SLCO5A1 and synaptic assembly genes contribute to impulsivity in juvenile myoclonic epilepsy

Naim Panjwani[1], Amy Shakeshaft[2,3] Delnaz Roshandel[1], Fan Lin [1], Amber Collingwood[2], Anna Hall[2], Katherine Keenan[1], Celine Deneubourg[2], Filippo Mirabella[2], Simon Topp[2], Jana Zarubova[4], Rhys H. Thomas[5,6], Inga Talvik[7], Marte Syvertsen[8], Pasquale Striano[9,10], Anna B. Smith[2], Kaja K. Selmer[11,12], Guido Rubboli[13,14], Alessandro Orsini[15], Ching Ching Ng[16], Rikke S. Møller[13,17], Kheng Seang Lim[18], Khalid Hamandi[19,20], David A. Greenberg[21], Joanna Gesche[22], Elena Gardella[13,17], Choong Yi Fong[23], Christoph P. Beier[22], Danielle M. Andrade[24], Heinz Jungbluth[25,26] , Mark P. Richardson[2,27], Annalisa Pastore[2], Manolis Fanto[2], *Deb K. Pal[2,3,27] & *Lisa J. Strug[1,28,29]

[1]Program in Genetics and Genome Biology, The Hospital for Sick Children, Toronto, Canada; [2]Department of Basic & Clinical Neurosciences, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK; [3]MRC Centre for Neurodevelopmental Disorders, King's College London; [4]Department of Neurology, Second Faculty of Medicine, Charles University and Motol University Hospital, Prague, Czech Republic; [5]Newcastle upon Tyne NHS Foundation Trust, Newcastle, United Kingdom; [6]Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle, UK; [7]Tallin Children's Hospital, Tallin, Estonia; [8]Department of Neurology, Drammen Hospital, Vestre Viken Health Trust, Oslo, Norway; [9]IRCCS Istituto 'G. Gaslini', Genova, Italy; [10]University of Genova, Genova, Italy; [11]Department of Research and Innovation, Division of Clinical Neuroscience, Oslo University Hospital, Norway; [12]National Centre for Epilepsy, Oslo University Hospital, Norway; [13]Danish Epilepsy Centre, Dianalund, Denmark; [14]University of Copenhagen, Denmark; [15]Pediatric Neurology, Azienda Ospedaliero-Universitaria Pisana, Pisa University Hospital, Italy; [16]Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia; [17]Department of Regional Health Research, University of Southern Denmark, Odense, Denmark; [18]Division of Neurology, Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia; [19]The Welsh Epilepsy Unit, Department of Neurology Cardiff & Vale University Health Board, UK; [20]Department of Psychological Medicine and Clinical Neuroscience, Cardiff University, Cardiff, UK; [21]Nationwide Children's Hospital, Ohio, USA; [22]Odense University Hospital, Odense, Denmark; [23]Division of Paediatric Neurology, Department of Pediatrics, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia; [24]Adult Epilepsy Genetics Program, Krembil Research Institute, University of Toronto, Canada; [25]Randall Centre for Cell and Molecular Biophysics, Muscle Signalling Section, Faculty of Life Sciences and Medicine, King's College London, London, UK; [26]Department of Paediatric Neurology, Neuromuscular Service, Evelina's Children Hospital, Guy's & St. Thomas' Hospital NHS Foundation Trust, London, UK; [27]King's College Hospital, London, UK; [28]Departments of Statistical Sciences and Computer Science and Division of Biostatistics, The University of Toronto, Toronto, Canada; [29]The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Canada.

*Correspondence to:

Professor Deb K Pal
Maurice Wohl Clinical Neurosciences Institute
Institute of Psychiatry, Psychology & Neuroscience,
King's College London
5 Cutcombe Road, London SE5 9RX
deb.pal@kcl.ac.uk
@palneurolab

Professor Lisa J Strug
Departments of Statistical Sciences and Computer Science
The University of Toronto
Senior Scientist
The Hospital for Sick Children
555 University Avenue
Toronto, ON Canada
Lisa.Strug@utoronto.ca

Introductory Paragraph

Elevated impulsivity is a key component of attention-deficit hyperactivity disorder (ADHD), bipolar disorder and epilepsy[1-5]. We performed a genome-wide association, colocalization and pathway analysis of impulsivity in juvenile myoclonic epilepsy (JME). We identify genome-wide associated SNPs at 8q13.3 ($p=7.5 \times 10^{-9}$) and 10p11.21 ($p=3.6 \times 10^{-8}$). The 8q13.3 locus colocalizes with *SLCO5A1* expression quantitative trait loci in cerebral cortex ($p=9.5 \times 10^{-3}$). *SLCO5A1* codes for a membrane-bound organic anion transporter[6] and upregulates synapse assembly/organisation genes[7]. Pathway analysis also demonstrates 9.3-fold enrichment for synaptic assembly genes (p=0.03) including *NRXN1, NLGN1* and *PTPRD*. RNAi knockdown of *Oatp30B*, the *Drosophila* homolog of *SLCO5A1*, causes both over-reactive startling behaviour ($p=8.7 \times 10^{-3}$) and increased seizure-like events ($p=6.8 \times 10^{-7}$). Polygenic risk score for ADHD correlates with impulsivity scores ($p=1.60 \times 10^{-3}$), demonstrating shared genetic contributions. *SLCO5A1* loss-of-function represents a novel impulsivity and seizure mechanism. Synaptic assembly genes may inform the aetiology of impulsivity in health and disease.

Main

Impulsivity is a heritable behavioural trait leading to actions that are "poorly conceived, prematurely expressed, unduly risky or inappropriate to the situation and that often result in undesirable consequences"[8]. Raised impulsivity is a key endophenotype of attention-deficit hyperactive disorder (ADHD)[1], bipolar disorder[2] and juvenile myoclonic epilepsy[3-5]. A previous genome-wide association study (GWAS) of impulsive personality traits (UPPS-P Sensation Seeking, Drug Experimentation and UPPS-P Negative Urgency) in 22,861 healthy individuals of European ancestry demonstrated two significant associated loci at 3p12.1 and 22q13.1[9]. Variants at the 3p12.1 locus correlated with predicted *Cell Adhesion Molecule–2* (*CADM2*) expression, in the putamen[10], and the 22q13.1 locus near *CACNA1I* has been previously implicated in schizophrenia[11]. CADM2 mediates synaptic signalling and is highly expressed in the human cerebral cortex and cerebellum[12]. Given impulsivity is elevated in neuropsychiatric disorders, there may be shared genetic architecture with impulsivity in the general population and/or across disorders, however to our knowledge there has been no GWAS of impulsivity in any neuropsychiatric disorder.

Impulsivity is elevated in different epilepsies, but the evidence across multiple dimensions of impulsivity is strongest in Juvenile Myoclonic Epilepsy (JME)[3-5], a common adolescent-onset syndrome characterized by awakening, generalized myoclonic, tonic-clonic and absence seizures often triggered by sleep deprivation. Trait impulsivity in JME is associated with the frequency of both myoclonic and absence seizures[3], but it is not clear if this indicates a causal relationship or a common mechanism regulating both impulsivity and seizures, though convergent lines of evidence

suggest the involvement of overlapping prefrontal-striatal networks in both JME and impulsivity[13-20]. Finding a shared aetiology would offer new therapeutic approaches for drug-resistant epilepsy.

The overall syndrome of JME has complex inheritance with few replicated susceptibility loci[21], and other loci with less support[22-24]. A major challenge for epilepsies of complex inheritance is to explain the wide variation in phenotypic expression and treatment response between individuals. Forty-percent experience antiseizure medication (ASM) resistance or intolerance[25]. In addition, no current ASM modifies the lifelong disease course of JME and the pharmacological options are sparse, especially for women[25]. Hence novel treatments based on genetic disease mechanisms, such as those emerging for monogenic channelopathy and mTOR pathway epilepsies, are urgently needed[26,27]. Our methodological approach is to carry out genome-wide analysis of endophenotypes in JME such as impulsivity and clinically relevant outcomes such as ASM resistance, a strategy with predicted advantages for reducing heterogeneity, increasing statistical power[28,29] and improving direct clinical translation for precision medicine.

We investigated the influence of 8,950,360 variants on impulsivity in European ancestry JME patients (n=324) and a mega-analysis with all ancestries (n=372), who self-rated their trait impulsivity using the Barratt Impulsivity Scale, eight-item BIS-Brief version[30]. We conducted a GWAS of BIS-Brief score in the European subset, adjusted for sex, genotyping batch, age at consent, population stratification, and seizure frequency (Extended Data Table 1). We discovered two genome-wide significant loci at chromosome 8 (rs73293634 (G/T)) and chromosome 10 (rs75042057 (T/G) (Fig. 1 and Table 1). Association of these two loci in a mega-analysis including all ancestry groups (Supplementary Fig. 1) provides stronger evidence of association as measured by the p-value (Table 1). The phenotypic variation explained (PVE) for rs73293634 was 10.1% in the European analysis. rs73293634 falls in an intergenic region near *SLCO5A1* and two individuals with large structural deletions that include *SLCO5A1* are reported in the Decipher Genomics database with seizures and a neurodevelopmental disorder (https://www.deciphergenomics.org/gene/SLCO5A1/patient-overlap/cnvs).

The significant genome-wide association on chromosome 10 (rs75042057) falls in intron 22 of *PARD3* (NM_001184785.2). The PVE by the SNP is 9.3%. Significant linkage (multipoint max LOD 4.23, alpha 0.34) was previously reported to this locus in French-Canadian families with idiopathic generalized epilepsy (IGE)[31], of which JME is a subtype.

<<Insert Figure 1 here >>

<< Insert Table 1 here>>

Since the GWAS-associated variants are not exonic, we next assessed whether the variants impact gene expression, and for which gene in which tissue of origin, by assessing colocalization of the

genome-wide significant peaks with expression quantitative trait loci (eQTL) in brain tissues. We used eQTLs from the Genotype-Tissue Expression project (GTEx) v8[12], PsychENCODE[32], and human fetal brains[33] and combined them with the GWAS summary statistics from the mega-analysis, for colocalization analysis adjusting for multiple hypothesis testing[34]. Colocalization analysis with eQTLs from GTEx brain and tibial nerve tissues for genes at the locus (chr8:69,650,000-70,000,000, hg38) shows significant colocalization with *SLCO5A1* in the cerebral cortex, and no colocalization with other genes in the region (Fig. 2a and Extended Data Fig. 1; Simple Sum colocalization $p=9.5 \times 10^{-3}$). The minor allele for the lead SNP rs73293634 (T) decreases expression in GTEx cerebral cortex (Fig. 2c). We found no significant colocalization in eQTLs from PsychENCODE[32] and fetal brains[33], although nearby variants in the locus in adult brains in PsychENCODE have, in general, a clear influence on *SLCO5A1* expression (Fig. 2b). According to BrainSpan[35,36], *SLCO5A1* is highly expressed prenatally, with expression dropping after birth but remains detectable throughout adulthood (Fig. 2d). We did not observe significant colocalization at the chromosome 10 locus with eQTLs from adult brains in GTEx[12], PsychENCODE[32] or fetal brains[33].

<<Insert Figure 2 here >>

<< Insert Figure 3 here >>

SLCO5A1 is a membrane-bound organic anion transporter with no known substrate[7] (Fig. 3). We performed a full BLASTp search of the SLCO5A1 polypeptide sequence (NP_112220.2) on *Drosophila melanogaster* to identify the closest matching sequence alignment. This revealed *Oatp30B* as the closest matching gene with a 37.66% identity and E-value of 2×10^{-150} (NP_995667.1). BLASTp of Oatp30B polypeptide sequence (Q9VLB3) across all species for conserved domains reveals this gene has conserved Major Facilitator Superfamily (MFS), OATP, and Kazal domains (Fig. 3 and 4a). We therefore used an effective RNAi transgenic line (Extended Data Fig. 2a) to assess the effect of pan-neuronal adult knockdown of *Oatp30B/SLCO5A1.* Flies with reduced *Oatp30B* levels displayed a small but significant shortening of their lifespan (Extended Data Fig. 2b) and a striking over-reaction to vibration stimuli applied through the automated Drosophila Arousal Tracking (DART) system[37], which elicit an otherwise modest activity response in two separate control fly genotypes (Fig. 4b). Additional analysis of locomotor behaviour clarifies that *Oatp30B* knockdown did not alter the speed of flies or the duration of each activity bout or the interval in between bouts of action (Extended Data Fig. 2c-e), indicating a specific defect in excessive response to stimuli. Furthermore, *Oatp30B* knockdown led to a dramatic increase in the frequency of seizure-like events (Fig. 4c) when exposed to hyperthermia, a trigger for seizures in *Drosophila*[38]. Recovery to full motility after seizure-like events was also significantly slower in flies with *Oatp30B* knockdown (Fig. 4d). These data establish a common causal link between *Oatp30B/SLCO5A1* downregulation, impulsivity-like startling behaviour, and susceptibility to seizure-like events.

The lead associated SNP at the *SLCO5A1* locus has MAF=4.1% and accounts for 10.1% PVE. We next sought to assess whether sub-GWAS significant signals could inform additional contributing genes or pathways and whether there were shared genetic contributions with other psychiatric or epilepsy phenotypes. We selected all variants displaying $P \leq 5 \times 10^{-4}$ and annotated these variants to the transcription start site of the nearest gene using the Ensembl Variant Effect Predictor (v94)[39], resulting in 855 unique genes; these were used as input both in a gene ontology (GO) enrichment analysis[40-42] and in FUMA's GENE2FUNC tool[43] (v1.3.7). Gene enrichment analysis using AmiGO[41] identified a 9.3-fold enrichment of associated genes from the presynaptic assembly and organisation gene set (nine out of 23 genes; p=0.03; GO:0099054). These genes were *NRXN1*, *NLGN1*, *NLGN4X, PTPRD*, *FZD1*, *PCLO*, *CNTN5*, *IL1RAPL1*, and *PTEN* (Table 2). The combined PVE for the lead variants annotated to these nine genes is 27.4% (34.1% with the addition of rs73293634 from the *SLCO5A1* locus and rs75042057 from the *PARD3* locus).

Investigation of these 855 genes revealed further specificity regarding localisation and function. First, a comparison with GTEx v8 54 tissue types revealed an enrichment of these genes in the cerebral cortex ($-\log_{10}P>5$). Second, we found significant enrichment of these 855 genes with two gene pathways in the Molecular Signatures Database (MSigDB v7.5.1)[44] identified as being active in murine whole brain[45] and neuronal progenitor cells[46]. Third, there was significant overlap with genes reported in the GWAS Catalog that contribute to phenotypes relevant to the predominance of JME seizures on awakening, impulsivity and metabolism: chronotype (68 out of 549 genes overlap, $p=1.14 \times 10^{-12}$), obesity-related traits (77 out of 706 overlap, $p=7.33 \times 10^{-12}$), general risk tolerance (33/247 overlap, $p=7.56 \times 10^{-7}$), and adventurousness (23/141, $p= 3.15 \times 10^{-6}$). Last, given impulsivity is a major component of ADHD, bipolar disorder and epilepsy, we tested and found that a higher ADHD polygenic risk score (PRS) was significantly associated with a higher BIS-Brief score ($p=1.60 \times 10^{-3}$). PRSs for bipolar disorder, generalized and focal epilepsy did not show significant evidence for association with BIS-Brief score (p=0.08, 0.33 and 0.96, respectively). Interestingly, genes nearest ADHD PRS SNPs were not enriched for synaptic assembly genes. Altogether this suggests that the impulsive trait seen in JME is an endophenotype that shares genetic architecture with impulsivity in the general population as well as with individuals diagnosed with ADHD.

This is the first GWAS of trait impulsivity in a neuropsychiatric disorder and we present evidence for the role of SLC05A1 in impulsivity and seizure susceptibility through triangulation[47] with GWAS, colocalization with gene expression and functional evaluation using an established model of seizure susceptibility in *Drosophila*[48]. There has only been one GWAS of impulsive traits in the general population, which identified genome-wide significant association with variants in the *CADM2* gene, encoding a cell adhesion protein from the SynCam Immunoglobulin superfamily of recognition molecules, important for synaptic organisation and specificity; and association of variants at the *CACNA1I* locus, which has been observed in previous studies with schizophrenia[9]. Our GWAS did not

show significant association with the previously reported associated variants at the *CADM2* and *CACNA1I* loci[9], (p=0.152 for rs139528938 and p=0.32 for rs4522708 which is a SNP with r^2=0.87 with the reported SNP, rs199694726, in our BIS dataset). Genome-wide summary statistics were not made available to make additional comparisons including assessing polygenic overlap with risk scores.

Previous expression studies show that *SLCO5A1* upregulates gene sets implicated in cell adhesion, synapse assembly and organisation, principally belonging to the cadherin superfamily[7]; and the enrichment for presynaptic assembly and organisation pathway in our dataset includes genes encoding trans-synaptically interacting proteins such as *NRXN1-NLGN1* and *PTPRD-IL1RAP1L* that are implicated in a wide range of neuropsychiatric disorders[49,50]. Genetic correlation between ADHD and the BIS-Brief score suggests converging genetic influences across ADHD and epilepsy. Taken together, these results support an important role for specific cell recognition molecules in the organisation of synaptic connections that have an influence in early adulthood as a mechanism for variation in impulsivity across health and disease[51].

While prefrontal-striatal inhibitory control networks are implicated in impulse control, specifically between mPFC and nucleus accumbens[18,20], a role for these limbic networks has only been hinted at in epilepsy. Striato-nigral circuits, preferentially involving the ventral striatum, have long ago been implicated in the *regulation* of generalised seizures in rodent models of generalised epilepsy[19]. Recently, an *initiating* role for cortico-striatal networks in absence seizures with generalized spike-and-wave discharges has been shown in the mouse model of the genetic epilepsy caused by haploinsufficiency of *STXBP1*[52], specifically by reduced cortical excitatory transmission onto striatal fast spiking interneurons. The impulsivity-like startling and the seizure-like phenotype of the *SLCO5A1/Oatp30B* knockdown in *Drosophila* suggests the genetic co-causality of impulsivity and seizures. This supports the idea that excitatory-inhibitory imbalance in the prefrontal-striatal network may predispose simultaneously to epilepsy and impulsivity substrates and invites new approaches to neuromodulation of generalised seizures.

Methods

Human Participants

We collected clinical and genetic data from the Biology of Juvenile Myoclonic Epilepsy (BIOJUME) consortium study (n=864)[25]. We obtained informed consent from all participants and ethical approval from UK Health Research Authority: South Central Oxford C Research Ethics Committee (16/SC/0266) and all other collaborating sites. The SickKids Research Ethics Committee of The Hospital for Sick Children (1000033784) gave ethical approval for this work.

Barratt Impulsivity Scale - Brief

We collected self-rating of trait impulsivity through the BIS-brief[3,30], an eight-item scale generating a total score ranging from 8 to 32.

Genotyping quality control

We genotyped participants' DNA in four batches (n=702) using the Illumina Omni 2.5 array and performed quality control (QC) using PLINK v1.90b6.18[53] and custom in-house scripts. Briefly, we removed individuals and variants with call rates below 90%; samples with sex-gender mismatches and/or high heterozygosity; males with heterozygous calls for X chromosome markers (non-pseudoautosomal region); and females with non-missing calls for markers on the Y chromosome. We retained heterozygous calls for mitochondrial markers in both sexes (i.e. due to heteroplasmy). We obtained an unrelated sample by using KING v.2.2.4 software's[54] --unrelated option (that is, those with estimated kinship coefficient less than 0.088). We corrected and updated the ped file with all found relationships, and identified markers with Mendelian errors using PEDSTATS 0.6.12[55]. We flagged 399 markers but did not remove those out of Hardy-Weinberg Equilibrium ($p < 10^{-4}$). We conducted principal component analysis adjusted using the kinship matrix output by KING using PC-AiR in the GENESIS v2.16.0 package[56].

We performed quality control on each genotyping batch separately, followed by removal of ambiguous A/T, G/C SNPs, chr0 SNPs, indels, monomorphic variants, and duplicate variants; and performed strand alignment using Will Rayner's alignment files (https://www.well.ox.ac.uk/~wrayner/strand/), then merged all batches. We re-analysed and removed cryptic relationships across batches. The final merged set contained 1,489,917 variants (we removed a large number of monomorphic variants), 695 individuals (241 males, 454 females) including 23 related pairs (for association analyses however, an unrelated set is selected).

Genotype imputation

We used the McCarthy Tools v4.3.0 to prepare the genotype data for imputation (https://www.well.ox.ac.uk/~wrayner/tools/HRC-1000G-check-bim-v4.3.0.zip) using TOPMED as the reference panel (r2@1.0.0) on the TOPMED imputation server[57-59]. We converted coordinates from hg37 to hg38 coordinates using strand files (https://www.well.ox.ac.uk/~wrayner/strand/InfiniumOmni2-5-8v1-4_A1-b38-strand.zip). We merged the pseudoautosomal region (PAR) using PLINK's --merge-x option and checked variants using the HRC checking tool. We removed a total of 282,660 variants due to no matches in the reference (but still analyzed for association with BIS afterwards), and 1,739,329 variants remained for imputation on the server. We used Eagle v2.4 for phasing, and minimac v4 v1.0.2 for imputation. We kept variants with imputation quality score $r^2>0.4$ and minor allele frequency (MAF) > 1% for analysis. A total of 8,950,360 variants remained for association analysis.

Genome-wide association analysis

We included for analysis 381 individuals who passed phenotype QC with complete BIS-Brief rating. From these, four failed genotyping QC, and one individual was removed due to cryptic relatedness (n=376). The mega-GWAS analysis consisted of a total of 372 unrelated individuals adjusted for sex, genotyping batch, and population stratification (Supplementary Fig. 2). The mega-GWAS was used for colocalization analysis of the genome-wide association peak on chromosome 8. We identified 329 patients as European ancestry (defined as within 6 standard deviations from the 1000 Genomes[60] European cluster in a principal component analysis). Among these, four patients were missing information on seizure frequency, so we used 324 individuals for the genome-wide association analysis. We adjusted for sex, genotyping batch, age at consent, population stratification, and the frequency of myoclonus or absence seizures. All analyses were conducted on the European GWAS unless noted otherwise. Chromosome X (non-pseudoautosomal region) was analysed with males coded as zero for the reference allele and two for the alternate allele.

Gene enrichment analysis

Variants with $p \leq 5 \times 10^{-4}$ were annotated to the gene with the nearest transcription start site using the Ensembl Variant Effect Predictor (v94)[39], which mapped to 855 unique genes. This gene set was used as input both in a gene ontology (GO) enrichment analysis[40-42] and in FUMA's GENE2FUNC tool[43] (v1.3.7), to test for enrichment in annotated pathways. The GO enrichment analysis revealed enrichment of nine presynaptic assembly genes in our dataset, which were followed up with further analyses.

Phenome-wide association study (PheWAS) analysis

We queried the top associated genome-wide variant and the top associated variant for each of the nine presynaptic assembly enriched genes across PheWAS databases: GWAS Atlas (https://atlas.ctglab.nl/), Global Biobank Engine[61], PheWeb[62], and Gene Atlas[63].

We used PheWeb portals:

- UK Biobank: https://pheweb.org/MGI-freeze2/
- Oxford Brain Imaging Genetics (BIG) Project: http://big.stats.ox.ac.uk/
- fastGWA: https://yanglab.westlake.edu.cn/resources/ukb_fastgwa/imp/
- https://pheweb.org/UKB-SAIGE/

Polygenic risk score (PRS) analysis

Clumping and thresholding was used to calculate attention deficit hyperactivity disorder (ADHD), bipolar disorder, generalized epilepsy and focal epilepsy PRS in individuals of European ancestry using PLINK v1.9[53]. Four PRS were considered and calculated, corresponding to a Bonferroni-corrected p-value for significance of 0.05/4=0.0125. The source of summary statistics used, variant

filtering, clumping and thresholding details are summarized in Extended Data Table 2. PRS values were generated by weighting selected SNPs after clumping and thresholding by the additive scale effect ($\log_{10}$ OR), and then summing over the variants. The PRS values were then centred to the mean. Association of PRSs with BIS-Brief was tested using linear regression with age, sex and frequency of absence/myoclonic seizure as covariates in the model.

Colocalization analysis

We used the Simple Sum 2[34] and COLOC2[64] colocalization methods as implemented in LocusFocus[65] (v1.4.9) to test for colocalization of the genome-wide peaks with eQTL analyses in brain tissues in GTEx v8[12], PsychENCODE[32], and fetal brains[33]. For the genome-wide associated locus on chromosome 8, we performed colocalization analysis on both the mega-GWAS and Europeans-only GWAS (Supplementary Fig. 2). The required significance threshold, after multiple testing of all colocalization datasets analyzed was 0.01.

Domain architecture of SLCO5A1

A BLAST search against the entire Protein Data Base (PDB) identified only one hit with a convincingly high E-value (1e-55) that pointed to the Chain L of the Kazal-like domain containing mice protein (7EEB). The search had a 26% identity and a coverage of 74%. After this hit, the other four identified sequences had E-values > 0.002, clearly distinguishing between significant and non-significant hits. 7EEB is a large complex containing several subunits, among which is *SLCO6C1,* which is the region scoring for *SLCO5A1*.

Phenotypic variance explained

To assess the phenotypic variance explained (PVE) by a SNP or a group of SNPs, we calculated the partial r^2 as the proportion of the residual sum of squares (RSS) reduced when adding the SNP (or group of SNPs) to the base regression model with all covariates.

siRNA probe design and knockdown of *Oatp30B* in *Drosophila melanogaster*

Drosophila

Flies were maintained and crossed at 18°C. All ageing was done in a controlled environment of 29°C and 60% humidity.

Stocks

ubiGal80^{ts} // *UAS-Oatp30B^{IR}* (GD12775) obtained from the VDRC // *w^{1118}*, *nSybGal4, TubGal4* and *UAS-GFP^{IR}* obtained from the BDSC.

Lifespan

Lifespan analysis was performed as previously reported[38]. All crosses were maintained at 18°C during the developmental stages of the progeny. Newly eclosed adult flies were collected within 5 days at 18°C. Females and males were pooled together and equally distributed within vials.

Motor behaviour assay

Single fly tracking was carried out as previously described[38]. In each of 3 experiments, up to 12 flies per genotype, aged 15 days (adult stage) at 29°C to allow RNAi expression and knock-down, were placed into individual round 6-wells arenas. The protocol used consisted of 6 stimuli events equally split during a period of 2 h and 15 min, the first one starting after 30 min of recording and the last one 30 min before the end of the protocol. Each stimuli event was composed of 5 vibrations of 200 ms spaced by 500 ms. The x/y position of each single fly was tracked and analysed using the DART software in order to evaluate the relative speed and activity before, during and after the stimuli event. The speed analysis was used for the "Stimuli Response Trace" and the general activity used to deduce "Active Speed", "Mean Bout Length" and "Inter-Bout Interval", using a custom-made modification of the DART software[37].

Heat-induced seizure assay

Flies aged 15 days at 29°C to allow RNAi expression and knock-down were isolated into new plastic vials without food for 10-20 min before immersion in a 42°C water bath for 120 seconds. Each tube was video recorded during and post immersion and seizures were defined as a period of brief leg twitches, convulsions and failure to maintain standing posture. Flies were, thereafter, allowed to recover at room temperature and the time to recover from seizure was calculated only for flies that had undergone seizures. All experiments were randomised and double-blinded.

RNA extraction and qPCR

RNA was extracted as previously reported[66] from 15 adult flies of both sexes, aged 15 days at 29°C to allow RNAi expression and knock-down, using TriZol (Thermo-Fischer). cDNA was generated using SuperScript III Reverse Transcriptase (Thermo-Fischer). Quantitative PCR was performed in combination with qPCRBIO SyGreen Blue mix (PCR Biosystems) on Quantstudio 7 from real-time PCR system (Thermo-Fischer). *eIF4a* was used as housekeeping control. The following oligos were used: *Oatp30B* Fw (GAATCCGACCAACCGCCTGA), *Oatp30B* Rv (ATGGATTCCTGCCGCCTGTG), *eIF4a* Fw (CGTGAAGCAGGAGAACTGG), *eIF4a* Rv (CATCTCCTGGGTCAGTTG).

Data Availability

eQTL data are available for download for GTEx (https://gtexportal.org/home), PsychENCODE (http://resource.psychencode.org/), and fetal brains

(https://doi.org/10.6084/m9.figshare.6881825). GWAS summary statistics for this study are available for download from our website (https://lab.research.sickkids.ca/strug/softwareandresources/).

Author Contributions

LJS and DKP contributed to conception and study design. NP, DR, AS, AC, AH, LJS and DKP contributed to data management and project administration. DA, JPC, CPB, CYF, JG, DAG, CD, FM, KH, KSL, JK, AO, MR, KKS, GR, MS, IT, RT, JZ, MPR, LJS and DKP contributed to acquisition of study data. NP, DR, AS, MF, LJS, and DKP contributed to analysis of data. NP, DR, AS, MF, LJS, and DKP contributed to drafting the manuscript. Members of the BIOJUME consortium are listed in the appendix.

Conflicts of Interest

DA, KKS, RT, and JZ report honoraria from UCB Pharma (manufacturer of levetiracetam) and RT reports honoraria from Sanofi (manufacturer of sodium valproate). KH reports honoraria from UCB Pharma, Eisai and GW Pharma. MS reports honoraria from UCB Pharma and Eisai. GR reports honoraria from UCB Pharma (manufacturer of levetiracetam), from EISAI (manufacturer of perampanel), from Angelini Pharma (manufacturer of cenobamate). All other authors report no conflicts of interest.

Acknowledgements: This work was supported by the Canadian Institutes of Health Research: Biology of Juvenile Myoclonic Epilepsy 201503MOP-342469 (DKP, LJS) and 201809FDN-407295 (LJS); UK Medical Research Council, Centre for Neurodevelopmental Disorders MR/N026063/1 (DKP, MPR); UK Medical Research Council, Programme Grant MR/K013998/1, (MPR); PhD stipend from UK Medical Research Council and the Sackler Institute for Translational Neurodevelopment (AS); NIHR Specialist Biomedical Research Centre for Mental Health of South London and Maudsley National Health Service Foundation Trust (DKP, MPR); UK Engineering and Physical Sciences Research Council, Centre for Predictive Modelling in Healthcare (EP/N014391/1 (MPR)); DINOGMI Department of Excellence of MIUR 2018–2022 (legge 232 del 2016 (PS)); Wales BRAIN Unit and Research Delivery Staff funded by Welsh Government through Health and Care Research Wales (KH); Biomarin srl, ENECTA srl, GW Pharmaceuticals, Kolfarma srl. and Eisai (PS); South-Eastern Regional Health Authority, Norway (Project Number 2016129 (JK)); The Research Council of Norway (Project Number 299266 (MS)); Epilepsy Research UK (RT, MR); Health & Care Research Wales (MR), Wales Gene Park (MR), Abertawe Bro Morgannwg University NHS R&D (MR); UCB (GR); Nationwide Children's Hospital (DAG); Odense University Hospital (JG); University of Southern Denmark (17/18517 (CPB)); Grants NC/V001051/1 from the NC3Rs (MF), European Union's Horizon 2020 Research and Innovation Programme (765912 - DRIVE - H2020-MSCA-ITN-2017 (HJ)) and Action Medical Research (GN2446 (HJ, MF)).

References

1. Dalley, J.W. & Robbins, T.W. Fractionating impulsivity: neuropsychiatric implications. *Nat Rev Neurosci* **18**, 158-171 (2017).
2. Ramirez-Martin, A., Ramos-Martin, J., Mayoral-Cleries, F., Moreno-Kustner, B. & Guzman-Parra, J. Impulsivity, decision-making and risk-taking behaviour in bipolar disorder: a systematic review and meta-analysis. *Psychol Med* **50**, 2141-2153 (2020).
3. Shakeshaft, A. *et al.* Trait impulsivity in Juvenile Myoclonic Epilepsy. *Ann Clin Transl Neurol* (2020).
4. Smith, A., Syvertsen, M. & Pal, D.K. Meta-analysis of response inhibition in juvenile myoclonic epilepsy. *Epilepsy Behav* **106**, 107038 (2020).
5. Wandschneider, B. *et al.* Risk-taking behavior in juvenile myoclonic epilepsy. *Epilepsia* **54**, 2158-65 (2013).
6. Hagenbuch, B. & Meier, P.J. The superfamily of organic anion transporting polypeptides. *Biochim Biophys Acta* **1609**, 1-18 (2003).
7. Sebastian, K. *et al.* Characterization of SLCO5A1/OATP5A1, a solute carrier transport protein with non-classical function. *PLoS One* **8**, e83257 (2013).
8. Daruna, J.H. & Barnes, P.A. A neurodevelopmental view of impulsivity and its relationship to the superfactors of personality. in *The impulsive client: theory, research and treatment* (eds. McCown, W.G., Johnson, J.L. & Shure, M.B.) 23-37 (American Psychological Association, 1993).
9. Sanchez-Roige, S. *et al.* Genome-Wide Association Studies of Impulsive Personality Traits (BIS-11 and UPPS-P) and Drug Experimentation in up to 22,861 Adult Research Participants Identify Loci in the CACNA1I and CADM2 genes. *J Neurosci* **39**, 2562-2572 (2019).
10. Barbeira, A.N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun* **9**, 1825 (2018).
11. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-7 (2014).
12. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-5 (2013).
13. O'Muircheartaigh, J. *et al.* Focal structural changes and cognitive dysfunction in juvenile myoclonic epilepsy. *Neurology* **76**, 34-40 (2011).
14. Keller, S.S. *et al.* Microstructural and volumetric abnormalities of the putamen in juvenile myoclonic epilepsy. *Epilepsia* **52**, 1715-24 (2011).
15. Landvogt, C., Buchholz, H.-G., Bernedo, V., Schreckenberger, M. & Werhahn, K.J. Alteration of dopamine D2/D3 receptor binding in patients with juvenile myoclonic epilepsy: Alteration of Dopamine D2/D3 Receptor Binding in JME. *Epilepsia* **51**, 1699-1706 (2010).
16. Ciumas, C. *et al.* Reduced dopamine transporter binding in patients with juvenile myoclonic epilepsy. *Neurology* **71**, 788-94 (2008).
17. Dalley, Jeffrey W., Everitt, Barry J. & Robbins, Trevor W. Impulsivity, Compulsivity, and Top-Down Cognitive Control. *Neuron* **69**, 680-694 (2011).
18. Dalley, J.W. & Roiser, J.P. Dopamine, serotonin and impulsivity. *Neuroscience* **215**, 42-58 (2012).
19. Deransart, C., Vercueil, L., Marescaux, C. & Depaulis, A. The role of basal ganglia in the control of generalized absence seizures. *Epilepsy Res* **32**, 213-23 (1998).

20. Cho, S.S. *et al.* Morphometric correlation of impulsivity in medial prefrontal cortex. *Brain Topogr* **26**, 479-87 (2013).
21. Santos, B.P.D. *et al.* Genetic susceptibility in Juvenile Myoclonic Epilepsy: Systematic review of genetic association studies. *PLoS One* **12**, e0179629 (2017).
22. International League Against Epilepsy Consortium on Complex Epilepsies. Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nat Commun* **9**, 5269 (2018).
23. Bai, D. *et al.* DNA variants in coding region of EFHC1: SNPs do not associate with juvenile myoclonic epilepsy. *Epilepsia* (2008).
24. Bailey, J.N. *et al.* Variant Intestinal-Cell Kinase in Juvenile Myoclonic Epilepsy. *N Engl J Med* **378**, 1018-1028 (2018).
25. Shakeshaft, A. *et al.* Sex-specific disease modifiers in juvenile myoclonic epilepsy. *Sci Rep* **12**, 2785 (2022).
26. Li, M. *et al.* Antisense oligonucleotide therapy reduces seizures and extends life span in an SCN2A gain-of-function epilepsy model. *J Clin Invest* **131**(2021).
27. Karalis, V. & Bateup, H.S. Current Approaches and Future Directions for the Treatment of mTORopathies. *Dev Neurosci* **43**, 143-158 (2021).
28. Hall, M.-H. & Smoller, J.W. A New Role for Endophenotypes in the GWAS Era: Functional Characterization of Risk Variants. *Harv Rev Psychiatry* **18**, 67-74 (2010).
29. Manchia, M. *et al.* The Impact of Phenotypic and Genetic Heterogeneity on Results of Genome Wide Association Studies of Complex Diseases. *PLoS One* **8**, e76295 (2013).
30. Steinberg, L., Sharp, C., Stanford, M.S. & Tharp, A.T. New tricks for an old measure: the development of the Barratt Impulsiveness Scale-Brief (BIS-Brief). *Psychol Assess* **25**, 216-26 (2013).
31. Kinirons, P. *et al.* A novel locus for idiopathic generalized epilepsy in French-Canadian families maps to 10p11. *Am J Med Genet A* **146A**, 578-84 (2008).
32. Wang, D. *et al.* Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**(2018).
33. O'Brien, H.E. *et al.* Expression quantitative trait loci in the developing human brain and their enrichment in neuropsychiatric disorders. *Genome Biol* **19**, 194 (2018).
34. Wang, F., Panjwani, N., Wang, C., Sun, L. & Strug, L.J. A flexible summary statistics-based colocalization method with application to the mucin cystic fibrosis lung disease modifier locus. *Am J Hum Genet* **109**, 253-269 (2022).
35. BrainSpan Atlas of the Developing Human Brain [Internet].
36. Sunkin, S.M. *et al.* Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Res* **41**, D996-D1008 (2013).
37. Faville, R., Kottler, B., Goodhill, G.J., Shaw, P.J. & van Swinderen, B. How deeply does your mutant sleep? Probing arousal to better understand sleep defects in Drosophila. *Sci Rep* **5**, 8454 (2015).
38. Mazaud, D. *et al.* Transcriptional Regulation of the Glutamate/GABA/Glutamine Cycle in Adult Glia Controls Motor Activity and Seizures in Drosophila. *J Neurosci* **39**, 5269-5283 (2019).
39. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122 (2016).
40. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-9 (2000).
41. Carbon, S. *et al.* AmiGO: online access to ontology and annotation data. *Bioinformatics* **25**, 288-9 (2009).
42. Gene Ontology, C. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res* **49**, D325-D334 (2021).

43. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **8**, 1826 (2017).
44. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425 (2015).
45. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766-70 (2008).
46. Mikkelsen, T.S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-60 (2007).
47. Lawlor, D.A., Tilling, K. & Davey Smith, G. Triangulation in aetiological epidemiology. *Int J Epidemiol* **45**, 1866-1886 (2016).
48. Parker, L., Howlett, I.C., Rusan, Z.M. & Tanouye, M.A. Seizure and epilepsy: studies of seizure disorders in Drosophila. *Int Rev Neurobiol* **99**, 1-21 (2011).
49. Uhl, G.R. & Martinez, M.J. PTPRD: neurobiology, genetics, and initial pharmacology of a pleiotropic contributor to brain phenotypes. *Ann N Y Acad Sci* **1451**, 112-129 (2019).
50. Hu, Z., Xiao, X., Zhang, Z. & Li, M. Genetic insights and neurobiological implications from NRXN1 in neuropsychiatric disorders. *Mol Psychiatry* **24**, 1400-1414 (2019).
51. Sanes, J.R. & Zipursky, S.L. Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. *Cell* **181**, 536-556 (2020).
52. Miyamoto, H. *et al.* Impaired cortico-striatal excitatory transmission triggers epilepsy. *Nat Commun* **10**, 1917 (2019).
53. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
54. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867-73 (2010).
55. Wigginton, J.E. & Abecasis, G.R. PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics* **21**, 3445-7 (2005).
56. Conomos, M.P., Miller, M.B. & Thornton, T.A. Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. *Genet Epidemiol* **39**, 276-93 (2015).
57. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat Genet* **48**, 1284-1287 (2016).
58. Fuchsberger, C., Abecasis, G.R. & Hinds, D.A. minimac2: faster genotype imputation. *Bioinformatics* **31**, 782-4 (2015).
59. Taliun, D. *et al.* Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* **590**, 290-299 (2021).
60. Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).
61. McInnes, G. *et al.* Global Biobank Engine: enabling genotype-phenotype browsing for biobank summary statistics. *Bioinformatics* **35**, 2495-2497 (2019).
62. Gagliano Taliun, S.A. *et al.* Exploring and visualizing large-scale genetic associations by using PheWeb. *Nat Genet* **52**, 550-552 (2020).
63. Canela-Xandri, O., Rawlik, K. & Tenesa, A. An atlas of genetic associations in UK Biobank. *Nat Genet* **50**, 1593-1599 (2018).
64. Dobbyn, A. *et al.* Landscape of Conditional eQTL in Dorsolateral Prefrontal Cortex and Co-localization with Schizophrenia GWAS. *Am J Hum Genet* **102**, 1169-1184 (2018).
65. Panjwani, N. *et al.* LocusFocus: Web-based colocalization for the annotation and functional follow-up of GWAS. *PLoS Comput Biol* **16**, e1008336 (2020).
66. Napoletano, F. *et al.* Polyglutamine Atrophin provokes neurodegeneration in Drosophila by repressing fat. *EMBO J* **30**, 945-58 (2011).

67. Pao, S.S., Paulsen, I.T. & Saier, M.H., Jr. Major facilitator superfamily. *Microbiol Mol Biol Rev* **62**, 1-34 (1998).
68. Walmsley, A.R., Barrett, M.P., Bringaud, F. & Gould, G.W. Sugar transporters from bacteria, parasites and mammals: structure-activity relationships. *Trends Biochem Sci* **23**, 476-81 (1998).
69. Madej, M.G., Dang, S., Yan, N. & Kaback, H.R. Evolutionary mix-and-match with MFS transporters. *Proc Natl Acad Sci U S A* **110**, 5870-4 (2013).
70. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet* **51**, 63-75 (2019).
71. Stahl, E.A. *et al.* Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* **51**, 793-803 (2019).
72. International League Against Epilepsy Consortium on Complex, E. Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nat Commun* **9**, 5269 (2018).
73. Leu, C. *et al.* Polygenic burden in focal and generalized epilepsies. *Brain* **142**, 3473-3481 (2019).

Figures

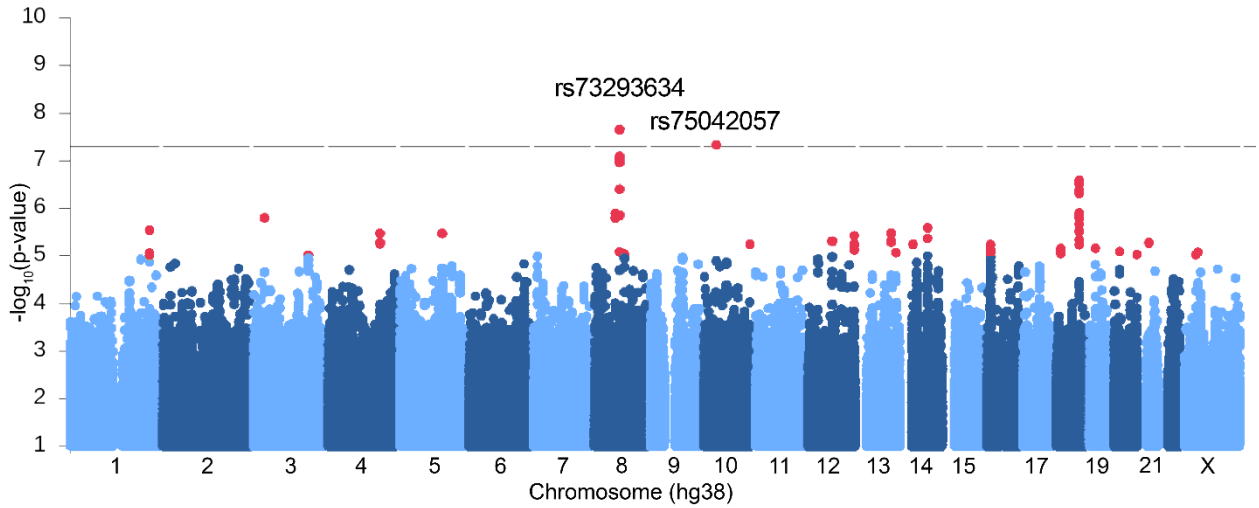


Fig. 1: Manhattan plot showing GWAS with BIS-Brief score. We found two significant genome-wide associations on chromosome 8 (rs73293634 (G/T)) and 10 (rs75042057 (T/G)) in the analysis of 324 European individuals with JME. Variants below $-\log_{10}P < 1$ were omitted in the plot.

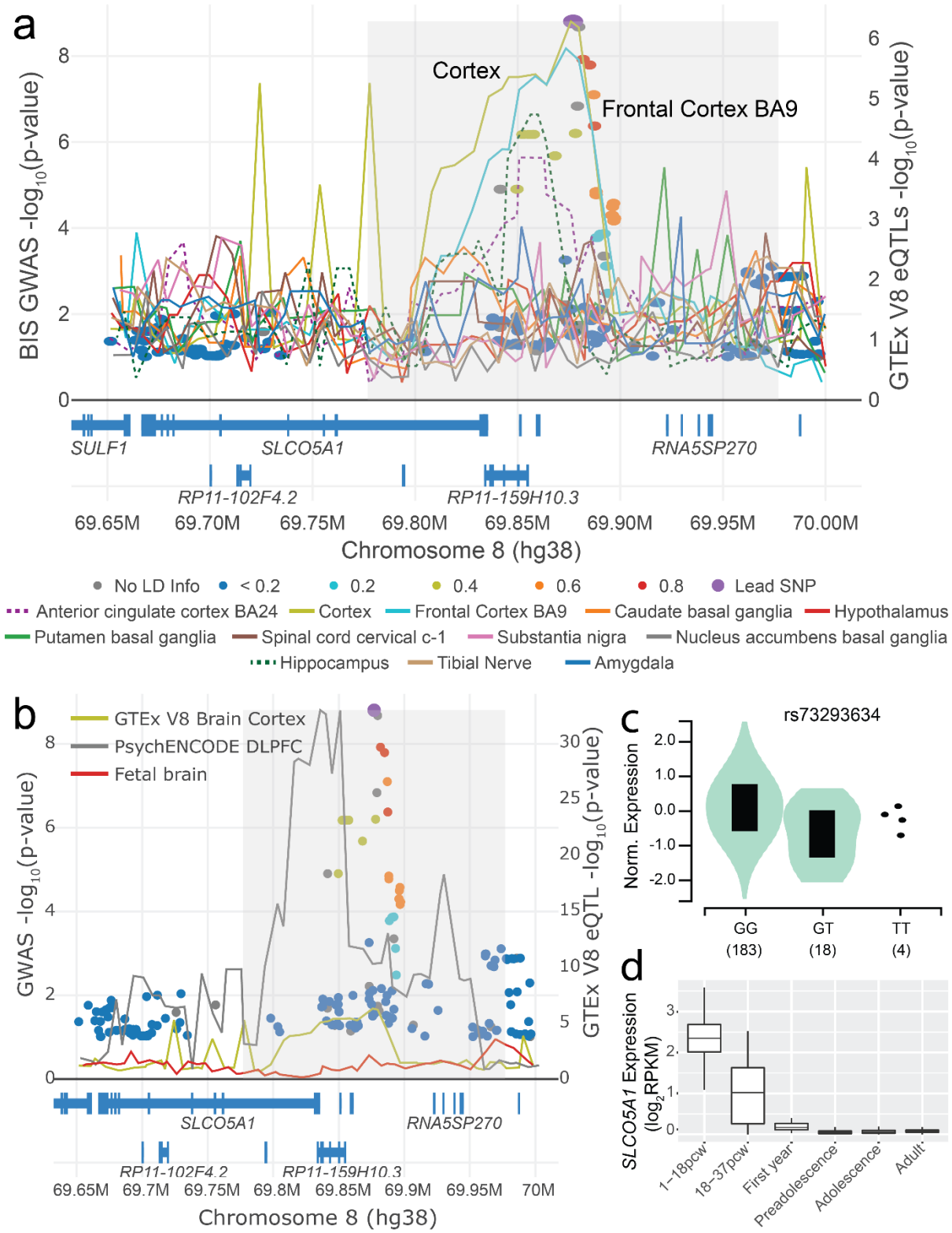


Fig. 2: **LocusFocus[65] plot for the GWAS with BIS (dots) and eQTLs in GTEx[12] brain and tibial nerve tissues for the *SLCO5A1* gene (lines). a**, Colocalization figure from LocusFocus for the *SLCO5A1* gene. Lines depict the minimum p-value trace in a sliding window for *SLCO5A1* eQTLs from GTEx, one line per tissue. Circles depict the GWAS with BIS, with the lead SNP in purple and pairwise LD with the lead SNP marked as shown in the legend, calculated using the 1000 Genomes Project[60] European subset. Significant colocalization is observed for *SLCO5A1* eQTLs in GTEx v8 for the cerebral cortex after increasing sample size in a mega-GWAS (n=367, $-\log_{10}$ Simple Sum 2^{34} P-value = 9.5×10^{-3}). Colocalization analysis with only the Europeans is provided in Extended Data Fig. 1. Colocalization was also tested for all other nearby genes shown in the figure, but no other genes' eQTLs colocalized with BIS GWAS (not shown). **b,** Colocalization analysis with PsychENCODE eQTLs in the dorsolateral prefrontal cortex (DLPFC) (n=1,866)[32], and

eQTLs derived from second trimester fetal brains (n=120)[33], with GTEx's brain cortex eQTL as in A provided for reference. Colocalization analysis results suggest no colocalization with either PsychENCODE (Simple Sum 2 P-value = 0.985) or fetal brain eQTLs (does not pass first stage test in Simple Sum 2 for having significant eQTLs in the region). **c,** Violin plot for the eQTL effect of rs73293634 SNP on *SLCO5A1* expression in the cerebral cortex from GTEx v8. **d**, Expression change of *SLCO5A1* from brains in various developmental stages from BrainSpan[35,36]. pcw, post conception weeks; preadolescence, 2-12 years old (inclusive); adolescence, 13-19 years old; adult, ≥ 20 years old (oldest samples are 40 years old).

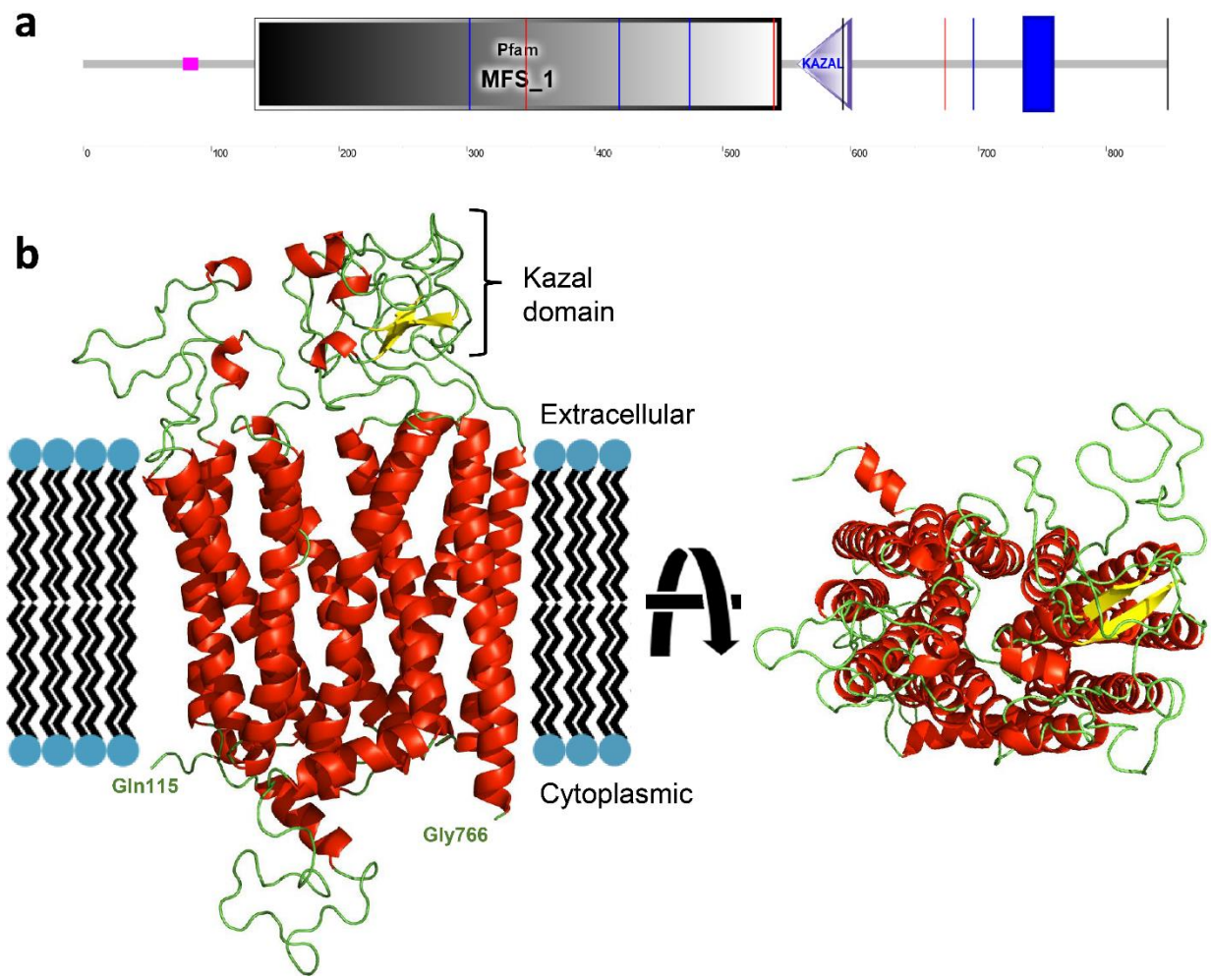


Fig. 3: Domain architecture of human SLCO5A1. a, Schematic representation of the protein with the indication of recognised domains. A SMART analysis to identify structural domains confirmed the presence of two modules, Major Facilitator Superfamily (MFS) and a Kazal domain, interspaced with potentially unstructured sequences. The MFS transporters are membrane proteins capable of transporting small solutes in response to chemiosmotic ion gradients[67,68]. They are represented in many organisms from Archaea to Homo sapiens. MFS proteins target a wide range of substrates, including ions, carbohydrates, lipids, amino acids and peptides, nucleosides and other small molecules and transport them in both directions across the membrane[69]. The Kazal domain is an evolutionary conserved module usually acting as a serine-protease inhibitor. **b**, Predicted model of the monomeric form of SLCO5A1 from amino acids 115-766, built using the SwissModel homology server (https://swissmodel.expasy.org) and utilising the template structure pdb:7eeb. Red: alpha helices; Yellow: Beta strands; Green: Loops.

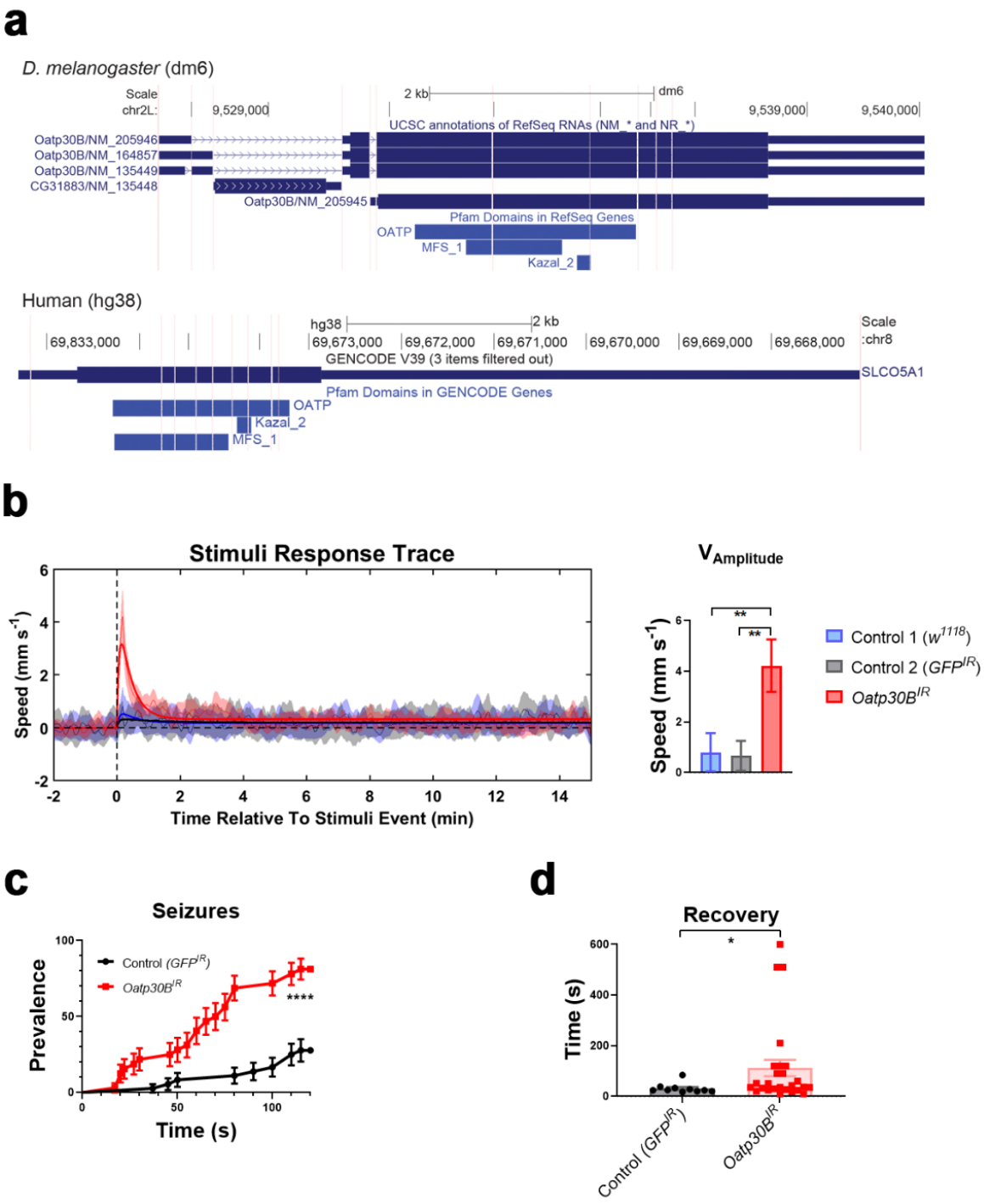


Fig. 4: a, UCSC browser exon view of *Oatp30B* in *D. melanogaster* (dm6 assembly) and *SLCO5A1* in human (hg38 assembly). Pfam tracks show Oatp30B and SLCO5A1 have conserved MFS, OATP and Kazal domains. **b, Startling reaction to trains of vibrations in flies with *Oatp30B* knock down.** The *UAS-Oatp30B^{IR}* (GD12775) transgenic or the control *UAS-GFP^{IR}* were driven with *nSyb-Gal4* and *Ubi-Gal80ts*. The *w^{1118}* strain is a control for the genetic background in absence of transgenes. Mean +/- SEM ** p<0.01, One Way ANOVA, Tukey's post-hoc test. Units are the vibration events experienced 6 times for each fly. n=174-210. **c, Increased seizure prevalence in flies with *Oatp30B* knock down.** The *UAS-Oatp30B^{IR}* (GD12775) transgenic or the control *UAS-GFP^{IR}* were driven with *nSyb-Gal4* and *Ubi-Gal80ts*. Percent +/- SE **** p<0.0001, Log-rank (Mantel-Cox) test, χ^2 24.68 for 1 df, n=34-36. **d, Increased post-seizure recovery time in flies with *Oatp30B* knock down.** The *UAS-Oatp30B^{IR}* (GD12775) transgenic or the control *UAS-GFP^{IR}* were driven with *nSyb-Gal4* and *Ubi-Gal80ts*. Mean +/- SEM * p<0.05, Mann Whitney non-parametric test, two tails, n=10-26. Only flies that displayed a seizure within 120 s as in Fig. 4b have been included in the analysis.

Tables

Table 1: Summary of genome-wide associated variants and chr18 region (rs515846) for the GWAS of BIS scores in JME (n=324). All observed sample allele frequencies are comparable to those seen in the European 1000 Genomes (phase 3)[60].

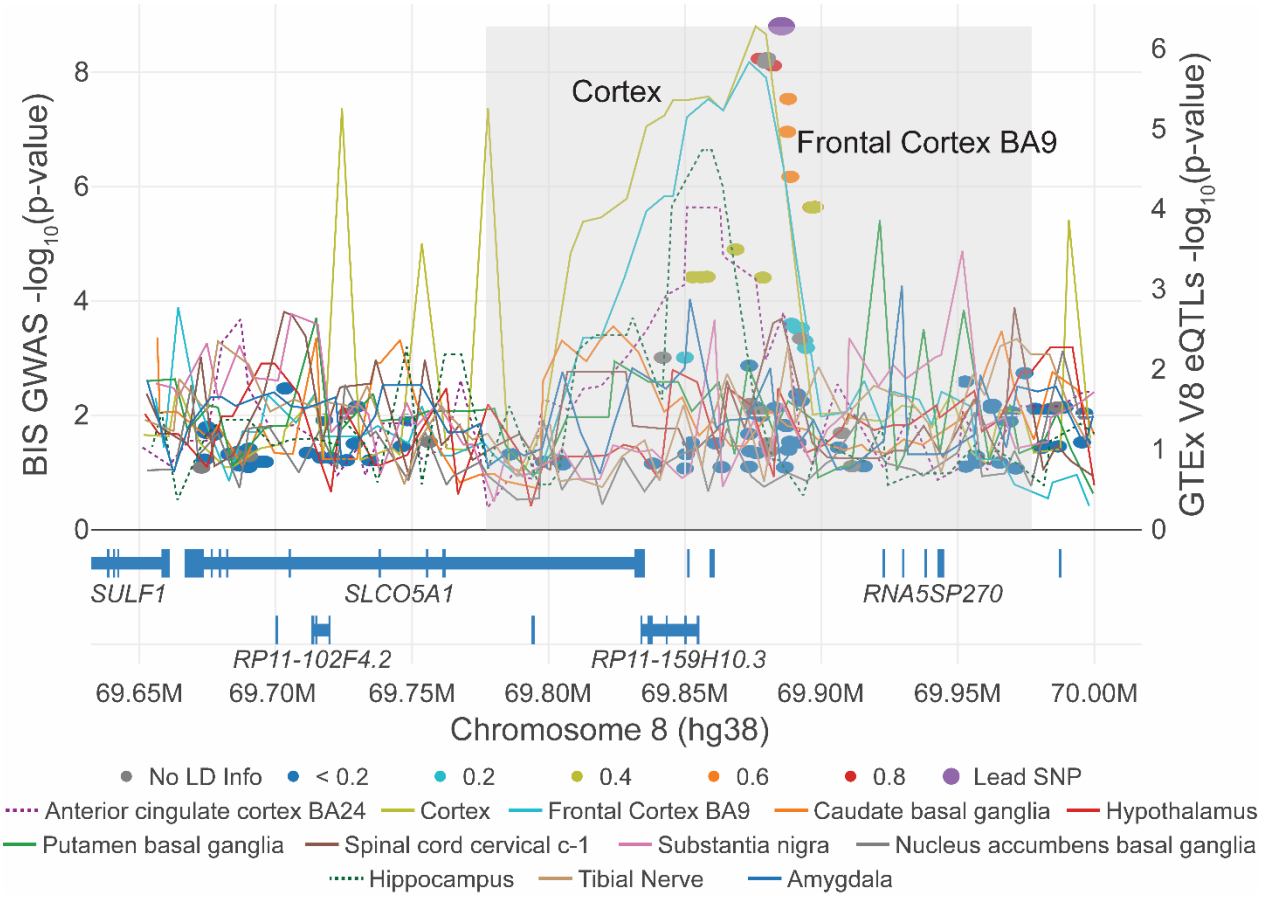
		European GWAS (n=324)				Mega-GWAS (n=372)			
Variant ID (hg38)	Imputation r^2	MAF	Beta	Standard error	P-value	MAF	Beta	Standard error	P-value
chr8:69,884,968* rs73293634 (G/T)	0.961	0.036 (T)	5.42	0.91	7.5×10^{-9}	0.041	4.55	0.79	1.61×10^{-8}
chr8:69,876,965 rs146866040 (A/G)	0.979	0.032 (G)	5.38	0.94	2.5×10^{-8}	0.031	5.60	0.90	1.57×10^{-9}
chr10:34,202,650 rs75042057 (T/G)	0.878	0.019 (G)	7.51	1.33	3.6×10^{-8}	0.022	6.60	1.19	4.99×10^{-8}

*The lead SNP for the mega-GWAS was rs146866040. The LD between them is 0.89.

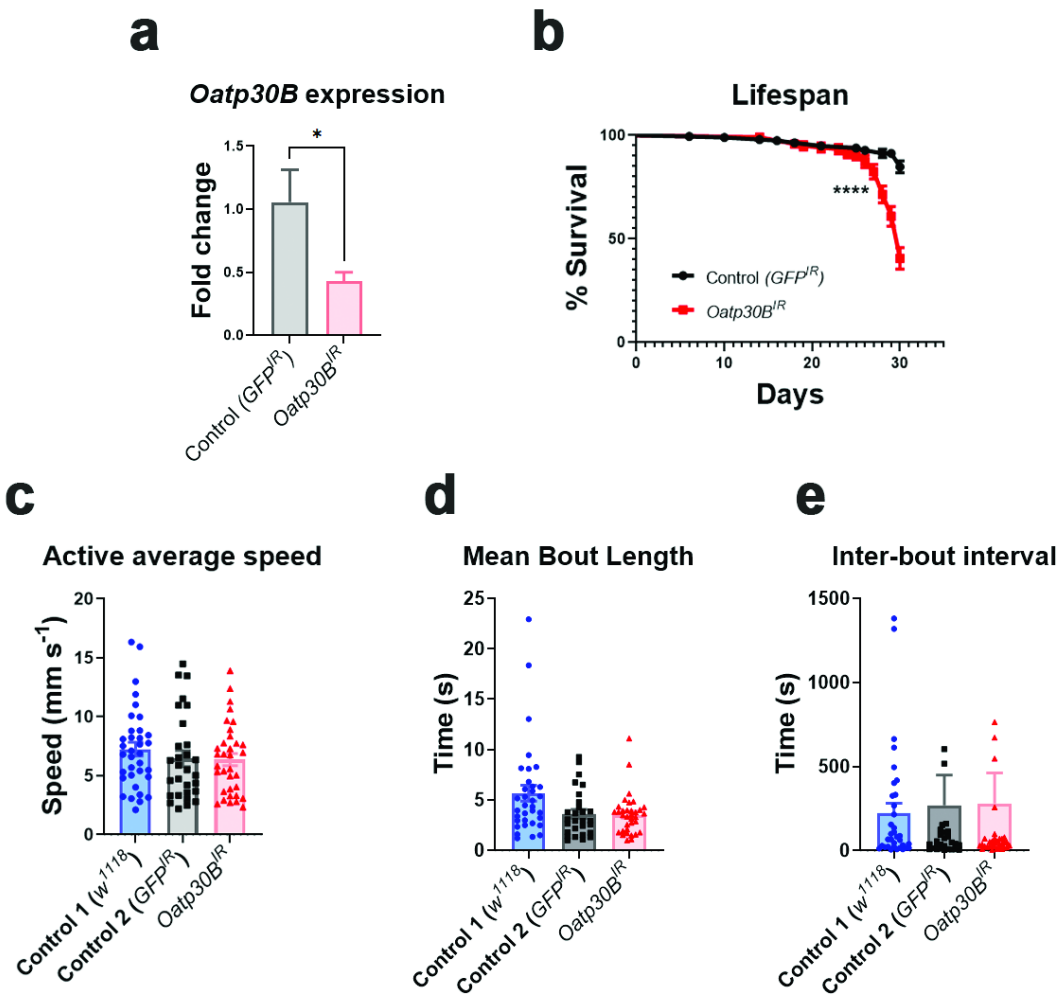
Table 2: List of top variants annotated to the nine presynaptic assembly genes enriched in the European GWAS of BIS in JME (n=324). *PTPRD*, Protein Tyrosine Phosphatase Receptor Type D; *FZD1*, Frizzled Class Receptor 1; *NLGN1*, Neuroligin 1; *NLGN4X*, Neuroligin 4 X-Linked; *PCLO*, Piccolo Presynaptic Cytomatrix Protein; *CNTN5*, Contactin 5; *NRXN1*, Neurexin 1; *IL1RAPL1*, Interleukin 1 Receptor Accessory Protein Like 1; *PTEN*, Phosphatase and Tensin Homolog.

Gene	Rsid	Beta	P-value	PVE
PTPRD	rs1781264	1.827	1.19E-04	0.042
FZD1	rs916716	1.428	8.16E-05	0.051
NLGN1	rs73177088	6.191	9.95E-04	0.044
NLGN4X	rs146813567	-2.898	3.06E-04	0.039
PCLO	rs4728525	-2.311	7.35E-04	0.052
CNTN5	rs150418619	5.806	1.06E-04	0.051
NRXN1	rs1040239	-1.573	1.21E-04	0.045
IL1RAPL1	rs5943492	1.039	8.73E-04	0.043
PTEN	rs112050451	5.158	1.27E-03	0.041

Extended Materials: Figures



Extended Data Fig. 1: Colocalization figure from LocusFocus[65] for the *SLCO5A1* gene. Same as in main Figure 3A, but circles depict the GWAS with BIS in the European subset (n=324). Colocalization analysis results reveal colocalization with GTEx[12] V8 brain cortex ($-\log_{10}$ Simple Sum P-value (SSP) = 1.36), although this colocalization does not pass the multiple testing significance threshold of $-\log_{10}$SSP=2.0 for testing colocalization across eQTLs from GTEx[12], PsychENCODE[32] and fetal brain eQTLs from [33].



Extended Data Fig. 2: a, Oatp30B mRNA levels as measured by qPCR. The *UAS-Oatp30B*^{IR} (GD12775) transgenic or the control *UAS-GFP*^{IR} were driven with *Tub-Gal4* and *Ubi-Gal80ts*. The graph reports data from 3 biological samples (n=10 flies, both males and females) and 5 technical replicates. Mean +/- SEM * P<0.05, Unpaired t-test, one-tailed. **b, Reduced lifespan in flies with *Oatp30B* knock-down.** The *UAS-Oatp30B*^{IR} (GD12775) transgenic or the control *UAS-GFP*^{IR} were driven with *nSyb-Gal4* and *Ubi-Gal80ts*. Percent +/- SE **** P<0.0001, Log-rank (Mantel-Cox) test, χ^2 51.74 for 1 df, n=119-194. **c, Unchanged speed during action.** The *UAS-Oatp30B*^{IR} (GD12775) transgenic or the control *UAS-GFP*^{IR} were driven with *nSyb-Gal4* and *Ubi-Gal80ts*. The *w*^{1118} strain is a control for the genetic background in absence of transgenes. **d, Unaffected duration of single action bouts.** Flies as in c. **e, Unaffected rest interval in between single action bouts.** Flies as in c.

Extended Materials: Tables

Extended Data Table 1: Univariate analysis with BIS. A total of 324 European individuals with JME were tested for association for each of the variables shown with BIS. We have previously shown the association of sex and seizure frequency with BIS (ref). Seizure frequency was used as a marker of controlled seizures and was defined as missing if there was no reported myoclonus frequency; otherwise, it was the maximum observed frequency for myoclonus or absence seizure as follows: daily seizures=3; weekly=2; less than weekly=1; none (currently or ever)=0. Cohort refers to the genotyping batch; samples recruited were genotyped in four batches at different time points.

Variable	Beta	Standard Error	P-value
Sex	-1.36	0.51	8.0×10^{-3}
Age at consent	0.042	0.028	0.13
Myoclonus or absence seizure frequency	1.09	0.25	2.2×10^{-5}
PC1	3.96	5.05	0.43
PC2	-4.48	5.42	0.41
PC3	4.30	4.39	0.33
Cohort 2	1.41	0.78	0.071
Cohort 3	1.49	0.92	0.10
Cohort 4	1.66	0.79	0.035

Extended Data Table 2: Methods applied for the computation of PRS.

Disease	Summary statistics Study	Sample size	Ethnicity	Variant filtering	BIOJUME Variant Filtering	Clumping + Thresholding Clumping	p-value Threshold
ADHD	Demontis et al., 2019[70]	19,099 cases & 34,194 controls	European	MAF ≤0.05 r^2 ≤0.8 Indels Ambiguous SNPs Multi-allelic SNPs MHC region	r^2 ≤0.5	r^2: 0.3 radius: 500 kb	0.1
Bipolar Disorder	Stahl et al., 2019[71]	20,352 cases and 31,358 controls	European	MAF ≤0.05 Info ≤0.9 Indels Ambiguous SNPs Multi-allelic SNPs MHC region except for the most significant associated SNP (rs36034627, Chr6:27,269,584, T>G)	r^2 ≤0.5	r^2: 0.1 radius: 500 kb	0.2
Generalized epilepsy*	ILAE Consortium, 2018[72]	15,212 cases & 29,677 controls	Majority European	MAF ≤0.02 r^2 ≤0.3 Ambiguous SNPs	r^2 ≤0.5	r^2: 0.1 radius: 500 kb	0.5
Focal epilepsy*	ILAE Consortium, 2018[72]	15,212 cases & 29,677 controls	Majority European	MAF ≤0.02 r^2 ≤0.3 Ambiguous SNPs	r^2 ≤0.5	r^2: 0.1 radius: 500 kb	0.5

MHC region: The major histocompatibility complex region, chr6:25-34Mbp

ILAE: International League Against Epilepsy

The methods applied to calculated generalized and focal epilepsy PRS are based on Leu et al. 2019[73].

Appendix: BIOJUME Consortium

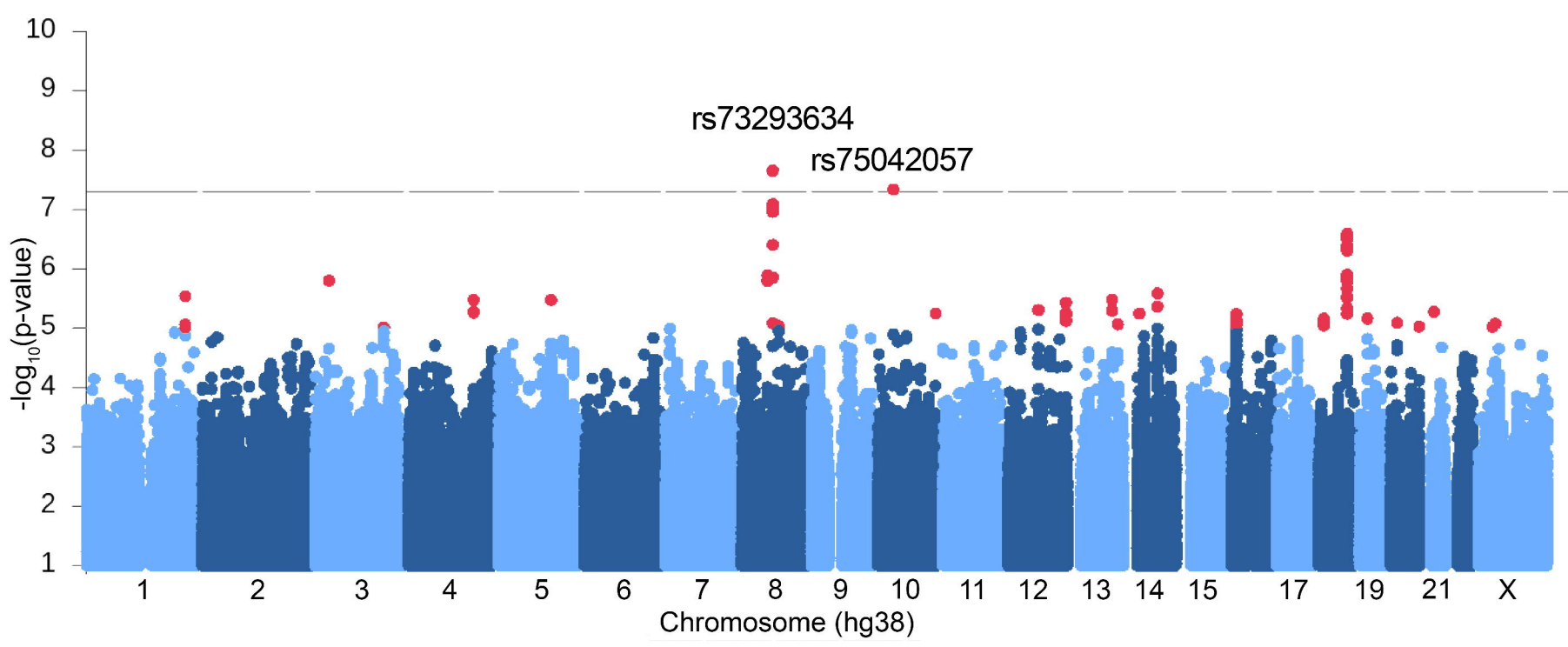
Sites and site investigators included in the BIOJUME consortium.

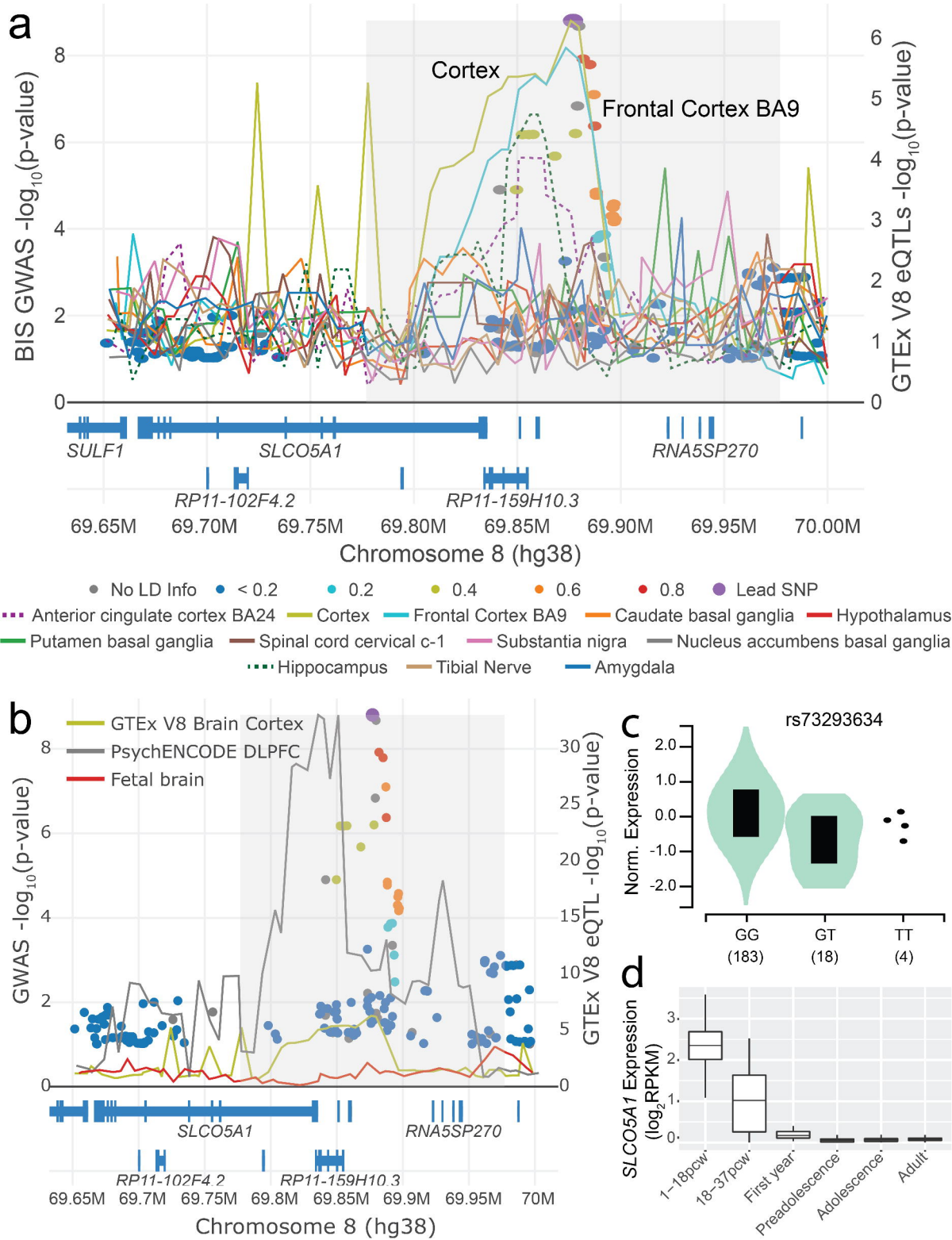
Country	**Site**	**Site principal investigators (PIs) and research staff**
Canada	SickKids Hospital, Toronto	Lisa Strug (PI), Naim Panjwani, Fan Lin
	Toronto Western Hospital	Danielle M. Andrade
Czech Republic	Charles University	Jana Zarubova (PI), Zuzana Šobíšková, Cechovaz Pracoviste, Michaela Kajsova
Denmark	Danish National Epilepsy Centre	Guido Rubboli (PI), Rikke S. Møller, Elena Gardella
	Syddansk Universitet	Christoph P. Beier (PI), Joanna Gesche, Maria Miranda
Estonia	Tallinn Children's Hospital	Inga Talvik (PI)
Italy	Commissione Genetica	Pasquale Striano (PI), Alessandro Orsini
	University of Catania	Andrea Pratico (PI)
Malaysia	University of Malaya	Choong Yi Fong (PI), Ching Ching Ng, Kheng Seang Lim
Norway	Vestre Viken Health Trust	Kaja K. Selmer, Marte Syvertsen (Co-PIs)
UK	Airedale NHS Foundation Trust	Pronab Bala (PI), Amy Kitching
	Ashford and St. Peter's Hospitals NHS Foundation Trust	Kate Irwin (PI), Lorna Walding, Lynsey Adams
	Bradford Teaching Hospitals NHS Foundation Trust	Uma Jegathasan (PI), Rachel Swingler, Rachel Wane
	Brighton and Sussex University Hospitals NHS Trust	Julia Aram (Co-PI), Nikil Sudarsan (Co-PI), Dee Mullan, Rebecca Ramsay, Vivien Richmond, Mark Sargent, Paul Frattaroli
	Calderdale and Huddersfield Foundation Trust	Matthew Taylor (PI), Marie Home, Sal Uka, Susan Kilroy, Tonicha Nortcliffe, Halima Salim, Kelly Holroyd
	Cardiff & Vale University Health Board	Khalid Hamandi (PI), Alison McQueen, Dympna Mcaleer

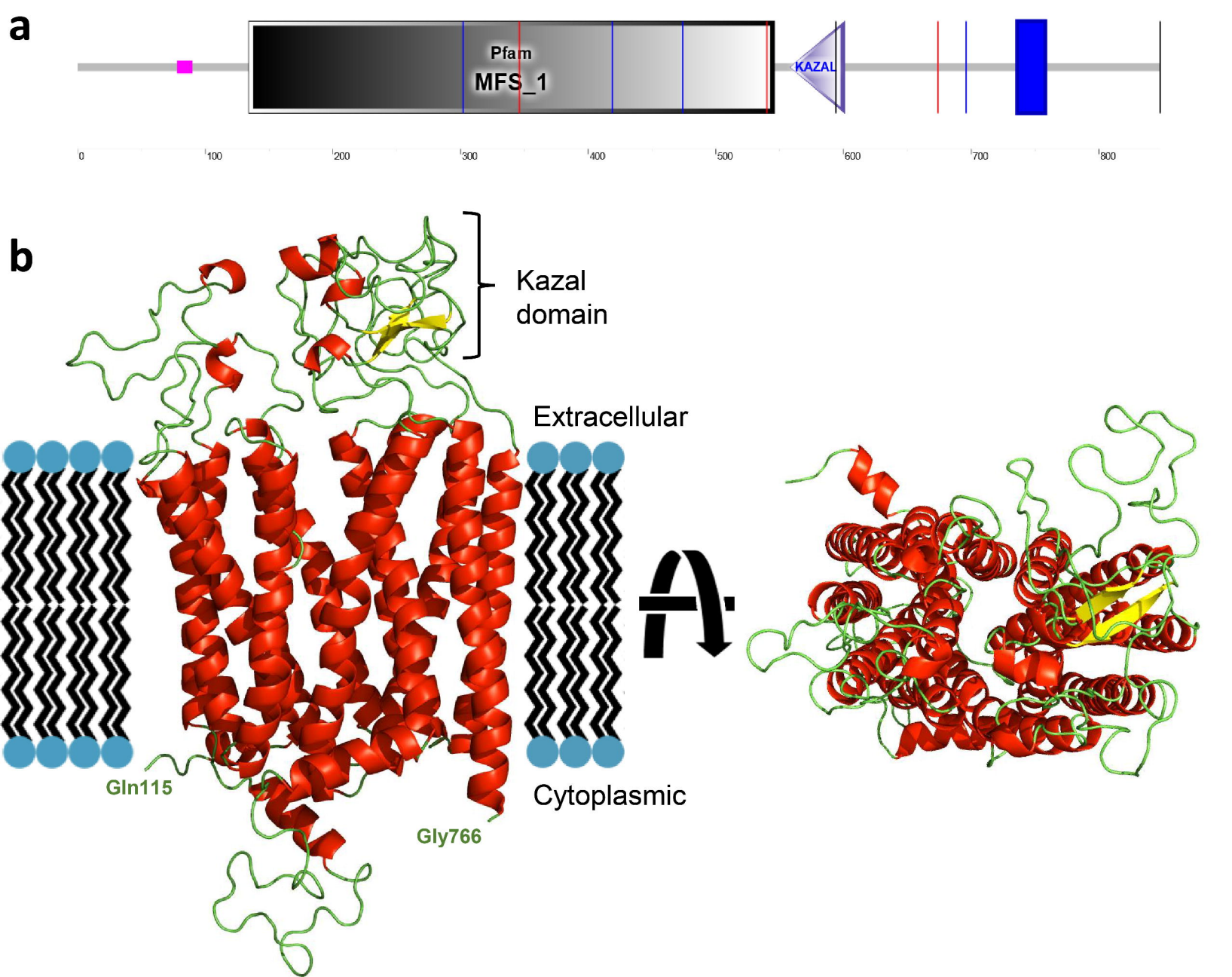
	County Durham and Darlington NHS Foundation Trust	Dina Jayachandran (PI), Dawn Egginton
	Croydon Health Services NHS Trust	Bridget MacDonald (PI), Michael Chang
	Cwm Taf Morgannwg University Health Board	David Deekollu (Co-PI), Alok Gaurav (Co-PI), Caroline Hamilton, Jaya Natarajan
	Dartford and Gravesham NHS Trust	Shane Delamont (PI), Carmel Stuart, Imogen Hayes
	East and North Hertfordshire NHS Trust	Inyan Takon (PI), Janet Cotta
	East Kent Hospitals University NHS Foundation Trust	Nick Moran (PI), Jeremy Bland
	East Lancashire Hospitals NHS Trust	Rosemary Belderbos (PI), Heather Collier, Joanne Henry, Matthew Milner, Sam White
	Guy's and St Thomas' NHS Foundation Trust	Michalis Koutroumanidis (PI), Javier Peña Ceballos, William Stern
	King's College Hospital NHS Foundation Trust	Mark P. Richardson (Co-PI), Jennifer Quirk (Co-PI), Javier Peña Ceballos, Anastasia Papathanasiou
	King's College London	Deb K. Pal (PI), Mark P. Richardson, Holly Crudgington, Anna Hall, Amber Collingwood, Amy Shakeshaft, Ioannis Stavropoulos, Anna Smith, Robert McDowall, Sophie Bayley
	Kingston Hospital NHS Foundation Trust	Dora Lozsadi (PI), Andrew Swain, Charlotte Quamina, Jennifer Crooks
	Lancashire Teaching Hospitals NHS Foundation Trust	Tahir Majeed (PI), Sonia Raj, Shakeelah Patel, Michael Young
	Leeds Teaching Hospitals NHS Trust	Melissa Maguire (Co-PI), Munni Ray (Co-PI), Caroline Peacey, Linetty Makawa, Asyah Chhibda, Eve Sacre, Shanaz Begum
	Manchester University NHS Foundation Trust	Lap Yeung (Co-PI), Claire Holliday, Louise Woodhead, Karen Rhodes

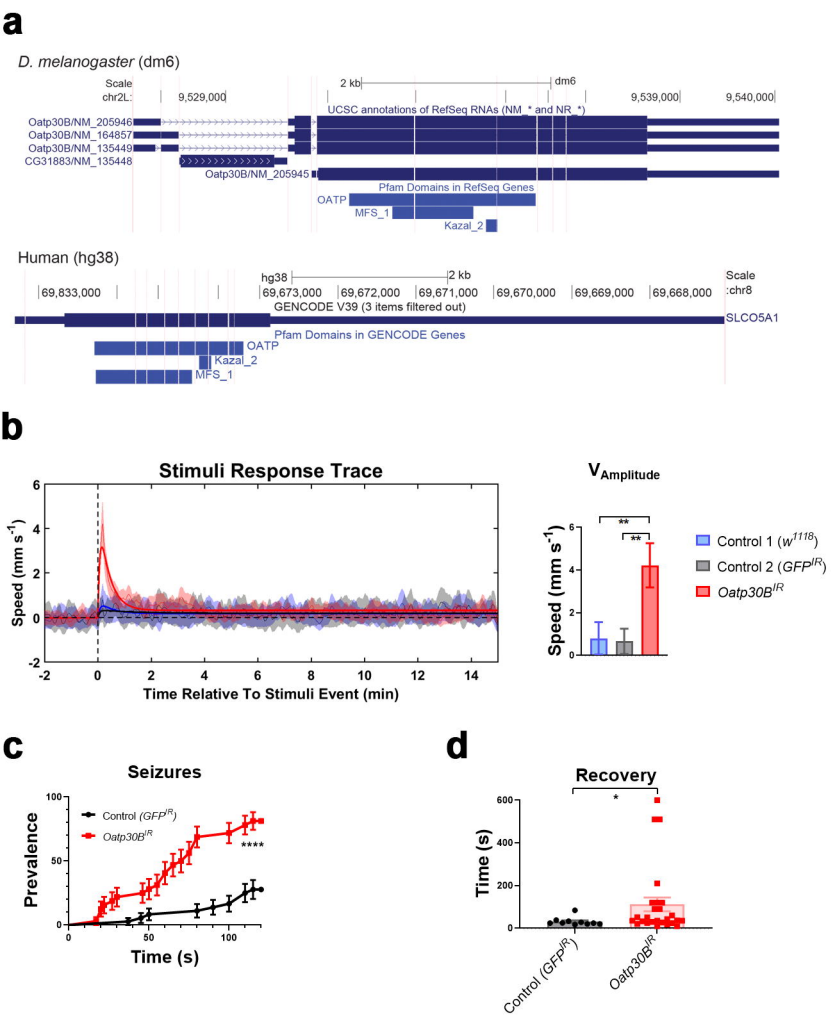
	Newcastle upon Tyne Hospitals NHS Foundation Trust	Rhys Thomas (Co-PI), Shan Ellawela (Co-PI), Joanne Glenton, Verity Calder, John Davis, Paul McAlinden, Sarah Francis
	NHS Grampian	Karen Lanyon (Co-PI), Graham Mackay (CoPI), Elma Stephen (Co-PI), Coleen Thow, Margaret Connon
	NHS Tayside	Martin Kirkpatrick (PI), Susan MacFarlane, Anne Macleod, Debbie Rice
	North Tees and Hartlepool NHS Foundation Trust	Siva Kumar (PI), Carolyn Campbell, Vicky Collins
	Nottingham University Hospitals NHS Trust	William Whitehouse (PI), Christina Giavasi (PI), Boyanka Petrova, Thomas Brown, Catie Picton, Michael O'Donoghue, Charlotte West, Helen Navarra
	Portsmouth Hospitals NHS Trust	Seán J. Slaght (PI), Catherine Edwards, Andrew Gribbin, Liz Nelson, Stephen Warriner
	Royal Free London NHS Foundation Trust	Heather Angus-Leppan (PI), Loveth Ehiorobo, Bintou Camara, Tinashe Samakomva
	Salford Royal NHS Foundation Trust	Rajiv Mohanraj (PI), Vicky Parker
	Sandwell & West Birmingham Hospitals NHS Trust	Rajesh Pandey (PI), Lisa Charles, Catherine Cotter
	Sheffield Children's NHS Foundation Trust	Archana Desurkar (PI), Alison Hyde, Rachel Harrison
	Sheffield Teaching Hospitals NHS Foundation Trust	Markus Reuber (PI), Rosie Clegg, Jo Sidebottom, Mayeth Recto, Patrick Easton, Charlotte Waite, Alice Howell, Jacqueline Smith, Rosie Clegg
	Southport and Ormskirk Hospital NHS Trust	Shyam Mariguddi (PI), Zena Haslam
	St George's University Hospitals NHS Foundation Trust	Elizabeth Galizia (PI), Hannah Cock, Mark Mencias, Samantha Truscott, Deirdre Daly, Hilda Mhandu, Nooria Said

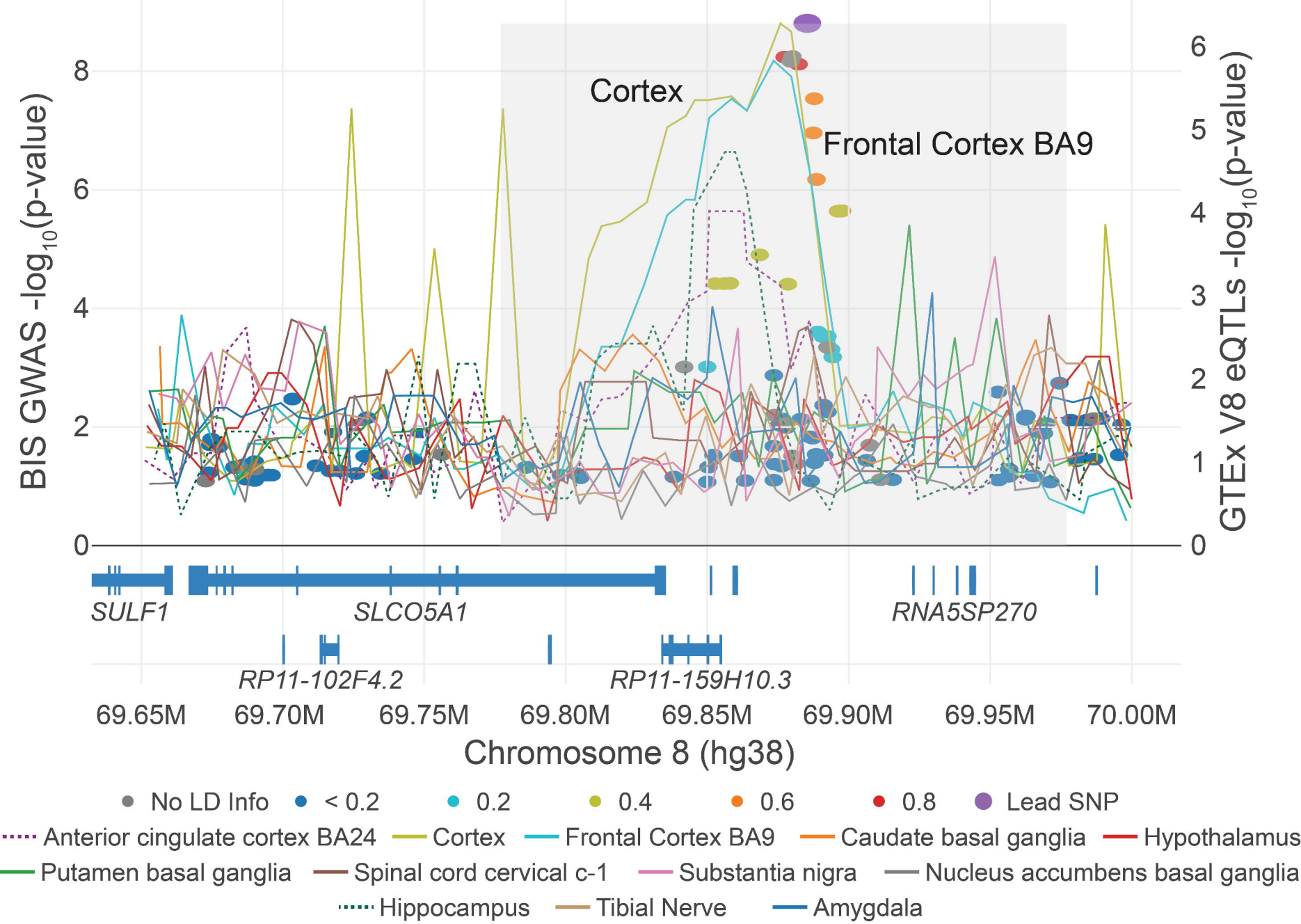
	Swansea University Medical School and Swansea Bay University Healthboard	Mark Rees (PI), Seo-Kyung Chung, Owen Pickrell, Beata Fonferko-Shadrach, Mark Baker
	Taunton & Somerset NHS Foundation Trust	Amy Whiting (PI), Kirsty O'Brien
	The Mid Yorkshire Hospitals NHS Trust	Fraser Scott (Co-PI), Naveed Ghaus (Co-PI), Gail Castle, Jacqui Bartholomew, Ann Needle, Julie Ball, Andrea Clough
	The Royal Wolverhampton NHS Trust	Shashikiran Sastry (PI), Charlotte Busby
	The Walton Centre NHS Foundation Trust	Amit Agrawal (PI), Debbie Dickerson, Almu Duran
	University Hospitals Birmingham NHS Foundation Trust	Muhammad Khan (PI), Laura Thrasyvoulou, Eve Irvine, Sarah Tittensor, Jacqueline Daglish
	University Hospitals of Derby and Burton NHS Foundation Trust	Sumant Kumar (PI), Claire Backhouse, Claire Mewies
	University Hospitals Plymouth NHS Trust	Rahul Bharat (PI), Sarah-Jane Sharman
	Walsall Healthcare NHS Trust	Darwin Pauldhas (PI), Sharon Kempson, Lisa Richardson, Lynn Hawkins
	West Suffolk NHS Foundation Trust	Arun Saraswatula (PI), Helen Cockerill
USA	Nationwide Children's Hospital, Ohio	David A. Greenberg (PI)

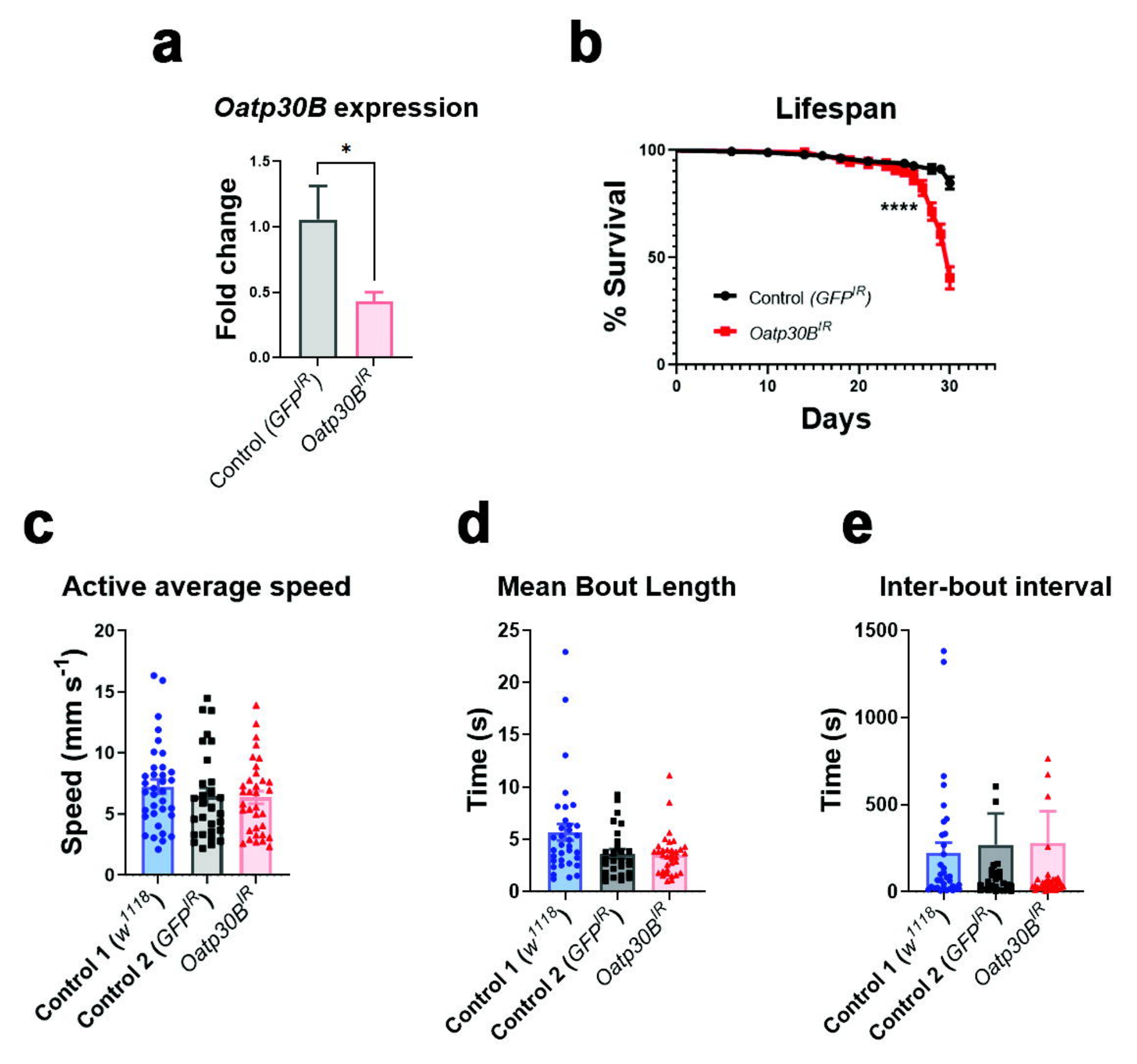

rs73293634
rs75042057
-log10(p-value)
10
9
8
7
6
5
4
3
2
1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
17
19
21
X
Chromosome (hg38)


a
BIS GWAS -log10(p-value)
GTEx V8 eQTLs -log10(p-value)
Cortex
Frontal Cortex BA9
SULF1
SLCO5A1
RNA5SP270
RP11-102F4.2
RP11-159H10.3
69.65M
69.70M
69.75M
69.80M
69.85M
69.90M
69.95M
70.00M
Chromosome 8 (hg38)
No LD Info
< 0.2
0.2
0.4
0.6
0.8
Lead SNP
Anterior cingulate cortex BA24
Cortex
Frontal Cortex BA9
Caudate basal ganglia
Hypothalamus
Putamen basal ganglia
Spinal cord cervical c-1
Substantia nigra
Nucleus accumbens basal ganglia
Hippocampus
Tibial Nerve
Amygdala
b
GTEx V8 Brain Cortex
PsychENCODE DLPFC
Fetal brain
GWAS -log10(p-value)
GTEx V8 eQTL -log10(p-value)
SLCO5A1
RNA5SP270
RP11-102F4.2
RP11-159H10.3
69.65M
69.7M
69.75M
69.8M
69.85M
69.9M
69.95M
70M
Chromosome 8 (hg38)
c
rs73293634
Norm. Expression
GG
(183)
GT
(18)
TT
(4)
d
SLCO5A1 Expression
(log2RPKM)
1–18pcw.
18–37pcw.
First year
Preadolescence
Adolescence
Adult


a
Pfam
MFS_1
KAZAL
0
100
200
300
400
500
600
700
800
b
Kazal
domain
Extracellular
Cytoplasmic
Gln115
Gly766


a
D. melanogaster (dm6)
Scale
chr2L:
9,529,000
2 kb
dm6
9,539,000
9,540,000
UCSC annotations of RefSeq RNAs (NM_* and NR_*)
Oatp30B/NM_205946
Oatp30B/NM_164857
Oatp30B/NM_135449
CG31883/NM_135448
Oatp30B/NM_205945
Pfam Domains in RefSeq Genes
OATP
MFS_1
Kazal_2
Human (hg38)
hg38
2 kb
Scale
:chr8
69,833,000
69,673,000
69,672,000
69,671,000
69,670,000
69,669,000
69,668,000
GENCODE V39 (3 items filtered out)
SLCO5A1
Pfam Domains in GENCODE Genes
OATP
Kazal_2
MFS_1
b
Stimuli Response Trace
Speed (mm s-1)
Time Relative To Stimuli Event (min)
VAmplitude
Speed (mm s-1)
**
**
Control 1 (w1118)
Control 2 (GFPIR)
Oatp30BIR
c
Seizures
Prevalence
Control (GFPIR)
Oatp30BIR
****
Time (s)
d
Recovery
Time (s)
*
Control (GFPIR)
Oatp30BIR


BIS GWAS -log10(p-value)
GTEx V8 eQTLs -log10(p-value)
Cortex
Frontal Cortex BA9
SULF1
SLCO5A1
RNA5SP270
RP11-102F4.2
RP11-159H10.3
69.65M
69.70M
69.75M
69.80M
69.85M
69.90M
69.95M
70.00M
Chromosome 8 (hg38)
No LD Info
< 0.2
0.2
0.4
0.6
0.8
Lead SNP
Anterior cingulate cortex BA24
Cortex
Frontal Cortex BA9
Caudate basal ganglia
Hypothalamus
Putamen basal ganglia
Spinal cord cervical c-1
Substantia nigra
Nucleus accumbens basal ganglia
Hippocampus
Tibial Nerve
Amygdala


a
Oatp30B expression
Fold change
1.5
1.0
0.5
0.0
*
Control (GFP^IR)
Oatp30B^IR
b
Lifespan
% Survival
100
50
0
****
Control (GFP^IR)
Oatp30B^IR
0
10
20
30
Days
c
Active average speed
Speed (mm s^-1)
20
15
10
5
0
Control 1 (w^1118)
Control 2 (GFP^IR)
Oatp30B^IR
d
Mean Bout Length
Time (s)
25
20
15
10
5
0
Control 1 (w^1118)
Control 2 (GFP^IR)
Oatp30B^IR
e
Inter-bout interval
Time (s)
1500
1000
500
0
Control 1 (w^1118)
Control 2 (GFP^IR)
Oatp30B^IR