). Physical activity and cardiovascular risk factors among Chilean young men and women. *Rev. Med. Chile* 138, 1209–1216.

Báez, S., Tsuchiya, Y., Calvo, A., Pruyas, M., Nakamura, K., Kiyohara, C., et al. (2010). Genetic variants involved in gallstone formation and capsaicin metabolism, and the risk of gallbladder cancer in Chilean women. *World J. Gastroenterol.* 16, 372–378. doi: 10.3748/wjg.v16.i3.372

Bailey, L. B., and Gregory, J. F. III. (1999). Polymorphisms of Methylenetetrahydrofolate Reductase and Other Enzymes: metabolic significance, risks and impact on folate requirement. *J. Nutr.* 129, 919–922. doi: 10.1093/jn/129.5.919

Benn, M. (2009). Apolipoprotein B levels, APO B alleles, and risk of ischemic cardiovascular disease in the general population, a review. *Atherosclerosis* 206, 17–30. doi: 10.1016/j.atherosclerosis.2009.01.004

Benn, M., Stene, M. C. A., Nordestgaard, B. G., Jensen, G. B., Steffensen, R., and Tybjaerg-Hansen, A. (2008). Common and Rare Alleles in Apolipoprotein B Contribute to Plasma Levels of Low-Density Lipoprotein Cholesterol in the General Population. *J. Clin. Endocrinol. Metab.* 93, 1038–1045. doi: 10.1210/jc.2007-1365

Caamaño Navarrete, F., Alarcón Hormazábal, M., and Delgado Floody, P. (2015). Niveles de obesidad, perfil metabólico, consumo de tabaco y presión arterial en jóvenes sedentarios. *Nutr. Hosp.* 32, 2000–2006.

Defesche, J. C., Gidding, S. S., Harada-Shiba, M., Hegele, R. A., and Santos, R. D. (2017). Familial hypercholesterolaemia. *Nat. Rev. Dis. Primers* 3:17093. doi: 10.1038/nrdp.2017.93

DEIS (2021). *Departamento de Estadísticas e Información de Salud. Ministerio de Salud. Informe Semanal de Defunciones por COVID-19 N 61*. Available online at: https://www.minsal.cl/wp-content/uploads/2021/08/Informe-Semanal-del-ago-12-2021.pdf (accessed August 12, 2021)

Eichner, J. E., Dunn, S. T., Perveau, G., Thompson, D. M., Stewart, K. E., and Stroehla, B. C. (2002). Apolipoprotein E polymorphism and cardiovascular disease: a HuGE review. *Am. J. Epidemiol.* 155, 487–495. doi: 10.1093/aje/155.6.487

ENS (2003). *Chilean Ministry of Health. National Population Health Survey*. Santiago: Ministerio de Salud de Chile.

Eyheramendy, S., Martinez, F. I., Manevy, F., Vial, C., and Repetto, G. M. (2015). Genetic structure characterization of Chileans reflects historical immigration patterns. *Nat. Commun.* 6:6472. doi: 10.1038/ncomms7472

Fuentes, M., Pulgar, I., Gallo, C., Bortolini, M. C., Canizales-Quinteros, S., Bedoya, G., et al. (2014). Gene geography of Chile: regional distribution of American, European and African genetic contributions. *Rev. Med. Chil.* 142, 281–289.

Ghogomu, S. M., Ngolle, N. E., Mouliom, R. N., and Asa, B. F. (2016). Association between the MTHFR C677T gene polymorphism and essential hypertension in South West Cameroon. *Genet. Mol. Res.* 15:28. doi: 10.4238/gmr.15017462

Guzmán, E. C., Hirata, M. H., Quintão, E. C., and Hirata, R. D. (2000). Association of the apolipoprotein B gene polymorphisms with cholesterol levels and response to fluvastatin in Brazilian individuals with high risk for coronary heart disease. *Clin. Chem. Lab. Med.* 38, 731–736. doi: 10.1515/CCLM.2000.103

Han, S. Y., Xu, Y. H., Gao, M. H., Wang, Y. S., Wang, J., Liu, Y. Y., et al. (2016). Serum Apolipoprotein E concentration and polymorphism influence serum lipid levels in Chinese Shandong Han population. *Medicine* 95:e5639. doi: 10.1097/MD.0000000000005639

Homsma, S. J. M., Huijgen, R., Middeldorp, S., Sijbrands, E. J. G., and Kastelein, J. J. P. (2008). Molecular screening for familial hypercholesterolaemia: consequences for life and disability insurance. *Eur. J. Hum. Genet.* 16, 14–17. doi: 10.1038/sj.ejhg.5201940

Innerarity, T. L., Mahley, R. W., Weisgraber, K. H., Bersot, T. P., Krauss, R. M., Vega, G. L., et al. (1990). Familial defective apolipoprotein B-100: a single mutation that causes hypercholesterolemia. *J. Lipid Res.* 31, 1337–1349. doi: 10.1016/s0022-2275(20)42605-7

Jiang, S., Chen, Q., Venners, S. A., Zhong, G., Hsu, Y. H., Xing, H., et al. (2013). Effect of simvastatin on plasma homocysteine levels and its modification by MTHFR C677T polymorphism in Chinese patients with primary hyperlipidemia. *Cardiovasc. Ther.* 31, e27–e33. doi: 10.1111/1755-5922.12002

Khan, T. A., Shah, T., Prieto, D., Zhang, W., Price, P., Fowkes, G. R., et al. (2013). Apolipoprotein E genotype, cardiovascular biomarkers and risk of stroke: systematic review and meta-analysis of 14,015 stroke cases and pooled analysis of primary biomarker data from up to 60,883 individuals. *Int. J. Epidemiol.* 42, 475–492. doi: 10.1093/ije/dyt034

Kolovou, G., Damaskos, D., Anagnostopoulou, K., and Cokkinos, D. V. (2009). Apolipoprotein E Gene Polymorphism and Gender. *Ann. Clin. Lab. Sci.* 39, 120–133.

Kreindl, C., Olivares, M., Brito, A., Araya, M., and Pizarro, F. (2014). Variación estacional del perfil lipídico en adultos aparentemente sanos de Santiago, Chile. *Arch. Latinoam. Nutr.* 64, 145–152.

Lagos, J., Zambrano, T., Rosales, A., and Salazar, L. A. (2015). APOE polymorphisms contribute to reduced atorvastatin response in Chilean Amerindian subjects. *Int. J. Mol. Sci.* 16, 7890–7899. doi: 10.3390/ijms16047890

Larifla, L., Armand, C., Bangou, J., Blanchet-Deverly, A., Numeric, P., Fonteau, C., et al. (2017). Association of APOE gene polymorphism with lipid profile and coronary artery disease in AfroCaribbeans. *PLoS One* 12:e0181620. doi: 10.1371/journal.pone.0181620

Leal-Ugarte, E., Peralta-Leal, V., Meza-Espinoza, J. P., Durán-González, J., Macías-Gómez, N., Bocanegra-Alonso, A., et al. (2019). Association of the MTHFR 677C>T polymorphism with obesity and biochemical variables in a young population of Mexico. *J. Med. Biochem.* 38, 461–467. doi: 10.2478/jomb-2018-0046

Liehn, E. A., Ponomariov, V., Diaconu, R., Streata, I., Ioana, M., Crespo-Avilan, G. E., et al. (2018). Apolipoprotein E in Cardiovascular Diseases: novel Aspects of an Old-fashioned Enigma. *Arch. Med. Res.* 49, 522–529. doi: 10.1016/j.arcmed.2018.08.008

Liew, S. C., and Gupta, E. D. (2015). Methylenetetrahydrofolate reductase (MTHFR) C677T polymorphism: epidemiology, metabolism and the associated diseases. *Eur. J. Hum. Genet.* 58, 1–10. doi: 10.1016/j.ejmg.2014.10.004

Liggett, S. B., Kelly, R. J., Parekh, R. R., Matkovich, S. J., Benner, B. J., Hahn, H. S., et al. (2007). A functional polymorphism of the Gaq (GNAQ) gene is associated with accelerated mortality in African-American heart failure. *Hum. Mol. Genet.* 16, 2740–2750. doi: 10.1093/hmg/ddm229

Liu, P. F., Ding, B., Zhang, J. Y., Mei, X. F., Li, F., Wu, P., et al. (2020). Association Between MTHFR C677T Polymorphism and Congenital Heart Disease. A PRISMA-Compliant Meta-Analysis. *Int. Heart J.* 61, 553–561. doi: 10.1536/ihj.19-389

Luo, Z., Lu, Z., Muhammad, I., Chen, Y., Chen, Q., Zhang, J., et al. (2018). Associations of the MTHFR rs1801133 polymorphism with coronary artery disease and lipid levels: a systematic review and updated meta-analysis. *Lipids Health Dis.* 17:191. doi: 10.1186/s12944-018-0837-y

Mahesh, M., Cheng, G., and Khalighi, K. (2019). Association of Methylenetetrahydrofolate Reductase *(MTHFR) A1298C* Polymorphism with Lower High-Density Lipoprotein Cholesterol Level. *Ann. Clin. Lab. Sci.* 49, 232–236.

Maitland-van der Zee, A. H., Lynch, A., Boerwinkle, E., Arnett, D. K., Davis, B. R., Leiendecker-Foster, C., et al. (2008). Interactions between the single nucleotide polymorphisms in the homocysteine pathway (MTHFR 677C>T, MTHFR 1298 A>C, and CBSins) and the efficacy of HMG-CoA reductase inhibitors in preventing cardiovascular disease in high-risk patients of hypertension: the GenHAT study. *Pharmacogenet. Genomics* 18, 651–656. doi: 10.1097/FPC.0b013e3282fe1759

McKay, C. P., Friedmann, E. I., Gómez-Silva, B., Cáceres-Villanueva, L., Andersen, D. T., and Landheim, R. (2003). Temperature and moisture conditions for life in the extreme arid region of the Atacama desert: four years of observations including the El Niño of 1997-1998. *Astrobiology* 3, 393–406.

Nakazone, M. A., De Marchi, M. A., Pinhel, M. A. S., Barros, C. F. D. C., Júlio, M. A. F., Pinheiro, A., et al. (2009). Effects of APOE, APOB and LDLR variants on serum lipids and lack of association with xanthelasma in individuals from Southeastern Brazil. *Genet. Mol. Biol.* 32, 227–233. doi: 10.1590/s1415-47572009005000028

Nazzal, C., and Alonso, F. T. (2013). Younger Women Have a Higher Risk of In-Hospital Mortality Due to Acute Myocardial Infarction in Chile. *Rev. Esp. Cardiol.* 66, 104–109. doi: 10.1016/j.rec.2012.07.007

Nitsche, F., Alliende, M. A., Santos, J. L., Pérez, F., Santa María, L., Hertrampf, E., et al. (2003). Frequency of C677T polymorphism of 5, 10-methylenetetrahydrofolate reductase (MTHFR) in Chilean mothers of spina bifida cases and controls. *Rev. Med. Chil.* 131, 1399–1404.

Niu, C., Luo, Z., Yu, L., Yang, Y., Chen, Y., Luo, X., et al. (2017). Association of the ApoB rs693 and rs17240441 polymorphisms with plasma ApoB and lipid levels: a meta-analysis. *Lipids Health Dis.* 16:166. doi: 10.1186/s12944-017-0558-7

Peacock, R., Dunning, A., Hamsten, A., Tornvall, P., Humphries, S., and Talmud, P. (1992). Apolipoprotein B gene polymorphisms, lipoproteins and coronary atherosclerosis: a study of young myocardial infarction survivors and healthy population-based individuals. *Atherosclerosis* 92, 151–164. doi: 10.1016/0021-9150(92)90274-k

Roco, A., Quiñones, L. A., Sepúlveda, P., Donoso, H., Lapostol, C., Alarcón, R., et al. (2015). Prevalence of seven cardiovascular-related genetic polymorphisms in a Chilean mestizo healthy population. *Acta Cardiol.* 70, 528–535. doi: 10.1080/AC.70.5.3110513

Rojas, C., Ramírez, H., Salazar, L. A., Kalergis, A. M., Gálvez, A. S., and Escobar-Vera, J. (2019). Characterization of LDLR rs5925 and PCSK9 rs505151 genetic variants frequencies in healthy subjects from Northern Chile: influence on plasma lipid levels. *J. Clin. Lab. Anal.* 33:e23001. doi: 10.1002/jcla.23001

Ruiz-Linares, A., Adhikari, K., Acuña-Alonzo, V., Quinto-Sanchez, M., Jaramillo, C., Arias, W., et al. (2014). Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. *PLoS Genet.* 10:e1004572. doi: 10.1371/journal.pgen.1004572

Salazar, L. A., Melo, C. E., Cavalli, S. A., Hinuy, H. M., Hirata, M. H., and Hirata, R. D. C. (2001). Micrométodo para extração de DNA genômico útil no diagnóstico molecular da hipercolesterolemia familial. *Rev. Bras. Anal. Clin.* 33, 111–116.

Santana Bezerra, H., Severo de Assis, C., Dos Santos Nunes, M. K., Wanderley de Queiroga Evangelista, I., Modesto Filho, J., Neta Alves Pegado Gomes, C., et al. (2019). The MTHFR promoter hypermethylation pattern associated with the A1298C polymorphism influences lipid parameters and glycemic control in diabetic patients. *Diabetol. Metab. Syndr.* 11:4. doi: 10.1186/s13098-019-0399-9

Smolková, B., Bonassi, S., Buociková, V., Dušinská, M., Horská, A., Kuba, D., et al. (2015). Genetic determinants of quantitative traits associated with cardiovascular disease risk. *Mutat. Res.* 778, 18–25. doi: 10.1016/j.mrfmmm.2015.05.005

Tamburus, N. Y., Verlengia, R., Kunz, V. C., César, M. C., and Silva, E. (2018). Apolipoprotein B and angiotensin-converting enzyme polymorphisms and aerobic interval training: randomized controlled trial in coronary artery disease patients. *Braz. J. Med. Biol. Res.* 51:e6944. doi: 10.1590/1414-431X20186944

Tavintharan, S., Lim, S. C., Chan, Y. H., and Sum, C. F. (2007). Apolipoprotein E genotype affects the response to lipid-lowering therapy in Chinese patients with type 2 diabetes mellitus. *Diabetes Obes. Metab.* 9, 81–86. doi: 10.1111/j.1463-1326.2006.00577.x

Tran, P., Leclerc, D., Chan, M., Pai, A., Hiou-Tim, F., Wu, Q., et al. (2002). Multiple transcription start sites and alternative splicing in the methylenetetrahydrofolate reductase gene result in two enzyme isoforms. *Mamm. Genome* 13, 483–492. doi: 10.1007/s00335-002-2167-6

Utermann, G. (1987). Apolipoprotein E polymorphism in health and disease. *Am. Heart J.* 113, 433–440. doi: 10.1016/0002-8703(87)90610-7

Varleta, P., Acevedo, M., Casas-Cordero, C., Berríos, A., and Navarrete, C. (2020). Low Cardiovascular Disease Awareness in Chilean Women: insights from the ESCI Project. *Glob. Heart* 15, 55. doi: 10.5334/gh.534

Wilson, P. W., Myers, R. H., Larson, M. G., Ordovas, J. M., Wolf, P. A., and Schaefer, E. J. (1994). Apolipoprotein E Alleles, Dyslipidemia and Coronary Heart Disease. The Framingham Offspring Study. *JAMA* 272, 1666–1671.

World Medical Association Inc (2008). *DECLARATION OF HELSINKI. Ethical Principles for Medical Research Involving Human Subjects*. France: World Medical Association Inc.

Wu, K., Bowman, R., Welch, A. A., Luben, R. N., Wareham, N., Khaw, K. T., et al. (2007). Apolipoprotein E polymorphisms, dietary fat and fibre and serum lipids: the EPIC Norfolk study. *Eur. Heart J.* 28, 2930–2936. doi: 10.1093/eurheartj/ehm482

Ye, P., Shang, Y., and Ding, X. (2003). The influence of apolipoprotein B and E gene polymorphisms on the response to simvastatin therapy in patients with hyperlipidemia. *Chin. Med. Sci. J.* 18, 9–13.

Zhou, X., Lin, H., Zhang, S., Ren, J., Wang, Z., Zhang, Y., et al. (2016). Effects of climatic factors on plasma lipid levels: a 5-year longitudinal study in a large Chinese population. *J. Clin. Lipidol.* 10, 1119–1128. doi: 10.1016/j.jacl.2016.06.009

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 25 November 2021
doi: 10.3389/fphar.2021.661443

Preliminary Pharmacogenomic-Based Predictive Models of Tamoxifen Response in Hormone-dependent Chilean Breast Cancer Patients

Carla Miranda[1], Macarena Galleguillos[1], Roberto Torres[2], Karla Tardón[2], Dante D. Cáceres[3], Kuen Lee[1,4], María A. Redal[5,6], Nelson M. Varela[1,6] and Luis A. Quiñones[1,6]**

[1]Laboratory of Chemical Carcinogenesis and Pharmacogenetics, Department of Basic-Clinical Oncology (DOBC), Faculty of Medicine, University of Chile, Santiago, Chile, [2]National Cancer Institute, Santiago, Chile, [3]Institute of Population Health, Faculty of Medicine, University of Chile, Santiago, Chile, [4]Faculty of Medicine, University of Chile, Santiago, Chile, [5]Genetic Division, Department of Medicine, Hospital de Clínicas José de San Martín, Buenos Aires, Argentina, [6]Latin American Network for Implementation and Validation of Clinical Pharmacogenomics Guidelines (RELIVAF-CYTED), Madrid, Spain

OPEN ACCESS

Edited by:
Chonlaphat Sukasem,
Mahidol University, Thailand

Reviewed by:
Umamaheswaran Gurusamy,
University of California, San Francisco, United States
Chalirmporn Atasilp,
Thammasat University, Thailand
Monpat Chamnanphon,
Srinakharinwirot University, Thailand

***Correspondence:**
Luis A. Quiñones
lquinone@uchile.cl
Nelson M. Varela
nvarela@med.uchile.cl

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: 30 January 2021
Accepted: 25 October 2021
Published: 25 November 2021

Citation:
Miranda C, Galleguillos M, Torres R, Tardón K, Cáceres DD, Lee K, Redal MA, Varela NM and Quiñones LA (2021) Preliminary Pharmacogenomic-Based Predictive Models of Tamoxifen Response in Hormone-dependent Chilean Breast Cancer Patients. Front. Pharmacol. 12:661443. doi: 10.3389/fphar.2021.661443

Tamoxifen (TAM), a selective oestrogen receptor modulator, is one of the most used treatments in oestrogen receptor-positive (ER+) early and metastatic breast cancer (BC) patients. The response to TAM has a high degree of inter-individual variability. This is mainly due to genetic variants in *CYP2D6* gene, as well as other genes encoding proteins involved in the TAM pharmacokinetic and/or pharmacodynamic. Therefore, prediction of the TAM response using these genetic factors together with other non-genetic variables may be relevant to improve breast cancer treatment. Thus, in this work, we used genetic polymorphisms and clinical variables for TAM response modelling. One hundred sixty-two ER + BC patients with 2 years of TAM treatment were retrospectively recruited, and the genetic polymorphisms *CYP2D6*4, CYP3A4*1B (CYP3A4*1.001), CYP3A5*3, UGT2B7*2, UGT2B15*2, SULT1A1*2*, and *ESRA V364E* were analyzed by PCR-RFLP. Concomitantly, the therapeutic response was obtained from clinical records for association with genotypes using univariate and multivariate biostatistical models. Our results show that *UGT2B15*1/*2* genotype protects against relapse (OR = 0.09; $p = 0.02$), *CYP3A5*3/*3* genotype avoids endometrial hyperplasia (OR = 0.07; $p = 0.01$), *SULT1A1*1/*2* genotype avoids vaginal bleeding (OR = 0.09; $p = 0.03$) and *ESRA 364E/364E* genotype increases the probability of vaginal bleeding (OR = 5.68; $p = 0.02$). Logistic regression models, including genomic and non-genomic variables, allowed us to obtain preliminary predictive models to explain relapse ($p = 0.010$), endometrial hyperplasia ($p = 0.002$) and vaginal bleeding ($p = 0.014$). Our results suggest that the response to TAM treatment in ER + BC patients might be associated with the presence of the studied genetic variants in *UGT2B15*, *CYP3A5, SULT1A1* and *ESRA* genes. After clinical validation protocols, these models might be used to help to predict a percentage of BC relapse and adverse reactions, improving the individual response to TAM-based treatment.

Keywords: breast cancer, pharmacokinetics, pharmacodynamics, polymorphism, ADR, Relapse

1 INTRODUCTION

Breast cancer (BC) is the second leading cause of cancer death in women (American Cancer Society, 2017) and is caused by many factors, such as age, breast density, heredity, and exposure to oestrogens (Kumar et al., 2010). Breast cancer (BC) patients with oestrogen receptor positivity (ER+) are treated in addition to surgery, chemotherapy, and/or radiation therapy with hormone therapy (Maughan et al., 2010; (McDonnell and Wardell 2010; NCI 2012). Tamoxifen (TAM), a type of hormonal therapy known as a selective oestrogen receptor modulator, behaves as an oestrogen antagonist in the breast and as an agonist in the endometrium (Lewis and Jordan 2005; Terán and Teppa 2005; Hertz et al., 2009; NCI 2012), and despite the new hormonal agents, TAM remains the drug of choice for any BC stage (Davies et al., 2011) since it increases the disease-free time in pre- or postmenopausal women, reducing the annual death rate by 34% (MINSAL, 2005; Untch and Thomssen, 2010). However, the response among patients is variable in terms of efficacy and side effects.

TAM undergoes extensive biotransformation in the liver by phase I (CYP) and phase II enzymes. Phase I generates, among other metabolites, N-desmethyltamoxifen (N-desmethylTAM), 4-hydroxy tamoxifen (4-hydroxyTAM), and 4-hydroxy-N-desmethyltamoxifen (endoxifen) (Brauch et al., 2009). Endoxifen and 4-hydroxyTAM are the main and unique active metabolites of TAM due to their high affinity to ERs and their activity being 30- to 100-fold greater than TAM itself (Wakeling and Slater 1980; Johnson et al., 2004; Fernández-Santander et al., 2007; Brauch et al., 2009). *In vitro* studies have determined that these metabolites are more effective in reducing cell proliferation (Jordan et al., 1977; Clarke et al., 2003; Borges et al., 2006). The two most abundant metabolites in plasma are N-desmethylTAM and endoxifen (Clarke et al., 2003). Subsequently, these metabolites undergo biotransformation by phase II enzymes, the main enzymes involved being SULT1A1, UGT2B7, and UGT2B15 (Nishiyama et al., 2002). TAM has a long half-life, and steady-state concentrations are obtained at 4 weeks for TAM and at 8 weeks for N-desmethylTAM (Buckley and Goa, 1989; Lien et al., 1992; Kisanga et al., 2004).

Differences have been found between the response to TAM and the histological type, degree of differentiation of the breast tumor cells, age, and state of menopause in the patient. TAM has an overall response rate of almost 75% in patients with ER+ and PR + BC. In the case of adjuvant treatment, TAM reduces the risk of relapse by 25% and mortality by 17% (Fisher et al., 1998; Davies et al., 2011). On the other hand, the long-term safety of TAM is well elucidated, and possible harmful effects on bone metabolism, endometrial cancer, thromboembolic diseases, and cognitive disorders due to chronic oestrogen deprivation in brain tissue have been described (Fisher et al., 1998; Cuzick et al., 2002; Cuzick, 2005). A 1% increase in the incidence of endometrial cancer has been observed due to its oestrogen-inducing effect in the uterus and thromboembolic events in 1–2% of cases.

Despite several studies having been conducted, after 3 decades, there are still differences in the TAM response that have not been explained. It is known that these variations in the response to drugs are multifactorial, the result of the interaction between multiple and complex genetic, physiological and environmental factors (Ma and Lu 2011). A potential explanation is the presence of genetic variants in the genes encoding biotransformation enzymes affecting their efficacy and safety (Zhou et al., 2008). In this respect, undoubtedly genetic variation in the *CYP2D6* driving to ultrarapid (XN), intermediate metabolizer or poor metabolizer phenotypes (e.g. *3, *4, *5, *9, *10, *17 haplotypes) may affects plasma concentrations of TAM and its metabolites, as CPIC has stablished in the updated guidelines (Goetz *et a*l, 2018). *CYP2D6*4* (rs3892097) is associated with decreased side effects, such as hot flashes, disease-free time, and survival rate (Goetz et al., 2005; Jin et al., 2005; Schroth et al., 2007; Bijl et al., 2009; Ramón y Cajal et al., 2010; Rae et al., 2012), and decreased serum levels of TAM metabolites (Gjerde et al., 2007; Kiyotani et al., 2010; Lim et al., 2011; Madlensky et al., 2011). In addition, *CYP3A4*1B* (currently named *CYP3A4*1.001*, rs2740574) is associated with a 3-fold higher risk of endometrial cancer in TAM-treated BC patients (Chu et al., 2007). *CYP3A5*3* (rs776746) is associated with characteristics of the tumor in BC postmenopausal women treated with TAM (Tucker et al., 2005). On the other hand, *CYP2C9* *2 and *3 variants reduce the concentration of TAM metabolites, affecting the therapeutic response (Mürdter et al., 2011).

Phase II enzymes also affect TAM response. *SULT1A1*2* (rs9282861) increases the risk of relapse in TAM-treated BC patients (Wegman et al., 2005). *UGT2B15*2* (rs1902023) decreases the risk of relapse of BC (Nowell et al., 2005). Furthermore, the presence of *UGT2B15*2* and *SULT1A1*2* is associated with a lower risk of relapse and a significant reduction in survival time in patients with BC treated with TAM (Nowell et al., 2005).

Finally, several mutations have been identified in the oestrogen receptor α gene (*ESRA*) in BC patients (Herynk et al., 2004); however, their effect on the efficacy and safety of TAM treatment has not been elucidated. In this respect, using SIFT and PolyPhen bioinformatic tools, it was established that the single nucleotide polymorphism (SNP) *ESRA* (OMIM* *133430) *V364E* (rs121913044; 1453T>A; Val364Glu) generates a deleterious change with possible damage to the N-terminus of the hormone-binding domain in the ER, giving rise to a 40 times lower affinity for oestrogen (Flicek et al., 2011). This SNP also exhibits a dependence on oestrogen for binding to an *ERE*, although it maintains its negative dominant activity entirely. Therefore, *ESRA* V364E is highly active and capable of repressing ER-mediated transcription, both when *ESRA* V364E and normal ESRA proteins are present together in cells and even without DNA binding (Wrenn and Katzenellenbogen 1993; Ince et al., 1995; McInerney et al., 1996).

Although in the last decade several studies have investigated genetic variants in TAM-metabolizing enzymes that might determine the differences in the response to treatment in patients with BC, there are still controversies about their relationship with the response (Brauch and Schwab, 2014; Brewer et al., 2014; Province et al., 2014; Binkhorst et al., 2015). To address this question, we aimed to study the association among response (relapse and ADR) and 7 genetic

variants in genes encoding proteins involved in the pharmacokinetics and pharmacodynamics of TAM (*CYP2D6*4, CYP3A4*1B (CYP3A4*1.001), CYP3A5*3, SULT1A1*2, UGT2B7*2, UGT2B15*2* and *ESRA* V364E) in women with hormone-dependent BC and under adjuvant treatment with TAM. The main goal is to generate predictive models to approximate the response of patients according to their genetic-metabolic characteristics.

2 PATIENTS AND METHODS

2.1 Study Design

A retrospective case-control study was carried out from January 2014 to January 2015 at the Polyclinic of Oncology of the National Institute of Cancer (INC). The sample size was determined according to the frequency of carriers with the variant allele carriers in the population under study, using PS Power and Sample Size Calculations Version 3.0, January 2009, considering an 80% power; $\alpha = 0.05$; OR = 2.0; and the less frequent *CYP3A4*1B* (*CYP3A4*1.001*, rs2740574), according to the literature (Fuentes and Silveyra, 2019).

2.2 Patients

One hundred sixty-two (162) patients were enrolled. The inclusion criteria were as follows: 1) patients with histologically confirmed BC, 2) in adjuvant treatment with 20 mg of tamoxifen (Novadex®) daily for at least 24 months; 2) > 18 years-old, 3) positive ER and negative HER-2 status, 4) cell differentiation degree I to III. The exclusion criteria were as follows: 1) patients diagnosed with *in situ* BC; 2) negative ER, negative PR, triple-negative or Her-2 positive patients; 3) concomitant treatments with vitamin K antagonists (warfarin, acenocoumarol or dicoumarol), serotonin reuptake antagonist antidepressants (fluoxetine or paroxetine), mitomycin, nitonavir, primidone, or chemotherapy (fluorouracil, methotrexate or cyclophosphamide), 4) chronic unbalanced systemic pathology or other active cancers and, 5) aromatase inhibitor treatment or LHRH agonist. The events (relapses and ADR) were evaluated after 6 months of TAM treatment. All the patients signed a written consent form and agreed to be included in this study.

Toxicity was assessed based on the Common Terminology Criteria for Adverse Events, v4.03. Vaginal bleeding was defined as any vaginal bleeding in postmenopausal patients (more than 1 year without menstruation or oophorectomized) or any abnormal bleeding, such as unexpected bleeding or in quantity or habitual bleeding in premenopausal patients. Endometrial hyperplasia was defined as an endometrial thickness greater than 5 mm detected by transvaginal echotomography in postmenopausal patients or an abnormal thickness of the endometrium according to the phase of the menstrual cycle in premenopausal patients.

The ethnicity of the study group was approached by the percentage of Amerindian-Caucasian admixture ($\%M_{a\text{-}c}$) based on the ABO system (Valenzuela et al., 1987; Acuña et al., 2000).

The patients were gathered into two groups: the control group include patients who do not present the analyzed ADRs (relapse, hot flashes, etc.), and the cases group include patients who have the ADRs.

2.3 Genotyping Analyses

Potentially functional SNPs encoding the proteins related to the TAM response were obtained from the PharmGKB database (Whirl-Carrillo et al., 2012), the NCBI dbSNP database (https://www.ncbi.nlm.nih.gov/snp/), the SNPinfo Web Server (https://snpinfo.niehs.nih.gov), the Ensembl® genome database project (https://www.ensembl.org/index.html) and the level of evidence for each SNP (**Supplementary Table S1**). Genomic DNA was isolated from the peripheral blood samples of the subjects using a Genomic DNA Extraction Blood DNA Kit (FAVORGEN®, BIOTECH CORP, Headquarters, Taiwan, China) and from buccal mucosa cells using a MasterAmp™ Buccal Swab Kit (Epicentre®, an Illumina company, Madison, USA). DNA samples were quantified at 260/280 nm using a Nanodrop spectrophotometer (model DS-11, FX Series Spectrophotometer, USA). *CYP2D6*4* (rs3892097), *CYP3A4*1B* (*CYP3A4*1.001*, rs2740574), *CYP3A5*3* (rs776746), *SULT1A1*2* (rs9282861), *UGT2B7*2* (rs7439366), *UGT2B15*2* (rs1902023), and *ESRA* V364E (rs121913044) were analyzed using polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) (Cavalli et al., 2001; Schur et al., 2001; Hajdinjak and Zagradišnik 2004; Lee et al., 2005; Kagaya et al., 2007; Arslan et al., 2011). The primers and restriction enzymes used are presented in **Supplementary Table S2**. Each assay contained four controls: one sample for each genotype as a positive control and one negative sample with nuclease-free pure water to volume. For quality assurance purposes, we randomly chose 20% of the samples for repetition of the analysis. When analyses were not coincident, we excluded the samples.

2.4 Statistical Analyses

Statistical analyses were performed using GraphPad Prism 4.0 and STATA 11.1, with $p < 0.05$ considered statistically significant. The results are expressed as the mean ± standard deviation (SD), number, percentage, or frequency where appropriate. The Shapiro-Wilk test was used to determine quantitative variable distributions. Mean values between the two groups were compared using the unpaired t-test for variables with a normal distribution, while the Mann-Whitney U test was used for variables with non-normal distributions. Frequencies of qualitative variables were also calculated. The chi-square test with Fischer's exact test was used to investigate differences in genotypic and allelic frequencies between the two groups.

We checked for Hardy-Weinberg equilibrium (HWE) in our sample even though the conditions for HWE are not applicable because it is not a random sampling in a random-mating population, a control population or the general population (Namipashaki et al., 2015) and is a group with a selection bias based on the disease (i.e., SNPs can also be related to cancer). The polymorphic variants *CYPD6**4, *CYP3A4*1B* (*CYP3A4*1.001*), *CYP3A5**3, *SULT1A1**2, and accomplish HWE, while *UGT2B15**2 and *ESRA V364E* do not (cut-off chi^2: 3.84).

TABLE 1 | Characteristics of patients (n = 162).

Variables	N*	(%)	$\bar{x} \pm SD$
Anthropometric characteristics			
Age (years)	160	(98.77)	58 ± 13
Weight, (Kg)	160	(98.77)	69 ± 15
Height, (m)	156	(96.30)	1.55 ± 0.06
BMI (Kg/m^2)	156	(96.30)	29 ± 6
Socio-genetic gradient			
Blood type			
AB	3	(1.85)	
A	21	(12.96)	
B	8	(4.94)	
O	66	(40,74)	
N.D	64	(39.51)	
Number of members in the family			3 ± 2
Socioeconomic (income/member)			
<$CLP135,000 (U$ 200)	33	(21.15)	
$CLP135,001–500,000 (U$ 200–750)	91	(58.33)	
$CLP 500,001–1,000,000 (U$ >750–1,450)	27	(17.31)	
>$CLP 2,000,000 (U$ >2,900)	5	(3.21)	
N.D	6	(3.85)	
Risk factor's			
Alcoholic Habit Presence	0	(100)	
Presence of Smoking Habit	48	(30.19)	
Presence family history of some cancer	100	(62.89)	
Presence Family History of breast or ovary cancer	43	(27.04)	
Gynecological Characteristics			
Menarche age (years)	153	(94.44)	13 ± 2
Number of Gestations	160	(98.77)	3 ± 2
Number of deliveries	160	(98.77)	3 ± 2
Number of Abortions	160	(98.77)	1 ± 1
Breastfeeding time (months)	137	(84.57)	24 ± 28
Oral Contraceptive Treatment (months)	154	(95.06)	39 ± 69
Menopausal status	68	(41.98)	
Premenopause			
Postmenopause	94	(58.02)	
Treatment with HRT (months)	21	(22,34)	11 ± 46
Pathological Features			
Age of diagnosis (years)	160	(98.77)	54 ± 13
Cancer stage at diagnosis			
I	59	(36.42)	
II	83	(51.23)	
III	20	(12.35)	
Tumor Histology			
In situ Ductal carcinoma (DCis)	4	(2.47)	
Invasive Ductal Carcinoma (IDC)	139	(85.80)	
Invasive Lobular Carcinoma (ILC)	8	(4.94)	
Others, (IBC, IPC, etc.)	9	(5.56)	
N.D.	2	(1.23)	
Cell Differentiation Degree			
G1	41	(25.31)	
G2	78	(48.15)	
G3	25	(15.43)	
N.D.	18	(11.11)	
Treatment before TAM			
Surgery	12	(7.41)	
Surgery + radiotherapy	37	(22.84)	
Surgery + chemotherapy	13	(8.02)	
Surgery + chemotherapy + radiotherapy	36	(22.22)	
No treatment	64	(39.51)	

N.D: no data; TAM: tamoxifen; SD: standard deviation; IBC: inflammatory breast cancer; IPC: intracystic papillary carcinoma; HRT: hormone replacement therapy.

[a]*some data for patients were lost from the clinical files, so the numbers are lower than the total.*

TABLE 2 | Relapse and Adverse drug reactions (ADRs) in patients.

Clinical response	N	(%)
Relapse		
No	148	(91.36)
Yes	9	(5.56)
N.D.	5	(3.09)
ADRs		
Endometrial cancer		
No	161	(99.38)
Yes	1	(0.62)
Endometrial hyperplasia		
No	147	(90.74)
Yes	15	(9.26)
Vaginal bleeding		
No	154	(95.06)
Yes	8	(4.94)
Phlebitis		
No	160	(98.77)
Yes	2	(1.23)
Headache		
No	157	(96.91)
Yes	5	(3.09)
Nausea		
No	156	(96.30)
Yes	6	(3.70)
Hot flash		
No	62	(38.27)
Yes	100	(61.73)
Cramps		
No	142	(87.65)
Yes	20	(12.35)
Bone pain		
No	140	(86.42)
Yes	22	(13.58)
Urticaria		
No	158	(97.53)
Yes	4	(2.47)

ADR, adverse drug reaction, evaluated with Common Terminology Criteria for Adverse Events [CTCAE], 2010. N.D: No data. Control groups include patients who do not present the analyzed ADRs (relapse, hot flashes, etc., i.e. "No"), and Cases groups include patients who have the ADRs (i.e. "Yes").

The polymorphisms were evaluated using co-dominant (wild type vs heterozygote vs variant), dominant (wild type vs. heterozygote/variant), and recessive (wild type/heterozygote vs. variant) inheritance models. Bi-variable analyses were performed to determine the association between the events (relapses and ADR), considered the dependent variables, and the characteristics of the patients, considered the independent variables. Thus, the independent variables considered were BMI, blood group, smoking habit, socioeconomic status, treatment with oral contraceptives, hormone replacement therapies, age at menarche, menopausal status, family history of cancer, age at diagnosis, cancer stage, and presence of genetic variants (exposure variable).

Multivariate logistic regression models and multivariate linear regression analyses were employed to investigate associations between genetic or non-genetic characteristics and relapse or ADR. The logistic multivariate models were adjusted stepwise using a forward and backward procedure to include potentially relevant variables to derive statistical association models characterized by pseudo R^2. All association studies were assayed by choosing parameters with a better statistical association for each analysis. The odds ratio (OR), or coefficients (Coef.), and 95% confidence intervals (CIs) are reported in the multivariate logistic regression models.

2.5 Ethics

The study was carried out under strict ethical procedures recommended by the Ethics Committee of the Faculty of Medicine of the University of Chile (July 24, 2013) and the

TABLE 3 | Genotype and allele frequencies of the studied polymorphisms in patients.

Polymorphisms	N Frequency	
Enzymes involved in the activation of TAM		
CYP2D6		
*1/*1 (GG)	121	(0.747)
*1/*4 (GA)	37	(0.228)
*4/*4 (AA)	4	(0.025)
*1 (G)	279	(0.861)
*4 (A)	45	(0.139)
**CYP3A4*		
*1/*1 (AA)	144	(0.889)
*1/*1B (AG)	17	(0.105)
*1B/*1B (GG)	1	(0.006)
*1 (A)	305	(0.941)
*1B (G)	19	(0.059)
CYP3A5		
*1/*1 (AA)	5	(0.031)
*1/*3 (AG)	59	(0.364)
*3/*3 (GG)	98	(0.605)
*1 (A)	69	(0.213)
*3 (G)	255	(0.787)
Enzymes involved in the elimination of TAM and its metabolites		
SULT1A1		
*1/*1 (GG)	33	(0.204)
*1/*2 (GA)	88	(0.543)
*2/*2 (AA)	41	(0.253)
*1 (G)	154	(0.475)
*2 (A)	170	(0.525)
UGT2B7		
*1/*1 (TT)	19	(0.117)
*1/*2 (TC)	72	(0.444)
*2/*2 (CC)	71	(0.438)
*1 (T)	110	(0.340)
*2 (C)	214	(0.660)
UGT2B15		
*1/*1 (AA)	20	(0.123)
*1/*2 (AC)	94	(0.580)
*2/*2 (CC)	48	(0.296)
*1 (A)	134	(0.414)
*2 (C)	190	(0.586)
Estrogen receptor. TAM therapeutic target		
ESRA V364E		
364V/364V (TT)	101	(0.623)
364V/364E (TA)	33	(0.204)
364E/364E (AA)	28	(0.173)
364V (T)	235	(0.725)
364E (A)	89	(0.275)

TAM: tamoxifen; CYP: Cytochrome P450; SULT: sulfotransferase; UGT: Uridine 5′-diphospho-Glucuronosyltransferase, ESR: estrogen receptor.
**CYP3A4*1B is currently CYP3A4*1.001 according to PharmGKB (pharmgkb.org).*

TABLE 4 | Univariable logistic regression analysis of risk of severe ADRs according to genotypes.

	Cases	Controls	OR**	95%CI***	*p-value*[#]
Efficacy					
Relapse					
UGT2B15					
*1/*1 (AA)	4	15	Ref	—	—
*1/*2 (AC)	2	85	0.088	0.015–0.525	**0.008**
*2/*2 (CC)	3	40	0.281	0.056–1.407	0.123
*1/*1 (AA)	4	16	Ref		
*2/*2 (CC) + *1/*2 (AC)	5	125	0.160	0.031–0.910	**0.019**
*1/*1 (AA) + *1/*2 (AC)	6	100	Ref		
*2/*2 (CC)	3	40	1.305	0.193–6.188	0.509
Safety					
Endometrial hyperplasia					
CYP3A5					
*1/*1 (AA)	2	3	Ref.	—	—
*1/*3 (AG)	9	48	0.281	0.041–1.930	0.197
*3/*3 (GG)	4	89	0.067	0.009–0.524	**0.010**
*1/*1 (AA)	2	3	Ref.		
*3/*3 (GG) + *1/*3 (AG)	13	137	0.142	0.015–1.891	0.074
*1/*1 (AA) + *1/*3 (AG)	11	51	Ref.		
*3/*3 (GG)	4	89	0.208	0.047–0.758	**0.007**
Vaginal bleeding					
SULT1A1					
*1/*1 (GG)	4	28	Ref.	—	—
*1/*2 (GA)	1	81	0.086	0.009–0.806	**0.032**
*2/*2 (AA)	3	38	0.552	0.115–2.668	0.460
*1/*1 (GG)	4	28	Ref.		
*2/*2 (AA) + *1/*2 (GA)	4	119	0.235	0.042–1.361	0.057
*1/*1 (GG) + *1/*2 (GA)	5	109	Ref.		
*2/*2 (AA)	3	38	1.668	0.254–9.300	0.357
ESRA V364E					
364V/364V (TT)	3	96	Ref.	—	—
364V/364E (TA)	1	29	1.10	0.111–11.017	0.933
364E/364E (AA)	4	22	5.81	1.214–27.882	**0.028**
364V/364V (TT)	3	96	Ref.	—	—
364E/364E (AA) + 364V/364E (TA)	5	51	3.137	0.579–20.848	0.114
364V/364V (TT) + 364V/364E (TA)	4	125	Ref.	—	—
364E/364E (AA)	4	22	5.68	0.966–32.397	**0.028**

**ADR, adverse drug reaction, evaluated with CTCAE4.03; **OR, odds ratio; ***95% CI, 95% confidence interval; [#]Logistic regression. Only statistically significant associations are shown (p < 0.05).*

Research Ethics Committee of the Northern Metropolitan Health Service (June 3, 2013), in accordance with the procedures suggested in the Declaration of Helsinki (WMA, 2013) and according to Chilean Laws 20.120, 20.584, and 19.628 and the guidelines of Good Clinical Practices. The treatment schedule for patients was according to the cancer stage (I-III), histology, cell differentiation degree, and presence and relative abundance of 3 selected differentially abundant proteins (ER+, PR+, Her2-), which involved surgery followed by radiotherapy and/or chemotherapy according to the Breast Cancer Clinical Guide, 2nd Ed. Santiago, Chile.

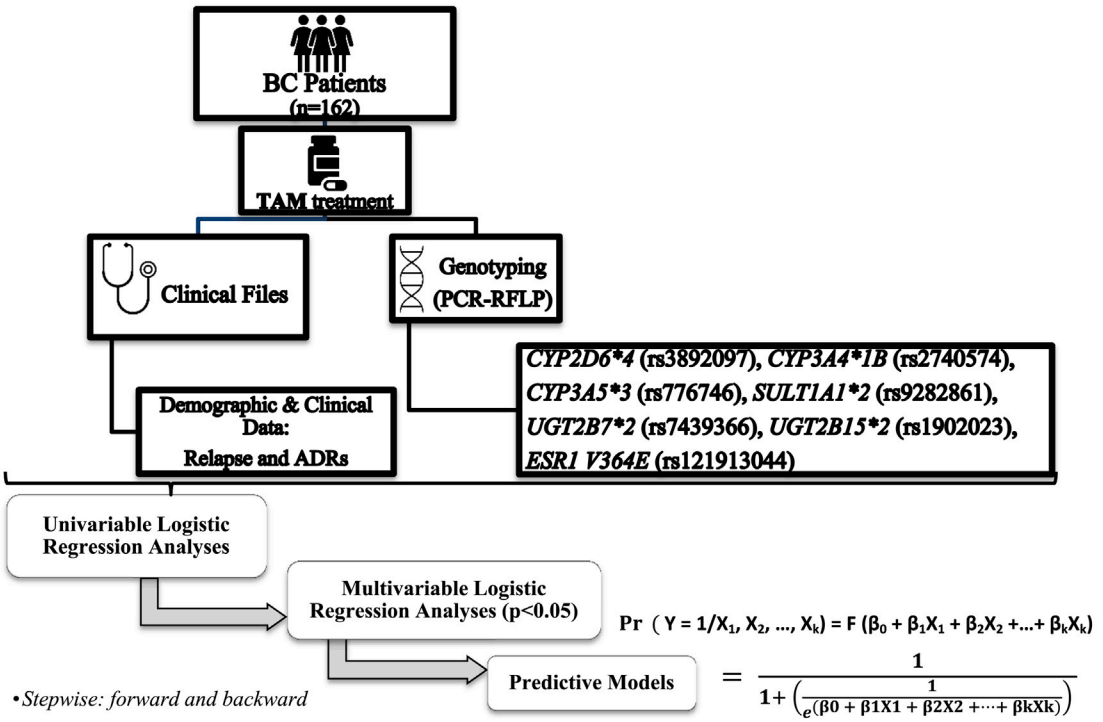
3 RESULTS

The baseline characteristics of the patients (58 ± 13 years; BMI: 29 ± 6) are shown in **Table 1**, and the characteristics of the patients according to genotype are shown in **Supplementary Table S3**. Sociogenetic gradients, risk factors, gynecological characteristics, and pathological features for descriptive analyses can also be observed. Of the total patients, 68 were premenopausal, and 94 were postmenopausal. In relation to the cancer stage at diagnosis, the majority of patients were in stage II (51.23%), the predominant histology type was invasive ductal carcinoma (86.88%), and the cell differentiation degree was predominantly G2 (54.17%). A 49.3% $M_{a\text{-}c}$ was found.

Table 2 shows the therapeutic response characteristics of the patients. Relapse was found in 9 patients (5.73%). The most severely observed ADRs among patients were endometrial hyperplasia (9.20%) and vaginal bleeding (4.91%), and the most frequent were hot flashes (61.96%), bone pain (13.50%), and cramps (12.27%).

The genotypic and allelic frequencies for the analyzed polymorphisms are shown in **Table 3**.

We performed univariable logistic regression of the risk of relapse (**Supplementary Tables S4, S5**), endometrial hyperplasia (**Supplementary Tables S6, S7**), and vaginal bleeding (**Supplementary Tables S8, S9**) in association with genetic



- **Ln(Relapse)=** -21.870+18.823(SEII)+18.776(SEIII)+21.888(SEIV)-0.944(HBC)-0.402(G)+0.526(A)-0.121(AgeDiag)+5.153(menop) +0.510(SII)+5.353(SIII)-0.272(*SULT1A1**1/*2)+0.809(*SULT1A1**2/*2)-4.585(*UGT2B15**1/*2)-1.150 (*UGT2B15**2/*2), p = 0.010; Pseudo R^2 = 0.495
- **Ln(Endometrial hyperplasia)**= -17.793+0.147(BMI)+ 0.477(HBOC)-1.313(menop)-3.446(SII)-3.132(SIII)+13.387(IDC)+17.433(OH)+1.512 (*CYP2D6**1/*4)-2.096(*CYP3A4**1/*1B)+1.892(*CYP3A5**1/*3)-1.037 (*CYP3A5**3/*3)-1.631 (*SULT1A1**1/*2)-0.526 (*SULT1A1**2/*2)-1.071(*UGT2B15**1/*2)+1.374(*UGT2B15**2/*2)-1.701*(ESR1364V/364E)*+1.332 (*ESR1364E/364E),* p = 0.002; Pseudo R^2 = 0.42
- **Ln(vaginal bleeding)**= -4.011+0.051(BMI)-0.008(OCT)-0.745(menop)+1.876(*CYP2D6*1/*4*)+1.948(*CYP3A4*1/*1B*)-3.805(*SULT1A1*1/*2*) -1.181(*SULT1A1**2/*2)+0.305(*ESR1364V/364E*) + 2.902 *(ESR1364E/364E)*, p = 0.014; Pseudo R^2 = 0.34
- **Ln(Phlebitis)**=-50.071-3.068(BMI)+7.682(M)-46.463(G)+57.795(P)-2.49(L)+4.938(Menop), p = 0.036; Pseudo R^2 = 0.66
- **Ln(Headache)**=4.56+3.55(HBOC)+1.101(P)-0.089(L)-0.21(AgeDiag)+1.483(SII)+1.565(*CYP3A5**1/*3)-1.676 (*UGT2B7**1/*2)-4.406 (*UGT2B7**2/*2), p = 0.028; Pseudo R^2 = 0.42
- **Ln(Nausea)**=-1.243+0.710(Nm)+1.033(P)-1.125(AgeDiag)-1.415(*CYP2D6**1/*4)-1.126(*SULT1A1**1/*2)-0.526 (*ESR1* 364V/364E), p = 0.013; Pseudo R^2 = 0.36
- **Ln(hot flashes)**=-1.436+0.171(Nm)+0.304(HBOC)-0.159(M)+0.006(L)-0.586(Menop)+2.679(*CYP3A5**1/*3)+2.191(*CYP3A5**3/*3)+1.852 (*UGT2B7**1/*2)+1.745(*UGT2B7**2/*2), p = 0.023; Pseudo R^2 = 0.12
- **Ln(Cramps)**=0.606-0.056(BMI)+2.083(HBOC)-0.013(OCT)-1.036(A)-0.014(L)-0.012(HRT)-0.327(*CYP2D6**1/*4)-1.281 (*CYP3A4**1/*1B)-1.330(*CYP3A5**1/*3)-2.503(*CYP3A5**3/*3)+1.612(*UGT2B7**1/*2)+1.033(*UGT2B15**1/*2)+1.695 (*UGT2B15**2/*2), p = 0.053; Pseudo R^2 = 0.25
- **Ln(Bone pain)**=-4.936+0.06(BMI)-0.288(G)+0.048(AgeDiag)-0.317(*CYP2D6**1/*4)+0.837(*CYP2D6**4/*4)-1.368 (*ESR1 364V/364E*)-2.075 (*ESR1 364E/364E*), p = 0.017; Pseudo R^2 = 0.14
- **Ln(Urticaria)**=2.917-0.363(BMI)-0.086(L)+1.701(Menop) +3.234(*CYP3A4**1/*1B)+3.499(*UGT2B7**1/*2), p = 0.039; Pseudo R^2 = 0.340

FIGURE 1 | Research Scheme and obtained multivariate predictive models. ADRs, adverse drug reactions; BC: Breast Cancer; SULT: Sulfotransferase; UGT: UDP-glucuronosyl transferase; CYP: Cytochrome P-450; ESRA: Estrogen Receptor 1; SE: Socioeconomic status (per capita income), I: <$135.000 CLP (U$ 200); II: $135.001-$500.000 CLP (U$ 200–750); III: $500.001-$1.000.000 CLP (U$ >750–1,450); IV: >2.000.000 CLP >$CLP 2,000,000 (U$ >2,900); HBC: Family History of Breast Cancer; HBOC: Family History of breast or ovary cancer; M: Menarche age; G: Number of Gestations; A: Number of Abortions; P: Number of deliveries; L: Breastfeeding time; Age diag: Age of diagnosis, years; OCT: Oral Contraceptive Treatment; HRT: Treatment with HRT for menopause; Menop: postmenopause; S: Cancer stage at diagnosis (Stage I-IV); DCis: Ductal carcinoma in situ; IDC: Invasive Ductal Carcinoma; ILC: Invasive Lobular Carcinoma; OH: Others Histology (OH).

and non-genetic variables. In **Table 4**, only statistically significant results for the univariable logistic regression analysis (risk of relapse, endometrial hyperplasia, and vaginal bleeding) are shown. The results show that *UGT2B15*2* A/C (rs1902023) in a dominant model of inheritance was associated with relapse; *CYP3A5*3* A/G (rs776746) in a recessive model of inheritance was associated with endometrial hyperplasia; and *SULT1A1*2* G/A (rs9282861) in a dominant model of inheritance and *ESRA* V364E (rs121913044) in a recessive model of inheritance were associated with vaginal bleeding.

After using stepwise forward and backward procedures, multivariate logistic regression analyses for the risk of relapse and ADRs, including genetic and non-genetic variables, were performed. We obtained significant models for endometrial hyperplasia, vaginal bleeding, phlebitis, headache, nausea, hot flash, cramps, bone pain, and urticaria (**Figure 1**). **Supplementary Tables S10–S12** show significant multivariate logistic regression analyses or logit models for relapse (**Supplementary Table S10**), endometrial hyperplasia (**Supplementary Table S11**) and vaginal bleeding (**Supplementary Table S12**) with Pseudo R^2 = 0.495 (p = 0.01), Pseudo R^2 = 0.42 (p = 0.002) and Pseudo R^2 = 0.34 (p = 0.014), respectively.

Associations among genetic variants with any adverse effect *versus* no adverse effect or with some grades of adverse effects were also assessed, but no statistical significance was observed (**Supplementary Tables S13, S14**).

4 DISCUSSION

Patient response to TAM has been investigated for a long time. In this respect, differences of approximately 25–50% in patients are observed (Fisher et al., 1998; Davies et al., 2011). This is an important issue due to the harmful effects on bone metabolism, cancer of the endometrium, thromboembolic diseases, and cognitive disorders caused by chronic oestrogen deprivation in brain tissue (Fisher et al., 1998; Cuzick et al., 2002; Cuzick, 2005). Moreover, various side effects similar to those seen in menopause may occur, such as hot flashes, weight gain, vaginal dryness, nausea, and elevated liver transaminases (Cuzick, 2005; Carpenter and Miller 2005).

Although it is well known that genetic variants in *CYP2D6* could explain between 10–20% of these cases and its use as a biomarker for TAM is recommended by the FDA, there is still no agreement on its clinical utility for predicting results in BC, the studies reveal contradictory results and the role of other pharmacokinetic and pharmacodynamic proteins have been poorly studied (Brauch and Schwab, 2014; Brewer et al., 2014; Province et al., 2014; Binkhorst et al., 2015; Goetz et al., 2018). Therefore, the present research focused on studying the association between 7 genetic variants in several genes encoding proteins involved in the pharmacokinetics and pharmacodynamics of TAM and response (relapse) and ADRs in TAM-treated hormone-dependent BC women and contributing to the field, generating predictive models that approximate the response of patients according to their genetic-metabolic characteristics. CYP2D6*3 and *17 alleles were previously found to be absent in different sub-groups of the Chilean population (Roco et al., 2012; Varela et al., 2015) therefore we exclude them from this study and other potentially relevant variants could not be studied, unfortunately.

In this study, we recorded ADRs after 6 months of TAM treatment. This is because several other factors influence TAM ADRs, especially before TAM active metabolites reach steady state, which is after 2 months of treatment (Buckley and Goa, 1989; Lien et al., 1992; Kisanga et al., 2004). Therefore, we considered that after 6 months of treatment, the ADRs might be associated mainly with TAM. Similarly, after 2 years of treatment, relapses can be properly evaluated (Cheng et al., 2012).

The study patients have a 49.3% $M_{a\text{-}c}$, which agrees with that described by Valenzuela *et al.* in 1987, in relation to admixture and socioeconomic stratum, because main of the enrolled population of patients (94%) belong to a low socioeconomic stratum with higher Mapuche ancestry; this was an expected result according to the place of recruitment of the patients.

The study of the association between genetic variants and clinical parameters showed that the presence of *UGT2B15*1/*2* protects against relapse (OR = 0.09; p = 0.008), *CYP3A5*3/*3* protects against presenting endometrial hyperplasia (OR = 0.07, p = 0.01), and *SULT1A1*1/*2* protects against vaginal bleeding (OR = 0.09; p = 0.03) and that *ESRA* 364E/364E is a risk factor for vaginal bleeding (OR = 5.68; p = 0.03). In this respect, it seems to be relevant to perform more studies regarding the *ESRA* V364E variant, a very poorly studied variant that can be determined in patients with ER + BC to prevent vaginal bleeding caused by TAM therapy, which is also a suspicious sign of endometrial cancer.

In our study, some analyses showed associations without statistical significance, which may be due to the low number of occurrences of the event in each subgroup, lack of clinical data in any of the contrasting groups, or because the variant explains a smaller part of the response, which could be clarified by increasing the sample size in future studies. Thus, in small groups, Fisher's exact test was used to obtain proper results.

The biostatistical tools used in this field have established that non-significant trends, but with $p < 0.2$ and physiological significance, can be included in predictive models and in multi-variable logistic regression analysis to help predict the response (relapses) and adverse reactions to treatment (Núñez et al., 2011). Therefore, preliminary predictive models to predict the response (efficacy and safety) in TAM-treated BC patients were generated. The efficacy result, analyzed as relapse, led to a model that explains 49.5% of the possibility of presenting a relapse (p = 0.01) by including *SULT1A1*2* and *UGT2B15*2* and a series of relevant non-genomic variables. These results were expected since both the SULT1A1 and UGT2B15 enzymes have specificity for 4-hydroxyTAM, even though the genetic variant *UGT2B15*2* presents a nucleotide change located in a substrate binding site that is associated with a decrease in catalytic activity and alterations in the Km and Vmax values (Lévesque et al., 1997). Similarly, *SULT1A1*2* has a lower catalytic activity than the wild-type allele (Arslan et al., 2011). Thus, these variants would produce a decrease in the elimination of 4-hydroxyTAM.

Moreover, our results are similar to those published by Nowell and his colleagues (Nowell et al., 2002 and Nowell et al., 2005), who conducted a retrospective study in 162 TAM-treated BC patients from Arkansas who received TAM to determine whether genetic variability (*CYP2D6*4, SULT1A1*2,* and *UGT2B15*2*) in the TAM metabolic pathway was associated with recurrence. They found that high activity genotypes *UGT2B15 *1/*1* and combined genotypes *UGT2B15*1/*1* and *SULT1A1*2/*2* lead to an increased risk of disease recurrence and that *SULT1A1*2/*2* causes an approximately 3 times higher risk of death than *SULT1A1*1/*1* and **1/*2*. Conversely, Wegman et al. (2005) found that the genetic variants *CYP2D6*4/*4,* and *SULT1A1*1/*1* cause an increase in the disease-free relapse time in patients with premenopausal and postmenopausal BC treated with 40 mg TAM using proportional hazards Cox regression adjusted for age, tumor size, and lymph nodes. However, Ahern et al. (2011) and Lash et al. (2011) in homologous models found that the genetic variants *CYP2D6*4, UGT1A8*3, UGT2B7*2,* and *UGT2B15*2* did not modify the relapse rate of the disease in patients with ER+/TAM + BC. On the other hand, Sensorn et al., 2013 investigated genetic variant *CYP3A4*1B* (−392A>G, actually CYP3A4*1.001 G/A)/*18 (878T>C), CYP3A5*3 (6986 G>A), *ABCB1 3435 C>T*, *ABCC2*1C* (−24C > T) and *ABCC2 68231 A>G* to evaluate the risk of recurrence within 3 years among thirty Thai women after receiving tamoxifen adjuvant therapy but they could not find *CYP3A4*1B/*18*, and did not observe a statistical association between *CYP3A5*3, ABCC2*1C*, and *ABCC2 68231 A>G* with clinical outcome.

Additionally, preliminary predictive models were generated to predict ADRs. The endometrial hyperplasia model explained 41.6% of the possibility of presenting the event (p = 0.002). The model includes the *CYP2D6*4, CYP3A4*1B (CYP3A4*1.001), CYP3A5*3, SULT1A1*2, UGT2B15*2* and *ESRA* V364E genotypes plus several relevant non-genomic factors. Likewise, a preliminary predictive model was generated that explains 33.7% of the possibility of presenting vaginal bleeding (p = 0.014) by including the *CYP2D6*4, CYP3A4*1B (CYP3A4*1.001), SULT1A1*2,* and *ESRA* V364E genotypes and relevant non-genomic factors. Currently, there is no evidence regarding the association of genetic variants with TAM-induced vaginal bleeding or endometrial thickening. Chu et al. (2007) found that women with BC who carried *CYP3A4*1B (CYP3A4*1.001),* had a 3-fold increased risk of developing endometrial cancer after treatment with TAM. The phlebitis model explained 66% of the possibility of presenting this ADR (p = 0.036), a significant finding that has not been previously reported. In the same way, preliminary predictive models were generated to predict the appearance of headache (42.2%; p = 0.028), nausea (36.3%; p = 0.013), and hot flashes (12.1%; p = 0.023) in patients. Our results differ from those reported by Goetz et al. (2005), where *CYP2D6*4/*4* tends to contribute to a higher risk of disease relapse and a lower incidence of hot flashes in a similar study. No other studies were found in the literature describing associations with headache and nausea. Finally, predictive models to predict the appearance of cramps (24.6%; p = 0.053), bone pain (14.1%; p = 0.017) and urticaria (34%; p = 0.039) were obtained. We were not able to find any similar study in the scientific literature.

Altogether, our results show that some specific genetic variants in genes encoding proteins involved in the TAM pharmacokinetic and/or pharmacodynamic can influence the efficacy of TAM therapy in BC.

Since there are still discrepancies between the findings, which may be due to the genetic differences present in the subjects analyzed, it is crucial to continue with similar and larger studies that confirm the association of *CYP2D6*4, CYP2C9*2* and **3, CYP3A4*1B (CYP3A4*1.001), CYP3A5*3, SULT1A1*2, UGT2B15*2* and *ESR1* V364E and response to treatment with TAM. Differences in allele frequencies of variants influencing drug's pharmacokinetics across populations are considered to result in the interethnic differences in pharmacokinetics observed. For example, the lower mean body weights of populations of East Asian in relation to European populations are well established to contribute to differences in the clearance or volume of distribution of some drugs (Lu et al., 2018). On the other hand, in relation to that scarce information is available in Latin American populations, but allele frequencies heterogeneously differ from Asian, Caucasian and African populations (Mendoza et al., 2001; López et al., 2005; Roco et al., 2012; Friedrich et al., 2014), giving rise an idea of the phenotypic response and patient outcome.

Despite our analysis, the present study has some limitations. Even though we had a good sample size for combinatorial analyses, a small number of patients examined in each subgroup comparison could mask potential associations, especially for low-frequency polymorphisms, particularly in the multivariate analyses. The small number of variants analyzed could restrict the statistical power and predictive potential of models, considering that some other candidate genes/polymorphisms were not evaluated in this study (e.g. *CYP2D6 *2, *5, *6, *10* and **41*, *CYP2C9*2, CYP2C19 *2* and **17, UGT1A10 139Lys, UGT1A8* 173Gly/277Cys and *ABCC2*), which could still be relevant. In some cases, missing clinical data was relevant, giving rise to a possible differential misclassification bias affecting the estimated associations between potentially relevant combinations of risk factors and adverse reactions. Other environmental factors, such as contaminants and stress, or social determinants were not explored in this study, which could potentially affect the associations.

However, our findings and preliminary predictive models proposed can be used to prevent, in a good percentage of patients, the negative effects of TAM therapy. Additionally, they could help in the decision-making of a change in therapy to aromatase inhibitors (AIs) (Bernhard et al., 2015).

Certainly, for routine clinical application, the models must be validated in terms of their specificity, sensitivity, and predictive value, so it is necessary to carry out a second stage of research that involves treatment intervention in the patient according to their genotypic profile and relevant non-genomic variables derived from this study. It should be considered that the predictive models generated in this study after being validated could be used in the Chilean population. They could be used in another

population with the same percentage of Caucasian aboriginal mix.

5 CONCLUSION

In our study, *UGT2B15*2* A/C is likely to have a clinically significant impact on recurrence risk, *CYP3A5*3* A/G on endometrial hyperplasia, and *SULT1A1*2* G/A, and *ESRA* V364E on vaginal bleeding in Chilean patients with breast cancer who receive tamoxifen adjuvant therapy. The obtained predictive models showed that the response to TAM treatment in BC patients is partly associated with some of the genetic variants studied. This, after a clinical validation protocol, might help to predict a percentage of disease relapse and negative effects of TAM therapy, improving the efficacy and safety of TAM pharmacotherapeutic treatment.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the University of Chile (July 24, 2013), the Northern Metropolitan Health Service, National Cancer Institute (May 8, 2013), and Research Ethics Committee of the North Metropolitan Health Service (June 3, 2012. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CM: The conception of the research, enrolment of patients, experimental analyses, analysis of data and, writing the manuscript. MG: Experimental analyses and analyses of data. RT: Enrolment of patients. KT: Enrolment of patients. DC: Analysis of data. KL: Analysis of data. MR: Analysis of data and writing the manuscript. NV: Analysis of data and writing the manuscript. LQ: The conception of the research, analysis of data and, writing the manuscript.

ACKNOWLEDGMENTS

The authors wish to thank the breast cancer patients from the "Instituto Nacional del Cáncer" for their altruistic collaboration in pursuit of the common welfare. We also thank the Latin American Society of Pharmacogenomics and Personalized Medicine (SOLFAGEM) by sponsoring this article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.661443/full#supplementary-material

REFERENCES

Acuña, M., Llop, E., and Rothhammer, F. (2000). Composición genética de la población chilena: las comunidades urales de los valles de Elqui, Limarí y Choapa. *Rev. Méd. Chile* 128 (6), 593–600. doi:10.4067/S0034-98872000000600004

Ahern, T. P., Christensen, M., Cronin-Fenton, D. P., Lunetta, K. L., Søiland, H., Gjerde, J., et al. (2011). Functional Polymorphisms in UDP-Glucuronosyl Transferases and Recurrence in Tamoxifen-Treated Breast Cancer Survivors. *Cancer Epidemiol. Biomarkers Prev.* 20 (9), 1937–1943. doi:10.1158/1055-9965.EPI-11-0419

American Cancer Society (2017). "What Is Breast Cancer?" American Cancer Society. Cancer Facts and Figures. Atlanta, GA: American Cancer Society. Available at: http://www.cancer.org/cancer/breast-cancer/about/what-is-breast-cancer.html (Accessed June 20, 2021).

Arslan, S., Silig, Y., and Pinarbasi, H. (2011). Sulfotransferase 1A1 Arg(213)His Polymorphism and Prostate Cancer Risk. *Exp. Ther. Med.* 2 (6), 1159–1162. doi:10.3892/etm.2011.334

Bernhard, J., Luo, W., Ribi, K., Colleoni, M., Burstein, H. J., Tondini, C., et al. (2015). Patient-reported Outcomes with Adjuvant Exemestane versus Tamoxifen in Premenopausal Women with Early Breast Cancer Undergoing Ovarian Suppression (TEXT and SOFT): a Combined Analysis of Two Phase 3 Randomised Trials. *Lancet Oncol.* 16, 848–858. doi:10.1016/S1470-2045(15)00049-2

Bijl, M. J., van Schaik, R. H., Lammers, L. A., Hofman, A., Vulto, A. G., van Gelder, T., et al. (2009). The CYP2D6*4 Polymorphism Affects Breast Cancer Survival in Tamoxifen Users. *Breast Cancer Res. Treat.* 118 (1), 125–130. doi:10.1007/s10549-008-0272-2

Binkhorst, L., Mathijssen, R. H., Jager, A., and van Gelder, T. (2015). Individualization of Tamoxifen Therapy: Much More Than Just CYP2D6 Genotyping. *Cancer Treat. Rev.* 41 (3), 289–299. doi:10.1016/j.ctrv.2015.01.002

Borges, S., Desta, Z., Li, L., Skaar, T. C., Ward, B. A., Nguyen, A., et al. (2006). Quantitative Effect of CYP2D6 Genotype and Inhibitors on Tamoxifen Metabolism: Implication for Optimization of Breast Cancer Treatment. *Clin. Pharmacol. Ther.* 80 (1), 61–74. doi:10.1016/j.clpt.2006.03.013

Brauch, H., Mürdter, T. E., Eichelbaum, M., and Schwab, M. (2009). Pharmacogenomics of Tamoxifen Therapy. *Clin. Chem.* 55 (10), 1770–1782. doi:10.1373/clinchem.2008.121756

Brauch, H., and Schwab, M. (2014). Prediction of Tamoxifen Outcome by Genetic Variation of CYP2D6 in Post-Menopausal Women with Early Breast Cancer. *Br. J. Clin. Pharmacol.* 77 (4), 695–703. doi:10.1111/bcp.12229

Brewer, N. T., Defrank, J. T., Chiu, W. K., Ibrahim, J. G., Walko, C. M., Rubin, P., et al. (2014). Patients' Understanding of How Genotype Variation Affects Benefits of Tamoxifen Therapy for Breast Cancer. *Public Health Genomics* 17 (1), 43–47. doi:10.1159/000356565

Buckley, M. M., and Goa, K. L. (1989). Tamoxifen. A Reappraisal of its Pharmacodynamic and Pharmacokinetic Properties, and Therapeutic Use. *Drugs* 37 (4), 451–490. doi:10.2165/00003495-198937040-00004

Carpenter, R., and Miller, W. R. (2005). Role of Aromatase Inhibitors in Breast Cancer. *Br. J. Cancer* 93 (Suppl. 1), S1–S5. doi:10.1038/sj.bjc.6602688

Cavalli, S. A., Hirata, M. H., and Hirata, R. D. (2001). Detection of MboII Polymorphism at the 5' Promoter Region of CYP3A4. *Clin. Chem.* 47 (2), 348–351. doi:10.1093/clinchem/47.2.348

Cheng, L., Swartz, M. D., Zhao, H., Kapadia, A. S., Lai, D., Rowan, P. J., et al. (2012). Hazard of Recurrence Among Women After Primary Breast cancer Treatment–a 10-Year Follow-Up Using Data from SEER-Medicare. *Cancer Epidemiol Biomarkers Prev.* 21 (5), 800–809. doi:10.1158/1055-9965.EPI-11-1089

Chu, W., Fyles, A., Sellers, E. M., McCready, D. R., Murphy, J., Pal, T., et al. (2007). Association between CYP3A4 Genotype and Risk of Endometrial Cancer Following Tamoxifen Use. *Carcinogenesis* 28 (10), 2139–2142. doi:10.1093/carcin/bgm087

Clarke, R., Liu, M. C., Bouker, K. B., Gu, Z., Lee, R. Y., Zhu, Y., et al. (2003). Antiestrogen Resistance in Breast Cancer and the Role of Estrogen Receptor Signaling. *Oncogene* 22 (47), 7316–7339. doi:10.1038/sj.onc.1206937

Cuzick, J. (2005). Aromatase Inhibitors for Breast Cancer Prevention. *J. Clin. Oncol.* 23 (8), 1636–1643. doi:10.1200/JCO.2005.11.027

Cuzick, J., ForbesEdwards, J. R., Edwards, R., Baum, M., Cawthorn, S., Coates, A., et al. (2002). First Results from the International Breast Cancer Intervention Study (IBIS-I): A Randomised Prevention Trial. *Lancet* 360 (9336), 817–824. doi:10.1016/s0140-6736(02)09962-2

Davies, C., Davies, C., Godwin, J., Gray, R., Clarke, M., Cutter, D., et al. (2011). Relevance of Breast Cancer Hormone Receptors and Other Factors to the Efficacy of Adjuvant Tamoxifen: Patient-Level Meta-Analysis of Randomised Trials. *Lancet* 378 (9793), 771–784. doi:10.1016/S0140-6736(11)60993-8

Fernández-Santander, A., Gómez, A. T., Bernal, A. T., Gómez-Gallego, F., Santiago Dorrego, C., and Moya, F. B. (2007). *Aspectos Farmacogenéticos En El Tratamiento Del Cáncer de Mama Con Tamoxifeno. Biological Sciences*. Madrid, Spain: ADEMAS Com, FundaciÓn Tejerina. Available at: http://books.google.com/books?hl=en& lr=& id=GanArxPeUeAC& oi=fnd& pg=PA5& dq=Aspectos+farmacogenéticos+en+el+tratamiento+del+cáncer+de+mama+con+Tamoxifeno& ots=P2jmC77GWg& sig=GWxp_taJbt5AJiCdE3MHLyfUZUg.

Fisher, B., Costantino, J. P., Wickerham, D. L., Redmond, C. K., Kavanah, M., Cronin, W. M., et al. (1998). Tamoxifen for Prevention of Breast Cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J. Natl. Cancer Inst.* 90 (18), 1371–1388. doi:10.1093/jnci/90.18.1371

Flicek, P., Amode, M. R., Barrell, D., Beal, K., Brent, S., Chen, Y., et al. (2011). Ensembl 2011. *Nucleic Acids Res.* 39 (Database issue), D800–D806. doi:10.1093/nar/gkq1064

Friedrich, D. C., Genro, J. P., Sortica, V. A., Suarez-Kurtz, G., de Moraes, M. E., Pena, S. D., et al. (2014). Distribution of CYP2D6 Alleles and Phenotypes in the Brazilian Population. *PLoS One* 9 (10), e110691. doi:10.1371/journal.pone.0110691

Fuentes, N., and Silveyra, P. (2019). Estrogen Receptor Signaling Mechanisms. *Adv. Protein Chem. Struct. Biol.* 116, 135–170. doi:10.1016/bs.apcsb.2019.01.001

Gjerde, J., Hauglid, M., Breilid, H., Lundgren, S., Varhaug, J. E., Kisanga, E. R., et al. (2007). Effects of CYP2D6 and SULT1A1 Genotypes Including SULT1A1 Gene Copy Number on Tamoxifen Metabolism. *Ann. Oncol.* 19 (1), 56–61. doi:10.1093/annonc/mdm434

Goetz, M. P., Rae, J. M., Suman, V. J., Safgren, S. L., Ames, M. M., Visscher, D. W., et al. (2005). Pharmacogenetics of Tamoxifen Biotransformation Is Associated with Clinical Outcomes of Efficacy and Hot Flashes. *J. Clin. Oncol.* 23 (36), 9312–9318. doi:10.1200/JCO.2005.03.3266

Goetz, M. P., Sangkuhl, K., Guchelaar, H. J., Schwab, M., Province, M., Whirl-Carrillo, M., et al. (2018). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2D6 and Tamoxifen Therapy. *Clin. Pharmacol. Ther.* 103 (5), 770–777. doi:10.1002/cpt.1007

Hajdinjak, T., and Zagradišnik, B. (2004). Prostate Cancer and Polymorphism D85Y in Gene for Dihydrotestosterone Degrading Enzyme UGT2B15: Frequency of DD Homozygotes Increases with Gleason Score. *Prostate* 59 (4), 436–439. doi:10.1002/pros.20024

Hertz, D. L., McLeod, H. L., and Hoskins, J. M. (2009). Pharmacogenetics of Breast Cancer Therapies. *Breast* 18 (Suppl. 3), S59–S63. doi:10.1016/S0960-9776(09)70275-9

Herynk, M. H., and Fuqua, S. A. (2004). Estrogen Receptor Mutations in Human Disease. *Endocr. Rev.* 25 (6), 869–898. doi:10.1210/er.2003-0010

Ince, B. A., Schodin, D. J., Shapiro, D. J., and Katzenellenbogen, B. S. (1995). Repression of Endogenous Estrogen Receptor Activity in MCF-7 Human Breast Cancer Cells by Dominant Negative Estrogen Receptors. *Endocrinology* 136 (8), 3194–3199. doi:10.1210/endo.136.8.7628351

Jin, Y., Desta, Z., Stearns, V., Ward, B., Ho, H., Lee, K. H., et al. (2005). CYP2D6 Genotype, Antidepressant Use, and Tamoxifen Metabolism during Adjuvant Breast Cancer Treatment. *J. Natl. Cancer Inst.* 97 (January), 30–39. doi:10.1093/jnci/dji005

Johnson, M. D., Zuo, H., Lee, K. H., Trebley, J. P., Rae, J. M., Weatherman, R. V., et al. (2004). Pharmacological Characterization of 4-Hydroxy-N-Desmethyl Tamoxifen, a Novel Active Metabolite of Tamoxifen. *Breast Cancer Res. Treat.* 85 (2), 151–159. doi:10.1023/B:BREA.0000025406.31193.e8

Jordan, V. C., Collins, M. M., Rowsby, L., and Prestwich, G. (1977). A Monohydroxylated Metabolite of Tamoxifen with Potent Antioestrogenic Activity. *J. Endocrinol.* 75 (2), 305–316. doi:10.1677/joe.0.0750305

Kagaya, H., Inoue, K., Miura, M., Satoh, S., Saito, M., Tada, H., et al. (2007). Influence of UGT1A8 and UGT2B7 Genetic Polymorphisms on Mycophenolic Acid Pharmacokinetics in Japanese Renal Transplant Recipients. *Eur. J. Clin. Pharmacol.* 63 (3), 279–288. doi:10.1007/s00228-006-0248-2

Kisanga, E. R., Gjerde, J., Guerrieri-Gonzaga, A., Pigatto, F., Pesci-Feltri, A., Robertson, C., et al. (2004). Tamoxifen and Metabolite Concentrations in Serum and Breast Cancer Tissue during Three Dose Regimens in a Randomized Preoperative Trial. *Clin. Cancer Res.* 10 (7), 2336–2343. doi:10.1158/1078-0432.ccr-03-0538

Kiyotani, K., Mushiroda, T., Imamura, C. K., Hosono, N., Tsunoda, T., Kubo, M., et al. (2010). Significant Effect of Polymorphisms in CYP2D6 and ABCC2 on Clinical Outcomes of Adjuvant Tamoxifen Therapy for Breast Cancer Patients. *J. Clin. Oncol.* 28 (8), 1287–1293. doi:10.1200/JCO.2009.25.7246

Lash, T. L., Cronin-Fenton, D., Ahern, T. P., Rosenberg, C. L., Lunetta, K. L., Silliman, R. A., et al. (2011). CYP2D6 Inhibition and Breast Cancer Recurrence in a Population-Based Study in Denmark. *J. Natl. Cancer Inst.* 103 (6), 489–500. doi:10.1093/jnci/djr010

Lee, S. J., Bell, D. A., Coulter, S. J., Ghanayem, B., and Goldstein, J. A. (2005). Recombinant CYP3A4*17 Is Defective in Metabolizing the Hypertensive Drug Nifedipine, and the CYP3A4*17 Allele May Occur on the Same Chromosome as CYP3A5*3, Representing a New Putative Defective CYP3A Haplotype. *J. Pharmacol. Exp. Ther.* 313 (1), 302–309. doi:10.1124/jpet.104.078758

Lévesque, E., Beaulieu, M., Green, M. D., Tephly, T. R., Bélanger, A., and Hum, D. W. (1997). Isolation and Characterization of UGT2B15(Y85): A UDP-Glucuronosyltransferase Encoded by a Polymorphic Gene. *Pharmacogenetics* 7 (4), 317–325. doi:10.1097/00008571-199708000-00007

Lewis, J. S., and Jordan, V. C. (2005). Selective Estrogen Receptor Modulators (SERMs): Mechanisms of Anticarcinogenesis and Drug Resistance. *Mutat. Res.* 591 (1–2), 247–263. doi:10.1016/j.mrfmmm.2005.02.028

Lien, E. A., Johannessen, D. C., Aakvaag, A., and Lønning, P. E. (1992). Influence of Tamoxifen, Aminoglutethimide and Goserelin on Human Plasma IGF-I Levels in Breast Cancer Patients. *J. Steroid Biochem. Mol. Biol.* 41 (3–8), 541–543. doi:10.1016/0960-0760(92)90380-2

Lim, J. S., Chen, X. A., Singh, O., Yap, Y. S., Ng, R. C., Wong, N. S., et al. (2011). Impact of CYP2D6, CYP3A5, CYP2C9 and CYP2C19 Polymorphisms on Tamoxifen Pharmacokinetics in Asian Breast Cancer Patients. *Br. J. Clin. Pharmacol.* 71 (5), 737–750. doi:10.1111/j.1365-2125.2011.03905.x

López, M., Guerrero, J., Jung-Cook, H., and Alonso, M. E. (2005). CYP2D6 Genotype and Phenotype Determination in a Mexican Mestizo Population. *Eur. J. Clin. Pharmacol.* 61 (10), 749–754. doi:10.1007/s00228-005-0038-2

Lu, S., Nand, R. A., Yang, J. S., Chen, G., and Gross, A. S. (2018). Pharmacokinetics of CYP2C9, CYP2C19, and CYP2D6 Substrates in Healthy Chinese and European Subjects. *Eur. J. Clin. Pharmacol.* 74, 285–296. doi:10.1007/s00228-017-2375-3

Ma, Q., and Lu, A. Y. (2011). Pharmacogenetics, Pharmacogenomics, and Individualized Medicine. *Pharmacol. Rev.* 63 (2), 437–459. doi:10.1124/pr.110.003533

Madlensky, L., Natarajan, L. S. T., Tchu, S., Pu, M. S. F., Mortimer, J., Flatt, S. W., et al. (2011). Tamoxifen Metabolite Concentrations, CYP2D6 Genotype, and Breast Cancer Outcomes. *Clin. Pharmacol. Ther.* 89 (5), 718–725. doi:10.1038/clpt.2011.32

Maughan, K. L., Lutterbie, M. A., and Ham, P. S. (2010). Treatment of Breast Cancer. *Am. Fam. Physician* 81 (June), 1339–1346.

McDonnell, D. P., and Wardell, S. E. (2010). The Molecular Mechanisms Underlying the Pharmacological Actions of ER Modulators: Implications for New Drug Discovery in Breast Cancer. *Curr. Opin. Pharmacol.* 10 (6), 620–628. doi:10.1016/j.coph.2010.09.007

McInerney, E. M., Ince, B. A., Shapiro, D. J., and Katzenellenbogen, B. S. (1996). A Transcriptionally Active Estrogen Receptor Mutant Is a Novel Type of Dominant Negative Inhibitor of Estrogen Action. *Mol. Endocrinol.* 10 (12), 1519–1526. doi:10.1210/mend.10.12.8961262

Mendoza, R., Wan, Y. J., Poland, R. E., Smith, M., Zheng, Y., Berman, N., et al. (2001). CYP2D6 Polymorphism in a Mexican American Population. *Clin. Pharmacol. Ther.* 70 (6), 552–560. doi:10.1067/mcp.2001.120675

MINSAL (2005). *Guía Clínica Cáncer de Mama En Personas de 15 Años y Más*. 1st Ed. Santiago: Ministerio de Salud de Chile.

Mürdter, T. E., Schroth, W., Bacchus-Gerybadze, L., Winter, S., Heinkele, G., Simon, W., et al. (2011). Activity Levels of Tamoxifen Metabolites at the Estrogen Receptor and the Impact of Genetic Polymorphisms of Phase I and II Enzymes on Their Concentration Levels in Plasma. *Clin. Pharmacol. Ther.* 89 (5), 708–717. doi:10.1038/clpt.2011.27

Namipashaki, A., Razaghi-Moghadam, Z., and Ansari-Pour, N. (2015). The Essentiality of Reporting Hardy-Weinberg Equilibrium Calculations in Population-Based Genetic Association Studies. *Cel J.* 17, 187. doi:10.22074/cellj.2016.3711

NCI (2012). Terapia Hormonal Para El Cáncer de Seno. Available at: http://www.cancer.gov/espanol/recursos/hojas-informativas/tratamiento/terapia-hormonal-seno (Accessed June 20, 2021).

Nishiyama, T., Ogura, K., Nakano, H., Ohnuma, T., Kaku, T., Hiratsuka, A., et al. (2002). Reverse Geometrical Selectivity in Glucuronidation and Sulfation of Cis- and Trans-4-hydroxytamoxifens by Human Liver UDP-Glucuronosyltransferases and Sulfotransferases. *Biochem. Pharmacol.* 63 (10), 1817–1830. doi:10.1016/s0006-2952(02)00994-2

Nowell, S., Sweeney, C., Winters, M., Stone, A., Lang, N. P., Hutchins, L. F., et al. (2002). Association between Sulfotransferase 1A1 Genotype and Survival of Breast Cancer Patients Receiving Tamoxifen Therapy. *J. Natl. Cancer Inst.* 94 (21), 1635–1640. doi:10.1093/jnci/94.21.1635

Nowell, S. A., Ahn, J., Rae, J. M., Scheys, J. O., Trovato, A., Sweeney, C., et al. (2005). Association of Genetic Variation in Tamoxifen-Metabolizing Enzymes with Overall Survival and Recurrence of Disease in Breast Cancer Patients. *Breast Cancer Res. Treat.* 91 (3), 249–258. doi:10.1007/s10549-004-7751-x

Núñez, E., Steyerberg, E. W., and Núñez, J. (2011). Regression Modeling Strategies. *Rev. Esp Cardiol.* 64 (6), 501–507. doi:10.1016/j.recesp.2011.01.019

Province, M. A., Goetz, M. P., Brauch, H., Flockhart, D. A., Hebert, J. M., Whaley, R., et al. (2014). CYP2D6 Genotype and Adjuvant Tamoxifen: Meta-Analysis of Heterogeneous Study Populations. *Clin. Pharmacol. Ther.* 95 (2), 216–227. doi:10.1038/clpt.2013.186

Rae, J. M., Drury, S., Hayes, D. F., Stearns, V., Thibert, J. N., Haynes, B. P., et al. (2012). CYP2D6 and UGT2B7 Genotype and Risk of Recurrence in Tamoxifen-Treated Breast Cancer Patients. *J. Natl. Cancer Inst.* 104 (6), 452–460. doi:10.1093/jnci/djs126

Ramón y Cajal, T., Altés, A., Paré, L., del Rio, E., Alonso, C., Barnadas, A., et al. (2010). Impact of CYP2D6 Polymorphisms in Tamoxifen Adjuvant Breast Cancer Treatment. *Breast Cancer Res. Treat.* 119 (1), 33–38. doi:10.1007/s10549-009-0328-y

Kumar, V., Abbas, A. K., and Aster, J. C. (2010). *Robbins y Cotran. Patología estructural y funcional*. 9th Ed. Barcelona: Elsevier España SA.

Roco, A., Quiñones, L., Agúndez, J. A., García-Martín, E., Squicciarini, V., Miranda, C., et al. (2012). Frequencies of 23 Functionally Significant Variant Alleles Related with Metabolism of Antineoplastic Drugs in the Chilean Population: Comparison with Caucasian and Asian Populations. *Front. Genet.* 3, 229. doi:10.3389/fgene.2012.00229

Schroth, W., Antoniadou, L., Fritz, P., Schwab, M., Muerdter, T., Zanger, U. M., et al. (2007). Breast Cancer Treatment Outcome with Adjuvant Tamoxifen Relative to Patient CYP2D6 and CYP2C19 Genotypes. *J. Clin. Oncol.* 25 (33), 5187–5193. doi:10.1200/JCO.2007.12.2705

Schur, B. C., Bjerke, J., Nuwayhid, N., and Wong, S. H. (2001). Genotyping of Cytochrome P450 2D6*3 and *4 Mutations Using Conventional PCR. *Clin. Chim. Acta* 308 (1–2), 25–31. doi:10.1016/s0009-8981(01)00422-3

Sensorn, I., Sirachainan, E., Chamnanphon, M., Pasomsub, E., Trachu, N., Supavilai, P., et al. (2013). Association of CYP3A4/5, ABCB1 and ABCC2 Polymorphisms and Clinical Outcomes of Thai Breast Cancer Patients Treated with Tamoxifen. *Pharmgenomics Pers Med.* 6 (1), 93–98. doi:10.2147/PGPM.S44006

Terán, J., and Teppa, A. (2005). Moduladores Selectivos de Los Receptores Estrogénicos (SERMs): Bioquímica, Farmacología y Aplicación En Guinecología. *Ginecol. Obstet. Mex* 73, 424–435.

Tucker, A. N., Tkaczuk, K. A., Lewis, L. M., Tomic, D., Lim, C. K., and Flaws, J. A. (2005). Polymorphisms in Cytochrome P4503A5 (CYP3A5) May Be Associated with Race and Tumor Characteristics, but Not Metabolism and Side Effects of Tamoxifen in Breast Cancer Patients. *Cancer Lett.* 217 (1), 61–72. doi:10.1016/j.canlet.2004.08.027

Untch, M., and Thomssen, C. (2010). Clinical Practice Decisions in Endocrine Therapy. *Cancer Invest.* 28 (January), 4–13. doi:10.3109/07357907.2010.501637

Valenzuela, C., Acuña, M., and Harb, Z. (1987). Gradiente Sociogenetico En La Población Chilena. *Rev. Med. Chil* 115, 295–299.

Varela, N., Quiñones, L. A., Stojanova, J., Garay, J., Cáceres, D., Cespedes, S., et al. (2015). Characterization of the CYP2D6 Drug Metabolizing Phenotypes of the Chilean Mestizo Population through Polymorphism Analyses. *Pharmacol. Res.* 101, 124–129. doi:10.1016/j.phrs.2015.07.020

Wakeling, A. E., and Slater, S. R. (1980). Estrogen-Receptor Binding and Biologic Activity of Tamoxifen and its Metabolites. *Cancer Treat. Rep.* 64 (6–7), 741–744.

Wegman, P., Vainikka, L., Stål, O., Nordenskjöld, B., Skoog, L., Rutqvist, L. E., et al. (2005). Genotype of Metabolic Enzymes and the Benefit of Tamoxifen in Postmenopausal Breast Cancer Patients. *Breast Cancer Res.* 7 (3), R284–R290. doi:10.1186/bcr993

Whirl-Carrillo, M., McDonagh, E. M., Hebert, J. M., Gong, L., Sangkuhl, K., Thorn, C. F., et al. (2012). Pharmacogenomics Knowledge for Personalized Medicine. *Clin. Pharmacol. Ther.* 92 (4), 414–417. doi:10.1038/clpt.2012.96

WMA (2013). World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *JAMA* 310 (20), 2191–2194. doi:10.1001/jama.2013.281053

Wrenn, C. K., and Katzenellenbogen, B. S. (1993). Structure-Function Analysis of the Hormone Binding Domain of the Human Estrogen Receptor by Region-specific Mutagenesis and Phenotypic Screening in Yeast. *J. Biol. Chem.* 268 (32), 24089–24098. doi:10.1016/s0021-9258(20)80497-9

Zhou, S. F., Di, Y. M., Chan, E., Du, Y. M., Chow, V. D., Xue, C. C., et al. (2008). Clinical Pharmacogenetics and Potential Application in Personalized Medicine. *Curr. Drug Metab.* 9 (8), 738–784. doi:10.2174/138920008786049302

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 06 December 2021
doi: 10.3389/fphar.2021.674117

Effect of *CYP3A4*, *CYP3A5*, *MDR1* and *POR* Genetic Polymorphisms in Immunosuppressive Treatment in Chilean Kidney Transplanted Patients

Stephania Contreras-Castillo[1†], Anita Plaza[2†], Jana Stojanova[3,4], Gustavo Navarro[2], Rodolfo Carmona[2], Fernando Corvalán[1], Leslie Cerpa[1,5], Christopher Sandoval[1], Daniel Muñoz[6], Marina Leiva[2], Luis E. Castañeda[7], Nayaret Farias[8], Carolina Alvarez[8], Gabriel Llull[8], Sergio Mezzano[2], Leopoldo Ardiles[2], Nelson Varela[1,5], María S. Rodríguez[8], Claudio Flores[2], Juan Pablo Cayún[1,5], Paola Krall[2,9]* and Luis A. Quiñones[1,5]**

[1]Laboratory of Chemical Carcinogenesis and Pharmacogenetics (CQF), Department of Basic and Clinical Oncology, Faculty of Medicine, University of Chile, Santiago, Chile, [2]Laboratory of Nephrology, Universidad Austral de Chile, Valdivia, Chile, [3]Interdisciplinary Centre for Health Studies (CIESAL), Universidad de Valparaíso, Valparaíso, Chile, [4]Department of Clinical Pharmacology and Toxicology, St. Vincent's Hospital, Sydney, NSW, Australia, [5]Latin American Network for Implementation and Validation of Clinical Pharmacogenomics Guidelines (RELIVAF-CYTED), Madrid, Spain, [6]Pharmacy Institute, Faculty of Sciences, Universidad Austral de Chile, Valdivia, Chile, [7]Program of Human Genetics, Institute of Biomedical Sciences, Faculty of Medicine, University of Chile, Santiago, Chile, [8]Transplantation Unit, San Juan de Dios Hospital, Santiago, Chile, [9]Departament of Pediatrics and Child Surgery, Faculty of Medicine, University of Chile, Santiago, Chile

Cyclosporine (CsA) and tacrolimus (TAC) are immunosuppressant drugs characterized by a narrow therapeutic range and high pharmacokinetic variability. The effect of polymorphisms in genes related to the metabolism and transport of these drugs, namely *CYP3A4*, *CYP3A5*, *MDR1* and *POR* genes, has been evaluated in diverse populations. However, the impact of these polymorphisms on drug disposition is not well established in Latin American populations. Using *TaqMan*® probes, we determined the allelic frequency of seven variants in *CYP3A4*, *CYP3A5*, *MDR1* and *POR* in 139 Chilean renal transplant recipients, of which 89 were treated with CsA and 50 with TAC. We tested associations between variants and trough and/or 2-hour concentrations, normalized by dose (C_0/D and C_2/D) at specific time points post-transplant. We found that *CYP3A5*3/*3* carriers required lower doses of TAC. In TAC treated patients, most *CYP3A5*3/*3* carriers presented higher C_0/D and a high proportion of patients with C_0 levels outside the therapeutic range relative to other genotypes. These results reinforce the value of considering *CYP3A5* genotypes alongside therapeutic drug monitoring for TAC treated Chilean kidney recipients.

Keywords: polymorphisms, pharmacogenetics, kidney transplant, cyclosporine, tacrolimus

OPEN ACCESS

Edited by:
Amit V. Pandey,
University of Bern, Switzerland

Reviewed by:
Salvador F. Aliño,
University of Valencia, Spain
Umamaheswaran Gurusamy,
University of California San Francisco, United States

***Correspondence:**
Juan Pablo Cayún
jcayun@ug.uchile.cl
Paola Krall
paola.krall@uchile.cl
Luis A. Quiñones
lquinone@uchile.cl

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Pharmacogenetics and Pharmacogenomics, a section of the journal Frontiers in Pharmacology

Received: *28 February 2021*
Accepted: *08 November 2021*
Published: *06 December 2021*

Citation:
Contreras-Castillo S, Plaza A, Stojanova J, Navarro G, Carmona R, Corvalán F, Cerpa L, Sandoval C, Muñoz D, Leiva M, Castañeda LE, Farias N, Alvarez C, Llull G, Mezzano S, Ardiles L, Varela N, Rodríguez MS, Flores C, Cayún JP, Krall P and Quiñones LA (2021) Effect of CYP3A4, CYP3A5, MDR1 *and* POR *Genetic Polymorphisms in Immunosuppressive Treatment in Chilean Kidney Transplanted Patients. Front. Pharmacol. 12:674117. doi: 10.3389/fphar.2021.674117*

INTRODUCTION

Chronic kidney disease is a global public health problem which in 2017 constituted the second and fifth cause of death in Central and Andean Latin America, respectively (Bikbov et al., 2020). These patients require dialysis or kidney transplant as renal replacement therapies. Patients that undergo transplantation require multimodal immunosuppression where a calcineurin inhibitor, either cyclosporine (CsA) or tacrolimus (TAC), is the mainstay of regimens. These drugs exhibit a

narrow therapeutic index: overexposure may result in outcomes such as viral infections, nephrotoxicity and post-transplant diabetes, and underexposure puts patients at risk of allograft rejection (Cotovio et al., 2013; Egeland et al., 2017). Blood concentration levels of calcineurin inhibitors exhibit high inter- and intra-individual variability and associate with clinical outcomes (Webster et al., 2005; Rodrigo et al., 2016). Consequently, dosing is individualized based on measured blood concentrations, or therapeutic drug monitoring.

CsA and TAC are metabolized by the cytochrome P450 enzyme system, primarily the CYP3A4 and CYP3A5 isoforms. Their absorption is influenced by the P-glycoprotein transporter, ABCB1/MDR1. CYP3A4 and CYP3A5 activities are modulated by cytochrome P450 oxidoreductase enzyme, POR (Jiang et al., 2008; Jonge et al., 2011). The genes of these proteins harbor common genetic polymorphisms that are known to influence CsA and TAC disposition, and to affect clinical outcomes (Hesselink et al., 2014; Traynor et al., 2015; Sun et al., 2017; Gelder et al., 2020).

The *CYP3A5*3* (rs776746, c.219-273-1A > G) polymorphism can cause alternative splicing that generates a truncated enzyme with reduced activity. The polymorphic allele (G) is more common than the wild type allele (*A*) in most populations (Kuehl et al., 2001). The impact of this polymorphism on TAC disposition is well established, as evidenced by its inclusion in an international dosing guideline from the Clinical Pharmacogenetics Implementation Consortium (CPIC) (Birdwell et al., 2015). CsA is less significantly affected, likely due to different relative affinities of CYP3A5 for TAC and CsA. *CYP3A4*1B* (rs2740574, c.−392A > G) is purported to cause increased enzymatic activity (Amirimani et al., 2003). The gene encoding P-glycoprotein, *MDR1*, contains several common polymorphisms, however, the impact of individual polymorphisms or common haplotypes has exhibited divergent results (Staatz et al., 2010). *POR*28* C > T at nucleotide 1508 (rs1057868) has exhibited reduced CYP3A5 activity (Elens et al., 2013).

We investigated the association of common polymorphisms in *CYP3A4*, *CYP3A5*, *MDR1* and *POR* genes with blood concentrations and dose requirements in kidney transplant recipients treated with CsA or TAC from two hospitals in Chile. While the impact of these polymorphisms has been evaluated, the populations studied have not been from the Latin American region. Our work adds further support to the existing body of literature and is the first report in a particular admixed Latin American population. In addition, data from observational work adds support for genotyping relevant polymorphisms in these patients in order to more quickly achieve levels within a target range, and keep within range over the course of prolonged treatment.

MATERIALS AND METHODS

Patients and Study Design

We performed two observational studies involving transplant units at two hospitals in Chile, where 85–90% of the transplantations were performed with deceased donors. The first study was of a retrospective cohort design and involved adults treated with CsA, transplanted at San Juan de Dios Hospital (HSJD), Santiago, between 2002 and 2013 (n = 89, CsA cohort). Clinical data for this cohort were collected at weeks 1, 2 and 4 post-transplant. The second study was of a cross-sectional design and involved adults treated with TAC, transplanted at Valdivia Base Hospital (HBVAL) between 2003 and 2018 (n = 50, TAC cohort). Data for this cohort were collected at a single time point, 6 months to 14 years post-transplant.

Inclusion criteria for both studies were kidney transplant, CsA or TAC as the primary immunosuppressant. All TAC patients used diltiazem as a tacrolimus-sparing agent. Patients receiving an alternative primary immunosuppressant after 3 months, or receiving drugs that inhibit or induce CYP3A5 significantly, were excluded. CsA and TAC doses were routinely monitored and adjusted based on C_0 and/or C_2 levels to achieve the respective therapeutic range (CsA C_2 180–250 ng/ml: TAC C_0 5–10 ng/ml). Levels were monitored beyond week 1 post-transplant to ensure steady state had been reached.

Ethics

Ethical approval was obtained from the corresponding Institutional Review Board (Faculty of Medicine of University of Chile and Health Service for HSJD, and the hospital committee and Health Service for HBVAL). This study was performed according to the Declaration of Helsinki, Good Clinical Practice and Chilean Legislation (laws 20.120, 20.584 and 19.628). All patients signed an informed consent form prior to providing a sample for genotyping. Clinical data collection, including blood concentration levels and doses, was retrospective.

Drug Measurement and Quantification

CsA and TAC blood levels were determined by a homogeneous immunoassay at both HSJD and HBVAL (Cobas E-411, Roche Diagnostics). Blood samples were obtained pre-dose for CsA or TAC (C_0) and 2 h after administration for CsA (C_2). CsA and TAC levels were registered and normalized by dose (C_0/D) to perform analyses between genotypes.

Genotype Analyses

Blood samples were collected between 2014 and 2018. For genotyping, 3–5 ml of venous blood was centrifuged for 25 min at 2500xg and 10°C to obtain the buffy coat. DNA was isolated from the peripheral blood mononuclear cells using the High Pure PCR Template Preparation kit (Roche® Diagnostics Gmbh, Mannheim, Germany) or the Whole Blood Genomic DNA Purification kit (ThermoFisher®). Purified DNA was stored at −20°C until use. Commercial *TaqMan*® probes (ThermoFisher®) were used to determine seven polymorphisms. The thermocyclers used were Stratagene mx3000p (Agilent®) or RotorGeneQ (Qiagen®) for HSJD and HBVAL, respectively. We randomly selected 40% of the samples to validate genotype results obtained by the TaqMan assay. We used direct sequencing and/or PCR-RFLP for validation and concordance between assays was 100%. Confirmed samples representing different genotypes (homozygous reference allele,

TABLE 1 | Baseline characteristics of patients.

	Total (*n* = 139)
Cohort	
HBVAL	50 (36.0%)
HSJD	89 (64.0%)
Sex	
Female	67 (48.2%)
Male	72 (51.8%)
Age (years)	
Mean (SD)	41.8 (14.0)
Median [Min, Max]	42.0 [11.0–71.0]
Weight (Kg)	
Mean (SD)	65.1 (11.6)
Median [Min, Max]	65.5 [31.0, 91.0]
Height (mt)	
Mean (SD)	1.62 (0.100)
Median [Min, Max]	1.63 [1.30, 1.86]
BMI (Kg/mt2)	
Mean (SD)	24.6 (3.43)
Median [Min, Max]	24.1 [18.3, 34.2]
Cold ischemia (hr)	
Mean (SD)	20.1 (6.19)
Median [Min, Max]	21.3 [1.00, 33.0]

HBVAL: hospital base valdivia; HSJD: Hospital San Juan de Dios; BMI: body mass index.

TABLE 2 | Genotype frequencies of patients.

	Total (*n* = 139)
*CYP3A4*1B* rs2740574–392A > G	
AA	115 (84%)
AG	22 (16%)
GG	0 (0%)
*CYP3A5*3* rs776746 6986A > G	
AA	10 (7.2%)
AG	48 (35%)
GG	81 (58%)
*CYP3A4*22* rs35599367 C191T	
CC	133 (96%)
CT	5 (3.6%)
TT	0 (0%)
MDR1 3435 rs1045642 C3435T	
CC	46 (33.1%)
CT	65 (46.8%)
TT	28 (20.1%)
MDR1 1236 rs1128503 C1236T	
CC	32 (23.0%)
CT	71 (51.1%)
TT	36 (25.9%)
MDR1 2677 rs2032582 C2677T	
CC	88 (64%)
CT	38 (8.7%)
TT	12 (28%)
*POR*28* rs1057868 1508C > T	
CC	72 (53%)
CT	53 (39%)
TT	11 (8.1%)

heterozygous, homozygous variant allele) were used as internal controls in each assay. Each assay contained four controls: one sample representing each genotype as positive controls (with the exception of very rare genotypes), and one negative sample with nuclease-free pure water to volume.

Statistical Analyses

Analyses were performed using R version 4.02 (R studio version March 1, 1056). C_0 and C_2 levels were adjusted by dose (C_0/D and C_2/D), and dose was normalized by weight. Normality and homoscedasticity were tested for these variables using Shapiro-Wilks and Levene tests, respectively. To meet normality assumptions, square-root transformations were applied to C_0/D and C_2/D, and log10 transformation was applied to TAC-D.

Repeated measures ANOVA was performed to evaluate the differences between genotypes along 1, 2 and 4 weeks after transplantation on continuous variables (C_0/D, Dose and eGFR). Pairwise comparisons between genotypes and weeks were analyzed by t-tests corrected by Bonferroni adjustment. TAC-D was compared between genotypes by ANOVA, and pairwise comparisons between genotypes were analyzed by t-tests corrected by Bonferroni adjustment. We also compared TAC-D between G-carriers (G/G + A/G) and A/A *CYP3A5* genotypes, and between A-carriers (A/A + A/G) and G/G *CYP3A5* genotypes using ANOVA. We also compared TAC-D adjusted by weight between G-carriers and A/A *CYP3A5* genotypes, and between A-carriers and G/G *CYP3A5* genotypes using Kruskal-Wallis tests.

Differences in proportion of patients in the therapeutic range between genotype groups was performed using the Chi-square test or Fisher test. Estimated Glomerular Filtration (eGFR) was calculated using weight, self-reported ethnicity and serum creatinine as variables ("transplantr" package version 2.0). Statistical evaluations were considered in the context of the functional consequence of each polymorphism (increased activity expected, decreased activity expected or controversial consequence). For all analyses, two-sided p-values ≤ 0.05 were considered significant.

RESULTS

Demographic Characteristics of Patients

We studied 139 adult Chilean kidney transplant patients receiving CsA (64%) or TAC (36%). Demographic and clinical data are shown in **Table 1**. Most patients had undergone their first transplant (97.8% for the CsA-cohort and 91.1% for the TAC-cohort). A majority received organs from deceased donors (98.9% for CsA and 86.5% for TAC). The etiology of renal failure was diverse, and for most patients it was unknown or unregistered.

All patients were genotyped for *CYP3A4*1B, CYP3A5*3, CYP3A4*22, MDR1 1236 C > T, MDR1 2677 A > T, MDR1 3435 C > T* and *POR*28* polymorphisms. Genotype frequencies are shown in **Table 2**. None of the genotypes exhibited deviation from Hardy-Weinberg equilibrium.

Genetic Association Between Polymorphisms and Drug Levels

We tested associations between polymorphisms and drug levels normalized by dose. Data from CsA treated patients were

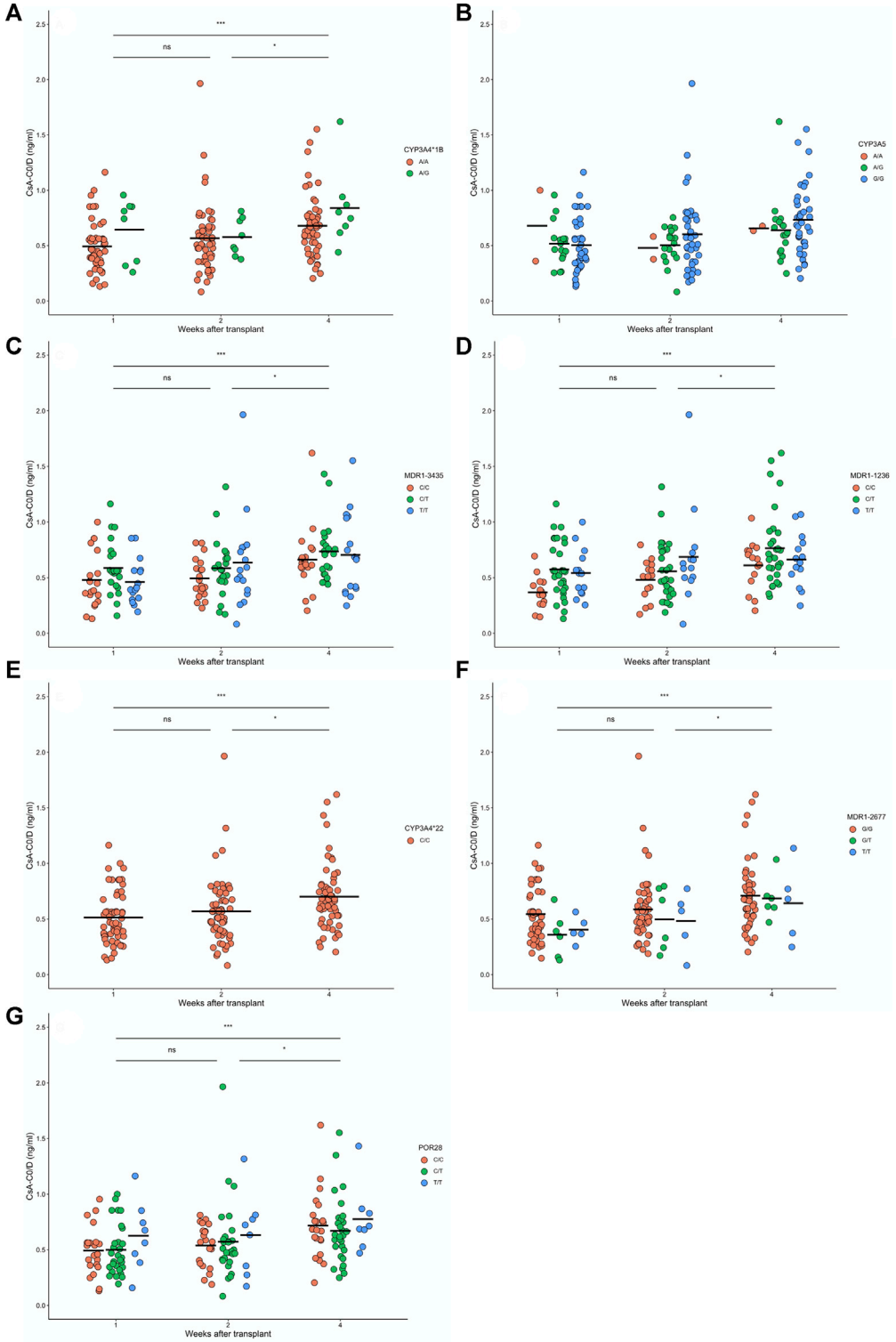


FIGURE 1 | Levels of cyclosporine (CsA) normalized by dose before dose administration (CsA-C0/D) after 1, 2 and 4 weeks of renal transplant for 139 patients with different **(A)** CYP3A4*1B, **(B)** CYP3A5, **(C)** MDR1-3435, **(D)** MDR1-1236, **(E)** CYP3A4*22, **(F)** MDR1-2677, and **(G)** POR28 genotypes. For CsA-C0/D, non-differences were found between genotypes, but significance differences were found along time after transplantation (see **Supplementary Table S1**). Paired week comparisons using Bonferroni-adjusted t- tests are indicated as non-significant (ns: $p > 0.05$) or significant (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$).

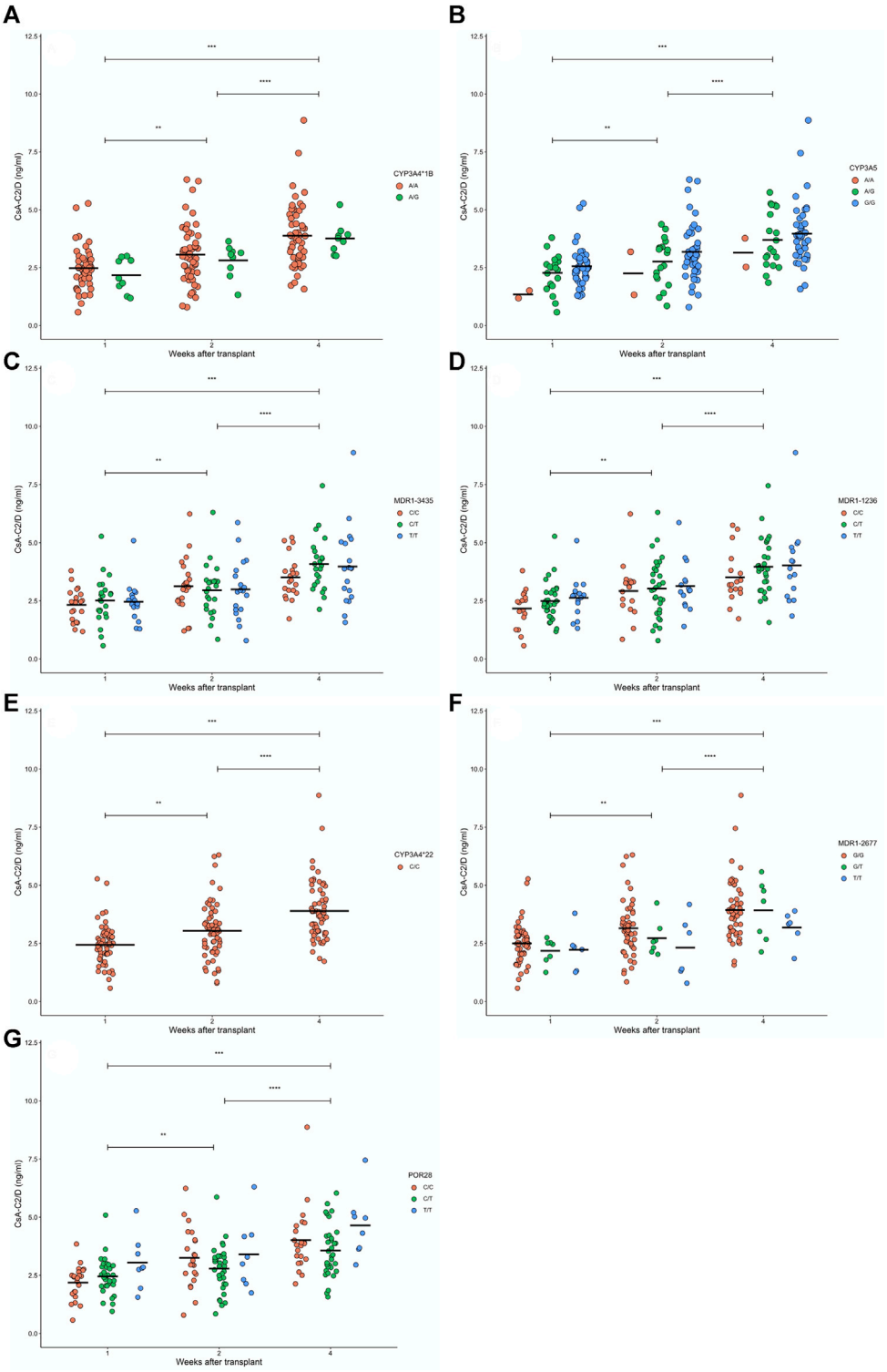


FIGURE 2 | Levels of cyclosporine (CsA) normalized by dose after 2 h of administration (CsA-C0/D) after 1, 2 and 4 weeks of renal transplant for 139 patients with different **(A)** CYP3A4*1B, **(B)** CYP3A5, **(C)** MDR1-3435, **(D)** MDR1-1236, **(E)** CYP3A4*22, **(F)** MDR1-2677, and **(G)** POR28 genotypes. For CsA-C0/D, non-differences were found between genotypes, but significance differences were found along time after transplantation (see **Supplementary Table S1**). Paired week comparisons using Bonferroni-adjusted t- tests are indicated as non-significant (ns: $p > 0.05$) or significant (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$).

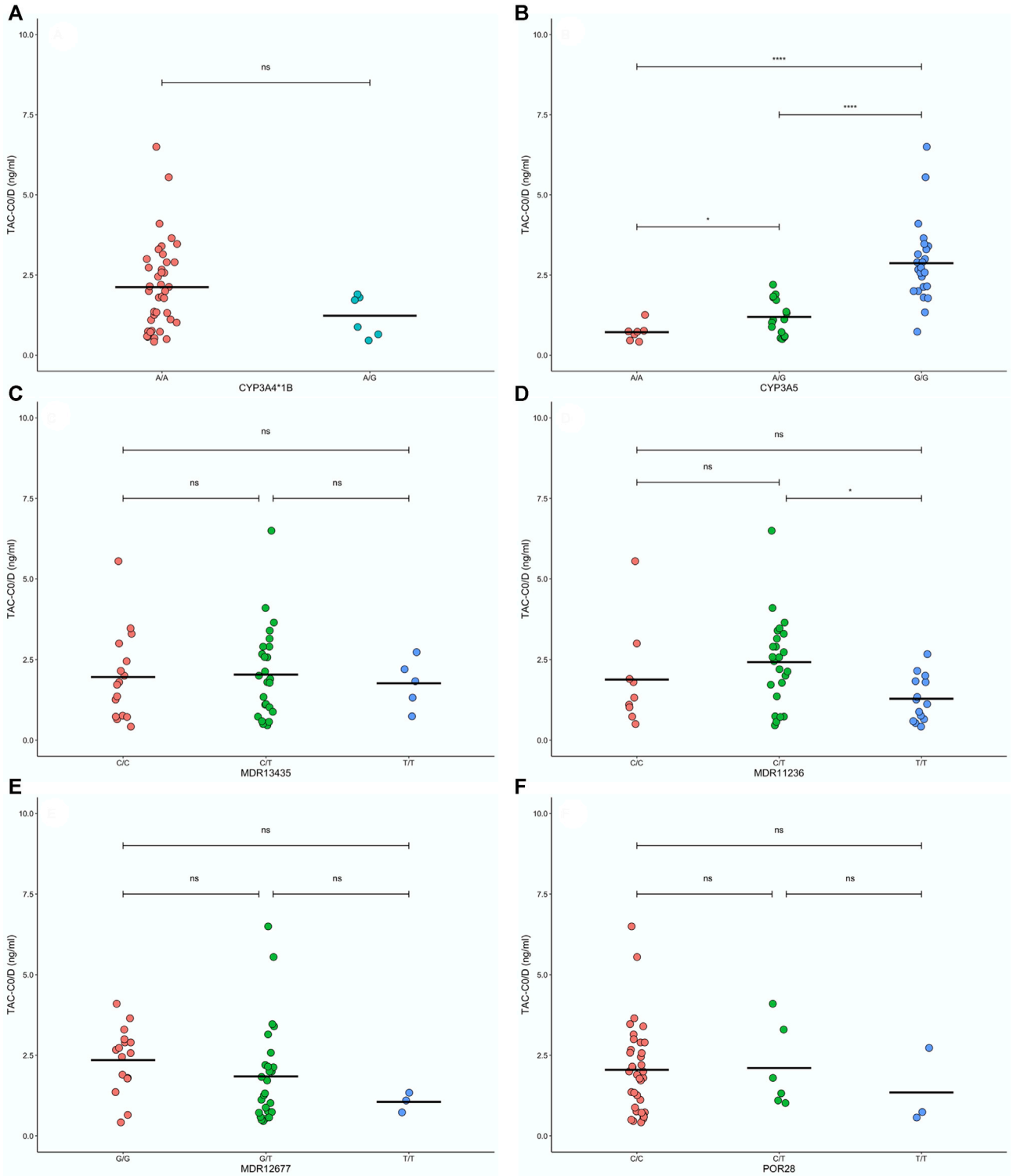


FIGURE 3 | Levels of tacrolymus (TAC) normalized by dose before administration (TACC0/D) for 50 patients with different **(A)** CYP3A4*1B, **(B)** CYP3A5, **(C)** MDR1-3435, **(D)** MDR1-1236, **(E)** MDR1-2677, and **(F)** POR28 genotypes. Paired comparisons between genotypes using Bonferroni-adjusted t- tests are they results are indicated as nonsignificant (ns: $p > 0.05$) or significant (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$).

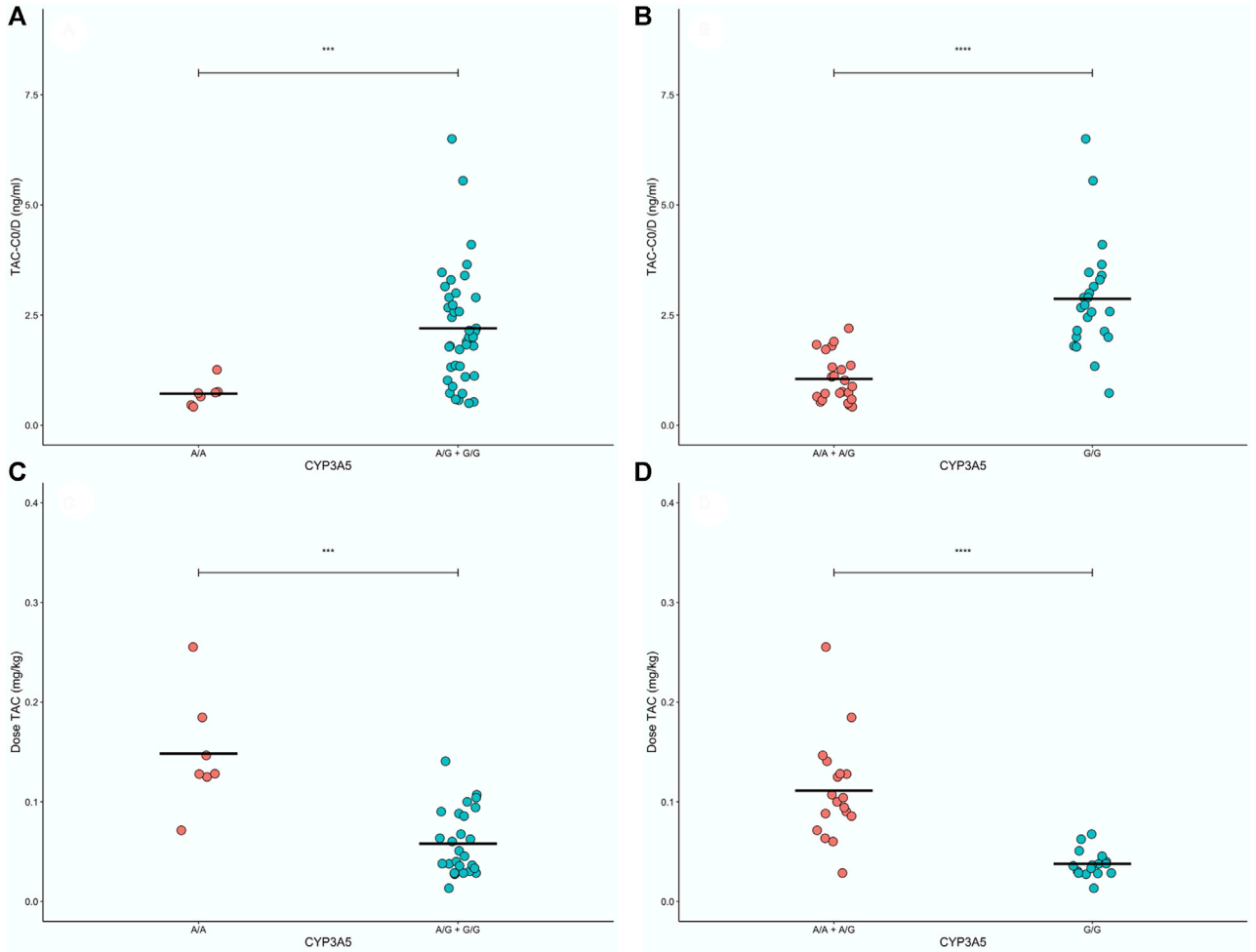


FIGURE 4 | Levels of tacrolimus (TAC) normalized by dose before administration (TACC0/D) and tacrolimus dose standardized by patient weight according to dominant **(A,C)** and recessive **(B,D)** modes for CYP3A5 gene, respectively. ANOVA results are indicated as significant differences between grouped genotypes (***$p <$ 0.001, and ****$p < 0.0001$).

obtained at week 1, 2 and 4 after transplantation. We tested associations between genotypes and C_0/D (**Figure 1**) and C_2/D (**Figure 2**), at the different weeks of follow-up. We also tested combinations of *CYP3A* polymorphisms for associations, according to previous reports (Staatz et al., 2010; Okubo et al., 2013). We defined two groups: 1) low enzymatic activity: *CYP3A4* (**1A/*1A*), *CYP3A4* (**22/*22*), *CYP3A5* (**3/*3*); and 2) intermediate or high enzymatic activity: *CYP3A4* (**1A/*1B*) and *CYP3A5* (**1/*1 or *1/*3*) (data not shown). Carriers of the *MDR1 1236 C/C* genotype exhibited lower mean C_0/D at week 1 compared to carriers of C/T (**Figure 1D**). No other genotype or genotype combination exhibited associations with C_0/D or C_2/D.

For TAC treated patients, we tested associations between genotypes and C_0/D 6 months to 14 years post-transplant (**Figure 3**). *CYP3A5*3 A/A* carriers and *A/G* carriers each exhibited lower mean C_0/D levels compared to *G/G* carriers (**Figure 3B**, $p < 0.05$ for both). Carriers of *MDR1 1236 T/T* exhibited lower mean C_0/D levels compared to *C/T* carriers ($p = 0.012$, **Figure 3D**). No other genotype exhibited associations with C_0/D.

We explored inheritance models for *CYP3A5*3*, *MDR1 1236* and *2677*. For *CYP3A5*3* the strongest model suggested dominant inheritance, where *G/G* carriers exhibited 3-fold lower mean C0/D levels than *A/A* + *A/G* carriers (**Figure 4**). *MDR1 1236 C/C* + *C/T* carriers exhibited higher mean C_0/D levels compared to *T/T* carriers, suggesting a recessive model ($p = 0.026$, **Supplementary Figures S1A,B**). *MDR1 2677 C/C* carriers exhibited higher mean C_0/D compared to *C/T* + *T/T* carriers, suggesting a dominant mode ($p = 0.017$, **Supplementary Figures S2A,B**). The most common haplotype for *MDR1* is *TTT* (carriage of the polymorphic base T at *1236, 2677* and *3435*). We did not find a difference in mean C_0/D between *MDR1 TTT* carriers and those with alternative haplotypes (**Supplementary Figure S3**).

TABLE 3 | Proportion of patients in therapeutic range per genotype.

	CYP3A4*1B rs2740574–392A > G				CYP3A4*22 rs35599367 C191C		
	Wt (AA)	**Het (AG)**	***p*-value**		**Wt (CC)**	**Het (CT)**	***p*-value**
In range	23 (71.8%)	5 (83.3%)	1	In range	28 (68.3%)	2 (100%)	1
Out range	9 (28.2%)	1 (16.7%)		Out range	13 (31.7%)	0 (0%)	
	CYP3A5*3 rs776746 6986 A > G				**CYP3A5*3 rs776746 6986 A > G**		
	Wt (AA)	**Het (AG) + Mut (GG)**	***p*-value**		**Wt (AA) + Het (AG)**	**Mut (GG)**	***p*-value**
In range	5 (83.3%)	25 (67.6%)	0.649	In range	17 (85.0%)	13 (56.5%)	0.053
Out range	1 (16.7%)	12 (32.4%)		Out range	3 (15.0%)	10 (43.5%)	
	CYP3A5*3 rs776746 6986 A > G						
	Wt (AA)	**Het (AG)**	**Mut (GG)**	***p*-value**			
In range	5 (83.3%)	12 (85.7%)	13 (56.5%)	0.168			
Out range	1 (16.6%)	2 (14.3%)	10 (43.5%)				
	POR*28 rs1057868 1508C > T				**POR*28 rs1057868 1508C > T**		
	Wt (CC)	**Het (CT) + Mut (TT)**	***p*-value**		**Wt (CC) + Het (CT)**	**Mut (TT)**	***p*-value**
In range	22 (66.7%)	8 (80%)	0.696	In range	27 (67.5%)	3 (100%)	0.5418
Out range	11 (33.3%)	2 (20%)		Out range	13 (32.5%)	0 (0%)	
	POR*28 rs1057868 1508C > T						
	Wt (CC)	**Het (CT)**	**Mut (TT)**	***p*-value**			
In range	22 (66.7%)	5 (71.4%)	3 (100%)	0.841			
Out range	11 (33.3%)	2 (28.6%)	0 (0%)				
	MDR1 rs1045642 C3435T				**MDR1 rs1045642 C3435T**		
	Wt (CC)	**Het (CT) + Mut (TT)**	***p*-value**		**Wt (CC) + Het (CT)**	**Mut (TT)**	***p*-value**
In range	8 (61.5%)	22 (73.3%)	0.4854	In range	27 (69.2%)	3 (75%)	1
Out range	5 (38.5%)	8 (26.7%)		Out range	12 (30.8%)	1 (25%)	
	MDR1 rs1045642 C3435T						
	Wt (CC)	**Het (CT)**	**Mut (TT)**	***p*-value**			
In range	8 (61.5%)	19 (73.1%)	3 (75.0%)	0.780			
Out range	5 (38.5%)	7 (26.9%)	1 (25.0%)				
	MDR1 rs1128503 C1236T				**MDR1 rs1128503 C1236T**		
	Wt (CC)	**Het (CT) + Mut (TT)**	***p*-value**		**Wt (CC) + Het (CT)**	**Mut (TT)**	***p*-value**
In range	6 (75%)	24 (68.6%)	1 (%)	In range	21 (67.7%)	9 (75%)	0.7272
Out range	2 (25%)	11 (31.4%)		Out range	10 (32.3%)	3 (25%)	
	MDR1 rs1128503 C1236T						
	Wt (CC)	**Het (CT)**	**Mut (TT)**	***p*-value**			
In range	6 (75%)	15 (65.2%)	9 (75%)	0.907			
Out range	2 (25%)	8 (34.8%)	3 (25%)				
	MDR1 2677 rs2032582 C2677T/A				**MDR1 2677 rs2032582 G2677T**		
	Wt (CC)	**Het (CT) + Mut (TT)**	***p*-value**		**Wt (CC) + Het (CT)**	**Mut (TT)**	***p*-value**
In range	12 (66.7%)	18 (72.0%)	0.7466	In range	29 (70.7%)	1 (50%)	0.5183
Out range	6 (33.3%)	7 (28.0%)		Out range	12 (29.2%)	1 (50%)	
	MDR1 2677 rs2032582 G2677T						
	Wt (CC)	**Het (CT)**	**Mut (TT)**	***p*-value**			
In range	12 (66.7%)	17 (73.9%)	1 (50%)	0.646			
Out range	6 (33.3%)	6 (26.1%)	1 (50%)				
	MDR1 TTT haplotype	**MDR1 non-TTT haplotype**	***p*-value**				
In range	10 (66.7%)	20 (71.4%)	0.7422				
Out range	5 (33.3%)	8 (28.6%)					

p-*value: Fisher'exact test.*

Genetic Association Between Polymorphisms and Dose Requirements

For CsA treated patients, we tested associations between genotypes and weight adjusted dose requirement at week 1, 2 and 4 and found no associations (**Supplementary Figures S4A–G**). For TAC treated patients, we found *CYP3A5*3* A/A carriers had the highest weight normalized dose requirement, followed by A/G carriers and G/G carriers (ANOVA $p < 0.05$, **Supplementary Figures S5,S6**).

Genetic Association Between Polymorphisms and Proportion of TAC Patients in Therapeutic Range

TAC treated patients were categorized as being inside or outside the therapeutic range at 6 months post-transplant (C_0 5–10 ng/ml), and we tested the influence of genotypes. *CYP3A5*3* GG carriers were more frequently out of range compared to A/G and A/A carriers combined ($p = 0.053$, **Table 3**). CsA patients were not tested as they receive frequent drug monitoring in the period corresponding to the data available (within 1-month post-transplant).

DISCUSSION

Our aim was to test associations between genetic polymorphisms involved in the absorption and metabolism of CsA and TAC, and drug disposition variables in two groups of Chilean kidney transplant patients, adding to an existing body of work performed in predominantly Caucasian, Asian and African American populations (Staatz et al., 2010). Consistent with existing work, we found *CYP3A5*3/*3* (G/G) carriers required lower doses of TAC, presented higher C_0/D and a higher proportion of patients with C_0 levels outside the therapeutic range relative to other genotypes (Birdwell et al., 2015). Of note, in the TAC group the A allele is more frequent in patients within the therapeutic range, possibly due to a more stable drug disposition when at least one functional allele is expressed. As expected, no significant associations were found for this allele in the CsA cohort, likely due to a lower affinity of CYP3A5 for this drug.

We found *MDR1* 2677 C/C associated with higher TAC trough levels. This is consistent with *MDR1*-2677 C resulting in decreased protein expression and causing reduced drug efflux. *MDR1* 1236 T/T associated with lower TAC levels and higher doses, although the functional consequence of this variant is unclear and divergent results have been published (Staatz et al., 2010). The *MDR1 TTT* haplotype, combining polymorphic variants at positions 1236, 2677 and 3435, did not associate with C_0 or TAC dose. We therefore conclude *MDR1* variants do not significantly affect drug disposition, or may have a minor impact likely to be clinically irrelevant (Phuthong et al., 2017; Robertsen et al., 2018).

Our work has a number of limitations. Firstly, both cohorts received limited follow-up. Further, we did not take into account multi-drug immunosuppressive regimens that may have varied between patients (secondary immunosuppressant and/or corticosteroid). Thirdly, we did not perform a sample size calculation as for a Latin American population the effect size for the impact of each polymorphism on the variables studied, and for each immunosuppressant studied, is uncertain, and the unpredictable frequency of kidney transplants at both hospitals during the enrolment period.

The *CYP3A5* genotyping to guide a tacrolimus starting dose is a clinically relevant practice that is currently undertaken in several centers around the world. It may indeed prove a cost-effective addition to the clinical tool kit to ensure that patients, that already receive therapeutic drug monitoring, achieve levels within the therapeutic range faster. Further, it could flag patients that require closer monitoring to ensure levels within the therapeutic range at 6 months post-transplant and beyond, such as those out of range in our cohort. The ultimate aim is to ensure patients receive adequate immunosuppression, keep their graft patent and avoid adverse effects related to overexposure. Data on cost effectiveness and important clinical endpoints are deficient, especially for lower/middle income countries, including those in Latin America. Further work on these aspects is warranted.

DATA AVAILABILITY STATEMENT

Raw data can be accessed here: https://figshare.com/articles/dataset/Datasets/14128547/1

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Boards (Faculty of Medicine of University of Chile and Health Service for HSJD, and the hospital committee and Health Service for HBVAL). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

SC-C, AP, FC, LCe, GL, and CS performed experiments. SC-C, JC, PK, JS, RC, GN, ML, and CF interpreted data. SC-C, JC and LCa analyzed data. MR, LA, LCa, DM, NF, NV, CA, SM, RC, GN, ML, and CF provided advice with the clinical design, facilitated access to the clinic, provided clinical feedback and discussion. SC-C, JC, PK, JS, and LQ wrote manuscripts. LQ and PK conceived and supervised the project. All authors contributed to editing and reviewing the manuscript.

ACKNOWLEDGMENTS

The authors wish to thank the patients from the San Juan de Dios Hospital (HSJD) and Hospital Base Valdivia (HBVAL) for their altruistic collaboration in pursuit of the common welfare. We also thank the Latin American Society of Pharmacogenomics and Personalized Medicine (SOLFAGEM) for sponsoring this article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.674117/full#supplementary-material

Supplementary Figure S1 | C0 of TAC patients associated to MDR1 1236 polymorphism. **(A)** Co-dominant model. **(B)** Dominant model. **(C)** Recessive model. Significance level: *p*-value < 0.05. See genetic polymorphisms details in **Supplementary Table S1**.

Supplementary Figure S2 | C0 of TAC patients associated to MDR1 2677 polymorphism. **(A)** Co-dominant model. **(B)** Dominant model. **(C)** Recessive model. Significance level: *p*-value < 0.05. See genetic polymorphisms details in **Supplementary Table S1**.

Supplementary Figure S3 | C0 of TAC patients associated to MDR1 TTT haplotype. MDR1 haplotype configured by C3435T, C1236T and C2677T/A. Significance level: *p*-value < 0.05. See genetic polymorphisms details in **Supplementary Table S1**.

Supplementary Figure S4 | TAC-Dose associations to CYP3A5 polymorphism. **(A)** Co-dominant model. **(B)** Dominant model. **(C)** Recessive model. TAC: tacrolimus. Significance level: *p*-value < 0.05. See genetic polymorphisms details in **Supplementary Table S1**.

Supplementary Figure S5 | CsA Dose/weight per genotype. Repeated-measured ANOVA. W1: week 1. W2: week 2. W4: week 4. Significance level: *p*-value < 0.05. See genetic polymorphisms details in **Supplementary Table S1**.

Supplementary Figure S6 | TAC Dose/weight per genotype. ANOVA test and T test. Significance level: *p*-value < 0.05. See genetic polymorphisms details in **Supplementary Table S1**.

Supplementary Figure S7 | TAC-Dose associated to MDR1 TTT haplotype. MDR1 haplotype configured by C3435T, C1236T, and C2677T/A. Significance level: *p*-value < 0.05. See genetic polymorphisms details in **Supplementary Table S1**.

REFERENCES

Amirimani, B., Ning, B., Deitz, A. C., Weber, B. L., Kadlubar, F. F., and Rebbeck, T. R. (2003). Increased Transcriptional Activity of the CYP3A4*1B Promoter Variant. *Environ. Mol. Mutagen* 42 (4), 299–305. doi:10.1002/em.10199

Bikbov, B., Purcell, C. A., Levey, A. S., Smith, M., Abdoli, A., Abebe, M., et al. (2020). Global, Regional, and National Burden of Chronic Kidney Disease, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *The Lancet* 395 (10225), 709–733. doi:10.1016/S0140-6736(20)30045-3

Birdwell, K. A., Decker, B., Barbarino, J. M., Peterson, J. F., Stein, C. M., Sadee, W., et al. (2015). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for CYP3A5 Genotype and Tacrolimus Dosing. *Clin. Pharmacol. Ther.* 98 (1), 19–24. doi:10.1002/cpt.113

Cotovio, P., Neves, M., Rodrigues, L., Alves, R., Bastos, M., Baptista, C., et al. (2013). New-Onset Diabetes after Transplantation: Assessment of Risk Factors and Clinical Outcomes. *Transpl. Proc* 45 (3), 1079–1083. doi:10.1016/j.transproceed.2013.03.009

Egeland, E. J., Robertsen, I., Hermann, M., Midtvedt, K., Størset, E., Gustavsen, M. T., et al. (2017). High Tacrolimus Clearance Is a Risk Factor for Acute Rejection in the Early Phase after Renal Transplantation. *Transplantation* 101 (8), e273–79. doi:10.1097/TP.0000000000001796

Elens, L., Nieuweboer, A. J., Clarke, S. J., Charles, K. A., de Graan, A. J., Haufroid, V., et al. (2013). Impact of POR*28 on the Clinical Pharmacokinetics of CYP3A Phenotyping Probes Midazolam and Erythromycin. *Pharmacogenet Genomics* 23 (3), 148–155. doi:10.1097/FPC.0b013e32835dc113

Gelder, T. V., Meziyerh, S., Swen, J. J., Aiko, P., de Vries, J., and Moes, D. J. A. R. (2020). The Clinical Impact of the C0/D Ratio and the CYP3A5 Genotype on Outcome in Tacrolimus Treated Kidney Transplant Recipients. *Front. Pharmacol.* 11, 1–6. doi:10.3390/genes11101205

Hesselink, D. A., Bouamar, R., Elens, L., van SchaikVan Schaik, R. H., and van Gelder, T. (2014). The Role of Pharmacogenetics in the Disposition of and Response to Tacrolimus in Solid Organ Transplantation. *Clin. Pharmacokinet.* 53 (2), 123–139. doi:10.1007/s40262-013-0120-3

Jiang, Z. P., Wang, Y. R., Xu, P., Liu, R. R., Zhao, X. L., and Chen, F. P. (2008). Meta-Analysis of the Effect of MDR1 C3435T Polymorphism on Cyclosporine Pharmacokinetics. *Basic Clin. Pharmacol. Toxicol.* 103 (5), 433–444. doi:10.1111/j.1742-7843.2008.00300.x

Jonge, H. D., Metalidis, C., Naesens, M., Lambrechts, D., Dirk, R., and Kuypers, J. (2011). The P450 Oxidoreductase *28 SNP Is Associated with Low Initial Tacrolimus Exposure and Increased Dose Requirements in CYP3A5-Expressing Renal Recipients. *Pharmacogenomics* 12 (9), 1281–1291. doi:10.2217/pgs.11.77

Kuehl, P., Zhang, J., Lin, Y., Lamba, J., Assem, M., Schuetz, J., et al. (2001). Sequence Diversity in CYP3A Promoters and Characterization of the Genetic Basis of Polymorphic CYP3A5 Expression. *Nat. Genet.* 27 (4), 383–391. doi:10.1038/86882

Okubo, M., Murayama, N., Shimizu, M., Shimada, T., Guengerich, F. P., and Yamazaki, H. (2013). CYP3A4 Intron 6 C>T Polymorphism (CYP3A4*22) Is Associated with Reduced CYP3A4 Protein Level and Function in Human Liver Microsomes. *J. Toxicol. Sci.* 38 (3), 349–354. doi:10.2131/jts.38.349

Phuthong, S., Settheetham-Ishida, W., Natphopsuk, S., Settheetham, D., and Ishida, T. (2017). Haplotype Analysis of MDR1 and Risk for Cervical Cancer in Northeastern Thailand. *Asian Pac. J. Cancer Prev.* 18 (7), 1815–1819. doi:10.22034/APJCP.2017.18.7.1815

Robertsen, I., Debord, J., Åsberg, A., Marquet, P., and Woillard, J. B. (2018). A Limited Sampling Strategy to Estimate Exposure of Everolimus in Whole Blood and Peripheral Blood Mononuclear Cells in Renal Transplant Recipients Using Population Pharmacokinetic Modeling and Bayesian Estimators. *Clin. Pharmacokinet.* 57, 1459–1469. doi:10.1007/s40262-018-0646-5

Rodrigo, E., Segundo, D. S., Fernández-Fresnedo, G., López-Hoyos, M., Benito, A., Ruiz, J. C., et al. (2016). Within-Patient Variability in Tacrolimus Blood Levels Predicts Kidney Graft Loss and Donor-Specific Antibody Development. *Transplantation* 100 (11), 2479–2485. doi:10.1097/TP.0000000000001040

Staatz, C. E., Goodman, L. K., and Tett, S. E. (2010). Effect of CYP3A and ABCB1 Single Nucleotide Polymorphisms on the Pharmacokinetics and Pharmacodynamics of Calcineurin Inhibitors: Part II. *Clin. Pharmacokinet.* 49 (3), 207–221. doi:10.2165/11317550-000000000-00000

Sun, B., Guo, Y., Gao, J., Shi, W., Fan, G., Li, X., et al. (2017). Influence of CYP3A and ABCB1 Polymorphisms on Cyclosporine Concentrations in Renal Transplant Recipients. *Pharmacogenomics* 18 (16), 1503–1513. doi:10.2217/pgs-2017-0127

Traynor, C., Conlon, P., Phelan, P. J., O'Kelly, P., Elens, L., Mccormack, M., et al. (2015). Association of CYP3A Variants with Kidney Transplant Outcomes. *Ren. Fail.* 37 (4), 562–566. doi:10.3109/0886022X.2015.1007013

Webster, A. C., Woodroffe, R. C., Taylor, R. S., Chapman, J. R., and Craig, J. C. (2005). Tacrolimus Versus Ciclosporin as Primary Immunosuppression for Kidney Transplant Recipients: Meta-Analysis and Meta-Regression of Randomised Trial Data. *BMJ* 331 (7520), 810. doi:10.1136/bmj.38569.471007.AE

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership