frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 13 October 2025
DOI 10.3389/fmolb.2025.1669085

OPEN ACCESS

EDITED BY
Hailin Tang,
Sun Yat-sen University Cancer Center (SYSUCC), China

REVIEWED BY
Hengrui Liu,
University of Cambridge, United Kingdom
Lei Li,
University of South China, China

*CORRESPONDENCE
Brian D. Piening,
brian.piening@providence.org

†These authors have contributed equally to this work

RECEIVED 18 July 2025
ACCEPTED 16 September 2025
PUBLISHED 13 October 2025

CITATION
Wagner JT, Welle JT, Lucas Beckett IA, Emery KR, Cosgrove BA, Olszewski K, Wagner N, Bower TC, Yuan LC, Shull EM, Jade K, Clemens J, Magis AT, Campbell MB, Gordon OK, Bifulco CB and Piening BD (2025) Design and validation of a clinical whole genome sequencing-based assay for patient screening in a large healthcare system. *Front. Mol. Biosci.* 12:1669085. doi: 10.3389/fmolb.2025.1669085

Design and validation of a clinical whole genome sequencing-based assay for patient screening in a large healthcare system

Josiah T. Wagner[1†], John T. Welle[1†], Isabelle A. Lucas Beckett[1,2], Kate R. Emery[1], Benjamin A. Cosgrove[1], Krzysztof Olszewski[3], Nick Wagner[1], Tucker C. Bower[1], Li Chi Yuan[1], Eric M. Shull[1], Kathleen Jade[2], Jon Clemens[1], Andrew T. Magis[2], Mary B. Campbell[1], Ora K. Gordon[1], Carlo B. Bifulco[1,4] and Brian D. Piening[1,4]*

[1]Providence Institute for Clinical Innovation, Providence Health, Portland, OR, United States, [2]Institute For Systems Biology, Seattle, WA, United States, [3]Fabric Genomics, Oakland, CA, United States, [4]Earle A. Chiles Research Institute, Portland, OR, United States

Introduction: Population genetic screening is rapidly emerging as a key methodology in the clinical laboratory for detecting actionable genomic conditions in asymptomatic patients. While current clinical methods are largely focused on targeted gene panels, the increasing efficiency of next-generation sequencing (NGS) platforms permits the use of whole genome sequencing (WGS) for routine clinical applications. The key advantage of WGS is that the complete genome produced by a single sequencing event can form the basis for a patient's genomic health care record for reanalysis throughout a patient's lifetime.

Methods: We developed a scalable clinical WGS-based lab developed procedure (LDP) for heritable disease gene testing and pharmacogenomics (PGx). We performed extensive validation across a range of blood, saliva, and reference specimens.

Results: The clinical deliverable for the WGS LDP was 78 genes associated with actionable genomic conditions and 4 PGx genes. The validation cohort consisted of samples from 188 study participants that were orthogonally sequenced at commercial reference laboratories and additional reference materials. The WGS LDP demonstrated excellent sensitivity, specificity, and accuracy.

Conclusion: The deployed LDP was then used to sequence over 2,000 patients as part of a broader clinical implementation study ("Geno4ME"). Our findings support WGS as a viable method for broad clinical screening.

KEYWORDS

genomic medicine, pharmacogenomics (PGx), clinical germline testing, lab developed procedure, population screening

Introduction

Identification of individuals at risk for heritable genetic conditions or suboptimal drug dosing provides opportunities for life-saving medical intervention and improvement in treatment outcomes. Accumulating population genomics evidence suggests that a significant number of individuals carry a clinically actionable genetic variant that is associated with an increased risk of disease, yet are unaware of their genetic risks until symptomatic (Manickam et al., 2018; Abul-Husn et al., 2016; Win et al., 2017; Haer-Wigman et al., 2019). For genes with a well-established association with inherited cancer, cardiovascular disease, or other condition, published clinical guidelines provide specific disease-risk reduction plans, enhanced screenings for early disease detection, and/or treatments (Musunuru et al., 2020; Sonkin et al., 2024; Daly et al., 2021; Weiss et al., 2021). Additionally, recent insights suggest that up to 98%–99% of individuals have one or more genetic variants that can impact drug efficacy and safety (Dunnenberger et al., 2015). Pharmacogenomics (PGx), which focuses on the identification of genomic variants in an individual that modulate pharmacokinetics and pharmacodynamics of specific classes of drugs, can have direct implications for prescribing guidelines (Relling and Evans, 2015). The high prevalence of individuals with actionable genomic variants has resulted in growing interest in genomic screening tools for broad population health applications (Phillips et al., 2018). This interest, combined with advancements in cost-effective sequencing technologies, have generated population-focused research initiatives in diverse populations as part of the field of genomic medicine. Focusing on mixed populations of healthy patients and patients with cardiovascular disease, ClinSeq and MedSeq were among the first large-scale genome sequencing research studies at the intersection of population health and genomic medicine (Biesecker et al., 2009; Vassy et al., 2014). These landmark studies highlighted the challenges of using large scale sequencing for clinical care, such as variant classification of novel and known variants, and the complexity of returning results to patients and providers (Biesecker, 2012). Importantly, these studies revealed the potential for patient data re-analysis over their lifetime as new phenotypes and genetic understanding become available. Recently, the *All of Us* Research Program has been a leading whole-genome sequencing effort to perform population genomics analysis in unselected populations. However, this nationwide program is a public research program and is disconnected from a participant's primary provider and clinical care. Thus, there remains a need for scalable genomic screening tools with gene panels that can be rapidly expanded, quickly interpreted, and incorporated into a patient's clinical care.

A comprehensive genomic assay evaluable for both heritable genomic conditions and PGx in populations should ideally cover a large and diverse panel of gene loci. While most modern genomic screening tools have been effective at characterizing small genomic variants (approximately less than 50 bp in size), extensive evidence suggests that large copy number variants in the human population are a contributor to heritable disorders (Zhang et al., 2009). Thus, an ideal procedure should be able to characterize multiple variant types, including single-nucleotide variants (SNVs), multi-nucleotide variants/polymorphisms (MNVs), insertions, deletions, and copy-number variants (CNVs). High-throughput methods for simultaneously characterizing single nucleotide variants in multiple genes, such as MALDI-TOF and SNV array-based methods, have been applied in a population health context (Uffelmann et al., 2021; East et al., 2021; Stanssens et al., 2004). To characterize larger variants, microarray-based comparative genomic hybridization (array CGH or aCGH) has been applied clinically (Bejjani and Shaffer, 2006; Park et al., 2011; Wayhelova et al., 2019). Characterization of diverse variant types using these methods can require multiple assays to acquire a comprehensive genomic profile and often there is difficulty with resolving novel and/or complex variants. To address these limitations, whole-exome sequencing (WES) and whole-genome sequencing (WGS) have been applied in a variety of population genomic health applications, such as healthy population screening (Foss et al., 2022; Lindor et al., 2017; Vassy et al., 2017; Williams, 2022), unselected research cohort screening (Manickam et al., 2018; Biesecker et al., 2009; Vassy et al., 2014; Bick et al., 2024; Perkins et al., 2018; Gonzalez-Garay et al., 2013; Carey et al., 2016), and newborn screening (Woerner et al., 2021; Chen et al., 2023; Holm et al., 2018). Both WES and WGS allow for single-nucleotide resolution of variants and have been demonstrated to capture a significant number of gene- and exon-level CNVs with potential clinical importance (Hehir-Kwa et al., 2015; Conrad et al., 2010). WGS, when compared to WES, has reduced variant allele capture bias, can potentially profile more complex chromosomal rearrangements, and can be expanded to noncoding and intergenic regions (Steyaert et al., 2018). Recent advances in PCR-free WGS have led to increased variant detection sensitivity and retention of complex genotypes (e.g., repetitive regions) when compared to conventional WGS (Dolzhenko et al., 2017; Zhou et al., 2022). Therefore, a PCR-free WGS-based assay is ideal for comprehensively evaluating variation in coding regions while providing opportunity for future region expansion as additional loci of interest become evident.

Although using a WGS-based assay as a primary method of diagnosis holds significant promise as a tool for rare diseases and large-scale population health genomics (Bick et al., 2024; Liu et al., 2019), few studies have evaluated the practicality and performance of WGS for returning clinically-actionable results in a healthcare setting. In addition, consistent and accurate classification of variants between reporting laboratories remains a challenge, especially for large population health programs (Harrison et al., 2022). While software tools that can aggregate population, genomic, protein, and disease-specific information to assist in variant classification have been developed (Ravichandran et al., 2019; Preston et al., 2022; Kim et al., 2024; Gall et al., 2022), the degree to which these tools enhance workflow productivity and accuracy is still being established. Here, we developed a clinical PCR-free WGS-based lab developed procedure (LDP) for hereditable disease testing and PGx. We validated our WGS-based assay for variant detection, variant pathogenicity classification, and PGx interpretation against orthogonal panel testing at outside reference laboratories using a large cohort. Additionally, we determined if DNA originating from either blood or saliva specimens impacted WGS assay performance. Our results support the feasibility and high diagnostic accuracy of using a WGS-based assay as a core population health tool for evaluating clinically actionable genomic variants.

Methods

Participant selection and sample collection

This study was reviewed and approved by the Providence Institutional Review Board (approval number STUDY2020000637). The validation of the WGS LDP was performed in support of the "Genomic Medicine for Everyone" (Geno4ME) clinical implementation study, involving WGS of Providence patients, return of selected clinical results, and genome banking for future research (Lucas Beckett et al., 2025). Study participants were patients in the Providence system and informed consent was obtained from all participants using an in-house developed automated electronic consent platform. Information on participants' personal/family history of cancer and/or cardiovascular conditions, race and ethnicity, as well as current medications taken was obtained using a self-reported survey. Whole blood was collected using venipuncture and stabilized in standard clinical use EDTA tubes. Saliva was collected and stored using DNA Genotek Oragene-DNA saliva DNA collection kit (DNA Genotek #OGR-600). In total, 120 whole blood and 70 saliva specimens were collected in the WGS Validation cohort, with 60 participants providing paired whole blood and saliva samples for cross-sample validation (N = 189 unique participants).

Genomic condition gene selection for analysis

For selected genomic conditions, we evaluated variant pathogenicity in 78 genes determined to be clinically actionable and reportable by the American College of Medical Genetics and Genomics (ACMG) as secondary findings in clinical exome and genome sequencing and/or National Comprehensive Cancer Network (NCCN) guidelines (Daly et al., 2021; Kalia et al., 2017). The complete list of genes and genomic conditions analyzed in this study are shown in Figure 1.

Sample genomic DNA extraction, sequencing, QC, and variant calling

Genomic DNA for whole genome sequencing was extracted using the Qiagen QIAsymphony DSP Midi Kit (catalog 937,255). Whole genome next-generation sequencing (NGS) libraries were prepared from 300 to 500 ng gDNA with the Illumina DNA PCR-Free Prep, Tagmentation kit (catalog 20041795). Sequencing was performed on the Illumina NovaSeq 6,000 with 24 libraries loaded per S4 flow cell and a target depth of 30X coverage. As a sequencing quality control, the Illumina PhiX Control v3 Library was sequenced on every WGS with error rates of less than 1% considered passing. In addition, every 100 samples a germline variant quality control sample containing known germline DNA variants was extracted prepared as a WGS library, sequenced, and annotated in the same manner as other patient samples. Intra-run variation (within-run) was determined by preparing and sequencing three WGS replicates of sample gme-wes-25 and sequencing the libraries on the same WGS run. Inter-run (between-run) variation was determined by preparing and sequencing three WGS replicates of DNA from control sample NA12878 and sequencing across three separate WGS sequencing runs. Variant calling and pharmacogenetic genotyping were performed using standard analysis pipelines on the Illumina DRAGEN (Dynamic Read Analysis for GENomics) Bio- IT Platform (v3.9.5) with the following flags: --enable-map-align true, --enable-map-align-output true, --enable-duplicate-marking true, --enable-sort true, --vc-combine-phased-variants-distance 3, --vc-enable-roh true, --vc-enable-baf true, --vc-enable-phasing true, --enable-variant-caller true, --enable-cnv true, --cnv-enable-self-normalization true, --cnv-enable-tracks true, --vc-enable-roh true, --vc-enable-baf true, --vc-enable-phasing true, --enable-variant-caller true, --enable-cnv true, --cnv-enable-self-normalization true, --cnv-enable-tracks true, --cnv-segmentation-mode slm, --cnv-interval-width 250, --enable-sv true. Variants with variant allele frequencies (VAFs) less than 10% were excluded from further analysis. This VAF threshold was chosen based on the estimated minimum supporting variant read depth required at a target coverage of 30X to reliably call a variant. The hg19 (GRCh37) genome build was selected as the reference genome for this study based on the availability of select analysis tools at the time of assay development. Regions for gene variant analysis were extracted using coordinates on the hg19 reference derived from the longest Refseq transcript (O'Leary et al., 2016; Church et al., 2011). These regions for analysis were extended by an additional 1,000 bases upstream and downstream of the first (5′ UTR) and last (3′ UTR) exons, respectively. Exon-level and whole gene-level CNV and structural variation (SV) was detected using the DRAGEN and/or Manta CNV callers included as part of the Illumina DRAGEN Bio-IT Platform. Coverage metrics were calculated using deepTools2 (Ramírez et al., 2016). Reads not mapping to hg19 were discarded from further analysis.

Validation of WGS variant calling method to variants from patient electronic medical records

Initial validation of the WGS DRAGEN variant calling pipeline was performed by comparison of WGS DRAGEN variants to clinically significant genomic variants obtained from patient electronic health records (EHRs). Genomic DNA for WGS was extracted, sequenced, and analyzed as described above for 30 bio-banked patient blood samples collected under IRB protocol STUDY2018000254, an internally funded Providence registry of germline pathogenic/likely pathogenic (P/LP) carriers of hereditary cancer risk. The registry utilizes the Progeny database to curate and maintain highly annotated screening, disease and genetic testing data for participants in the biorepository. Sample selection was drawn from the Progeny database of de-identified samples. Participants provided DNA samples for future use in discovery and all available EHR genomic testing information (EHR Comparison validation group). Samples were selected based on variety of genes to be within the Geno4ME suite broad variant profile (truncating, missense, indel and genomic rearrangements, with available EHR genomic testing information [EHR Comparison validation group]). A positive variant match was determined by the presence of an expected gene coding and/or protein change as described in the patient EHR to a variant in the WGS DRAGEN pipeline that passed all quality filters.

10.3389/fmolb.2025.1669085

CANCER

Gene	Disease
APC	Familial adenomatous polyposis
ATM	Ataxia-telangiectasia (includes ATM-related cancers)
AXIN2	Oligodontia-colorectal cancer syndrome
BMPR1A	Juvenile polyposis
BRCA1	Hereditary breast and ovarian cancer syndrome
BRCA2	
BRIP1	BRIP1-related cancers
CDH1	Hereditary diffuse gastric cancer syndrome
CDK4	Hereditary cutaneous melanoma
CDKN2A	Hereditary melanoma-pancreatic cancer syndrome
CHEK2	CHEK2-related cancers including breast, colon and other sites
EPCAM	Lynch syndrome (hereditary nonpolyposis colorectal cancer syndrome)
GREM1	Hereditary mixed polyposis syndrome
MEN1	Multiple endocrine neoplasia type 1
MLH1	Lynch syndrome (hereditary nonpolyposis colorectal cancer syndrome)
MSH2	
MSH3	MSH3-associated polyposis
MSH6	Lynch syndrome (hereditary nonpolyposis colorectal cancer syndrome)
MUTYH	MYH-associated polyposis; adenomas, multiple colorectal, FAP type 2; colorectal adenomatous polyposis, autosomal recessive, with pilomatricomas
NBN	NBN-related cancers
NF1	Neurofibromatosis type 1
NF2	Neurofibromatosis type 2
NTHL1	NTHL1-associated polyposis
PALB2	PALB2-related cancers
PMS2	Lynch syndrome (hereditary nonpolyposis colorectal cancer syndrome)
POLD1	POLD1-related cancers
POLE	POLE-related cancers
PTEN	PTEN hamartoma tumor syndrome
RAD51C	RAD51C-related cancers
RAD51D	RAD51D-related cancers
RB1	Retinoblastoma
RET	Multiple endocrine neoplasia type 2
RET	Familial medullary thyroid cancer
SDHAF2	Hereditary paraganglioma-pheochromocytoma syndrome
SDHB	
SDHC	
SDHD	
SMAD4	Juvenile polyposis and Hereditary hemorrhagic telangiectasia
STK11	Peutz-Jeghers syndrome
TP53	Li-Fraumeni syndrome
TSC1	Tuberous sclerosis complex
TSC2	
VHL	Von Hippel-Lindau syndrome
WT1	WT1-related Wilms tumor

CARDIOVASCULAR AND CONNECTIVE TISSUE

Disease		Gene
Hyperlipidemia	Familial hypercholesterolemia	LDLR
		PCSK9
	Familial hypercholesterolemia and Familial hypobetalipoproteinemia	APOB
Cardiomyopathy	Arrhythmogenic right ventricular cardiomyopathy	DSC2
		DSG2
		DSP
		PKP2
		TMEM43
	Hypertrophic cardiomyopathy, dilated cardiomyopathy	ACTC1
		GLA
		LMNA
		MYBPC3
		MYH7
		MYL2
		MYL3
		PRKAG2
		TNNI3
		TNNT2
		TPM1
Arrhythmia	Brugada syndrome	SCN5A
	Catecholaminergic polymorphic ventricular tachycardia	RYR2
	Romano-Ward long-QT syndrome types 1, 2, and 3	SCN5A
		KCNH2
		KCNQ1
Aneurysm/Connective tissue	Ehlers-Danlos syndrome, vascular type	COL3A1
	Familial thoracic aortic aneurysms and dissections (FTAAD)	ACTA2
		MYH11
	Loeys-Dietz syndrome	SMAD3
		TGFBR1
		TGFBR2
	Marfan syndrome	FBN1
		TGFBR1

OTHER

Disease		Gene
	Malignant hyperthermia susceptibility	CACNA1S
		RYR1
	Ornithine transcarbamylase deficiency	OTC
	Wilson disease	ATP7B

FIGURE 1
Genes tested for the Geno4ME LDP heritable disorder panel and their associated diseases.

Classification of variants

To classify variants in genes with an associated genetic condition, pathogenicity was assessed based on criteria from the joint American College of Medical Genetics and Genomics and Association for Molecular Pathology Standards and Guidelines (ACMG/AMP criteria from 2015) (Richards et al., 2015). PP5 or BP6 criteria were not considered for final variant pathogenicity classification as recommended by Biesecker et al. (2018). Prior to manual variant curation, variants were initially classified using the

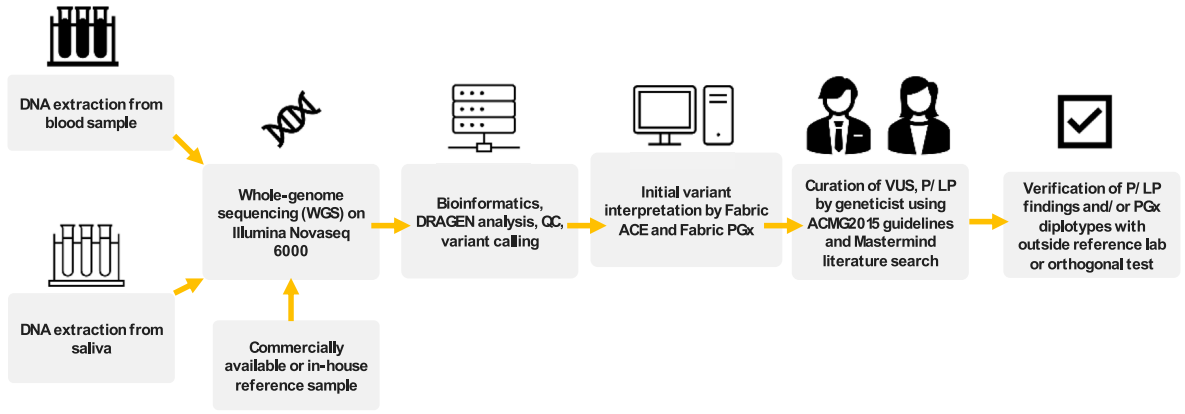

FIGURE 2
Overview of workflow for WGS sample processing, interpretation, and comparison to outside references for assay validation. This complete workflow is referred to as Geno4ME LDP.

Artificial Intelligence Classification Engine (ACE), an automated ACMG classification algorithm in the Fabric Enterprise platform (Version 6.18.X) (De La Vega et al., 2021). All variants, including CNVs and SVs, were uploaded to the Fabric Enterprise platform for automated interpretation prior to manual curation or interpretation. Variants were targeted for manual curation if initially classified as P/LP by ACE and/or the interpretation for the associated genetic condition in the ClinVar genomic annotation aggregation database was P, LP, conflicting, or not provided (Landrum et al., 2014). Evaluation of variant population frequency, *in silico* predictions of variant effect, and statistical support for pathogenicity using the ACMG/AMP criteria were performed using tools in the Fabric Enterprise platform. Estimation of population variant allele frequency was based on the gnomAD database (Chen et al., 2024). Literature review for variants identified for manual curation was assisted by Mastermind Genomic Search Engine (Chunn et al., 2020). For putative variants of uncertain significance (VUS) and P/LP variants, alignment quality of the region was manually inspected using Integrative Genomics Viewer (IGV) for read and alignment quality (Robinson et al., 2011). The overall workflow for WGS and interpretation we termed the Geno4ME LDP as is henceforth referred to below. The Geno4ME LDP method was used for comparison to all reference methods (Figure 2).

Geno4ME LDP orthogonal validation with outside reference methodologies

Sensitivity, specificity, and accuracy of the Geno4ME LDP for variant calling was determined by comparison to a CLIA-certified commercial molecular laboratory using an outside reference method (OS-ORM). The OS-ORM was based on hybrid-capture NGS for all 78 genes associated with a genomic condition surveyed in this study were covered by OS-ORM Panel A, OS-ORM Panel B, or both (Supplementary Table S1). Because not all VUS were reported by the outside provider due to differences in outside provider classification or ORM panel return of results (RoR) criteria, only variants identified by the OS-ORM to be P/LP were considered to be true positives for comparison. For genes where a different transcript was selected between the Geno4ME LDP and ORM, gene variants were remapped to the Geno4ME LDP selected transcript (Supplementary Table S2). In total, Geno4ME LDP variant calling results for 188 samples (119 whole blood and 69 saliva from the WGS Validation ORM group, Supplementary Spreadsheet SA) were validated by comparison the OS-ORM. Sensitivity, specificity, and accuracy of the Geno4ME LDP for variant calling was further tested by comparison to germline data from a previously-described cohort of 25 cancer patient reference samples sequenced by a clinically-validated WES assay (WES Comparison Validation Group, Supplementary Spreadsheet SA) (Bigelow et al., 2022).

TABLE 1 Expected SV/CNV in control NIBSC samples tested using the Geno4ME LDP method.

Sample ID	Expected SV/CNV
NIBSC 1	None
NIBSC 2	None
NIBSC 3	*MSH2* deletion exons 1–6, heterozygous
NIBSC 4	*MSH2* deletion exon 7, heterozygous
NIBSC 5	*MSH2* deletion exons 1–2, heterozygous
NIBSC 6	*MSH2* deletion exon 1, heterozygous
NIBSC 7	*MLH1* exon 13 amplification (three or more copies)

Reference DNA samples for CNV caller validation Geno4ME LDP were obtained from the National Institute for Biological Standards and Control (NIBSC, UK Stem Cell Bank Blanche Lane South Mimms Potters Bar Herts. EN6 3QG, NIBSC code: 11/218-XXX). Seven purified human genomic DNA samples with or without CNV variants in *MLH1* and *MSH2* were tested using the Geno4ME LDP and compared to the known copy number genotypes provided by the manufacturer (Table 1). Accuracy of the Geno4ME LDP

TABLE 2 List of drugs and their associated PGx genes assayed.

Drug(s)	Drug type	Associated gene(s)
Warfarin	Anticoagulant/antiplatelet	*VKORC1*, *CYP2C9*, *CYP4F2*, rs12777823
Clopidogrel	Anticoagulant/antiplatelet	*CYP2C19*
Lansoprazole Omeprazole Pantoprazole	Proton pump inhibitors	*CYP2C19*
Citalopram Escitalopram	Selective serotonin reuptake inhibitors	*CYP2C19*

TABLE 3 Summary of samples used for validation in this work.

Validation group	N Individuals	N samples	Purpose
EHR Comparison	30	30	Validation of DRAGEN WGS variant calling
WGS Validation ORM	188	188	Validation of Geno4ME LDP variant calling and PGx diplotyping to OS-ORM
WGS Validation SB[a]	60	120	Validation of Geno4ME LDP variant calling and PGx diplotyping between Saliva and Blood samples
WGS Validation CNV	N/A	7	Validation of CNV calling using the Geno4ME LDP to previously characterized DNA samples
WES Comparison	24	24	Validation of Geno4ME LDP variant calling to orthogonal WES-ORM
WGS Validation PGx	N/A	18	Validation of Geno4ME LDP PGx using previously characterized Coriell cell lines

[a]59 samples for the Saliva-Blood comparison overlap with the WGS validation ORM validation group.

TABLE 4 Coverage over genome metrics calculated by validation group.

Validation group	N samples	Mean	Median	SD	Min	Max
EHR Comparison	30	35.8X	35.6X	7.37X	20.9X	50.2X
WGS Validation (ORM + SB)	249	36.5X	37.0X	9.4X	16.5X	62.3X
WGS Validation CNV	7	40.4X	38.4X	6.8X	34.2X	53.9X
WES Comparison	24	37.2X	35.2X	8.7X	24.9X	54.7X
WGS Validation PGx	18	40.0X	41.9X	6.0X	30.7X	48.5X

was measured by identification of expected exon CNVs alterations characterized by the manufacturer.

Geno4ME LDP PGx genotyping and PGx phenotyping

Five PGx genes with gene-drug prescribing guidelines were selected based on the published joint recommendations from Clinical Pharmacogenetics Implementation Consortium (CPIC) and the U.S. Food and Drug Administration (Hicks et al., 2015; Lima et al., 2021; Scott et al., 2013; US Food and Drug Administration, 2020; Johnson et al., 2017). For these five PGx genes, variants were selected based on recommendations from AMP, the College of American Pathologists (CAP), and/or CPIC (Johnson et al., 2017; Pratt et al., 2019). The pre-selected defining PGx variants, the logic to define genotypes as well as the genotypes to phenotype mapping were implemented into Fabric Genomics cloud platform and used as part of the Geno4ME LDP (henceforth referred to as Geno4ME LDP PGx). The PGx panel considered for validation included 7 gene-drug pairs that were selected based on FDA and CPIC guidelines (Table 2). For *CYP2C19*, both Tier 1 (*2,*3, and*17) and Tier 2 (*4A,*4B,*5,*6,*7,*9,*10, and*35) alleles were included per AMP/CAP recommendation, *CYP2C9* Tier 1 alleles (*2,*3,*5,*6,*8, and*11), *VKORC1* (c.-1639G>A, rs9923231), *CYP4F2* (*3), and the single variant rs12777823 (*CYP2C* cluster) were included as recommended in the CPIC guideline for warfarin. The genotype to phenotype mapping was based on PharmGKB,

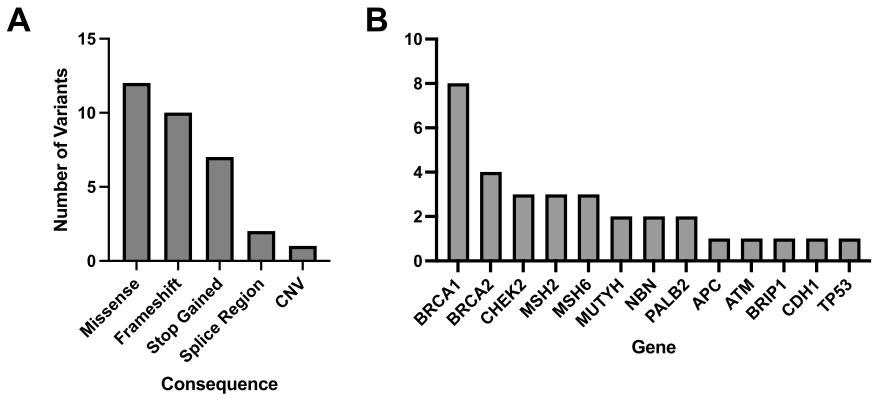

FIGURE 3
Molecular consequences of all matched variants identified in patient EHRs and compared to the DRAGEN WGS variant calling. A total of 33 variants across 30 patient EHRs had variant matches in the DRAGEN WGS variant calling pipeline and the majority were missense **(A)**. Out of the variants identified in patient EHRs, 13 genes were represented **(B)**.

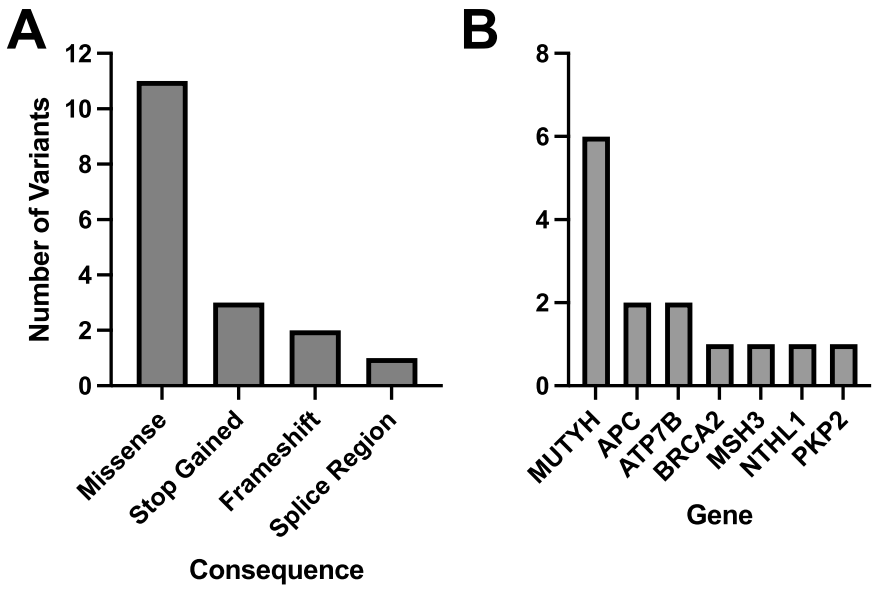

FIGURE 4
Molecular consequences and distribution of genes with matching P/LP classifications for all matched variants identified in the WGS Validation cohort. The WGS Validation cohort was tested using the Geno4ME LDP and OS-ORM in parallel. A total of 17 variants across 17 participants had matches to the ORM and the majority were missense **(A)**. Out of 14 P/LP variants that matched between the WGS Validation cohort and the OS-ORM, 7 genes were represented and the most frequent gene with P/LP variants was *MUTYH* **(B)**.

CPIC, and PharmVar annotations (Supplementary Table S3) (Johnson et al., 2017; Gaedigk et al., 2021; Whirl-Carrillo et al., 2021). *CYP2C19*, *CYP2C9*, and *CYP4F2* alleles negative for assayed variants were designated as *1.

Validation of Geno4ME LDP PGx genotyping

Validation of *CYP2C19*, *CYP2C9*, *CYP4F2, VKORC1* genotyping by Geno4ME LDP PGx was performed by comparison to a CLIA-certified commercial molecular laboratory using an outside reference method for PGx (OS-ORM PGx). The outside reference method for PGx validation was MassARRAY genotyping (Invitae). Accuracy of Geno4ME LDP PGx genotyping against the ORM PGx was performed using the same 188 samples used for validating the Geno4ME LDP variant call concordance (WGS Validation ORM group, Supplementary Spreadsheet SA). In addition, accuracy of Geno4ME LDP PGx genotyping was further validated by comparing Geno4ME LDP PGx to 18 previously characterized cell lines/DNA samples (WGS Validation PGx validation group). The following cell lines/DNA samples were obtained from the NHGRI Sample Repository for Human Genetic Research at the Coriell Institute for Medical Research: NA7019, NA7029 NA07439, NA10847, NA12717

10.3389/fmolb.2025.1669085

TABLE 5 Concordance of the Geno4ME LDP and the ORM for identifying gene SNVs/indels in the WGS Validation cohort.

Cohort	WGS validation ORM
N samples	188
Experimental Method	Geno4ME LDP
Reference Method	Hybridization-based NGS
Positive genes confirmed by Reference Method	17/17
Negative genes confirmed by Reference Method	14,382/14,382
Effect size	**Percent (95% CI)**
Sensitivity	100% (80.5%–100%)
Specificity	100% (100%–100%)
Positive Predictive Value	100% (80.5%–100%)
Negative Predictive Value	100% (100%–100%)

NA17641, NA18524, NA19109, NA23275, NA24008, HG00436, HG00589, HG01190, NA07348, NA12003, NA12878, NA19207, and NA19785. These cell lines/DNA samples were used to validate the Geno4ME LDP PGx for *CYP2C19, CYP2C9, CYP4F2*, and *VKORC1* diplotypes based on previously described diplotypes (Pratt et al., 2016). Genotypes not supported by the Geno4ME LDP PGx were not evaluated for accuracy when comparing to the NHGRI Coriell samples. The single variant rs12777823 was not supported by the OS-ORM PGx or NHGRI Coriell reference method and was therefore not compared to the Geno4ME LDP PGx method. Overall, the total number of individual samples used for Geno4ME LDP PGx validation was 206.

Statistics and plots

Sensitivity, specificity, Positive Predictive Value (PPV), Negative Predicative Value (NPV), and 95% confidence intervals (Cis) were calculated using the MedCalc Diagnostic Test evaluation calculator online tool (https://www.medcalc.org/calc/diagnostic_test.php). Graphs were created using Graphpad Prism (v10.1.1) or Rstudio (v2023.06.01, R version 4.3.1). For comparisons of the Geno4ME LDP for gene variant calling to the ORM and WES-RM, a positive gene match was determined if all variants detected within a gene matched between the Geno4ME LDP and reference method. Negative genes were determined if no variants were reported in either the Geno4ME LDP or reference method. Due to assay limitations, *MSH3* exon 1 and the *PMS2* regions with high homology to its pseudogene *PMS2CL* (exon 9 and exons 11–15) were excluded for comparisons of the Geno4ME LDP to the OS-ORM and WES-RM. A summary of validation group sample sizes and their purposes is shown in Table 3.

Results

Cohort composition and sequencing characteristics

For the WGS Validation ORM and WGS Validation SB validation groups the average age at time of collection was 56.5 years (SD = 13 years, N = 189 individuals). Out of these 189 participants, 157 self-reported sex as female and 32 as male. 60 participants provided matched saliva and blood samples. The average WGS sequencing coverage over human genome, separated by validation group, is shown in Table 4. Median coverage over genome per group ranged from 35.2–41.9X median per-base coverage. Samples used for inter-run variation analysis had a mean coverage over genome of 34.2X (SD = 1.7, CV = 4.9%). Samples used for intra-run variation analysis had a mean coverage over genome of 40.9X (SD = 8.8, CV = 21.4%). Full sequencing metrics for individual samples and by validation group are shown in Supplementary Figures S1–S3 and Supplementary Spreadsheet SA.

Concordance of DRAGEN WGS variant calling method to variants from patient EHRs

A total of 33 variants identified across 30 patient EHRs were used for comparison to the WGS DRAGEN variant caller (Supplementary Spreadsheet SB). Most identified variants were missense (Figure 3A) and the most frequent gene represented was *BRCA1* (Figure 3B).

Patient HR23 was found to have a large, multi-exon heterozygous *BRCA1* deletion identified by DRAGEN (*BRCA1* c.678_4740del, del *BRCA1* exons 10–15/24) and this deletion had 99% overlap with the deletion described by EHR (*BRCA1* c.671_4675del). Overall, all 33 variants identified in patient EHRs were identified using the WGS DRAGEN variant caller.

Concordance of Geno4ME LDP variant calling with the ORM

Within the WGS Validation ORM group, an average of 453.2 total variants per sample (SD = 141.7) were evaluated by the Geno4ME LDP. Most identified variants were SNVs, MNVs, or indels (N = 188, Supplementary Figure S4). Using the Geno4ME LDP, a total of 17 genes across 17 different participants were positive with a P/LP variant and most of these P/LP variants were missense (Figure 4A; Supplementary Spreadsheet SC). The results of the Geno4ME LDP were compared to the OS-ORM and were found to match with 100% concordance (Table 5; Supplementary Figure S5). Overall, 7 different genes with P/LP variants were represented in the WGS Validation ORM group (Figure 4B).

Performance of Geno4ME LDP for initial classification of variants

Across the 188 samples in the WGS Validation ORM validation group, 13,096 SNV, MNV, or indel variants were initially evaluated

10.3389/fmolb.2025.1669085

TABLE 6 Performance of the ACE autoclassification software for initial classification of variants in the WGS validation ORM group heritable disorder gene panel.

	Final classification (ACE + ACMG)					
		B	LB	VUS	LP	P	Total[a]
ACE Autoclassification	B	11,533	0	0	0	0	11,533
	LB	0	1,101	0	0	0	1,101
	VUS	250	0	197	0	8[b]	455
	LP	0	0	1	0	3	4
	P	0	0	0	0	3	3
	Total[a]	11,783	1,101	198	0	14	

[a]The number of samples represented is N = 188 across the WGS Validation ORM group.
[b]Two instances of the established risk allele *APC* c.3920T>A p.Ile1307Lys are included as Pathogenic.

TABLE 7 Variant concordance between the Geno4ME LDP and WES-ORM in the WES Comparison cohort.

Cohort	WES comparison
N samples	25
Experimental Method	Geno4ME LDP
Reference Method	WES
Positive genes confirmed by Reference Method	77/78
Negative genes confirmed by Reference Method	874/874
Effect size	**Percent (95% CI)**
Sensitivity	98.7% (93%–100%)
Specificity	100% (99.6%–100%)
Positive Predictive Value	100% (95.3%–100%)
Negative Predictive Value	99.9% (99.2%–100%)

and classified by ACE in the Fabric Enterprise platform (Table 6). Of the 455 variants classified by ACE as VUS, 250 were reclassified as benign following curation. Most of these 250 variants reclassified from VUS to benign consisted of two missense variants that were interpreted by automated pipelines as MNVs without a presence in population databases (dbSNP reference alleles rs386638457 and rs386643884) instead of separate, high population frequency SNVs. Eight variants initially classified by ACE as VUS were reclassified to pathogenic following manual curation. One variant, the in-frame insertion variant *WT1* c.378_392dup p.Ala127_Pro131dup, was classified by ACE as LP but after manual curation was reclassified as VUS. Overall, no variants classified by ACE as B/LB were later found to be LP/P following comparison to variant classification results from the ORM or manual review.

Concordance of Geno4ME LDP and WES-RM variant calling

A total of 78 variants across 25 samples were identified using the Geno4ME LDP and/or the WES-ORM within the WES Validation ORM validation group (Table 7; Supplementary Spreadsheet SD). Most of the identified variants were missense (Figure 5A). Of the 78 variants identified, 63 unique variants were represented across 26 different genes (Figure 5B). All variants were identified in both the Geno4ME LDP and WES-ORM for each sample, except for a *TP53* c.854A>T p.Glu285Val variant in sample gme-wes-19 that was present at 27% VAF only in the previously sequenced WES data. Investigation of the raw WGS DRAGEN variant calling data for the sample revealed two supporting reads for the *TP53* c.854A>T p.Glu285Val variant (depth = 15) but a low variant quality score (Qual = 5.18) that resulted in filtering by the Fabric Enterprise platform. Review of the patient medical record revealed that a matched solid tumor specimen that was sequenced using hybrid-capture NGS and contained the *TP53* variant at 5% VAF. Using the WES-ORM as the reference method, the Sensitivity and Specificity of Geno4ME LDP for variant calling was 98.72% and 100%, respectively (Table 7; Supplementary Figure S6).

Performance of the Geno4ME LDP for identification of large CNVs

Seven control NIBSC samples were tested using the Geno4ME LDP to evaluate CNV caller accuracy at single and multi-exon resolution. All five samples with CNVs characterized by the sample provider were identified using one or both GenoME LDP CNV calling algorithms to affect the same expected gene exons with the same zygosity and CNV type (Supplementary Spreadsheet SE). Two samples expected to be negative for CNV alterations were also negative by Geno4ME LDP. Overall, the concordance between the CNV alterations described by the sample provider for the seven control NIBSC samples and the Geno4ME LDP was 100%.

10.3389/fmolb.2025.1669085

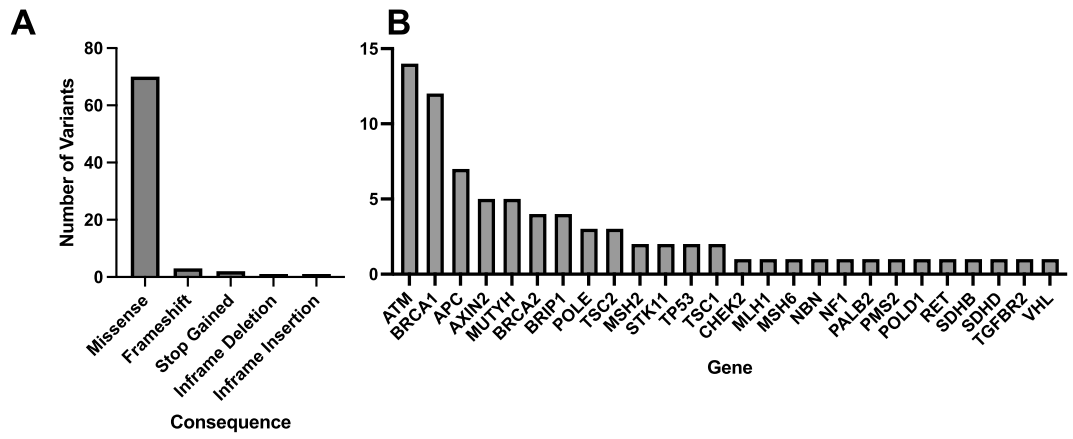

FIGURE 5
Molecular consequences and genes represented for of all matched variants identified comparing Geno4ME LDP to WES-ORM. A total of 78 variants were identified across all samples, with the majority being missense **(A)**, and 26 unique genes were represented **(B)**.

TABLE 8 PGx variant concordance of the Geno4ME LDP method and DNA derived from previously characterized NHGRI coriell samples.

Cohort size	Number of matching samples	Gene	Concordant variants
N = 18	18/18 (100%)	*CYP2C19*	18/18 (100%)
		CYP2C9	16/16 (100%)
		CYP4F2	14/14 (100%)
		VKORC1	18/18 (100%)

Concordance of Geno4ME LDP PGx with NHGRI coriell samples

Eighteen samples obtained from NHGRI Coriell with known *CYP2C19*, *CYP2C9*, *CYP4F2*, and *VKORC1* diplotypes (WGS Validation PGx group) were evaluated using the Geno4ME LDP PGx. Of these samples, six diplotype calls across five NHGRI Coriell samples could not be evaluated due to the diplotype being unknown or not supported by the Geno4ME LDP PGx pipeline (Supplementary Spreadsheet SF). For diplotypes supported by the Geno4ME LDP PGx, the overall concordance between the Geno4ME LDP PGx and the previously characterized NHGRI Coriell samples was 100% (Table 8).

Concordance of Geno4ME LDP PGx to ORM PGx

Performance of the Geno4ME LDP PGx was further evaluated by comparing Geno4ME LDP PGx results of the WGS Validation ORM validation group to the OS-ORM PGx (N = 188). The most common diplotype for each gene was *CYP2C19**1/*1, CYP2C9*1/*1, CYP4F2*1/*1, and *VKORC1* GA (Table 9). Most participants in the WGS Validation ORM validation group had normal metabolizer status for *CYP2C19* and *CYP2C9* (Table 10). Slightly more than half of participants were classified as warfarin resistant based on *CYP4F2* diplotype and/or were warfarin sensitive based on *VKORC1* diplotype. For one participant, gme-039, the *CYP2C19* diplotype was identified using the Geno4ME LDP method as*4A/*17 and the OS-ORM PGx as*4/*17. However, the discrepancy was determined to be due to nomenclature differences of reporting*4 subalelles where*4A and*4 have the same core alleles and phenotype. The overall concordance between the Geno4ME LDP and ORM PGx diplotyping for *CYP2C19*, *CYP2C9*, *CYP4F2*, and *VKORC1* was 100% (Table 11; Supplementary Spreadsheet SG).

Concordance between blood and saliva samples using the Geno4ME LDP

Of the 60 participants with matched blood and saliva samples that were processed using the Geno4ME LDP (WGS Validation SB validation group), 32 matched pairs were positive for P/LP or VUS variants and 28 were screen-negative (Table 12). Within the 32 positive matched pairs, the majority of the 47 matching variants were missense (Figure 6A). These 47 variants were identified across 26 different genes and the overall concordance between DNA derived from blood or saliva was 100% (Figure 6B; Supplementary Spreadsheet SH). For these same 60 participants, Geno4ME LDP PGx diplotype results for *CYP2C19*, *CYP2C9*, *CYP4F2*, *VKORC1*, and rs12777823 were compared and had an overall concordance of 100% (Table 13; Supplementary Spreadsheet SI).

TABLE 9 Distribution of PGx gene diplotypes from the WGS validation cohort identified using the Geno4ME LDP.

Gene	Diplotype	n	Percent
CYP2C19	*1/*1	73	38.8%
	*1/*17	50	26.6%
	*1/*2	35	18.6%
	*2/*17	13	6.9%
	*17/*17	8	4.3%
	*2/*2	5	2.7%
	*1/*3	1	0.5%
	*2/*4B	1	0.5%
	*4A/*17	1	0.5%
	*8/*17	1	0.5%
CYP2C9	*1/*1	127	67.6%
	*1/*2	36	19.1%
	*1/*3	20	10.6%
	*2/*3	3	1.6%
	*2/*2	2	1.1%
CYP4F2	*1/*1	88	46.8%
	*1/*3	85	45.2%
	*3/*3	15	8.0%
VKORC1	GA	109	58.0%
	GG	61	32.4%
	AA	18	9.6%
rs12777823	GG	132	70.2%
	GA	53	28.2%
	AA	3	1.6%

TABLE 10 PGx metabolizer status distribution from the WGS validation cohort.

Gene	Metabolizer status	n	Percent
CYP2C19	Normal metabolizer	73	38.8%
	Intermediate metabolizer	51	27.1%
	Rapid metabolizer	50	26.6%
	Ultrarapid metabolizer	8	4.3%
	Poor Metabolizer	6	3.2%
CYP2C9	Normal metabolizer	127	67.6%
	Intermediate metabolizer	56	29.8%
	Poor metabolizer	5	2.7%
CYP4F2	Warfarin resistant	100	53.2%
	Normal warfarin sensitivity	88	46.8%
VKORC1	Warfarin sensitive	127	67.6%
	Normal warfarin sensitivity	61	32.4%
rs12777823	Normal warfarin sensitivity	132	70.2%
	Warfarin sensitive	56	29.8%

Discussion

Rapid increases in sequencer capacity and throughput, automation and AI assistance for variant detection and interpretation, and significant decreases in per base sequencing costs are making it feasible to use WGS for routine clinical applications. Given that a majority of individuals carrying an actionable P/LP variant are unaware of their carrier status (Zawatsky et al., 2021), and that accumulating evidence suggests that identifying individuals for testing based on medical history alone insufficiently captures the majority of individuals with monogenic risk for a heritable genetic condition (Manickam et al., 2018; Abul-Husn et al., 2016; Murray et al., 2020), the introduction of WGS-based screening of patient populations has high potential for improving health outcomes. Furthermore, several recent studies have demonstrated the utility of population-based high-throughput genomic screening for facilitating early disease detection, risk management, and reducing adverse drug reactions (Buchanan et al., 2020; Grzymski et al., 2020; Rao et al., 2023; Swen et al., 2023).

To address this opportunity for improving health outcomes, we show that the Geno4ME LDP is a highly accurate clinical procedure for evaluating actionable genomic variants. We found 100% of previously tested variants obtained from patient EHRs to be identified using the DRAGEN WGS variant caller. The Geno4ME LDP had high overall agreement with well-characterized reference samples and with samples orthogonally tested by the ORM or WES-RM. When compared to variants identified using the OS-ORM PGx, the Geno4ME LDP had 100% sensitivity and specificity for identifying variants of interest. In addition, the Geno4ME LDP had 100% accuracy when compared to CNV and PGx reference samples. One notable discrepancy was the presence of a predicted pathogenic *TP53* c.854A>T p.Glu285Val in one WES sample but absent using Geno4ME LDP. The low level VAF of the *TP53* variant in the matched tumor sample compared to the control blood sample suggested that the variant may represent chimerism or, more likely, clonal hematopoiesis of indeterminate potential (CHIP). However, the patient was deceased at the time of Geno4ME LDP testing and therefore additional sample could not be obtained to definitively determine the origin of the *TP53* variant in the WES data. This discrepancy demonstrates the need for genomic screening clinical tests to be understood in the context of the patient's/participant's medical history, and like all germline testing the ability to assess

TABLE 11 Concordance of the Geno4ME LDP method and the OS-ORM PGx for identifying PGx diplotypes in the WGS validation cohort.

Cohort	Experimental method	Reference method	Gene tested	Number of concordant results	Percent concordance
WGS Validation ORM	Geno4ME LDP	MALDI-TOF Genotyping	*CYP2C19*	188	100%
			CYP2C9	188	100%
			VKORC1	188	100%
			CYP4F2	188	100%

TABLE 12 Small variant concordance of the WGS-Fabric method between DNA derived from blood or saliva for the same participants.

Cohort	Outcome	Number of matching samples	Concordant variants
WGS Validation SB	Positive	32/32 (100%)	47/47 (100%)
	Negative	28/28 (100%)	NA

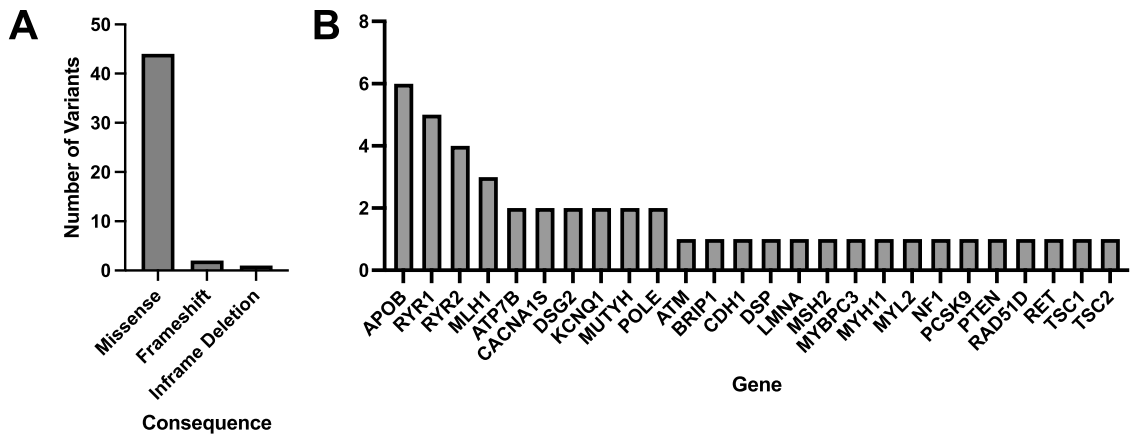

FIGURE 6
Molecular consequences and genes represented for all matched variants identified in the WGS Validation cohort comparing DNA derived from blood or saliva using Geno4ME LDP. A total of 47 variants were identified across all samples, with the majority being missense **(A)**, and 26 unique genes were represented **(B)**.

TABLE 13 PGx variant concordance of the Geno4ME LDP method between DNA derived from Blood or Saliva for the same participants.

Cohort	Number of matching samples	Gene	Concordant variants
WGS Validation SB	60/60 (100%)	*CYP2C19*	60/60 (100%)
		CYP2C9	60/60 (100%)
		CYP4F2	60/60 (100%)
		VKORC1	60/60 (100%)
		rs12777823	60/60 (100%)

fibroblast cell lines if needed, particularly with *TP53* variants. Notably, detection of low-level chimerism or CHIP likely requires an increased read depth that was outside the scope of this population genomics study. Future iterations of the Geno4ME LDP could benefit from increased coverage or other highly sensitive detection methods to detect CHIP, which is now being recognized to have possible implications for clinical decision making (Libby et al., 2019; Zhang et al., 2025; Uddin et al., 2022).

The feasibility of using a clinical lab procedure such as Geno4ME LDP for large populations requires scalability at both the levels of sample collection and data analysis. We found identical variant concordance between samples obtained using saliva or blood. Our work agrees with previous data suggesting that once oral microbiome reads are excluded, saliva-derived and blood-derived genomic DNA are generally comparable in quality for clinical WGS workflows (Yao et al., 2020; Kvapilova et al., 2024). Since saliva can be obtained via at-home collections without the need for a phlebotomist, the sample type would be more conducive to population health applications. At the level of data analysis, use of ACE in the variant curation process rapidly screened variants of interest and allowed for reduced hands-on time for manually curating variants. For example, out of 113,096 evaluated variants in the WGS Validation ORM group heritable disorder gene panel, ACE was able to reduce the number of variants requiring manual evaluation to 250 (0.02% of all variants evaluated in the group) without loss of P/LP identification sensitivity. In addition, the automated Geno4ME LDP PGx accurately diplotyped all samples, providing predicted metabolizer status for all PGx genes analyzed in this study without manual intervention.

Upfront testing of the whole genome using a method such as Geno4ME LDP allows for simultaneous testing of multiple genomic features while also allowing for future incorporation of additional screened genes and/or polygenic risk tests through region unmasking. Because the Geno4ME LDP was validated for all variant types of interest, new genes can rapidly be added to the panel in the future. Reasons for reanalysis using the Geno4ME LDP could include new gene-disease associations, additional information availability for a patient, and advances in variant calling techniques (Robertson et al., 2022). In principle, reanalysis of gene variants in the Geno4ME LDP for newly or previously selected genes would be achieved by simply passing previously generated WGS variant data through Fabric Enterprise platform with updated population, predictive, and curated databases. Future reanalysis of participants tested using the Geno4ME LDP as part of the Geno4ME clinical implementation study (N = 2,017 participants) may provide valuable insight into the usefulness and challenges of iterative population genomic testing (Lucas Beckett et al., 2025).

We acknowledge that there are remaining challenges when using short-read sequencing for achieving comprehensive population genomics testing. Recent work suggests that complex structural variants represent a substantial component of actionable genomic conditions (Collins et al., 2020). In addition, highly polymorphic genes such as relevant PGx gene *CYP2D6*, pseudogene regions, and highly repetitive regions continue to present analysis challenges (Kane, 2021; Morton et al., 2020; C et al., 2020). Although WGS tools for genotyping *CYP2D6* are available (Chen et al., 2021), the clinical importance of *CYP2D6* would necessitate comprehensive validation if it were to be part of the Geno4ME LDP, and may require alternative sequencing methods to capture complex alleles (Pratt et al., 2021). A similar alternative sequencing strategy may also be necessary for detection of alleles in *PMS2* that are confounded by its pseudogene, *PMS2CL* (Bouras et al., 2024; Gould et al., 2018). We also recognize that while we were able to confirm accurate detection of several CNVs in available reference material, the Geno4ME LDP could benefit from CNV detection validation at other regions to ensure the behavior of CNV detection is reproducible for other genes in the panel. Future comparisons of the Geno4ME LDP to well-characterized complex reference samples will determine the limits of the method for detecting actionable SV and polymorphic variation as part of clinical testing.

Data availability statement

The raw datasets generated and/or analyzed during the current study are not publicly available due to protocol and Informed Consent language that mandate of IRB to review projects and require a data use agreement (DUA), but are available from the corresponding author on reasonable request. Analyzed data is provided within the manuscript and supplementary information files.

Ethics statement

The studies involving humans were approved by the Providence Institutional Review Board (Approval number: STUDY2020000637). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JsW: Software, Investigation, Writing – review and editing, Writing – original draft, Formal Analysis, Visualization, Data curation, Validation, Methodology, Conceptualization. JhW: Writing – original draft, Formal Analysis, Conceptualization, Methodology, Validation, Writing – review and editing, Investigation. IL: Investigation, Formal Analysis, Visualization, Validation, Data curation, Writing – review and editing. KE: Visualization, Project administration, Writing – review and editing. BC: Writing – review and editing, Data curation, Software. KO: Software, Writing – review and editing. NW: Data curation, Software, Writing – review and editing. TB: Data curation, Writing – review and editing, Software. LY: Writing – review and editing, Investigation. ES: Data curation, Visualization, Writing – review and editing, Software. KJ: Writing – review and editing, Visualization. JC: Writing – review and editing, Project administration. AM: Writing – review and editing, Resources, Project administration, Data curation. MC: Project administration, Writing – review and editing. OG: Conceptualization, Funding acquisition, Writing – review and editing. CB: Methodology, Supervision, Visualization, Resources, Funding acquisition, Conceptualization, Validation, Writing – review and editing. BP: Validation, Conceptualization,

Methodology, Supervision, Investigation, Funding acquisition, Resources, Visualization, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported by internal funding from Providence Health.

Conflict of interest

Author KO was employed by Fabric Genomics.

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2025.1669085/full#supplementary-material

References

Abul-Husn, N. S., Manickam, K., Jones, L. K., Wright, E. A., Hartzel, D. N., Gonzaga-Jauregui, C., et al. (2016). Genetic identification of familial hypercholesterolemia within a single US health care system. *Science* 354 (6319), aaf7000. doi:10.1126/science.aaf7000

Bejjani, B. A., and Shaffer, L. G. (2006). Application of array-based comparative genomic hybridization to clinical diagnostics. *J. Mol. Diagnostics* 8 (5), 528–533. doi:10.2353/jmoldx.2006.060029

Bick, A. G., Metcalf, G. A., Mayo, K. R., Lichtenstein, L., Rura, S., Carroll, R. J., et al. (2024). Genomic data in the all of us research program. *Nature* 627, 340–346. doi:10.1038/s41586-023-06957-x

Biesecker, L. G. (2012). Opportunities and challenges for the integration of massively parallel genomic sequencing into clinical practice: lessons from the ClinSeq project. *Genet. Med.* 14 (4), 393–398. doi:10.1038/gim.2011.78

Biesecker, L. G., Mullikin, J. C., Facio, F. M., Turner, C., Cherukuri, P. F., Blakesley, R. W., et al. (2009). The ClinSeq project: piloting large-scale genome sequencing for research in genomic medicine. *Genome Res.* 19 (9), 1665–1674. doi:10.1101/gr.092841.109

Biesecker, L. G., Harrison, S. M., and ClinGen Sequence Variant Interpretation Working Group (2018). The ACMG/AMP reputable source criteria for the interpretation of sequence variants. *Genet. Med.* 20 (12), 1687–1688. doi:10.1038/gim.2018.42

Bigelow, E., Saria, S., Piening, B., Curti, B., Dowdell, A., Weerasinghe, R., et al. (2022). A random forest genomic classifier for tumor agnostic prediction of response to Anti-PD1 immunotherapy. *Cancer Inf.* 21, 11769351221136081. doi:10.1177/11769351221136081

Bouras, A., Lefol, C., Ruano, E., Grand-Masson, C., and Wang, Q. (2024). PMS2 or PMS2CL? Characterization of variants detected in the 3′ of the PMS2 gene. *Genes Chromosomes Cancer* 63 (1), e23193. doi:10.1002/gcc.23193

Buchanan, A. H., Lester Kirchner, H., Schwartz, M. L. B., Kelly, M. A., Schmidlen, T., Jones, L. K., et al. (2020). Clinical outcomes of a genomic screening program for actionable genetic conditions. *Genet. Med.* 22 (11), 1874–1882. doi:10.1038/s41436-020-0876-4

Cheetham, S. W., Faulkner, G. J., and Dinger, M. E. (2020). Overcoming challenges and dogmas to understand the functions of pseudogenes. *Nat. Rev. Genet.* 21 (3), 191–201. doi:10.1038/s41576-019-0196-1

Carey, D. J., Fetterolf, S. N., Davis, F. D., Faucett, W. A., Kirchner, H. L., Mirshahi, U., et al. (2016). The geisinger MyCode community health initiative: an electronic health record–linked biobank for precision medicine research. *Genet. Med.* 18 (9), 906–913. doi:10.1038/gim.2015.187

Chen, X., Shen, F., Gonzaludo, N., Malhotra, A., Rogert, C., Taft, R. J., et al. (2021). Cyrius: accurate CYP2D6 genotyping using whole-genome sequencing data. *Pharmacogenomics J.* 21 (2), 251–261. doi:10.1038/s41397-020-00205-5

Chen, T., Fan, C., Huang, Y., Feng, J., Zhang, Y., Miao, J., et al. (2023). Genomic sequencing as a first-tier screening test and outcomes of newborn screening. *JAMA Netw. Open* 6 (9), e2331162. doi:10.1001/jamanetworkopen.2023.31162

Chen, S., Francioli, L. C., Goodrich, J. K., Collins, R. L., Kanai, M., Wang, Q., et al. (2024). A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* 625 (7993), 92–100. doi:10.1038/s41586-023-06045-0

Chunn, L. M., Nefcy, D. C., Scouten, R. W., Tarpey, R. P., Chauhan, G., Lim, M. S., et al. (2020). Mastermind: a comprehensive genomic association search engine for empirical evidence curation and genetic variant interpretation. *Front. Genet.* 11, 577152. doi:10.3389/fgene.2020.577152

Church, D. M., Schneider, V. A., Graves, T., Auger, K., Cunningham, F., Bouk, N., et al. (2011). Modernizing reference genome assemblies. *PLoS Biol.* 9 (7), e1001091. doi:10.1371/journal.pbio.1001091

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 (7809), 444–451. doi:10.1038/s41586-020-2287-8

Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., et al. (2010). Origins and functional impact of copy number variation in the human genome. *Nature* 464 (7289), 704–712. doi:10.1038/nature08516

Daly, M. B., Pal, T., Berry, M. P., Buys, S. S., Dickson, P., Domchek, S. M., et al. (2021). Genetic/familial high-risk assessment: breast, ovarian, and pancreatic, version 2.2021, NCCN clinical practice guidelines in oncology. *J. Natl. Compr. Canc Netw.* 19 (1), 77–102. doi:10.6004/jnccn.2021.0001

De La Vega, F. M., Chowdhury, S., Moore, B., Frise, E., McCarthy, J., Hernandez, E. J., et al. (2021). Artificial intelligence enables comprehensive genome interpretation and nomination of candidate diagnoses for rare genetic diseases. *Genome Med.* 13, 153–19. doi:10.1186/s13073-021-00965-0

Dolzhenko, E., van Vugt, J. J. F. A., Shaw, R. J., Bekritsky, M. A., van Blitterswijk, M., Narzisi, G., et al. (2017). Detection of long repeat expansions from PCR-Free whole-genome sequence data. *Genome Res.* 27 (11), 1895–1903. doi:10.1101/gr.225672.117

Dunnenberger, H. M., Crews, K. R., Hoffman, J. M., Caudle, K. E., Broeckel, U., Howard, S. C., et al. (2015). Preemptive clinical pharmacogenetics implementation: current programs in five US medical centers. *Annu. Rev. Pharmacol. Toxicol.* 55, 89–106. doi:10.1146/annurev-pharmtox-010814-124835

East, K. M., Kelley, W. V., Cannon, A., Cochran, M. E., Moss, I. P., May, T., et al. (2021). A state-based approach to genomics for rare disease and population screening. *Genet. Med.* 23 (4), 777–781. doi:10.1038/s41436-020-01034-4

Foss, K. S., O'Daniel, J. M., Berg, J. S., Powell, S. N., Cadigan, R. J., Kuczynski, K. J., et al. (2022). The rise of population genomic screening: characteristics of current programs and the need for evidence regarding optimal implementation. *J. Pers. Med.* 12 (5), 692. doi:10.3390/jpm12050692

Gaedigk, A., Casey, S. T., Whirl-Carrillo, M., Miller, N. A., and Klein, T. E. (2021). Pharmacogene variation consortium: a global resource and repository for pharmacogene variation. *Clin. Pharmacol. Ther.* 110 (3), 542–545. doi:10.1002/cpt.2321

Gall, B. J., Smart, T. B., Munch, R., Kolluri, S., Tadepally, H., Lim, K. P. H., et al. (2022). Assessment of an automated approach for variant interpretation in screening for monogenic disorders: a single-center study. *Mol. Genet. and Genomic Med.* 10 (12), e2085. doi:10.1002/mgg3.2085

Gonzalez-Garay, M. L., McGuire, A. L., Pereira, S., and Caskey, C. T. (2013). Personalized genomic disease risk of volunteers. *Proc. Natl. Acad. Sci.* 110 (42), 16957–16962. doi:10.1073/pnas.1315934110

Gould, G. M., Grauman, P. V., Theilmann, M. R., Spurka, L., Wang, I. E., Melroy, L. M., et al. (2018). Detecting clinically actionable variants in the 3′ exons of PMS2 *via* a reflex workflow based on equivalent hybrid capture of the gene and its pseudogene. *BMC Med. Genet.* 19 (1), 176. doi:10.1186/s12881-018-0691-9

Grzymski, J., Elhanan, G., Morales Rosado, J. A., Smith, E., Schlauch, K. A., Read, R., et al. (2020). Population genetic screening efficiently identifies carriers of autosomal dominant diseases. *Nat. Med.* 26 (8), 1235–1239. doi:10.1038/s41591-020-0982-5

Haer-Wigman, L., van der Schoot, V., Feenstra, I., Vulto-van Silfhout, A. T., Gilissen, C., Brunner, H. G., et al. (2019). 1 in 38 individuals at risk of a dominant medically actionable disease. *Eur. J. Hum. Genet.* 27 (2), 325–330. doi:10.1038/s41431-018-0284-2

Harrison, S. M., Austin-Tse, C. A., Kim, S., Lebo, M., Leon, A., Murdock, D., et al. (2022). Harmonizing variant classification for return of results in the all of us research program. *Hum. Mutat.* 43 (8), 1114–1121. doi:10.1002/humu.24317

Hehir-Kwa, J. Y., Pfundt, R., and Veltman, J. A. (2015). Exome sequencing and whole genome sequencing for the detection of copy number variation. *Expert Rev. Mol. Diagn* 15 (8), 1023–1032. doi:10.1586/14737159.2015.1053467

Hicks, J. K., Bishop, J. R., Sangkuhl, K., Müller, D. J., Ji, Y., Leckband, S. G., et al. (2015). Clinical pharmacogenetics implementation consortium (CPIC) guideline for CYP2D6 and CYP2C19 genotypes and dosing of selective serotonin reuptake inhibitors. *Clin. Pharmacol. and Ther.* 98 (2), 127–134. doi:10.1002/cpt.147

Holm, I. A., Agrawal, P. B., Ceyhan-Birsoy, O., Christensen, K. D., Fayer, S., Frankel, L. A., et al. (2018). The BabySeq project: implementing genomic sequencing in newborns. *BMC Pediatr.* 18, 225–10. doi:10.1186/s12887-018-1200-1

Johnson, J. A., Caudle, K. E., Gong, L., Whirl-Carrillo, M., Stein, C. M., Scott, S. A., et al. (2017). Clinical pharmacogenetics implementation consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 update. *Clin. Pharmacol. and Ther.* 102 (3), 397–404. doi:10.1002/cpt.668

Kalia, S. S., Adelman, K., Bale, S. J., Chung, W. K., Eng, C., Evans, J. P., et al. (2017). Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2. 0): a policy statement of the American college of medical genetics and genomics. *Genet. Med.* 19 (2), 249–255. doi:10.1038/gim.2016.190

Kane, M. (2021). CYP2D6 overview: allele and phenotype frequencies. In: *Medical genetics summaries*. Bethesda, MD: National Center for Biotechnology Information.

Kim, J., Naqvi, A. S., Corbett, R. J., Kaufman, R. S., Vaksman, Z., Brown, M. A., et al. (2024). AutoGVP: a dockerized workflow integrating ClinVar and InterVar germline sequence variant classification. *Bioinformatics* 40 (3), btae114. doi:10.1093/bioinformatics/btae114

Kvapilova, K., Misenko, P., Radvanszky, J., Brzon, O., Budis, J., Gazdarica, J., et al. (2024). Validated WGS and WES protocols proved saliva-derived gDNA as an equivalent to blood-derived gDNA for clinical and population genomic analyses. *BMC genomics* 25 (1), 187. doi:10.1186/s12864-024-10080-0

Landrum, M. J., Lee, J. M., Riley, G. R., Jang, W., Rubinstein, W. S., Church, D. M., et al. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic acids Res.* 42 (D1), D980–D985. doi:10.1093/nar/gkt1113

Libby, P., Sidlow, R., Lin, A. E., Gupta, D., Jones, L. W., Moslehi, J., et al. (2019). Clonal hematopoiesis: crossroads of aging, cardiovascular disease, and cancer: JACC review topic of the week. *J. Am. Coll. Cardiol.* 74 (4), 567–577. doi:10.1016/j.jacc.2019.06.007

Lima, J. J., Thomas, C. D., Barbarino, J., Desta, Z., Van Driest, S. L., El Rouby, N., et al. (2021). Clinical pharmacogenetics implementation consortium (CPIC) guideline for CYP2C19 and proton pump inhibitor dosing. *Clin. Pharmacol. and Ther.* 109 (6), 1417–1423. doi:10.1002/cpt.2015

Lindor, N. M., Thibodeau, S. N., and Burke, W. (2017). "Whole-genome sequencing in healthy people. *Mayo Clinic Proc.* 92, 159–172.

Liu, H.-Y., Zhou, L., Zheng, M. Y., Huang, J., Wan, S., Zhu, A., et al. (2019). Diagnostic and clinical utility of whole genome sequencing in a cohort of undiagnosed Chinese families with rare diseases. *Sci. Rep.* 9 (1), 19365–11. doi:10.1038/s41598-019-55832-1

Lucas Beckett, I. A., Emery, K. R., Wagner, J. T., Jade, K., Cosgrove, B. A., Welle, J., et al. (2025). Geno4ME study: implementation of whole genome sequencing for population screening in a large healthcare system. *npj Genomic Med.* 10 (1), 50. doi:10.1038/s41525-025-00508-1

Manickam, K., Buchanan, A. H., Schwartz, M. L. B., Hallquist, M. L. G., Williams, J. L., Rahm, A. K., et al. (2018). Exome sequencing–based screening for BRCA1/2 expected pathogenic variants among adult biobank participants. *JAMA Netw. Open* 1 (5), e182140. doi:10.1001/jamanetworkopen.2018.2140

Morton, E. A., Hall, A. N., Kwan, E., Mok, C., Queitsch, K., Nandakumar, V., et al. (2020). Challenges and approaches to genotyping repetitive DNA. *G3 Genes, Genomes, Genet.* 10 (1), 417–430. doi:10.1534/g3.119.400771

Murray, M. F., Evans, J. P., and Khoury, M. J. (2020). DNA-Based population screening: potential suitability and important knowledge gaps. *JAMA* 323 (4), 307–308. doi:10.1001/jama.2019.18640

Musunuru, K., Hershberger, R. E., Day, S. M., Klinedinst, N. J., Landstrom, A. P., Parikh, V. N., et al. (2020). Genetic testing for inherited cardiovascular diseases: a scientific statement from the American heart association. *Circulation Genomic Precis. Med.* 13 (4), e000067. doi:10.1161/HCG.0000000000000067

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic acids Res.* 44 (D1), D733–D745. doi:10.1093/nar/gkv1189

Park, S.-J., Jung, E. H., Ryu, R. S., Kang, H. W., Ko, J. M., Kim, H. J., et al. (2011). Clinical implementation of whole-genome array CGH as a first-tier test in 5080 pre and postnatal cases. *Mol. Cytogenet.* 4 (1), 12–10. doi:10.1186/1755-8166-4-12

Perkins, B. A., Caskey, C. T., Brar, P., Dec, E., Karow, D. S., Kahn, A. M., et al. (2018). Precision medicine screening using whole-genome sequencing and advanced imaging to identify disease risk in adults. *Proc. Natl. Acad. Sci.* 115 (14), 3686–3691. doi:10.1073/pnas.1706096114

Phillips, K. A., Deverka, P. A., Hooker, G. W., and Douglas, M. P. (2018). Genetic test availability and spending: where are we now? Where are we going? *Health Aff.* 37 (5), 710–716. doi:10.1377/hlthaff.2017.1427

Pratt, V. M., Everts, R. E., Aggarwal, P., Beyer, B. N., Broeckel, U., Epstein-Baak, R., et al. (2016). Characterization of 137 genomic DNA reference materials for 28 pharmacogenetic genes: a GeT-RM collaborative project. *J. Mol. diagnostics* 18 (1), 109–123. doi:10.1016/j.jmoldx.2015.08.005

Pratt, V. M., Cavallari, L. H., Del Tredici, A. L., Hachad, H., Ji, Y., Moyer, A. M., et al. (2019). Recommendations for clinical CYP2C9 genotyping allele selection: a joint recommendation of the association for molecular pathology and college of American pathologists. *J. Mol. Diagnostics* 21 (5), 746–755. doi:10.1016/j.jmoldx.2019.04.003

Pratt, V. M., Cavallari, L. H., Del Tredici, A. L., Gaedigk, A., Hachad, H., Ji, Y., et al. (2021). Recommendations for clinical CYP2D6 genotyping allele selection: a joint consensus recommendation of the association for molecular pathology, college of American pathologists, Dutch pharmacogenetics working group of the royal Dutch pharmacists association, and the european society for pharmacogenomics and personalized therapy. *J. Mol. Diagnostics* 23 (9), 1047–1064. doi:10.1016/j.jmoldx.2021.05.013

Preston, C. G., Wright, M. W., Madhavrao, R., Harrison, S. M., Goldstein, J. L., Luo, X., et al. (2022). ClinGen variant curation interface: a variant classification platform for the application of evidence criteria from ACMG/AMP guidelines. *Genome Med.* 14 (1), 6. doi:10.1186/s13073-021-01004-8

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., et al. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44 (W1), W160–W165. doi:10.1093/nar/gkw257

Rao, N. D., Kaganovsky, J., Malouf, E. A., Coe, S., Huey, J., Tsinajinne, D., et al. (2023). Diagnostic yield of genetic screening in a diverse, community-ascertained cohort. *Genome Med.* 15 (1), 26–12. doi:10.1186/s13073-023-01174-7

Ravichandran, V., Shameer, Z., Kemel, Y., Walsh, M., Cadoo, K., Lipkin, S., et al. (2019). Toward automation of germline variant curation in clinical cancer genetics. *Genet. Med.* 21 (9), 2116–2125. doi:10.1038/s41436-019-0463-8

Relling, M. V., and Evans, W. E. (2015). Pharmacogenomics in the clinic. *Nature* 526 (7573), 343–350. doi:10.1038/nature15817

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., 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 (5), 405–424. doi:10.1038/gim.2015.30

Robertson, A. J., Tan, N. B., Spurdle, A. B., Metke-Jimenez, A., Sullivan, C., and Waddell, N. (2022). Re-analysis of genomic data: an overview of the mechanisms and complexities of clinical adoption. *Genet. Med.* 24 (4), 798–810. doi:10.1016/j.gim.2021.12.011

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29 (1), 24–26. doi:10.1038/nbt.1754

Scott, S., Sangkuhl, K., Stein, C. M., Hulot, J. S., Mega, J. L., Roden, D. M., et al. (2013). Clinical pharmacogenetics implementation consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update. *Clin. Pharmacol. and Ther.* 94 (3), 317–323. doi:10.1038/clpt.2013.105

Sonkin, D., Thomas, A., and Teicher, B. A. (2024). Cancer treatments: past, present, and future. *Cancer Genet.* 286, 18–24. doi:10.1016/j.cancergen.2024.06.002

Stanssens, P., Zabeau, M., Meersseman, G., Remes, G., Gansemans, Y., Storm, N., et al. (2004). High-throughput MALDI-TOF discovery of genomic sequence polymorphisms. *Genome Res.* 14 (1), 126–133. doi:10.1101/gr.1692304

Steyaert, W., Callens, S., Coucke, P., Dermaut, B., Hemelsoet, D., Terryn, W., et al. (2018). Future perspectives of genome-scale sequencing. *Acta Clin. Belg.* 73 (1), 7–10. doi:10.1080/17843286.2017.1413809

Swen, J. J., van der Wouden, C. H., Manson, L. E., Abdullah-Koolmees, H., Blagec, K., Blagus, T., et al. (2023). A 12-gene pharmacogenetic panel to prevent adverse drug reactions: an open-label, multicentre, controlled, cluster-randomised crossover implementation study. *Lancet* 401 (10374), 347–356. doi:10.1016/S0140-6736(22)01841-4

Uddin, M. M., Zhou, Y., Bick, A. G., Burugula, B. B., Jaiswal, S., Desai, P., et al. (2022). Longitudinal profiling of clonal hematopoiesis provides insight into clonal dynamics. *Immun. and Ageing* 19 (1), 23. doi:10.1186/s12979-022-00278-9

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., et al. (2021). Genome-wide association studies. *Nat. Rev. Methods Prim.* 1 (1), 59. doi:10.1038/s43586-021-00056-9

US Food and Drug Administration (2020). *Table of pharmacogenomic biomarkers in drug labeling*. Silver Spring, MD: US Food and Drug Administration. Available online at: https://www.fda.gov/drugs/science-and-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling (Accessed July 2, 2020).

Vassy, J. L., Lautenbach, D. M., McLaughlin, H. M., Kong, S. W., Christensen, K. D., Krier, J., et al. (2014). The MedSeq project: a randomized trial of integrating whole genome sequencing into clinical medicine. *Trials* 15, 85–12. doi:10.1186/1745-6215-15-85

Vassy, J. L., Christensen, K. D., Schonman, E. F., Blout, C. L., Robinson, J. O., Krier, J. B., et al. (2017). The impact of whole-genome sequencing on the primary care and outcomes of healthy adult patients: a pilot randomized trial. *Ann. Intern. Med.* 167 (3), 159–169. doi:10.7326/M17-0188

Wayhelova, M., Smetana, J., Vallova, V., Hladilkova, E., Filkova, H., Hanakova, M., et al. (2019). The clinical benefit of array-based comparative genomic hybridization for detection of copy number variants in Czech children with intellectual disability and developmental delay. *BMC Med. Genomics* 12 (1), 111–11. doi:10.1186/s12920-019-0559-7

Weiss, J. M., Gupta, S., Burke, C. A., Axell, L., Chen, L. M., Chung, D. C., et al. (2021). NCCN guidelines® insights: genetic/familial high-risk assessment: colorectal, version 1.2021. *J. Natl. Compr. Cancer Netw. JNCCN* 19 (10), 1122–1132. doi:10.1164/jnccn.2021.0048

Whirl-Carrillo, M., Huddart, R., Gong, L., Sangkuhl, K., Thorn, C. F., Whaley, R., et al. (2021). An evidence-based framework for evaluating pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* 110 (3), 563–572. doi:10.1002/cpt.2350

Williams, M. S. (2022). Population screening in health systems. *Annu. Rev. Genomics Hum. Genet.* 23, 549–567. doi:10.1146/annurev-genom-111221-115239

Win, A. K., Jenkins, M. A., Dowty, J. G., Antoniou, A. C., Lee, A., Giles, G. G., et al. (2017). Prevalence and penetrance of major genes and polygenes for colorectal cancer. *Cancer Epidemiol. Biomarkers Prev.* 26 (3), 404–412. doi:10.1158/1055-9965.EPI-16-0693

Woerner, A. C., Gallagher, R. C., Vockley, J., and Adhikari, A. N. (2021). The use of whole genome and exome sequencing for newborn screening: challenges and opportunities for population health. *Front. Pediatr.* 9, 663752. doi:10.3389/fped.2021.663752

Yao, R. A., Akinrinade, O., Chaix, M., and Mital, S. (2020). Quality of whole genome sequencing from blood *versus* saliva derived DNA in cardiac patients. *BMC Med. genomics* 13, 11–10. doi:10.1186/s12920-020-0664-7

Zawatsky, C. L. B., Shah, N., Machini, K., Perez, E., Christensen, K. D., Zouk, H., et al. (2021). Returning actionable genomic results in a research biobank: analytic validity, clinical implementation, and resource utilization. *Am. J. Hum. Genet.* 108 (12), 2224–2237. doi:10.1016/j.ajhg.2021.10.005

Zhang, F., Gu, W., Hurles, M. E., and Lupski, J. R. (2009). Copy number variation in human health, disease, and evolution. *Annu. Rev. genomics Hum. Genet.* 10, 451–481. doi:10.1146/annurev.genom.9.081307.164217

Zhang, Z., Ji, Q., Zhang, Z., Lyu, B., Li, P., Zhang, L., et al. (2025). Ultra-sensitive detection of melanoma NRAS mutant ctDNA based on programmable endonucleases. *Cancer Genet.* 294, 47–56. doi:10.1016/j.cancergen.2025.02.008

Zhou, G., Zhou, M., Zeng, F., Zhang, N., Sun, Y., Qiao, Z., et al. (2022). Performance characterization of PCR-Free whole genome sequencing for clinical diagnosis. *Medicine* 101 (10), e28972. doi:10.1097/MD.0000000000028972