MULTIVARIATE GENOME-WIDE ASSOCIATION STUDY OF LANGUAGE-RELATED HUMAN BRAIN WHITE MATTER TRACTS DEMONSTRATES HIGH POLYGENICITY

Anonymous Author 1[*1,2], Anonymous Author 2[2,3], and Anonymous Author 3[3]

[*]Corresponding Author: name@domain.com
[1]This Department, University X, City, Country

1. Background

Human brain development and language are closely connected (Kuhl, 2010). White matter connections play a critical role in shaping the brain organization and function as it supports the interactions between brain regions. Prior studies linked genetic variation to language revealed a complex, and polygenic genetic architecture (Eising et al., 2022, Mekki et al., 2022). Our aim is to investigate and elucidate the neurobiological underpinnings of human brain connection support of language.

2. Methods

We used individuals from the UK Biobank cohort with both diffusion MRI and genotyping data. We excluded participants with unusual heterozygosity, high missingness, and sex mismatches. We further restricted our analyses to individuals with white British ancestry in order to avoid any possible confounding effects related to ancestry. This resulted in 31,465 individuals passing the sample QC. Using PLINK, we excluded variants with minor allele frequency < 0.01, and imputation quality INFO scores < 0.8. Multiallelic variants were also removed. We considered 35 well-known language-related white matter tracts defined from the probabilistic atlas (Rojkova et al., 2016) and extracted a set of image derived phenotypes (IDPs). We estimated the SNP-based heritability of the IDPs using GCTA and performed a multivariate genome-wide association study using MOSTest (Van der meer et al., 2020). We controlled for covariates including age, sex, genotype array type, MRI assessment center and the first ten genetic principal components capturing population genetic diversity. We performed a rank-based

inverse normalization to ensure that the distribution of the IDPs are normally distributed.

3. Results

3.1. Language-related brain structural connectivities are heritable.

All but the left Fronto Insular tract 4 phenotype showed significant SNP-based heritability (FDR-corrected $p<0.05$), ranging from 7.6% for the Left Frontal Inferior longitudinal fasciculus to 61.5% for the Corpus callosum.

3.2. Language-related brain structural connectivities are highly polygenic.

There were 268 independent genome-wide significant loci associated with different aspects of language-related brain structural connectivities. Of 173 previously reported dyslexia-associated genes (Doust et al., 2022), 38 showed genome-wide significance.

3.3. Enrichment analyses point to neurodevelopmental processes.

Significant functional enrichments of the identified genes were found in 75 biological systems encompassing different brain organizations, including the pathways of neurogenesis, neuronal differentiation, and embryonic brain expression were identified.

3.4. The neurobiological development of language-related human brain structural connectivities is active during the early to late prenatal period.

We found relatively higher mRNA expression of human language structural connectivities associated genes during early-prenatal ($p=1.46e\text{-}5$), to late prenatal ($p=3.76e{-}8$), from 9 ($p=1.23e{-}6$) to 21 ($p=2.03e{-}6$) post conception week (FDR-corrected $p < 0.05$).

4. Conclusion

In this work, we investigated the genetic architecture of language-related brain white matter tracts using state of the art genomic strategies and highlighted new candidate genes. This preliminary work represents a step forward towards understanding how genes influence the language network brain structures, complementing behavioral and brain functional studies.

References

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. Neuron, 67(5), 713-727.

Eising, E., Mirza-Schreiber, N., De Zeeuw, E. L., Wang, C. A., Truong, D. T., Allegrini, A. G., ... & Fisher, S. E. (2022). Genome-wide analyses of individual differences in quantitatively assessed reading-and language-related skills in up to 34,000 people. Proceedings of the National Academy of Sciences, 119(35), e2202764119.

Mekki, Y., Guillemot, V., Lemaître, H., Carrión-Castillo, A., Forkel, S., Frouin, V., & Philippe, C. (2022). The genetic architecture of language functional connectivity. Neuroimage, 249, 118795.

Rojkova, K., Volle, E., Urbanski, M., Humbert, F., Dell'Acqua, F., & Thiebaut de Schotten, M. (2016). Atlasing the frontal lobe connections and their variability due to age and education: a spherical deconvolution tractography study. Brain Structure and Function, 221, 1751-1766.

Van der Meer, Dennis, et al. "Understanding the genetic determinants of the brain with MOSTest." Nature communications 11.1 (2020): 3512.

Doust, C., Fontanillas, P., Eising, E., Gordon, S. D., Wang, Z., Alagöz, G., ... & Luciano, M. (2022). Discovery of 42 genome-wide significant loci associated with dyslexia. Nature genetics, 54(11), 1621-1629.