



Article

Transcript Variants of Genes Involved in Neurodegeneration Are Differentially Regulated by the APOE and MAPT Haplotypes

Sulev Koks ^{1,2,*} , Abigail L. Pfaff ^{1,2}, Vivien J. Bubb ³ and John P. Quinn ³ 

¹ Perron Institute for Neurological and Translational Science, Perth, WA 6009, Australia; Abigail.Pfaff@murdoch.edu.au

² Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Perth, WA 6150, Australia

³ Department of Pharmacology and Therapeutics, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool L69 3BX, UK; jillbubb@liverpool.ac.uk (V.J.B.); jqinn@liverpool.ac.uk (J.P.Q.)

* Correspondence: sulev.koks@perron.uwa.edu.au; Tel.: +61-(0)-8-6457-0313

Abstract: Genetic variations at the Apolipoprotein E (ApoE) and microtubule-associated protein tau (MAPT) loci have been implicated in multiple neurodegenerative diseases, but their exact molecular mechanisms are unclear. In this study, we performed transcript level linear modelling using the blood whole transcriptome data and genotypes of the 570 subjects in the Parkinson's Progression Markers Initiative (PPMI) cohort. ApoE, MAPT haplotypes and two SNPs at the SNCA locus (rs356181, rs3910105) were used to detect expression quantitative trait loci eQTLs associated with the transcriptome and differential usage of transcript isoforms. As a result, we identified 151 genes associated with the genotypic variations, 29 *cis* and 122 *trans* eQTL positions. Profound effect with genome-wide significance of ApoE e4 haplotype on the expression of TOMM40 transcripts was identified. This finding potentially explains in part the frequently established genetic association with the APOE e4 haplotypes in neurodegenerative diseases. Moreover, MAPT haplotypes had significant differential impact on 23 transcripts from the 17q21.31 and 17q24.1 loci. MAPT haplotypes had also the largest up-regulating (256) and the largest down-regulating (−178) effect sizes measured as β values on two different transcripts from the same gene (LRRC37A2). Intronic SNP in the SNCA gene, rs3910105, differentially induced expression of three SNCA isoforms. In conclusion, this study established clear association between well-known haplotypic variance and transcript specific regulation in the blood. APOE e4 and MAPT H1/H2 haplotypic variants are associated with the expression of several genes related to the neurodegeneration.

Keywords: Apolipoprotein E (ApoE); microtubule associated protein (MAPT); synuclein alpha (SNCA); eQTL; TOMM40; KANSL1; mitochondria; Parkinson's Progression Markers Initiative (PPMI)



Citation: Koks, S.; Pfaff, A.L.; Bubb, V.J.; Quinn, J.P. Transcript Variants of Genes Involved in Neurodegeneration Are Differentially Regulated by the APOE and MAPT Haplotypes. *Genes* **2021**, *12*, 423. <https://doi.org/10.3390/genes12030423>

Academic Editor: Patrick O'Donoghue

Received: 7 January 2021
Accepted: 10 March 2021
Published: 15 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Apolipoprotein E (ApoE) epsilon 4 haplotype (e4) is the major risk factor for the Alzheimer disease (AD) [1]. Three main haplotype versions, e4, e3 and e2, exist for ApoE and these are defined by the single nucleotide polymorphisms (SNPs) rs429358 and rs7412. The effect of e4 is dosage-dependent on the risk of AD and at the same time e2 variant of ApoE exhibits protective effect for late onset AD [2]. This protective effect has been verified in the animal models overexpressing e4 or e3 variants of human ApoE genes [3]. The increased risk for AD with ApoE e4 is also evident in diverse populations and it is considered that all the genetic risk for AD in this region is accounted for by the ApoE e4 haplotype [4]. In addition, ApoE variants are not only involved in AD, but several studies also report an association between ApoE e4 and Parkinson's Disease (PD) [5]. The association between ApoE and PD is complex and seems to be population specific to certain PD subtypes or specific symptoms [5–7]. It is most commonly associated with cognitive decline in PD or Lewy body dementia [6,7]. A recent detailed analysis on three longitudinal cohorts

identified ApoE e4 to be associated with cognitive progression of PD [8]. Despite decades of research, the molecular reasons for the ApoE e4 association are not yet clear. Genetic variations in the TOMM40 gene have been described to modify the effects of ApoE variants due to linkage disequilibrium (LD) [9]. Indeed, several subsequent studies confirmed that ApoE and TOMM40 have a combined effect on the heritability of AD and analysis of both genes improves the precision to estimate the age of onset of AD [9–12]. While there is a strong genetic link between TOMM40 and ApoE the precise molecular interaction between ApoE and TOMM40 is not understood. The strong functional association could also be exhibited on their expression as the adjacent and close to one another on human chromosome 19 suggestive of overlapping transcriptional regulation.

Microtubule-associated protein tau (MAPT) is considered a significant risk factor for neurodegenerative diseases, including PD [13–19]. Tau protein is biochemically well-characterised and its role in aggregation and neurodegeneration was well established [18]. Recent genome wide association studies (GWAS) have repeatedly identified the MAPT locus as a significant independent risk locus for PD, confirming early genetic association studies indicating the link between MAPT and PD [20–22]. Genetic variation of the MAPT locus is described by the H1 and H2 haplotypes, that can be identified by different SNPs [23,24]. This is a large region in the genome with a LD spanning around 1.6 megabases without any recombination. The haplotype diversity is caused by the 970 kb inversion polymorphism, H1 has direct orientation and H2 has an inverted orientation [25,26]. H1 haplotypes are associated with AD, amyotrophic lateral sclerosis, PD and other neurodegenerative diseases [25]. H1 haplotype occurs in all populations with reasonable recombination and variability. H2 haplotype is of European origin and shows very limited variation in sequence [25,26]. This inversion contains many genes like MAPT, CRHR1, NSF, KANSL1, LRRC37A, LRRC37A2, ARL17A, ARL17B [27]. While the associations of MAPT haplotypes and neurodegenerative diseases are well established, the functional mechanism behind these associations is not clear.

The purpose of current study was to analyse the effects of the known variants at the ApoE, MAPT and SNCA loci (rs356181, rs3910105) on transcriptional regulation. We utilised data from the Parkinson's Progression Markers Initiative (PPMI) project where the genotypes of 570 subjects and their blood transcriptome data are available. We aimed to identify what transcripts or genes are differentially regulated by the genetic variants at these loci. As a result, we identified profound regulatory effect of ApoE haplotypes on TOMM40 expression and differential transcriptional regulation by the H1 or H2 haplotypes at the MAPT locus. In addition to the *cis* regulatory loci, we identified many *trans* locations that are also controlled by the variations in these haplotypes.

2. Materials and Methods

2.1. Datasets

In this study we utilized the Parkinson's Progression Markers Initiative (PPMI) cohort data that were downloaded from www.ppmi-info.org/data (accessed on 19 January 2021). The PPMI is a longitudinal cohort to follow Parkinson's patients and to describe the course of Parkinson's disease. The dataset contains whole transcriptome data from the blood together with genetic and clinical data. The clinical dataset already contained genotypes or haplotypes ApoE, Apo e4, MAPT, rs356181 and rs356181 loci. Genotyping data available for the PPMI cohort are based on the NeuroX SNP chip results. ApoE e2, e3 and e4 haplotypes were defined by the rs429358 CT and rs7412 CT genotypes. TT is for e2, TC is for e3 and CC is the e4 haplotype. MAPT haplotypes were defined by rs17652121, rs8070723, rs1052587, rs16940799 and rs17652748 genotypes where minor alleles for each position were counted and H1H1, H1H2 and H2H2 were detected. rs356181 and rs356181 are based on genotyping chip results.

Whole blood RNAseq data were downloaded from the PPMI website and transcript-based annotation was used for further analysis. This is the release of Phase 1 and Phase 2 PPMI RNA-seq data that were already processed, mapped to reference genome hg 19

and with the counts data generated. Briefly, FASTQ files were mapped to hg19 (GRCh37) by STAR on GENCODE v19, counts were created for genes and transcripts using FeatureCounts and abundance estimates (transcripts per Million, TPM) via Salmon. All this prerequisite work was done as a part of ongoing PPMI RNA-seq analysis and was conducted by HudsonAlpha Institute for Biotechnology, Institute of Translational Genomics of the University of Southern California and The Translational Genomics Research Institute, TGen. As the RNAseq data were annotated with the hg19 version of human genome all other annotations were also based on hg19. In this analysis all subjects, PD (314) and SWEDD (49) cases and controls (156), were combined and transcript expression signals were tabulated after importing Salmon files in the R to prepare them for the eQTL analysis. Altogether, 95,309 transcripts were used for the analysis in combination with the five genetic loci APOE, APOE_e4, MAPT, SNCA rs356181 and SNCA rs3910105.

2.2. eQTL Analysis

Matrix eQTL was used to calculate the genetic loci regulating the expression transcript variants [28]. We used additive linear model with covariates, age and sex, with FDR threshold 0.05. During eQTL analysis, local (*cis*) and distant (*trans*) quantitative loci were called, distant locus threshold was set on 1M bp. Raw results were used for *circos* plotting and plotting using R *ggbio* and *ggpubr* packages [29,30].

The correction for multiple testing of eQTL was performed using FDR and only results that remained significant after FDR correction are reported here. For pairwise comparisons between the genotype Wilcoxon test was used and *p*-values were challenged with the Bonferroni multiple comparison test.

3. Results

3.1. General Analysis for All Markers

Matrix eQTL analysis identified 151 eQTL loci, 29 in *cis* and 122 in *trans* position as statistically significant with the FDR below 0.05 (Supplementary Table S1). This ranged from largest effect on gene expression was β value 256 of MAPT haplotype on the LRRC37A2-201 transcript. The second largest was MAPT effect on KANSL1-002 transcript (β 153). The highest down-regulating effect on gene expression was again from MAPT haplotypes on the expression of LRRC37A2-001 (β value -178). MAPT haplotypes had strong negative effects on the KANSL1-006 transcript, LRRC37A2-201 transcript and on other transcripts in the MAPT locus. We identified transcript specific regulatory effect in all studied loci. Overall summary of the identified eQTLs is given in Table 1 and detailed information is in Supplementary Table S1. Supplementary Table S1 also contains the exact Ensembl codes for the transcript variants. We built a *circos* plot based on Supplementary Table S1 to show all the combined identified effects (Figure 1). More detailed *circus* plot is available as Supplementary Figure S1.

Table 1. The overview of the eQTL loci, number of transcripts and their effect sizes as β values for blood transcriptome. Explanation: Variants—variants under the study, N—number of statistically significant transcripts identified for particular variant, Min β —minimal β value, Max β —maximum β value, Mean β —mean β value, SD—standard deviation, SE—standard error of mean.

| Variants | N | Min β | Max β | Mean β | SD | SE |
|-----------|----|-------------|-------------|--------------|------|-----|
| ApoE | 18 | -18.7 | 0.7 | -1.9 | 4.9 | 1.2 |
| ApoE e4 | 58 | -6.7 | 15.9 | 1.0 | 3.3 | 0.4 |
| MAPT | 47 | -178.5 | 256.6 | 1.9 | 60.8 | 8.9 |
| rs356181 | 19 | -20.5 | 52.6 | 3.9 | 14.6 | 3.4 |
| rs3910105 | 9 | -21.2 | 5.3 | -4.1 | 9.7 | 3.2 |

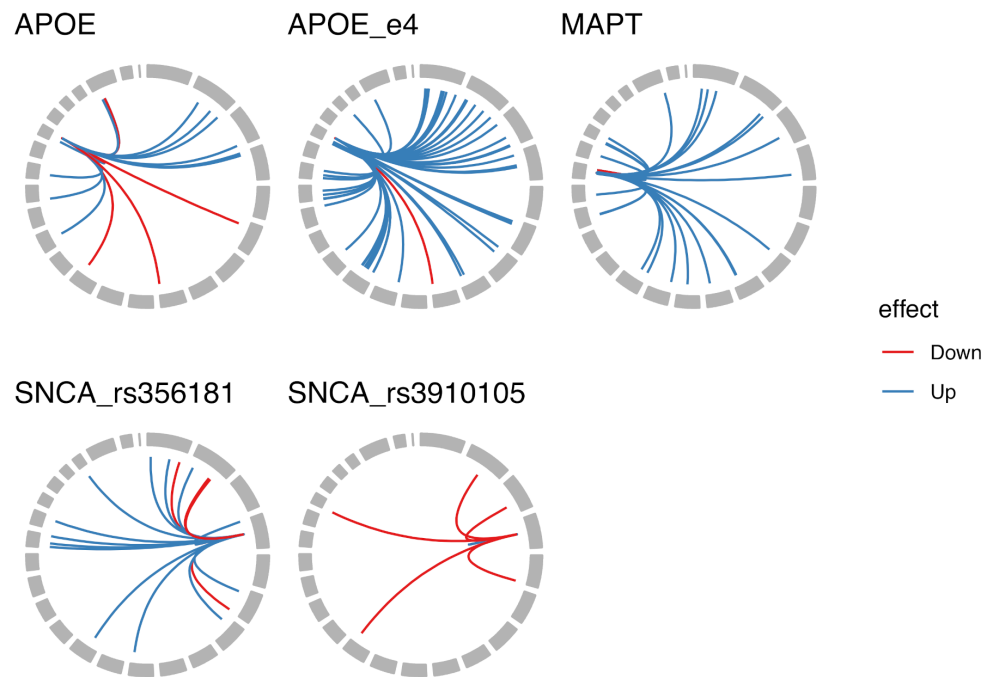


Figure 1. Combined Circos plot describing the regulatory effects of all studied loci. Each locus is shown in a separate panel with *cis* and *trans* activities combined into single circos plot. Down or up-regulating effects are shown with red and blue colour, respectively. Some loci had dominantly up-regulating effects, while the one locus had mainly down-regulating effects.

ApoE locus was analysed by using all three haplotypes and their combinations separately and only addressing the effect of e4 haplotype. ApoE e4 alone gave much larger number of eQTL targets than the separate e2, e3, e4 ApoE haplotypes combined demonstrating the overwhelming effect of ApoE e4 haplotype (Figure 1, “APOE” and “APOE_e4”). ApoE demonstrated more suppressing or downregulating effects than ApoE e4, however most of the identified eQTLs had stimulating effects on the target transcripts. Six different combinations of the ApoE haplotypes possessed strong and statistically significant effect on the expression of the TOMM40 gene (Figure 2). In Figure 2 we show the effect on the TOMM40 gene where the expression signals from different transcripts were combined. ApoE e3e3 had the strongest expression signal followed by e2e3 haplotype. The lowest expression was caused by e4e4 haplotype and any haplotype combination with at least single e4 genotype showed very low expression level of TOMM40 compared to any other haplotype combinations. We identified that ApoE e3 allele is profoundly stimulating for TOMM40 expression, whereas e4 is significantly suppressive. The e2 allele also seems to have repressive effects but not as significant as the e4 variant. Our analysis showed that the main effect of the ApoE haplotypes comes from the stimulating effect of the e3 allele on TOMM40 expression.

SNCA rs3910105 had a dominating repressive effect to its targets that is evident from Figure 1 and Table 1. This feature is quite unique compared to other variants used in the present study.

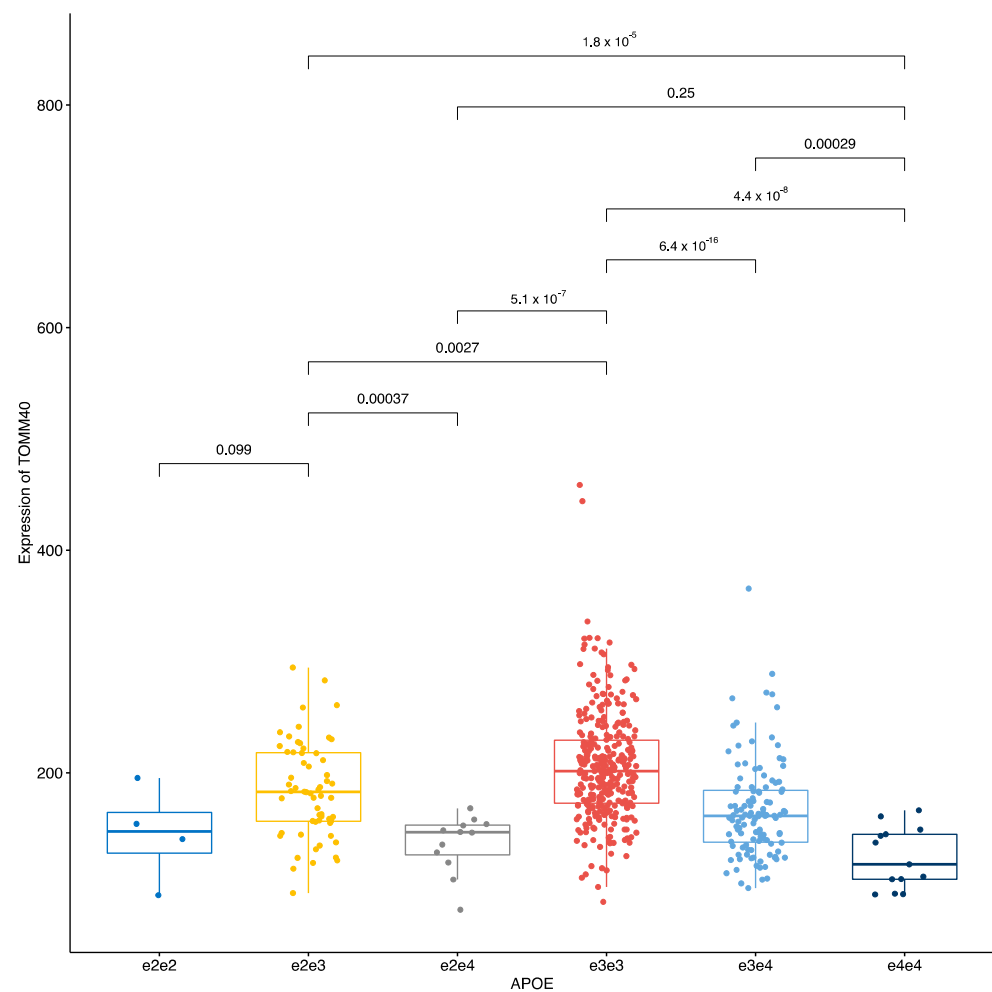


Figure 2. ApoE haplotypes differentially regulate the expression of the TOMM40 gene. The gene signal in this figure is aggregated from all RNA isoforms. ApoE e4 haplotype has statistically significantly lower expression than other haplotypes. Even single e4 allele was sufficient to significantly reduce expression of the TOMM40 gene. The bars in this figure and in all other figures show pairwise comparisons and respective *p*-values after the Wilcoxon test.

3.2. Analysis of the Individual Markers on Transcripts

3.2.1. All ApoE Haplotypes

As described in previous section, ApoE haplotypes possessed a strong stimulating activity on the TOMM40 gene and this effect was caused by the e3 variant. Haplotype e3e3 had the highest level of expression and any haplotype with single e3 (e2e3 and e3e4) were the next highest expressing variants. Here, we analyzed the effect of these haplotypes to the transcript level expression profile (Figure 3). Indeed, like with the gene level TOMM40 expression, e3 again showed the strongest stimulating effect on two TOMM40 transcripts, TOMM40-001 and TOMM40-005 (Figure 3). Transcript level analysis showed that while combination e3e4 had significantly downregulated expression of both isoforms, e2e3 had similar expression with the e3e3 haplotype. This is in contrast with the gene level analysis shown at the Figure 2 where e2e3 had a statistically significant reduction in expression compared to e3e3. Transcript based analysis shows that e2 variant may possess some transcript specific stimulating effect. ApoE e4 variant had the lowest expression level for transcripts.

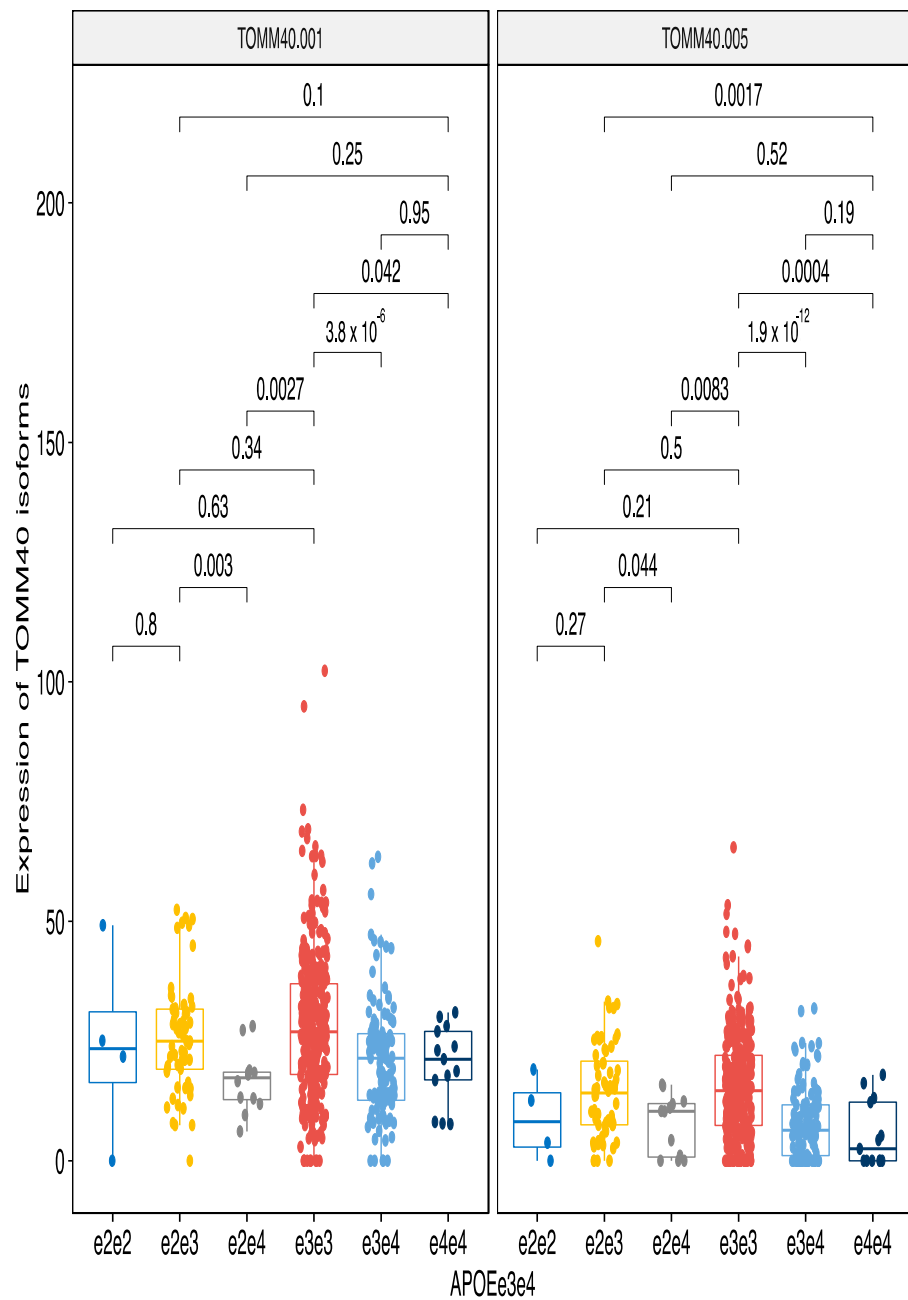


Figure 3. ApoE haplotypes regulate expression of the TOMM40 transcripts 001 and 005. ApoE e4 haplotype has statistically significantly lower expression than other haplotypes. Even a single e4 allele was sufficient to significantly reduce expression of the TOMM40 isoforms. The bars show pairwise comparisons and respective *p*-values after the Wilcoxon test.

3.2.2. ApoE e4 Effect

As ApoE e4 had a strong suppressive effect on combined TOMM40 gene expression (Figure 2), we next performed the analysis using individuals grouped by the number of e4 alleles they possessed. There were three combined haplotypes; those with no e4 alleles, with one e4 or two e4 alleles. As a result, we identified a robust suppressive effect of e4 on the expression of all TOMM40 transcripts that were significantly affected in the matrix eQTL analysis (Figure 4).

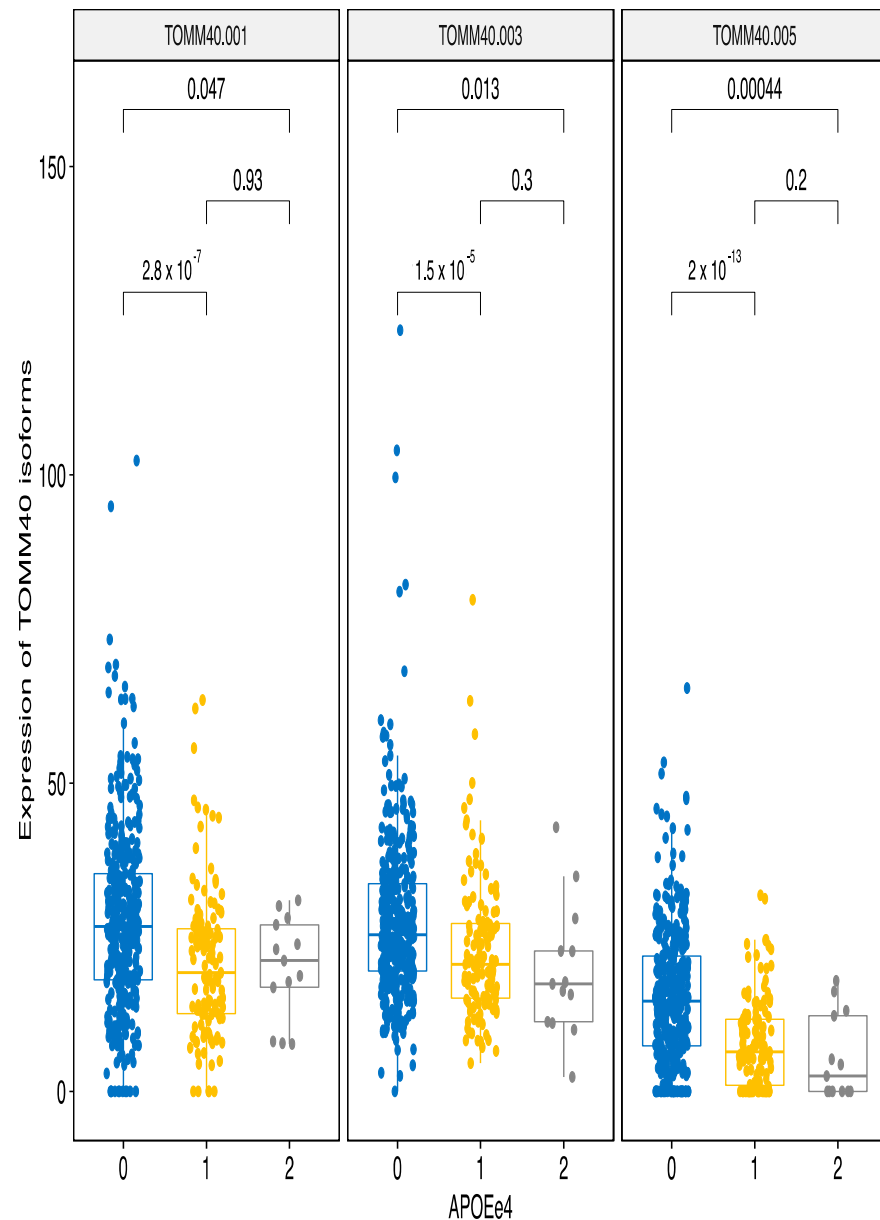


Figure 4. ApoE e4 haplotypes down-regulate expression of the TOMM40 transcripts 001, 003 and 005. Even a single e4 allele is sufficient to significantly reduce expression of the TOMM40 isoforms. The bars indicate pairwise comparisons and respective *p*-values after the Wilcoxon test.

Interestingly, ApoE e4 also had a significant repressive effect on the expression of sarcosine dehydrogenase transcript SARDH-003 (Table S1). The SARDH gene encodes an enzyme that localizes in the mitochondrial matrix and is responsible for oxidative demethylation of sarcosine. It is remarkable that ApoE e4 had very selective repressive effect only to four transcripts of two genes that all are involved in mitochondrial function. TOMM40 has seven transcript variants from which six are protein coding and SARDH gene has eight transcript variants with seven of them being protein coding. All other ApoE e4 effects were activating for their target genes. Therefore, ApoE e4 genetic effect might be caused by its involvement in regulating the transcripts of TOMM40 and SARDH genes.

3.2.3. MAPT Haplotypes

Genetic variation at the MAPT locus altogether affected 47 genomic regions with 18 *cis* and 29 *trans* effects. Transcripts of the genes LRRC37A2, KANSL1, ARL17B, LRRC37A,

ARHGAP27, PLEKHM1, MAPT, CRHR1, ARL17A were affected by the MAPT haplotype by *cis* regulation. All these genes were from the locus 17q21.31. Moreover, four transcripts of the gene LRRC37A3 in the 17q24.1 were also significantly affected (Supplementary Table S1). The β values of the model ranged from -178 to 256 indicating a robust regulatory effect on the genes. In many cases different transcripts of the same gene had opposite effects in a sense that some transcripts were up-regulated and the others were downregulated. For instance, the transcript CRHR1-202 was upregulated by the H2H2 haplotype, while the CRHR1-201 was downregulated by the H2H2. On the other hand, a single transcript of ARL17A was up-regulated by the H2H2 haplotype (β 2.3) and two transcripts of the ARL17B were both downregulated by H2H2 (β values -59 and -19).

The effects of the MAPT haplotypes on expression were the strongest of all the variants analyzed in this study. The largest upregulation by H2H2 haplotype was identified for the transcripts LRRC37A2-201, KANSL1-002 and LRRC37A-003 (Figure 5) with H2 acting in a dose—dependent manner.

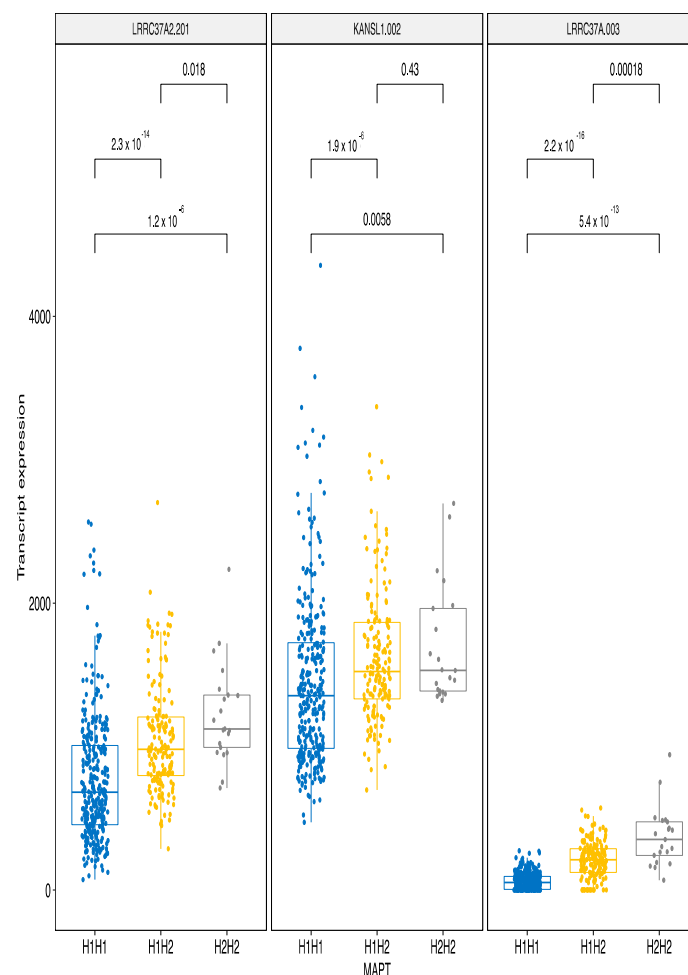


Figure 5. MAPT H2H2 haplotype upregulates several transcripts, three with the largest β values are illustrated here. Even a single H2 variant is sufficient to increase the expression of given transcripts. The bars indicate pairwise comparisons and respective p -values after the Wilcoxon test.

The most significant repressive effect of MAPT haplotypes was also evident for the locus 17q21.31. The most significant and the largest downregulation by H2H2 haplotype was identified for the transcripts LRRC37A2-001, KANSL1-006 and LRRC37A-002 (Figure 6). These are different transcripts but from the same genes that had the strongest up-regulating effect. This finding indicates MAPT haplotypes have a bidirectional effect that is transcript isoform specific indicating a molecular specificity of the H2H2 haplotype in regulating

the locus. The single transcript from *MAPT* gene itself was downregulated by the H2H2 haplotype. Two transcripts from another promising target for neurodegenerative disease, *PLEKHM1*, were both down-regulated by the H2H2 haplotype.

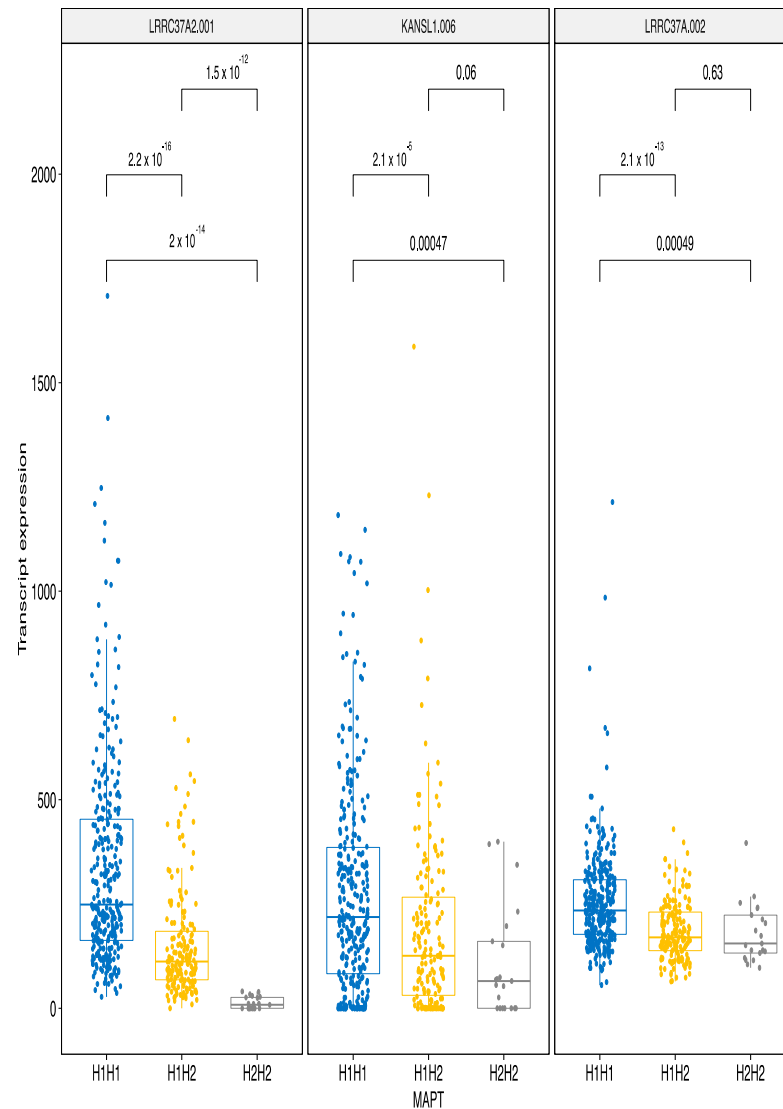


Figure 6. *MAPT* H2H2 haplotype downregulates several transcripts. *LRR37A2*, *KANSL1* and *LRR37A* transcripts had the highest negative β values (down-regulating effect) and are illustrated here. These are different transcripts from the same genes that are shown in Figure 5 and were up-regulated by the same haplotype. Even a single H2 variant is sufficient to downregulate the expression of given transcripts. The bars indicate pairwise comparisons and respective p -values after the Wilcoxon test.

3.2.4. *SNCA* Locus rs356181

The SNP rs356181 is an intronic variant in the gene *LOC105377329*, that is in close proximity to *SNCA* with potential association of its expression. We identified that variation at rs356181 modified expression of *SNCA* transcripts. *SNCA*-007 is downregulated by the CT and TT genotypes, whereas *SNCA*-008 is upregulated by CT and TT genotypes (Figure 7).

3.2.5. *SNCA* Locus rs3910105

SNP rs3910105 is an intronic variant in the *SNCA* gene. This SNP regulated nine different targets, four were in *cis* location. Three targets were different *SNCA* transcripts and the fourth gene was *MMRN1*, adjacent to *SNCA* (Figure 8). Our analysis indicated

that CC, CT and TT genotypes differentially regulated these transcripts. Both MMRN1 and SNCA-008 were downregulated by CT and TT genotypes. However, SNCA-007 and SNCA-009 were significantly upregulated by CT and TT variants in the rs3910105. In regard to the SNCA gene transcripts, the regulatory pattern is reverse compared to rs356181 that indicates specific regulatory pattern of these SNPs.

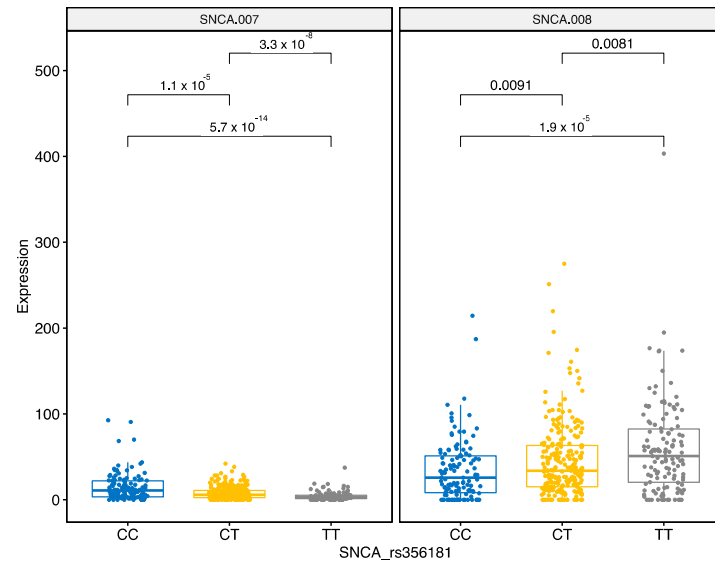


Figure 7. Different genotypes of the SNP rs356181 (CC, CT or TT) differentially regulate expression of two SNCA transcripts. The bars indicate pairwise comparisons between CC, CT and TT genotypes and respective *p*-values after the Wilcoxon test.

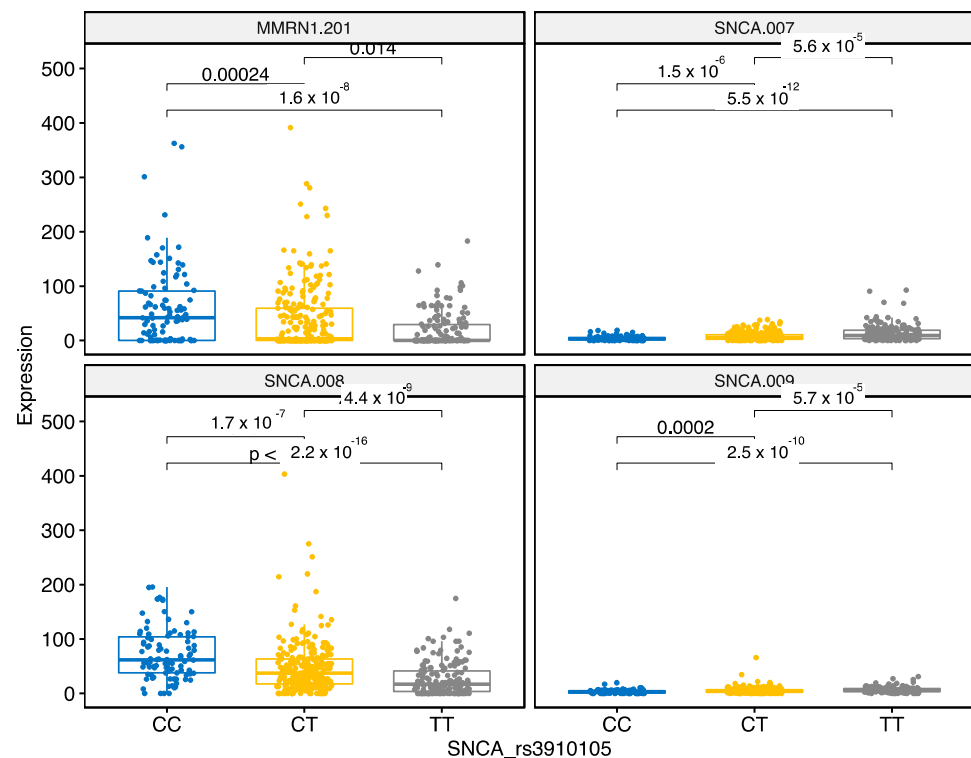


Figure 8. Different genotypes of the SNP rs3910105 (CC, CT or TT variants) differentially regulate expression of three SNCA transcripts and MMRN1. The bars indicate pairwise comparisons and respective *p*-values after the Wilcoxon test.

4. Discussion

In this study we describe the eQTL analysis performed using PPMI blood RNAseq data and genotypes for ApoE, MAPT, rs3910105 and rs356181. We identified that these genetic loci possess highly significant regulatory functions for at least 151 transcripts across the genome. Our results highlight new findings that may explain in part the functional consequences of these well-defined genetic variants in neurodegenerative diseases.

We identified that the TOMM40 gene expression (Figure 2) and transcript levels (Figures 3 and 4) are dependent on the ApoE haplotypes. Subjects with ApoE e3e3 haplotypes had the highest level of TOMM40 expression and even those with a single e3 allele had significantly increased TOMM40 expression in the blood. This is the first time that a connection between the ApoE haplotypes and TOMM40 gene and transcript expression has been established. The genetic link between ApoE and TOMM40 has been well established [9–11,31–33]. Most of these studies have looked for genetic interaction between ApoE and TOMM40 polyT polymorphisms to explain the differences in the AD sub-phenotypes and in the ages of the disease onset [10]. There are few studies where this genetic interaction is analysed in the context of transcriptional regulation and cis-regulatory effect caused by polyT polymorphism has been described [34]. In that sense, we are not the first to describe functional regulatory interaction between TOMM40 and ApoE, but our finding links ApoE haplotypes to the levels of the TOMM40 gene and transcripts. More precisely, we identified the e3 haplotype as a high-expression haplotype and e4 as low expression haplotype. We also identified that e4 haplotype regulates expression of another mitochondrial gene, SARDH. Our results connect ApoE haplotypes directly to the function of mitochondria and help to explain the neurodegeneration caused by the ApoE variants as the impact of mitochondrial dysfunction in neurodegeneration is well-established [35].

The model for mitochondrial mechanisms and changes in mitochondrial functions underpinning the ApoE effects was proposed several years ago [31]. Recent genetic studies have also indicated that ApoE haplotypes are also related to PD and other neurodegenerative conditions and are responsible for the cognitive decline in these patients [8]. The role of TOMM40 in mitochondrial dysfunction is further supported by existing literature where a reduction in protein expression levels has been found in neurodegenerative diseases [36]. All these findings are consistent with the genetic effect of ApoE haplotypes being mediated by its ability to regulate the expression of TOMM40 and to induce (or avoid) mitochondrial dysfunction. Protective e3e3 haplotype of ApoE induced significant expression of TOMM40 and its transcripts. This expression difference could be the mechanism behind the protective effect of ApoE e3e3.

We cannot exclude the possibility that the e3e3 effect is caused by the unbalanced distribution of the e3 allele as it is the most abundant version of the ApoE haplotypes. This is unlikely the case, as we also see the significant differences with heterozygous haplotypes e2e3 and e3e4 where the numbers of the subjects are larger. However, authors feel that extra experimental work is needed to get biochemical evidence for this association.

MAPT haplotypes H1H1, H1H2 and H2H2 are possibly the most studied haplotypes in neurodegenerative disease research. While this haplotype and locus is interesting as it hosts the MAPT gene that encodes tau protein, the exact molecular mechanism why the MAPT locus is associated with neurodegenerative disease is not understood [27]. Our report identified new potential candidates that are regulated by the variants at this locus. The haplotypic variation induced statistically significant changes in expression of many genes at this locus. Affected transcripts were from LRRC37A2, KANSL1, ARL17B, LRRC37A, ARHGAP27, PLEKHM1, MAPT, CRHR1 and ARL17A genes that reside in the locus 17q21.31 and also from the LRRC37A3 gene that is in the adjacent locus 17q24.1. Transcriptional differences were bidirectional in many cases. Some transcripts of LRRC37A2, KANSL1 and LRRC37A were upregulated whereas others were downregulated. This specificity indicates functional impact behind eQTL analysis. Additional support for real functional regulatory activity comes from the size of the effect that we measured with β values. The largest (up to 256) and the lowest β values (−178) were related to the regulation

of aforementioned genes and transcripts. This large variation in β values indicates distinct regulatory effect that H1 or H2 genotypes on specific targets. The genes that are affected by this haplotypic variation are all functional candidates for the pathogenetic model of neurodegeneration. KANSL1 and MAPT are primary candidates, but PLEKHM1 is often a neglected one. PLEKHM1 is a well characterized protein that is involved in the endosomal system and autophagy and is necessary for autophagosome-lysosome fusion [37]. Endosomal pathway and autophagy are both pathways that are severely affected during neurodegeneration making PLEKHM1 an interesting candidate for further studies.

Our analysis also involved two SNPs from the locus of SNCA and we identified several targets regulated by these variants. Both SNPs affected the differential expression of SNCA transcripts and rs3910105 also affected the MMRN1 gene that is adjacent to the SNCA gene. MMRN1 is elastin microfibril interface 4 and is able to carry platelet factor V. A recent study identified that MMRN1 expression was increased in SNCA knockout mice that was attributed to the genetic modification in the SNCA locus [38]. However, our study identified regulatory link between the intronic SNP in SNCA gene and MMRN1 gene expression. Additional experimental studies are needed to identify the nature of this interaction.

Strengths of our study include the large study cohort consisting of data from 570 individuals and the DNA and RNA were from the same tissue, blood. As we had both RNA and DNA from the same persons and from identical tissues our study had minimal confounding factors. During the linear modelling we adjusted for the age and sex of the participants. Using the PPMI dataset has additional advantage as the research community can access cell lines from some of these individuals for functional verification experiments.

5. Conclusions

In conclusion, we identified new target genes that are regulated by the common variation relevant to the risk of neurodegenerative diseases. ApoE haplotypes significantly regulate transcription of the TOMM40 gene and the e3 variant strongly upregulates expression of TOMM40, whereas e4 downregulates its expression. This regulatory connection with TOMM40 can explain the genetic effects of ApoE that has been described decades ago.

MAPT H2H2 haplotype is associated with profound differential effect on the transcripts of the genes that have functional impact on neurodegeneration. This differential expressional control over KANSL1, LRRC37A2 and LRRC37A transcripts has very large biological effect and therefore may explain the molecular mechanism for the MAPT haplotypes in neurodegenerative processes.

Supplementary Materials: The following are available online at <https://www.mdpi.com/2073-4425/12/3/423/s1>, Figure S1: Detailed circos plots of all analyzed loci; Table S1: All Identified eQTL loci in our analysis. Legend to Figure S1: The figure is detailed circos plot of all analysed genetic variants and their target transcripts. This figure is based on Table S1 and contains all the Table S1 data. Each locus is on separate circos plot. Plots show the ideogram with chromosome numbers, the position of the analysed genetic variants as a blue bar between ideogram and dot plot. Dot plot shows β values of the linear modelling to describe the effect size that connect genetic variant with the transcript. Dots are sized according to their β effect size and coloured according to the *cis* (green) or *trans* effects (magenta). Connecting lines link the genetic variants with its target transcripts and colour indicates up-regulation (blue) or downregulation (red) of the transcript. Legend for Table S1: snps—single nucleotide polymorphism, Ensembl TX—Transcript ID in the Ensembl database; txabbr—Transcript abbreviation in the Ensembl database; statistic—*t*-Test Statistic; *p* value—*p*-value from the ANOVA testing; FDR—False Discovery Rate; β —value of β coefficient, standardised regression coefficient; ensembl_gene_id—Gene ID in the Ensembl database; description—gene name; hgncgene—HUGO Gene Nomenclature Committee gene symbol; chromband—Chromosome number and band; position—position of the effect, *cis* or *trans*.

Author Contributions: Conceptualization, S.K.; methodology, A.L.P. and S.K.; formal analysis, S.K.; data interpretation, A.L.P., V.J.B., J.P.Q. and S.K.; writing—original draft preparation, S.K.; writing—

review and editing, A.L.P., V.J.B., J.P.Q. and S.K.; funding acquisition, A.L.P. and S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by MSWA, The Michael J. Fox Foundation, Shake It Up Australia and The Perron Institute.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Human Ethics Research Office of The University of Western Australia (protocol code RA/4/20/5308 approved on 05.08.2019)

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study by PPMI.

Data Availability Statement: Raw data are available from the PPMI website (www.ppmi-info.org/data, accessed on 19 January 2021).

Acknowledgments: This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data, accessed on 19 January 2021). For up-to-date information on the study, visit www.ppmi-info.org, accessed on 19 January 2021. PPMI is sponsored and partially funded by The Michael J. Fox Foundation for Parkinson's Research. PPMI—a public-private partnership—is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including Abbvie, Allergan, Amathus therapeutics, Avid Radiopharmaceuticals, Biogen Idec, Biogen, Bristol-Myers Squibb, Celgene, Denali, GE Healthcare, Genentech, GlaxoSmithKline, Janssen neuroscience, Lilly, Lundbeck, Merck, Meso Scale Discovery, Pfizer, Piramal, Prevail Therapeutics, Roche, Sanofi Genzyme, Servier, Takeda, Teva, UCB, Verily and Voyager Therapeutics.

Conflicts of Interest: The authors declare no conflict of interest related to this study. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Corder, E.H.; Saunders, A.M.; Strittmatter, W.J.; Schmechel, D.E.; Gaskell, P.C.; Small, G.W.; Roses, A.D.; Haines, J.L.; Pericak-Vance, M.A. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* **1993**, *261*, 921–923. [[CrossRef](#)] [[PubMed](#)]
2. Corder, E.H.; Saunders, A.M.; Risch, N.J.; Strittmatter, W.J.; Schmechel, D.E.; Gaskell, P.C., Jr.; Rimmler, J.B.; Locke, P.A.; Conneally, P.M.; Schmechel, K.E.; et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat. Genet.* **1994**, *7*, 180–184. [[CrossRef](#)] [[PubMed](#)]
3. Buttini, M.; Orth, M.; Bellosta, S.; Akeefe, H.; Pitas, R.E.; Wyss-Coray, T.; Mucke, L.; Mahley, R.W. Expression of human apolipoprotein E3 or E4 in the brains of Apoe^{-/-} mice: Isoform-specific effects on neurodegeneration. *J. Neurosci.* **1999**, *19*, 4867–4880. [[CrossRef](#)]
4. Jun, G.; Vardarajan, B.N.; Buross, J.; Yu, C.E.; Hawk, M.V.; Dombroski, B.A.; Crane, P.K.; Larson, E.B.; Mayeux, R.; Haines, J.L.; et al. Comprehensive search for Alzheimer disease susceptibility loci in the APOE region. *Arch. Neurol.* **2012**, *69*, 1270–1279. [[CrossRef](#)] [[PubMed](#)]
5. Sun, R.; Yang, S.; Zheng, B.; Liu, J.; Ma, X. Apolipoprotein E Polymorphisms and Parkinson Disease with or without Dementia: A Meta-Analysis Including 6453 Participants. *J. Geriatr. Psychiatry Neurol.* **2019**, *32*, 3–15. [[CrossRef](#)]
6. Iwaki, H.; Blauwendraat, C.; Leonard, H.L.; Kim, J.J.; Liu, G.; Maple-Grødem, J.; Corvol, J.C.; Pihlstrøm, L.; van Nimwegen, M.; Hutten, S.J.; et al. Genomewide association study of Parkinson's disease clinical biomarkers in 12 longitudinal patients' cohorts. *Mov. Disord.* **2019**, *34*, 1839–1850. [[CrossRef](#)]
7. Sabir, M.S.; Blauwendraat, C.; Ahmed, S.; Serrano, G.E.; Beach, T.G.; Perkins, M.; Rice, A.C.; Masliah, E.; Morris, C.M.; Pihlstrom, L.; et al. Assessment of APOE in atypical parkinsonism syndromes. *Neurobiol. Dis.* **2019**, *127*, 142–146. [[CrossRef](#)]
8. Tan, M.M.; Lawton, M.A.; Jabbari, E.; Reynolds, R.H.; Iwaki, H.; Blauwendraat, C.; Kanavou, S.; Pollard, M.I.; Hubbard, L.; Malek, N.; et al. Genome-Wide Association Studies of Cognitive and Motor Progression in Parkinson's Disease. *Mov. Disord.* **2020**. [[CrossRef](#)] [[PubMed](#)]
9. Roses, A.D.; Lutz, M.W.; Amrine-Madsen, H.; Saunders, A.M.; Crenshaw, D.G.; Sundseth, S.S.; Huentelman, M.J.; Welsh-Bohmer, K.A.; Reiman, E.M. A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease. *Pharm. J.* **2010**, *10*, 375–384. [[CrossRef](#)]
10. Chiba-Falek, O.; Gottschalk, W.K.; Lutz, M.W. The effects of the TOMM40 poly-T alleles on Alzheimer's disease phenotypes. *Alzheimers Dement.* **2018**, *14*, 692–698. [[CrossRef](#)]

11. Roses, A.D.; Lutz, M.W.; Crenshaw, D.G.; Grossman, I.; Saunders, A.M.; Gottschalk, W.K. TOMM40 and APOE: Requirements for replication studies of association with age of disease onset and enrichment of a clinical trial. *Alzheimers Dement.* **2013**, *9*, 132–136. [[CrossRef](#)]
12. Lutz, M.W.; Crenshaw, D.; Welsh-Bohmer, K.A.; Burns, D.K.; Roses, A.D. New Genetic Approaches to AD: Lessons from APOE-TOMM40 Phylogenetics. *Curr. Neurol. Neurosci. Rep.* **2016**, *16*, 48. [[CrossRef](#)]
13. Deutschlander, A.B.; Konno, T.; Soto-Beasley, A.I.; Walton, R.L.; van Gerpen, J.A.; Uitti, R.J.; Heckman, M.G.; Wszolek, Z.K.; Ross, O.A. Association of MAPT subhaplotypes with clinical and demographic features in Parkinson's disease. *Ann. Clin. Transl. Neurol.* **2020**, *7*, 1557–1563. [[CrossRef](#)]
14. Mamah, C.E.; Lesnick, T.G.; Lincoln, S.J.; Strain, K.J.; de Andrade, M.; Bower, J.H.; Ahlskog, J.E.; Rocca, W.A.; Farrer, M.J.; Maraganore, D.M. Interaction of alpha-synuclein and tau genotypes in Parkinson's disease. *Ann. Neurol.* **2005**, *57*, 439–443. [[CrossRef](#)]
15. Kwok, J.B.; Teber, E.T.; Loy, C.; Hallupp, M.; Nicholson, G.; Mellick, G.D.; Buchanan, D.D.; Silburn, P.A.; Schofield, P.R. Tau haplotypes regulate transcription and are associated with Parkinson's disease. *Ann. Neurol.* **2004**, *55*, 329–334. [[CrossRef](#)] [[PubMed](#)]
16. Skipper, L.; Wilkes, K.; Toft, M.; Baker, M.; Lincoln, S.; Hulihan, M.; Ross, O.A.; Hutton, M.; Aasly, J.; Farrer, M. Linkage disequilibrium and association of MAPT H1 in Parkinson disease. *Am. J. Hum. Genet.* **2004**, *75*, 669–677. [[CrossRef](#)]
17. Nakano, I. Frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP 17). *Ryoikibetsu Shokogun Shirizu* **1999**, *27 Pt 2*, 20–23.
18. Buée, L.; Delacourte, A. Comparative biochemistry of tau in progressive supranuclear palsy, corticobasal degeneration, FTDP-17 and Pick's disease. *Brain Pathol.* **1999**, *9*, 681–693. [[CrossRef](#)] [[PubMed](#)]
19. Dumanchin, C.; Camuzat, A.; Campion, D.; Verpillat, P.; Hannequin, D.; Dubois, B.; Saugier-Verber, P.; Martin, C.; Penet, C.; Charbonnier, F.; et al. Segregation of a missense mutation in the microtubule-associated protein tau gene with familial frontotemporal dementia and parkinsonism. *Hum. Mol. Genet.* **1998**, *7*, 1825–1829. [[CrossRef](#)]
20. Bandres-Ciga, S.; Ahmed, S.; Sabir, M.S.; Blauwendraat, C.; Adarmes-Gómez, A.D.; Bernal-Bernal, I.; Bonilla-Toribio, M.; Buiza-Rueda, D.; Carrillo, F.; Carrión-Claro, M.; et al. The Genetic Architecture of Parkinson Disease in Spain: Characterizing Population-Specific Risk, Differential Haplotype Structures, and Providing Etiologic Insight. *Mov. Disord.* **2019**, *34*, 1851–1863. [[CrossRef](#)] [[PubMed](#)]
21. Vandrovcova, J.; Pittman, A.M.; Malzer, E.; Abou-Sleiman, P.M.; Lees, A.J.; Wood, N.W.; de Silva, R. Association of MAPT haplotype-tagging SNPs with sporadic Parkinson's disease. *Neurobiol. Aging* **2009**, *30*, 1477–1482. [[CrossRef](#)]
22. Nalls, M.A.; Blauwendraat, C.; Vallerga, C.L.; Heilbron, K.; Bandres-Ciga, S.; Chang, D.; Tan, M.; Kia, D.A.; Noyce, A.J.; Xue, 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.* **2019**, *18*, 1091–1102. [[CrossRef](#)]
23. Pittman, A.M.; Myers, A.J.; Abou-Sleiman, P.; Fung, H.C.; Kaleem, M.; Marlowe, L.; Duckworth, J.; Leung, D.; Williams, D.; Kilford, L.; et al. Linkage disequilibrium fine mapping and haplotype association analysis of the tau gene in progressive supranuclear palsy and corticobasal degeneration. *J. Med. Genet.* **2005**, *42*, 837–846. [[CrossRef](#)] [[PubMed](#)]
24. Pittman, A.M.; Myers, A.J.; Duckworth, J.; Bryden, L.; Hanson, M.; Abou-Sleiman, P.; Wood, N.W.; Hardy, J.; Lees, A.; de Silva, R. The structure of the tau haplotype in controls and in progressive supranuclear palsy. *Hum. Mol. Genet.* **2004**, *13*, 1267–1274. [[CrossRef](#)]
25. Zody, M.C.; Jiang, Z.; Fung, H.C.; Antonacci, F.; Hillier, L.W.; Cardone, M.F.; Graves, T.A.; Kidd, J.M.; Cheng, Z.; Abouelleil, A.; et al. Evolutionary toggling of the MAPT 17q21.31 inversion region. *Nat. Genet.* **2008**, *40*, 1076–1083. [[CrossRef](#)]
26. Stefansson, H.; Helgason, A.; Thorleifsson, G.; Steinthorsdottir, V.; Masson, G.; Barnard, J.; Baker, A.; Jonasdottir, A.; Ingason, A.; Gudnadottir, V.G.; et al. A common inversion under selection in Europeans. *Nat. Genet.* **2005**, *37*, 129–137. [[CrossRef](#)] [[PubMed](#)]
27. Soto-Beasley, A.I.; Walton, R.L.; Valentino, R.R.; Hook, P.W.; Labbé, C.; Heckman, M.G.; Johnson, P.W.; Goff, L.A.; Uitti, R.J.; McLean, P.J.; et al. Screening non-MAPT genes of the Chr17q21 H1 haplotype in Parkinson's disease. *Parkinsonism Relat. Disord.* **2020**, *78*, 138–144. [[CrossRef](#)]
28. Shabalin, A.A. Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **2012**, *28*, 1353–1358. [[CrossRef](#)] [[PubMed](#)]
29. Yin, T.; Cook, D.; Lawrence, M. ggbio: An R package for extending the grammar of graphics for genomic data. *Genome Biol.* **2012**, *13*, R77. [[CrossRef](#)] [[PubMed](#)]
30. Kassambara, A. ggpubr: 'ggplot2' Based Publication Ready Plots; STHDA, Marseille, France 2020. Available online: <https://rpkgs.datanovia.com/ggpubr/> (accessed on 18 January 2020).
31. Roses, A.; Sundseth, S.; Saunders, A.; Gottschalk, W.; Burns, D.; Lutz, M. Understanding the genetics of APOE and TOMM40 and role of mitochondrial structure and function in clinical pharmacology of Alzheimer's disease. *Alzheimers Dement.* **2016**, *12*, 687–694.
32. Chiba-Falek, O.; Linnertz, C.; Guyton, J.; Gardner, S.D.; Roses, A.D.; McCarthy, J.J.; Patel, K. Pleiotropy and allelic heterogeneity in the TOMM40-APOE genomic region related to clinical and metabolic features of hepatitis C infection. *Hum. Genet.* **2012**, *131*, 1911–1920. [[CrossRef](#)]
33. Roses, A.D. An inherited variable poly-T repeat genotype in TOMM40 in Alzheimer disease. *Arch. Neurol.* **2010**, *67*, 536–541. [[CrossRef](#)]

34. Linnertz, C.; Anderson, L.; Gottschalk, W.; Crenshaw, D.; Lutz, M.W.; Allen, J.; Saith, S.; Mihovilovic, M.; Burke, J.R.; Welsh-Bohmer, K.A.; et al. The cis-regulatory effect of an Alzheimer's disease-associated poly-T locus on expression of TOMM40 and apolipoprotein E genes. *Alzheimers Dement.* **2014**, *10*, 541–551. [[CrossRef](#)] [[PubMed](#)]
35. Billingsley, K.J.; Barbosa, I.A.; Bandrés-Ciga, S.; Quinn, J.P.; Bubb, V.J.; Deshpande, C.; Botia, J.A.; Reynolds, R.H.; Zhang, D.; Simpson, M.A.; et al. Mitochondria function associated genes contribute to Parkinson's Disease risk and later age at onset. *NPJ Parkinsons Dis.* **2019**, *5*, 8. [[CrossRef](#)] [[PubMed](#)]
36. Heinemeyer, T.; Stemmet, M.; Bardien, S.; Neethling, A. Underappreciated Roles of the Translocase of the Outer and Inner Mitochondrial Membrane Protein Complexes in Human Disease. *DNA Cell Biol.* **2019**, *38*, 23–40. [[CrossRef](#)] [[PubMed](#)]
37. McEwan, D.G.; Dikic, I. PLEKHM1: Adapting to life at the lysosome. *Autophagy* **2015**, *11*, 720–722. [[CrossRef](#)]
38. Chaprov, K.D.; Goloborshcheva, V.V.; Tarasova, T.V.; Teterina, E.V.; Korokin, M.V.; Soldatov, V.O.; Pokrovskiy, M.V.; Kucheryanu, V.G.; Morozov, S.G.; Ovchinnikov, R.K. Increased Expression of the Multimerin-1 Gene in α -Synuclein Knockout Mice. *Dokl. Biol. Sci.* **2020**, *494*, 260–263. [[CrossRef](#)] [[PubMed](#)]