

# SCIENTIFIC REPORTS



OPEN

## Common BACE2 Polymorphisms are Associated with Altered Risk for Alzheimer's Disease and CSF Amyloid Biomarkers in APOE $\epsilon$ 4 Non-Carriers

Matt Huentelman<sup>1,2</sup>, Matthew De Both<sup>1,2</sup>, Wayne Jepsen<sup>1,2</sup>, Ignazio S. Piras<sup>1,2</sup>, Joshua S. Talboom<sup>1,2</sup>, Mari Willeman<sup>1,2</sup>, Eric M. Reiman<sup>2,3,4,5</sup>, John Hardy<sup>6</sup> & Amanda J. Myers<sup>7</sup>

It was recently suggested that beta-site amyloid precursor protein (APP)-cleaving enzyme 2 (BACE2) functions as an amyloid beta (A $\beta$ )-degrading enzyme; in addition to its better understood role as an APP secretase. Due to this finding we sought to understand the possible genetic risk contributed by the *BACE2* locus to the development of late-onset Alzheimer's disease (AD). In this study, we report that common single nucleotide polymorphism (SNP) variation in *BACE2* is associated with altered AD risk in apolipoprotein E gene (APOE) epsilon 4 variant ( $\epsilon$ 4) non-carriers. In addition, in  $\epsilon$ 4 non-carriers diagnosed with AD or mild cognitive impairment (MCI), SNPs within the *BACE2* locus are associated with cerebrospinal fluid (CSF) levels of A $\beta$ 1-42. Further, SNP variants in *BACE2* are also associated with BACE2 RNA expression levels suggesting a potential mechanism for the CSF A $\beta$ 1-42 findings. Lastly, overexpression of BACE2 *in vitro* resulted in decreased A $\beta$ 1-40 and A $\beta$ 1-42 fragments in a cell line model of A $\beta$  production. These findings suggest that genetic variation at the *BACE2* locus modifies AD risk for those individuals who don't carry the  $\epsilon$ 4 variant of APOE. Further, our data indicate that the biological mechanism associated with this altered risk is linked to amyloid generation or clearance possibly through BACE2 expression changes.

The amyloid (A $\beta$ ) hypothesis of Alzheimer's disease (AD) has primarily been driven by the observation that genetic variability that alters amyloid precursor protein (APP) metabolism increases deposition of A $\beta$ <sup>1</sup>. Classically, the causative mutations were in *APP* itself and in the presenilin genes – PSEN1 and 2. A $\beta$  is generated by the sequential actions of  $\beta$ -secretase (BACE1) and  $\gamma$ -secretase, of which the presenilins are the major component<sup>2</sup>. More recently, reduced function variants of both ADAM10<sup>3</sup> and ADAM17<sup>4</sup>, the enzymes involved in the alternate  $\alpha$ -secretase cleavage of APP, have been shown to increase the risk of AD as well. Finally, it is clear that the homologue of BACE1, BACE2 can also cleave both APP and amyloid beta (A $\beta$ )<sup>5,6</sup>.

The actions of BACE2 may be involved in altering pathogenic A $\beta$  fragment concentrations in multiple ways. It is known that BACE2 can cleave APP at the  $\beta$ -secretase site (albeit with much lower efficiency than BACE1) and near the  $\alpha$ -secretase site as well<sup>7-9</sup>. Fluhner *et al.* demonstrated that the cleavage of APP via BACE2 results in production of soluble APP- $\alpha$  and reduced production of amyloidogenic species suggesting that APP metabolism by BACE2 acts in an opposing fashion to BACE1 in cells that express both enzymes<sup>8</sup>. However, more recent research suggests that the primary activity of BACE2 occurs at a " $\theta$ -secretase" site (Phe<sup>10</sup>) located within the A $\beta$  domain

<sup>1</sup>Neurogenomics Division, The Translational Genomics Research Institute (TGen), Phoenix, AZ, USA. <sup>2</sup>The Arizona Alzheimer's Consortium, Phoenix, AZ, USA. <sup>3</sup>Banner Alzheimer's Institute, Phoenix, AZ, USA. <sup>4</sup>Neurodegenerative Disease Research Center, BioDesign Institute, Arizona State University, Tempe, AZ, USA. <sup>5</sup>Department of Psychiatry, University of Arizona, Tucson, AZ, USA. <sup>6</sup>Department of Molecular Neuroscience, Reta Lilla Weston Laboratories, and UK Dementia Research Institute at University College London, Institute of Neurology, London, United Kingdom. <sup>7</sup>Department of Psychiatry & Behavioral Sciences, Programs in Neuroscience and Human Genetics and Genomics and Center on Aging, Miller School of Medicine, University of Miami, Miami, FL, USA. Correspondence and requests for materials should be addressed to M.H. (email: [mhuentelman@tgen.org](mailto:mhuentelman@tgen.org))

in APP. In 2006, Sun *et al.* identified this novel  $\theta$ -secretase site and demonstrated that it results in the prevention of A $\beta$  production<sup>5</sup>. Lastly, Abdul-Hay *et al.* demonstrated that BACE2 can degrade A $\beta$  with an efficiency second only to insulin degrading enzyme (IDE)<sup>6</sup>.

Although two studies, one in the Finnish population<sup>11</sup> and one in Down syndrome, have both implicated BACE2 in disease, to date neither exome sequencing or genome-wide association studies have reported common or rare pathogenic variability in either BACE1 or BACE2 as altering risk for disease. However, Wang *et al.* recently reported that mutations in the juxtamembrane helix (JH) domain of BACE2 results in a “conditional”  $\beta$ -secretase activity that leads to elevated A $\beta$  production thereby establishing a potential genetic basis for BACE2 variation in AD risk<sup>12</sup>.

With this in mind we examined the *BACE2* locus, using a clinically characterized and neuropathology confirmed sample set (TGenII) as our starting point. We have previously shown that these confirmed samples increase the power to detect genetic associations. This is presumably because they avoid the misclassification of controls and the misdiagnosis of cases<sup>13</sup>. Next, we attempted to replicate findings in the Alzheimer’s Disease Neuroimaging Initiative (ADNI) cohort by examining the cerebrospinal fluid (CSF) biomarker data. Additionally, we examined the *BACE2* locus for an expression quantitative trait (eQTL) to determine if variants in the region could be associated with *BACE2* expression levels. Lastly, we confirmed that alteration of *BACE2* expression influenced A $\beta$  levels *in vitro*.

## Methods

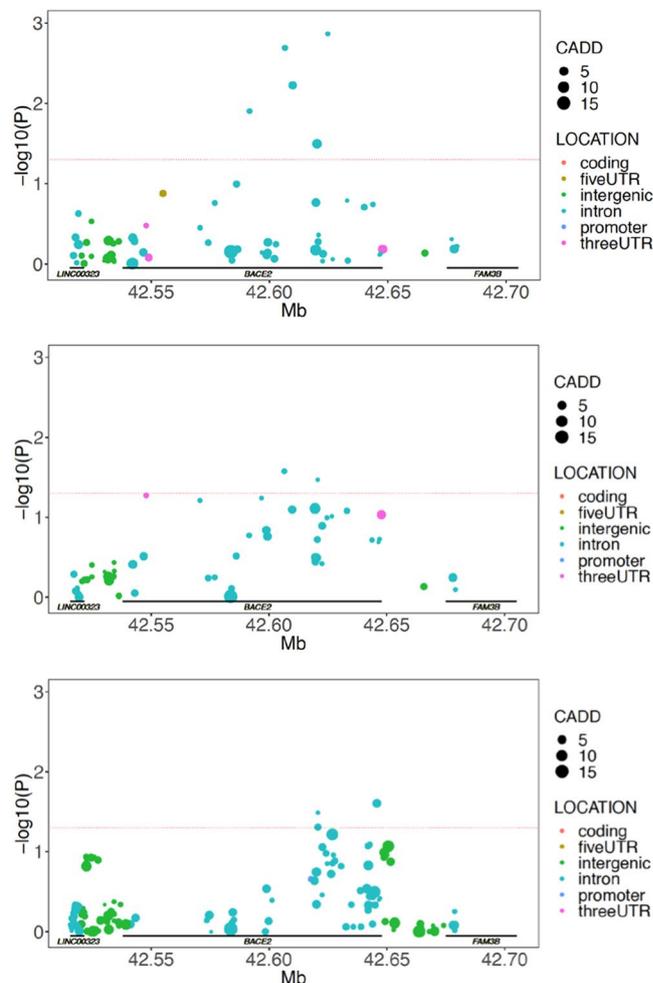
**“TGenII” clinically characterized and neuropathologically verified cohort.** Subjects from the TGenII neuropath cohort were collected as previously described<sup>14</sup>. We selected cases and controls as subjects of European ancestry who were verified pathologically as having AD or not, respectively. Subjects were greater than age 65 at death. Board-certified neuropathologists performed Braak staging based on the degree of neurofibrillary tangles<sup>15</sup>, and/or CERAD<sup>16</sup> classification on the extent of neuritic plaques. Any samples with a history of other known neurological disease and/or clinical history of stroke, cerebrovascular disease, and/or Lewy bodies were excluded.

**ADNI subjects.** Subjects from the ADNI cohort (adni.loni.usc.edu) were collected as described previously<sup>17</sup>. Study participants were given both clinical and cognitive assessments. The mini-mental state exam (MMSE)<sup>18</sup> was used to screen for and score cognitive impairment along with the clinical dementia rating (CDR)<sup>19</sup> to determine level of dementia. AD subjects were determined as clinically according to the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA) criteria<sup>20</sup> and having MMSE scores between 20–26 and CDR between 0.5 and 1.0. Mild cognitive impairment (MCI) subjects had MMSE scores between 24–30, CDR = 0.5, reported memory complaints but no other signs of impairment of other cognitive functions, and were non-demented. Healthy normal controls (NC) had MMSE of 24–30, CDR = 0, and showed no signs of dementia. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer’s disease (AD). For up-to-date information, see [www.adni-info.org](http://www.adni-info.org).

**Single nucleotide polymorphism (SNP) genotyping.** The TGenII cohort was genotyped on the Affymetrix Genome-Wide Human SNP 6.0 Array (Affymetrix, Inc., Santa Clara, CA, USA) with standard manufacturer’s protocols (Affymetrix Genome-Wide Human SNP Nsp/Sty 6.0 User Guide; Rev. 1, 2007). The ADNI cohort was genotyped on the Illumina Human610-Quad BeadChip (Illumina, Inc. San Diego, CA) according to the manufacturer’s protocols (Infinium HD Assay; Super Protocol Guide; Rev. A, May 2008). Both SNP datasets are free and publicly available either via the main study sites or through the Alzheimer’s Disease Genetic Consortium (ADGC or <https://www.nigads.org/datasets/ng00028>).

**BACE2 locus imputation.** Imputation was conducted with IMPUTE2 software (v2.1.2) using the multi-population reference panel approach<sup>10,21</sup>. We used the June 2011 release of the 1000 Genomes (1000genomes.org) Phase 1. We imputed a 5 megabase region encompassing *BACE2* on chromosome 21 between positions 40,000,000–45,000,000 in hg19 coordinates. For the phasing stage we used 80 haplotypes (-k option) as a template. For both the phasing and imputation rounds of IMPUTE2, we performed 10 burn in iterations (-burnin option) and performed 30 Markov chain Monte Carlo (MCMC) iterations (-iter option). Since we performed the multi-population approach, we set the “effective size” option (-Ne) to the suggested value of 20000. In addition, we utilized the strand alignment procedure (-fix\_strand\_g option) to minimize genotype strand discrepancies between the reference and study panels. GTOOL was used to convert IMPUTE2 GEN output to PED format.

**SNP and haplotype association testing.** SNPs with minor allele frequencies less than 1% or with genotyping rate below 95% were filtered prior to analysis. For the initial hypothesis testing of BACE1 and BACE2 a Fisher’s exact test was utilized and 1000 max(T) permutations in PLINK<sup>22</sup> were performed. After determining BACE2 warranted further follow-up, we assessed population structure and incorporated the results as covariates in regression models. We used ADMIXTURE<sup>23</sup> (v1.04) with K = 3 by selecting a subset of directly genotyped SNPs (i.e., not imputed) with >99% call rates, minor allele frequency >0.3, pairwise R<sup>2</sup> < 0.01. Q1 and Q2 vector solutions and sex were included as covariates in the regressions, and APOE  $\epsilon$ 4 carrier status was included as a covariate and genotype interaction term. Haplotypes were called with PLINK using the default parameters. Correction for multiple testing during SNP analysis was conducted using the Bonferroni method with independent SNPs considered to be those with r<sup>2</sup> < 0.80.



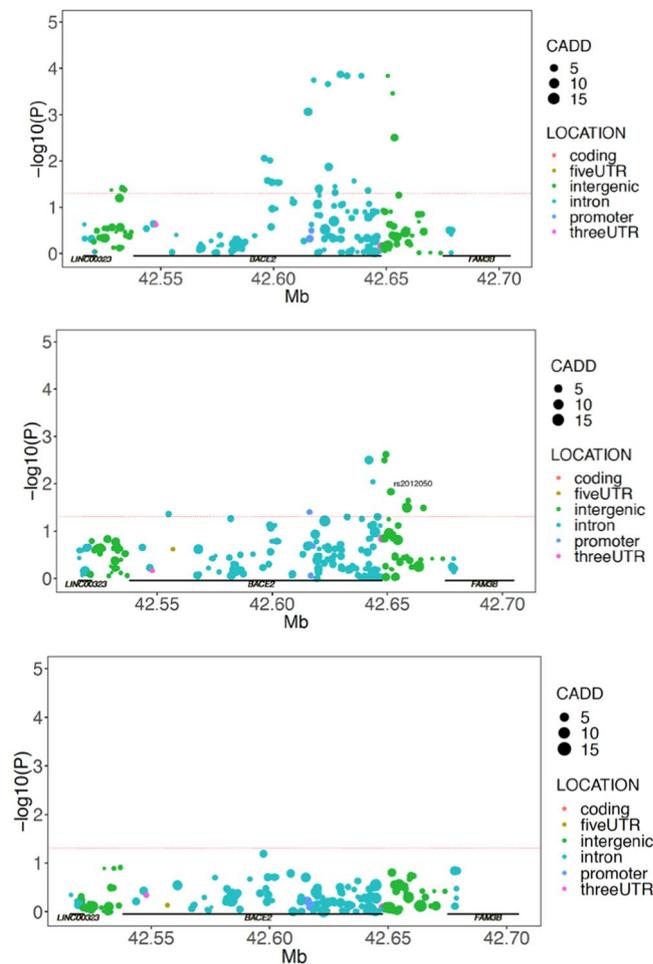
**Figure 1.** AD risk is associated with SNPs within the *BACE2* locus in *APOE*  $\epsilon 4$  non-carriers. *BACE2* association in the clinically characterized and neuropathologically verified cohort. In all panels, SNP uncorrected P-values are plotted as colored circles based on their location within the genetic locus, the size of the circle is representative of the CADD score for the SNP. A horizontal red dashed line represents nominal significance of  $p = 0.05$ . (Top panel) *BACE2* association in all discovery subjects ( $N = 1599$ ; 1014 cases; 585 controls). (Middle) *BACE2* interaction p-values with *APOE* carrier status (Bottom) *BACE2* interaction p-values with *APOE* carrier status including the first two population structure vectors and sex.

**A $\beta$  CSF measures.** A $\beta$  CSF levels were used from the ADNI dataset. The approach for CSF collection and measurement of the A $\beta$  biomarker was reported previously<sup>17</sup>. The approach for testing A $\beta$ 1-42 fragment association with *BACE2* was carried out in a similar fashion to the AD association; however, instead of using case/control as binary phenotypes, we performed a linear regression with A $\beta$ 1-42 levels as the phenotype. The imputed *BACE2* SNPs were tested against A $\beta$ 1-42 CSF measures with sex and ADMIXTURE  $K = 3$  Q1 and Q2 values as covariates.

**In vitro A $\beta$  ELISA measurements.** *BACE2* overexpression was achieved using the Sleeping Beauty Transposon system<sup>24</sup>. BE(2)-m17 neuroblastoma cells were plated in uncoated, 6-well plates (14e6 cells/well) and media exchanged every 24 hours. Seventy-two-hours after plating-, supernatant was collected and supplemented with 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) to a final concentration of 1.0 mM. Samples were clarified, decanted, and stored at  $-80^{\circ}\text{C}$  until analysis. Extracellular A $\beta$  levels were quantified using the Human  $\beta$ Amyloid [1-40] (ThermoFisher Scientific) and the Human A $\beta$  [1-42] ELISAs (ThermoFisher Scientific), according to the manufacturer's instructions.

### Highlights.

- SNPs within the *BACE2* locus were associated with altered Alzheimer's disease risk preferentially in *APOE*  $\epsilon 4$  non-carriers.
- In individuals with diagnosed AD or MCI who do not carry an *APOE*  $\epsilon 4$  allele, SNPs within *BACE2* were associated with differential A $\beta$ 1-42 levels in the CSF.

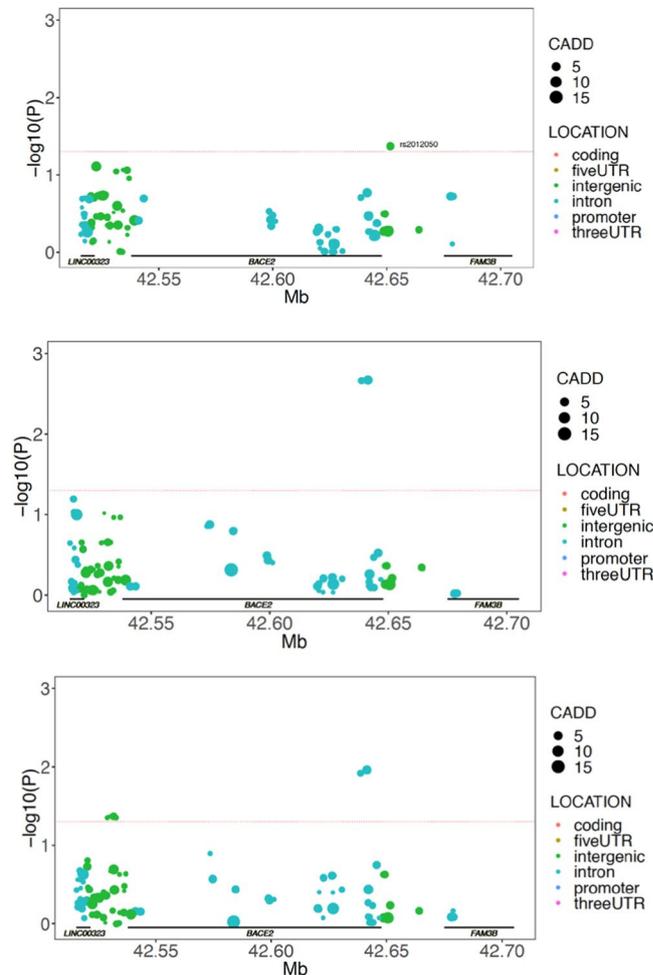


**Figure 2.** CSF A $\beta$ 1-42 levels are associated with *BACE2* locus SNPs in AD and MCI Individuals from the ADNI cohort who are APOE  $\epsilon$ 4 non-carriers. *BACE2* association with A $\beta$ 1-42 CSF levels in ADNI cohort AD, MCI, and control NA individuals including APOE carrier status, the first two population structure vectors, and sex as interaction terms. In all panels, uncorrected SNP P-values are plotted as colored circles based on their location within the genetic locus, the size of the circle is representative of the CADD score for the SNP. A horizontal red dashed line represents nominal significance of  $p = 0.05$ . (Top panel) *BACE2* interaction p-values in ADNI AD individuals. (Middle) *BACE2* interaction p-values in ADNI MCI individuals. (Bottom) *BACE2* interaction p-values in ADNI control NA individuals.

- SNPs within *BACE2* are associated with *BACE2* mRNA levels in control individuals via eQTL analysis and this effect is modified by APOE status.
- Overexpression of *BACE2* in a cell line that naturally produces A $\beta$  peptides results in lower amounts of A $\beta$ 1-40 and A $\beta$ 1-42.
- Considered together, these data suggest that genetic variation at the *BACE2* locus is associated with AD risk and that it may be functionally linked to altered A $\beta$  processing.

## Results and Discussion

We initiated our study in a cohort of clinically characterized and neuropathologically verified AD cases and controls ( $n = 1599$ ; 1014 cases, 585 controls) known as “TGenII”. This cohort is advantageous for genetic study over living, clinically diagnosed AD cohorts due to the absence of the mis-assignment of cases and controls. We examined the *BACE2* locus in three steps. First, we examined the association at the locus using a simple logistic regression case:control analysis examining genotyped and imputed SNPs (Fig. 1, top panel). The SNP rs8134992 was identified as the most significant with an uncorrected p-value of 0.00136. Next, we examined this association after including APOE carrier status (e.g. 0, 1, or 2 APOE  $\epsilon$ 4 alleles) as an interaction term in the additive logistic regression model (Fig. 1, middle panel). This yielded a significant association at rs2898441 with an uncorrected p-value of 0.0267. In a final analysis of this association we included the first two population structure co-variables and sex as interaction terms in addition to APOE carrier status (Fig. 1, bottom panel). This resulted in rs9978431 as the SNP with the most significant p-value at 0.0249. All of the data for these comparisons can be found in Supplemental Tables 1–3. The interaction with APOE status was further analyzed by performing

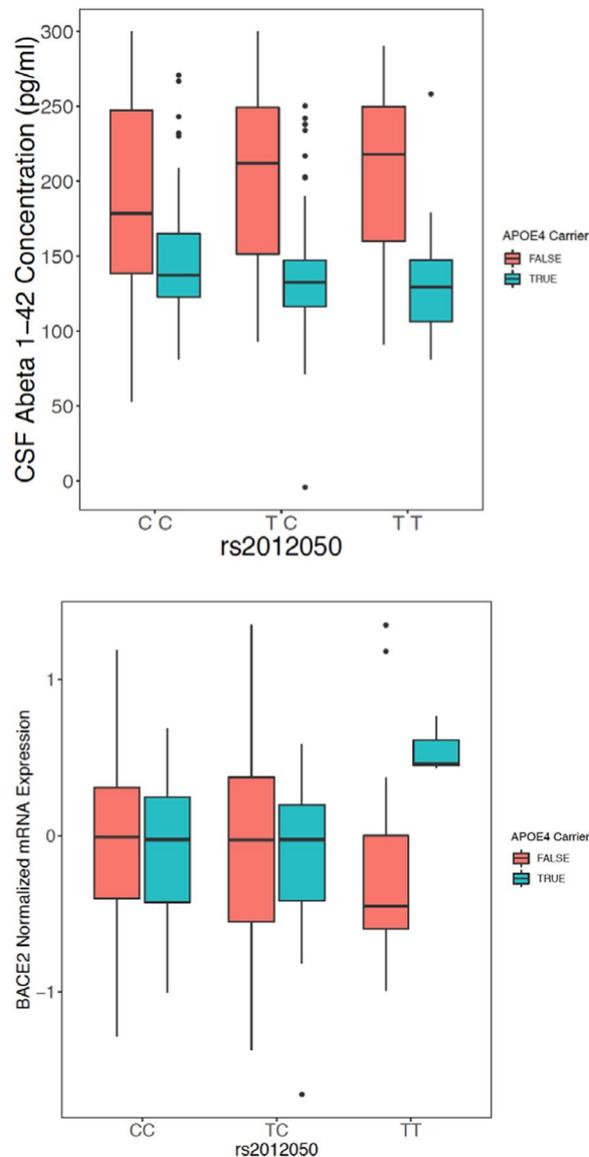


**Figure 3.** BACE2 locus SNPs are associated with BACE2 gene expression levels in cortex samples from control donors. eQTL p-values are indicated for BACE2 SNPs across three different array probes for BACE2 expression. In all panels, uncorrected SNP P-values are plotted as colored circles based on their location within the genetic locus, the size of the circle is representative of the CADD score for the SNP. A horizontal red dashed line represents nominal significance of  $p = 0.05$ . (Top panel) BACE2 eQTL p-values in control individuals for probe GI21040361-A. (Middle) BACE2 eQTL p-values in control individuals for probe ILMN1669323. (Bottom) BACE2 eQTL p-values in control individuals for probe ILMN2326712.

a post-hoc examination of the association signal across the locus by splitting the cohort into APOE carriers and non-carriers. This demonstrated that the association within *BACE2* was only significant in the APOE  $\epsilon 4$  non-carriers (Supplemental Fig. 1). In an attempt to relate potential functional impact to the SNPs we utilized CADD scoring<sup>25</sup>. There were no SNPs with CADD scores over 15 suggesting that the associated SNPs do not have direct impact at the protein function level.

These association results, in a well characterized neuropathology-confirmed cohort, suggest that there are common SNPs within the *BACE2* locus that alter AD risk preferentially in individuals who do not carry the APOE  $\epsilon 4$  genetic risk for the disease. This is important to note due to the role of *BACE2* in amyloid processing. Due to this finding, we next hypothesized that genetic variation at *BACE2* would be associated with A $\beta$ 1-42 biomarker levels due to its ability to degrade the peptide or via its theta-secretase actions which could lead to lowered A $\beta$ 1-42 production by shunting APP metabolism down a non-amyloidogenic path.

Using the ADNI cohort and a linear regression analysis incorporating the first two population structure co-variates, sex, and APOE carrier status as interaction terms, we demonstrated an association with CSF A $\beta$  1-42 levels that was significant in the MCI and AD individuals (Fig. 2). The strongest associated SNPs (rs2837994 in AD at  $p = 0.000183$  and rs9980146 in MCI at  $p = 0.0207$ , both uncorrected) were located in the same region as defined above for the *BACE2* genetic association. Of note, there was no significant association in the NA control individuals in ADNI suggesting that this locus may be preferentially relevant for CSF A $\beta$ 1-42 biomarker levels during disease processes. As before, we utilized a post-hoc analysis to identify which APOE group was associated with the CSF A $\beta$ 1-42 changes and found that the association was again only significant in APOE  $\epsilon 4$  non-carriers. This is in agreement with the genetic associations we detected in the first phase of our study. The full results of this analysis can be found in Supplemental Tables 4–6.

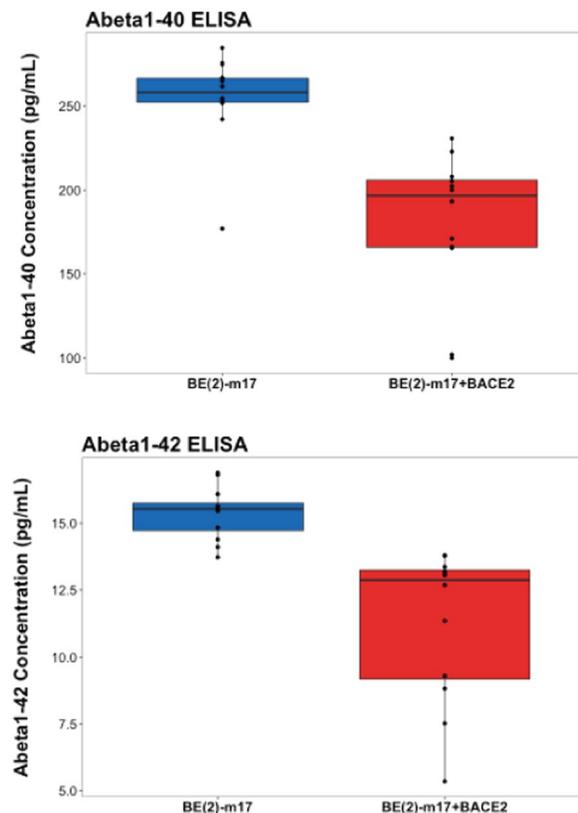


**Figure 4.** SNP rs2012050 is associated with altered BACE2 gene expression and A $\beta$ 1-42 CSF levels. The relationship between rs2012050 genotype (x-axis) and either A $\beta$ 1-42 CSF levels (Top panel) or BACE2 gene expression (Bottom) is shown. APOE carrier status is indicated by the color of the boxplot; red = APOE  $\epsilon$ 4 non-carriers, blue = APOE  $\epsilon$ 4 carriers. Note the higher levels of A $\beta$ 1-42 in the CSF from rs2012050 T/T individuals who are APOE  $\epsilon$ 4 non-carriers as well as a lower expression of the BACE2 gene in individuals with the same genetic characteristics.

The work in the ADNI cohort further supports the association at the *BACE2* locus and its specificity in APOE  $\epsilon$ 4 non-carriers. No variants identified during the CADD scoring approach clearly have the potential to alter protein function. To that end, we predicted that BACE2 gene expression levels may be altered by the associated variants; thus, in turn, altered levels of BACE2 mRNA and protein would lead to an alteration in A $\beta$ 1-42.

To examine this we utilized an expression quantitative trait locus (eQTL) analysis approach to identify variants within the *BACE2* locus that may be associated with *BACE2* RNA levels. We utilized array-based gene expression data from the TGenII cohort cortical brain tissue samples from control individuals only (to avoid any confounding influence of disease on gene expression). Three probes that are specific to BACE2 were examined and all three demonstrated an eQTL association that was modified by APOE carrier status (Fig. 3). Of particular note was SNP rs2012050 ( $p = 0.0425$ , uncorrected) which was also identified as significant in the ADNI MCI cohort CSF A $\beta$ 1-42 analysis. APOE  $\epsilon$ 4 non-carriers of the minor allele of rs2012050 demonstrate lower BACE2 gene expression and also higher CSF A $\beta$ 1-42 levels (Fig. 4). This is in agreement with the potential actions of BACE2 on A $\beta$ 1-42. All of the eQTL data can be found in Supplemental Tables 7–9.

Next we overexpressed BACE2 in a cell line that we showed previously expresses the necessary genetic components and produces A $\beta$  fragments *in vitro*<sup>26</sup>. Notably, this cell line expresses BACE2 at very low levels (less than 0.5 FPKM based on RNA sequencing data, not shown). We overexpressed BACE2 in this cell line and demonstrated that



**Figure 5.** BACE2 overexpression results in lower levels of A $\beta$  1-40 and 1-42 peptides in cell culture supernatant from BE(2)-m17 cells. AB1-40 and 1-42 levels were measured in the cell culture supernatant from wild-type (blue box plots) or cells with BACE2 overexpression (red box plots).

this resulted in lower amounts of A $\beta$ 1-40 and 1-42 fragments in the cell culture media (Fig. 5). This final experiment confirmed other reports in the literature demonstrating that BACE2 can degrade A $\beta$  peptides<sup>6</sup>. Further, these data suggest that the alteration of BACE2 expression may result in the CSF level alterations we observed.

Collectively, this work supports previous analyses suggesting that genetic variability in BACE2 alters mRNA expression and plays a role in AD. Understanding the precise role of BACE2 in APP/A $\beta$  metabolism is clearly important because drugs designed as BACE1 inhibitors also demonstrate cross-reactivity with BACE2. The effect of BACE2 inhibition on APP processing would likely be anti-amyloidogenic since it would lead to less A $\beta$  being formed<sup>8</sup>, but inhibition of BACE2 may also reduce its cleavage of A $\beta$ . However, it is noteworthy that Sun *et al.*<sup>5</sup> suggested increasing BACE2 activity was a valid approach for AD therapy given its high propensity to degrade A $\beta$ . This suggests that inhibition of BACE2 is not a desirable approach. A further complication, when trying to predict the effects of BACE2 inhibition on disease etiology, is that BACE2 expression is high in both neurons and glial cells. This means that with the gliosis inherent in the AD disease process, BACE2 expression may become more pronounced as the disease progresses. Indeed, it is possible that such a BACE2 effect on pathogenesis during AD disease progression may underlie the APOE effect we report herein. Importantly, our data suggest that BACE2 has important roles in APP/A $\beta$  metabolism. Indeed, this has wide reaching implications 1) for assessment of BACE inhibition in AD treatment, and 2) in the interpretation of AD rodent models where the Swedish (“BACE1 preferring”) mutation is used to model the disease.

### Data Availability

All SNP genotypes and CSF biomarker levels are available openly via the respective study sites or through the ADGC.

### References

- Selkoe, D. J. & Hardy, J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol Med* **8**, 595–608, <https://doi.org/10.15252/emmm.201606210> (2016).
- Thinakaran, G. & Koo, E. H. Amyloid precursor protein trafficking, processing, and function. *J Biol Chem* **283**, 29615–29619, <https://doi.org/10.1074/jbc.R800019200> (2008).
- Kim, M. *et al.* Potential late-onset Alzheimer's disease-associated mutations in the ADAM10 gene attenuate {alpha}-secretase activity. *Hum Mol Genet* **18**, 3987–3996, <https://doi.org/10.1093/hmg/ddp323> (2009).
- Hartl, D. *et al.* A rare loss-of-function variant of ADAM17 is associated with late-onset familial Alzheimer disease. *Mol Psychiatry*. <https://doi.org/10.1038/s41380-018-0091-8> (2018).
- Sun, X., He, G. & Song, W. BACE2, as a novel APP theta-secretase, is not responsible for the pathogenesis of Alzheimer's disease in Down syndrome. *FASEB J* **20**, 1369–1376, <https://doi.org/10.1096/fj.05-5632com> (2006).

6. Abdul-Hay, S. O., Sahara, T., McBride, M., Kang, D. & Leissring, M. A. Identification of BACE2 as an avid ss-amyloid-degrading protease. *Mol Neurodegener* **7**, 46, <https://doi.org/10.1186/1750-1326-7-46> (2012).
7. Basi, G. *et al.* Antagonistic effects of beta-site amyloid precursor protein-cleaving enzymes 1 and 2 on beta-amyloid peptide production in cells. *J Biol Chem* **278**, 31512–31520, <https://doi.org/10.1074/jbc.M300169200> (2003).
8. Fluhner, R. *et al.* A non-amyloidogenic function of BACE-2 in the secretory pathway. *J Neurochem* **81**, 1011–1020 (2002).
9. Yan, R., Munzner, J. B., Shuck, M. E. & Bienkowski, M. J. BACE2 functions as an alternative alpha-secretase in cells. *J Biol Chem* **276**, 34019–34027, <https://doi.org/10.1074/jbc.M105583200> (2001).
10. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906–913, <https://doi.org/10.1038/ng2088> (2007).
11. Myllykangas, L. *et al.* Chromosome 21 BACE2 haplotype associates with Alzheimer's disease: a two-stage study. *J Neurol Sci* **236**, 17–24, <https://doi.org/10.1016/j.jns.2005.04.008> (2005).
12. Wang, Z. *et al.* BACE2, a conditional beta-secretase, contributes to Alzheimer's disease pathogenesis. *JCI Insight* **4**, <https://doi.org/10.1172/jci.insight.123431> (2019).
13. Escott-Price, V., Myers, A. J., Huentelman, M. & Hardy, J. Polygenic risk score analysis of pathologically confirmed Alzheimer disease. *Ann Neurol* **82**, 311–314, <https://doi.org/10.1002/ana.24999> (2017).
14. Corneveaux, J. J. *et al.* Association of CR1, CLU and PICALM with Alzheimer's disease in a cohort of clinically characterized and neuropathologically verified individuals. *Hum Mol Genet* **19**, 3295–3301, <https://doi.org/10.1093/hmg/ddq221> (2010).
15. Braak, H. & Braak, E. Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiol Aging* **16**, 271–278, discussion 278–284 (1995).
16. Morris, J. C. *et al.* The consortium to establish a registry for Alzheimer's disease (CERAD): I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* **39**, 1159–1165, <https://doi.org/10.1212/WNL.39.9.1159> (1989).
17. Shaw, L. M. *et al.* Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann Neurol* **65**, 403–413, <https://doi.org/10.1002/ana.21610> (2009).
18. Cockrell, J. R. & Folstein, M. F. Mini-Mental State Examination (MMSE). *Psychopharmacol Bull* **24**, 689–692 (1988).
19. Morris, J. C. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* **43**, 2412–2414 (1993).
20. McKhann, G. *et al.* Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* **34**, 939–944 (1984).
21. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* **1**, 457–470, <https://doi.org/10.1534/g3.111.001198> (2011).
22. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575, <https://doi.org/10.1086/519795> (2007).
23. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**, 1655–1664, <https://doi.org/10.1101/gr.094052.109> (2009).
24. Hou, X., Du, Y., Deng, Y., Wu, J. & Cao, G. Sleeping Beauty transposon system for genetic etiological research and gene therapy of cancers. *Cancer Biol Ther* **16**, 8–16, <https://doi.org/10.4161/15384047.2014.986944> (2015).
25. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* **46**, 310–315, <https://doi.org/10.1038/ng.2892> (2014).
26. Macias, M. P. *et al.* A cellular model of amyloid precursor protein processing and amyloid-beta peptide production. *J Neurosci Methods* **223**, 114–122, <https://doi.org/10.1016/j.jneumeth.2013.11.024> (2014).

## Acknowledgements

This manuscript is dedicated to the memory of our dear friend and colleague Jason Corneveaux who first identified the BACE2 association. His contributions to our laboratory were innumerable and significant and he continues to be sorely missed by our research team. We thank the patients and their families for their selfless donations. Many data and biomaterials were collected from several National Institute on Aging (NIA) and National Alzheimer's Coordinating Center (NACC, grant #U01 AG016976) funded sites. Amanda J. Myers, PhD (University of Miami, Department of Psychiatry) and John A. Hardy, PhD (Reta Lila Weston Institute, University College London) collected and prepared the series. Marcelle Morrison-Bogorad, PhD., Tony Phelps, PhD and Walter Kukull PhD are thanked for helping to co-ordinate this collection. The directors, pathologist and technicians involved include: National Institute on Aging: Ruth Seemann, John Hopkins Alzheimer's Disease Research Center (NIA grant # AG05146): Juan C. Troncoso, MD, Dr. Olga Pletnikova, University of California, Los Angeles (NIA grant # P50 AG16570): Harry Vinters, MD, Justine Pomakian, The Kathleen Price Bryan Brain Bank, Duke University Medical Center (NIA grant #AG05128, NINDS grant # NS39764, NIMH MH60451 also funded by Glaxo Smith Kline): Christine Hulette, MD, Director, John F. Ervin, Stanford University: Dikran Horoupian, MD, Ahmad Salehi, MD, PhD, Massachusetts Alzheimer's Disease Research Center (P50 AG005134): E. Tessa Hedley-Whyte, MD, MP Frosch, MD, Karlotta Fitch, University of Michigan (NIH grant P50-AG053760): Dr. Roger Albin, Lisa Bain, Eszter Gombosi, University of Kentucky (NIH #AG05144): William Markesbery, MD, Sonya Anderson, Mayo Clinic, Jacksonville: Dennis W. Dickson, MD, Natalie Thomas, Washington University, St Louis Alzheimer's Disease Research Center (NIH #P50AG05681): Dan McKeel, MD, John C. Morris, MD, Eugene Johnson, Jr., PhD, Virginia Buckles, PhD, Deborah Carter, University of Washington, Seattle (NIH #P50 AG05136): Thomas Montine, MD, PhD, Aimee Schantz, MEd., Boston University Alzheimer's Disease Research Center (NIH grant P30-AG13846): Ann C. McKee, Carol Kubilus Banner Sun Health Research Institute Brain Donation Program of Sun City, Arizona (NIA #P30 AG19610; Arizona Alzheimer's Disease Core Center, Arizona Department of Health Services, contract 211002, Arizona Alzheimer's Research Center; Arizona Biomedical Research Commission, contracts 4001, 0011, 05\_901 and 1001 to the Arizona Parkinson's Disease Consortium; Michael J. Fox Foundation for Parkinson's Research): Thomas G. Beach, MD, PhD, Lucia I. Sue, Geidy E. Serrano; Emory University: Bruce H. Wainer, MD, PhD, Marla Gearing, PhD, University of Texas, Southwestern Medical School: Charles L. White, III, M.D., Roger Rosenberg, Marilyn Howell, Joan Reisch, Rush University Medical Center, Rush Alzheimer's Disease Center (NIH #AG10161): David A. Bennett, M.D. Julie A. Schneider, MD, MS, Karen Skish, MS, PA (ASCP)MT, Wayne T Longman, University of Miami Brain Endowment Bank (supported in part by HHSN-271-2013-00030C and the McGowan Endowment): Deborah C. Mash, MD, Margaret J Basile, Mitsuko Tanaka, Oregon Health & Science University: Randy Wotljer, PhD. Additional tissues include samples from the following sites: Newcastle Brain Tissue Resource (funding via the Medical Research Council, local NHS trusts and Newcastle University): C.M. Morris, MD, Ian G McKeith, Robert H Perry MRC London Brain Bank for Neurodegenerative Diseases (funding via the Medical Research Council): Simon Lovestone, M.D. PhD, Safa

Al-Sarraj, MD, Claire Troakes, The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend Werk, International Parkinson Fonds, Internationale Stichting Alzheimer Onderzoek): Inge Huitinga, MD, Marleen Rademaker, Michiel Kooreman, Institut de Neuropatologia, Servei Anatomia Patologica, Universitat de Barcelona: Isidre Ferrer, MD, PhD, Susana Casas Boluda. The authors acknowledge funding support from NIH-NIA grant R01-AG041232 (Multiple-PI; to Matt Huentelman and Amanda Myers) and the State of Arizona DHS (Arizona Alzheimer's Consortium, PI; Eric Reiman). The neuropathology confirmed series collection was funded by NIH EUREKA grant R01-AG034504 to AJM as well as NIH intramural funds (to JH and AJM). This work was also supported by the UK Dementia Research Institute which receives its funding from DRI Ltd, funded by the UK Medical Research Council, Alzheimer's Society and Alzheimer's Research UK. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

### Author Contributions

Matt Huentelman: conceived of the study, generated data, analyzed data, wrote the manuscript, and procured funding for the work. Matthew De Both: analyzed data, and wrote the manuscript. Wayne Jepsen: generated data, analyzed data, and wrote the manuscript. Ignazio S. Piras: analyzed data, and wrote the manuscript. Joshua S. Talboom: analyzed data and wrote the manuscript. Mari Willeman: analyzed data and wrote the manuscript. Eric M. Reiman: conceived of the study and procured funding for the work. John Hardy: conceived of the study, wrote the manuscript, and procured funding for the work. Amanda J. Myers: conceived of the study, generated data, analyzed data, wrote the manuscript, and procured funding for the work.

### Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-45896-4>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019