Title: ***Multi-Omic Analysis of Ischemic Stroke Clots Identifies Complex Traits Associated with Etiology***

Authors: Briana A. Santo, Seyyed Mostafa M. Janbeh Sarayi, Kerry E. Poppenberg, Kenneth V. Snyder, Adnan H. Siddiqui, Vincent M. Tutino

Background: Identification of ischemic stroke etiology (TOAST class) is essential for treatment and secondary prevention. Most studies focus on investigation of clot biology using various omic analysis.

Hypothesis: Canonical correlation analysis (CCA) of paired clot transcriptomic and histomic features will identify etiology-specific white blood cell (WBC) image-omic correlates.

Methods: 31 clots (LAA=5, CE=3, OTH=3, CR=5) with paired H&E histology whole-slide images and transcriptomic data[1] were analyzed. A computational histopathology pipeline[2] for WBC segmentation and characterization was used to quantify 46 image features (e.g., textures representative of heterochromatin to euchromatin ratio) from >100,000 WBC instances. Features were then used to classify WBCs into 1 of 10 validated classes and calculate clot class frequency distributions. SparseCCA was used to test associations among image features and expression signatures. GO enrichment analysis was used to identify biological processes (BPs) associated with (+)-correlated genes.

Results: Canonical covariates (e.g., histology canonical variable [CV] 1 and transcriptome CV1) were strongly and significantly correlated (R>0.9, p<0.001). Pairwise correlation analysis of CVs and original feature sets (e.g., to find original WBC image features captured in histology CVs) identified 3 CVs highly enriched in LAA (CV6), CE (CV14), and OTH (CV17) that represented distinct WBC subtypes in histology with unique GO BPs (Fig. 1).

Conclusion: For the first time, joint analysis of clot histology and gene expression data has identified significant structure-expression covariates representing biologically-interpretable traits. Together with routine clinical data, these biomarkers may facilitate robust etiology identification, reducing CR diagnoses and aiding secondary stroke prevention.

References

[1] Tutino VM, et al. *JNIS*. 2022.
[2] Patel TR, Santo BA, et al. *SVIN*. 2022.

Figures

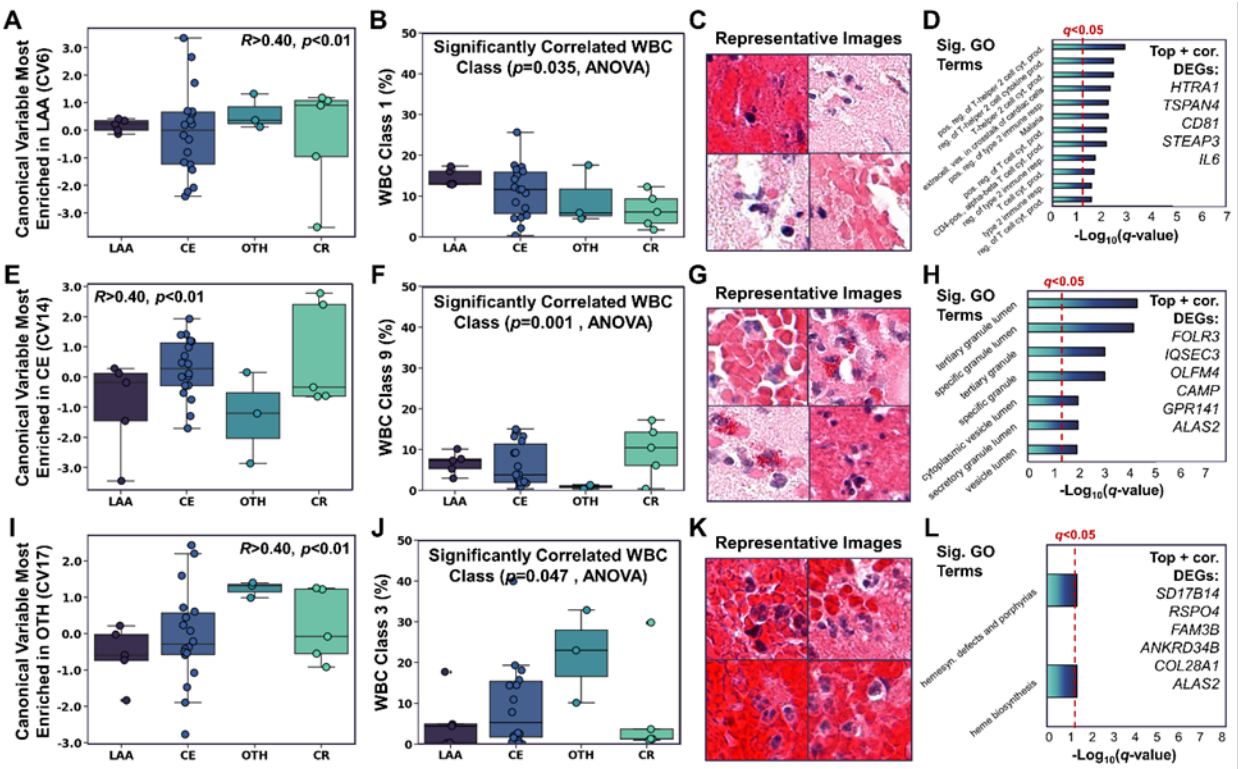

Figure 1: Enriched WBC image and expression features among TOAST classes. A-D). Canonical covariate (CV) 6 was significantly enriched in LAA **(A)**, with WBC Class 1 in histology **(B)** showing cells reminiscent of leukocytes. There was high correlation of histology **(C)** with T-Cell GO biological processes **(D)**. The frequency distribution of WBC Class 1 was significantly different among TOAST classes (p=0.035), and representative cells were reminiscent of leukocytes with dark, round nuclei and a thin rim of cytoplasm visible. **E-H).** CV14 was enriched in CE **(E)** and was found to represent neutrophils and granular cells on histology (WBC Class 9, p=0.001 among TOAST groups) **(F)**. In representative images, polymorphonuclear cells corresponding to neutrophils are readily apparent, along with eosin-rich granular structures. These WBCs **(G)** were positively correlated with genes enriching in ontology terms related to granules **(H)**. **I-L).** In OTH, we found that CV17 was significantly enriched **(I)** and represented WBCs embedded in RBC-rich regions of thrombi, Class 3 **(J)**. From histology **(K)**, WBC Class 3 was highly correlated with heme biosynthesis genes **(L)**, and was significantly different among TOAST groups with prevalence in OTH (p=0.047). (Abbreviations: CV=canonical variable, GO=gene ontology, DEG=differentially expressed genes, ANOVA=analysis of variance, sig=significant, cor.=correlated)