

# Ancestry-specific gene expression in peripheral monocytes mediates risk of neurodegenerative disease

Aaron Z Wagen<sup>1,5</sup>, Regina H Reynolds<sup>1,5</sup>, Jia Nee Foo<sup>6</sup>, Aine Fairbrother-Browne, Emil K Gustavsson<sup>1,5</sup>, Sarah Galgiano-Turin, Nicholas W Wood<sup>3,5</sup>, Cornelis Blauwendraat<sup>4</sup>, Sonia Gandhi<sup>1,3,5</sup>, Mina Ryten<sup>1,5,7</sup>

<sup>1</sup>Department of Genetics and Genomic Medicine, Great Ormond Street Institute of Child Health, University College London, London, UK

<sup>2</sup>The Francis Crick Institute, 1 Midland Road, London, UK

<sup>3</sup>Department of Clinical and Movement Neurosciences, UCL Queen Square Institute of Neurology, Queen Square, London, UK

<sup>4</sup>Center for Alzheimer's and Related Dementias, National Institute on Aging and National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

<sup>5</sup>Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase, MD 20815, USA.

<sup>6</sup>Lee Kong Chian School of Medicine, Nanyang Technological University Singapore, 11 Mandalay Road, Singapore 308232, Singapore

<sup>7</sup>Dementia Research Institute, Department of Clinical Neuroscience, Cambridge University, Cambridge UK

Licence : CC-BY 4.0

## Abstract

It is hypothesised that peripheral immune states responding to regional environmental triggers contribute to central neurodegeneration. Region-specific genetic selection pressures require this hypothesis to be assessed in an ancestry specific manner. Here we utilise genome-wide association studies and expression quantitative trait loci from African, East Asian and European ancestries to show that genes causing neurodegeneration are preferentially expressed in innate rather than adaptive immune cells, and that expression of these genes mediates the risk of neurodegenerative disease in monocytes in an ancestry-specific manner.

## Main

The role of ancestry-specific risk is increasingly recognised in neurodegenerative diseases (NDDs). Rectifying the historical focus on European populations, recent efforts have explored the clinical and genetic diversity of conditions such as Alzheimer's (AD) and Parkinson's diseases (PD), and Frontotemporal dementia (FTD)<sup>1-4</sup>. This includes appreciation of the differences in the prevalence and pathology of NDDs across populations<sup>5,6</sup>, as well as differences in the genetic architecture of both monogenic and complex forms of disease<sup>3,4,7,8</sup>. Central to this has been the discovery of novel disease loci in ancestry-specific genome-wide association studies (GWAS) of NDDs<sup>7-11</sup>, and of variability of gene expression in post-mortem brains<sup>6,12</sup>. Despite these studies being limited by relatively small case numbers, they have identified novel risk loci with unexpectedly large effect sizes, emphasising the significant effect ancestry can have in neurodegeneration.

Simultaneously, the immune system is increasingly implicated in NDDs, including AD, PD and FTD. Within the CNS, the role of microglia, the parenchymal myeloid cells of the CNS, as well as immune related macro-glial states are gaining prominence<sup>13-15</sup>. Furthermore, both positron emission tomography (PET) based imaging approaches, as well as cerebrospinal fluid (CSF) biochemical analyses have demonstrated increases in inflammatory biomarkers in NDDs, with levels correlated to clinical outcomes. This central immune response is mirrored peripherally, and there is suggestive evidence that the peripheral response is fundamentally involved in disease pathogenesis<sup>16,17</sup>. Clinically, NDDs can commence with a peripheral prodrome, for example, anosmia and constipation in PD. Epidemiological data shows that NDDs have followed viral pandemics throughout history<sup>18</sup>, and that recent viral infection increases the risk of NDD<sup>19,20</sup>. Interestingly, this appears to occur without evidence of replicating viral particles, raising the possibility of a virally-triggered peripheral immune state that affects CNS function. Indeed, the concept of brain-immune crosstalk has been advanced with the potential to explain how environmental factors might contribute to neurodegenerative risk<sup>21</sup>, and underpinning recent therapeutic strategies to treat these diseases<sup>22-24</sup>.

In this work, we hypothesise that genes causally linked to AD, PD and FTD are themselves active in the peripheral immune response, and that this activity contributes to the risk of disease in an ancestry-specific manner. There is already some evidence to support this idea in PD, where studies in mice have shown that alpha-synuclein (*SNCA*), the major constituent of Lewy bodies and a cause of PD, has a protective role against viral infection centrally<sup>25</sup>, and can mediate antigen presentation and the inflammatory response in the peritoneum<sup>26</sup>. Similarly, common pathogenic variants in leucine-rich repeat kinase-2 (*LRRK2*) modulate inflammation in response to CNS and systemic infections<sup>27,28</sup>, as does a pathogenic *GBA1* variant with respect to viral encephalitis<sup>29</sup>. This association is also seen in AD and FTD, where, respectively, *TREM2*<sup>30,31</sup> and *C9orf72*<sup>32</sup> have peripheral immunomodulatory effects. Genetic analyses have linked the heritability of NDDs with peripheral immune cell types<sup>33,34</sup>. However, the risk of NDDs caused by the expression of genes in peripheral immune cells across diverse ancestries has not been investigated.

To study this, we systematically explored the role of 39 NDD genes causally implicated in AD, PD or FTD across peripheral immune cells (Supplementary Table 1, Methods). Given that both ancestry and activation state are well recognised to affect gene expression in

peripheral immune cells, we used a multi-ancestry study of 22 different cell-types to study gene expression<sup>35</sup>. This dataset was derived from 222 African (n = 82), East Asian (n = 60) and European (n = 80) donors, either at baseline or after activation with SARS-CoV-2 or Influenza A viruses.

We found that 32 of the 39 NDD-causing genes (82%) were expressed across peripheral immune cells (Figure 1a). Using phenotypic QTL data, we noted that 29 of these 32 genes (91%) were also causally associated with a white blood cell count metric in peripheral blood. Indeed, 20-83% of the NDD-causing genes, including *APOE*, *SNCA* and *MAPT*, were loci that influenced macrophage, monocyte, lymphocyte or total white cell count in at least one ancestry (Figure 1b, Supplementary Figure 1).

Next, we explored the expression patterns of NDD-causing genes by cell type, activation state, and ancestry. Over 80% of these genes showed maximal expression in innate immune cell types, most commonly in monocytes (16 out of 34 cell types) and NK cells (8 out of 34 cell types, Figure 1d). Genes with maximal expression in myeloid cell types (including *APOE*, *PSEN1*, *TREM2*, *LRRK2* and *C9orf72* in monocytes; and *APP* and *PSEN2* in dendritic cells) showed greatest cell-type specificity (mean  $\tau$ -statistic = 0.82, Figure 1e). In contrast, genes with maximal expression in NK cells were more broadly expressed (mean  $\tau$ -statistic 0.63, p value for two sample T-test = 0.01). Three NDD genes (*TREM2*, *APOE* and *LRRK2*) showed high specificity ( $\tau$ -statistic > 0.9) to unstimulated activation states as compared to those treated with influenza or SARS-CoV-2 (Figure 1f). Notably, these three genes all had maximal expression in monocytes. Myeloid cells also showed the greatest number of genes differentially expressed across ancestries (European relative to African ancestry) with 10 genes identified, including *APP*, *BIN1*, *GBA1*, *RAB32*, *C9orf72* and *GRN* (Figure 1c). Taken together, these results show that genes central to NDDs are preferentially expressed in peripheral immune cells of myeloid lineage, and that these cells contain the greatest proportion of cell type- and ancestry-specific NDD gene expression.

We then explored whether the expression of these genes in the periphery might be causally linked with disease. To do this, we applied a well-established genetic approach, colocalisation analysis, to explore the possibility that NDD gene expression in the peripheral immune system might be a driver of disease risk<sup>36</sup>. This analysis uses a Bayesian approach to ascertain the posterior probabilities that association signals of two traits at a genetic locus share a causal variant (posterior probability of hypothesis 4, PPH4), or whether these association signals are distinct (posterior probability of hypothesis 3, PPH3, Methods). Crucially, these analyses were conducted in an ancestry-specific manner, as we reasoned that the peripheral immune system would be the first to encounter geographically restricted microbes and pollutants, leading to variable environmental selection pressures<sup>37</sup>. This required use of very recently produced ancestry-specific GWAS, as well as corresponding ancestry-aware functional genomic annotations, namely expression quantitative trait loci (eQTL). We utilised the eQTL available from the same multi-ancestry study used to explore gene expression above, due to the availability of data across multiple populations, cell types and activation states<sup>35</sup>.

We first explored ancestry-specific AD GWASs with multi-ancestry eQTL data. When using a European AD GWAS<sup>38</sup>, we identified a colocalising signal in *BIN1* in monocytes (PPH4 = 0.93, Supplementary Table 2, Supplementary Fig. 2). There was a positive correlation

between effect sizes of the GWAS and eQTL single nucleotide polymorphisms (SNPs) at this locus, suggesting that increased *BIN1* expression in peripheral monocytes correlated with increased risk of AD. While the role of *BIN1* in modulating AD is known within the CNS<sup>39</sup>, this result suggests a causative role for the gene in AD the periphery. Noting that *BIN1* is not a significant locus in the East Asian AD GWAS<sup>11</sup>, we found no colocalising signal at *BIN1* when this GWAS was tested against the multi-ancestry eQTL (Supplementary Table 3).

Focusing on PD, with a GWAS derived from individuals of European ancestry<sup>40</sup>, we identified a colocalising signal within the PD-associated gene *LRRK2*, mediating the risk of PD and gene expression in SARS-CoV-2-treated CD14 monocytes (Figure 2a, Supplementary Table 4, Supplementary Figure 3). To check whether this result was ancestry specific, we undertook the same analysis with an East Asian PD GWAS<sup>9</sup>, finding that the result was underpowered at this locus ( $PPH3 + PPH4 < 0.75$ , Figure 2a, Supplementary Table 5, Supplementary Figure 4, Methods). Given this, we checked the East Asian GWAS against an East-Asian-specific eQTL dataset<sup>41</sup>, finding the association signals underlying *LRRK2* expression in monocytes were distinct from those underlying PD risk in all monocyte cell types tested ( $PPH3 > 0.99$ , Supplementary Table 6, Supplementary Figure 5).

Focusing on the *SNCA* locus, arguably the most important gene in PD, there was insufficient power when the multi-ancestry eQTL was tested against both the European and East Asian PD GWASs ( $PPH4 + PPH3 < 0.75$ , Figure 2a, Supplementary Table 4, 5). However, noting that the  $PPH4/PPH3$  ratio was high at 3.5, we checked this locus using an East-Asian-specific PD eQTL and GWAS. We found that in this East-Asian-specific analysis that there was a significant colocalisation in *SNCA* expression in CD16+ monocytes (Figure 2b, Supplementary Table 6, Supplementary Figure 5). Using the multi-ancestry eQTL with an African PD GWAS, there was a suggestive but underpowered colocalising signal at *BIN1* in CD14+ monocytes ( $PPH4=0.71$ ,  $PPH3+PPH4=0.73$ ,  $PPH4/PPH3$  ratio=39.9). We could not validate this result in an African-specific eQTL dataset given the lack of power in the available datasets (Supplementary Table 8).

As expected, at the *LRRK2* locus, there was a positive correlation between effect sizes in European populations, implying that increased *LRRK2* expression associates with increased PD risk (Figure 2b). In contrast, at the *SNCA* locus in East Asian populations, the correlation between effect sizes was negative, suggesting that decreased *SNCA* expression in CD16+ monocytes was associated with increased PD risk. While this result appears to be counter-intuitive given that increased *SNCA* in the CNS is associated with PD, this finding is in keeping with multiple existing genomic and proteomic studies, which show that PD is associated with a decreased *SNCA* RNA and protein in blood and CSF<sup>33,42-44</sup>.

In summary, we show that genes causally implicated in NDDs mediate the risk of these diseases partially through effects in the peripheral immune system. Crucially, the expression of these genes as well as the risk mediated by them were ancestry dependent. This suggests not only that neurodegeneration has a peripheral component, but that across different ancestries regional selection pressures may have differentially modulated disease mechanisms. The results also suggest that genes that have a detrimental effect within the CNS may be protective peripherally<sup>25,26</sup>. In addition to emphasising the importance of ancestry-specific functional genetic annotations, this work defines innate immune cells, and specifically monocytes, as potential targets for treatments for neurodegeneration.

## Methods

### Neurodegenerative gene list curation

Genes causally implicated in AD were identified as described by Neuner et al<sup>45</sup>, selecting the 18 genes with increased confidence of causation. Fourteen genes causally implicated in PD with high or very high confidence were selected as reported by Blauwendraat et al<sup>46</sup>, with the addition of *RAB32* whose role in PD has recently emerged<sup>47,48</sup>. Causative FTD genes were selected from Antonioni et al<sup>49</sup>, including the 6 genes that contribute  $\geq 1\%$  of disease frequency.

### eQTL preprocessing

We systematically searched PubMed for eQTL studies of peripheral blood from the last 10 years with the following features: i) included more than 150 participants; ii) specifically stated ancestry; iii) divided cells into clearly defined subtypes by single cell sequencing or cytometric approaches; and iv) included an activation state generated either with an immunogenic trigger or through the inclusion of participants with immunological disease. The Aquino et al dataset<sup>35</sup> was the only dataset identified that included participants from African, Asian and European ancestry. Noting that it analysed 22 different cell types, and included activation state data with samples stimulated with either SARS-CoV-2 virus or influenza A virus, it was selected as the baseline dataset for this study. For colocalisation analyses, the Aquino et al dataset was supplemented by use of Ota et al<sup>41</sup>, as a large dataset specific to East Asian ancestry. The only African-specific ancestry eQTLs that were identified included two datasets from African and African admixed populations, one exploring monocytes<sup>50</sup>, and the other macrophages<sup>51</sup>.

Where relevant, summary statistics were processed by lifting over genomic locations from GRCh37 to GRCh38. SNP genomic coordinates were mapped to Reference SNP cluster IDs (rsIDs) using the `SNPlocs.Hsapiens.dbSNP144.GRCh38` package, and ancestry-specific minor allele frequencies were imported from the `MafDb.1Kgenomes.phase3.GRCh38` package<sup>52</sup>.

### Blood count phenotype-QTL processing

Significant associations between genetic variation and blood count metrics were derived from a multi-ancestry phenotype QTL<sup>53</sup>. Significant traits between blood cell metrics of interest (neutrophil, lymphocyte, monocyte and total white blood cell count) and NDD genes were exported using Open Targets (<https://genetics.opentargets.org/>, RRID:SCR\_014622)<sup>54</sup>.

### Gene expression analysis

The gene expression and differential gene expression datasets used in this study were generated by Aquino et al<sup>35</sup>. To facilitate comparison between cell types, the raw transcript per million (TPM) for each gene was normalised by the maximum TPM per gene across cell types and activation states, giving the metric of proportion of expression per gene. Noting that cells in Aquino et al were treated with both influenza A and SARS-CoV-2 viruses, this



was also normalised using the above approach. Of each cell type that showed maximal expression of a gene, the maximal cell state was defined as the viral exposure that resulted in the maximal expression of the gene for that celltype.

To measure the cell-type specificity of gene expression, the  $\tau$ -statistic was calculated using the method defined by Yanai et al<sup>55</sup>:

$$\tau = \frac{\sum_{i=1}^N (1-x_i)}{N-1},$$

where N is the number of tissues and  $x_i$  is the expression profile component normalised by the maximal component value. For each cell type that contained the highest expression per gene, a further  $\tau$ -statistic was calculated to measure cell-state specificity of response to viral stimulation.

Differential gene expression analysis results were utilised from Aquino et al. They compared European relative to African populations using a linear regression model, adjusted for individual donor age, cellular mortality and cellular composition as described<sup>35</sup>.

## GWAS preprocessing

GWAS summary statistics from 3 ancestry-specific studies of PD were used: European<sup>40</sup>, African<sup>10</sup>, and East Asian<sup>9</sup> ancestry. Two ancestry-specific GWAS studies were used for AD: European<sup>38</sup> and Asian<sup>11</sup>. Where relevant, summary statistics were processed by lifting over genomic locations from GRCh37 to GRCh38. SNP genomic coordinates were mapped to rsIDs using the SNPlocs.Hsapiens.dbSNP144.GRCh38 package (<https://bioconductor.org/packages/release/data/annotation/html/SNPlocs.Hsapiens.dbSNP144.GRCh38.html>)<sup>56</sup>, and ancestry-specific minor allele frequencies were imported from the MafDb.1Kgenomes.phase3.GRCh38 package (<https://www.bioconductor.org/packages/release/data/annotation/html/MafDb.1Kgenomes.phase3.GRCh38.html>)<sup>52</sup>.

## Colocalisation analysis

To evaluate the probability that GWAS loci and eQTLs share a single causal variant, a colocalisation analysis was performed using coloc (version 5.1.0.1, <https://cran.r-project.org/package=coloc>) and colochelpR (version 0.99.1, <http://dx.doi.org/10.5281/zenodo.5011869>)<sup>36,57</sup>. GWAS loci within 1 Mb significant GWAS SNPs were explored. The prior probability that any random SNP in the region is associated with the GWAS ( $p_1$ ) or eQTL ( $p_2$ ) was set to the default  $10^{-4}$ , whereas the prior probability that any random SNP in the region is associated with both traits ( $p_{12}$ ) was set to  $10^{-5}$ . Using a Bayesian approach, posterior probabilities were calculated for four different hypotheses that together sum to 1. Hypothesis 1 and 2 measure associations of the GWAS locus or the eQTL locus, where there is no sufficient power in the two studies to compare the two (as defined by PPH4 + PPH3 < 0.75). Hypothesis 3 (PPH3) measures the probability that the traits have distinct causal variants at a locus, while hypothesis 4 (PPH4) measures the probability that a locus is colocalised as a result of a single causal variant.

## Statistical analysis

All statistical analysis was undertaken in R version 4.2.0 (RRID:SCR\_001905, <https://www.r-project.org/>).

## Bibliography

1. Ben-Shlomo, Y. *et al.* The epidemiology of Parkinson's disease. *The Lancet* **403**, 283–292 (2024).
2. Dehghani, N., Bras, J. & Guerreiro, R. How understudied populations have contributed to our understanding of Alzheimer's disease genetics. *Brain* **144**, 1067–1081 (2021).
3. Koros, C. *et al.* The Landscape of Monogenic Parkinson's Disease in Populations of Non-European Ancestry: A Narrative Review. *Genes* **14**, 2097 (2023).
4. Reitz, C., Pericak-Vance, M. A., Foroud, T. & Mayeux, R. A global view of the genetic basis of Alzheimer disease. *Nat. Rev. Neurol.* **19**, 261–277 (2023).
5. Naslavsky, M. S. *et al.* Global and local ancestry modulate APOE association with Alzheimer's neuropathology and cognitive outcomes in an admixed sample. *Mol. Psychiatry* **27**, 4800–4808 (2022).
6. Schlesinger, D. *et al.* African ancestry protects against Alzheimer's disease-related neuropathology. *Mol. Psychiatry* **18**, 79–85 (2013).
7. Kim, J. J. *et al.* Multi-ancestry genome-wide association meta-analysis of Parkinson's disease. *Nat. Genet.* **56**, 27–36 (2024).
8. Lake, J. *et al.* Multi-ancestry meta-analysis and fine-mapping in Alzheimer's disease. *Mol. Psychiatry* **28**, 3121–3132 (2023).
9. Foo, J. N. *et al.* Identification of Risk Loci for Parkinson Disease in Asians and Comparison of Risk Between Asians and Europeans: A Genome-Wide Association Study. *JAMA Neurol.* **77**, 746–754 (2020).
10. Rizig, M. *et al.* Genome-wide Association Identifies Novel Etiological Insights Associated with Parkinson's Disease in African and African Admixed Populations. *medRxiv* 2023.05.05.23289529 (2023) doi:10.1101/2023.05.05.23289529.

11. Shigemizu, D. *et al.* Ethnic and trans-ethnic genome-wide association studies identify new loci influencing Japanese Alzheimer's disease risk. *Transl. Psychiatry* **11**, 1–10 (2021).
12. Benjamin, K. J. M. *et al.* Analysis of gene expression in the postmortem brain of neurotypical Black Americans reveals contributions of genetic ancestry. *Nat. Neurosci.* **27**, 1064–1074 (2024).
13. Bartels, T., De Schepper, S. & Hong, S. Microglia modulate neurodegeneration in Alzheimer's and Parkinson's diseases. *Science* **370**, 66–69 (2020).
14. Sadick, J. S. *et al.* Astrocytes and oligodendrocytes undergo subtype-specific transcriptional changes in Alzheimer's disease. *Neuron* **110**, 1788-1805.e10 (2022).
15. D'Sa, K. *et al.* Alpha-synuclein aggregates trigger anti-viral immune pathways and RNA editing in human astrocytes. 2024.02.26.582055 Preprint at <https://doi.org/10.1101/2024.02.26.582055> (2024).
16. Berriat, F., Lobsiger, C. S. & Boillée, S. The contribution of the peripheral immune system to neurodegeneration. *Nat. Neurosci.* **26**, 942–954 (2023).
17. Castellani, G., Croese, T., Peralta Ramos, J. M. & Schwartz, M. Transforming the understanding of brain immunity. *Science* **380**, eabo7649 (2023).
18. Limphaibool, N., Iwanowski, P., Holstad, M. J. V., Kobylarek, D. & Kozubski, W. Infectious Etiologies of Parkinsonism: Pathomechanisms and Clinical Implications. *Front. Neurol.* **10**, (2019).
19. Kettunen, P., Koistinaho, J. & Rolova, T. Contribution of CNS and extra-CNS infections to neurodegeneration: a narrative review. *J. Neuroinflammation* **21**, 152 (2024).
20. Levine, K. S. *et al.* Virus exposure and neurodegenerative disease risk across national biobanks. *Neuron* **111**, 1086-1093.e2 (2023).
21. Schwartz, M. & Deczkowska, A. Neurological Disease as a Failure of Brain–Immune Crosstalk: The Multiple Faces of Neuroinflammation. *Trends Immunol.* **37**, 668–679 (2016).
22. Rosenzweig, N. *et al.* PD-1/PD-L1 checkpoint blockade harnesses monocyte-derived



- macrophages to combat cognitive impairment in a tauopathy mouse model. *Nat. Commun.* **10**, 465 (2019).
23. Sarazin, M. *et al.* The path to next-generation disease-modifying immunomodulatory combination therapies in Alzheimer's disease. *Nat. Aging* **4**, 761–770 (2024).
  24. Schwartz, M. Can immunotherapy treat neurodegeneration? *Science* **357**, 254–255 (2017).
  25. Beatman, E. L. *et al.* Alpha-Synuclein Expression Restricts RNA Viral Infections in the Brain. *J. Virol.* **90**, 2767–2782 (2015).
  26. Alam, M. M. *et al.* Alpha synuclein, the culprit in Parkinson disease, is required for normal immune function. *Cell Rep.* **38**, 110090 (2022).
  27. Shutinoski, B. *et al.* Lrrk2 alleles modulate inflammation during microbial infection of mice in a sex-dependent manner. *Sci. Transl. Med.* **11**, eaas9292 (2019).
  28. Fava, V. M. *et al.* Pleiotropic effects for Parkin and LRRK2 in leprosy type-1 reactions and Parkinson's disease. *Proc. Natl. Acad. Sci.* **116**, 15616–15624 (2019).
  29. Melamed, S. *et al.* Innate immune response in neuronopathic forms of Gaucher disease confers resistance against viral-induced encephalitis. *Acta Neuropathol. Commun.* **8**, 144 (2020).
  30. Griciuc, A. & Tanzi, R. E. The role of innate immune genes in Alzheimer's disease. *Curr. Opin. Neurol.* **34**, 228 (2021).
  31. Matos, A. de O., Dantas, P. H. dos S., Queiroz, H. A. G. de B., Silva-Sales, M. & Sales-Campos, H. TREM-2: friend or foe in infectious diseases? *Crit. Rev. Microbiol.* **50**, 1–19 (2024).
  32. McCauley, M. E. *et al.* C9orf72 in myeloid cells suppresses STING-induced inflammation. *Nature* **585**, 96–101 (2020).
  33. Reynolds, R. H. *et al.* Local genetic correlations exist among neurodegenerative and neuropsychiatric diseases. *Npj Park. Dis.* **9**, 70 (2023).
  34. Lona-Durazo, F., Reynolds, R. H., Scholz, S. W., Ryten, M. & Gagliano Taliun, S. A. Regional genetic correlations highlight relationships between neurodegenerative disease

- loci and the immune system. *Commun. Biol.* **6**, 1–11 (2023).
35. Aquino, Y. *et al.* Dissecting human population variation in single-cell responses to SARS-CoV-2. *Nature* **621**, 120–128 (2023).
  36. Wallace, C. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genet.* **16**, e1008720 (2020).
  37. Brodin, P. & Davis, M. M. Human immune system variation. *Nat. Rev. Immunol.* **17**, 21–29 (2017).
  38. Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019).
  39. Sudwarts, A. *et al.* BIN1 is a key regulator of proinflammatory and neurodegeneration-related activation in microglia. *Mol. Neurodegener.* **17**, 33 (2022).
  40. Nalls, M. A. *et al.* Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* **18**, 1091–1102 (2019).
  41. Ota, M. *et al.* Dynamic landscape of immune cell-specific gene regulation in immune-mediated diseases. *Cell* **184**, 3006–3021.e17 (2021).
  42. Craig, D. W. *et al.* RNA sequencing of whole blood reveals early alterations in immune cells and gene expression in Parkinson's disease. *Nat. Aging* **1**, 734–747 (2021).
  43. Irmady, K. *et al.* Blood transcriptomic signatures associated with molecular changes in the brain and clinical outcomes in Parkinson's disease. *Nat. Commun.* **14**, 3956 (2023).
  44. Mollenhauer, B. *et al.*  $\alpha$ -Synuclein and tau concentrations in cerebrospinal fluid of patients presenting with parkinsonism: a cohort study. *Lancet Neurol.* **10**, 230–240 (2011).
  45. Neuner, S. M., Tcw, J. & Goate, A. M. Genetic architecture of Alzheimer's disease. *Neurobiol. Dis.* **143**, 104976 (2020).
  46. Blauwendraat, C., Nalls, M. A. & Singleton, A. B. The genetic architecture of Parkinson's disease. *Lancet Neurol.* **19**, 170–178 (2020).
  47. Gustavsson, E. K. *et al.* RAB32 Ser71Arg in autosomal dominant Parkinson's disease:

- linkage, association, and functional analyses. *Lancet Neurol.* **23**, 603–614 (2024).
48. Hop, P. J. *et al.* Systematic rare variant analyses identify RAB32 as a susceptibility gene for familial Parkinson's disease. *Nat. Genet.* **56**, 1371–1376 (2024).
49. Antonioni, A. *et al.* Frontotemporal Dementia, Where Do We Stand? A Narrative Review. *Int. J. Mol. Sci.* **24**, 11732 (2023).
50. Quach, H. *et al.* Genetic adaptation and neandertal admixture shaped the immune system of human populations. *Cell* **167**, 643-656.e17 (2016).
51. Nédélec, Y. *et al.* Genetic ancestry and natural selection drive population differences in immune responses to pathogens. *Cell* **167**, 657-669.e21 (2016).
52. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
53. Chen, M.-H. *et al.* Trans-ethnic and Ancestry-Specific Blood-Cell Genetics in 746,667 Individuals from 5 Global Populations. *Cell* **182**, 1198-1213.e14 (2020).
54. Ochoa, D. *et al.* The next-generation Open Targets Platform: reimaged, redesigned, rebuilt. *Nucleic Acids Res.* **51**, D1353–D1359 (2023).
55. Yanai, I. *et al.* Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21**, 650–659 (2005).
56. Hervé Pagès. SNPlocs.Hsapiens.dbSNP144.GRCh38. (2017).
57. Reynolds, R. H. RHR Reynolds/colochelpR: v0.99.1. *Zenodo* (2021)  
doi:10.5281/zenodo.5011869.

## Code availability

The complete code utilised in this work can be accessed at [https://github.com/aaronwagen/peripheral\\_immune\\_neurodegeneration](https://github.com/aaronwagen/peripheral_immune_neurodegeneration), including knitted markdowns with integrated analysis, results, and packages utilised.

## Data availability

Summary statistics for Foo et al, Rizig et al and Shigemizu et al were supplied by the authors. Otherwise, all datasets utilised are from public sources (Supplementary Table 9).

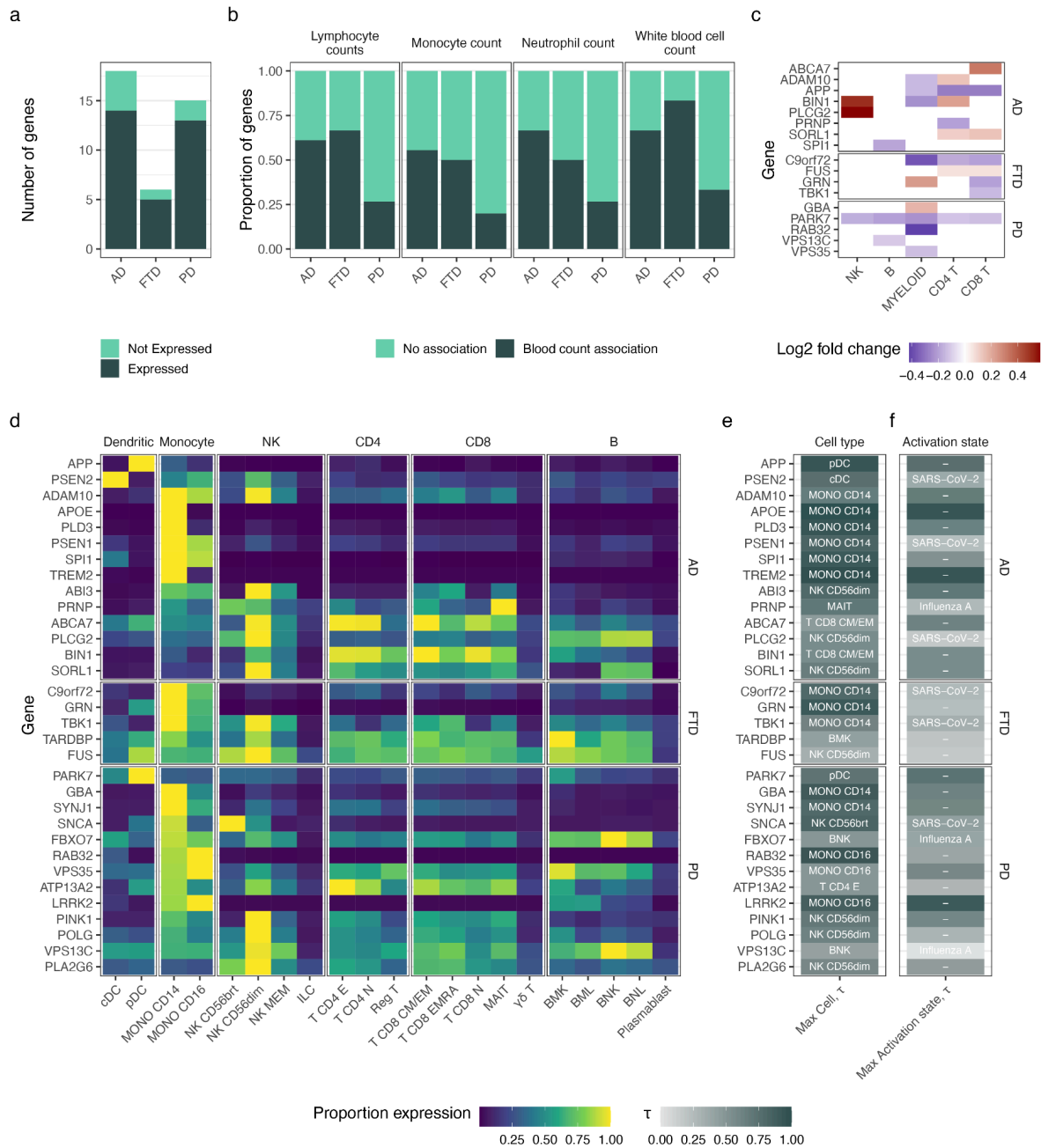
## Acknowledgements

This research was funded in part by Aligning Science Across Parkinson's [ASAP-000509 and ASAP-000463] through the Michael J. Fox Foundation for Parkinson's Research (MJFF). This work utilized the computational resources of the NIH HPC Biowulf cluster (<http://hpc.nih.gov>). This work was supported in part by the Intramural Research Program of the National Institutes of Health including: the Center for Alzheimer's and Related Dementias, within the Intramural Research Program of the National Institute on Aging and the National Institute of Neurological Disorders and Stroke.

AZW was supported through the award of a Clinical Research Fellowship funded by the Wolfson Foundation and Eisai Ltd. SG was supported by Wellcome (100172/Z/12/2) and is currently an MRC Senior Clinical Fellow (MR/T008199/1). MR was supported by the UK Medical Research Council (MRC) through her award of Tenure-track Clinician Scientist Fellowship (MR/N008324/1).

## Competing interests

RHR is currently employed by CoSyne Therapeutics (Lead Computational Biologist). All work performed for this publication was performed in her own time, and not as a part of her duties as an employee.

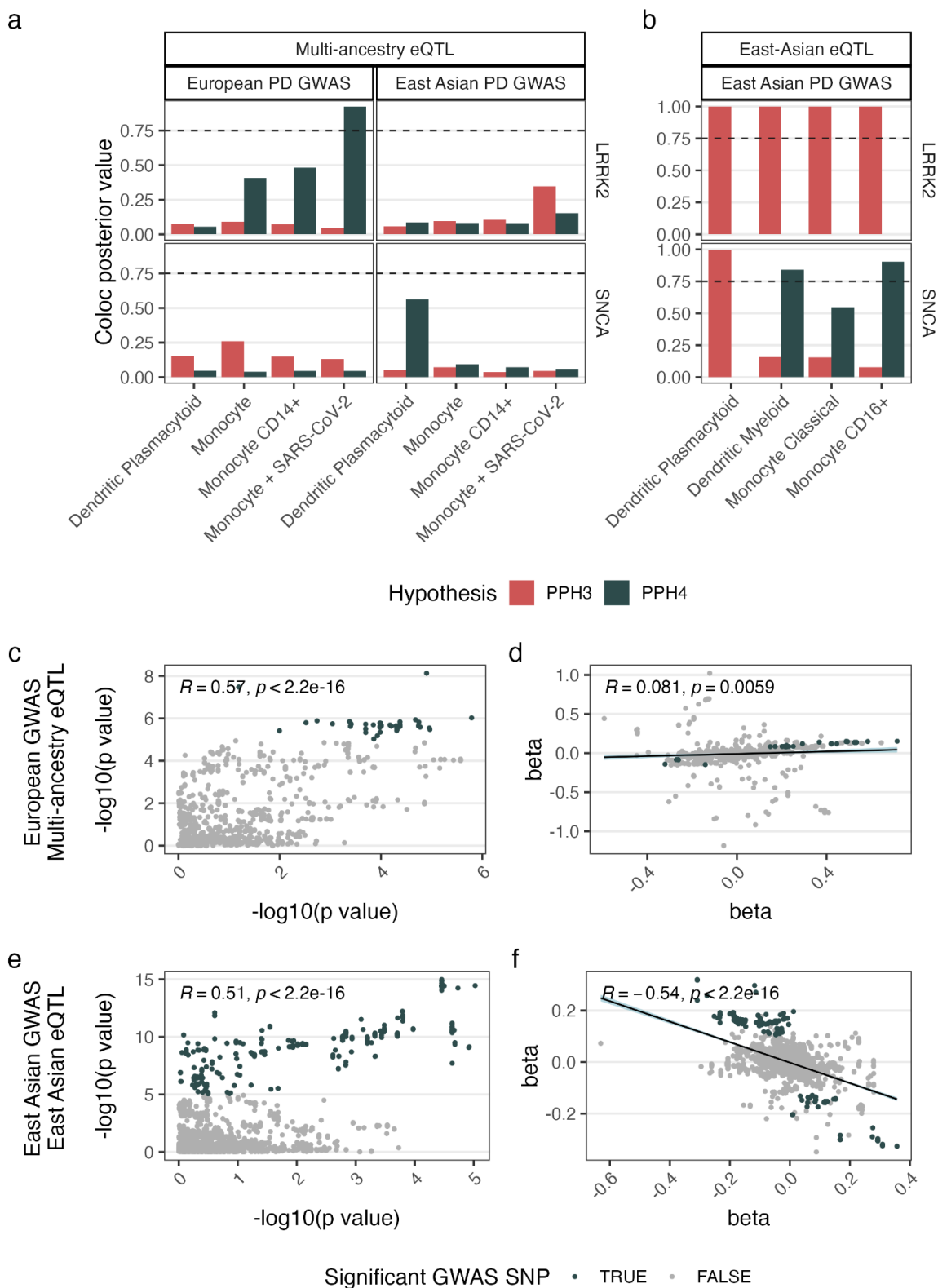


**Fig. 1: Genes causing neurodegeneration are expressed in peripheral immune cells in a cell-type, cell-state and ancestry specific manner.**

**a**, Number of genes causing neurodegeneration expressed in multi-ancestry study of peripheral blood cells (Aquino et al<sup>35</sup>). **b**, Proportion of genes expressed in peripheral blood cells (14 AD, 13 PD and 5 FTD) that have an association with lymphocyte, monocyte, neutrophil and total white blood cell count (left-to-right, from Chen et al<sup>53</sup>). **c**, Results of differential gene expression analysis, showing genes significantly differentially expressed in blood cell lineages in European relative to African ancestry. Empty tile denotes no significant differential expression. **d** Heatmap showing proportion of gene expression of genes causing neurodegeneration across cell types and lineages in multi-ancestry study. The cell-state with maximal gene expression is shown, with expression represented as proportion of maximal expression per gene over all cell types. **e**, Column showing cell-type specificity of genes, with labels showing the cell-type with greatest gene expression, and shaded by  $\tau$ -statistic

(darker suggests greater specificity). **f**, Column showing cell-state specificity, where, for the cell-type that showed maximal expression for each gene, the label shows the exposure that caused the maximal expression (or '-' where maximal expression was in unstimulated cells). Shading shows the  $t$ -statistic of the labelled activation state compared to the other activations.





**Fig 2. Expression of PD genes in peripheral immune cells mediate risk of PD in an ancestry dependent manner. a,** Colocalisation of *LRRK2* (top) and *SNCA* (bottom) between multi-ancestry eQTL and PD GWAS from European (left) and East Asian (right) ancestry. PPH3 measures discrete risk, PPH4 measures colocalised risk (significant colocalisation = posterior hypothesis > 0.75, represented by the dashed line). **b,**

Colocalisation of *SNCA* using eQTL and GWAS from East Asian ancestry.

**c-f**, Locus analysis of significant colocalising signals between PD GWASs and eQTLs.

Results from *LRRK2* locus using European PD GWAS and multi-ancestry eQTL in SARS-CoV-2 treated monocytes, showing  $-\log_{10}(p \text{ values})$  (**c**) and betas (**d**). Results from *SNCA* locus using East Asian PD GWAS and East Asian eQTL in CD16+ monocytes, showing correlation of p-values (**e**) and beta values (**f**). Text shows Pearson's correlation coefficient and p value, with beta plots also showing the regression line of best fit shaded with 95% confidence intervals.