



A Comparative Evaluation of Tools to Predict Metabolite Profiles From Microbiome Sequencing Data

Xiaochen Yin¹, Tomer Altman², Erica Rutherford¹, Kiana A. West¹, Yonggan Wu¹, Jinlyung Choi¹, Paul L. Beck³, Gilaad G. Kaplan³, Karim Dabbagh¹, Todd Z. DeSantis¹ and Shoko Iwai^{1*}

¹ Second Genome Inc., Brisbane, CA, United States, ² Altman Analytics LLC, San Francisco, CA, United States, ³ Department of Medicine, University of Calgary, Calgary, AB, Canada

OPEN ACCESS

Edited by:

Stefanie Widder,
Medical University of Vienna, Austria

Reviewed by:

Himel Mallick,
Merck, United States
Cecilia Noecker,
University of California,
San Francisco, United States

*Correspondence:

Shoko Iwai
shoko@secondgenome.com

Specialty section:

This article was submitted to
Systems Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 17 August 2020

Accepted: 16 November 2020

Published: 04 December 2020

Citation:

Yin X, Altman T, Rutherford E,
West KA, Wu Y, Choi J, Beck PL,
Kaplan GG, Dabbagh K, DeSantis TZ
and Iwai S (2020) A Comparative
Evaluation of Tools to Predict
Metabolite Profiles From Microbiome
Sequencing Data.
Front. Microbiol. 11:595910.
doi: 10.3389/fmicb.2020.595910

Metabolomic analyses of human gut microbiome samples can unveil the metabolic potential of host tissues and the numerous microorganisms they support, concurrently. As such, metabolomic information bears immense potential to improve disease diagnosis and therapeutic drug discovery. Unfortunately, as cohort sizes increase, comprehensive metabolomic profiling becomes costly and logistically difficult to perform at a large scale. To address these difficulties, we tested the feasibility of predicting the metabolites of a microbial community based solely on microbiome sequencing data. Paired microbiome sequencing (16S rRNA gene amplicons, shotgun metagenomics, and metatranscriptomics) and metabolome (mass spectrometry and nuclear magnetic resonance spectroscopy) datasets were collected from six independent studies spanning multiple diseases. We used these datasets to evaluate two reference-based gene-to-metabolite prediction pipelines and a machine-learning (ML) based metabolic profile prediction approach. With the pre-trained model on over 900 microbiome-metabolome paired samples, the ML approach yielded the most accurate predictions (i.e., highest F1 scores) of metabolite occurrences in the human gut and outperformed reference-based pipelines in predicting differential metabolites between case and control subjects. Our findings demonstrate the possibility of predicting metabolites from microbiome sequencing data, while highlighting certain limitations in detecting differential metabolites, and provide a framework to evaluate metabolite prediction pipelines, which will ultimately facilitate future investigations on microbial metabolites and human health.

Keywords: metabolome, human microbiome, computational prediction, metabolic potential, Next Generation Sequence

INTRODUCTION

The scientific community has only recently begun to realize and fully appreciate the significant role of the microbiome in human health (Turnbaugh et al., 2007; Integrative HMP (iHMP) Research Network Consortium, 2014). Increased access to high-throughput sequencing technologies has facilitated a record number of metagenomic- and metatranscriptomic-based investigations of host tissues and the microbial communities they support, which have begun to shed light on the pivotal

impact of this ecosystem's eubiosis in human health. Omics technologies empower mechanistic and therapeutic discovery relating to disease onset, progression, and treatment (Metwally and Haller, 2019; Zhang et al., 2019). Small molecules are key factors in all host-microbe interactions, which can be synthesized, metabolized, and even modified by specific microbial taxa. The downstream effects of such metabolite modulation have been implicated in biological processes germane to human health (Zierer et al., 2018; Descamps et al., 2019). For example, researchers have shown that the microbial metabolite trimethylamine-N-oxide (TMAO) is a predictive marker of cardiometabolic diseases (Ussher et al., 2013; Ufnal et al., 2015; Li et al., 2018), as have secondary bile acid metabolites on immune system homeostasis and glucose and lipid metabolism (Molinerio et al., 2019; Song et al., 2019), and microbial-derived gamma-aminobutyric acid (GABA) as a neurotransmitter of the central nervous system (Bravo et al., 2011; Strandwitz et al., 2019; Zheng et al., 2019).

Untargeted metabolomic analyses greatly facilitate the detection and characterization of a wide range of metabolites, affording researchers a comprehensive understanding of the metabolic pathways invoked within a microbial community. Such techniques have bolstered and accelerated mechanistic studies and biomarker identification strategies across a variety of diseases (Li et al., 2016; Franzosa et al., 2019; Glinton and Elsea, 2019; Urpi-Sarda et al., 2019). With massively large amounts of raw microbiome sequencing data being deposited into public sequence repositories at ever-increasing rates, we hypothesized that it is possible to predict metabolic profiles based solely on the sequencing data from a microbial community. After all, an accurate and reliable *in silico* means of predicting metabolic capacity from nucleic acid sequences would embolden drug discovery by generating testable hypotheses *sans* cost-prohibitive upstream metabolome profiling analyses.

Recent years have seen advances in linking microbiome sequencing data to metabolome data. One such strategy relies on the network of connections linking a given gene to reactions and compounds in a database. These linkages are used to infer molecular compound identities from the genetic information housed within a microbial community. A method called predicted relative metabolomic turnover (PRMT) was used to predict metabolites from a coastal marine metagenomics dataset, and the predicted metabolites correlated strongly with environmental factors (Larsen et al., 2011). MIMOSA was later developed to predict metabolic potential in a given microbial community and identify the microbial taxa most responsible for the synthesis/consumption of key metabolites (Noecker et al., 2016). Capitalizing on plentiful abundances of gene-to-metabolite data housed in repositories like KEGG, these utilities promote the generation of testable hypotheses and identification of potential drug targets (Chang et al., 2019). MIMOSA has been successfully applied in a number of studies to identify the microbial origin of certain metabolites (Stewart et al., 2018; Sharon et al., 2019). Meanwhile, interested in metabolites that directly associate with genes regardless of the reaction network and not limited to the KEGG database, we developed Mangosteen: a metabolome prediction pipeline

dependent upon relationships between KEGG/BioCyc reactions and the molecular compounds directly associated with those reactions. Both MIMOSA and Mangosteen are reference-based, and as such, they rely heavily on the completeness and accuracy of the database queried. As the vast majority of microbial taxa belonging to the human gut microbiome remain unknown (Turnbaugh et al., 2007; Sunagawa et al., 2013), predictions from these reference databases provide a partial view of the metabolic capacity housed within a community.

Mallick et al. (2019) devised MelonnPan, which exploits machine learning (ML) to predict metabolomic potential. *Via* MelonnPan, a model trained from paired microbiome and metabolome datasets can be used to predict metabolites from a novel microbiome dataset *san a priori* knowledge regarding relationships between genes and metabolites. This approach circumvents the limitations of the reference-based methods discussed above and has been used to generate promising results between two inflammatory bowel disease (IBD) cohorts (Franzosa et al., 2019; Mallick et al., 2019).

Despite the development and refinement of these pioneering pipelines, to date there has not been a thorough comparison of reference-based and ML-based techniques. Ergo, we comparatively analyzed the performance of two reference-based methods, Mangosteen and the compound prediction component of MIMOSA, and the ML-based MelonnPan approach in conducting microbiome metabolite predictions (Figure 1). A detailed evaluation was performed on occurrence, abundance, and between-group differences against metabolome data acquired *via* empirical measurements (Figure 1B).

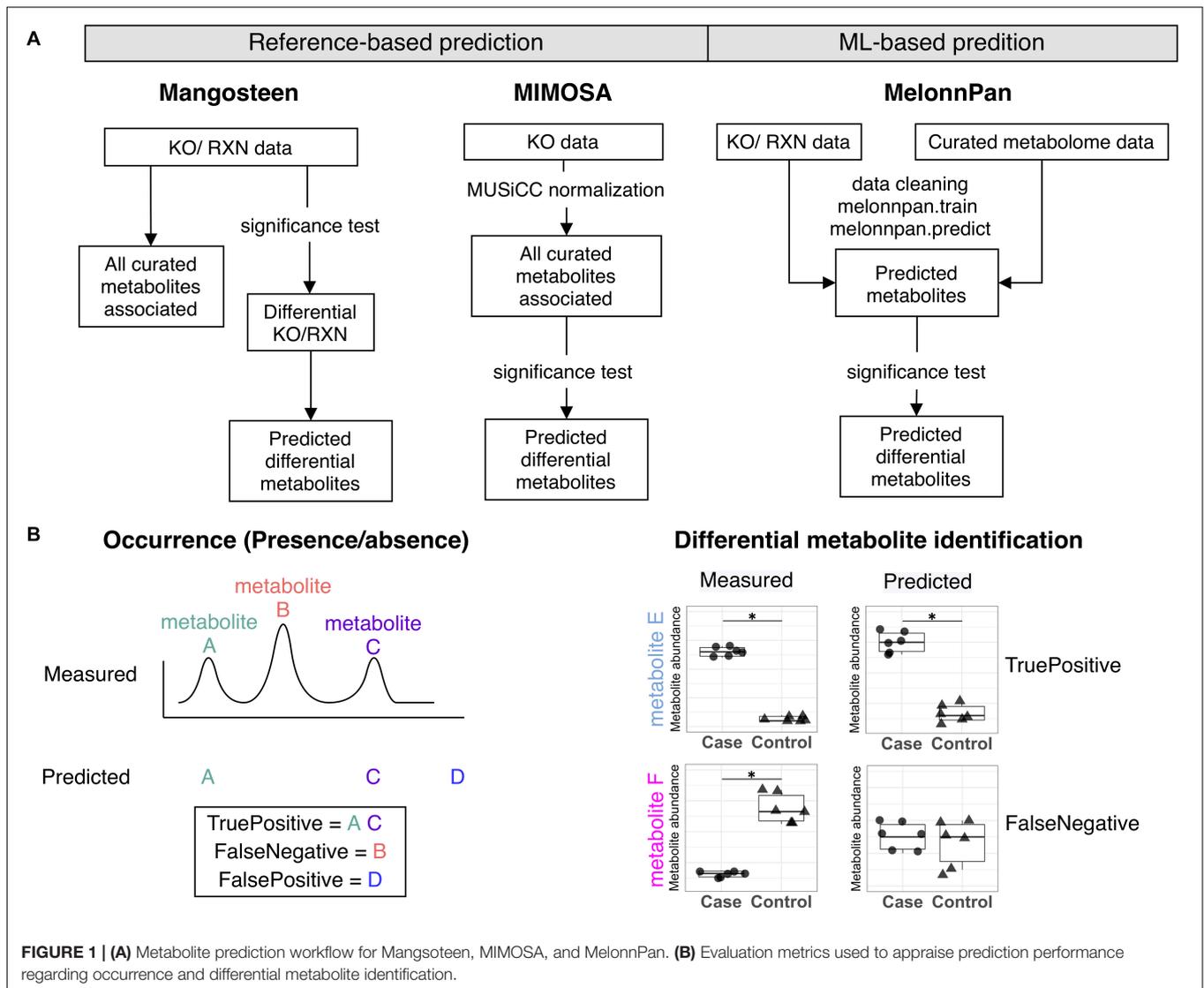
MATERIALS AND METHODS

Data Collection

A PubMed search querying the keywords “gut microbiota” OR “gut microbiome” OR “fecal microbiota” OR “fecal microbiome” AND “metabolome” AND “humans” resulted in 171 papers (Supplementary Figure 1). Filtering was applied to remove studies that (1) were not original research articles, (2) were conducted *in vitro* or in newborn and/or infant subjects, (3) consisted of samples other than stool or intestinal tissue, (4) had microbiome data other than NGS sequencing, (5) had targeted metabolomics data, and (6) had fewer than 30 microbiome-metabolome paired samples. Notably, as different metabolome generation and processing methods have significant impact on results, we only selected studies which applied untargeted metabolome profiling with multiple liquid chromatography-mass spectrometry (LC-MS) methods or nuclear magnetic resonance (NMR) for a comprehensive view of the metabolite pool. Eighteen studies were retained and six were used in this evaluation due to data accessibility (Table 1).

Microbiome Sequencing Data Pre-processing

Raw sequencing data were either downloaded from NCBI Sequence Read Archive (SRA) for public studies or generated in-house using the Illumina platform (Table 1). We analyzed



16S rRNA amplicon Illumina sequencing data using DADA2 (Callahan et al., 2016) and subjected to functional composition prediction *via* Piphillin with identity set at 0.97 (Iwai et al., 2016; Narayan et al., 2020). For 16S rRNA amplicon pyrosequencing data, we compared it to StrainSelect (strainselect.secondgenome.com) using USEARCH (Edgar, 2010). We assigned a strain operational taxonomic unit (OTU) to sequences matching a unique strain with a global alignment identity $\geq 99\%$ and with the highest identity to a single strain. To ensure specificity of these strain matches, a difference $\geq 0.25\%$ between the identity of the best and second-best match was required (e.g., 99.75 vs. 99.5). All remaining non-strain sequences were quality filtered and dereplicated with USEARCH. We then clustered the resulting unique sequences at 97% using UPARSE (*de novo* OTU clustering) and determined a representative consensus sequence per *de novo* OTU. Piphillin (Iwai et al., 2016; Narayan et al., 2020) was then applied to the OTU table to infer community function.

For metagenomic and metatranscriptomics reads, adapter sequences and low-quality ends were trimmed with Trimmomatic ($<Q20$; Bolger et al., 2014). We then removed contaminant sequences, e.g., PhiX174 and sequencing primers, using Bowtie2 (Langmead and Salzberg, 2012). For metatranscriptomic data, all rRNA sequences from all three domains of life were identified and removed from consideration using SortMeRNA 2.0 (Kopylova et al., 2012). Host sequences were omitted using Kraken (Wood and Salzberg, 2014), which used exact matches of raw shotgun sequences to k-mers derived from the human reference genome. Filtered DNA sequences were mapped against a custom database built from KEGG (May 2019 release; Kanehisa et al., 2011) and BioCyc (version 23.0; Karp et al., 2017). Specifically, we collected protein sequences of bacteria, fungi and viruses from KEGG as well as protein sequences in MetaCyc and all PGDBs in BioCyc followed by a de-replication step at 0.99 identity and 0.99 alignment coverage using CD-HIT (Li and Godzik, 2006; Fu et al., 2012). A search

TABLE 1 | Characteristics of datasets included for prediction and evaluation.

Disease area – sample type	Study	Treatment group (counts of biospecimens in each group)	Metabolome dataset		Microbiome dataset		Contrast for differential analysis
			Profiling technology	Target region	Profiling technology	Data availability	
Healthy – Stool samples	Maier et al., 2017	High resistant starch intervention (14) Low resistant starch intervention (14) Baseline control (13)	SolariX Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR-MS; Bruker Daltonik GmbH)	16S rRNA gene: V4-V6	HiSeq 2000 (Illumina)	EBI-ENA accession: ERP104494	High resistant starch intervention: Baseline control
Colorectal cancer (CRC) – Biopsy samples	Hale et al., 2018	Diseased tissue (28) Healthy tissue distal to diseased tissue (35) Healthy tissue proximal to diseased tissue (20)	Quadruple time-of-flight mass spectrometer (Agilent Technologies 6550 Q-TOF)	16S rRNA gene: V3-V5	MiSeq (Illumina)	NCBI SRA BioProject: PRJNA445346	Diseased tissue: Healthy tissue distal to diseased tissue
Autism spectrum disorder (ASD) – Stool samples	Kang et al., 2018	ASD (23) Healthy control (21)	Varian Direct Drive (VNMRS) 600 MHz spectrometer (Agilent Technologies)	16S rRNA gene: V2-V3	Genome Sequencer FLX-Titanium System (Roche)	Qitta: study ID 11169	ASD: Healthy control
Inflammatory bowel disease (IBD) – Stool samples	Franzosa et al., 2019	Crohn's disease (88) Ulcerative colitis (76) Non-IBD Control (56)	Q Exactive Hybrid Quadrupole-Orbitrap mass spectrometer; Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific)	Whole genome	HiSeq 2500 (Illumina)	NCBI SRA BioProject: PRJNA400072	Crohn's disease: Non-IBD Control
	Lloyd-Price et al., 2019	Crohn's disease (139) Non-IBD Control (86)	Q Exactive/Exactive Plus orbitrap mass spectrometers (Thermo Fisher Scientific)	Whole genome	HiSeq2000; HiSeq 2500 (Illumina)	NCBI SRA BioProject: PRJNA398089	Crohn's disease: Non-IBD Control
	SG_IBD (generated in this study)	Active ulcerative colitis (10) Inactive ulcerative colitis (19) Healthy control (15)	Q Exactive orbitrap mass spectrometers (Thermo Fisher Scientific)	Whole transcriptome	HiSeq2500 (Illumina)	NCBI SRA BioProject: PRJNA668188	Active ulcerative colitis: Healthy control
				Whole transcriptome	NextSeq 550 (Illumina)		

for translated DNA sequences was executed using Diamond (Buchfink et al., 2014) and hits that spanned ≥ 20 amino acids with $\geq 80\%$ similarity were retained. Upon identifying multiple hits, reads were equally split between the best hits.

Metabolite Identifier Curation

Detailed LC-MS/NMR conditions, software and library databases for metabolite data generation per each study are listed in **Supplementary Table 1**. These annotated metabolites were assigned unique identifiers corresponding to an in-house chemical dictionary so as to facilitate comparison between studies. KEGG, BioCyc, the PubChem Identifier Exchange Service¹, and the Chemical Translation Service² were used to identify and convert metabolites.

Metabolite Prediction Using Mangosteen

A relationship table linking KEGG orthologs (KO) and BioCyc reactions (RXNs) to the compounds housed in KEGG (release date 2019 May) and BioCyc (version 23.0) databases was

compiled. Specifically, if KOs have corresponding KEGG reactions which produce or consume KEGG compounds, we directly linked those KOs to KEGG compounds. An example is that K24443 corresponds to two reactions R02428, R02526, which produce or consume compounds C02753, C00001, C00502, C01114, and C00545. Thus, K24443 is linked to those compounds. There are cases when KOs do not have corresponding reactions but KEGG Enzyme numbers, and we would link those enzymes to reactions and further to compounds. An example is that K00046 has no corresponding reactions but the enzyme number EC:1.1.1.69, through which we linked to reactions R01738, R01740 and the reactions produce or consume compounds C00257, C00003, C01062, C00004, C00080, C00006, C00005, thus those compounds are linked to K00046. Manual effort was also made to link transporter KOs (BRITE class: ko02000) to their associated compounds. For the BioCyc database, RXNs and their associated compounds were inferred from “core_description_of_RXNs” directly, which included reactions from both MetaCyc and all PGDBs in BioCyc. This collection of reactions is referred to as “BioCyc reactions” in the remainder of the text. Reaction directionality was not taken into consideration when constructing the relationship table

¹<https://pubchem.ncbi.nlm.nih.gov/idxexchange/idxexchange.cgi>

²<https://cts.fiehnlab.ucdavis.edu/batch>

because many biochemical reactions are reversible depending on the reactant and product concentrations.

We inferred all metabolites that linked to KOs or RXNs based on the relationship table for coverage evaluation. If a KO or RXN links to multiple metabolites, all were reported. For differential metabolite prediction, we first identified differential microbial functions (KOs/RXNs) between case and control samples *via* DESeq2 analysis (adjusted *p*-value threshold = 0.2; Love et al., 2014) following prevalence filtering at 5%. The list of differential microbial functions was used as the input for Mangosteen and all compounds linked to them were reported to be differential compounds.

Metabolite Prediction Using MIMOSA

Using the source code from MIMOSA³, we updated the community metabolic network template to the May 2019 version of the KEGG database. Microbiome functional abundance data were normalized by means of MUSiCC using default settings (Manor and Borenstein, 2015), as is recommended prior to applying MIMOSA for prediction. We then calculated the community-metabolic-potential (CMP) of each sample, which served as the predicted metabolome.

Metabolite Prediction Using MelonnPan

In lieu of applying the default model included in the MelonnPan package (Mallick et al., 2019), we constructed a model based on metabolome and microbial function abundance data from six previous investigations (Table 1). Because multiple microbiome sequencing datatypes exist for two studies (shotgun metagenomic and metatranscriptomic data in Lloyd et al. study; Lloyd-Price et al., 2019; 16S rRNA gene amplicon and metatranscriptomic data in SG_IBD study), we chose to leave-one-study-out cross-validation to minimize potential overfitting for MelonnPan. We retained all metabolites and microbial KOs or RXNs with over 10% prevalence and mean relative abundance over 0.01% in accordance with MelonnPan operating procedures. From the model that resulted, we mitigated the potential for spurious results by omitting metabolites predicted by only one KO or RXN from consideration (model size = 1).

Evaluation Metric

Predicted community metabolite profiles were compared to empirically measured metabolite profiles, and both occurrence and abundance were appraised. We then calculated the precision, recall, and F1 score for each dataset per pipeline. Occurrence evaluation only considers the presence or absence of a metabolite in a given microbial community. Abundance evaluation considers both the results of Procrustes analyses conducted across the predicted and measured abundance matrices (Peres-Neto and Jackson, 2001) and the identification of differential metabolite described below.

Prevalence filtering was applied to measured metabolome data, as well as MIMOSA and MelonnPan predicted metabolome data, and metabolites detected or predicted in fewer than 5% of the samples were omitted from consideration. Data resulting

from empirically measured- and MelonnPan predicted-profiles were log transformed after imputing zeros with the minimum non-zero value per metabolite. We then conducted Student's *t*-tests to identify differential metabolites between case and control samples, followed by Benjamini-Hochberg FDR correction (adjusted *p*-value threshold = 0.2). As the abundance matrix generated with MIMOSA includes both positive and negative values, we elected to apply Wilcoxon rank sum tests with Benjamini-Hochberg FDR correction (adjusted *p*-value threshold = 0.2). Wilcoxon signed rank tests were used to comparatively evaluate precision, recall, and F1 scores across pipelines and databases between paired datasets.

Random sampling analysis was also performed and compared to the pipeline prediction. For coverage evaluation, the same number as the predicted metabolites were randomly selected from the in-house chemical dictionary and compared with the pipeline predicted results. For differential metabolite identification, gene labels were randomly shuffled prior to DESeq2 and used for prediction for Mangosteen. For MIMOSA and MelonnPan, predicted metabolite labels were shuffled in the predicted matrices, followed by the same procedure for differential metabolite identification. The results were further evaluated against the measured metabolites and performance was represented in precision, recall and F1 score. The procedure was repeated for 99 times.

RESULTS

Selected Studies Include Diverse Metabolome and Microbiome Sequencing Technologies

Among the 6 studies that fit our selection criteria as described in Supplementary Figure 1, we considered the datasets from three IBD studies (Franzosa et al., 2019; Lloyd-Price et al., 2019), one colorectal cancer (CRC) investigation (Hale et al., 2018), one autism spectrum disorder study (Kang et al., 2018), and one dietary intervention (Maier et al., 2017; Table 1). Many of the samples originated from stool (*n* = 648 from five studies), while a smaller subset (*n* = 83) arose from intestinal tissue samples of the CRC study. All samples were subjected to metabolomic analysis *via* either mass spectrometry (MS; five studies) or NMR spectroscopy (NMR; one study). For microbiome profiling, half of the datasets originate from 16S rRNA gene amplicon sequencing (four datasets), followed by shotgun metagenomic analysis (two datasets), and metatranscriptomic sequencing (two datasets).

Pipeline Paradigm Determines Metabolite Prediction

Empirically measured metabolome profiles from each of the aforementioned datasets were compared directly with pipeline predictions. Only peaks corresponding to known compounds were retained, which resulted in a total of 1,998 metabolites spanning six studies. We observed significant differences in metabolite numbers across different metabolomics

³<https://github.com/borenstein-lab/MIMOSA>

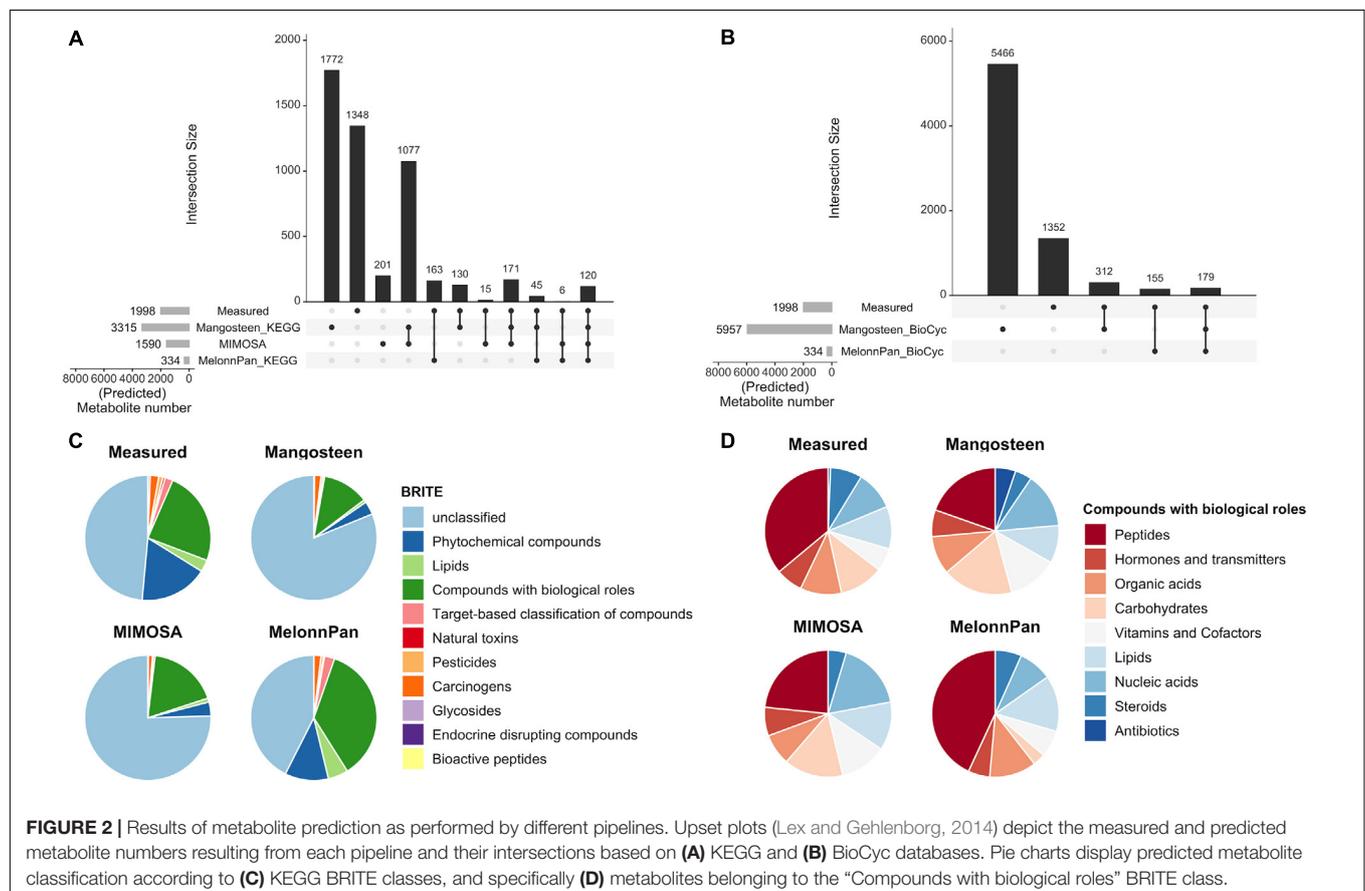
technologies. Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR-MS) identified the greatest number of metabolites ($n = 1,273$) and NMR the least ($n = 62$; **Supplementary Figure 2A**). Of all metabolites detected, 1,887 corresponded to KEGG-specific compound identifiers and 987 to BioCyc-specific compound identifiers. Quality-controlled metagenomic and metatranscriptomic data were mapped against KEGG and BioCyc databases and Piphillin (Iwai et al., 2016; Narayan et al., 2020) was applied to 16S rRNA sequencing data to infer communities' functional profiles. Ultimately, we identified 11,527 KOs and 8,427 RXNs across eight datasets spanning six studies (**Supplementary Figure 2B**), which collectively served as the input for each prediction pipeline.

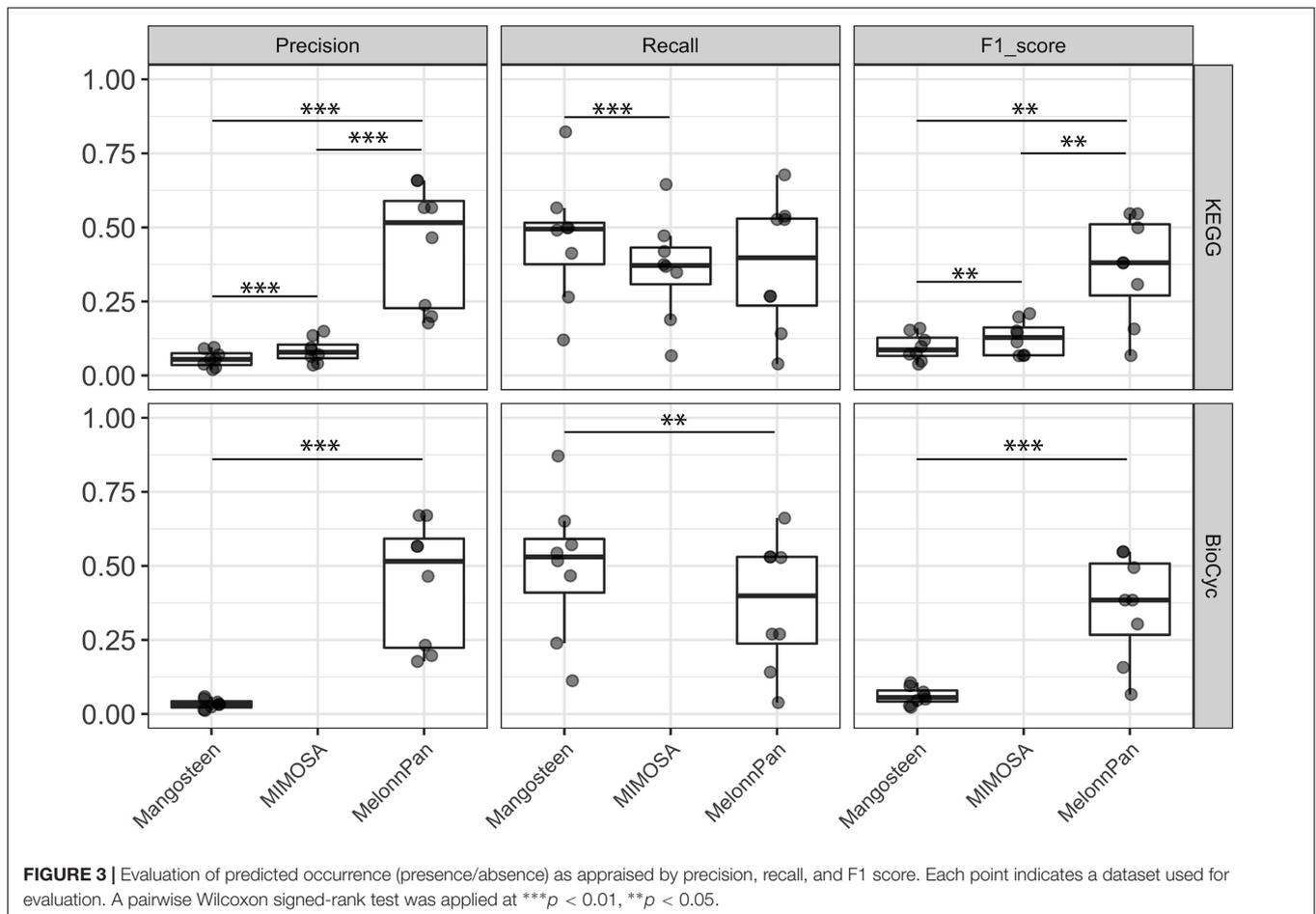
The Mangosteen pipeline identified 3,315 KEGG- and 5,957 BioCyc-associated metabolites across all studies (Mangosteen-K and Mangosteen-B, respectively). While Mangosteen-B predicted significantly fewer unique microbial functions than Mangosteen-K (**Supplementary Figure 3A**), it linked to significantly more metabolites (**Supplementary Figure 3B**). MIMOSA predicted a total of 1,590 metabolites using the KEGG database, 1,077 of which were shared with Mangosteen-K (**Figure 2A**). In comparison, the ML-based MelonnPan pipeline predicted only 334 metabolites upon interrogating models built on microbial functions mapped to KEGG (MelonnPan-K) or BioCyc (MelonnPan-B; **Figures 2A,B** and **Supplementary Figure 4**).

Metabolites bearing KEGG identifications were then assigned to their corresponding BRITE classifications (Kanehisa et al., 2011). Melonnpan-K predicted significantly greater fractions of metabolites deemed “Compounds with biological roles” (BRITE ID: 08001), “Phytochemical compounds” (08003), “Lipids” (08002), and “Target-based classification of compounds” (08010) than the reference-based approaches (**Figure 2C**). Considering only the “Compounds with biological roles” category, MelonnPan-K predicted the greatest fractions of metabolites resembling “Peptides,” “Organic acids,” “Lipids,” and “Steroids” and smallest fractions of those resembling “Nucleic acids,” “Vitamins and Cofactors,” “Hormones and transmitters,” and “Carbohydrates” (**Figure 2D**). Mangosteen-K was the only pipeline to predict metabolites belonging to the “Antibiotics” category of classification.

ML-Based MelonnPan Yielded the Highest F1 Score in Predicting Metabolite Occurrence

Predicted metabolite profiles were compared to empirically measured metabolite profiles. MelonnPan exhibited the greatest precision (MelonnPan-K: mean 0.44, MelonnPan-B: mean 0.44) in predicting the presence/absence of metabolites in the human gut, followed by MIMOSA (0.08) and Mangosteen (Mangosteen-K: 0.06, Mangosteen-B: 0.03; **Figure 3**). Recall





was significantly greater with Mangosteent-K, as opposed to MIMOSA, and with Mangosteent-B as opposed to MelonnPan-B. Taking into consideration both precision and recall, F1 scores demonstrated the superiority of ML-based MelonnPan predictions (MelonnPan-K: 0.36, MelonnPan-B: 0.36). KEGG-referenced Mangosteent exhibited a small but significant increase in both precision (Mangosteent-K vs. Mangosteent-B: mean 0.06 vs. 0.03, $p = 0.01$) and overall F1 scoring (0.1 vs. 0.06, $p = 0.01$) compared to its BioCyc-referenced counterpart. With respect to ML-based MelonnPan, these two reference databases used for model building performed equally well and yielded no significant differences in precision, recall, or F1 score. We also compared the prediction results to random sampling, which randomly predicted the same number of metabolites, and all pipelines showed significantly better performance regarding precision, recall and F1 score ($p < 0.01$).

None of the Predicted Metabolite Profiles Retain High Levels of Similarity to the Empirically Measured Metabolome

As both MIMOSA and MelonnPan output metabolite abundance predictions, we examined the extent of similarity between the predicted and empirically measured metabolite abundance

profiles. Euclidean distance from the predicted and measured metabolomes were calculated to facilitate Procrustes analyses (Peres-Neto and Jackson, 2001). Very little similarity was observed between the metabolomes (correlation coefficient range: 0.04–0.26, $p > 0.05$), with squared m12 values (a measure of fit between two datasets; low value indicates high similarity; Peres-Neto and Jackson, 2001) ranging from 0.93 to 1 (Supplementary Figure 5). With respect to Procrustes analysis, MIMOSA- and MelonnPan-generated prediction profiles did not significantly differ when compared to measured metabolome, and neither did KEGG-referenced and BioCyc-referenced MelonnPan predictions.

All Pipelines Performed Poorly in Identifying Differential Metabolites

We evaluated these pipelines on their ability to identify differentially abundant metabolites between case and control groups as shown in Table 1. The predicted differential metabolites identified in each pipeline were compared to differential metabolites identified in experimental data. ML-based MelonnPan predictions yielded the highest precision and F1 scores (but still low and not significant) compared to the two reference-based prediction methods (MelonnPan-K Precision mean: 0.11, F1 score: 0.11; MelonnPan-B: Precision:

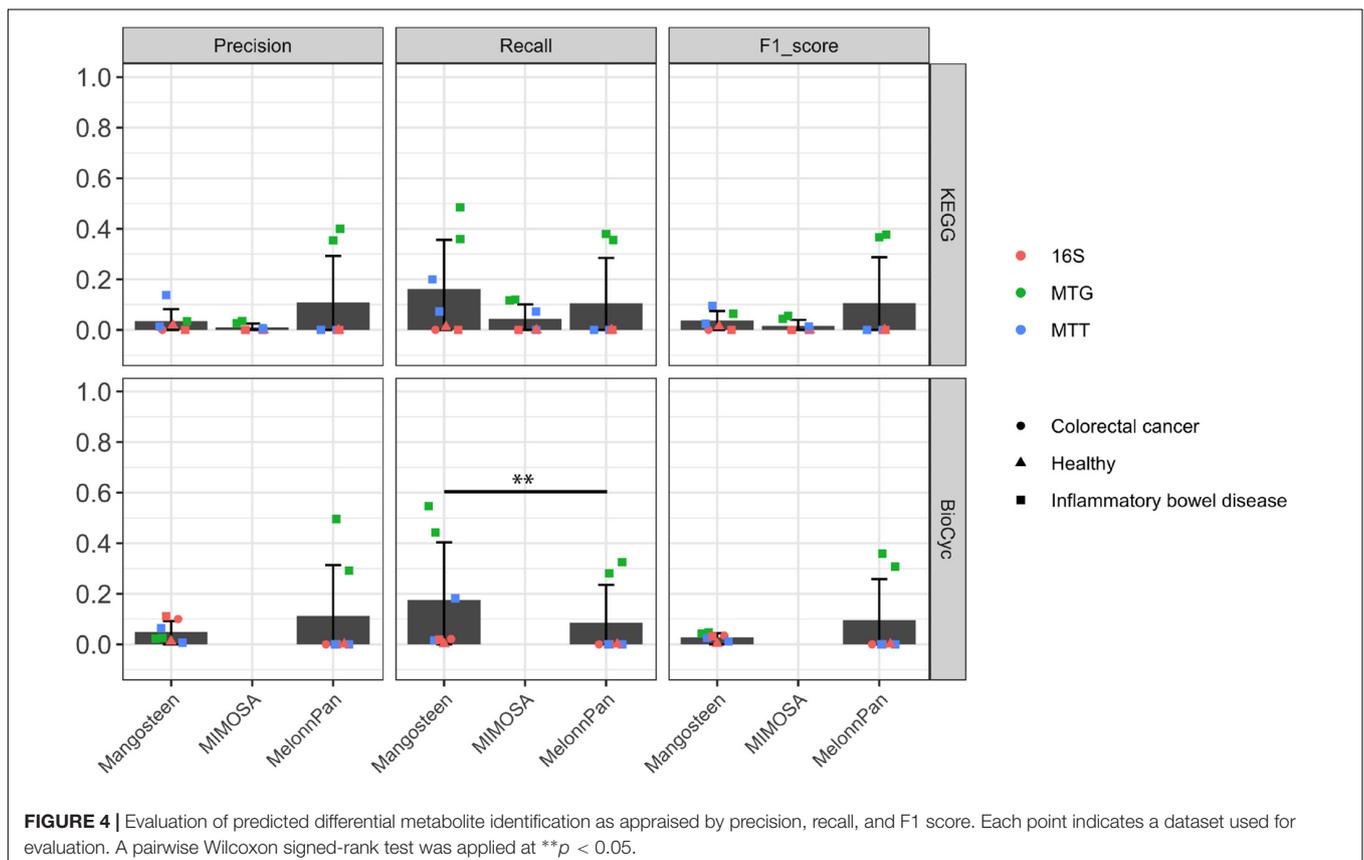
0.11, F1 score: 0.10; **Figure 4**). We also compared the results to random sampling, where differential metabolites were identified from either shuffled microbial abundance table (for Mangosteen) or predicted metabolite abundance tables (for MIMOSA and MelonnPan). Despite of the overall low F1 score, Mangosteen showed significantly higher F1 scores ($p < 0.05$) compared to random prediction while MIMOSA and MelonnPan did not (MIMOSA: $p = 0.2$, MelonnPan-K: $p = 0.37$, and MelonnPan-B: $p = 0.42$).

DISCUSSION

Gaining insight into the metabolite pool of a microbial community is of paramount importance to understand its ecological role(s), not to mention its potential as a source of therapeutics and other invaluable molecules (Jia et al., 2008; Patel and Ahmed, 2015; Wishart, 2016). As an alternative to cost- and resource-prohibitive full-scale metabolome profiling, metabolome prediction from *a priori* microbiome sequencing data affords researchers the ability to generate hypotheses in a rapid, cost-effective, and relatively reliable fashion. We systematically compared the ability of metabolite predictions by two published tools alongside a newly developed reference-based pipeline using more than 900 paired microbiome-metabolome stool/intestinal tissue samples from six different studies on various human diseases. Resulting metabolite profile predictions

were compared to one another and contrasted alongside an empirically measured metabolome dataset, focusing primarily on occurrence (i.e., presence vs. absence) and identification of differentially abundant metabolites.

When used to predict the presence or absence of given metabolites within a community, the ML-based MelonnPan approach significantly outperformed its reference-based competitors, with respect to overall precision and F1 score. However, the reference-based Mangosteen and MIMOSA pipelines did exhibit high recall despite their lower precision and F1 scores. As recall is calculated based on true-positive and false-negative tallies, the high number of metabolites predicted *via* these reference-based pipelines likely contributed to the elevated recall values. In addition, any given KO may participate in any number of distinct reactions, bearing the potential to dramatically increase the number of associated metabolites. With regard to reference-based methods, MIMOSA surpassed Mangosteen-K in both precision and F1 score. This speaks to the benefit of considering the directionality as well as the network of connections between reactions within a community. Although Mangosteen showed worse performance in metabolite prediction compared to MIMOSA, it was able to link more metabolites and could be useful for mining metabolites outside of known metabolic network. While the type of microbiome data examined also affects prediction performance, we lacked an appropriate number of independent datasets to evaluate this aspect in our study (one study for metagenomic to metatranscriptomic



and another for 16S rRNA amplicon to metatranscriptomic sequencing comparisons). A pipeline's ability to accurately distinguish differential metabolites between case and control groups is typically a sound foundation from which hypotheses are structured and biomarker and therapeutics discovery initiatives expound. Such evaluations are predicated upon accurate predictions of metabolite abundance, and as such, are much more stringent than occurrence prediction. While we did not observe promising results in any of the pipelines, MelonnPan predictions yielded the highest mean F1 scores (around 0.1 in both MelonnPan-K and MelonnPan-B).

Machine-learning-based MelonnPan outperformed the reference-based methods in predicting both occurrence and differential metabolites. In contrast to the reference-based pipelines, this powerful technique considers the impact of host metabolism and hitherto undiscovered interactions that do not exist in reference databases. Like all ML strategies, MelonnPan relies heavily on the accuracy of training data. As such, while untargeted metabolome profiling strategies generate the most comprehensive overview of the metabolites present within a sample, coverage is unavoidably limited due to the inherent technical biases associated with either MS or NMR (Christians et al., 2011; Vuckovic, 2012). Hence, we were mindful to include metabolome data generated by multiple platforms to expand and diversify the training set and thus improve coverage of human gut-associated metabolites. In the current evaluation, although we tried our best to include as many paired samples as possible (over 900 pairs) to train the ML model, variance in prediction performance was observed across studies, which suggests more training sets are still needed to obtain a general and robust prediction model.

While conducting the differential metabolite evaluation, we observed the highest F1 score in the MelonnPan-predicted metagenomic dataset of Franzosa et al. (2019) as well as the metagenomic dataset of Lloyd-Price et al. (2019). Similar results were observed while evaluating MelonnPan with datasets generated in three independent IBD studies (for a disease-specific model; **Supplementary Figure 6**). Of all the datasets evaluated, these two were generated in the most similar manner, i.e., metabolome data were obtained from 4 LC-MS methods (targeting polar metabolites, metabolites of intermediate polarity, and lipids) while microbiome metagenomic sequencing data were generated using an Illumina HiSeq platform. Numerous factors and processes contribute significantly to variances observed between studies (e.g., sample preparation, instrument settings, and user variation) in both metabolome and microbiome sequencing-based investigations (Gika et al., 2010; Tulipani et al., 2013; Clooney et al., 2016). Thus, the high F1 scores observed in Franzosa et al. and Lloyd-Price et al.'s metagenomic datasets are likely due to similar data generation techniques. Furthermore, it stands to reason that while one dataset is used in a training model, prediction for a new study of similar datatypes is more accurate. This could also explain the superior performance reported in Mallick et al. (2019) MelonnPan paper, wherein training and testing datasets were generated using the exact same microbiome sequencing and metabolomic technologies. We also noticed metatranscriptomic-based prediction performed worse than metagenomic-based prediction for differential metabolite

identification. This again could be due to the dataset type in the training set because we observed higher RTSI scores (a measurement to quantify the representativeness of new samples with respect to training datasets from MelonnPan; the higher the value, the more accurate prediction is; Mallick et al., 2019) when using metagenomic data for prediction compared to metatranscriptomics in the Lloyd-Price et al.'s study. The choice of training datasets will also impact the computing demand for ML-based methods and should be taken into consideration. From our experiences, MelonnPan took average 24 CPU hours to train with the current datasets (~900 paired samples) while the prediction step finished in less than 10 min, similar to the reference-based MIMOSA and Mangosteen prediction. Ultimately, the training time could be eliminated if one general model is built and used repeatedly for prediction. This information is extremely useful for future metabolite prediction applications using ML-based approaches like MelonnPan and reminds the user of the importance of choosing appropriate training and prediction datasets with the goal of optimizing downstream prediction indices. With more and more metabolomic data being deposited into public domains, the ability to construct either datatype depend or independent ML models will expand rapidly and researchers will be able to choose appropriate datasets for training and prediction. Moreover, advances in untargeted metabolomic technologies with broader coverage, superior resolution, and greater accuracy (Ghaste et al., 2016; Ortmayr et al., 2016; Týčová et al., 2017; Bingol, 2018) will collectively lead to significantly improved training data, and in turn, more accurate ML-based predictions.

As opposed to ML-based and data-driven approaches, reference-based methods can potentially identify/predict metabolites whose concentration or specific chemical properties precluded them from detection *via* conventional untargeted metabolome profiling as well as ML-based prediction (Karu et al., 2018). Reference-based prediction pipelines make use of well-curated information in massively large repositories to infer possible metabolites within a sample. As such, the integrity and comprehensiveness of the database to be queried factors significantly into the process. KEGG (Kanehisa et al., 2011) and BioCyc (Caspi et al., 2011) are two widely used databases that house curated metabolic pathway information from the annotated genomes. With regard to Mangosteen predictions, we linked significantly more metabolites from RXNs collected from BioCyc than the KEGG database (**Supplementary Figure 3**), but this did not help BioCyc-based prediction to achieve greater F1 scores. The greater number of observed compound connections with BioCyc is consistent with previous reports showing BioCyc houses a greater number of compounds associated with reactions ("all reaction substrates") than KEGG (Altman et al., 2013). There are also many assumptions germane to reference-based prediction that are not accurate when describing the human gut microbiome (e.g., gene and/or transcript abundance is reflective of enzymatic activity, microbial communities are well-mixed, steady-state systems *sans* compartmental effects; Kumar et al., 2019). Nevertheless, *de novo* metabolite prediction is only one possible use of these reference-based tools, as the full functionality of MIMOSA is to identify metabolites which show strong evidence of their microbial producers/consumers

and Mangosteen is better at identifying all linked metabolites when given a microbial feature (KO or RXN), either differential ones between case and control or any features the researchers are interested in. Ultimately, caution and common sense go a long way in interpreting the predicted metabolite profiles resulting from reference-based pipelines.

There are several limitations in the current evaluation that need to be taken into consideration when researchers try to choose the optimal prediction tool for their research. While we treated the empirically measured, untargeted metabolome as a “gold standard” and directly compared all pipeline prediction results to this profile, biases in technology render even this portrayal imperfect in its attempt to unveil the true and complete metabolite pool (Christians et al., 2011; Vuckovic, 2012; Karu et al., 2018). In reality, predicted metabolites that were treated as false positives might not actually be false, and in turn, the actual performance of these prediction pipelines might supersede what we report here. Additional profiling of these “false positive” metabolites or technological advances in untargeted metabolomics will help to mitigate the issue on “false positive” predictions. In addition, although we tried to select the metabolome data in a uniform way, there are variations in extraction methods, chromatography conditions, model of the machines as well as software and databases between studies (**Supplementary Table 1**), which are currently not considered during evaluation. Secondly, the current evaluation only focuses on the human microbiome from stool and intestinal tissues, the performance of these software on other types of microbiome data, such as environmental samples, could vary and warrants further evaluation efforts. Moreover, emerging tools that integrate microbiome and metabolome data could add more interpretability beyond metabolites prediction and help researchers to understand the origin of metabolites and its association with microbiome, such as Annotation of Metabolite Origins via Networks (AMON; Shaffer et al., 2019), neural network based mmvec (Morton et al., 2019), MIMOSA2 (MIMOSA2, 2020), Generalized correlation analysis for Metabolome and Microbiome (GRaMM; Liang et al., 2019), etc.

Predicting metabolic capacity and metabolite diversity is of paramount consequence to better understanding and manipulating the function(s) of microbial communities, all of which bears immense significance to improving human health and empowering environmental microbiome research (Turnbaugh and Gordon, 2008; van Dam and Bouwmeester, 2016; Van Treuren and Dodd, 2020). Ultimately, this evaluation serves as a framework and launching point for future initiatives interested in exploiting metabolite prediction as a cost-effective means of generating cogent hypotheses.

REFERENCES

- Altman, T., Travers, M., Kothari, A., Caspi, R., and Karp, P. D. (2013). A systematic comparison of the MetaCyc and KEGG pathway databases. *BMC Bioinformatics* 14:112. doi: 10.1186/1471-2105-14-112
- Bingol, K. (2018). Recent advances in targeted and untargeted metabolomics by NMR and MS/NMR methods. *High Throughput* 7:9. doi: 10.3390/ht7020009

DATA AVAILABILITY STATEMENT

Metadata, 16S rRNA gene amplicon and metatranscriptomic sequencing data generated in this study (SG_IBD) have been deposited in the NCBI database under BioProject ID PRJNA668188.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Calgary, Conjoint Health Research Ethics Board (ID:18142 and 14-2429). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SI, TD, and KD conceived and supervised the project. XY, TA, TD, and SI designed the study. XY, ER, KW, YW, JC, PB, and GK performed data collection and processing. XY conducted the analysis and wrote the first draft of the manuscript. All authors edited and approved the manuscript.

FUNDING

This work was supported by Second Genome Inc., NIH Small Business Innovation Research funding (grant number: R44DA043954), Canadian Institute of Health Research CIHR Chronic Inflammation Team Grant and IMAGINE SPOR Network.

ACKNOWLEDGMENTS

We sincerely thank all researchers who made their data available for this project. Special thanks to Dr. Janet Jansson and Colin Joseph Brislawn at Pacific Northwest National Laboratory, and Drs. Nicholas Chia and Patricio Jeraldo at Mayo Clinic. We thank Myron LaDuc for scientific editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.595910/full#supplementary-material>

- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Bravo, J. A., Forsythe, P., Chew, M. V., Escaravage, E., Savignac, H. M., Dinan, T. G., et al. (2011). Ingestion of *Lactobacillus* strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16050–16055. doi: 10.1073/pnas.1102999108

- Buchfink, B., Xie, C., and Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869
- Caspi, R., Altman, T., Dreher, K., Fulcher, C. A., Subhraveti, P., Keseler, I. M., et al. (2011). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* 40, D742–D753. doi: 10.1093/nar/gkr1014
- Chang, Y. L., Rossetti, M., Vlamakis, H., Casero, D., Sunga, G., Harre, N., et al. (2019). A screen of Crohn's disease-associated microbial metabolites identifies ascorbate as a novel metabolic inhibitor of activated human T cells. *Mucosal Immunol.* 12, 457–467. doi: 10.1038/s41385-018-0022-7
- Christians, U., Klawitter, J., Hornberger, A., and Klawitter, J. (2011). How unbiased is non-targeted metabolomics and is targeted pathway screening the solution? *Curr. Pharm. Biotechnol.* 12, 1053–1066. doi: 10.2174/138920111795909078
- Clooney, A. G., Fouhy, F., Sletor, R. D., O' Driscoll, A., Stanton, C., Cotter, P. D., et al. (2016). Comparing apples and oranges? Next generation sequencing and its impact on microbiome analysis. *PLoS One* 11:e0148028. doi: 10.1371/journal.pone.0148028
- Descamps, H. C., Herrmann, B., Wiredu, D., and Thaisz, C. A. (2019). The path toward using microbial metabolites as therapies. *EBioMedicine* 44, 747–754. doi: 10.1016/j.ebiom.2019.05.063
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461
- Franzosa, E. A., Sirota-Madi, A., Avila-Pacheco, J., Fornelos, N., Haiser, H. J., Reinker, S., et al. (2019). Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat. Microbiol.* 4, 293–305. doi: 10.1038/s41564-018-0306-4
- Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565
- Ghaste, M., Mistrik, R., and Shulaev, V. (2016). Applications of fourier transform ion cyclotron resonance (FT-ICR) and orbitrap based high resolution mass spectrometry in metabolomics and lipidomics. *Int. J. Mol. Sci.* 17:816. doi: 10.3390/ijms17060816
- Gika, H. G., Theodoridis, G. A., Earll, M., Snyder, R. W., Sumner, S. J., and Wilson, I. D. (2010). Does the mass spectrometer define the marker? A comparison of global metabolite profiling data generated simultaneously via UPLC-MS on two different mass spectrometers. *Anal. Chem.* 82, 8226–8234. doi: 10.1021/ac1016612
- Glinton, K. E., and Elsea, S. H. (2019). Untargeted metabolomics for autism spectrum disorders: current status and future directions. *Front. Psychiatry* 10:647. doi: 10.3389/fpsy.2019.00647
- Hale, V. L., Jeraldo, P., Mundy, M., Yao, J., Keeney, G., Scott, N., et al. (2018). Synthesis of multi-omic data and community metabolic models reveals insights into the role of hydrogen sulfide in colon cancer. *Methods* 149, 59–68. doi: 10.1016/j.ymeth.2018.04.024
- Integrative HMP (iHMP) Research Network Consortium (2014). The integrative human microbiome project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. *Cell Host Microbe* 16, 276–289. doi: 10.1016/j.chom.2014.08.014
- Iwai, S., Weinmaier, T., Schmidt, B. L., Albertson, D. G., Poloso, N. J., Dabbagh, K., et al. (2016). Piphillin: improved prediction of metagenomic content by direct inference from human microbiomes. *PLoS One* 11:e0166104. doi: 10.1371/journal.pone.0166104
- Jia, W., Li, H., Zhao, L., and Nicholson, J. K. (2008). Gut microbiota: a potential new territory for drug targeting. *Nat. Rev. Drug Discov.* 7, 123–129. doi: 10.1038/nrd2505
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2011). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 40, D109–D114. doi: 10.1093/nar/gkr988
- Kang, D.-W., Ilhan, Z. E., Isern, N. G., Hoyt, D. W., Howsmon, D. P., Shaffer, M., et al. (2018). Differences in fecal microbial metabolites and microbiota of children with autism spectrum disorders. *Anaerobe* 49, 121–131. doi: 10.1016/j.anaerobe.2017.12.007
- Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., et al. (2017). The BioCyc collection of microbial genomes and metabolic pathways. *Brief. Bioinform.* 20, 1085–1093. doi: 10.1093/bib/bbx085
- Karu, N., Deng, L., Slae, M., Guo, A. C., Sajed, T., Huynh, H., et al. (2018). A review on human fecal metabolomics: methods, applications and the human fecal metabolome database. *Anal. Chim. Acta* 1030, 1–24. doi: 10.1016/j.aca.2018.05.031
- Kopylova, E., Noé, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28, 3211–3217. doi: 10.1093/bioinformatics/bts611
- Kumar, M., Ji, B., Zengler, K., and Nielsen, J. (2019). Modelling approaches for studying the microbiome. *Nat. Microbiol.* 4, 1253–1267. doi: 10.1038/s41564-019-0491-9
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Larsen, P. E., Collart, F. R., Field, D., Meyer, F., Keegan, K. P., Henry, C. S., et al. (2011). Predicted Relative Metabolomic Turnover (PRMT): determining metabolic turnover from a coastal marine metagenomic dataset. *Microb. Inform. Exp.* 1:4. doi: 10.1186/2042-5783-1-4
- Lex, A., and Gehlenborg, N. (2014). Sets and intersections. *Nat. Methods* 11:779. doi: 10.1038/nmeth.3033
- Li, R., Li, F., Feng, Q., Liu, Z., Jie, Z., Wen, B., et al. (2016). An LC-MS based untargeted metabolomics study identified novel biomarkers for coronary heart disease. *Mol. Biosyst.* 12, 3425–3434. doi: 10.1039/C6MB00339G
- Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158
- Li, X. S., Wang, Z., Cajka, T., Buffa, J. A., Nemet, I., Hurd, A. G., et al. (2018). Untargeted metabolomics identifies trimethyllysine, a TMAO-producing nutrient precursor, as a predictor of incident cardiovascular disease risk. *JCI Insight* 3:e99096. doi: 10.1172/jci.insight.99096
- Liang, D., Li, M., Wei, R., Wang, J., Li, Y., Jia, W., et al. (2019). Strategy for intercorrelation identification between metabolome and microbiome. *Anal. Chem.* 91, 14424–14432. doi: 10.1021/acs.analchem.9b02948
- Lloyd-Price, J., Arze, C., Ananthakrishnan, A. N., Schirmer, M., Avila-Pacheco, J., Poon, T. W., et al. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* 569, 655–662. doi: 10.1038/s41586-019-1237-9
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8
- Maier, T. V., Lucio, M., Lee, L. H., VerBerkmoes, N. C., Brislaw, C. J., Bernhardt, J., et al. (2017). Impact of dietary resistant starch on the human gut microbiome, metaproteome, and metabolome. *mBio* 8:e1343–17. doi: 10.1128/mBio.01343-17
- Mallick, H., Franzosa, E. A., Mclver, L. J., Banerjee, S., Sirota-Madi, A., Kostic, A. D., et al. (2019). Predictive metabolomic profiling of microbial communities using amplicon or metagenomic sequences. *Nat. Commun.* 10:3136. doi: 10.1038/s41467-019-10927-1
- Manor, O., and Borenstein, E. (2015). MUSiCC: a marker genes based framework for metagenomic normalization and accurate profiling of gene abundances in the microbiome. *Genome Biol.* 16:53. doi: 10.1186/s13059-015-0610-8
- Metwaly, A., and Haller, D. (2019). Multi-omics in IBD biomarker discovery: the missing links. *Nat. Rev. Gastroenterol. Hepatol.* 16, 587–588. doi: 10.1038/s41575-019-0188-9
- MIMOSA2 (2020). MIMOSA2. Available online at: <https://borenstein-lab.github.io/MIMOSA2shiny/> (accessed September 24, 2020).
- Moliner, N., Ruiz, L., Sánchez, B., Margolles, A., and Delgado, S. (2019). Intestinal bacteria interplay with bile and cholesterol metabolism: implications on host physiology. *Front. Physiol.* 10:185. doi: 10.3389/fphys.2019.00185
- Morton, J. T., Aksenov, A. A., Nothias, L. F., Foulds, J. R., Quinn, R. A., Badri, M. H., et al. (2019). Learning representations of microbe-metabolite interactions. *Nat. Methods* 16, 1306–1314. doi: 10.1038/s41592-019-0616-3
- Narayan, N. R., Weinmaier, T., Laserna-Mendieta, E. J., Claesson, M. J., Shanahan, F., Dabbagh, K., et al. (2020). Piphillin predicts metagenomic composition and dynamics from DADA2-corrected 16S rDNA sequences. *BMC Genomics* 21:56. doi: 10.1186/s12864-020-6537-9

- Noecker, C., Eng, A., Srinivasan, S., Theriot, C. M., Young, V. B., Jansson, J. K., et al. (2016). Microbiome taxonomic and metabolomic profiles elucidates mechanistic links between ecological and metabolic variations. *mSystems* 1:e00013–15. doi: 10.1128/mSystems.00013-15
- Ortmayr, K., Causon, T. J., Hann, S., and Koellensperger, G. (2016). Increasing selectivity and coverage in LC-MS based metabolome analysis. *TrAC Trends Anal. Chem.* 82, 358–366. doi: 10.1016/j.trac.2016.06.011
- Patel, S., and Ahmed, S. (2015). Emerging field of metabolomics: big promise for cancer biomarker identification and drug discovery. *J. Pharm. Biomed. Anal.* 107, 63–74. doi: 10.1016/j.jpba.2014.12.020
- Peres-Neto, P. R., and Jackson, D. A. (2001). How well do multivariate data sets match? The advantages of a procrustean superimposition approach over the Mantel test. *Oecologia* 129, 169–178. doi: 10.1007/s004420100720
- Shaffer, M., Thurimella, K., Quinn, K., Doenges, K., Zhang, X., Bokatzian, S., et al. (2019). AMON: annotation of metabolite origins via networks to integrate microbiome and metabolome data. *BMC Bioinformatics* 20:614. doi: 10.1186/s12859-019-3176-8
- Sharon, G., Cruz, N. J., Kang, D. W., Gandal, M. J., Wang, B., Kim, Y. M., et al. (2019). Human gut microbiota from autism spectrum disorder promote behavioral symptoms in mice. *Cell* 177, 1600–1618.e17. doi: 10.1016/j.cell.2019.05.004
- Song, X., Sun, X., Oh, S. F., Wu, M., Zhang, Y., and Zheng, W. (2019). Microbial bile acid metabolites modulate gut ROR γ + regulatory T cell homeostasis. *Nature* 577, 410–415. doi: 10.1038/s41586-019-1865-0
- Stewart, C. J., Hasegawa, K., Wong, M. C., Ajami, N. J., Petrosino, J. F., Piedra, P. A., et al. (2018). Respiratory syncytial virus and rhinovirus bronchiolitis are associated with distinct metabolic pathways. *J. Infect. Dis.* 217, 1160–1169. doi: 10.1093/infdis/jix680
- Strandwitz, P., Kim, K. H., Terekhova, D., Liu, J. K., Sharma, A., Levering, J., et al. (2019). GABA-modulating bacteria of the human gut microbiota. *Nat. Microbiol.* 4, 396–403. doi: 10.1038/s41564-018-0307-3
- Sunagawa, S., Mende, D. R., Zeller, G., Izquierdo-Carrasco, F., Berger, S. A., Kultima, J. R., et al. (2013). Metagenomic species profiling using universal phylogenetic marker genes. *Nat. Methods* 10, 1196–1199. doi: 10.1038/nmeth.2693
- Tulipani, S., Llorach, R., Urpi-Sarda, M., and Andres-Lacueva, C. (2013). Comparative analysis of sample preparation methods to handle the complexity of the blood fluid metabolome: when less is more. *Anal. Chem.* 85, 341–348. doi: 10.1021/ac302919t
- Turnbaugh, P. J., and Gordon, J. I. (2008). An invitation to the marriage of metagenomics and metabolomics. *Cell* 134, 708–713. doi: 10.1016/j.cell.2008.08.025
- Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., and Gordon, J. I. (2007). The human microbiome project. *Nature* 449, 804–810. doi: 10.1038/nature06244
- Týčová, A., Ledvína, V., and Klepárník, K. (2017). Recent advances in CE-MS coupling: instrumentation, methodology, and applications. *Electrophoresis* 38, 115–134. doi: 10.1002/elps.201600366
- Ufnal, M., Zadlo, A., and Ostaszewski, R. (2015). TMAO: a small molecule of great expectations. *Nutrition* 31, 1317–1323. doi: 10.1016/j.nut.2015.05.006
- Urpi-Sarda, M., Almanza-Aguilera, E., Llorach, R., Vázquez-Fresno, R., Estruch, R., Corella, D., et al. (2019). Non-targeted metabolomic biomarkers and metabolotypes of type 2 diabetes: a cross-sectional study of PREDIMED trial participants. *Diabetes Metab.* 45, 167–174. doi: 10.1016/j.diabet.2018.02.006
- Ussher, J. R., Lopaschuk, G. D., and Arduini, A. (2013). Gut microbiota metabolism of l-carnitine and cardiovascular risk. *Atherosclerosis* 231, 456–461. doi: 10.1016/j.atherosclerosis.2013.10.013
- van Dam, N. M., and Bouwmeester, H. J. (2016). Metabolomics in the rhizosphere: tapping into belowground chemical communication. *Trends Plant Sci.* 21, 256–265. doi: 10.1016/j.tplants.2016.01.008
- Van Treuren, W., and Dodd, D. (2020). Microbial contribution to the human metabolome: implications for health and disease. *Annu. Rev. Pathol. Mech. Dis.* 15, 345–369. doi: 10.1146/annurev-pathol-020117-043559
- Vuckovic, D. (2012). Current trends and challenges in sample preparation for global metabolomics using liquid chromatography–mass spectrometry. *Anal. Bioanal. Chem.* 403, 1523–1548. doi: 10.1007/s00216-012-6039-y
- Wishart, D. S. (2016). Emerging applications of metabolomics in drug discovery and precision medicine. *Nat. Rev. Drug Discov.* 15, 473–484. doi: 10.1038/nrd.2016.32
- Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15:R46. doi: 10.1186/gb-2014-15-3-r46
- Zhang, X., Li, L., Butcher, J., Stintzi, A., and Figeys, D. (2019). Advancing functional and translational microbiome research using meta-omics approaches. *Microbiome* 7:154. doi: 10.1186/s40168-019-0767-6
- Zheng, P., Zeng, B., Liu, M., Chen, J., Pan, J., Han, Y., et al. (2019). The gut microbiome from patients with schizophrenia modulates the glutamate-glutamine-GABA cycle and schizophrenia-relevant behaviors in mice. *Sci. Adv.* 5:eaa8317. doi: 10.1126/sciadv.aau8317
- Zierer, J., Jackson, M. A., Kastenmüller, G., Mangino, M., Long, T., Telenti, A., et al. (2018). The fecal metabolome as a functional readout of the gut microbiome. *Nat. Genet.* 50, 790–795. doi: 10.1038/s41588-018-0135-7

Conflict of Interest: XY, ER, KW, YW, JC, KD, TD, and SI are employees of Second Genome Inc., which provides salaries and stock options. Second Genome Inc. is an independent therapeutics company with products in development to treat intestinal disorders and other human diseases. TA is an employee of Altman Analytics LLC.

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.

Copyright © 2020 Yin, Altman, Rutherford, West, Wu, Choi, Beck, Kaplan, Dabbagh, DeSantis and Iwai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.