Untargeted metabolomics profiling and modeling for novel biomarker and pathway discovery in ulcerative colitis

Isabel Prado-Tucker, Xiangping Lin, PhD, Xinyue Zhang, PhD, Michael Snyder, PhD
Dept. of Genetics, Stanford University, Stanford, CA 94305

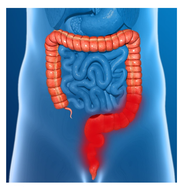
Abstract

Ulcerative colitis (UC) is a subtype of inflammatory bowel disease (IBD) characterized by superficial chronic inflammation of the lower colon. The disease prevalence is 156 to 291 in 100,000 Americans. While several genetic and exposomic risk factors have been found, the pathogenesis of UC on the molecular level is largely unknown and no cure exists. We hypothesized that we could identify novel biomarkers and pathways for UC by extensive metabolomic analysis with classification models and pathway analysis.

We explored the metabolic landscape of UC by analyzing metabolic data from The Inflammatory Bowel Disease Multi'omics Database (IBDMDB). A stratified random sample (StRS) of gender and diagnosis (UC/non-IBD) was used to select 120 participants. Metabolomic profiling was performed on the stool samples of these participants using liquid chromatography-mass spectrometry (LC-MS) with reverse-phase C8 column and positive ionization mode. Missing value imputation and feature selection were performed to prepare the data. Sparse partial least squares discriminant analysis (sPLS-DA) after principal component analysis (PCA), pathway analysis, and biomarker analysis were performed separately.

The sPLS-DA classification model had an accuracy of 0.93 in the test set when identifying UC and non-IBD. The most significant pathway changes discovered include carnitine shuttle (p=1.75E-02), bile acid biosynthesis (p=1.77E-02), C21-steroid hormone (p=1.78E-02), de novo fatty acid biosynthesis (p=1.94E-02), and fatty acid activation (p=2.03E-02). 10 significant biomarkers were identified; however, further testing is required to identify those metabolites. Our results suggest metabolomics is a promising approach to identify novel biomarkers and pathways for UC, an important step in measuring their response to known risk factors and tracking disease progression. Next steps include identifying the 10 significant metabolites and exploring how these metabolites and pathways are linked to UC.

Introduction



"Reasons You May Need An Ileostomy" by Werner Woodruff is licensed under CC BY-SA 2.0.

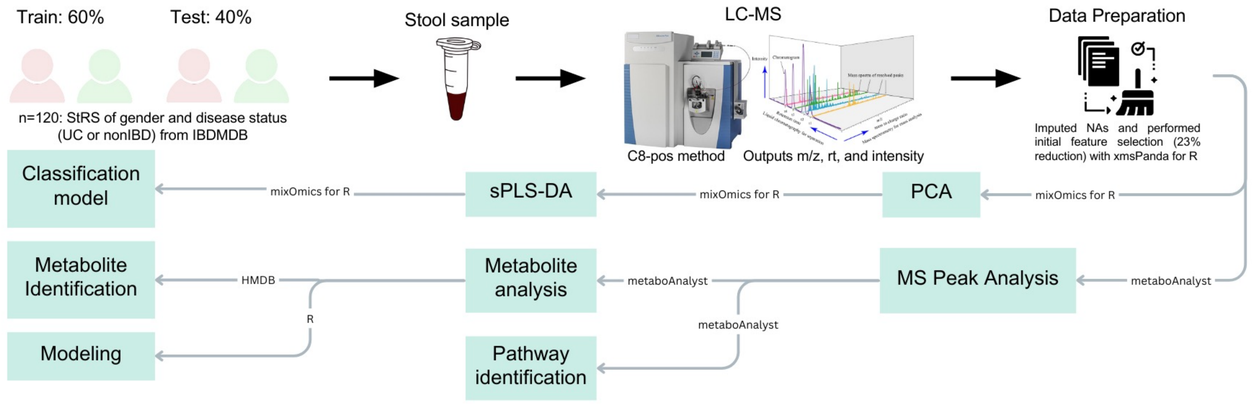
Ulcerative colitis

- Subtype of IBD
- Prevalence is 156 to 291 in 100,000 Americans and increasing with westernization and industrialization
- Superficial chronic inflammation of the colon, most often concentrated by the rectum
- Disease course presents as flares
- Genetic and exposomic risk factors
- No known cure

Metabolomics

- Study of metabolites, small (<1.5 kDa) molecules in metabolism
- Untargeted metabolomics: comprehensive method to identify all present metabolites, rather than a select few
- Promising tool for ulcerative colitis because it can identify biomarkers that aid in the understanding of all stages of the disease course

Methods & Materials



Results

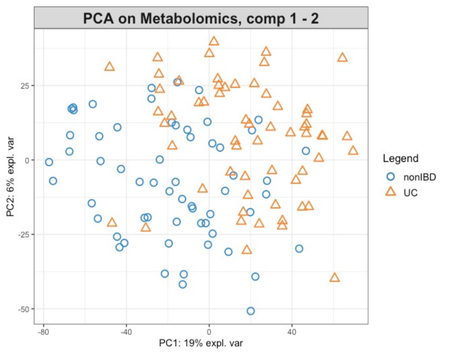


Figure 1. Unsupervised PCA model.

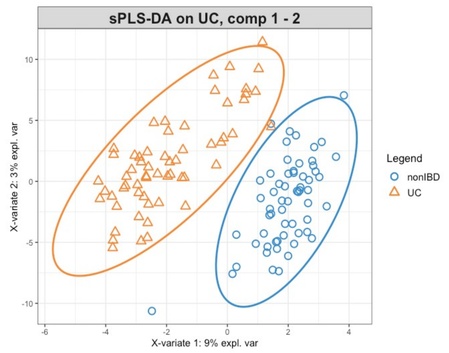


Figure 2. Supervised sPLS-DA model.

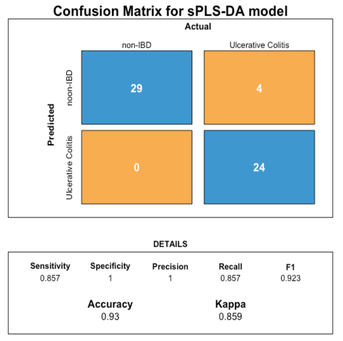


Figure 3. Confusion matrix for sPLS-DA model with accuracy of 0.93 on the test set.

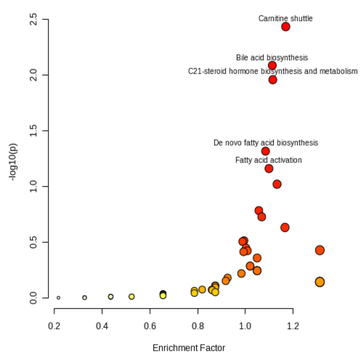


Figure 4. Significant metabolic pathways associated with ulcerative colitis.

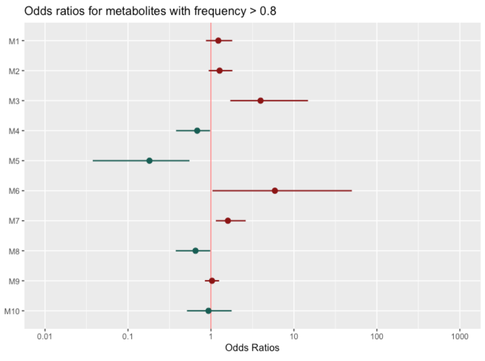


Figure 5. Odds ratios of the top 10 significant and stable metabolites (frequency > 0.80).

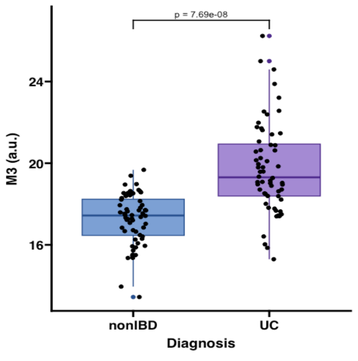


Figure 6. Comparison of intensity of the one of the significant metabolites with odds ratio 3.96 in UC and non-IBD.

Discussion

Our results support our hypothesis that novel biomarkers and pathways for UC can be identified through extensive metabolomic analysis with classification models and pathway analysis. Using an untargeted metabolomic approach, we were able to determine several significant associations between metabolites and UC.

The sPLS-DA classification model had a high accuracy of 0.93, suggesting that metabolomics provides valuable insight into UC.

We identified 5 significant metabolite pathways: carnitine shuttle (p=1.75E-02), bile acid biosynthesis (p=1.77E-02), C21-steroid hormone (p=1.78E-02), de novo fatty acid biosynthesis (p=1.94E-02), and fatty acid activation (p=2.03E-02). Some of these pathways, especially bile acid biosynthesis, have been linked to ulcerative colitis in previous literature.

Ten significant metabolites were identified; however, further testing is required to accurately identify the metabolites.

Conclusions

Our results suggest metabolomics is a promising approach to identify novel biomarkers and pathways for UC. Future steps for this research include identifying the 10 significant metabolites and exploring how these metabolites and pathways are linked to UC, to better understand the pathogenesis of UC and to develop new treatments.

Acknowledgements

This research was made possible thanks to the Stanford Institutes of Medical Research Program (SIMR), the Doris Duke Foundation, Essey Afewerki, and members of the Snyder Lab at Stanford University.

DORIS DUKE CHARITABLE FOUNDATION