

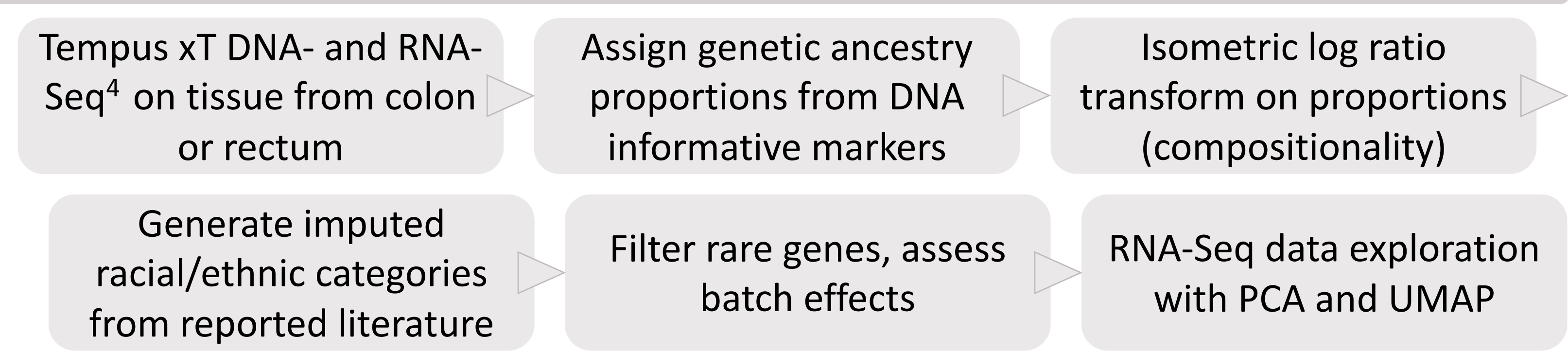
David Hein, MS<sup>1</sup>; Brooke Rhead PhD<sup>2</sup>; Yannick Pouliot, PhD<sup>2</sup>; Justin Guinney, PhD<sup>2</sup>; Francisco De La Vega DSc<sup>2</sup>; and Nina Sanford, MD<sup>1</sup>  
<sup>1</sup>The University of Texas Southwestern Medical Center, Dallas, TX; <sup>2</sup>Tempus Labs, Chicago, IL

## Introduction

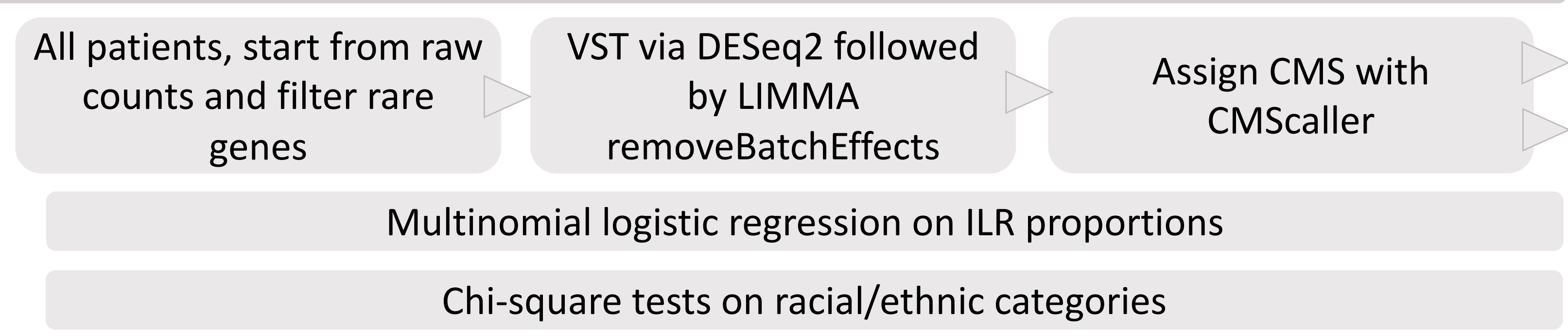
There is a growing incidence of colorectal cancer (CRC) among young adults and persistent disparities in outcomes by race/ethnicity across all ages. Gene expression signatures, as well as consensus molecular subtypes (CMS) derived from these, have been proposed to predict prognosis and therapy response in CRC. However, it is unclear whether gene expression and/or CMS are associated with racial disparities observed in CRC. We assessed whether race or genetic ancestry are associated with CMS or gene expression patterns in a deidentified cohort of 1956 CRC patients.

## Methods

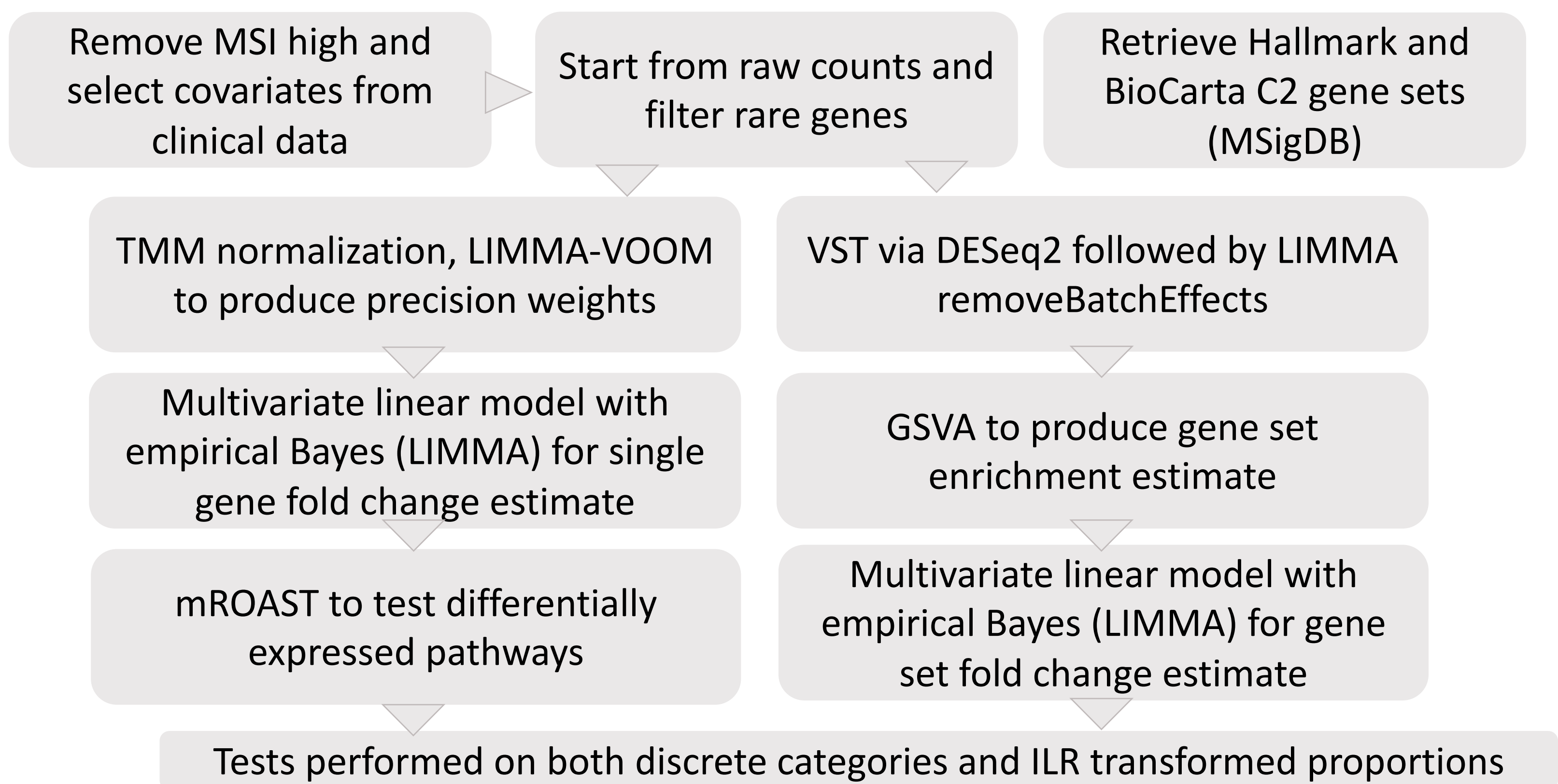
### Data Preparation



### Consensus Molecular Subtype Analysis

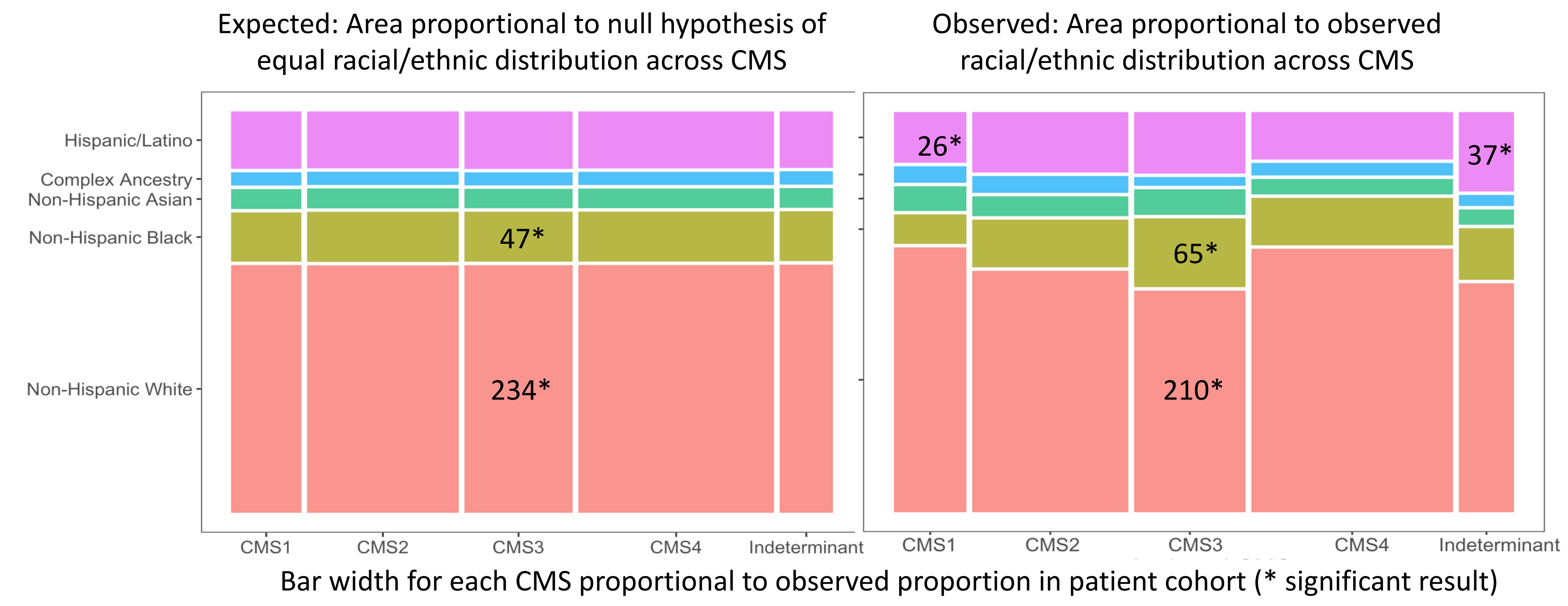


### Gene Set Analysis

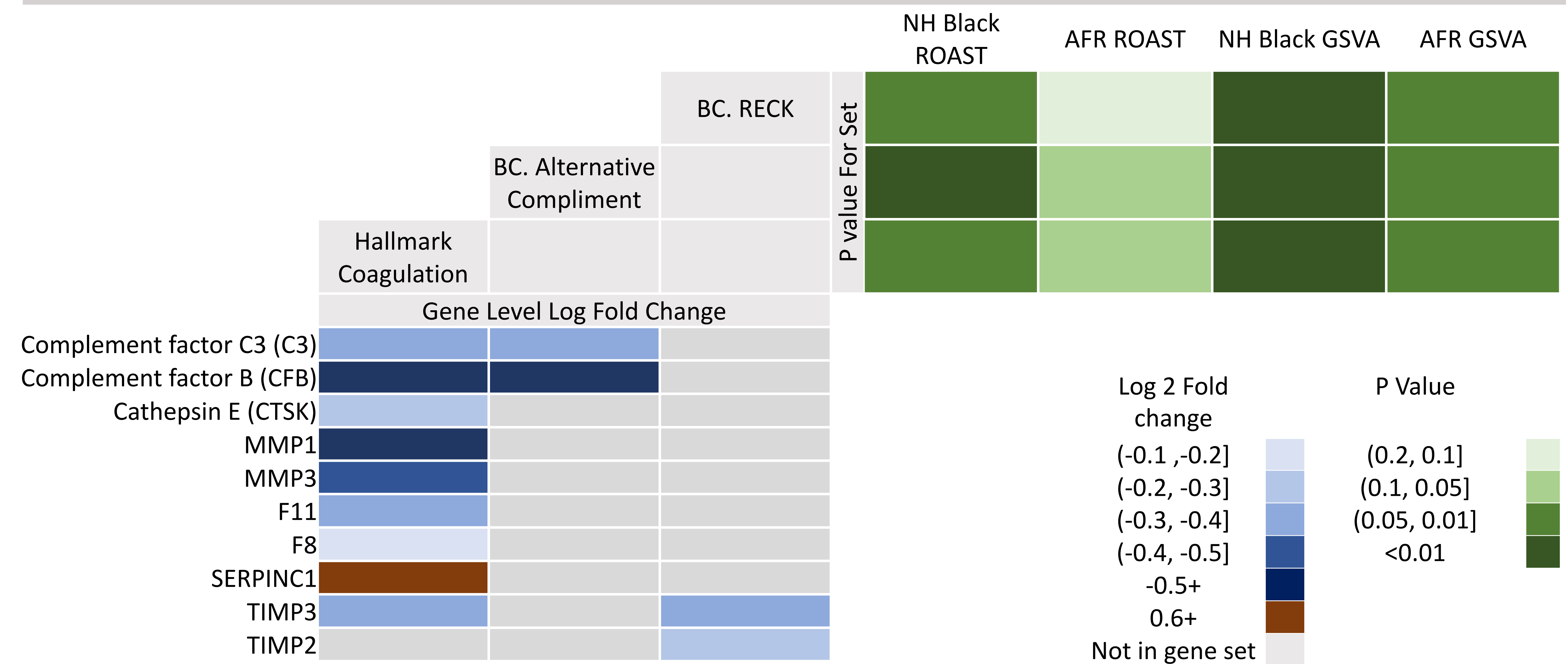


## Results

### CMS Results



### Gene Set Results



## Conclusions and Open Questions

We found that NH Black patients and AFR ancestry were associated with higher rates of CMS3, which is associated with *KRAS* mutation and was previously reported to be more common among Black patients. The association of Indeterminate with Hispanic/Latino ancestry highlights the need to use diverse patient cohorts when training unsupervised learning models to improve prognosis prediction in non-White patients. Finally, our gene set analysis reveals under expression in several cancer related gene sets among NH Black patients and those with increased AFR ancestry.

## Contact

David Hein  
 UT Southwestern Medical Center, Department of Radiation Oncology  
 David.hein@utsouthwestern.edu  
 Github: DavidHein96

## References

1. Beaubier N, et al. Clinical validation of the tempus xT next-generation targeted oncology sequencing assay. *Oncotarget*. 2019
2. Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015
3. Eide PW, et al. CMScaller: an R package for consensus molecular subtyping of colorectal cancer pre-clinical models. *Sci Rep*. 2017
4. Liberzon A, et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011
5. Hänzelmann S, GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013