Background: The standard approaches to understand gene expression and regulation in the brain include identification of differentially and co-expressed genes. To this end, the Accelerating Medicines Partnership Alzheimer's Disease (AMP-AD) consortium has produced multiple, large RNA-seq datasets from several postmortem brain regions. Separately, the ENCODE project has produced DNAse Hypersensitivity (DHS) samples for various brain regions. We have integrated these large datasets into transcriptional regulatory networks (TRN), providing a directional and mechanistic list of putative transcription factors for nearly all expressed genes in the brain. **Methods:** We reprocessed all ENCODE brain DHS samples at scale, generating footprints—signatures of occupancy by DNA binding proteins—using the Wellington and HINT algorithms. We assembled motifs from JASPAR2016, HOCOMOCO, UniPROBE, and SwissRegulon, removing redundant motifs with Tomtom and intersecting our footprints with all possible overlapping motifs. This resulted in a total of 1,530 motifs mapping to 1,515 different transcription factors. We developed and utilized Transcriptional Regulatory Network Analysis (TReNA), available as an R Bioconductor package. Gene regulatory regions considered in our model were obtained through Genehancer, thus enabling the inclusion of all known enhancer regions. TReNA utilizes an ensemble of machine learning techniques, including lassopv, square root lasso (flare) and randomForest to prioritize transcription factors based on the expression levels in RNA-seq for each target gene. The scores from the aforementioned techniques are scaled and normalized into a composite score, thereby ranking all associated transcription factors for each target gene. **Results:** We have identified transcriptional regulators for AD genes identified through GWAS. We have also identified putative targets for the AD-associated transcription factor MEF2C. We have identified multiple microglia-enriched transcription factors that regulate many differentially and co-expressed genes in AD, in particular, the AD-associated transcription factor SPI1. **Conclusions:** These resulting models can be applied to other datasets that generate lists of differentially or co-expressed genes as well as provide testable hypotheses for non-coding variants of interest. We are actively engaged in testing several hypotheses through experimental means and have made these TRNs publically available.

O3-03-02 SYSTEMS BIOLOGY RANKING OF CANDIDATE ALZHEIMER'S DRIVER GENES IDENTIFIES NEW GENETIC DRIVERS OF ALZHEIMER'S DISEASE ETIOLOGY

Benjamin A. Logsdon[1], Thanneer M. Perumal[1], Kenneth Daily[1], Solveig K. Sieberts[1], Larsson Omberg[1], Lara M. Mangravite[1], Christoph Preuss[2], Gregory W. Carter[2], [1]*Sage Bionetworks, Seattle, WA, USA;* [2]*The Jackson Laboratory, Bar Harbor, ME, USA. Contact e-mail: ben.logsdon@sagebase.org*

Background: Late onset Alzheimer's disease (LOAD) is a debilitating illness with no known disease modifying treatment. Identification of new AD biology will be key to finding effective treatments. We present a technology platform for integrating genomic analytic outputs to prioritize and understand systems biology evidence for specific genes driving LOAD. This includes a web application to enable users to quickly view multiple systems biology analyses on a single gene from the AMP-AD consortia including: coexpression network analyses, differential expression analyses, eQTL analyses, tissue specific expression, and druggability. To help investigators better prioritize AD drivers for validation, we present a new data driven ranking system of genes based on their probability that they drive LOAD. **Methods:** We build a predictive model of AD "driverness" to score genes with a reference set of 27 known AD genetic loci with a lasso penalized regression model. We build gene level AD RNAseq features using three broad classes 1) genome wide coexpression network topological summaries across three AMP-AD RNA-seq studies (ROSMAP, Mayo RNAseq, MSBB) and seven tissue types, 2) differential expression gene sets from these data, and 3) network topological summaries for 42 AD associated modules from these data. We select the top 1000 scored genes from the model to screen for variants with evidence in IGAP stage 1+2 at a lower multiple hypothesis testing burden. **Results:** We identify 667 variants of which 404 pass a Bonferroni correction ($P < 7x10^{-5}$). We identify six new genomic regions associated with Alzheimer's disease: *HBEGF*, *RABEP1*, *DNAH17*, *C1S*, *SNX1*, and *PFKFB3*. Furthermore, a missense exonic variant in *SNX1*, *rs1802376,* has a p-value in IGAP stage 1+2 of $6x10^{-6}$, a CADD score of 28.3, and a replication p-value in the ADSP exome data of $2.6x10^{-3}$. This provides evidence that integration of multiple systems biology resources can provide insights into new Alzheimer's disease loci and mechanisms. **Conclusions:** We provide a framework for assessment and integration of systems biology outputs from the AMP-AD consortia. This is a new community resource for AD researchers to discover available data on a gene of interest and high priority targets identified by the AMP-AD consortia.

O3-03-03 EPIGENOME-WIDE ASSOCIATION STUDIES IMPLICATE GENES INVOLVED IN GLIAL CELL FUNCTION AND VIRAL RESPONSE IN CEREBRAL WHITE MATTER HYPERINTENSITIES

Myriam Fornage[1], Hieab H. H. Adams[2], Joshua C. Bis[3], Eric Boerwinkle[4], Ian J. Deary[5], Hans J. Grabe[6], Lifang Hou[7], M. Arfan Ikram[8], Min A. Jhun[9], Lenore J. Launer[10], Michelle Luciano[5], Sudha Seshadri[11], Jennifer A. Smith[12], Nona Sotoodehnia[3], Alexander Teumer[13], Qiong Yang[14], [1]*The University of Texas Health Science Center at Houston, Houston, TX, USA;* [2]*Erasmus Medical Center, Rotterdam, Netherlands;* [3]*University of Washington, Seattle, WA, USA;* [4]*University of Texas Health Science Center at Houston, Houston, TX, USA;* [5]*University of Edinburgh, Edinburgh, United Kingdom;* [6]*Department of Psychiatry, University Medicine Greifswald and German Center for Neurodegenerative Disease, Rostock/Greifswald, Greifswald, Germany;* [7]*Northwestern University Fienberg School of Medicine, Chicago, IL, USA;* [8]*Department of Epidemiology, Erasmus University Medical Center, Rotterdam, Netherlands;* [9]*University of Michigan, Ann Arbor, MI, USA;* [10]*National Institute on Aging, Bethesda, MD, USA;* [11]*University of Texas Health Sciences Center, San Antonio, TX, USA;* [12]*University of Michigan School of Public Health, Ann Arbor, MI, USA;* [13]*German Center for Cardiovascular Research, Greifswald, Germany;* [14]*Boston University School of Public Health, Boston, MA, USA. Contact e-mail: myriam.fornage@uth.tmc.edu*

Background: Cerebral white matter hyperintensities (WMH) detected on MRI are strong radiological correlates of age-related cognitive decline and represent an early marker of Alzheimer's disease (AD), independent of AD pathology. Despite a recognized high heritability, the molecular basis of WMH susceptibility has not been fully characterized. In particular, the contribution of epigenetic modifications, such as DNA methylation, has not been explored. **Methods:** We conducted a meta-analysis of epigenome-wide association studies of WMH burden in up to 5679 elderly participants of European and African ancestry from 9 population-based cohorts using blood-derived DNA methylation measured

on the Infinium HumanMethylation450 BeadChip. Cohort-specific associations between burden of WMH and DNA methylation beta values were estimated using linear mixed-effect models and combined in a sample-size weighted fixed-effect meta-analysis. In addition, we used two different approaches to identify differentially methylated regions (DMRs), which may be more informative than individual loci. Bonferroni correction and False Discovery Rates were used to account for the multiple tests. **Results:** Single-site analyses identified a CpG site on chromosome 11 significantly associated with WMH (cg24202936, P= $1.2x10^{-7}$). Region-based analyses, which leverage the correlations between nearby CpG sites identified 3 DMRs across 3 genes significantly associated with WMH burden: *PRMT1*, *BTBD17*, and *IFITM10* (P= $1.4x10^{-10}$, $2.3x10^{-8}$, and $3.6x10^{-7}$, respectively). *PRMT1* encodes the Protein Arginine N-methylase 1, which methylates histones in genes involved in glioblastomagenesis. Mice lacking this gene are characterized by severe defects in oligodendrocyte maturation processes. The function of *BTBD17* and *IFITM10* is not characterized but both are expressed in the brain and exhibit changes in expression in response to viral infection. **Conclusions:** Consistent with our previously reported genetic association studies, this genome-wide DNA methylation analysis supports a role of genes involved in glial cell function in WMH etiology. It also suggests a novel role of genes involved in viral response. Analyses are in progress to examine whether these changes in DNA methylation are under genetic control and are associated with cognitive outcomes.

O3-03-04 A HIGH RESOLUTION CAPTURE-C PROMOTER INTERACTOME IMPLICATES CAUSAL GENES AT ALZHEIMER'S DISEASE GWAS LOCI

Alessandra Chesi[1], Elisabetta Manduchi[1,2], Michelle Leonard[1], Li-San Wang[2], Sumei Lu[1], Kenyaita Hodge[1], Gerard Schellenberg[2], Matthew Johnson[1], Andrew Wells[1,2], Struan Grant[1,2], [1]*Children's Hospital of Philadelphia, Philadelphia, PA, USA;* [2]*University of Pennsylvania, Philadelphia, PA, USA. Contact e-mail: chesia@email.chop.edu*

Background: Genome wide association studies (GWAS) have revealed many loci for Alzheimer's disease (AD). However, GWAS just reports genomic signals and not necessarily the precise localization of effector genes, with eQTLs making strong inferences to only a subset of such loci. Chromatin conformation capture-based techniques that detect contacts between distant regions of the genome offer an opportunity to understand GWAS signals that principally reside in non-coding regions, thus likely influencing regulatory elements. **Methods:** To move beyond analyzing one locus at a time and to improve on the low resolution of available Hi-C data, we developed a massively parallel, high resolution Capture-C based method to simultaneously characterize the physical genome-wide interactions of all human promoters in any cell type. We applied this 'SPATIaL-seq' (genome-Scale, Promoter-focused Analysis of chromaTIn Looping) method to human iPSC-derived neural precursor cells (NPCs), a model relevant to AD. We designed a custom Agilent SureSelect library targeting both ends of DpnII restriction fragments that overlap promoters of protein-coding and noncoding transcripts, totaling 36,691 RNA baited fragments. Following sequencing, we investigated significant interactions at varying resolutions. In parallel, we generated ATAC-seq open chromatin maps to filter for informative proxy single nucleotide polymorphisms (SNPs) (r^2>0.4) for each of the 21 AD common independent sentinel SNPs reported to date beyond the *APOE* locus (Lambert et al. *Nat Genet*), yielding overall 70 SNPs (harbored in 29 DpnII fragments) for ten of these loci. **Results:** By querying our promoter 'interactome' data in NPCs at a four DpnII fragment resolution (median distance between interacting regions ~148 kb, median region size =1,440bp) we observed contacts to "open" promoters relevant to six of the original GWAS loci. Some 'nearest' genes were supported e.g. *SORL1*, while at other loci more distant genes were implicated e.g. *CHRNA2* at '*CLU*' and *SBNO2* at '*ABCA7*'. **Conclusions:** We observed informative contacts between key proxy SNPs and putative effector genes for 29% of AD GWAS loci in this particular cellular context. Further efforts in other relevant cell types could shed light on other loci. Follow-up functional studies are required to validate these findings.

O3-03-05 INTEGRATIVE NETWORK ANALYSIS IDENTIFIES RELATIONSHIPS BETWEEN METABOLOMICS, GENOMICS, AND RISK FACTORS FOR AD

Burcu F. Darst[1,2], Qiongshi Lu[1,3], Rebecca L. Koscik[1,4], Erin Jonaitis[1,5], Lindsay R. Clark[5,6,7], Kirk J. Hogan[1,5,8], Sterling C. Johnson[1,5,9,10], Corinne D. Engelman[1,2,5,11], [1]*University of Wisconsin, Madison, Madison, WI, USA;* [2]*Population Health Sciences, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA;* [3]*Department of Biostatistics and Medical Informatics, Madison, WI, USA;* [4]*The Wisconsin Alzheimer's Institute, University of Wisconsin, Madison, Madison, WI, USA;* [5]*Wisconsin Alzheimer's Institute, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA;* [6]*University of Wisconsin School of Medicine and Public Health, Madison, WI, USA;* [7]*Division of Geriatrics and Gerontology, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA;* [8]*Anesthesiology, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA;* [9]*Geriatric Research Education and Clinical Center, William S. Middleton Memorial Veterans Hospital, Madison, WI, USA;* [10]*Wisconsin Alzheimer's Disease Research Center, Madison, WI, USA;* [11]*Wisconsin Alzheimer's Disease Research Center, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA. Contact e-mail: bdarst@wisc.edu*

Background: Although Alzheimer's disease (AD) is known to have relatively high heritability, few genetic variants have been identified to date, and those that have been identified tend to have very small effect sizes. Genetic factors may only convey AD risk in individuals with certain environmental exposures, some of which can be measured by metabolomics, suggesting that a multi-omics approach could be informative for this complex trait. **Methods:** We performed an integrated network analysis to investigate relationships between metabolomics, genomics, and AD risk factors using the Wisconsin Registry for Alzheimer's Prevention (WRAP), a cohort of participants who were dementia free at enrollment and enriched with a parental history of AD. Analyses included 1,111 Caucasian participants with whole blood expression data on 11,376 genes (imputed from genome-wide genotyping using PrediXcan), 1,097 fasting plasma metabolites, and up to 19 AD risk factors (some individuals were missing particular measures, which are described in Table 1). After adjusting each of the 12,493 variables for sex and age using mixed models with random effects for sibling relationships, residuals were used to test all 78,031,278 possible pairwise Spearman correlations. Significant correlations meeting a Bonferroni-adjusted *P*-value=6.4e-10 were used to develop an undirected graphical network, which specifically focused on relationships involving AD risk factors and