

Genome-wide Association Study of Amyloid Imaging-derived Subtypes in ADNI



Mengyuan Kan, PhD¹, Zixuan Wen, MA¹, Jingxuan Bao, PhD¹, Elizabeth Mamourian, MS¹, Tianhua Zhai, PhD¹, Kwangsik Nho, PhD², Gyungah Jun, PhD³, Duygu Tosun, PhD⁴, Heng Huang, PhD⁵, Andrew J. Saykin, PsyD², Paul Thompson, PhD⁶, Christos Davatzikos, PhD⁷, Li Shen, PhD^{1*}

DEPARTMENT of
BIostatistics
EPIDEMIOLOGY &
INFORMATICS

¹Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, PA, USA; ²Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN, USA; ³Department of Biostatistics and Department of Medicine (Biomedical Genetics), Boston University, Boston, MA, USA; ⁴Department of Radiology and Imaging Sciences, University of California, San Francisco, CA, USA; ⁵Department of Computer Science, University of Maryland, College Park, MD, USA; ⁶Imaging Genetics Center, Keck School of Medicine, University of South California, Marina del Rey, CA, USA; ⁷Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA *Correspondence: li.shen@pennmedicine.upenn.edu

Introduction

Alzheimer's disease (AD) shows substantial variability in amyloid deposition patterns across individuals. Treating individuals solely based on clinical diagnosis (AD, mild cognitive impairment, or cognitively normal) as a single phenotype may dilute genetic signals and limit discovery of subtype-specific mechanisms. Data-driven subtyping offers biologically informed phenotypes for genetic studies. Subtype and Stage Inference (SuStain) is a progression-based modeling framework that can identify distinct disease subtypes and stage from imaging biomarkers. The Alzheimer's Disease Neuroimaging Initiative (ADNI), with harmonized amyloid imaging and genome-wide genotyping data, provides an ideal dataset to perform GWAS of imaging-derived amyloid subtypes.

Methods

Subtype identification: Amyloid PET scans were processed using standardized pipelines, and regional amyloid burden was quantified as standardized uptake value ratios (SUVRs) across predefined cortical regions of interest (ROIs). Higher-level regional measures—including subcortical, frontal, cingulate, parietal, temporal, occipital, and sensorimotor cortices—were derived by averaging SUVRs across ROIs within each anatomical region. These high-level regional features were then used as inputs to event-based SuStain modeling, enabling identification of two disease subtypes and probabilistic assignment of each individual based on their posterior subtype and stage probabilities.

GWAS: Genotyped data from three ADNI phases were downloaded and merged. Quality control retained common SNPs with a variant call rate >0.98, sample call rate >0.95, Hardy-Weinberg Equilibrium p-value >10⁻⁶, and minor allele frequency (MAF) ≥0.01. Individuals were included if their genetic sex matched reported sex and they had no up-to-third-degree relationships with other participants. Genotype imputation was performed using the Michigan Imputation Server with 1000 Genomes Project phase 3 genotype data as reference for individuals of European ancestry. GWAS of imaging-derived subtypes was performed using logistic regression in PLINK2, adjusting for age, sex, years of education, and the first ten principal components.

Genetic loci were determined based on index SNPs (p-value <10⁻⁶) with more than one nominally associated SNPs (p-value <10⁻⁴) in linkage disequilibrium (r² ≥0.02) within 250kb using PLINK clump function.

Functional annotation: Functional annotation of SNPs, including the genomic regions they map to and their predicted functions, was performed using ANNOVAR.

Results

1,072 individuals with genotype data and SuStain subtypes available (Subtype-0: N=937; Subtype-1: N=135) and 8,416,387 common autosomal variants were included in GWAS. No variants reached genome-wide significance; however, two genetic loci on chromosome 12 and 20 within/near two protein-coding genes showed suggestive significance (p-value < 10⁻⁶) (Figure 1).

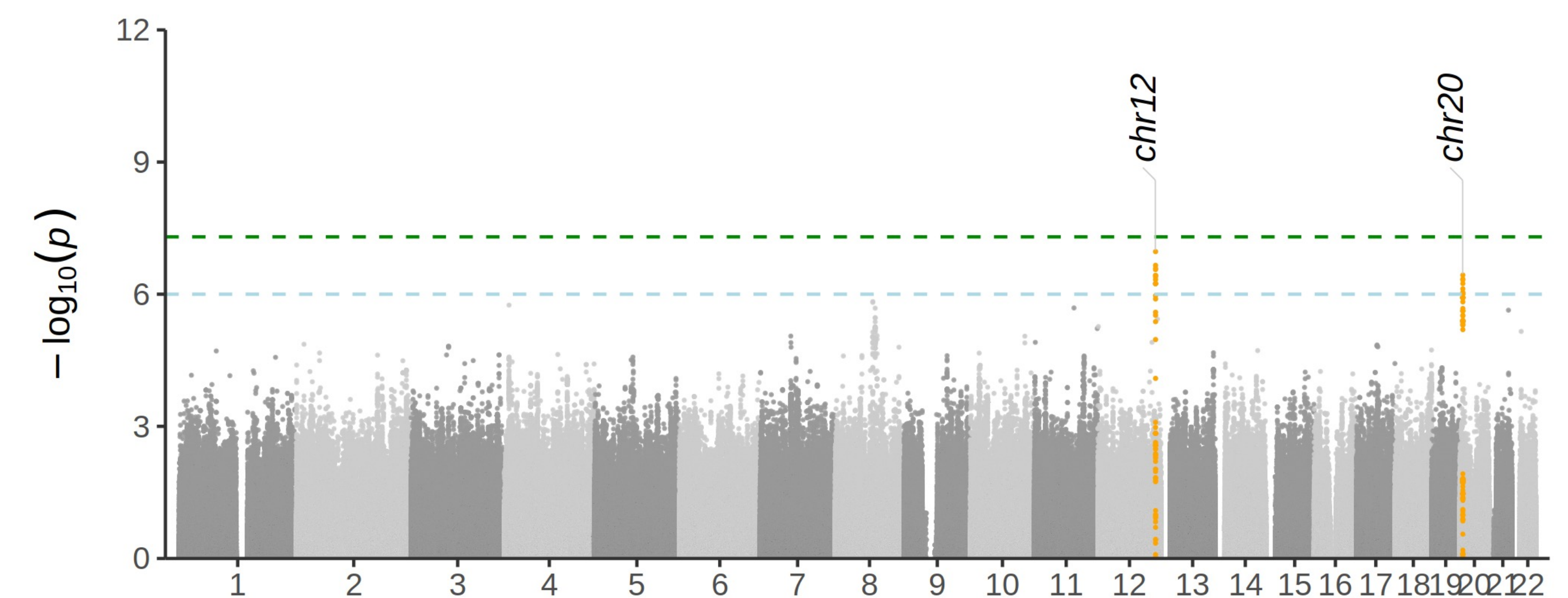


Figure 1. GWAS comparing two amyloid imaging-derived SuStain subtypes in ADNI. Genetic loci with index SNPs (p-value <10⁻⁶) are annotated. Green and blue horizontal dashed lines indicate the genome-wide significant threshold (p-value = 5 × 10⁻⁸) and the suggestive significant threshold (p-value = 10⁻⁶), respectively.

Conclusion

We identified amyloid imaging-derived SuStain subtypes independent of diagnostic categories, and subtype-related genetic variants may support further prediction of AD progression.

Acknowledgements This work is supported in part by NIH R01 AG071470, U01 AG066833, U01 AG068057, and U19 AG074879.