HHS Public Access

Author manuscript

Hum Mutat. Author manuscript; available in PMC 2022 July 01.

Published in final edited form as:

Hum Mutat. 2022 January ; 43(1): 16–29. doi:10.1002/humu.24289.

A family study implicates *GBE1* in the etiology of autism spectrum disorder

Miriam Fanjul-Fernández[1,2], **Natasha J Brown**[#,3,4,5], **Peter Hickey**[#,6,7], **Peter Diakumis**[#,8], **Haloom Rafehi**[7,9], **Kiymet Bozaoglu**[2,10], **Cherie C Green**[11,12], **Audrey Rattray**[11], **Savannah Young**[10], **Dana Alhuzaimi**[10], **Hayley S Mountford**[13], **Greta Gillies**[10], **Vesna Lukic**[9], **Tanya Vick**[5], **Keri Finlay**[5], **Bradley P Coe**[14], **Evan E Eichler**[14,15], **Martin B Delatycki**[1,2,10], **Sarah J Wilson**[11,16,17], **Melanie Bahlo**[6,7], **Ingrid E Scheffer**[11,17,18,19], **Paul J Lockhart**[*,2,10]

[1]Victorian Clinical Genetics Services, Parkville, Victoria, Australia.

[2]Department of Paediatrics, The University of Melbourne, Melbourne, Victoria, Australia.

[3]Victorian Clinical Genetics Services, Murdoch Children's Research Institute Victoria, Parkville, Victoria, Australia.

[4]Royal Children's Hospital Department of Paediatrics, The University of Melbourne, Melbourne, Victoria, Australia.

[5]Barwon Health, Geelong, Victoria, Australia.

[6]The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria, Australia.

[7]Department of Medical Biology, University of Melbourne, Melbourne, Victoria, Australia.

[8]University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer, Melbourne, Victoria, Australia.

[9]Population Health and Immunity Division, The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria, Australia.

[*]Address for correspondence: Paul J Lockhart, Bruce Lefroy Centre for Genetic Health Research, Murdoch Children's Research Institute, Flemington Road, Parkville, Victoria, Australia. paul.lockhart@mcri.edu.au, Phone: +61383416322, Fax: +61383416390.
[#]Contributed equally

Conflict of Interest: E.E.E. is on the scientific advisory board (SAB) of DNAnexus, Inc. The Department of Molecular and Human Genetics at Baylor College of Medicine receives revenue from clinical genetic testing conducted at Baylor Genetics. The other authors declare that they have no competing interests.

Web Resources:
1000 Genomes: https://www.internationalgenome.org/
ESP: https://evs.gs.washington.edu/EVS/
gnomAD: https://gnomad.broadinstitute.org/
KEGG: https://www.genome.jp/kegg/
DAVID: https://david.ncifcrf.gov/
Allen Human Brain Atlas: https://portal.brain-map.org/
SFARI: https://gene.sfari.org/
AutDB: http://autism.mindspec.org/autdb/Welcome.do
UK10K: https://www.uk10k.org/
UCSC: https://genome.ucsc.edu
Autism Genome Project (AGP): https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000267.v5.p2
OMIM: https://www.omim.org/
BrainSpan: http://www.brainspan.org/

[10]Bruce Lefroy Centre for Genetic Health Research, Murdoch Children's Research Institute, Melbourne, Victoria, Australia.

[11]Department of Medicine, University of Melbourne, Austin Health, Melbourne, Victoria, Australia.

[12]Department of Psychology and Counselling, School of Psychology and Public Health, La Trobe University, Melbourne, Victoria, Australia.

[13]Department of Biological and Medical Sciences, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, UK.

[14]Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington, USA.

[15]Howard Hughes Medical Institute, University of Washington School of Medicine, Seattle, Washington, USA.

[16]Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Victoria, Australia.

[17]Florey Institute, Melbourne, Victoria, Australia.

[18]Department of Paediatrics, The University of Melbourne, Royal Children's Hospital, Melbourne, Victoria, Australia.

[19]Murdoch Children's Research Institute, Melbourne, Victoria, Australia.

Abstract

Autism spectrum disorders (ASD) are neurodevelopmental disorders with an estimated heritability of >60%. Family-based genetic studies of ASD have generally focused on multiple small kindreds, searching for *de novo* variants of major effect. We hypothesised that molecular genetic analysis of large multiplex families would enable the identification of variants of milder effect. We studied a large multigenerational family of European ancestry with multiple family members affected with ASD or the broader autism phenotype (BAP). We identified a rare heterozygous variant in the gene encoding 1,4-alpha-glucan branching enzyme 1 (*GBE1*) that was present in seven of seven individuals with ASD, nine of ten individuals with the BAP and none of four tested unaffected individuals. We genotyped a community-acquired cohort of 389 individuals with ASD and identified three additional probands. Cascade analysis demonstrated the variant was present in eleven of thirteen individuals with familial ASD/BAP and neither of the two tested unaffected individuals in these three families, also of European ancestry. The variant was not enriched in the combined UK10K ASD cohorts of European ancestry but heterozygous *GBE1* deletion was overrepresented in large ASD cohorts, collectively suggesting an association between *GBE1* and ASD.

Keywords

Autism Spectrum Disorder; Broader Autism Phenotype; genetics; linkage; whole exome sequencing; glycogen branching enzyme

BACKGROUND

Autism spectrum disorder (ASD) is a life-long neurodevelopmental disorder that affects approximately 1 in 70 children, with four males affected for every female, and has a spectrum of phenotypic severity (Chen, Penagarikano, Belgard, Swarup, & Geschwind, 2015). Individuals with ASD exhibit impairments in social communication and a combination of narrow and intense interests, repetitive or stereotyped behaviours and sensory interests or aversions. ASD has an estimated heritability of 60–93%, with concordance rates for autism inclusive of broader phenotypes of ~90% in monozygotic twins compared with ~60% in dizygotic twins and with little or no influence of shared familial environment (Sandin et al., 2017; Tick, Bolton, Happe, Rutter, & Rijsdijk, 2016). Sibling risk for an affected proband is estimated at ~20% (Georgiades et al., 2013). In families with two affected siblings, the chance of having a third affected child is in the order of 50% (Ozonoff et al., 2011).

Genetic and epidemiological data have suggested that personality traits similar but less severe than ASD are also heritable (Freitag, 2007). Individuals with this broader autism phenotype (BAP), characterised by subtle behavioural features that reflect the core domains of ASD, are distinguished from ASD by lack of impairment in daily function (Piven, Palmer, Jacobi, Childress, & Arndt, 1997). Prevalence of the BAP in the general population is unknown, however subthreshold ASD traits are continuously distributed in the general population (Constantino & Todd, 2005). Notably, BAP traits are more prevalent in first-degree relatives of individuals with ASD than in other groups (Bailey, Palferman, Heavey, & Le Couteur, 1998; Losh, Childress, Lam, & Piven, 2008) and up to 30% of relatives of children with ASD have the BAP (Ozonoff et al., 2014; Sasson et al., 2013). Therefore, genetic studies of multiplex ASD families that take into account BAP status in relatives could provide valuable insights into the genetic basis of this disorder (Pisula & Ziegart-Sadowska, 2015).

ASD shows significant genetic complexity and heterogeneity. Known causes of ASD include chromosomal abnormalities, single gene disorders, such as Rett syndrome and fragile X syndrome, and rare metabolic conditions (de la Torre-Ubieta, Won, Stein, & Geschwind, 2016). Dysregulation of the proteins encoded by many of these genes was found to disrupt synaptic transmission, transcriptional regulation and RNA processing functions (Yuen et al., 2017). Genetic studies have highlighted the importance of *de novo* mutations in ASD, with hundreds of genes associated with sporadic ASD. *De novo* copy number variants (CNVs) are found in up to 10% of individuals with sporadic ASD, and 3% of multiplex families (Robinson et al., 2016; Ronemus, Iossifov, Levy, & Wigler, 2014; Sebat et al., 2007). *De novo* single nucleotide variants (SNVs) in coding regions of genes are thought to underlie 5–20% of ASD (Geisheker et al., 2017; Iossifov et al., 2014). However, in total, *de novo* genetic variants only account for approximately 30% of ASD, with no single gene accounting for more than ~2% of the ASD population. While large cohorts have been assembled to confirm the hundreds of single genes and genomic regions that have been implicated in ASD and neurodevelopmental disorders more broadly (SPARK Consortium, 2018), additional genetic mechanisms, such as multifactorial inheritance models, may also contribute (Guo et al., 2018). A recent study has further demonstrated the complex genetic

architecture behind ASD, with 10% of affected individuals carrying a combination of pathogenic SNV and/or CNV alterations (Guo et al., 2019).

Family studies of ASD have previously utilised a recessive inheritance model, which is only likely to apply in rare and consanguineous families and thus only able to explain ~1% of affected individuals (Lim et al., 2013; Yu et al., 2013). In contrast, multigenerational families with multiple affected individuals with ASD or the BAP are consistent with a dominant, variably expressed inheritance model, whereby a gene of major effect predisposes individuals to ASD/BAP. Multiplex family studies utilising a dominant model have traditionally combined many small families, each comprising two or three individuals with ASD (Cukier et al., 2014; Toma et al., 2014; Yuen et al., 2015). It is likely that these studies have been confounded by genetic and phenotypic heterogeneity, incomplete penetrance and phenocopies. A large multiplex family in which numerous individuals have ASD or the BAP is more likely to be genetically homogeneous than multiple small families, reducing the underlying genetic heterogeneity and facilitating the identification of a gene of major effect.

We performed detailed phenotyping of a large multiplex, multigenerational family with many affected individuals with ASD or the BAP (Trevis et al., 2020). We then performed a genetic linkage study predicated on a shared genetic basis underpinning the ASD and BAP phenotypes. We identified a rare heterozygous variant in the gene encoding 1,4-alpha-glucan branching enzyme 1 (*GBE1*) within the linkage peak and subsequently identified the same variant in three additional unrelated families ascertained from a community-acquired cohort with ASD.

METHODS

Editorial Policies and Ethical Considerations

Participants, or parents or legal guardians in the case of minors or those with intellectual disability, provided informed consent. The study was approved through the Human Research Ethics Committees of Barwon Health (HREC 02/34, HREC 04/57) and the Royal Children's Hospital (HREC 25043), Victoria, Australia.

Phenotyping

Family B were recruited, the pedigree was constructed and deep phenotyping of ASD and the BAP was performed. A detailed description of the phenotyping methodology for determining ASD and the BAP has been previously published (Trevis et al., 2020). Family members who were either not assessed or not recruited to the study are not included in the pedigree, which is scrambled to preserve anonymity (Figure 1). In parallel a cohort of 389 probands with sporadic ASD was recruited from the Barwon region of Victoria, Australia, following referral from clinicians and community diagnostic services, to our research study. All diagnoses of ASD were made using the DSM-IV or DSM-5 criteria. Probands had molecular karyotype analysis, in addition to basic metabolic and fragile X DNA testing and were only included in the cohort if the results of these investigations did not identify an etiology for their ASD.

Linkage analysis

For all participants, DNA was extracted from whole blood or saliva and high density single nucleotide polymorphism (SNP) arrays were used to genotype samples (Illumina 660W Quad or HumanCore Exome BeadChip). Copy number analysis was performed using the R package Aroma Affymetrix (Bengtsson, Irizarry, Carvalho, & Speed, 2008; Bengtsson, Wirapati, & Speed, 2009) and PennCNV (Wang et al., 2007). For the CNV analysis, the control dataset was obtained from the same laboratory and generated simultaneously to reduce errors.

In order to reduce confounding factors, we excluded relatives descended from individual III-1 (Figure 1) from the linkage analysis, due to their partner, III-2, with the BAP, possibly contributing an alternate ASD susceptibility allele(s) to subsequent generations (Figure 1, Bilineal Sub-Family B contained in box with dashed line). The remainder of the pedigree was called Core Family B (Figure 1). Twenty-one individuals were genotyped using the 660W Quad or HumanCore Exome BeadChip (Illumina, San Diego, California, USA) and genotypes were called using Illumina software. Markers incompatible with Mendelian inheritance were removed using LINKDATAGEN (Bahlo & Bromhead, 2009) and parametric analysis was initially performed using MORGAN (George, Wijsman, & Thompson, 2005). Exact linkage analysis was subsequently performed using MERLIN and a dominant disease model with disease allele frequency 0.0001, incomplete penetrance (95%) and a 10% phenocopy rate. Inbreeding coefficients for genotyped family members were estimated independently by FEstim (Leutenegger et al., 2003). Relatedness checks to verify pedigree relationships were performed using PLINK (Purcell et al., 2007) and in-house software.

Whole Exome capture and sequencing

Peripheral blood or saliva derived gDNA was isolated from participants using standard techniques. Fourteen individuals from Core Family B (Figure 1), including 10 affected and three unaffected individuals, plus one non-phenotyped, presumed obligate carrier (II-2) underwent exome capture (SureSelect XT Human All Exon v5 +UT, Agilent, Santa Clara, California, USA). Paired-end 100-bp reads were sequenced on a HiSeq2000 platform, with a mean target coverage of 50x. Reads were aligned to the reference genome (UCSC hg19) with Novoalign and local re-alignment was performed with Genome Analysis Toolkit (GATK). Single nucleotide variants and small indels were called simultaneously across all sample BAM files with GATK-Hapotype Caller and annotated with ANNOVAR (H. Li, 2011). Variants were filtered according to the following criteria: i) MAF ≤5% from public databases (1000G, ESP6500, ExAC65000/gnomAD), ii) genomic region (exonic/splicing and UTRs) and iii) mutation type (missense, nonsense, frameshift, indel). Whole genome sequencing was performed by Illumina FastTrack Services in two severely affected cases (IV-11 and V-10, Figure 1) with a mean target coverage of 30x. Sequencing data was processed as above and exonic variants were called using GATK-Haplotype Caller. Variants of interest were validated in siblings and genotyped in extended family members by Sanger sequencing.

Variant analysis

We prioritized variants shared by ≥ 8 of 10 affected individuals with ASD or the BAP, not observed in >1 of 3 unaffected individuals and present in the obligate carrier. Final prioritization included consideration of gene function (KEGG, Ingenuity, DAVID), brain expression profile (Allen Human Brain Atlas), overlap with known/putative ASD genes (SFARI, AutDB) and other neurodevelopmental disorders (Neurocarta, Genotator), *in silico* protein-damaging predictions (Polyphen2, SIFT, CADD) and conservation status (GERP+ and PhyloP).

ASD cohort genotyping

The *GBE1* variant c.176T>C was assayed in the sporadic cohort using the PCR-restriction fragment length polymorphism method (RFLP). An approximately 300 bp gDNA fragment containing *GBE1* exon 2 was amplified by PCR and subsequently digested with BsrI restriction enzyme (NEB) at 60 °C for 1 hour. Digestion products were analysed by agarose gel electrophoresis and samples positive for the c.176T>C variant were validated by Sanger sequencing.

UK10K analysis

VCF files limited to genes that function in the glycogen pathway (*AGL, G6PC, GAA, GBE1, GYS1, GYS1, GYS2, PFKM, PGAM2, PGM1, PHKA1, PHKA2, PHKB, PHKG2, PYGL, PYGM, SLC37A4*) from UK10K datasets (SKUSE n=329, BIONED n=71, MGAS n=94, IMGSAC n=110) were downloaded and processed by an in-house pipeline (described in exome data analysis).

Variant frequencies for European controls (n=7935) were extracted from the ExAC dataset following removal of individuals with psychiatric disorders. Rare variants (MAF<0.05) in the glycogen pathway gene set that were predicted to be protein-damaging by at least one of SIFT, Polyphen2 or CADD *in silico* tools were assessed. Differences in variant frequencies between cases and controls were tested by Fisher's exact test with adjustment for multiple testing.

CNV analysis

An initial CNV analysis of *GBE1* was performed in a cohort of 29,085 children with developmental delay and 19,584 healthy controls as previously described (Coe et al., 2014). *GBE1* was also examined with the UCSC Genome browser investigating the ClinGen CNV track and the Database of Genomic Variation (DGV) track, examining structural variations. CNV analysis was also performed with PennCNV (Wang et al., 2007) on data from 2,611 families (total of 7,880 individuals), all genotyped with the Illumina 1M SNP chip, studied as part of the Autism Genome Project (AGP) Consortium (Stage I of this study, dbGaP Study Accession: phs000267.v4.p2). These samples are part of the dbGaP [1] project No.9078 ('Gene discovery in multiplex large families with autism spectrum disorders').

Molecular studies

Mammalian expression plasmid constructs encoding wildtype and variant *GBE1* were manufactured by Integrated DNA Technologies, USA and integrity was confirmed by DNA sequencing. Transient transfection and western blot analysis of HEK293T cells was performed using FuGENE (Promega, USA) in 6-well plates essentially as previously described (Taylor, Brody, & Lockhart, 2012). Primary antibodies were incubated overnight at 4°C. GBE1 was detected using rabbit anti-GBE1 antibody (1:1000, Invitrogen, USA, cat# PA5–26515) and β-Actin was used as a loading control (mouse anti-β-Actin, 1:10,000, Sigma Aldrich, Germany, cat#A2228). Antibody binding was revealed using peroxidase conjugated secondary antibodies goat anti-rabbit (1:5,000 dilution, 111-035-144, Jackson) or sheep anti-mouse (1:5,000 dilution, 515-035-003, Jackson) with enhanced luminol-based chemiluminescent (ECL) Western Blotting Substrate (Clarity ELC, Bio-Rad Laboratories). Chemiluminescence was detected using an Amersham Imager 680 (Cytiva Life Sciences, USA) and quantitated with the associated Image Quant TL software. Immunoblot analysis was performed to determine steady-state amounts of recombinant GBE1 under basal conditions and after inhibition of the ubiquitin proteasome system (UPS), the major intracellular protein degradation pathway for unstable or damaged proteins. Cells were cultured overnight in basal media or media supplemented with 10 mM MG-132, which we have previously shown is sufficient to inhibit the UPS without causing significant cellular toxicity in HEK293 cells (Lockhart et al., 2004).

RESULTS

Family B included 37 phenotyped individuals, 11 of whom had ASD, 15 the BAP, and 11 were unaffected (extended pedigree not shown to preserve anonymity). Thirty-three phenotyped individuals were directly related to II-2 and four phenotyped individuals were “married-in” (defined as non-blood relatives of II-2). One individual with the BAP (III-1, Figure 1) has a “married-in” partner (III-2) who also has the BAP, and therefore their three children with the BAP and two grandchildren with ASD may have bi-lineal inheritance of ASD susceptibility alleles (Bilineal Subfamily B). ‘Core Family B’, a subset of the entire family, included 28 phenotyped individuals (20 affected and 8 unaffected) where the segregation of ASD traits followed likely unilineal inheritance (Figure 1, Core Family B). We hypothesised that all or most (allowing for phenocopies) of Core Family B shared a genetic variant consistent with autosomal dominant inheritance and therefore we focused on identifying the genetic variant underlying ASD and the BAP in this branch of Family B. In addition to the proband, several affected individuals in Core Family B had a molecular karyotype, fragile X DNA and basic metabolic testing performed, and no abnormalities were identified.

Linkage analysis identifies a shared haplotype in chromosome 3

We performed non-parametric linkage analysis in 21 individuals from Core Family B using genotypes derived from high-density SNP arrays. Phenotypes were collapsed into affected (ASD/BAP), unaffected and unknown. A dominant disease model with reduced penetrance (95%), and incorporating a 10% phenocopy rate, was assumed. A maximum parametric LOD score of 2.9 was obtained for a double peak in chromosome 3 (Figure

2A). No other regions achieved a LOD score greater than 1. Haplotype analysis defined two adjacent regions of 7Mb and 14Mb at 3p14.1–3p13 (chr3:65,611,042–72,781,075; Region 1) and 3p13–3p11.2 (chr3:73,024,350–87,370,782) (Region 2), respectively (Figure 2B). Both regions were inferred to segregate with 6/6 individuals with ASD, 8/9 individuals with BAP, 0/2 unaffected individuals and 1/1 non-phenotyped obligate carrier (II-2). To test the robustness of collapsed phenotypes, we performed an additional linkage analysis with phenotypes of affected (ASD), unaffected and unknown. All members of the pedigree with BAP were recoded to unknown affectedness status. This produced the same dominant linkage peak on chromosome 3 (LOD of 2.62) and two additional minor peaks on chromosomes 10 and 17 (LOD ~1.4, Supp. Figure S1), suggesting that including BAP has enhanced the linkage results by increasing the evidence for the chromosome 3 and reducing the evidence for the chromosome 10 and 17 peaks.

Exome sequencing and variant analysis

We performed exome sequencing in six individuals with ASD, four with BAP, three unaffected individuals and one non-phenotyped obligate carrier of Core Family B. We achieved >50x mean sequence coverage of targeted exonic regions. Over seven million raw (unfiltered) variants were analysed using two complementary methods: interrogation of the candidate linkage region and genome-wide variant analysis.

Linkage-based filtering: Seven heterozygous exonic variants with a minor allele frequency ≤ 0.05 were identified within the two linkage regions. Five were private variants, carried by only one affected individual. The remaining two were missense variants in the genes encoding 1,4-alpha-glucanbranching enzyme (*GBE1*, [NM_000158.4: c.176T>C, p.Ile59Thr]; rs28763904) and endogenous Bornavirus-like nucleoprotein 2 (*EBLN2*, [NM_018029.4: c.492T>G, pTyr164*]; rs2231925). The *GBE1* variant was present in all 10 affected individuals (and the non-phenotyped obligate carrier), while the *EBLN2* variant was observed in 9/10 affected individuals. None of the three unaffected relatives carried either of these variants.

Genome-wide filtering: This approach provides an alternative method to linkage-based filtering, since phenocopies and incomplete penetrance can influence linkage signals. Variants were filtered as described in the methods and three heterozygous variants (MAF ≤ 0.01) were identified that were shared by >8 affected individuals. This included the *GBE1* c.176T>C and *EBLN2* c.492T>G variants in addition to a missense variant in the gene encoding melanoma inhibitory activity SH3 Domain ER Export Factor 2 on chromosome 14 (*MIA2*, [NM_001247989.1: c.1353T>G, p.Ile451Met; rs118177774]. This variant was identified in 10 affected individuals and was absent in the three unaffected individuals. Given that *MIA2* is predominantly expressed in the liver and is thought to encode a tumor suppressor (Xu et al., 2011), and the variant is predicted via *in silico* tools to be benign, we excluded it from further consideration.

We performed confirmation and segregation analysis of the *GBE1* and *EBLN2* variants in 22 family members in Core Family B. In total, 17 affected individuals (7x ASD, 10x BAP), four unaffected and one non-phenotyped obligate carrier (II-2, Figure 1) were tested. We

identified the *GBE1* variant in 7/7 individuals with ASD, 9/10 individuals with the BAP and the non-phenotyped obligate carrier (II-2). In addition, nine individuals with DNA available in sub-Family B were genotyped and none carried the *GBE1* variant. The *EBLN2* variant was observed in 6/7 individuals with ASD, 9/10 individuals with BAP and 0/4 unaffected individuals.

GBE1 encodes a glycogen branching enzyme and pathogenic variants cause autosomal recessive glycogen storage disease IV (GSD IV; MIM# 232500) (Magoulas & El-Hattab, 1993) and the allelic disorder, adult polyglucosan body neuropathy (APBN; MIM# 263570) (Klein, 1993). The gene does not appear to be subject to missense nor loss of function constraint (gnomAD Z = 0.27, pLI=0.00, (Lek et al., 2016)), which is expected for genes causing recessive disorders. Data from BrainSpan showed low to medium levels of *GBE1* expression in the developmental and adult brain transcriptome. Conservation analysis suggests evolutionary constraint (GERP++ 5.68) on the nucleotides spanning the c.176T>C variant and at the protein level Ile59 is highly conserved (Figure 2C), with eight of 10 *in silico* tools (Supp. Table S1) predicting pathogenicity (Kopanos et al., 2019). The c.176T>C variant is present in the European (non-Finnish) population at a frequency of 0.009 and is classified as a benign variant in the context of GSD IV and APBD. In contrast, *EBLN2* encodes a single exon gene with similarity to Borna disease virus and has not been associated with any human disease (Horie et al., 2010). Expression of *EBLN2* is very low or absent in the human brain and the gene is only present in primates. Approximately 40 termination/frameshift loss of function variants are reported throughout the gene, all classified as low confidence variants because it is a single-exon gene (LOFTEE predictor, gnomAD, (Cummings et al., 2020)). The c.492T>G variant is present in the European (non-Finnish) population at a frequency of 0.011 and is classified as likely benign based on low conservation (GERP++ −0.94) and *in silico* modelling tools (four benign vs no pathogenic predictions). To test for potential non-familial causal variants, genome-wide analysis of the six most severely affected individuals was performed but no compelling candidate variants in ASD-associated genes were identified (Supp. Table S2). Collectively, the data suggested that the *GBE1* c.176T>C variant (rs28763904) is the best candidate identified to explain the family disease model.

Identification of the *GBE1* variant in three additional families

We next explored whether the *GBE1* c.176T>C variant may contribute to the wider ASD population, initially by analysing our ASD cohort of 389 probands. We identified three probands heterozygous for the variant. Further cascade phenotyping and variant testing in these families identified the variant in 12/14 phenotyped affected individuals (Figure 3A–C) related to the three probands originally identified. In Family C, the variant was inherited from the proband's great grandmother and was present in five of the six affected individuals tested; and the variant was absent in both unaffected (assessed) individuals. In Family D, the variant was inherited from the affected mother and was present in all three siblings with ASD. In Family E, the proband inherited the *GBE1* variant from her mother, who is of unknown phenotype.

In summary, in the three additional families, the *GBE1* c.176T>C variant was observed in 6/7 individuals with ASD and 5/5 individuals with BAP, assuming that the married-in individual (Family D, I-1) has BAP caused by a different genetic/non-genetic factor. The variant was absent in both of the two unaffected, assessed individuals. No other medical history segregated with ASD in these families. There was no history of liver dysfunction, neuropathy, myopathy and cardiomyopathy in any family member to suggest a similar pattern of metabolic derangement seen in GSD IV or ABPN (Mochel et al., 2012; Paradas et al., 2014). All SNP-chipped individuals from the four families were examined for close relatedness with pairwise relatedness methods with XIBD (Henden, Wakeham, & Bahlo, 2016) but apart from verifying the very close relationships as documented by pedigrees we were unable to identify any evidence of distant relatedness (>8 generations) between any pair of SNP-chipped individuals (data not shown).

Molecular analysis of the *GBE1* variant

Protein modelling of the GBE1 p.Ile59Thr variant performed using mCSM (Pires, Ascher, & Blundell, 2014) suggested the variant was highly destabilising. To test this, we generated recombinant wildtype (WT) and variant (p.Ile59Thr) and analysed steady state protein levels in transiently transfected HEK293T cells. Two additional recombinant GBE1 proteins were analysed; p.Leu224Pro and p.Tyr329Ser, the expression of which has previously been shown to generate an unstable protein with reduced steady state levels (Froese et al., 2015). We performed immunoblot analysis to determine steady-state levels of recombinant GBE1 under basal conditions and after inhibition of the ubiquitin proteasome system (UPS), the major intracellular protein degradation pathway for short-lived or damaged proteins. Visual inspection of the blots suggested that all three variant forms of GBE1 were present at a reduced steady state level compared to the wildtype construct under basal conditions but this could be partially rescued following treatment with MG-132 (Figure 4). To quantitate this observation, the western blots were probed with an directed against β-Actin and the relative ratio of GBE1:β-Actin was determined. The analysis of two independent transfection experiments, performed in duplicate, demonstrated a significant increase in the steady state levels of p.Ile59Thr (3.7x ± 1.1, p = 0.02) after treatment with MG-132 compared to the vehicle. Similarly, increases were observed for p.Leu224Pro (4.4x ± 2.4, p = 0.06) and p.Tyr329Ser (2.8x ± 0.7, p = 0.01).

Contribution of *GBE1* variants to ASD

To further investigate the association of *GBE1* with ASD, we interrogated a subset of the UK10K cohort based on European ethnicity. We examined exome sequencing data of 604 individuals with ASD from the UK10K project (UK10K Consortium et al., 2015) for the *GBE1* c.176T>C variant and identified it in the heterozygous state in 16 individuals. The distribution by study cohort was 5/71 BIONED, 3/110 IMGSAC, 3/94 MGAS and 5/328 SKUSE. While the variant was overrepresented in the BIONED study cohort compared to 7,935 European controls from the ExAC nonpsychiatric dataset (BIONED, 5/71 cases versus 189/7935 European controls, OR=2.96, P = 0.034), the variant was not significantly overrepresented in the combined UK10K cohorts (ASD-all 16/604 cases versus 189/7935 European controls, OR=1.11, P = 0.68). Analysis of all potentially pathogenic, rare variants

in *GBE1* and the 16 other genes that function in the glycogen pathway identified in the UK10K dataset similarly did not identify any variants overrepresented in the ASD cases.

CNV analysis of 29,085 cases with a wide range of ASD, intellectual disability and other neurodevelopmental conditions (Coe et al., 2014) identified six exon intersecting deletions of *GBE1* in affected individuals but none in 19,584 controls (Figure 5A, p= 0.045; OR 95% C.I = [0.79-Inf]). Among these were five deletions that only affected *GBE1* (p=0.076, OR 95% C.I. 0.62-Inf). A second CNV analysis was performed in 7,745 individuals with ASD from the Autism Genome Project (AGP) Consortium (dbGaP Study Accession: phs000267.v4.p2). Twelve individuals with a CNV affecting multiple exons of *GBE1* were identified (Figure 5B). Eleven of these CNVs were deletions, with one duplication of 0.26Mb. The deletions were generally small (2.7–150kb) with one deletion event being shared by six of the 12 individuals. In contrast analysis of *GBE1* in gnomAD (21,694 alleles) identified only a single deletion, which spans 12 exons and is 149kb in size (chr3:81586904–81736023, data not shown). Of the individuals from the AGP with deletions, four met the definition of ASD used by the authors, with three meeting the definition of strict autism. Since the AGP study focused on simplex families ascertained for autism, no individual in this study can be truly considered to be a control. The individual who met the spectrum definition of autism but not the strict definition, inherited the CNV from their father (one of the 12 individuals identified with a CNV).

DISCUSSION

Recent investigations of the genetics of ASD have primarily focused on the analysis of trios to identify *de novo* genetic variation, with hundreds of ASD risk loci implicated in disease pathogenesis (de la Torre-Ubieta et al., 2016; Yuen et al., 2017). However, increasing evidence suggests that there may be several patterns of inheritance resulting in ASD. Specifically, more deleterious *de novo* mutations may explain simplex families with one or occasionally more children with ASD, whilst mutations of milder effect may be transmitted within multiplex families with variable expressivity (Guo et al., 2019; Robinson et al., 2014). Indeed, several studies have suggested that the analysis of extended, well-characterised families may provide significant insights into the genetics of ASD (Allen-Brady et al., 2009; Piven et al., 2013; Robinson et al., 2014). We have employed a large multiplex family approach, which includes individuals with the BAP classified as affected, to identify ASD susceptibility variants. By applying linkage and exome sequencing analysis to a large multigenerational family, we identified a rare missense variant (c.176T>C, p.Ile59Thr) in *GBE1* that was present in 16/17 affected individuals in Core Family B, where unilineal inheritance was observed. None of the four unaffected genotyped individuals have the variant, which is predicted to be damaging by multiple *in silico* algorithms, is highly conserved and located within a putative functional N-terminal helical segment that binds to oligosaccharides (Froese et al., 2015). In the subsequent cohort analysis of 389 ASD probands, three additional affected individuals were identified with the *GBE1* variant. Cascade analysis demonstrated a further 11/12 individuals with ASD/BAP and this specific variant; 0/2 unaffected genotyped individuals carried the variant.

Although the affected founders in all four families were female, there was no evidence of sex bias overall as equal numbers of affected males and affected females (14 each) have the variant. In a separate study, thirty individuals with the *GBE1* variant from the four families had IQ testing, with an average IQ of 105 and only three individuals with ASD were found to have intellectual disability (IQ< 70) (Trevis et al., 2020). These results reflect studies suggesting that high functioning ASD may be associated with inherited variants of milder effect that segregate in multiplex families (Chang, Gilman, Chiang, Sanders, & Vitkup, 2015; Guo et al., 2018; Robinson et al., 2014).

Molecular characterisation of the p.Ile59Thr variant protein in a mammalian overexpression system supported the *in silico* predictions that the variant resulted in a destabilised protein. Steady state levels were significantly reduced compared to the WT, but this phenotype could be rescued in cells treated with a UPS inhibitor. These observations raise the possibility that the mechanism of GBE1 association with ASD in the families is reduced activity or dosage (see below). The CNV analysis of two large cohorts provided additional evidence to support the potential association of reduced *GBE1* dosage with neurodevelopmental disorders. Focal deletions affecting the coding region of *GBE1* were observed in five of 29,085 cases with a diagnosis of ASD or, more broadly, intellectual disability and other neurodevelopmental conditions but in none of 19,584 controls. Similarly, focal deletions were identified in 11 of 7,745 individuals in the AGP study but only one of 10,738 individuals in gnomAD, a dataset that aims to exclude individuals with severe pediatric disease. These observations are not consistent with the prediction that *GBE1* is tolerant to monoallelic loss of function variants (pLI = 0.00). One possible explanation is that individuals with heterozygous loss of function variants in the gnomAD database present subtle autistic features or the BAP and are therefore considered unaffected.

No association of *GBE1* with ASD was identified in individuals of European ethnicity in the UK10K cohort. This lack of association may in part be due to the different criteria used to determine clinical status in the different UK10K ASD cohorts. In addition, the population controls used for our analyses were not analysed by UK10K. By necessity, we used the ExAC database, extracting genotypes from European controls only (n=7935), excluding individuals with psychiatric disorders. It is likely that individuals with relatively subtle autistic features or the BAP are present within this control dataset, potentially masking any true association. This limitation in population ‘control’ databases may have contributed to the observation that no genetic variants in *GBE1* have been previously associated with ASD in sequencing-based studies.

Many studies employ a trio design and specifically test for *de novo* mutations in the affected proband, while others filter variants aggressively based on minor allele frequency in the general population (Codina-Sola et al., 2015; Toma et al., 2014). While the *GBE1* c.176T>C, p.Ile59Thr variant is predicted to be protein-damaging by multiple *in silico* prediction tools, the allele frequency (0.009 in European populations) and tolerance of the gene to variation contribute to classification as likely benign by ACMG criteria (Richards et al., 2015). However, this classification is in the context of the rare and severe Mendelian disorders of GSD IV and APBN. The authors of the ACMG guidelines acknowledge that the recommendations do not address the interpretation and reporting of complex trait alleles,

instead they were implemented to evaluate variants of monogenic Mendelian diseases and therefore are less robust for classification or susceptibility alleles contributing to complex disorders such as ASD (Richards et al., 2015).

Interestingly, previous linkage and molecular studies have indicated that *GBE1* may be an ASD susceptibility gene. A linkage peak at 3p12 was found for a family with seven individuals with ASD (Allen-Brady et al., 2009) and similarly, a 3p14–3p12 peak was found in a full-genome scan of 109 sibling pairs with ASD from 91 families (Cantor et al., 2005). *GBE1* was identified within a risk haplotype at higher frequency among individuals with ASD compared to parental controls (Casey et al., 2012) and expression of the gene was reported to be upregulated in two lymphoblast cell lines derived from sporadic children affected by ASD (Baron, Liu, Hicks, & Gregg, 2006).

GBE1 encodes a glycogen branching enzyme that catalyses the transfer of a group of glucose units from the outer end of a chain to a previous point, generating a highly ramified structure. This activity is essential to increase the solubility of glycogen within cells, preventing osmotic damage. *GBE1* is highly expressed in liver and muscle and at low to moderate levels in brain (BrainSpan atlas). Expression in the brain increases at ~21–25 weeks post conception and appears to remain stable through to adulthood, which agrees with previous studies of the expression pattern of genes associated with ASD (Li et al., 2017). Within the brain, glycogen is present mainly in astrocytes, although neurons have machinery for its synthesis (Mamczur et al., 2015). However, glycogen metabolism in neurons is tightly regulated, as accumulation of glycogen in neurons induces neuronal apoptosis (Vilchez et al., 2007). Astrocytes play a key role in glycogen metabolism in the brain, releasing glucose and/or glycogen-derived lactate, which serves as an energy source for neurons (Belanger, Allaman, & Magistretti, 2011). The inhibition of this process has been shown to interfere with the plasticity of neural networks (Mozrzymas, Szczesny, & Rakus, 2011; Newman, Korol, & Gold, 2011), which could plausibly contribute to subtle autism phenotypes.

Biallelic mutations in *GBE1* cause two distinct autosomal recessive metabolic disorders: glycogen storage disease IV (GSD IV) and the adult polyglucosan body neuropathy (APBN). Both disorders display extensive phenotypic heterogeneity but are characterised by intracellular accumulation of a poorly branched form of glycogen known as polyglucosan. GSD IV is associated with no or very low enzyme activity, and affected individuals die from liver failure *in utero* or early childhood. In contrast, individuals with APBN retain low to moderate levels of GBE1 activity and develop neurological problems in adulthood, with mild cognitive impairment in 50% of cases (Mochel et al., 2012). This suggests that a dosage effect underlies the phenotypic presentation associated with pathogenic variants in *GBE1*. We hypothesise that ASD is one potential manifestation of dysregulated GBE1 function. Interestingly, a similar mechanism has been invoked for several ASD genes, where partial loss of function is associated with ASD, while complete disruption causes severe Mendelian disease (Yu et al., 2013). While accumulation of polyglucosan is detrimental in GSD/APBN (Dainese et al., 2013; Kakhlon et al., 2013), we hypothesise that a different mechanism, the inability of neurons to obtain sufficient energy from glycogen metabolism, could be contributing to *GBE1*-mediated ASD. The association of ASD with an imbalance in energy production has been previously suggested, with carnitine deficiency proposed to

exacerbate ASD symptoms by hampering the production of energy in cells, including the brain's neurons (Celestino-Soper et al., 2012).

This study utilised linkage-discovery methodology in a family with multiple individuals diagnosed with ASD or the BAP. The underlying rationale being that very large multiplex families (> 8 affected) where ASD traits appear dominantly inherited are rare, but more genetically homogeneous and therefore amenable to gene discovery. While this approach has proven effective in other complex disorders such as epilepsy (Helbig, Scheffer, Mulley, & Berkovic, 2008), this approach has received limited attention in ASD. Although the BAP is recognised as milder, related phenotypes often observed in family members of individuals with ASD (Bolton et al., 1994), these traits are continuously distributed in the general population (Constantino & Todd, 2005). Therefore, a potential limitation of this study is that individuals with BAP within the pedigree were incorrectly assumed to be carrying the shared susceptibility variant. Additionally, variants in linkage disequilibrium with the *GBE1* variant, including variants in non-coding regions, might also be contributing to the manifestation of the ASD phenotype within this family. Therefore, caution should be taken when interpreting this variant in other familial contexts. Moreover, we cannot discard the possibility that other genetic factors not interrogated in this study could be causative of ASD within this family. While there is significant cumulative genetic evidence suggesting *GBE1* is associated with familial ASD, the limitation of genetic studies is the lack of functional evidence supporting the disease association. The development of new tools and models, in particular, human induced pluripotent stem cells and CRISPR/Cas9 genome editing, provide the opportunity to validate the genetic studies and test underlying pathogenic mechanism(s).

In this study, we hypothesised that the underlying relative genetic homogeneity in a large multiplex family with ASD would allow us to identify an ASD susceptibility variant. We identify *GBE1* as a potential ASD susceptibility gene and provide multiple lines of evidence to support its pathogenicity. Deep phenotyping of individuals was critical to enable classification of both ASD and the BAP; misclassifying individuals with the BAP as unaffected would lead to very different and we believe, incorrect linkage and variant identification outcomes. Deciphering ASD genetics requires careful discrimination and consideration of all potential phenotypes associated with the variants contributing to familial forms of ASD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We thank the families for participating in our research. We acknowledge the UK10K datasets generated by the Biomarkers in Neurodevelopmental Disorders (BioNed) Study, Institute of Child Health & Great Ormond Street Hospital Autism Families Study, The Molecular Genetics of Autism Study and The International Molecular Genetic Study of Autism Consortium. The datasets used for the analysis described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number, phs000267.v4.p2. Submission of the data, phs000267.v4.p2, to dbGaP was provided by Dr. Bernie Devlin on behalf of the Autism Genome Project (AGP). Collection and submission of the data to dbGaP were supported by a grant from the Medical Research Council (G0601030) and the Wellcome Trust (075491/Z/04), Anthony P. Monaco, P.I., University of Oxford. AGP project references: 1) "Functional impact of global rare copy number variation in autism spectrum disorders",

The AGP Consortium, Nature, 466(7204):368-72 (2010), PMID: 20531469. 2) “A genome-wide scan for common alleles affecting risk for autism”, The AGP Consortium, Human Molecular Genetics, 19(20):4072-82 (2010), PMID: 20663923. 3) “Mapping autism risk loci using genetic linkage and chromosomal rearrangements”, The AGP Consortium, Nature Genetics, 39(3):319-328 (2007), PMID: 17322880. This work was funded in part by a US National Institutes of Health (NIH) grant (R01MH101221) and National Health and Medical Research Council Australia Project Grants (GNT1044175 and GNT1098255). KB was supported by an EH Flack Fellowship and IES was supported by a NHMRC Senior Practitioner Fellowship (1006110, 2011-2015; 1104831 2016-2020). MB was supported by an NHMRC Fellowship (GNT1102971) and E.E.E. is an investigator of the Howard Hughes Medical Institute. PJL was supported by a NHMRC Fellowship (GNT1032364) and the Vincent Chiodo Foundation. This work was made possible through the Victorian State Government Operational Infrastructure Support and Australian Government NHMRC IRIISS.

Grant numbers:

US National Institutes of Health grant R01MH101221

National Health and Medical Research Council Australia grant 1044175

National Health and Medical Research Council Australia grant 1098255

National Health and Medical Research Council Australia grant 1006110

National Health and Medical Research Council Australia grant 1104831

National Health and Medical Research Council Australia grant 1102971

National Health and Medical Research Council Australia grant 1032364

Data availability:

The data from study participants that support the findings of this study are available on reasonable request from the corresponding author. The data are not publicly available due to privacy and ethical restrictions. The UK10K and Autism Genome Project (AGP) Consortium datasets analysed during the current study are available in the EGA (https://www.ebi.ac.uk/ega/) and dbGAP repositories (https://www.ncbi.nlm.nih.gov/gap) subject to appropriate data use agreements being executed.

REFERENCES

Allen-Brady K, Miller J, Matsunami N, Stevens J, Block H, Farley M, … Coon H (2009). A high-density SNP genome-wide linkage scan in a large autism extended pedigree. Mol Psychiatry, 14(6), 590–600. doi:10.1038/mp.2008.14 [PubMed: 18283277]

Bahlo M, & Bromhead CJ (2009). Generating linkage mapping files from Affymetrix SNP chip data. Bioinformatics, 25(15), 1961–1962. doi:10.1093/bioinformatics/btp313 [PubMed: 19435744]

Bailey A, Palferman S, Heavey L, & Le Couteur A (1998). Autism: the phenotype in relatives. J Autism Dev Disord, 28(5), 369–392. doi:10.1023/a:1026048320785 [PubMed: 9813774]

Baron CA, Liu SY, Hicks C, & Gregg JP (2006). Utilization of lymphoblastoid cell lines as a system for the molecular modeling of autism. J Autism Dev Disord, 36(8), 973–982. doi:10.1007/s10803-006-0134-x [PubMed: 16845580]

Belanger M, Allaman I, & Magistretti PJ (2011). Brain energy metabolism: focus on astrocyte-neuron metabolic cooperation. Cell Metab, 14(6), 724–738. doi:10.1016/j.cmet.2011.08.016 [PubMed: 22152301]

Bengtsson H, Irizarry R, Carvalho B, & Speed TP (2008). Estimation and assessment of raw copy numbers at the single locus level. Bioinformatics, 24(6), 759–767. doi:10.1093/bioinformatics/btn016 [PubMed: 18204055]

Bengtsson H, Wirapati P, & Speed TP (2009). A single-array preprocessing method for estimating full-resolution raw copy numbers from all Affymetrix genotyping arrays including GenomeWideSNP 5 & 6. Bioinformatics, 25(17), 2149–2156. doi:10.1093/bioinformatics/btp371 [PubMed: 19535535]

Bolton P, Macdonald H, Pickles A, Rios P, Goode S, Crowson M, … Rutter M (1994). A case-control family history study of autism. J Child Psychol Psychiatry, 35(5), 877–900. doi:10.1111/j.1469-7610.1994.tb02300.x [PubMed: 7962246]

Cantor RM, Kono N, Duvall JA, Alvarez-Retuerto A, Stone JL, Alarcon M, … Geschwind DH (2005). Replication of autism linkage: fine-mapping peak at 17q21. Am J Hum Genet, 76(6), 1050–1056. doi:10.1086/430278 [PubMed: 15877280]

Casey JP, Magalhaes T, Conroy JM, Regan R, Shah N, Anney R, … Ennis S (2012). A novel approach of homozygous haplotype sharing identifies candidate genes in autism spectrum disorder. Hum Genet, 131(4), 565–579. doi:10.1007/s00439-011-1094-6 [PubMed: 21996756]

Celestino-Soper PB, Violante S, Crawford EL, Luo R, Lionel AC, Delaby E, … Beaudet AL (2012). A common X-linked inborn error of carnitine biosynthesis may be a risk factor for nondysmorphic autism. Proc Natl Acad Sci U S A, 109(21), 7974–7981. doi:10.1073/pnas.1120210109 [PubMed: 22566635]

Chang J, Gilman SR, Chiang AH, Sanders SJ, & Vitkup D (2015). Genotype to phenotype relationships in autism spectrum disorders. Nat Neurosci, 18(2), 191–198. doi:10.1038/nn.3907 [PubMed: 25531569]

Chen JA, Penagarikano O, Belgard TG, Swarup V, & Geschwind DH (2015). The emerging picture of autism spectrum disorder: genetics and pathology. Annu Rev Pathol, 10, 111–144. doi:10.1146/annurev-pathol-012414-040405 [PubMed: 25621659]

Codina-Sola M, Rodriguez-Santiago B, Homs A, Santoyo J, Rigau M, Aznar-Lain G, … Cusco I (2015). Integrated analysis of whole-exome sequencing and transcriptome profiling in males with autism spectrum disorders. Mol Autism, 6, 21. doi:10.1186/s13229-015-0017-0 [PubMed: 25969726]

Coe BP, Witherspoon K, Rosenfeld JA, van Bon BW, Vulto-van Silfhout AT, Bosco P, … Eichler EE (2014). Refining analyses of copy number variation identifies specific genes associated with developmental delay. Nat Genet, 46(10), 1063–1071. doi:10.1038/ng.3092 [PubMed: 25217958]

Constantino JN, & Todd RD (2005). Intergenerational transmission of subthreshold autistic traits in the general population. Biol Psychiatry, 57(6), 655–660. doi:10.1016/j.biopsych.2004.12.014 [PubMed: 15780853]

Cukier HN, Dueker ND, Slifer SH, Lee JM, Whitehead PL, Lalanne E, … Pericak-Vance MA (2014). Exome sequencing of extended families with autism reveals genes shared across neurodevelopmental and neuropsychiatric disorders. Mol Autism, 5(1), 1. doi:10.1186/2040-2392-5-1 [PubMed: 24410847]

Cummings BB, Karczewski KJ, Kosmicki JA, Seaby EG, Watts NA, Singer-Berk M, … MacArthur DG (2020). Transcript expression-aware annotation improves rare variant interpretation. Nature, 581(7809), 452–458. doi:10.1038/s41586-020-2329-2 [PubMed: 32461655]

Dainese L, Monin ML, Demeret S, Brochier G, Froissart R, Spraul A, … Mochel F (2013). Abnormal glycogen in astrocytes is sufficient to cause adult polyglucosan body disease. Gene, 515(2), 376–379. doi:10.1016/j.gene.2012.12.065 [PubMed: 23266647]

de la Torre-Ubieta L, Won H, Stein JL, & Geschwind DH (2016). Advancing the understanding of autism disease mechanisms through genetics. Nat Med, 22(4), 345–361. doi:10.1038/nm.4071 [PubMed: 27050589]

Freitag CM (2007). The genetics of autistic disorders and its clinical relevance: a review of the literature. Mol Psychiatry, 12(1), 2–22. doi:10.1038/sj.mp.4001896 [PubMed: 17033636]

Froese DS, Michaeli A, McCorvie TJ, Krojer T, Sasi M, Melaev E, … Yue WW (2015). Structural basis of glycogen branching enzyme deficiency and pharmacologic rescue by rational peptide design. Hum Mol Genet, 24(20), 5667–5676. doi:10.1093/hmg/ddv280 [PubMed: 26199317]

Geisheker MR, Heymann G, Wang T, Coe BP, Turner TN, Stessman HAF, … Eichler EE (2017). Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. Nat Neurosci, 20(8), 1043–1051. doi:10.1038/nn.4589 [PubMed: 28628100]

George AW, Wijsman EM, & Thompson EA (2005). MCMC multilocus lod scores: application of a new approach. Hum Hered, 59(2), 98–108. doi:10.1159/000085224 [PubMed: 15838179]

Georgiades S, Szatmari P, Zwaigenbaum L, Bryson S, Brian J, Roberts W, … Garon N (2013). A prospective study of autistic-like traits in unaffected siblings of probands with autism spectrum disorder. JAMA Psychiatry, 70(1), 42–48. doi:10.1001/2013.jamapsychiatry.1 [PubMed: 22945359]

Guo H, Duyzend MH, Coe BP, Baker C, Hoekzema K, Gerdts J, … Eichler EE (2019). Genome sequencing identifies multiple deleterious variants in autism patients with more severe phenotypes. Genet Med, 21(7), 1611–1620. doi:10.1038/s41436-018-0380-2 [PubMed: 30504930]

Guo H, Wang T, Wu H, Long M, Coe BP, Li H, … Xia K (2018). Inherited and multiple de novo mutations in autism/developmental delay risk genes suggest a multifactorial model. Mol Autism, 9, 64. doi:10.1186/s13229-018-0247-z [PubMed: 30564305]

Helbig I, Scheffer IE, Mulley JC, & Berkovic SF (2008). Navigating the channels and beyond: unravelling the genetics of the epilepsies. Lancet Neurol, 7(3), 231–245. doi:10.1016/S1474-4422(08)70039-5 [PubMed: 18275925]

Henden L, Wakeham D, & Bahlo M (2016). XIBD: software for inferring pairwise identity by descent on the X chromosome. Bioinformatics, 32(15), 2389–2391. doi:10.1093/bioinformatics/btw124 [PubMed: 27153693]

Horie M, Honda T, Suzuki Y, Kobayashi Y, Daito T, Oshida T, … Tomonaga K (2010). Endogenous non-retroviral RNA virus elements in mammalian genomes. Nature, 463(7277), 84–87. doi:10.1038/nature08695 [PubMed: 20054395]

Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, … Wigler M (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature, 515(7526), 216–221. doi:10.1038/nature13908 [PubMed: 25363768]

Kakhlon O, Glickstein H, Feinstein N, Liu Y, Baba O, Terashima T, … Lossos A (2013). Polyglucosan neurotoxicity caused by glycogen branching enzyme deficiency can be reversed by inhibition of glycogen synthase. J Neurochem, 127(1), 101–113. doi:10.1111/jnc.12277 [PubMed: 23607684]

Klein CJ (1993). Adult Polyglucosan Body Disease. In Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), GeneReviews((R)). Seattle (WA).

Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Albarca Aguilera M, Meyer R, & Massouras A (2019). VarSome: the human genomic variant search engine. Bioinformatics, 35(11), 1978–1980. doi:10.1093/bioinformatics/bty897 [PubMed: 30376034]

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, … Exome Aggregation C (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature, 536(7616), 285–291. doi:10.1038/nature19057 [PubMed: 27535533]

Leutenegger AL, Prum B, Genin E, Verny C, Lemainque A, Clerget-Darpoux F, & Thompson EA (2003). Estimation of the inbreeding coefficient through use of genomic data. Am J Hum Genet, 73(3), 516–523. doi:10.1086/378207 [PubMed: 12900793]

Li H (2011). Improving SNP discovery by base alignment quality. Bioinformatics, 27(8), 1157–1158. doi:10.1093/bioinformatics/btr076 [PubMed: 21320865]

Li J, Wang L, Guo H, Shi L, Zhang K, Tang M, … Xia K (2017). Targeted sequencing and functional analysis reveal brain-size-related genes and their networks in autism spectrum disorders. Mol Psychiatry, 22(9), 1282–1290. doi:10.1038/mp.2017.140 [PubMed: 28831199]

Lim ET, Raychaudhuri S, Sanders SJ, Stevens C, Sabo A, MacArthur DG, … Daly MJ (2013). Rare complete knockouts in humans: population distribution and significant role in autism spectrum disorders. Neuron, 77(2), 235–242. doi:10.1016/j.neuron.2012.12.029 [PubMed: 23352160]

Lockhart PJ, Lincoln S, Hulihan M, Kachergus J, Wilkes K, Bisceglio G, … Farrer MJ (2004). DJ-1 mutations are a rare cause of recessively inherited early onset parkinsonism mediated by loss of protein function. J Med Genet, 41(3), e22. doi:10.1136/jmg.2003.011106 [PubMed: 14985393]

Losh M, Childress D, Lam K, & Piven J (2008). Defining key features of the broad autism phenotype: a comparison across parents of multiple- and single-incidence autism families. Am J Med Genet B Neuropsychiatr Genet, 147B(4), 424–433. doi:10.1002/ajmg.b.30612 [PubMed: 17948871]

Magoulas PL, & El-Hattab AW (1993). Glycogen Storage Disease Type IV. In Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), GeneReviews((R)). Seattle (WA).

Mamczur P, Borsuk B, Paszko J, Sas Z, Mozrzymas J, Wisniewski JR, … Rakus D (2015). Astrocyte-neuron crosstalk regulates the expression and subcellular localization of carbohydrate metabolism enzymes. Glia, 63(2), 328–340. doi:10.1002/glia.22753 [PubMed: 25257920]

Mochel F, Schiffmann R, Steenweg ME, Akman HO, Wallace M, Sedel F, … Lossos A (2012). Adult polyglucosan body disease: Natural History and Key Magnetic Resonance Imaging Findings. Ann Neurol, 72(3), 433–441. doi:10.1002/ana.23598 [PubMed: 23034915]

Mozrzymas J, Szczesny T, & Rakus D (2011). The effect of glycogen phosphorolysis on basal glutaminergic transmission. Biochem Biophys Res Commun, 404(2), 652–655. doi:10.1016/j.bbrc.2010.12.033 [PubMed: 21146500]

Newman LA, Korol DL, & Gold PE (2011). Lactate produced by glycogenolysis in astrocytes regulates memory processing. PLoS One, 6(12), e28427. doi:10.1371/journal.pone.0028427 [PubMed: 22180782]

Ozonoff S, Young GS, Belding A, Hill M, Hill A, Hutman T, … Iosif AM (2014). The broader autism phenotype in infancy: when does it emerge? J Am Acad Child Adolesc Psychiatry, 53(4), 398–407 e392. doi:10.1016/j.jaac.2013.12.020 [PubMed: 24655649]

Ozonoff S, Young GS, Carter A, Messinger D, Yirmiya N, Zwaigenbaum L, … Stone WL (2011). Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study. Pediatrics, 128(3), e488–495. doi:10.1542/peds.2010-2825 [PubMed: 21844053]

Paradas C, Akman HO, Ionete C, Lau H, Riskind PN, Jones DE, … DiMauro S (2014). Branching enzyme deficiency: expanding the clinical spectrum. JAMA neurology, 71(1), 41–47. doi:10.1001/jamaneurol.2013.4888 [PubMed: 24248152]

Pires DE, Ascher DB, & Blundell TL (2014). mCSM: predicting the effects of mutations in proteins using graph-based signatures. Bioinformatics, 30(3), 335–342. doi:10.1093/bioinformatics/btt691 [PubMed: 24281696]

Pisula E, & Ziegart-Sadowska K (2015). Broader Autism Phenotype in Siblings of Children with ASD--A Review. Int J Mol Sci, 16(6), 13217–13258. doi:10.3390/ijms160613217 [PubMed: 26068453]

Piven J, Palmer P, Jacobi D, Childress D, & Arndt S (1997). Broader autism phenotype: evidence from a family history study of multiple-incidence autism families. Am J Psychiatry, 154(2), 185–190. doi:10.1176/ajp.154.2.185 [PubMed: 9016266]

Piven J, Vieland VJ, Parlier M, Thompson A, O'Conner I, Woodbury-Smith M, … Szatmari P (2013). A molecular genetic study of autism and related phenotypes in extended pedigrees. J Neurodev Disord, 5(1), 30. doi:10.1186/1866-1955-5-30 [PubMed: 24093601]

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, … Sham PC (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet, 81(3), 559–575. doi:10.1086/519795 [PubMed: 17701901]

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, … Committee ALQA (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med, 17(5), 405–424. doi:10.1038/gim.2015.30 [PubMed: 25741868]

Robinson EB, Samocha KE, Kosmicki JA, McGrath L, Neale BM, Perlis RH, & Daly MJ (2014). Autism spectrum disorder severity reflects the average contribution of de novo and familial influences. Proc Natl Acad Sci U S A, 111(42), 15161–15165. doi:10.1073/pnas.1409204111 [PubMed: 25288738]

Robinson EB, St Pourcain B, Anttila V, Kosmicki JA, Bulik-Sullivan B, Grove J, … Daly MJ (2016). Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. Nat Genet, 48(5), 552–555. doi:10.1038/ng.3529 [PubMed: 26998691]

Ronemus M, Iossifov I, Levy D, & Wigler M (2014). The role of de novo mutations in the genetics of autism spectrum disorders. Nat Rev Genet, 15(2), 133–141. doi:10.1038/nrg3585 [PubMed: 24430941]

Sandin S, Lichtenstein P, Kuja-Halkola R, Hultman C, Larsson H, & Reichenberg A (2017). The Heritability of Autism Spectrum Disorder. JAMA, 318(12), 1182–1184. doi:10.1001/jama.2017.12141 [PubMed: 28973605]

Sasson NJ, Lam KS, Childress D, Parlier M, Daniels JL, & Piven J (2013). The broad autism phenotype questionnaire: prevalence and diagnostic classification. Autism Res, 6(2), 134–143. doi:10.1002/aur.1272 [PubMed: 23427091]

Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, … Wigler M (2007). Strong association of de novo copy number mutations with autism. Science, 316(5823), 445–449. doi:10.1126/science.1138659 [PubMed: 17363630]

Consortium SPARK. (2018). SPARK: A US Cohort of 50,000 Families to Accelerate Autism Research. Neuron, 97(3), 488–493. doi:10.1016/j.neuron.2018.01.015 [PubMed: 29420931]

Taylor JM, Brody KM, & Lockhart PJ (2012). Parkin co-regulated gene is involved in aggresome formation and autophagy in response to proteasomal impairment. Exp Cell Res, 318(16), 2059–2070. doi:10.1016/j.yexcr.2012.05.011 [PubMed: 22652456]

Tick B, Bolton P, Happe F, Rutter M, & Rijsdijk F (2016). Heritability of autism spectrum disorders: a meta-analysis of twin studies. J Child Psychol Psychiatry, 57(5), 585–595. doi:10.1111/jcpp.12499 [PubMed: 26709141]

Toma C, Torrico B, Hervas A, Valdes-Mas R, Tristan-Noguero A, Padillo V, … Cormand B (2014). Exome sequencing in multiplex autism families suggests a major role for heterozygous truncating mutations. Mol Psychiatry, 19(7), 784–790. doi:10.1038/mp.2013.106 [PubMed: 23999528]

Trevis KJ, Brown NJ, Green CC, Lockhart PJ, Desai T, Vick T, … Wilson SJ (2020). Tracing Autism Traits in Large Multiplex Families to Identify Endophenotypes of the Broader Autism Phenotype. Int J Mol Sci, 21(21). doi:10.3390/ijms21217965

UK10K Consortium., Walter K, Min JL, Huang J, Crooks L, Memari Y, … Soranzo N (2015). The UK10K project identifies rare variants in health and disease. Nature, 526(7571), 82–90. doi:10.1038/nature14962 [PubMed: 26367797]

Vilchez D, Ros S, Cifuentes D, Pujadas L, Valles J, Garcia-Fojeda B, … Guinovart JJ (2007). Mechanism suppressing glycogen synthesis in neurons and its demise in progressive myoclonus epilepsy. Nat Neurosci, 10(11), 1407–1413. doi:10.1038/nn1998 [PubMed: 17952067]

Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, … Bucan M (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res, 17(11), 1665–1674. doi:10.1101/gr.6861907 [PubMed: 17921354]

Xu Y, Yang Y, Cai Y, Liu F, Liu Y, Zhu Y, & Wu J (2011). The X protein of hepatitis B virus activates hepatoma cell proliferation through repressing melanoma inhibitory activity 2 gene. Biochem Biophys Res Commun, 416(3–4), 379–384. doi:10.1016/j.bbrc.2011.11.046 [PubMed: 22120627]

Yu TW, Chahrour MH, Coulter ME, Jiralerspong S, Okamura-Ikeda K, Ataman B, … Walsh CA (2013). Using whole-exome sequencing to identify inherited causes of autism. Neuron, 77(2), 259–273. doi:10.1016/j.neuron.2012.11.002 [PubMed: 23352163]

Yuen RK, Thiruvahindrapuram B, Merico D, Walker S, Tammimies K, Hoang N, … Scherer SW (2015). Whole-genome sequencing of quartet families with autism spectrum disorder. Nat Med, 21(2), 185–191. doi:10.1038/nm.3792 [PubMed: 25621899]

Yuen RKC, Merico D, Bookman M, J LH, Thiruvahindrapuram B, Patel RV, … Scherer SW (2017). Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat Neurosci, 20(4), 602–611. doi:10.1038/nn.4524 [PubMed: 28263302]

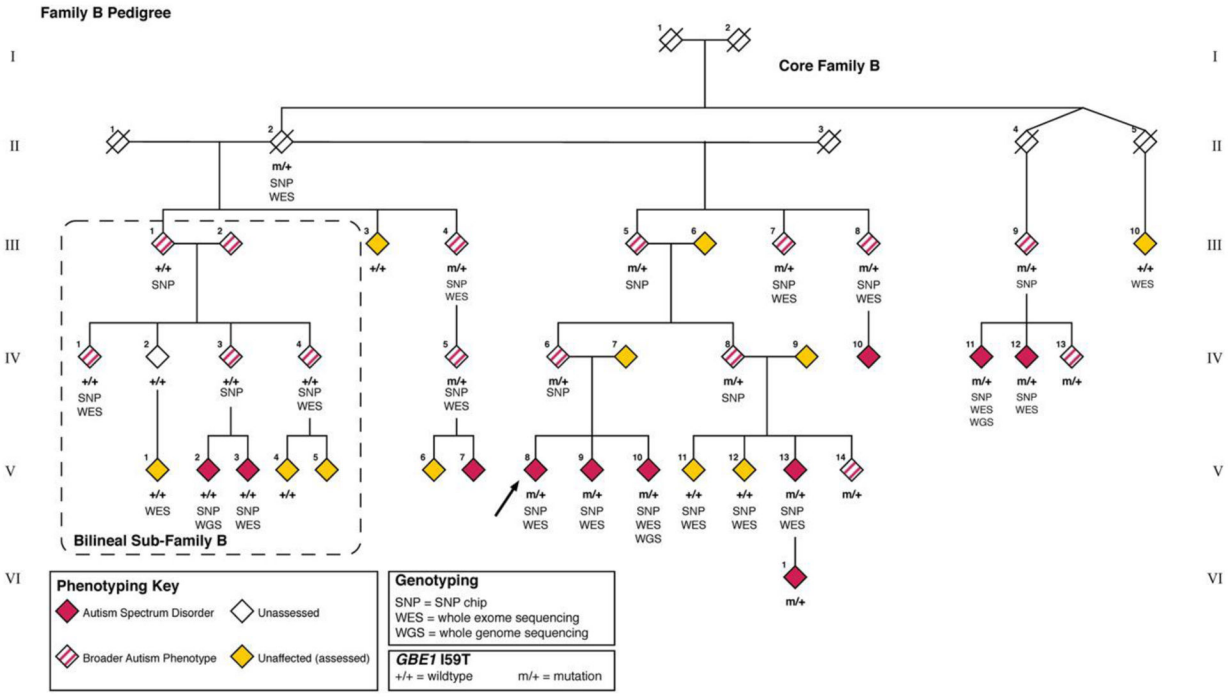

Figure 1:

De-identified and scrambled pedigree of Family B showing the unilineal Core family utilised for linkage analysis in this study and the excluded bilineal branch (indicated by the dashed box). The phenotyping key for ASD, BAP, unassessed and unaffected (assessed) individuals is indicated, as are the molecular studies performed on each sample and *GBE1* p.Ile59Thr variant status.

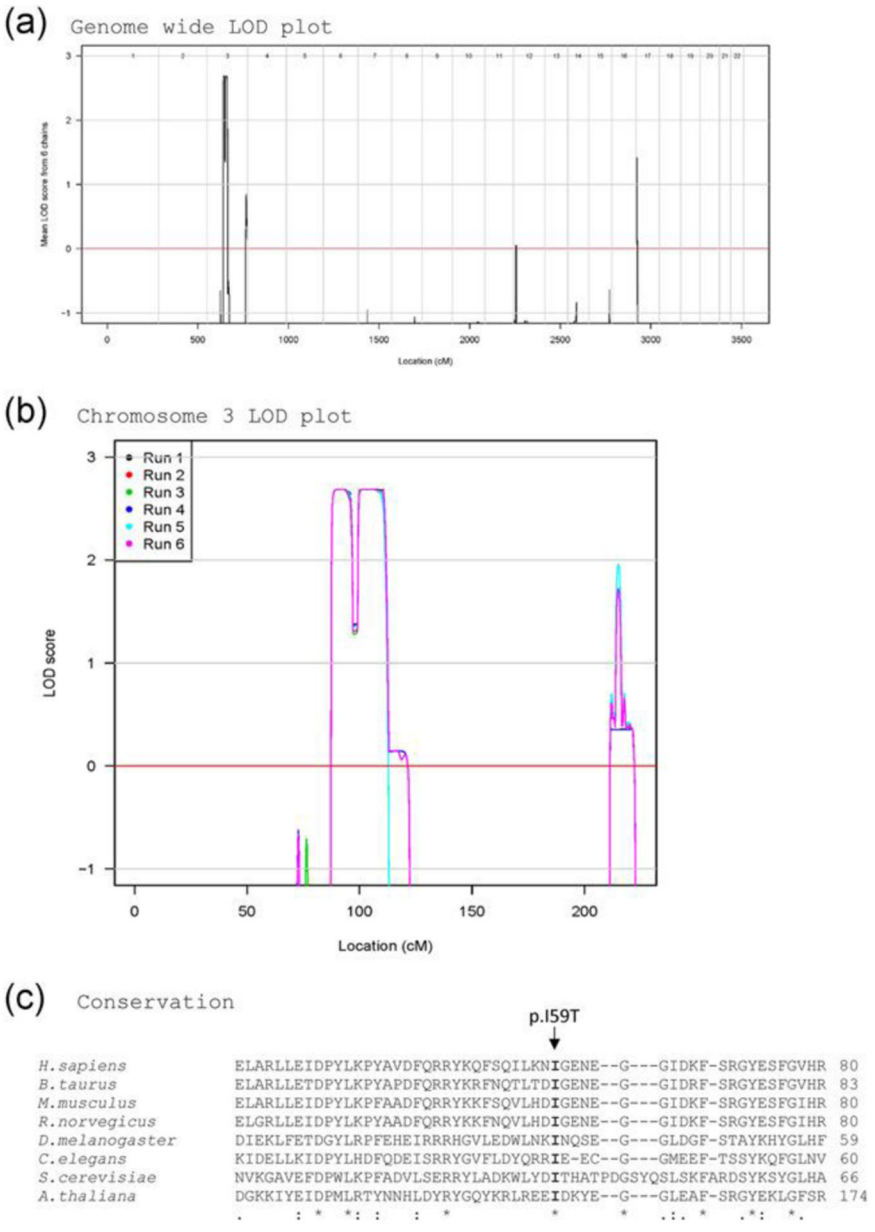

Figure 2:

Genome-wide linkage analysis plot for Core Family B (A) and chromosome 3 linkage plot (B). Data generated using the lm bayes function in MORGAN identified a region with a LOD score of 2.9 on chr3. (C) ClustalW alignment of the GBE1 amino acid sequence demonstrated evolutionary conservation of isoleucine at amino acid position 59.

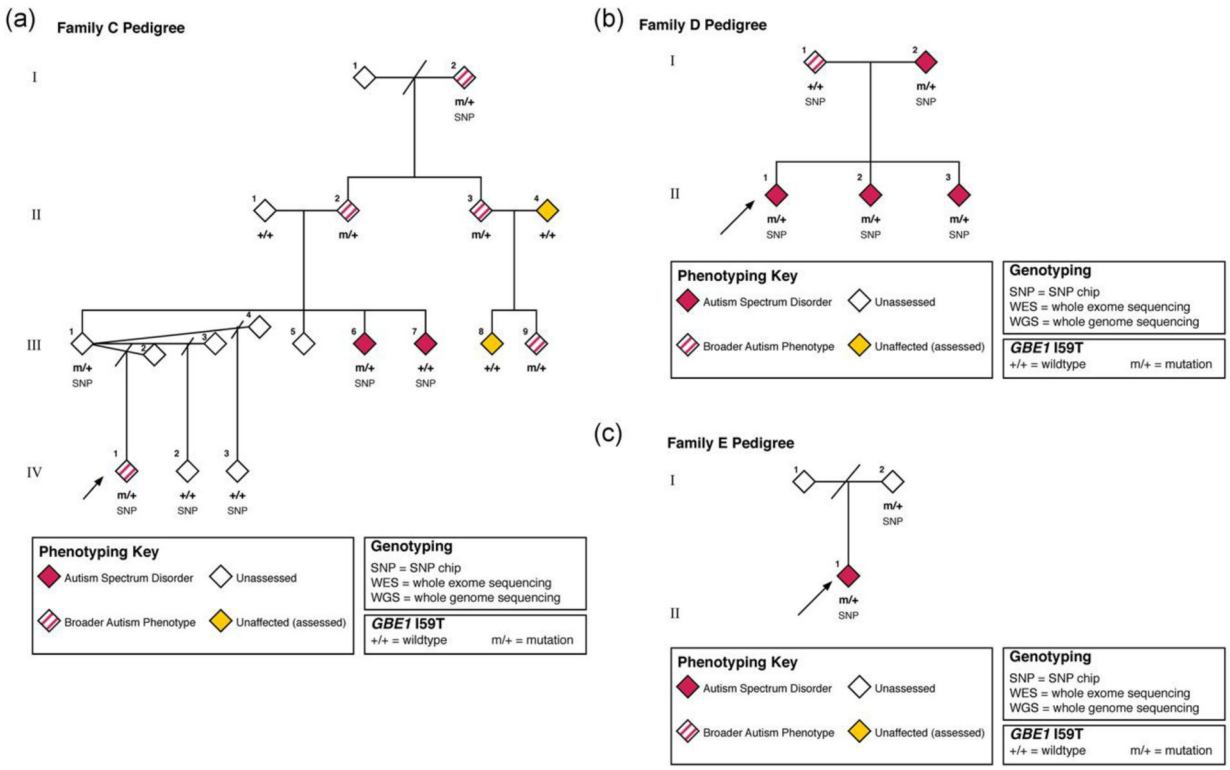

Figure 3:

De-identified and scrambled pedigrees of Family C (A), Family D (B) and Family E (C). The phenotyping key for ASD, BAP, unassessed and unaffected (assessed) individuals is indicated, as are the molecular studies performed on each sample and *GBE1* p.Ile59Thr variant status.

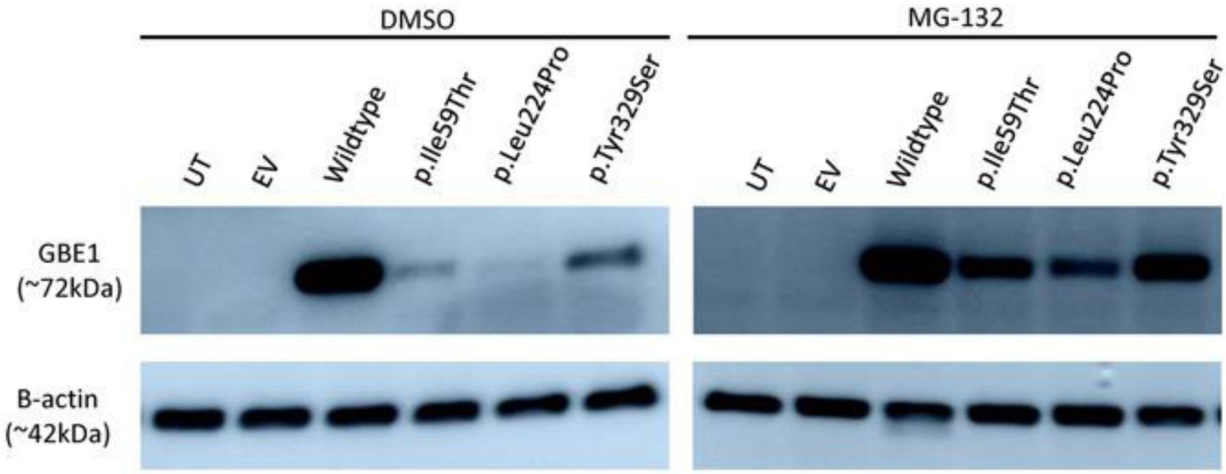

Figure 4.

HEK293 cells transfected with constructs encoding wildtype and variant GBE1 were grown in the absence or presence of 10 mM MG-132 for 14 hours. Total protein was isolated and analyzed by SDS-PAGE and immunoblotting with antibodies directed against GBE1 and β-actin. A representative image is shown, with approximate sizes in kDa indicated. UT=untransfected cells, EV=empty vector control.

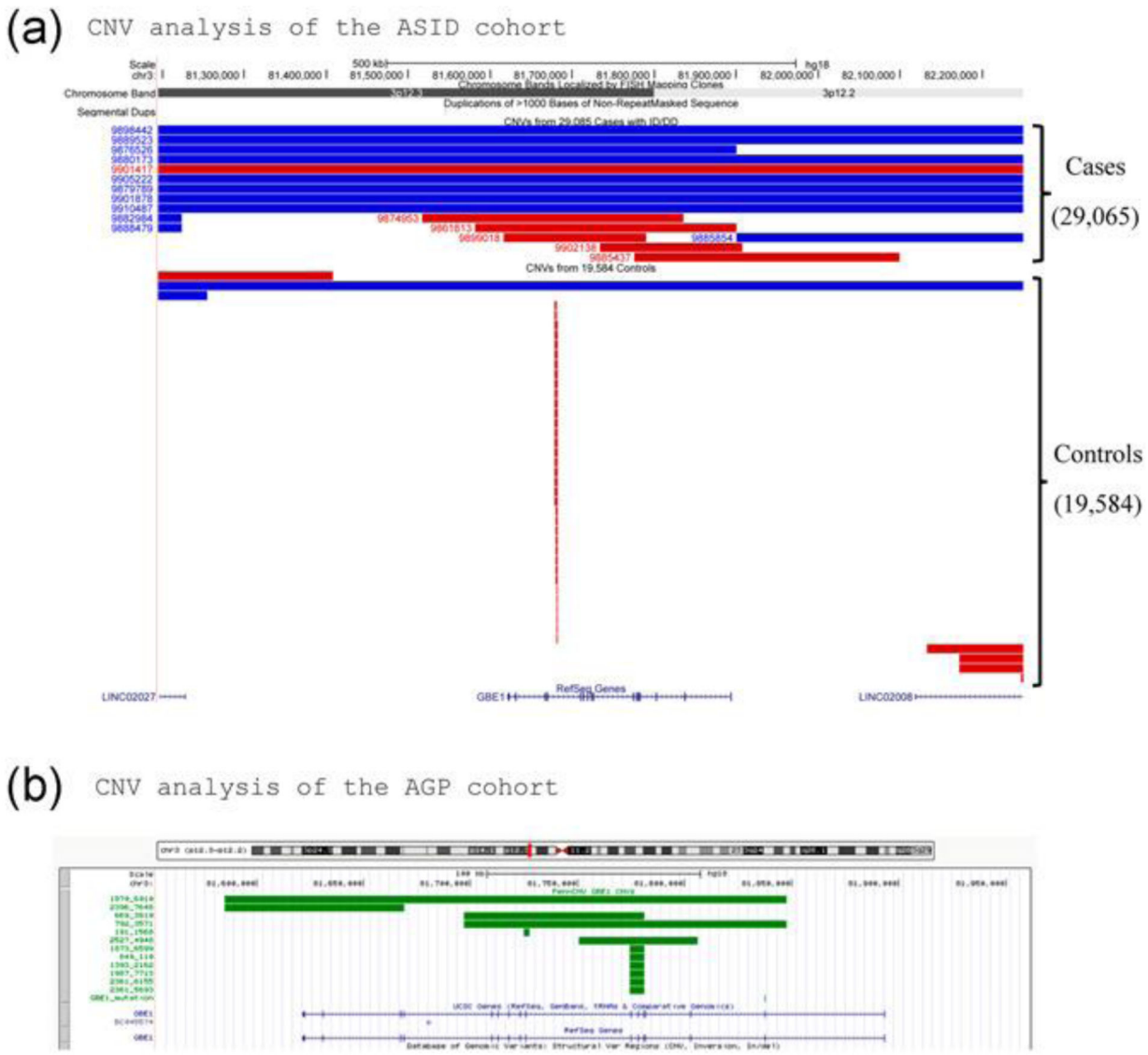

Figure 5.

CNV analysis. (A). Screenshot of UCSC Genome Browser displaying custom track of results of *GBE1* CNVs in a cohort of children with neurodevelopmental delay (top) and controls (bottom); deletions are indicated in red, duplications in blue (B). Screenshot of UCSC Genome Browser displaying custom track of results of *GBE1* deletions (green) in the AGP cohort.