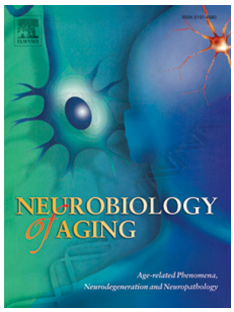


Failure to detect synergy between variants in transferrin and hemochromatosis and Alzheimer's disease in large cohort

Elizabeth Vance, Josue D. Gonzalez Murcia, Justin B. Miller, Alzheimer's Disease Genetic Consortium (ADGC), Lyndsay Staley, Paul K. Crane, Shubhabrata Mukherjee, John S.K. Kauwe

PII: S0197-4580(20)30027-0

DOI: https://doi.org/10.1016/j.neurobiolaging.2020.01.013

Reference: NBA 10764

To appear in: *Neurobiology of Aging*

Received Date: 9 April 2019

Revised Date: 13 November 2019

Accepted Date: 28 January 2020

Please cite this article as: Vance, E., Gonzalez Murcia, J.D., Miller, J.B., Alzheimer's Disease Genetic Consortium (ADGC), Staley, L., Crane, P.K., Mukherjee, S., Kauwe, J.S.K., Failure to detect synergy between variants in transferrin and hemochromatosis and Alzheimer's disease in large cohort, *Neurobiology of Aging* (2020), doi: https://doi.org/10.1016/j.neurobiolaging.2020.01.013.

Failure to detect synergy between variants in transferrin and hemochromatosis and Alzheimer's disease in large cohort

Elizabeth Vance[1], Josue D. Gonzalez Murcia[1], Justin B. Miller[1], Alzheimer's Disease Genetic Consortium (ADGC), Lyndsay Staley[1], Paul K. Crane[2], Shubhabrata Mukherjee[2], John S.K. Kauwe[1]
[1]Department of Biology, Brigham Young University
[2] Department of Medicine, University of Washington

Corresponding Author:
John S.K. Kauwe
kauwe@byu.edu
Department of Biology, Brigham Young University

Abstract

Alzheimer's disease (AD) is the most common cause of dementia and, despite decades of effort, there is no effective treatment. In the last decade, many association studies have identified genetic markers that are associated with AD status. Two of these studies suggest that an epistatic interaction between variants *rs1049296* in the Transferrin (*TF*) gene and *rs1800562* in the Homeostatic Iron Regulator (*HFE)* gene, commonly known as hemochromatosis, is in genetic association with AD. *TF* and *HFE* are involved in the transport and regulation of iron in the brain, and disrupting these processes exacerbates AD pathology through increased neurodegeneration and oxidative stress. However, by using a significantly larger dataset from the Alzheimer's Disease Genetic Consortium (ADGC), we fail to detect an association between *TF rs1049296* or *HFE rs1800562* with AD risk (*TF rs1049296* p=0.38 and *HFE rs1800562* p=0.40). In addition, logistic regression with an interaction term and a Synergy Factor Analysis (SFA) both failed to detect epistasis between *TF rs1049296* and *HFE rs1800562* (SF=0.94; p=0.48) in AD cases. Each of these analyses had sufficient statistical power (Power>0.99), suggesting that previously-reported associations may be the result of more complex epistatic interactions, genetic heterogeneity, or were false-positive associations due to limited sample sizes.

Keywords: Alzheimer's disease; transferrin; hemochromatosis; homeostatic iron regulator; synergy; epistasis

Introduction

Alzheimer's disease (AD) is the most common cause of dementia and inflicts an estimated 24 to 35 million people worldwide, with incidences predicted to increase dramatically as the population ages ("2018 Alzheimer's disease facts and figures," 2018). Although decades of research have been spent investigating the causes and architecture of this neurodegenerative disease, it still inflicts an estimated 5.7 million people in the United States alone. This number is projected to increase to 13.8 million by mid-century ("2018 Alzheimer's disease facts and figures," 2018). Association studies have accurately identified single-nucleotide polymorphisms (SNPs) associated with AD (D. Harold et al., 2009; Denise Harold et al., 2009; Hollingworth et al., 2011; J.-C. Lambert et al., 2009; J. C. Lambert et al., 2013; Seshadri et al., 2010; Shen et al., 2015; Shuai et al., 2015; Yan et al., 2015). However, these genetic loci account for only a fraction of AD heritability, (Ridge, Mukherjee, Crane, Kauwe, & Alzheimer's Disease Genetics, 2013) suggesting that much of AD's unexplained genetic make-up may be due to epistasis (Bullock et al., 2013; Combarros, Cortina-Borja, Smith, & Lehmann, 2009; M. T. Ebbert et al., 2014; Infante et al., 2004). Epistasis occurs when multiple genes interact to create a single phenotype (Cordell, 2002). These kinds of synergetic relationships play a critical role in the etiology of complex diseases, yet remain vastly understudied in AD pathology ("2018 Alzheimer's disease facts and figures," 2018; M. T. W. Ebbert, Ridge, & Kauwe, 2015; Raghavan & Tosto, 2017).

The Transferrin (*TF*) gene and the Homeostatic Iron Regulator (*HFE)* gene, commonly known as hemochromatosis, have been reported to show epistasis and play a role in the development of AD (Robson et al., 2004; Tisato et al., 2018). TFs are a group of non-heme iron-binding glycoproteins found in fluids and cells of vertebrates. The main role of *TF* is to maintain iron homeostasis in the body (Gkouvatsos, Papanikolaou, & Pantopoulos, 2012). In the brain, *TF* interacts with the Amyloid Precursor Protein (APP) (Belaidi et al., 2018) and tau (Jahshan, Esteves-Villanueva, & Martic-Milne, 2016), two of the major protein families implicated in AD pathology. Since iron is essential for oxygen transport, its mis-regulation in the brain can lead to oxidative stress and neurodegeneration (Dias, Junn, & Mouradian, 2013; Matak et al., 2016; Yarjanli, Ghaedi, Esmaeili, Rahgozar, & Zarrabi, 2017). *HFE* encodes for a transmembrane glycoprotein that binds to a *TF* receptor, subsequently regulating iron in the cell (Bennett, Lebron, & Bjorkman, 2000; Feder et al., 1996; Lebron et al., 1998). Mutations in *HFE* are associated with neurodegenerative diseases through increasing neuroinflammation and production of free radicals in the brain (Andersen, Johnsen, & Moos, 2014; Lull & Block, 2010). In addition, other studies suggest that *TF* and *HFE* are involved in the transport and regulation of iron in the brain, and disrupting these processes potentially affects AD pathology through increased neurodegeneration and oxidative stress (Ali-Rahmani, Schengrund, & Connor, 2014; Lehmann et al., 2006).

Robson et al. (2004) suggested that epistasis between *TF* variant *rs1049296* and *HFE* variant *rs1800562* is associated with AD. Although neither SNP alone was a risk factor for AD, the presence of both alleles resulted in a five times greater risk of developing AD. (Robson et al., 2004). Since the sample size for that study was relatively small (191 cases and 269 controls), a replication of these findings on a slightly larger dataset (1,161 cases and 1,342 controls) was conducted. A logistic regression analysis and a Synergy Factor Analysis (SFA) corroborated a significant association with AD risk among bi-allelic carriers of *rs1049296* and *rs1800562* (synergy factor= 2.71; p=0.0016) (Kauwe et al., 2010).

Our study expands on these previous studies and attempts to detect statistical epistasis between *TF rs1049296* and *HFE rs1800562* with respect to AD risk using 25,666 individuals (12,532 cases and

13,134 controls) from the Alzheimer's Disease Genetic Consortium (ADGC), which is an expansion of the dataset employed by Kauwe et al. (2010).

Material and Methods

Dataset and Filtering

Our analysis started with GWAS data from all 28,730 individuals in the Alzheimer's Disease Genetic Consortium (ADGC) dataset as described by Naj et al. (A. C. Naj et al., 2011). ADGC is a collection of 30 merged datasets spanning 1984 to 2012, and was established to help identify genetic markers of late onset AD. (Boehme, Mukherjee, Crane, & Kauwe, September 2014) (see Supplementary Table 1 for ADGC demographics). ADGC imputed the 30 datasets to the Haplotype Reference Consortium (HRC) reference panel, which includes 64,976 haplotypes and 39,235,157 SNPs (Loh et al., 2016; Adam C. Naj et al., 2017). Genotyped markers with a minor allele frequency less than 1% and a deviation from Hardy Weinberg Equilibrium (HWE) where $\alpha<10^{-6}$ were removed. All aspects of the study were approved by institutional review boards, and each applicant signed a written form of consent for their genetic data to be used for research purposes.

We followed the same filtering protocols established by Ridge et al. (Ridge et al., 2013) by genotyping markers with a minor allele frequency less than 1% and removing markers with a HWE p-value less than 10^{-6}. Principle components were calculated using Eigensoft (Patterson, Price, & Reich, 2006; Price et al., 2006) to account for population specific variations in allele distribution. After filtering, 12,532 cases and 13,134 control subjects contained genotypic information for *TF rs1049296* and *HFE rs1800562*.

Genetic Analyses

The main effects of *TF rs1049296* and *HFE rs1800562* on AD risk were measured using a multivariate nonparametric logistic regression analysis. Each SNP was first analyzed as a single term and then as an interaction term in a subsequent analysis. Similar to the Kauwe et al. (2010) study, we used the annotations in the ADGC dataset to include sex, age of onset, *APOE e4* allele status, cohort, and 10 principle components as covariates. In addition, we performed a chi-square analysis to determine odds ratios between AD status in each SNP as a single term and as an interaction term, respectively. Lastly, we performed a Synergy Factor Analysis (SFA) to calculate the size and significance of the interaction between *TF rs1049296* and *HFE rs1800562* and AD risk with minor allele non-carriers as the reference group (Cortina-Borja, Smith, Combarros, & Lehmann, 2009) (see Supplementary Table 3 for detailed SFA calculations). These analyses were performed for each of the 30 cohorts separately and for the entire ADGC dataset combined as a single cohort.

Furthermore, we calculated the power of analysis for the ADGC dataset using an online power tool available at https://www.dartmouth.edu/~eugened/power-samplesize.php (Demidenko, 2008; Demidenko, 2007). The previous analysis performed by Kauwe et al. (2010) had 0.31 power to detect an effect size of 1.14 at an alpha of 0.05 by using a sample size of 2,503. Our logistic regression model has power of >0.99 to detect a similar effect size of 1.14 at an alpha of 0.05 by using a sample size of 25,666 (see Supplemental Figure 1).

Results

The nonparametric logistic regression analysis using ADGC as one cohort demonstrated that when testing the main effects, neither *TF rs1049296* nor *HFE rs1800562* was associated with AD risk (*TF rs1049296* p=0.38; *HFE rs1800562* p=0.40). The logistic regression analyses including an interaction term for the two variants also failed to show significant association (p=0.23). Similarly,

the SFA analysis did not find epistasis between *TF rs1049296* and *HFE rs1800562* (SF=0.94; p=0.48).

We performed logistic regression on all 30 individual cohorts (see Supplemental Figure 2). We detected a significant epistatic association between the interaction term and AD status in the ACT cohort (p=0.038) and a suggested association in the ADC1 cohort (p=0.063). In addition, the individual effect of *HFE rs1800562* shows a suggested association with AD status in the ADC6 (p=0.099), WHICAP (p=0.052), ADC4 (p=0.076), and ROSMAP (p=0.094) cohorts. Furthermore, logistic regression for the individual effect of *TF rs1049296* determined a significant association with AD status in the WASHU cohort (p=0.016). However, none of these associations remained significant after a Bonferroni correction for multiple tests.

In addition, chi-squared analyses between terms and AD status demonstrated a non-significant likelihood for any single term or interaction. The odds ratio for *rs1049269* was 0.97 with a 95% confidence interval (CI) between 0.92 and 1.03, while *rs1800562* had an odds ratio of 1.06 with a CI of 0.98 to 1.15, and the interaction term had an odds ratio of 0.99 with a CI of 0.86 to 1.14. The odds ratios and confidence intervals for main effects and the interaction in each cohort are displayed in Supplemental Figure 3.

Discussion

We failed to detect evidence of epistasis between *TF rs1049296* and *HFE rs1800562* as a risk for AD in the ADGC dataset. These findings do not support the conclusions drawn in the previous reports by Robson et al. (2004) and Kauwe et al. (Kauwe et al., 2010). The cause for this variability among studies could be a result of genetic heterogeneity, the complex nature of epistasis, or false positives in these previous studies due to limited sample size.

Although recent literature suggests that much of the unidentified genetic makeup of AD is due to epistasis (Bullock et al., 2013; Combarros et al., 2009; M. T. Ebbert et al., 2014; Infante et al., 2004; Mez, 2016), the complex nature of these gene-gene interactions makes it difficult to define specific epistatic interactions when multiple genes could be involved (Gilbert-Diamond & Moore, 2011; Kouyos, Silander, & Bonhoeffer, 2007; Urbanowicz, Kiralis, Fisher, & Moore, 2012). Models for epistatic interactions are challenging to create because the models require large datasets to analyze combinations of variables simultaneously (Moore & Williams, 2009).

When an insufficient number of samples are used, results have poor statistical power, which leads to frequent false negatives in gene-gene interaction studies. Likewise, the numerous comparisons required to assess epistasis may generate false positive findings (Mackay & Moore, 2014). Inadequate sample size can also result in false positives and is identified through statistical power analyses (Christley, 2010). The experiments performed by Robson et al. (2004) and Kauwe et al. (Kauwe et al., 2010) used datasets with much fewer individuals than the dataset used in this manuscript, and consequently have lower statistical power than our analysis. Although it is difficult to assess the proper significance threshold for power calculations, our study has significantly more power than the Kauwe et al. (2010) study regardless of the alpha value employed in the power calculation (see Supplemental Figure 1). Our analysis attains a power of .80 with an alpha value of just .003, whereas the Kauwe et al. (2010) study would need a significance threshold of .55 to reach the same level of power. Current research suggests a phenomenon known as the “winner’s curse,” which occurs when the estimated effect of an association is inflated compared to the true genetic effect and the effects later measured in follow-up studies (Huang, Ritchie, Brozynska, & Inouye, 2018; Palmer & Pe'er, 2017). The level of power necessary to accurately detect epistasis is currently

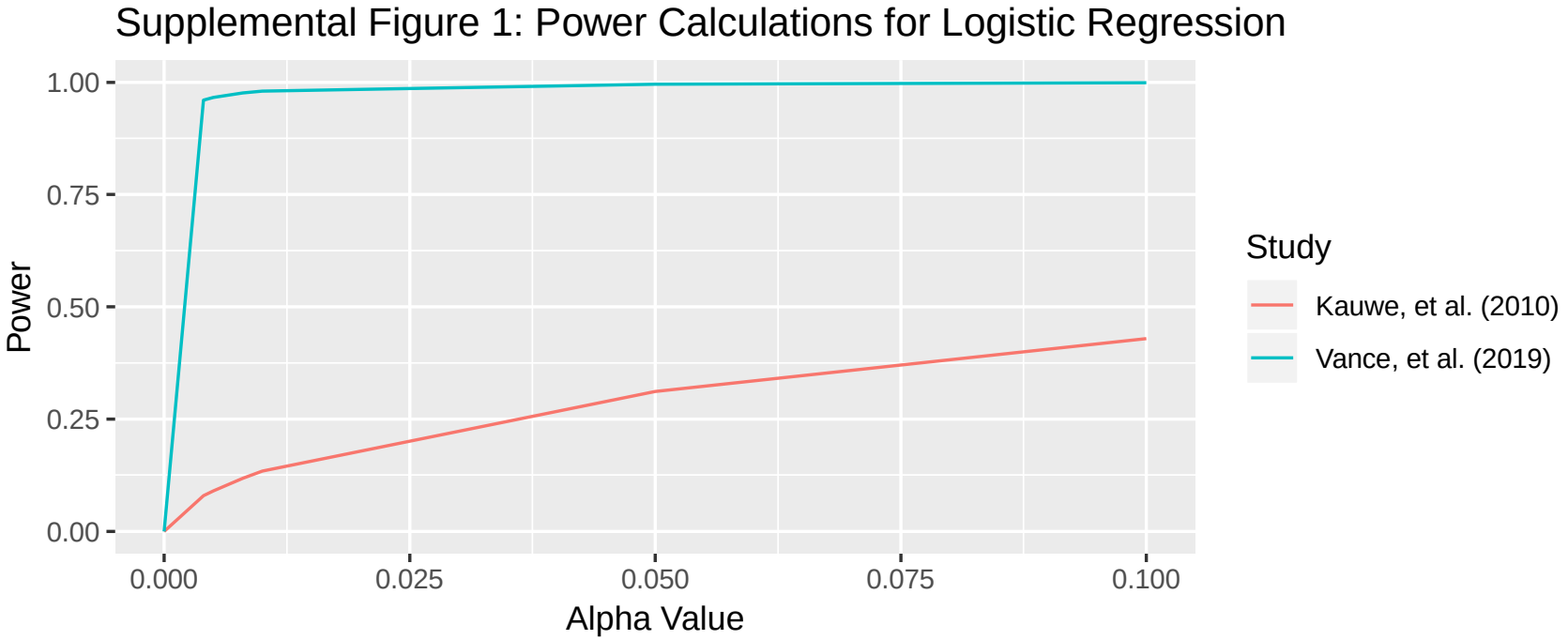
unknown, and as such, replication studies are a necessary part of validating epistasis. As our results show, statistical studies should be re-evaluated when larger datasets become available.

Heterogeneity in the genetic causes of AD is certainly present (Mez, 2016), and further erodes power to detect statistical epistasis. Similarly, combining various studies that use different diagnostic techniques could decrease our power to detect an epistatic signal if the classification criteria result in some patients being misclassified (Manchia et al., 2013). However, although the classification criteria for AD patients might vary depending on the sample, our analysis requires a large sample size in order to detect any synergetic relationship. Finally, even when statistical evidence for epistasis is detected, it does not necessarily indicate the presence of a physical biological interaction between the implicated proteins (M. T. W. Ebbert et al., 2015). Statistical patterns can be a product of a variety of underlying mechanisms. Therefore, the complexity of biological and statistical epistasis could also account for disparities in replication studies. Increasing sample sizes gives us better statistical power. Likewise, increasing the amount of multidimensional -omics data will help us focus our efforts on specific candidate interactions. For instance, we can use protein interaction networks and eQTLs to identify different loci that have similar effects on gene expression. This will help limit the search space of synergetic interactions. We anticipate that as more multidimensional -omics data become available, our ability to identify and understand the role of epistasis in AD risk will improve and help in the development of novel approaches to prevent and treat the disease.

Acknowledgements
ADGC grants: ADGC U01AG032984, NIAGADS U24AG041689, NCRAD U24AG21886, and NACC U01AG016976
See Supplemental Materials.

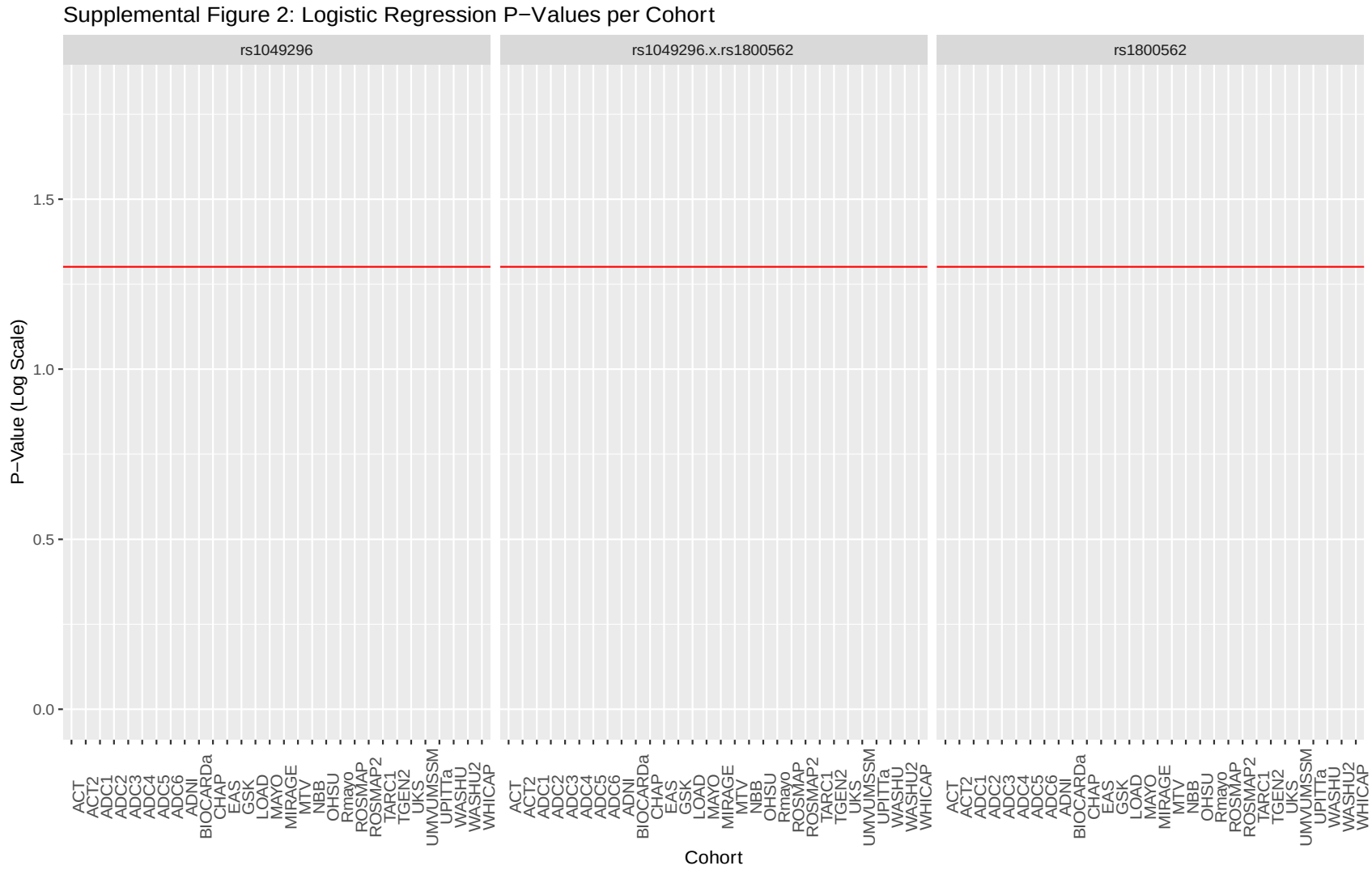
Declarations of Interest: none

Supplementary Figure 1 requires color



We computed power for the logisitic regression analyses used in our study (sample size = 25666) and the Kauwe et al. (2010) study (sample size = 2503). A range of alpha values is shown on the x−axis and the corresponding level of power is shown on the y−axis. With a sample size of 25666, our analysis attains a power of .80 with an alpha value of .003. With a sample size of 2503, the Kauwe et al. (2010) study would reach a power of .80 with an alpha level of .55. The power calculation tool we used can be found at https://www.dartmouth.edu/~eugened/power−samplesize.php (Demidenko, 2008; Demidenko, 2007).

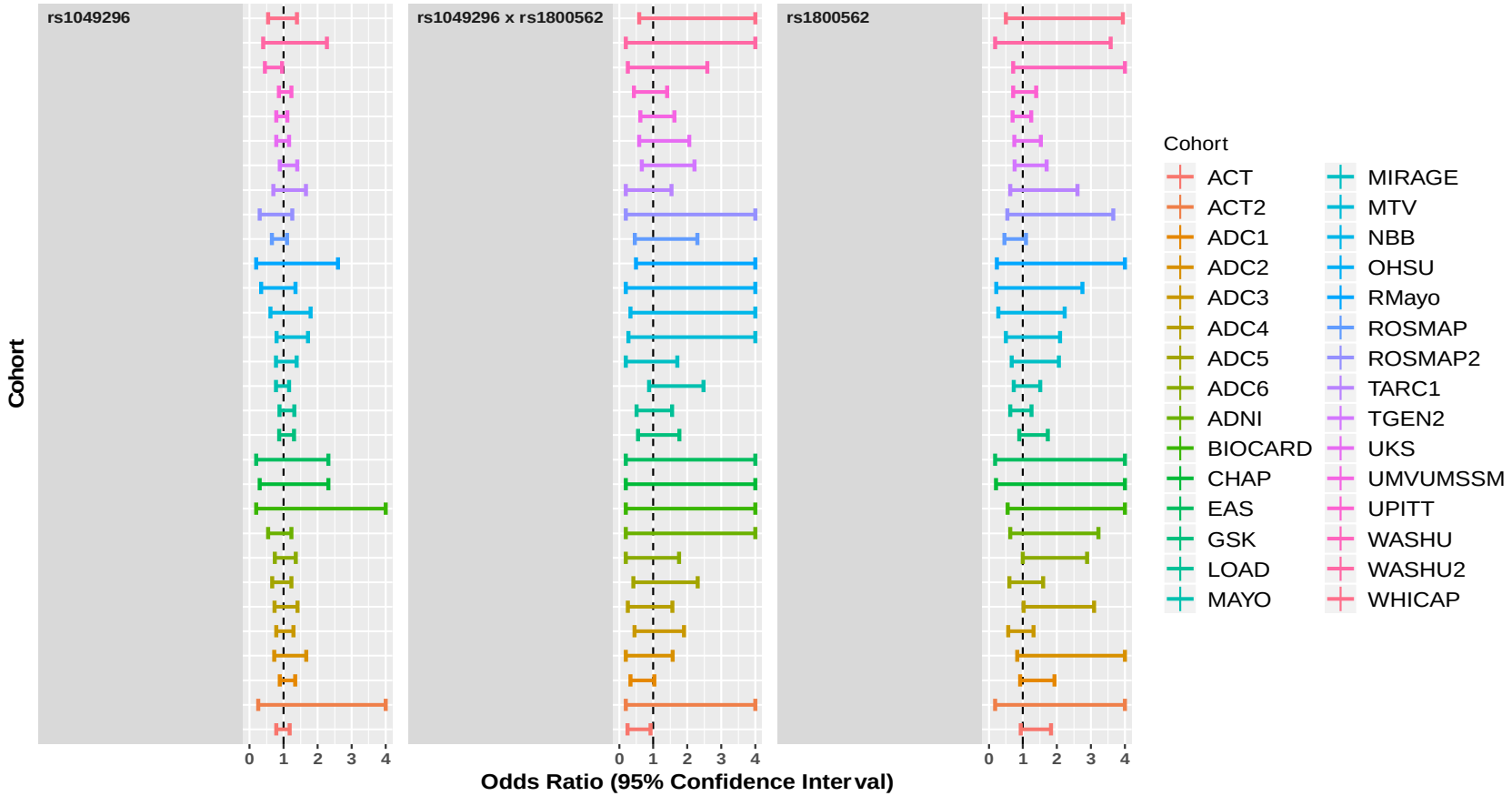
Supplementary Figure 2 requires color



We performed logistic regression on each cohort to determine the p−values for rs1049296, rs1800562, and the epistatic interaction of these variants. Each cohort is shown on the x−axis,and the p−value for each cohort is shown on the y−axis. The red line indicates the alpha value of 0.05. From our analysis, only the ACT and WASHU cohorts have significant p−values at these variants.

Supplementary Figure 3 requires color

Supplemental Figure 3: Logistic Regression Odds Ratios and Confidence Intervals per Cohort



We performed logistic regression on each cohort to determine the odds ratios for rs1049296, rs1800562, and the epistatic interaction of these variants. Each cohort is shown on the y−axis, and the odds ratio with the 95% confidence interval for each cohort is shown on the x−axis. Lower limit values are truncated to .2, while upper limit and odds ratio values are truncated to 4. The dashed line indicates an odds ratio value of 1. Although the ACT2, ADNI, BIOCARD, RMayo, and WHICAP cohorts have seemingly high odds ratio values (>2), it is important to note that the confidence intervals for these cohorts span the null value (1) and are not precise. Furthermore, the p−values for these cohorts suggest that the effects are not significant (see Figure 1).

Supplementary Table 1: ADGC Population Demographics

	N	Age at Onset	% Female	% APOE e4 positive
ADNI				
Cases	215	73.33	42.33%	65.58%
Controls	140		37.86%	25.00%
WASHU				
Cases	312	74.21	58.01%	51.60%
Controls	166		59.64%	25.90%
ROSMAP2				
Cases	62	NA	72.58%	16.13%
Controls	237		78.06%	33.33%
WASHU2				
Cases	30	74.91	53.33%	53.33%
Controls	65		44.62%	27.69%
EAS				
Cases	10	86.4	40.00%	20.00%
Controls	209		51.20%	14.35%
CHAP				
Cases	20	85.35	70.00%	30.00%
Controls	164		57.93%	17.68%
ADC3				
Cases	711	74.47	53.16%	49.93%
Controls	464		61.64%	21.55%
TARC1				
Cases	286	73.71	57.34%	51.05%
Controls	144		66.67%	21.53%
MTV				
Cases	241	73.35	57.68%	50.62%
Controls	194		61.86%	10.31%
MAYO				
Cases	616	NA	58.60%	55.68%
Controls	925		51.24%	24.97%
ACT				
Cases	479	83.88	61.80%	41.75%
Controls	1348		55.49%	18.32%
ACT2				
Cases	18	83.71	66.67%	44.44%
Controls	5		60.00%	60.00%
ADC4				
Cases	287	73.34	54.01%	41.11%
Controls	340		62.06%	22.94%
ADC6				
Cases	363	74.08	55.65%	27.27%
Controls	304		65.46%	23.68%
ADC1				
Cases	1503	72.37	52.50%	52.03%
Controls	543		60.96%	20.44%
ADC2				
Cases	546	73.45	49.45%	55.49%
Controls	121		68.60%	23.97%
ADC5				
Cases	273	73.33	53.48%	56.78%
Controls	496		64.52%	18.55%

RMAYO				
Cases	12	79.89	8.33%	25.00%
Controls	271		40.59%	15.87%
BIOCARD				
Cases	8	73.83	37.50%	37.50%
Controls	123		63.41%	28.46%
LOAD				
Cases	745	73.28	66.31%	72.62%
Controls	801		60.92%	27.47%
WASHU				
Cases	312	74.21	58.01%	51.60%
Controls	166		59.64%	25.90%
MIRAGE				
Cases	398	71.64	60.55%	46.48%
Controls	294		61.22%	31.97%
OHSU				
Cases	59	85.74	61.02%	45.76%
Controls	109		53.21%	16.51%
UPITT				
Cases	1267	72.93	63.22%	56.83%
Controls	834		63.31%	19.42%
UMVUMSSM				
Cases	1085	73.83	64.61%	56.41%
Controls	1112		61.06%	2.07%
WHICAP				
Cases	74	84.04	71.62%	21.62%
Controls	562		60.32%	20.46%
GSK				
Cases	796	74.58	57.79%	53.39%
Controls	764		64.40%	21.99%
NBB				
Cases	215	NA	72.09%	13.95%
Controls	85		57.65%	5.88%
ROSMAP				
Cases	364	85.59	68.13%	28.02%
Controls	853		70.34%	13.48%
TGEN2				
Cases	770	74.6	60.78%	47.27%
Controls	448		51.34%	14.29%
UKS				
Cases	767	72.24	56.06%	43.02%
Controls	973		47.79%	0.10%

Supplementary Table 2: Population Demographics (Kauwe et al., 2010)

	N	Average Age	% Female	% APOE e4 positive
WASHU				
Cases	331	76.6	62.00%	53.60%
Controls	385	77.7	61.00%	23.20%
MRC				
Cases	631	75.7	73.00%	62.20%
Controls	769	76.1	62.00%	23.40%
ADNI				
Cases	199	71.8	56.00%	65.00%
Controls	188	77.7	55.00%	27.80%

Supplementary Table 3: SFA Calculations

HFE rs1800562	TF rs1049296	Controls	Cases	Odds ratio	ln(OR)	var ln(OR)	se ln(OR)			lower	upper	alpha
-	-	8201	7832	Reference				**SF=**	**0.938229**	**0.7867**	**1.1190**	**0.05**
+	-	951	980	1.07905	0.07608	0.00232	0.04818	**se(ln(SF))=**	**0.089889**			
-	+	3174	2964	0.97783	-0.02241	0.00090	0.03003	**ln(SF)=**	**-0.063761**			
+	+	403	381	0.98995	-0.01010	0.00536	0.07318	**Z=**	**-0.709329**			
								p=	**0.47812**			

Reference

2018 Alzheimer's disease facts and figures. (2018). *Alzheimer's & Dementia, 14*(3), 367-429. doi:https://doi.org/10.1016/j.jalz.2018.02.001

Ali-Rahmani, F., Schengrund, C.-L., & Connor, J. R. (2014). HFE gene variants, iron, and lipids: a novel connection in Alzheimer's disease. *Frontiers in Pharmacology, 5*, 165. doi:10.3389/fphar.2014.00165

Andersen, H. H., Johnsen, K. B., & Moos, T. (2014). Iron deposits in the chronically inflamed central nervous system and contributes to neurodegeneration. *Cell Mol Life Sci, 71*(9), 1607-1622. doi:10.1007/s00018-013-1509-8

Belaidi, A. A., Gunn, A. P., Wong, B. X., Ayton, S., Appukuttan, A. T., Roberts, B. R., . . . Bush, A. I. (2018). Marked Age-Related Changes in Brain Iron Homeostasis in Amyloid Protein Precursor Knockout Mice. *Neurotherapeutics*. doi:10.1007/s13311-018-0656-x

Bennett, M. J., Lebron, J. A., & Bjorkman, P. J. (2000). Crystal structure of the hereditary haemochromatosis protein HFE complexed with transferrin receptor. *Nature, 403*(6765), 46-53. doi:10.1038/47417

Boehme, K. L., Mukherjee, S., Crane, P. K., & Kauwe, J. S. (September 2014). *ADGC 1000 genomes combined workflow (electronic document)*. Retrieved from.
http://kauwelab.byu.edu/Portals/22/adgc_combined_1000G_12032014.pdf

Bullock, J. M., Medway, C., Cortina-Borja, M., Turton, J. C., Prince, J. A., Ibrahim-Verbaas, C. A., . . . Morgan, K. (2013). Discovery by the Epistasis Project of an epistatic interaction between the GSTM3 gene and the HHEX/IDE/KIF11 locus in the risk of Alzheimer's disease. *Neurobiol Aging, 34*(4), 1309.e1301-1307. doi:10.1016/j.neurobiolaging.2012.08.010

Christley, R. (2010). *Power and Error: Increased Risk of False Positive Results in Underpowered Studies* (Vol. 3).

Combarros, O., Cortina-Borja, M., Smith, A. D., & Lehmann, D. J. (2009). Epistasis in sporadic Alzheimer's disease. *Neurobiol Aging, 30*(9), 1333-1349. doi:10.1016/j.neurobiolaging.2007.11.027

Cordell, H. J. (2002). Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Human Molecular Genetics, 11*(20), 2463-2468. doi:10.1093/hmg/11.20.2463

Cortina-Borja, M., Smith, A. D., Combarros, O., & Lehmann, D. J. (2009). The synergy factor: a statistic to measure interactions in complex diseases. *BMC Res Notes, 2*, 105. doi:10.1186/1756-0500-2-105

Demidenko, E. (2008). Sample size and optimal design for logistic regression with binary interaction. *Statist. Med., 27,* 36-46. doi:10/1002/sim.2980

Demidenko, E. (2007). Sample size determination for logistic regression revisited. *Statist.Med., 26,* 3385-3397. doi:10.1002/sim.2771

Dias, V., Junn, E., & Mouradian, M. M. (2013). The role of oxidative stress in Parkinson's disease. *Journal of Parkinson's disease, 3*(4), 461-491. doi:10.3233/JPD-130230

Ebbert, M. T., Ridge, P. G., Wilson, A. R., Sharp, A. R., Bailey, M., Norton, M. C., . . . Kauwe, J. S. (2014). Population-based analysis of Alzheimer's disease risk alleles implicates genetic interactions. *Biol Psychiatry, 75*(9), 732-737. doi:10.1016/j.biopsych.2013.07.008

Ebbert, M. T. W., Ridge, P. G., & Kauwe, J. S. K. (2015). Bridging the Gap between Statistical and Biological Epistasis in Alzheimer’s Disease. *BioMed Research International, 2015*, 7. doi:10.1155/2015/870123

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods, 39*(2), 175-191.

Feder, J. N., Gnirke, A., Thomas, W., Tsuchihashi, Z., Ruddy, D. A., Basava, A., . . . Wolff, R. K. (1996). A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet, 13*(4), 399-408. doi:10.1038/ng0896-399

Gilbert-Diamond, D., & Moore, J. H. (2011). Analysis of Gene-Gene Interactions. *Curr Protoc Hum Genet, 0 1*, Unit1.14. doi:10.1002/0471142905.hg0114s70

Gkouvatsos, K., Papanikolaou, G., & Pantopoulos, K. (2012). Regulation of iron transport and the role of transferrin. *Biochim Biophys Acta, 1820*(3), 188-202. doi:10.1016/j.bbagen.2011.10.013

Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M. L., . . . Williams, J. (2009). Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet, 41*(10), 1088-1093. doi:10.1038/ng.440

Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M. L., . . . Williams, J. (2009). Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet, 41*(10), 1088-1093. doi:10.1038/ng.440

Hollingworth, P., Harold, D., Sims, R., Gerrish, A., Lambert, J.-C., Carrasquillo, M. M., . . . Williams, J. (2011). Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet, 43*(5), 429-435. doi:10.1038/ng.803

Huang, Q. Q., Ritchie, S. C., Brozynska, M., & Inouye, M. (2018). Power, false discovery rate and Winner's Curse in eQTL studies. *Nucleic acids research, 46*(22), e133-e133. doi:10.1093/nar/gky780

Infante, J., Sanz, C., Fernández-Luna, J. L., Llorca, J., Berciano, J., & Combarros, O. (2004). Gene–gene interaction between interleukin-6 and interleukin-10 reduces AD risk. *Neurology, 63*(6), 1135-1136. doi:10.1212/01.Wnl.0000138570.96291.A8

Jahshan, A., Esteves-Villanueva, J. O., & Martic-Milne, S. (2016). Evaluation of ferritin and transferrin binding to tau protein. *J Inorg Biochem, 162*, 127-134. doi:10.1016/j.jinorgbio.2016.06.022

Kauwe, J. S. K., Bertelsen, S., Mayo, K., Cruchaga, C., Abraham, R., Hollingworth, P., . . . Alzheimer's Disease Neuroimaging, I. (2010). Suggestive synergy between genetic variants in TF and HFE as risk factors for Alzheimer's disease. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics, 153B*(4), 955-959. doi:10.1002/ajmg.b.31053

Kouyos, R. D., Silander, O. K., & Bonhoeffer, S. (2007). Epistasis between deleterious mutations and the evolution of recombination. *Trends in Ecology & Evolution, 22*(6), 308-315. doi:https://doi.org/10.1016/j.tree.2007.02.014

Lambert, J.-C., Heath, S., Even, G., Campion, D., Sleegers, K., Hiltunen, M., . . . Amouyel, P. (2009). Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet, 41*(10), 1094-1099. doi:10.1038/ng.439

Lambert, J. C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., . . . Amouyel, P. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet, 45*(12), 1452-1458. doi:10.1038/ng.2802

Lebron, J. A., Bennett, M. J., Vaughn, D. E., Chirino, A. J., Snow, P. M., Mintier, G. A., . . . Bjorkman, P. J. (1998). Crystal structure of the hemochromatosis protein HFE and characterization of its interaction with transferrin receptor. *Cell, 93*(1), 111-123.

Lehmann, D. J., Worwood, M., Ellis, R., Wimhurst, V. L. J., Merryweather-Clarke, A. T., Warden, D. R., . . . Robson, K. J. H. (2006). Iron genes, iron load and risk of Alzheimer's disease. *Journal of Medical Genetics, 43*(10), e52-e52. doi:10.1136/jmg.2006.040519

Loh, P.-R., Danecek, P., Palamara, P. F., Fuchsberger, C., A Reshef, Y., K Finucane, H., . . . L Price, A. (2016). Reference-based phasing using the Haplotype Reference Consortium panel. *Nature genetics, 48*(11), 1443-1448. doi:10.1038/ng.3679

Lull, M. E., & Block, M. L. (2010). Microglial activation and chronic neurodegeneration. *Neurotherapeutics, 7*(4), 354-365. doi:10.1016/j.nurt.2010.05.014

Mackay, T. F., & Moore, J. H. (2014). Why epistasis is important for tackling complex human disease genetics. *Genome Medicine, 6*(6), 42. doi:10.1186/gm561

Manchia, M., Cullis, J., Turecki, G., Rouleau, G. A., Uher, R., & Alda, M. (2013). The impact of phenotypic and genetic heterogeneity on results of genome wide association studies of complex diseases. *PLOS ONE, 8*(10), e76295. doi:10.1371/journal.pone.0076295

Matak, P., Matak, A., Moustafa, S., Aryal, D. K., Benner, E. J., Wetsel, W., & Andrews, N. C. (2016). Disrupted iron homeostasis causes dopaminergic neurodegeneration in mice. *Proceedings of the National Academy of Sciences of the United States of America, 113*(13), 3428-3435. doi:10.1073/pnas.1519473113

Mez, J. (2016). The executive prominent/memory prominent spectrum in Alzheimer’s disease is highly heritable. *41*, 115-121. doi:10.1016/j.neurobiolaging.2016.02.015

Moore, J. H., & Williams, S. M. (2009). Epistasis and Its Implications for Personal Genetics. *The American Journal of Human Genetics, 85*(3), 309-320. doi:https://doi.org/10.1016/j.ajhg.2009.08.006

Naj, A. C., Below, J., Zhao, Y., Chen, H.-H., van der Lee, S. J., Hamilton-Nelson, K. L., . . . Schellenberg, G. D. (2017). GENOME-WIDE RARE VARIANT IMPUTATION AND TISSUE-SPECIFIC TRANSCRIPTOMIC ANALYSIS IDENTIFY NOVEL RARE VARIANT CANDIDATE LOCI IN LATE-ONSET ALZHEIMER’S DISEASE: THE ALZHEIMER’S DISEASE GENETICS CONSORTIUM. *Alzheimer's & Dementia, 13*(7, Supplement), P189. doi:https://doi.org/10.1016/j.jalz.2017.07.045

Naj, A. C., Jun, G., Beecham, G. W., Wang, L. S., Vardarajan, B. N., Buros, J., . . . Schellenberg, G. D. (2011). Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet, 43*(5), 436-441. doi:10.1038/ng.801

Palmer, C., & Pe'er, I. (2017). Statistical correction of the Winner's Curse explains replication variability in quantitative trait genome-wide association studies. *PLoS genetics, 13*(7), e1006916-e1006916. doi:10.1371/journal.pgen.1006916

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genet, 2*(12), e190. doi:10.1371/journal.pgen.0020190

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet, 38*(8), 904-909. doi:10.1038/ng1847

Raghavan, N., & Tosto, G. (2017). Genetics of Alzheimer's Disease: the Importance of Polygenic and Epistatic Components. *Current neurology and neuroscience reports, 17*(10), 78-78. doi:10.1007/s11910-017-0787-1

Ridge, P. G., Mukherjee, S., Crane, P. K., Kauwe, J. S. K., & Alzheimer's Disease Genetics, C. (2013). Alzheimer's Disease: Analyzing the Missing Heritability. *PLOS ONE, 8*(11), e79771. doi:10.1371/journal.pone.0079771

Robson, K. J., Lehmann, D. J., Wimhurst, V. L., Livesey, K. J., Combrinck, M., Merryweather-Clarke, A. T., . . . Smith, A. D. (2004). Synergy between the C2 allele of transferrin and the C282Y allele of the haemochromatosis gene (HFE) as risk factors for developing Alzheimer's disease. *J Med Genet, 41*(4), 261-265.

Seshadri, S., Fitzpatrick, A. L., Ikram, M. A., DeStefano, A. L., Gudnason, V., Boada, M., . . . Consortium, E. (2010). Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA, 303*(18), 1832-1840. doi:10.1001/jama.2010.574

Shen, N., Chen, B., Jiang, Y., Feng, R., Liao, M., Zhang, L., . . . Liu, G. (2015). An Updated Analysis with 85,939 Samples Confirms the Association Between CR1 rs6656401 Polymorphism and Alzheimer's Disease. *Molecular Neurobiology, 51*(3), 1017-1023. doi:10.1007/s12035-014-8761-2

Shuai, P., Liu, Y., Lu, W., Liu, Q., Li, T., & Gong, B. (2015). Genetic associations of CLU rs9331888 polymorphism with Alzheimer's disease: A meta-analysis. *Neuroscience Letters, 591*, 160-165. doi:https://doi.org/10.1016/j.neulet.2015.02.040

Tisato, V., Zuliani, G., Vigliano, M., Longo, G., Franchini, E., Secchiero, P., . . . Gemmati, D. (2018). Gene-gene interactions among coding genes of iron-homeostasis proteins and APOE-alleles in cognitive impairment diseases. *PLoS One, 13*(3), e0193867. doi:10.1371/journal.pone.0193867

Urbanowicz, R. J., Kiralis, J., Fisher, J. M., & Moore, J. H. (2012). Predicting the difficulty of pure, strict, epistatic models: metrics for simulated model selection. *BioData Mining, 5*(1), 15. doi:10.1186/1756-0381-5-15

Yan, H., Kong, Y., He, B., Huang, M., Li, J., Zheng, J., . . . Shi, L. (2015). CYP2J2 rs890293 polymorphism is associated with susceptibility to Alzheimer's disease in the Chinese Han population. *Neuroscience Letters, 593*, 56-60. doi:https://doi.org/10.1016/j.neulet.2015.03.024

Yarjanli, Z., Ghaedi, K., Esmaeili, A., Rahgozar, S., & Zarrabi, A. (2017). Iron oxide nanoparticles may damage to the neural tissue through iron accumulation, oxidative stress, and protein aggregation. *BMC neuroscience, 18*(1), 51-51. doi:10.1186/s12868-017-0369-9

CRediT Author Statement

Elizabeth Vance: Conceptualization, Methodology, Formal analysis, Writing—Original Draft, Writing—Review & Editing, Visualization **Josue D. Gonzalez Murcia:** Methodology, Writing—Original Draft **Justin B. Miller:** Writing—Original Draft, Writing—Review & Editing, Supervision **Alzheimer's Disease Genetic Consortium:** Data Curation **Lyndsay Staley:** Methodology, Formal analysis **Paul K. Crane:** Validation, Data Curation **Shubhabrata Mukherjee:** Validation, Data Curation **John S.K. Kauwe:** Supervision, Data Curation