Integrated Multiomics for Deep Molecular Exploration from Minimally and Non-Invasive Samples

Bhaven Mehta[1], Andrea O'Hara[1], Tom Cohen[2], Ethan Stancliffe[2], Adam Richardson[2], David Corney[1], Haythem Latif[1]

[1]GENEWIZ from Azenta Life Sciences, South Plainfield, NJ 07080

[2]Panome Bio, St Louis, MO 63310

Abstract

The omics era has greatly expanded the repertoire of approaches available to unravel the complexity underpinning human health, with the ability to rapidly characterize genomes, epigenomes, transcriptomes, proteomes and metabolomes from a wide range of sample types. Minimally invasive (blood/plasma) and non-invasive (urine, saliva) samples are ideal for cancer screening but require highly sensitive methods paired with highly reliable biomarkers for detection.

Here we detail a proof-of-principle workflow which utilizes a single blood draw to rapidly produce a diverse set of multiomics results including genomics, epigenomics, transcriptomics, proteomics and metabolomics. Blood draws using heparin tubes were collected and processed within 24 hours of the primary blood draw to ensure high viability and yield of PBMCs, along with simultaneous plasma separation and collection. These PBMC and plasma samples are then processed such that whole exome sequencing, whole genome methylation sequencing, single cell RNA sequencing, proteomic analysis and metabolomic analysis can all be collected from patient blood draw.

Integrated analysis across multiple data modalities allows for holistic views of pathways and processes that are highly impacted. Genomics, alongside proteomics and transcriptomics, offers insights into genetic and functional potential. When combined with metabolomics, it enables a deeper understanding of phenotypic variations both within and across individuals in a patient cohort. All these datatypes can be produced within days of primary sample collection using minimal sample amounts from a single blood draw, to address the biological questions at hand. With increased accessibility of omics approaches, integrative workflows such as the one described continue to gain broader adoption and drive greater insights and innovation in human health applications.

Methods

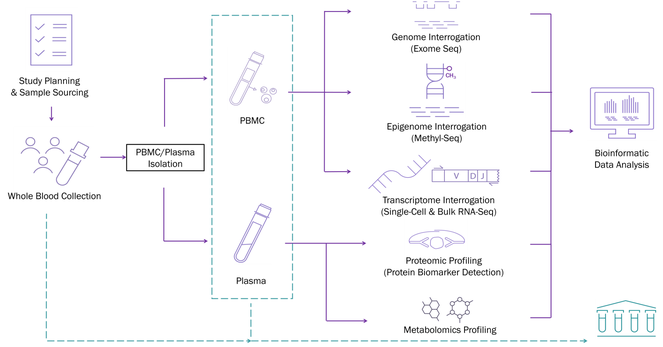
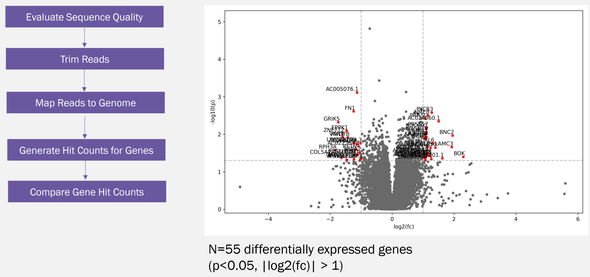
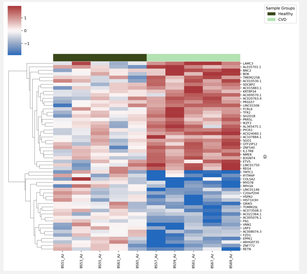


Figure 1. Schematic of liquid biopsy multiomics sample processing workflow. A 10mL blood draw yields ~4–5mL of plasma and ~0.5–3 x 10^6 cells per mL or roughly 5+ aliquots of plasma >500ml and 5+ aliquots of >1M cryopreserved cell pellets. For the workflow above, a single cell pellet is used for DNA isolation, with enough DNA (>2ug) for multiple genome/epigenome assays. One cryopreserved cell pellet was used for single-cell RNA seq (~1M cells), and one for bulk RNA-Seq, with remainder PBMC aliquots in long-term storage. A single aliquot of plasma can be split between Olink (~15ul) and SEER (~250ul) proteomics and LC/MS metabolomics (~50ul) assays, done simultaneously, with extra aliquots in long-term storage.

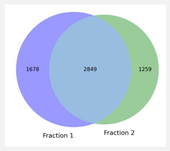
A. Transcriptomics Results

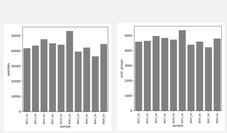


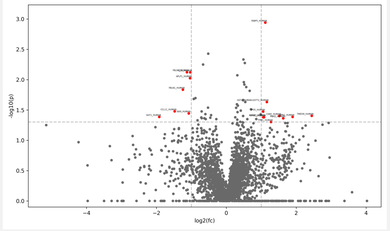
N=55 differentially expressed genes
(p<0.05, |log2(fc)| > 1)



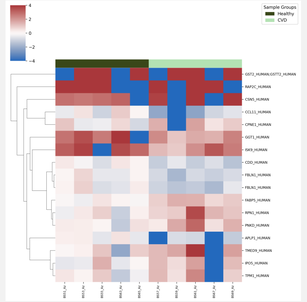
B. Proteomics Results



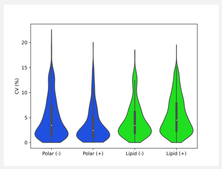


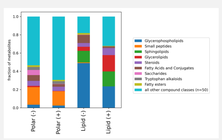


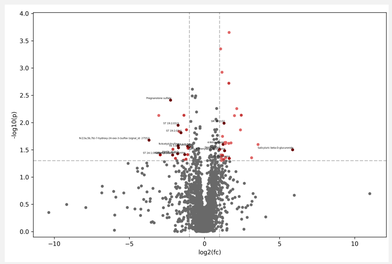
N=17 differentially expressed proteins (p<0.05, |log2(fc)| > 1)



C. Metabolomics Results







N=18 differential metabolites (p<0.05, |log2(fc)| > 1)

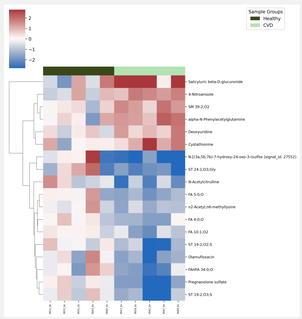
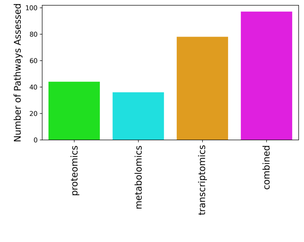
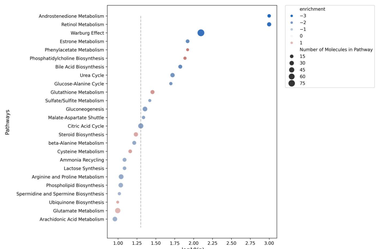


Figure 2. (A) RNA-Sequencing analysis identified 55 differentially expressed genes between groups. Bulk RNA-seq performed on the PBMCs collected from each sample. Results were normalized on a per-sample level and analyzed across all 10 samples. A total of 16,071 genes were profiled. Note: N=0 differentially expressed genes with (q<0.05). **(B) Proteomics results identified 17 differentially express proteins between groups.** Plasma samples were prepared with Seer Proteograph XT assay with two fractions produced. LC/MS/MS was profiled on each fraction and data was combined for peptide and protein analysis. Of the 5639 proteins detected, 3239 were found in 2 or more samples. **(C) Untargeted metabolomics analysis identified 18 differentially expressed metabolites between groups.** Four LC/MS (liquid chromatography mass spectrometry) assays were employed to profile the metabolome of the plasma samples. Of the 2607 metabolites profiled, 1925 metabolites were positively identified of the >50 polar metabolite and lipid classes surveyed.

A. Integrated Analysis: Pathway Analysis





B. Integrated Analysis: Network Integration

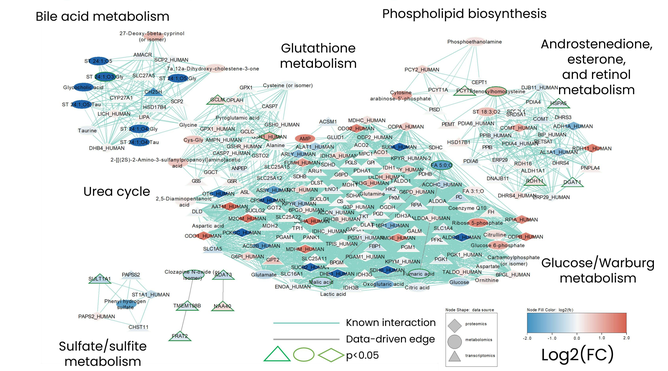


Figure 3. (A) Integrated pathway analysis. Thirteen metabolic pathways enriched in CVD when all three data modalities are combined. **(B) Integrated network analysis.** After correlating metabolite, protein, and gene expression levels across all samples to form a data driven network with known interactions (Metabolite-protein, Metabolite-gene, Gene-protein), the network is pruned to the most tightly connected metabolites/genes/proteins. The net integration of >20k molecular measurements yields 225 interconnected metabolites (42), proteins (73), and genes (110).

Conclusions

- Complex multiomics information can be analyzed from a single sample, including transcriptomics, proteomics & metabolomics.
- Multiomics workflows can be modified to accommodate a variety of input sample types, including urine, allowing for flexible sample processing.
- Integrated analysis allows for deeper insights into key pathways and networks for future discovery.
- Metabolomics analysis identified 18 differentially abundant metabolites; proteomics analysis identified 17 differentially expressed proteins; transcriptomics analysis identified 54 differentially expressed genes.
- Integrated analysis identified dysregulation of 13 pathways. Network analysis revealed these changes can be grouped into 6 modules, all of which have previously been correlated with CVD risk