Integrative genomics reveals pathogenic mediator of valproate-induced neurodevelopmental disability

Rahel Feleke,[1,†] Dana Jazayeri,[2,3,†] Maya Abouzeid,[1] Kim L. Powell,[2,4] Prashant K. Srivastava,[5,‡] Terence J. O'Brien,[2,4,‡] Nigel C. Jones[2,4,‡] and Michael R. Johnson[1,‡]

[†,‡]**These authors contributed equally to this work.**

1 Department of Brain Sciences, Imperial College London, London, UK

2 The Departments of Medicine and Neurology, The Royal Melbourne Hospital, The University of Melbourne, Parkville, Victoria, Australia

3 The ALIVE National Centre for Mental Health Research Translation, The Department of General Practice, Melbourne Medical School, The University of Melbourne, Parkville, Victoria, Australia

4 Department of Neuroscience, The Central Clinical School, Alfred Health, Monash University, Melbourne, Victoria, Australia

5 National Heart and Lung Institute, Imperial College London, London, UK

Correspondence to: Professor Michael R. Johnson

Department of Brain Sciences, Imperial College London, Room E419, Burlington Danes Building, 160 Du Cane Road, London W12 0NN, UK

E-mail: m.johnson@imperial.ac.uk

Running title: Valproate-induced neurodevelopment

Keywords: valproate; gene expression; neurodevelopment; disability

Abbreviations: ASM = Anti-seizure medication; ADHD = Attention Deficit Hyperactivity Disorder; All-E = All valproate exposed pups; All-N= All non-exposed pups; ASD = Autism Spectrum Disorder; BD = Bipolar Disorder; BP = Biological Process; CDG = Cross Disorders Group; DE = Differentially Expressed; DEG = Differentially Expressed Genes; DWLS = Dampened Weighted Least Squares; EPI= Epilepsy; GAERS = Genetic Absence Epilepsy Rats from Strasbourg; GGE = Genetic

Generalized Epilepsy; GO = Gene Ontology; GWAS = Genome-Wide Association Studies; HDAC = Histone DeACetylase; IQ = Intelligence Quotient; LDSC = Linkage Disequilibrium Score Regression; MCMs = Major Congenital Malformations; NEC = Non-Epileptic Control; PC = Principal Component; QC = Quality control; rno6 = Rattus norvegicus reference genome; scRNA-seq = Single-cell RNA-sequencing; SCZ = Schizophrenia; VPA = Sodium valproate; WebGestaltR =WEB-based Gene SeT AnaLysis Toolkit; WHR = Waist/Hip Ratio

ABSTRACT

Prenatal exposure to the anti-seizure medication sodium valproate (VPA) is associated with an increased risk of adverse postnatal neurodevelopmental outcomes, including lowered intellectual ability, autism spectrum disorder and attention-deficit hyperactivity disorder. In this study, we aimed to clarify the molecular mechanisms underpinning the neurodevelopmental consequences of gestational VPA exposure using integrative genomics. First, we assessed the effect of gestational VPA on fetal brain gene expression using a validated rat model of valproate teratogenicity that mimics the human scenario of chronic oral valproate treatment during pregnancy at doses which are therapeutically relevant to the treatment of epilepsy. Two different rat strains were studied - inbred Genetic Absence Epilepsy Rats from Strasbourg (GAERS), a model of genetic generalized epilepsy, and inbred Non-Epileptic Control rats. Female rats were fed standard chow or VPA mixed in standard chow for 2 weeks prior to conception and then mated with same-strain males. In the VPA-exposed rats maternal oral treatment was continued throughout pregnancy. Fetuses were extracted via C-section on gestational day 21 (one day prior to birth) and fetal brains were snap frozen and genome-wide gene expression data generated. We found that gestational VPA exposure via chronic maternal oral dosing was associated with substantial drug-induced differential gene expression in the pup brains, including dysregulated splicing, and observed that this occurred in the absence of evidence for significant neuronal gain or loss. The functional consequences of VPA-induced gene expression were explored using pathway analysis and integration with genetic risk data for psychiatric disease and behavioural traits. The set of genes down-regulated by VPA in the pup brains were significantly enriched for pathways related to neurodevelopment and synaptic function, and significantly enriched for heritability to human intelligence, schizophrenia and bipolar disorder. Our results provide a mechanistic link between chronic fetal VPA exposure and neurodevelopmental disability mediated by VPA-induced transcriptional dysregulation.

INTRODUCTION

Sodium valproate (VPA) is an anti-seizure medication (ASM) widely used in the treatment of epilepsy, in particular genetic generalized epilepsy (GGE) for which it is the most effective ASM,[1,2] as well as being used to treat non-epileptic conditions such as migraine and bipolar disorder.[3] In addition to the well-established approximate 10% risk of major congenital malformations (MCMs) associated with prenatal VPA exposure,[4,5] children born to mothers taking VPA are at an increased risk of adverse neurobehavioral outcomes including lowered intellectual ability, neurodevelopmental delay, autism spectrum disorder and attention-deficit hyperactivity disorder. [6–8]

We have previously reported the development and validation of a translational animal model of VPA-induced teratogenicity that mimics a number of the clinical features of human gestational VPA exposure, including chronic oral treatment of maternal rats before and after conception at therapeutically relevant blood levels (i.e., dosed orally to match VPA blood levels observed in humans).[9] This model was shown to recapitulate many of the developmental consequences of gestational VPA exposure, including brain maldevelopment, altered intravertebral distances and a significant developmental delay of vertebral arches. Here, we consider that the model also offers an opportunity to evaluate the molecular mechanisms underpinning VPA-associated neurodevelopmental disability.

In addition to its known effects on neuronal membrane excitability, VPA is an established histone deacetylase inhibitor and can induce gene expression changes both *in vitro* and *in vivo.*[10,11] Although these studies have established that VPA can induce gene expression changes in the central nervous system and neurons, their relevance to human neurodevelopmental outcomes is unclear due either to the use of non-therapeutically-relevant VPA concentrations or non-clinically utilized drug delivery systems (e.g., the use of sub-cutaneous or intraperitoneal delivery), as well as the use of gestational exposures that are not representative of the human condition such as acute exposures restricted to specific gestational time points.

We therefore aimed to advance insights into the mechanistic underpinnings of valproate-associated neurodevelopmental disability by evaluating brain gene expression changes of rat pups chronically exposed to VPA *in utero,* and by considering the consequences of these VPA-induced brain gene expression changes using integrative genomics. The influence of epilepsy and genetic background on VPA-induced changes in brain gene expression was examined by investigating both epileptic (Genetic Absence Epilepsy Rats from Strasbourg (GAERS), a model of GGE) and Non-Epileptic Control (NEC) dams.

MATERIALS AND METHODS

Experimental Design

Female rats from two different strains, inbred Non-Epileptic Controls (NEC) and Genetic Absence Epilepsy Rats from Strasbourg (GAERS), were obtained from the Department of Medicine, Royal Melbourne Hospital, University of Melbourne Biological Research Facility.[12] VPA was administered via food chronically in a clinically relevant manner. Rats were fed standard rodent diet pre-mixed with VPA 20g/kg or standard chow, as previously described.[9] Day 0 of pregnancy was marked by the presence of a plug. C-section was used to extract foetuses from the uterus on the 21st day of gestation. Pup brain tissues were then extracted using the QIAGEN Simultaneous Purification of Genomic DNA and total RNA from Animal Tissues protocol. External assessment including spinal measurements were used to examine the foetuses for any form of birth defects. The brain samples were extracted and immediately frozen and stored a −80 °C without reference to the malformation score.

Transcriptome sequencing and gene expression data processing

The mRNA Isolation Kit was used to extract total RNA from brain tissue samples following the manufacturer's instructions. Total RNA was examined for quantity and quality using the TruSeq Stranded Total RNA Library Prep Gold. Sequencing libraries were quantified and sequence reads per sample generated. Quality control (QC) of RNA-seq reads was conducted using FastQC[46]. Low quality reads and ribosome RNA reads were removed. Trimmed reads were then mapped to the Rattus norvegicus reference genome (rno6) using STAR (version 2.5). [47]

Fetal sex determination

Using a list of prenatally expressed Y-chromosome-encoded genes from a recently published study,[48] we were able to categorise the pup brains into male or female. The study identified these genes by comparing mRNA levels in different brain regions from human prenatal brain samples. In the current study, we assessed the expression of those genes using the rodent orthologs: USP9Y (Ubiquitin Specific Peptidase 9 Y-Linked), DDX3Y (DEAD-Box Helicase 3 Y-Linked), and UTY (Ubiquitously Transcribed Tetratricopeptide Repeat Containing, Y-Linked) and classified 14 samples as female and 16 samples as male (refer to **Table 1**).

Differential gene expression analysis

R-package edgeR[49]was used to perform differential expression analysis and a threshold of FDR < 0.05 was applied to identify significantly differentially expressed genes (DEGs). To explore mRNA expression differences between VPA-exposed and non-exposed pup brains (after regressing out any technical differences due to sex), the following comparisons were undertaken: i) E-NEC vs N-NEC; ii) E-GAERS vs N-GAERS; iii) All exposed pups (i.e., E-GAERS + E-NEC) vs all non-exposed pups (N-NEC + N-NEC). To inform gene expression changes in the rat pup brain relating to epilepsy we also considered the following two comparisons: iv) E-GAERS vs E-NEC; v) N-GAERS vs N-NEC. We report all changes but the primary aim of the work was to interrogate transcriptional changes following gestational VPA exposure.

Functional enrichment and pathway analysis

The R package WEB-based Gene SeT AnaLysis Toolkit (WebGestaltR, v 0.4.4),[50] which uses databases from the Gene Ontology Consortium, was used to perform functional term enrichment analysis of DEGs from each pairwise comparison. Default values for WebGestalt parameters were used, including overlap of 10 for minimum and maximum 500 genes and FDR-correction for multiple testing was performed using Benjamini and Hochberg (BH) method and significant pathways identified using a threshold of FDR < 0.05. In order to reduce redundancy across significant GO-terms the following steps were taken. First, for each of the analyses GO terms were filtered to exclude terms with ≥ 20 genes or ≤ 1000 genes. Second, using go_reduce() function from the r-utils R package (https://rhreynolds.github.io/rutils/articles/rutils.html) semantic similarity of GO terms were calculated. Third, a threshold of 0.9 was applied to get fewer GO-terms. This threshold was applied to the hierarchical tree generated by reduceSimMatric () function from the rrvgo R package (v 1.1.4), which uses a bottom-up clustering method.

LD score regression

LDSC was used to test for enrichment of heritability using published GWAS studies. First, annotation files were generated for each set of DEGs with rows corresponding to a SNP and a column for each sub-annotation. SNPs were then mapped to genes using dsSNP file NCBI Build 37 co-ordinates (build 147 and hg19), and values of 0 were given to SNPs not present in the file. Overall, fifteen annotation files were generated for all the comparisons (separate files for all-dysregulated, up-regulated and down-regulated DEGs from each comparison). Second, LDSC was run using data files from phase 3 of the 1000 Genome Project Phase 3 European population. As per LDSC Github Wiki

recommendation, LD scores were calculated for the annotations using 1cM window and the analysis was restricted to Hapmap3 SNPs. Third, LDSC python scripts ("munge_sumstats.py") were used to format the summary statistic files. Fourth, for the regression weights, LD calculated for HapMap3 SNPs were downloaded from the LDSC Github page (https://github.com/bulik/ldsc). Additionally, for the LD score files used for the LD score regression, the full baseline model was used. For all the LDSC analyses, an enrichment score and its associated P-value was calculated based on the proportion of total SNPs per annotations (column), after taking into account all other annotation. Annotation categories with a significant positive enrichment of SNP-heritability are then reported as a final result (a one tailed test). GWAS summary statistics were downloaded from The European Bioinformatics Institute (EMBL-EBI) and Psychiatric Genomic Consortium (PGC) Cross-Disorder Group. All subjects were of European ancestry. A full detailed list of all the summary statistic used in these analyses can be found in **Supplementary Table 4.**

Deconvolution

To deconvolute the bulk RNA-seq signature into its component single cell-type expression profiles, we anchored the analysis in a single-cell RNA-seq dataset from a recently published study that characterised the molecular architecture of the developing mouse brain.[51] Using this dataset, we were able to characterise 93 mouse brain tissues at single-cell resolution and generate transcriptomic atlas of the embryonic mouse brain from gastrulation up until birth. Here, using cells extracted on E(18), gene by cell matrix was constructed and cell types were identified, followed by further analysis using Seurat functions.[52] First, the dataset was normalised using the NormalizeData () function. This global scaling normalisation method normalises the gene expression values by multiplying the total expression of the cell by the number (n) of cells and then log transforms the result. A size factor of 10,000 was used for each cell. Second, cells were clustered by: a) Calculating the distance between two cells with similar expression using Euclidean algorithm and drawing edges. b) Then using the FindClusters() function the cells were clustered (parameters used: 30 principal components (PCs) and a resolution parameter of 2). Third, the clusters were then visualised using non-linear dimensionality reduction algorithm known as Uniform Manifold Approximation and Projection (UMAP, v 0.1.10). Finally, Wilcoxon rank sum test (FDR < 0.05) (implemented in the FindAllMarkers() function) was the used to identify genes differentially expressed (DE) in one cluster compared with all other clusters. Cell types were allocated by testing DEGs in a particular cell type for enrichment (Fisher's exact test) for cell-type markers from mouse datasets.[53] Next, the relative proportions of cell types in rat pups brain tissue samples were determined using weighted least squares-based deconvolution algorithms, DWLS. First, expression

matrix from all rat pup samples was normalised using count per million based normalisation methods. Second, using biomaRt R-package (v4.1), the orthologous mouse gene was inferred (Ensembl Genes 104). Finally, cells were deconvolved using weighted least squares approach.

Data and code availability

The raw and processed rat sequencing data generated in this study have been deposited in NCBI Gene Expression Omnibus database under the accession number GSE198756. All scripts used in this study can be found in GitHub (https://github.com/rahfel/VPA). Supplementary materials including full tables can also be found in GitHub ((https://github.com/rahfel/VPA).

RESULTS

Study workflow and data collection

A summary of the study workflow is shown in **Figure 1**. Female GAERS and NEC rats were both derived from the same original Wistar rat colony but selectively inbred to express, or not to express, absence seizures. [12,13] The animal facility was maintained on a light dark cycle of 12 hours light, 12 hours dark at a temperature range of 19 to 22 degrees. All females were fed a standard rodent diet as reported by Senn and colleagues [14] either pre-mixed with VPA 20g/kg of food or not (controls) for 2 weeks before mating with males of the same strain and continued throughout pregnancy. We have previously demonstrated that this dose of VPA administered orally results in significant seizure suppression in adult GAERS and achieves blood serum levels broadly equivalent to human therapeutic levels of VPA; across VPA-exposed GAERs and NEC dams, VPA levels ranged 120-380 μmol/L and 250-460 μmol/L respectively (t-test P=0.197).[9] The presence of a copulatory plug indicated day 0 of pregnancy. Once the plug was present, females were separated from the males for the course of their pregnancy. The following day was classified as day 1. The fetuses were extracted from the uterus via C-section on day 21, one day before expected birth. The dams were sacrificed and the fetus brains removed, snap frozen and stored at -80C. We have previously reported the results of the teratology studies on rat pups exposed to valproate *in utero*.[9] Pup brain samples were then chosen without reference to morphological score and were not significantly different in weight between the VPA-exposed and unexposed groups (Mean weights +/- SEM: NEC control 0.151g +/- 0.017; NEC VPA 0.155g +/- 0.014; GAERS Control 0.180g +/- 0.012; GAERS VPA 0.137g +/- 0.011, 2 WAY ANOVA (drug) p=0.15; (strain) p=0.71). Six litters were generated for each group and one or two pup brains

from each litter were randomly processed. Total RNA was extracted from 30 pup brains (**Table 1**) and genome-wide gene expression data was generated using RNA-sequencing (RNA-seq) (**Methods**).

Deconvolution analysis reveals no difference in cell-type composition between VPA-exposed and non-exposed pup brains

Previous pre-clinical studies on the effects of intraperitoneal administration of VPA have demonstrated VPA-induced apoptotic neurodegeneration in the developing rat brain and this has been suggested as a mechanism to explain VPA-associated cognitive impairment.[15,16] We therefore first sought to identify whether chronic oral VPA exposure used in our model was associated with significant changes in the proportion of cell-types in the rat pup brain. To this end, we applied a weighted least squares-based deconvolution algorithm *Dampened Weighted Least Squares* (DWLS)[17] which estimates cell-type composition in a bulk tissue RNA-seq dataset using prior information from an unrelated single-cell RNA-sequencing (scRNA-seq) signature from an analogous tissue. Taking this approach, we observed no significant difference in cell-type composition between valproate-exposed and unexposed pup brains, for either GAERS or NEC rats, suggesting chronic orally delivered gestational VPA exposure does not have a major effect on cell loss or gain (FDR-corrected Wilcoxon rank sum test >0.05; **Supplementary Figure 1; Supplementary Table 1**).

Differential gene expression analysis highlights transcriptional changes in VPA-exposed pup brains

Prior to undertaking differential gene expression analysis, we explored the principal components of variation between epileptic and non-epileptic pups and between VPA-exposed and non-exposed pups (**Supplementary Figure 2**). Using genome-wide gene expression profiles for each brain, we observed that epileptic and non-epileptic pups were broadly separated by the first principal component (PC1) and VPA-exposed and non-exposed pups by PC2. We therefore undertook differential gene expression analyses for the "case control" comparisons listed in **Table 2**.

We observed significantly (FDR < 0.05) differentially expressed genes in all five pairwise comparisons (summarised in **Table 2**; see **Supplementary Table 2** for a full list of differentially expressed genes for each comparison). Considering VPA-exposed vs non-exposed pup brains, we observed the largest number of differentially expressed genes (DEG) when comparing all VPA-exposed pups (All-E) consisting of VPA-exposed GAERS pups (E-GAERS) and VPA-exposed NEC (E-NEC) pups combined vs all non-exposed pups (All-N) consisting of non-exposed GAERS (N-GAERS) and non-exposed NEC (N-NEC) pups combined, where 3,470 genes were significantly (FDR

< 0.05) differentially expressed. Of these, 1,632 genes were significantly (FDR < 0.05) up-regulated and 1,838 down-regulated.

Figure 2 shows the overlaps in genes differentially expressed for each of the 3 pairwise VPA-exposed vs non-exposed comparisons (i.e., ALL-E vs ALL-N, E-NEC vs N-NEC and E-GAERS vs N-GAERS). We observed that the majority of genes differentially expressed by VPA in GAERS pups were also differentially expressed in NEC pups, suggesting that genetic epilepsy status is not a major determinant of the pattern of differential expression induced by VPA, as was previously the case for VPA-induced birth defects in this model.[9]

To investigate the functional consequences of VPA-induced differential gene expression in the pup brain, we first undertook pathway enrichment analysis. We found that genes down-regulated following VPA exposure were highly significantly enriched for functional processes related to modulation of synaptic function and neuronal processes (see **Figure 3** for summary of Biological Processes dysregulated by VPA and **Supplementary Tables 3a-3c** for full details of pathway enrichment analyses). For example, when considering all VPA-exposed pups vs unexposed pups (All-E vs All-N), among the set of genes down-regulated by VPA we observed highly significant enrichment for terms relating to glutamate receptor complex (FDR= $9.74x10^{-12}$), regulation of membrane potential (FDR= 0), synapse assembly (FDR= $2.17x10^{-11}$), regulation of actin cytoskeleton organisation (FDR= $1.48x10^{-02}$), post-synaptic membrane (FDR= $2.11x10^{-04}$), neurotransmitter receptor activity (FDR= $2.33x10^{-04}$) and axon guidance (FDR= 0). Interestingly, also among the pathways enriched in the down-regulated set of genes was regulation of insulin secretion (FDR= $9.5x10^{-03}$), which if replicated in the pancreas (not examined in this study) might suggest a possible drug-induced transcriptional mechanism for the increased incidence of impaired glucose control in patients treated with VPA.[18,19]

In contrast to the substantial enrichment of neuronal functions in genes down-regulated by VPA, the genes up-regulated by VPA were generally enriched for functional terms not directly related to neural processes namely mRNA splicing (FDR= $7.29x10^{-14}$), translation (FDR= $1.40x10^{-13}$), extra-cellular matrix organisation (FDR= $2.41x10^{-12}$) and cell cycle (FDR= $2.40x10^{-09}$). Consistent with VPA's known activity as a histone deacetylase inhibitor, we also saw enrichment of biological processes related to chromatin organization (FDR= $7.37x10^{-07}$) and chromatin assembly/disassembly (FDR = $1.45x10^{-6}$) among the genes upregulated by VPA.

In summary, genes up- and down-regulated by chronic *in utero* VPA exposure in the developing rat brain were represented by divergent functional categories, with genes down-regulated by VPA predominantly characterized by pathways relating to nervous system function.

Genes down-regulated by gestational VPA exposure are highly enriched for heritability for neurodevelopmental traits and disease

We next sought to investigate the relationship between genes differentially expressed in pup brains following gestational VPA and genetic risk for neurodevelopmental disease and behavioural traits. Since, in the context of rare variant genetic analyses, neurodevelopmental disability primarily arises from loss-of-function and deleterious mutations,[20] we hypothesised that genetic risk for neurodevelopmental disease may be enriched in the set of genes significantly down-regulated by VPA in the developing brain.

To test for enrichment of genetic association in sets of genes, we used stratified LD score regression (LDSC).[21] Given the broad range of adverse neurodevelopmental outcomes in children gestationally exposed to VPA, we considered genome-wide association studies (GWAS) for a broad range of psychiatric diseases and behavioural traits: attention deficit hyperactivity disorder (ADHD),[22] bipolar disorder (BD),[23] autism spectrum disorder (ASD),[24] schizophrenia (SCZ),[25] full-scale intelligence quotient (IQ),[26] epilepsy (EPI)[27] and "cross disorders group" (CDG), the latter representing a meta-analysis of eight individual psychiatric disorder GWAS namely anorexia nervosa, attention-deficit/hyperactivity disorder, autism spectrum disorder, bipolar disorder, major depression, obsessive-compulsive disorder, schizophrenia and Tourette syndrome [28]. As a negative control GWAS dataset not expected to be enriched in the set of genes differentially expressed in pup brains we used a GWAS for waist/hip ratio (WHR).[29]

Enrichment of genetic association analyses were run for all sets of genes significantly (FDR < 0.05) differentially expressed in the VPA-exposed pup brain (results summarised in **Figure 4**; full details in **Supplementary Table 4**). For genes down-regulated by VPA, we observed highly significant enrichment of heritability with bipolar disorder (All-E vs All-N, $FDR_{LDSC} = 1.16x10^{-08}$; E-NEC vs N-NEC, $FDR_{LDSC} = 2.93x10^{-09}$; E-GAERS vs N-GAERS, $FDR_{LDSC} = 1.55x10^{-03}$), schizophrenia (All-E vs All-N, $FDR_{LDSC} = 6.04x10^{-08}$; E-NEC vs N-NEC, $FDR_{LDSC} = 6.40x10^{-08}$; E-GAERS vs N-GAERS, $FDR_{LDSC} = 2.91x10^{-03}$), IQ (All-E vs All-N, $FDR_{LDSC} = 1.17x10^{-03}$; E-NEC vs N-NEC, $FDR_{LDSC} = 7.03x10^{-04}$; E-GAERS vs N-GAERS, $FDR_{LDSC} = 1.62x10^{-02}$) and cross disorders group (All-E vs All-N, $FDR_{LDSC} = 3.89x10^{-04}$; E-NEC vs N-NEC, $FDR_{LDSC} = 1.82x10^{-04}$; E-GAERS vs N-GAERS, $FDR_{LDSC} = 5.44x10^{-02}$). No significant enrichments of heritability were observed for these traits in unexposed GAERS pups compared to unexposed non-epileptic controls, and enrichments of association for WHR were non-significant across all comparisons (**Supplementary Table 4**).

There was no significant enrichment of heritability in the set of genes down-regulated in VPA-exposed pup brains to ADHD (All-E vs All-N, $FDR_{LDSC} = 0.69$; E-NEC vs N-NEC, $FDR_{LDSC} = 0.24$;

E-GAERS vs N-GAERS, FDR_{LDSC} = 0.88) autism spectrum disorder (All-E vs All-N, FDR_{LDSC} = 0.66; E-NEC vs N-NEC, FDR_{LDSC} = 0.68; E-GAERS vs N-GAERS, FDR_{LDSC} = 0.68) or epilepsy (All-E vs All-N, FDR_{LDSC} = 0.77; E-NEC vs N-NEC, FDR_{LDSC} = 0.57; E-GAERS vs N-GAERS, FDR_{LDSC} = 0.77).

For genes up-regulated by gestational VPA exposure, we did not observe a significant enrichment of genetic association to any neurodevelopmental disease or trait, other than a marginal enrichment to bipolar disorder (All-E vs All-N, $FDR_{LDSC} = 2.7x10^{-02}$; E-NEC vs N-NEC, $FDR_{LDSC} = 3.6x10^{-03}$; E-GAERS vs N-GAERS, FDR_{LDSC} = 0.84) and schizophrenia (All-E vs All-N, $FDR_{LDSC} = 4.3x10^{-02}$; E-NEC vs N-NEC, FDR_{LDSC} = 0.32; E-GAERS vs N-GAERS, FDR_{LDSC} = 0.56).

Taken together, these results suggest VPA exerts its adverse effects on fetal neurodevelopment predominantly via drug-induced down-regulation of genes highly relevant to neurodevelopment and nervous system function. The directionality of the effect is consistent with that observed from rare-variant analyses of genetic risk to neurodevelopmental disability where the predominant mechanism is a dominant negative effect from deleterious gene mutations.[30] Given the functional enrichment of chromatin assembly/disassembly terms among the genes upregulated by VPA, along with VPA's known activity as a histone deacetylase inhibitor, it seems likely that this transcriptional dysregulation is epigenetically encoded with potentially long-lasting consequences for human brain function and health even in the absence of persisting valproate exposure postnatally.

Alternatively spliced genes in VPA exposed pups compared to non-exposed pups

Among the pathways enriched in genes up-regulated by VPA we observed a highly significant (FDR = 0) enrichment for genes involved in splicing function. We therefore evaluated the role of VPA on differential splicing by first quantifying the read counts to individual gene exons and then comparing differential exon usage using EdgeR (**Methods**). Overall, there were 57 significantly (FDR < 0.05) alternatively spliced genes, among which 21 were also differentially expressed following VPA exposure (14 were down-regulated, 7 up-regulated) (**Supplementary Table 5**). Whilst there were no significant enrichments for known functional pathways or enrichment of genetic association to neurodevelopmental disease among the set of differentially spliced genes (data not shown), individual differentially spliced genes included genes for neuronal proteins such as Calmodulin Binding Transcription Activator 1 (CAMATA1; $FDR = 3.69x10^{-26}$) which is known to play a role in the regulation of glutamate levels and neuronal excitability, glutamate decarboxylase 2 (GAD2; $FDR = 9.03x10^{-25}$) which plays a role GABA-synthesis in neurons[31] and Forkhead box P4 (FOXP4; $FDR= 4.62x10^{-3}$) which is known to regulate neurogenesis and in which mutations are associated with speech

delay and congenital abnormalities.[32] These results suggest differential splicing as a potential further mechanism influencing behavioural outcomes following fetal VPA exposure.

DISCUSSION

The adverse neurodevelopmental consequences of chronic fetal valproate exposure have been well-documented and remain a significant limitation on the use of VPA in women who could become pregnant, with particular implications for the treatment of generalized epilepsy, where it remains the most effective treatment,[1,2] as well as other diseases in women of child-bearing potential such as bipolar disorder where it is an important therapy. The mechanisms by which VPA contributes to behavioural and cognitive disability in children following gestational exposure remain poorly defined, with existing studies limited by non-physiological drug administration and dosing schedules and relatively few studies having examined transcriptome-wide alterations in brain gene expression.[11] In this study, we utilized an established rat model of VPA-induced teratogenicity[9] that recapitulates human pre-natal valproate exposure and chronicity of oral dosing during pregnancy.

Among the proposed mechanisms for VPA-associated neurodevelopmental disability is VPA-induced apoptosis.[33] In the present study, using deconvolution analysis anchored in single-cell RNA-seq, we found no evidence for substantial shifts in the composition of the major cell-types of the brain (i.e., excitatory neurons, inhibitory neurons, astrocytes, oligodendrocytes, oligodendrocyte precursor cells, microglia) following VPA-exposure. Whilst our results suggest the adverse functional consequences of VPA-exposure reside in altered synaptic function, they do not exclude the possibility of VPA-induced variation in the proportion of cellular sub-types, which will require single-cell RNA-sequencing of large numbers of cells from VPA-exposed and non-exposed fetal brains to identify.

In contrast to our finding of no measurable difference in the proportion of the major cell-types of the mammalian brain, we observed significant VPA-induced differential gene expression in gestationally exposed pup brains. Pathway enrichment analysis showed a striking separation of enriched terms between the up-and down-regulated genes, with up-regulated genes relating predominantly to cell division, mRNA splicing, translation and extracellular matrix organisation, and down-regulated genes highly enriched for functional processes directly relevant to neurodevelopment including synapse assembly, post-synaptic membrane, regulation of neuronal membrane activity and synaptic transmission.

To further investigate the functional consequences of VPA-induced differential gene expression we integrated differentially expressed genes with GWAS summary statistics from a range of clinically relevant neurodevelopmental diseases and behavioural traits. We observed significant enrichments of

heritability in the set of genes down-regulated by VPA for bipolar disease, schizophrenia, IQ and cross disorder group. We did not see an enrichment for autism spectrum disorder (ASD) which was surprising given that some of the features of VPA-associated neurodevelopmental disability have been likened to autism, and maternal valproate exposure has been associated with autism-like behaviours in non-human primates.[34] This absence of a genetic association with ASD was in stark contrast to the enrichment for bipolar disorder, schizophrenia and IQ and suggests that VPA-associated neurodevelopmental disability may have specific clinical characteristics unique to VPA exposure and points to a requirement for more research to continue to define the clinical phenotype. Notably, the directionality of the effect of VPA-induced gene expression on brain function is consistent with that observed from rare-variant analyses of genetic risk to neurodevelopmental disease where the predominant mechanism is a dominant negative effect from deleterious mutation.[30] The functional enrichment of chromatin assembly/disassembly terms among the genes upregulated by VPA suggests the VPA-induced transcriptional dysregulation is epigenetically encoded and therefore with potential long-lasting consequences for human brain function and behaviour even following birth, prompting the need for further clinical research on the long-term outcomes of children born following fetal VPA exposure.

In addition to substantial differential gene expression, we found that chronic prenatal VPA exposure is associated with differential mRNA splicing in the brain. Interestingly, recent studies have shown evidence that implicates the regulatory role of neuron-specific alternative splicing in neurodevelopmental disorders[35,36] and alternative splicing in the brain is important for several neurological processes including cell differentiation, neurogenesis, synaptogenesis and in the generation of functional neuronal networks.[37] Among the genes observed to be significantly differentially spliced in the prenatal brain following VPA exposure were GAD2, which plays a role GABA-synthesis in neurons[31] and Foxp4, which is known to regulate neurogenesis and is associated with speech delay and congenital abnormalities.[32] These results suggest that alternative splicing may be an additional mechanism for adverse neurodevelopment in VPA-exposed fetal brains.

Apart from valproate, some studies have suggested a risk of neurodevelopmental complications from fetal exposure to other anti-seizure medications, although such risks remain to be fully characterised.[38,39] Additionally, studies have suggested that anti-depressant use during pregnancy may be associated with an increased risk of neurodevelopmental disorders including ASD and ADHD,[41] although not following exposure to antipsychotics.[42] These studies highlight the concern that exists for effects on neurodevelopment for several drugs and classes of drugs commonly prescribed during pregnancy. The research presented here, which demonstrates a robust brain transcriptional response to

VPA that is both functionally and genetically associated with relevant cognitive (IQ) and psychiatric (BP, SCZ, CDG) outcomes suggests that the rat model of chronic dosing followed by transcriptional assay in pup brains might provide a general approach to screening for drug-induced adverse neurodevelopmental effects. Moreover, the inference that the adverse behavioural and cognitive outcomes from gestational VPA exposure arise from transcriptional dysregulation, provides a potential system for testing drugs capable of reversing or ameliorating these changes. For example, as previously highlighted, VPA is a well-recognised histone deacetylase (HDAC) inhibitor, and pre-treatment with methionine has been shown to significantly reduce the incidence of spina bifida and other VPA-associated defects in mice, albeit at the expense of significantly increased embryo lethality.[43]

There are several limitations to our study. First, we did not measure serum VPA levels in the rats although we did follow the same oral dosing schedule previously described[9] which has been shown to achieve a blood level of 180-280 μmol/L and to have antiseizure effects.[44] Second, we did not undertake a dose-response curve across VPA dosages, which has the theoretical potential to establish if there is a "safe" dosage below which the transcriptional effects of VPA do not occur. This is an important clinical unknown highly relevant to women whose seizures are not controlled by any medication other than VPA. We note however, that the blood level achieved by oral dosing used in our study are below the standard therapeutic range for human epilepsy of 346 to 693 μmol/L, suggesting that our results are not an artefact of artificially high VPA levels and also pointing to effects on fetal brain transcription even at relatively low serum levels of valproate. This observation is in keeping with reports of impaired cognition and a 6-fold increase in educational intervention in children born to mothers taking low-dose valproate (<800mg per day) during pregnancy.[45] Third, we did not perform behavioural testing in exposed and un-exposed offspring in this model so we cannot confirm if the transcriptional effects of VPA on the brain directly correlate with changes in behaviour. However, previous research on this animal model has established that chronic oral dosing of valproate at the dosages employed in the present study do induce the expected developmental and morphological abnormalities in pups.[9]

Our study did not address the question of whether a genetically defined subpopulation of women can be identified who are at risk of having children with VPA-associated neurodevelopmental disability. Previous research has observed that women who had given birth to a malformed baby in their first VPA pregnancy are more likely to have a malformed child in their next compared to those who had taken VPA without fetal abnormalities.[4] This suggests that maternal factors, perhaps genetic factors, contribute to VPA-associated congenital malformations. However, we are not aware of similar

studies that have examined the offspring recurrence risk with respect to VPA-associated neurodevelopmental disability. Additionally, the extent to which variation in risk of VPA-associated neurodevelopmental disability (independent of VPA dosage) is explained by maternal genetic variation (e.g., using genome-wide association study methodology or exome-sequencing) is, to date, unexamined. As well as potential maternal genetic effects on risk (for example perhaps mediated by genetic variation in valproate clearance), one can hypothesise a potential effect from fetal genotype on risk as well, for example via genetic effects on VPA-induced neuronal gene expression (so-called "response eQTLs"). Determining if maternal or fetal genetic factors play a role in VPA-associated neurodevelopmental disability is therefore likely to require substantial further research.

In conclusion, the data presented here provide a mechanistic explanation for VPA-induced adverse neurodevelopment anchored in drug-induced transcriptional dysregulation. The extent to which these transcriptional effects are related to irreversible brain development and/or fixed epigenetically encoded changes, or which might be associated with a gradual restoration of normal brain transcription and function over time postnatally is unknown, but could potentially be explored using the experimental paradigm described in this study. Our research prompts the need for longer-term follow up of children born following gestational VPA exposure and the evaluation of other anti-seizure medications as well as dose-response curves in this model using transcriptional readouts.

ACKNOWLEDGEMENTS

This work was supported by the UKRI MRC (Award No: MR/S02638X/1) and by the NIHR Imperial Biomedical Research Centre (BRC) to MR Johnson, and a NHMRC Program (#APP1091593) Grant and NHMRC Investigator Grant (APP1176426) to TJ O'Brien. We acknowledge Ms. Emma Brain for technical assistances with the rat surgeries.

Supplementary material

Supplementary material is available at *Brain* online.

References

Marson AG, Al-Kharusi AM, Alwaidh M, et al. The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. *The Lancet*. 2007;369(9566):1016-1026. doi:10.1016/S0140-6736(07)60461-9

Marson A, Burnside G, Appleton R, et al. The SANAD II study of the effectiveness and cost-effectiveness of valproate versus levetiracetam for newly diagnosed generalised and unclassifiable epilepsy: an open-label, non-inferiority, multicentre, phase 4, randomised controlled trial. *The Lancet*. 2021;397(10282):1375-1386. doi:10.1016/S0140-6736(21)00246-4

Perucca E. Pharmacological and Therapeutic Properties of Valproate. *CNS Drugs*. 2002;16(10):695-714. doi:10.2165/00023210-200216100-00004

Vajda FJ, O'Brien TJ, Graham JE, Lander CM, Eadie MJ. Dose dependence of fetal malformations associated with valproate. *Neurology*. 2013;81(11):999-1003. doi:10.1212/WNL.0b013e3182a43e81

Tomson T, Battino D, Bonizzoni E, et al. Comparative risk of major congenital malformations with eight different antiepileptic drugs: a prospective cohort study of the EURAP registry. *The Lancet Neurology*. 2018;17(6):530-538. doi:10.1016/S1474-4422(18)30107-8

Bromley R, Weston J, Adab N, et al. Treatment for epilepsy in pregnancy: neurodevelopmental outcomes in the child. *Cochrane Database of Systematic Reviews*. 2014;2020(6). doi:10.1002/14651858.CD010236.pub2

Mill J, Petronis A. Pre- and peri-natal environmental risks for attention-deficit hyperactivity disorder (ADHD): the potential role of epigenetic processes in mediating susceptibility. *Journal of Child Psychology and Psychiatry*. 2008;49(10). doi:10.1111/j.1469-7610.2008.01909.x

Christensen J, Grønborg TK, Sørensen MJ, et al. Prenatal Valproate Exposure and Risk of Autism Spectrum Disorders and Childhood Autism. *JAMA*. 2013;309(16):1696. doi:10.1001/jama.2013.2270

Jazayeri D, Braine E, McDonald S, et al. A rat model of valproate teratogenicity from chronic oral treatment during pregnancy. *Epilepsia*. 2020;61(6):1291-1300. doi:10.1111/epi.16536

Chau DKF, Choi AYT, Yang W, Leung WN, Chan CW. Downregulation of glutamatergic and GABAergic proteins in valproric acid associated social impairment during adolescence in mice. *Behavioural Brain Research*. 2017;316. doi:10.1016/j.bbr.2016.09.003

Cui K, Wang Y, Zhu Y, et al. Neurodevelopmental impairment induced by prenatal valproic acid exposure shown with the human cortical organoid-on-a-chip model. *Microsystems & Nanoengineering*. 2020;6(1). doi:10.1038/s41378-020-0165-z

1. Casillas-Espinosa PM, Powell KL, Zhu M, et al. Evaluating whole genome sequence data from the Genetic Absence Epilepsy Rat from Strasbourg and its related non-epileptic strain. *PLOS ONE*. 2017;12(7). doi:10.1371/journal.pone.0179924
4. Marescaux C, Vergnes M, Depaulis A. Genetic absence epilepsy in rats from Strasbourg — A review. In: *Generalized Non-Convulsive Epilepsy: Focus on GABA-B Receptors*. Springer Vienna; 1992. doi:10.1007/978-3-7091-9206-1_4
7. Senn SM, Kantor S, Poulton IJ, et al. Adverse effects of valproate on bone: Defining a model to investigate the pathophysiology. *Epilepsia*. 2010;51(6). doi:10.1111/j.1528-1167.2009.02516.x
9. BITTIGAU P, SIFRINGER M, IKONOMIDOU C. Antiepileptic Drugs and Apoptosis in the Developing Brain. *Ann N Y Acad Sci*. 2003;993(1):103-114. doi:10.1111/j.1749-6632.2003.tb07517.x
11. Bittigau P, Sifringer M, Genz K, et al. Antiepileptic drugs and apoptotic neurodegeneration in the developing brain. *Proceedings of the National Academy of Sciences*. 2002;99(23):15089-15094. doi:10.1073/pnas.222550499
14. Tsoucas D, Dong R, Chen H, Zhu Q, Guo G, Yuan GC. Accurate estimation of cell-type composition from gene expression data. *Nature Communications*. 2019;10(1). doi:10.1038/s41467-019-10802-z
16. Avery LB, Bumpus NN. Valproic Acid Is a Novel Activator of AMP-Activated Protein Kinase and Decreases Liver Mass, Hepatic Fat Accumulation, and Serum Glucose in Obese Mice. *Molecular Pharmacology*. 2014;85(1). doi:10.1124/mol.113.089755
19. Sidhu HS, Srinivas R, Sadhotra A. Evaluate the effects of long-term valproic acid treatment on metabolic profiles in newly diagnosed or untreated female epileptic patients: A prospective study. *Seizure*. 2017;48. doi:10.1016/j.seizure.2017.03.007
22. Parenti I, Rabaneda LG, Schoen H, Novarino G. Neurodevelopmental Disorders: From Genetics to Functional Pathways. *Trends in Neurosciences*. 2020;43(8). doi:10.1016/j.tins.2020.05.004
24. Finucane HK, Bulik-Sullivan B, Gusev A, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics*. 2015;47(11). doi:10.1038/ng.3404
26. Demontis D, Walters RK, Martin J, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics*. 2019;51(1). doi:10.1038/s41588-018-0269-7
28. Stahl EA, Breen G, Forstner AJ, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nature Genetics*. 2019;51(5). doi:10.1038/s41588-019-0397-8
30. Grove J, Ripke S, Als TD, et al. Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*. 2019;51(3). doi:10.1038/s41588-019-0344-8

1. Ripke S, Walters JT, O'Donovan MC, Schizophrenia Working Group of the Psychiatric Genomics Consortium. Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *MedRxiv*. Published online 2020.
4. Savage JE, Jansen PR, Stringer S, et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*. 2018;50(7). doi:10.1038/s41588-018-0152-6
7. Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nature Communications*. 2018;9(1). doi:10.1038/s41467-018-07524-z
8. Lee PH, Anttila V, Won H, et al. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell*. 2019;179(7). doi:10.1016/j.cell.2019.11.020
11. Pulit SL, Stoneman C, Morris AP, et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Human Molecular Genetics*. 2019;28(1). doi:10.1093/hmg/ddy327
14. Johnson MR, Shkura K, Langley SR, et al. Systems genetics identifies a convergent gene network for cognition and neurodevelopmental disease. *Nature Neuroscience*. 2016;19(2):223-232. doi:10.1038/nn.4205
17. Walls AB, Nilsen LH, Eyjolfsson EM, et al. GAD65 is essential for synthesis of GABA destined for tonic inhibition regulating epileptiform activity. *Journal of Neurochemistry*. 2010;115(6). doi:10.1111/j.1471-4159.2010.07043.x
20. Rousso DL, Pearson CA, Gaber ZB, et al. Foxp-Mediated Suppression of N-Cadherin Regulates Neuroepithelial Character and Progenitor Maintenance in the CNS. *Neuron*. 2012;74(2). doi:10.1016/j.neuron.2012.02.024
23. Velez-Ruiz NJ, Meador KJ. Neurodevelopmental Effects of Fetal Antiepileptic Drug Exposure. *Drug Safety*. 2015;38(3). doi:10.1007/s40264-015-0269-9
24. Zhao H, Wang Q, Yan T, et al. Maternal valproic acid exposure leads to neurogenesis defects and autism-like behaviors in non-human primates. *Translational Psychiatry*. 2019;9(1). doi:10.1038/s41398-019-0608-1
28. Porter RS, Jaamour F, Iwase S. Neuron-specific alternative splicing of transcriptional machineries: Implications for neurodevelopmental disorders. *Molecular and Cellular Neuroscience*. 2018;87. doi:10.1016/j.mcn.2017.10.006
31. WANG Y, LIU J, HUANG B, et al. Mechanism of alternative splicing and its regulation. *Biomedical Reports*. 2015;3(2). doi:10.3892/br.2014.407

37. Su CH, D D, Tarn WY. Alternative Splicing in Neurogenesis and Brain Development. *Frontiers in Molecular Biosciences*. 2018;5. doi:10.3389/fmolb.2018.00012
38. Vossler DG. Comparative Risk of Major Congenital Malformations With 8 Different Antiepileptic Drugs: A Prospective Cohort Study of the EURAP Registry. *Epilepsy Currents*. 2019;19(2). doi:10.1177/1535759719835353
39. Meador KJ, Loring DW. Developmental effects of antiepileptic drugs and the need for improved regulations. *Neurology*. 2016;86(3):297-306. doi:10.1212/WNL.0000000000002119
40. Meador KJ, Cohen MJ, Loring DW, et al. Two-Year-Old Cognitive Outcomes in Children of Pregnant Women With Epilepsy in the Maternal Outcomes and Neurodevelopmental Effects of Antiepileptic Drugs Study. *JAMA Neurology*. 2021;78(8):927. doi:10.1001/jamaneurol.2021.1583
41. Boukhris T, Sheehy O, Mottron L, Bérard A. Antidepressant Use During Pregnancy and the Risk of Autism Spectrum Disorder in Children. *JAMA Pediatrics*. 2016;170(2). doi:10.1001/jamapediatrics.2015.3356
42. Wang Z, Chan AYL, Coghill D, et al. Association Between Prenatal Exposure to Antipsychotics and Attention-Deficit/Hyperactivity Disorder, Autism Spectrum Disorder, Preterm Birth, and Small for Gestational Age. *JAMA Internal Medicine*. 2021;181(10). doi:10.1001/jamainternmed.2021.4571
43. Ehlers K, Elmazar MMA, Nau H. Methionine Reduces the Valproic Acid-Induced Spina Bifida Rate in Mice without Altering Valproic Acid Kinetics. *The Journal of Nutrition*. 1996;126(1):67-75. doi:10.1093/jn/126.1.67
44. Al-Roubaie Z, Guadagno E, Ramanakumar A v., Khan AQ, Myers KA. Clinical utility of therapeutic drug monitoring of antiepileptic drugs. *Neurology: Clinical Practice*. 2020;10(4):344-355. doi:10.1212/CPJ.0000000000000722
45. Baker GA, Bromley RL, Briggs M, et al. IQ at 6 years after in utero exposure to antiepileptic drugs: A controlled cohort study. *Neurology*. 2015;84(4):382-390. doi:10.1212/WNL.0000000000001182
46. Andrews S. FastQC: a quality control tool for high throughput sequence data. . Published online 2010.
47. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1). doi:10.1093/bioinformatics/bts635
48. Reinius B, Jazin E. Prenatal sex differences in the human brain. *Molecular Psychiatry*. 2009;14(11):988-989. doi:10.1038/mp.2009.79
49. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1). doi:10.1093/bioinformatics/btp616
50. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Research*. 2019;47(W1):W199-W205. doi:10.1093/nar/gkz401

51. la Manno G, Siletti K, Furlan A, et al. Molecular architecture of the developing mouse brain. *Nature*. 2021;596(7870). doi:10.1038/s41586-021-03775-x
52. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*. 2015;33(5). doi:10.1038/nbt.3192
53. Zeisel A, Hochgerner H, Lönnerberg P, et al. Molecular Architecture of the Mouse Nervous System. *Cell*. 2018;174(4). doi:10.1016/j.cell.2018.06.021

FIGURE LEGENDS

Figure 1 Overview of experimental design and data processing. Two different genetic strains of rat were studied, Genetic Absence Epilepsy Rats from Strasbourg (GAERS) rats and Non-Epileptic (NEC) control rats. Whole brain samples were extracted from valproate-exposed (E-GAERS and E-NEC) and non-exposed (N-GAERS and N-NEC) pups (total n=30). Pup brains were snap frozen in liquid nitrogen and genome-wide gene expression assayed using RNA-sequencing (RNA-seq). Data pre-processing steps included quality control for sequence reads alignment to the rat genome and library size normalization. Differential gene expression and differential splicing analyses were carried out with downstream pathway and heritability enrichment analyses.

Figure 2 Summary of differential gene expression results for the three valproate-exposed vs non-exposed comparisons. Total number of significantly (FDR <0.05) differentially expressed genes (DEGs) for each comparison are shown (single brown dot corresponding to the first three histograms). The number of genes overlapping between each of the case vs control comparisons are shown in black (2-way) or red (3-way). The bars are arranged left to right based on the highest to least number of overlaps respectively.

Figure 3 Pathway enrichment for valproate-exposed vs non-exposed pups. Biological process (BP) gene ontology (GO) parental terms enriched in genes significantly (FDR <0.05) differentially expressed between All-exposed (All-E) vs All non-exposed (All-N) pups for down-regulated (blue bars) and up-regulated (red-bars) genes. For full details including enrichment FDR-values for parental and child terms including Cellular Compartment and Molecular Function GO pathway enrichments see **Supplementary Tables 3a-c**.
*-*log10FDR* values have been signed to indicate enrichment in the down regulated or up-regulated genes.

Figure 4 Heritability enrichment in differentially expressed genes. Linkage Disequilibrium Score Regression **(**LDSC) was used to test for enrichment of heritability in genes significantly differentially expressed for each of the 5 case control comparisons shown. Enrichment -log10(FDR) for the enrichment of genetic association for each trait is indicated by the horizontal bars coloured by the comparison group from which the differentially expressed genes were identified. Vertical line indicates FDR values at 0.05. Genome-wide association studies used for the enrichment analysis are indicated on

the vertical axis. Abbreviations: ASD: Autism Spectrum Disorder, ADHD: Attention Deficit Hyperactivity Disorder, BD: Bipolar Disorder, SCZ: Schizophrenia, IQ: Intelligence Quotient, EPI: Epilepsy, WHR: Waist-to-Hip-Ratio, CDG: Cross-Disorder Group.

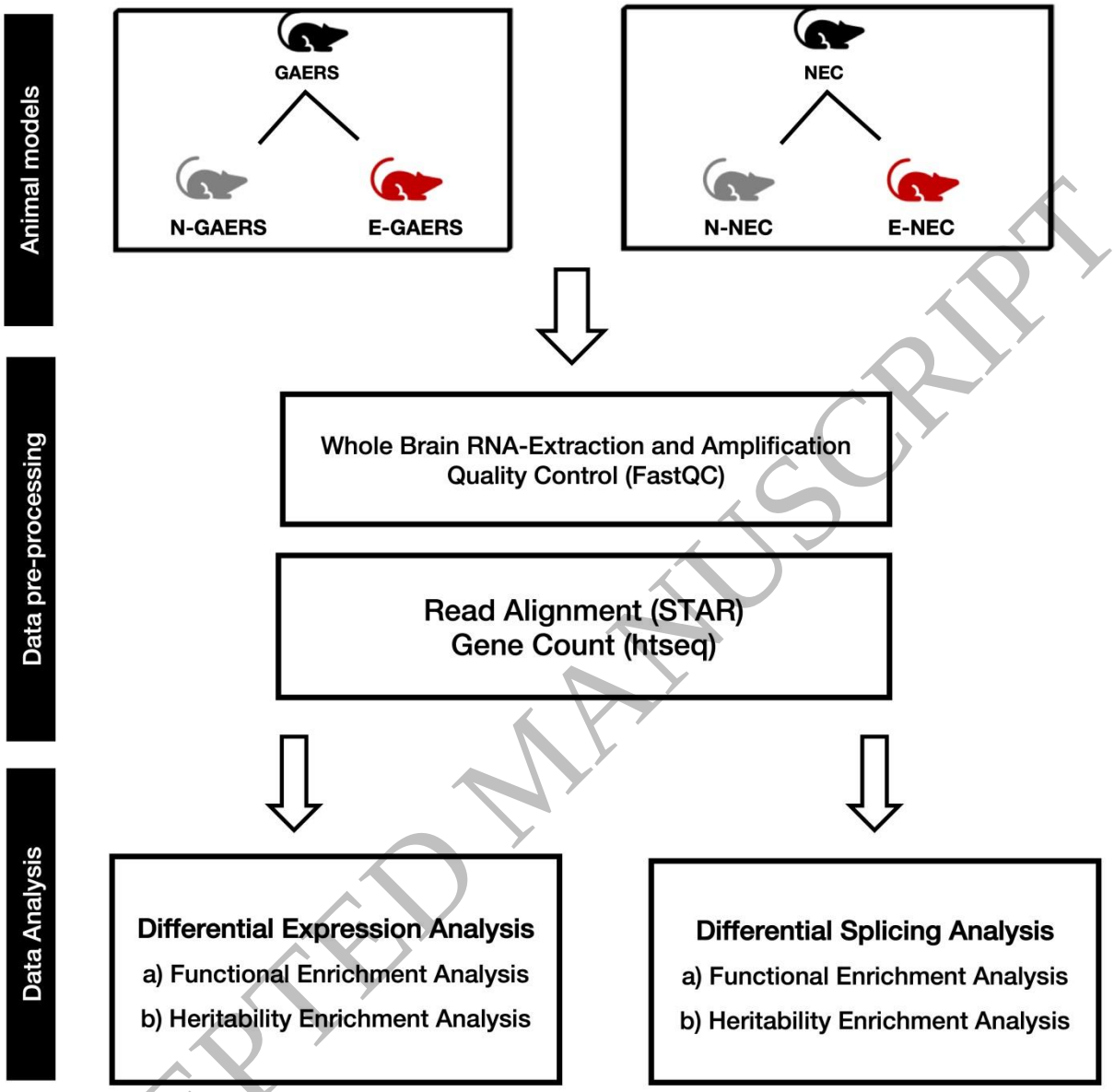

Figure 1
559x471 mm (x DPI)

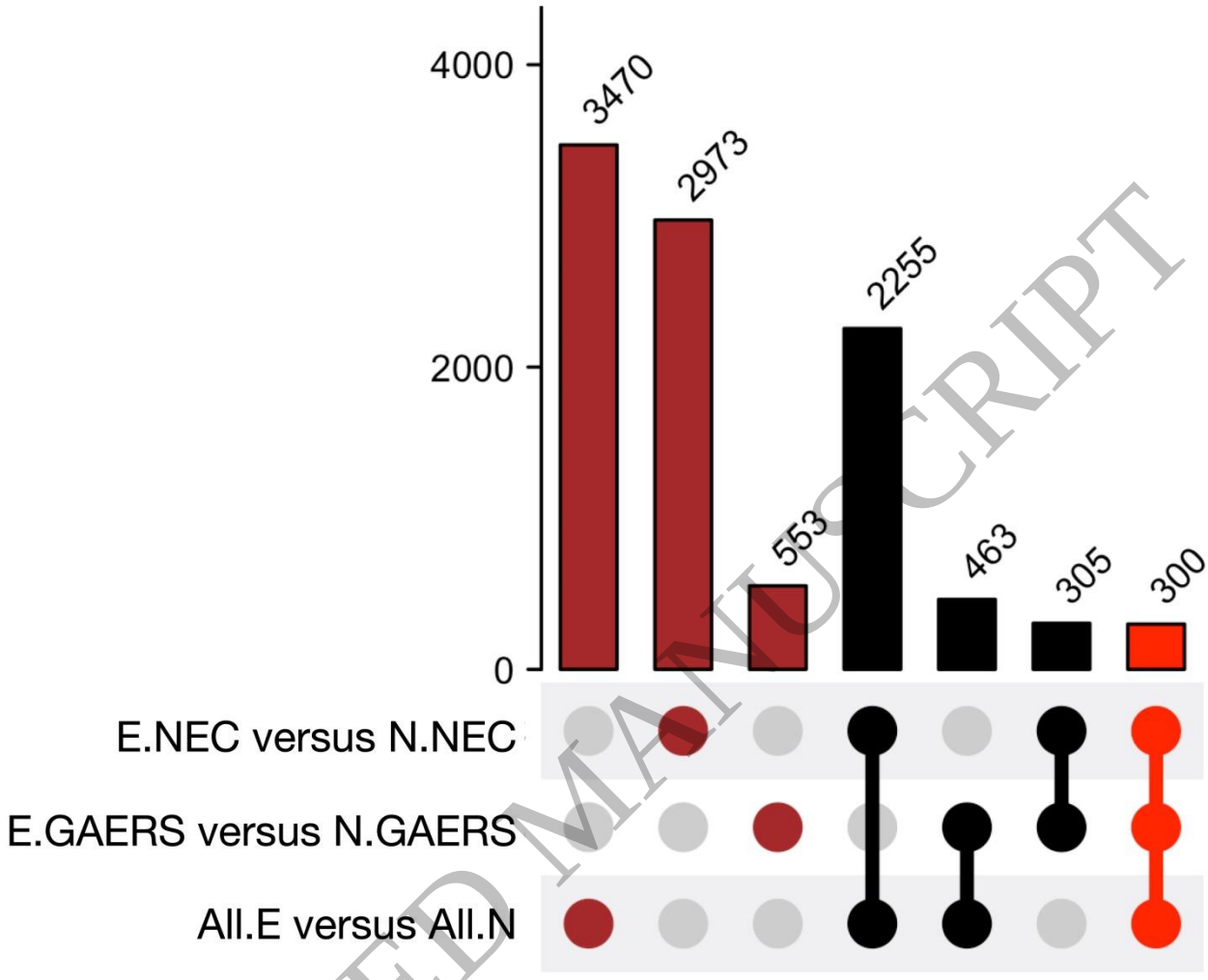

Figure 2
559x474 mm (x DPI)

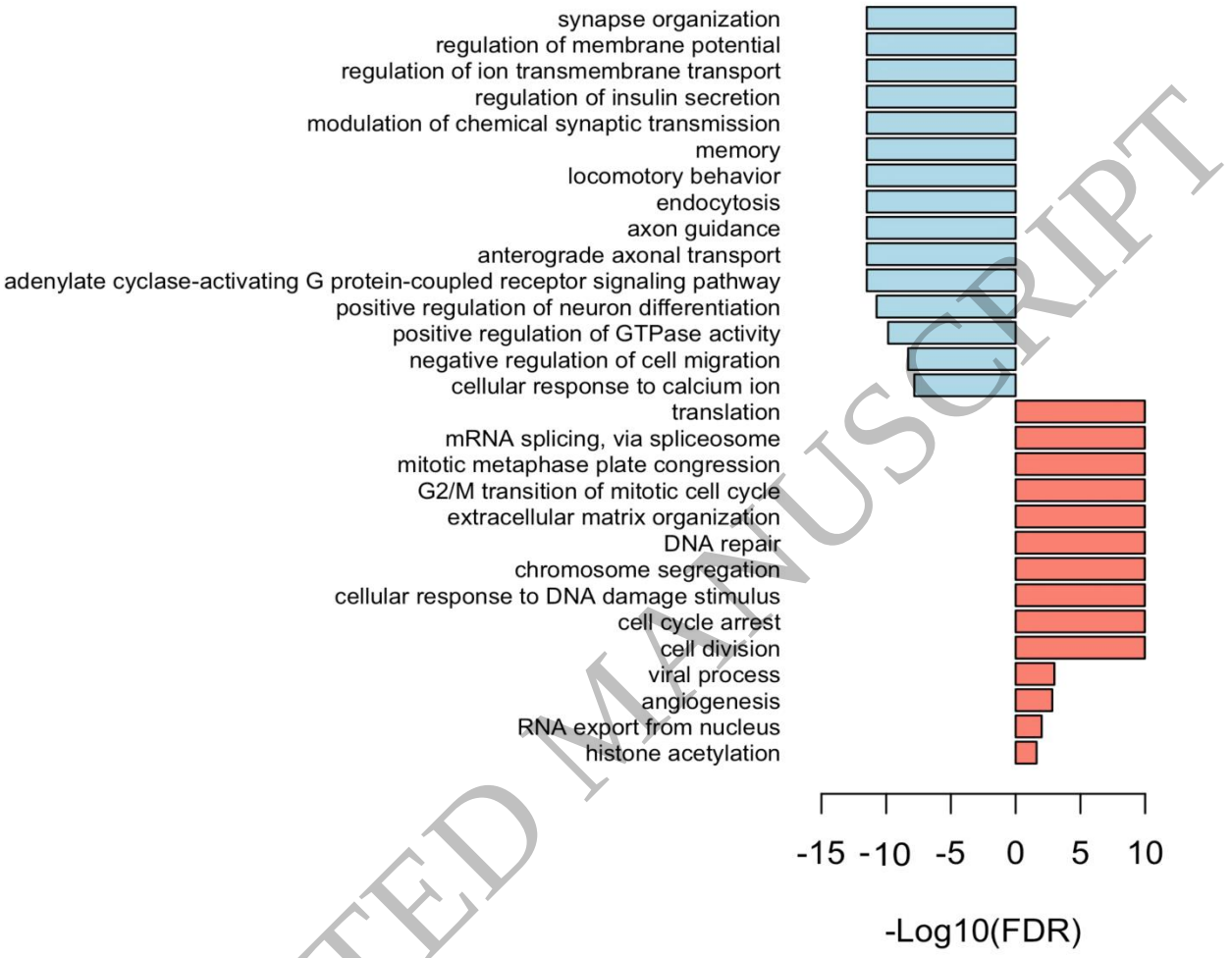

Figure 3
559x314 mm (x DPI)

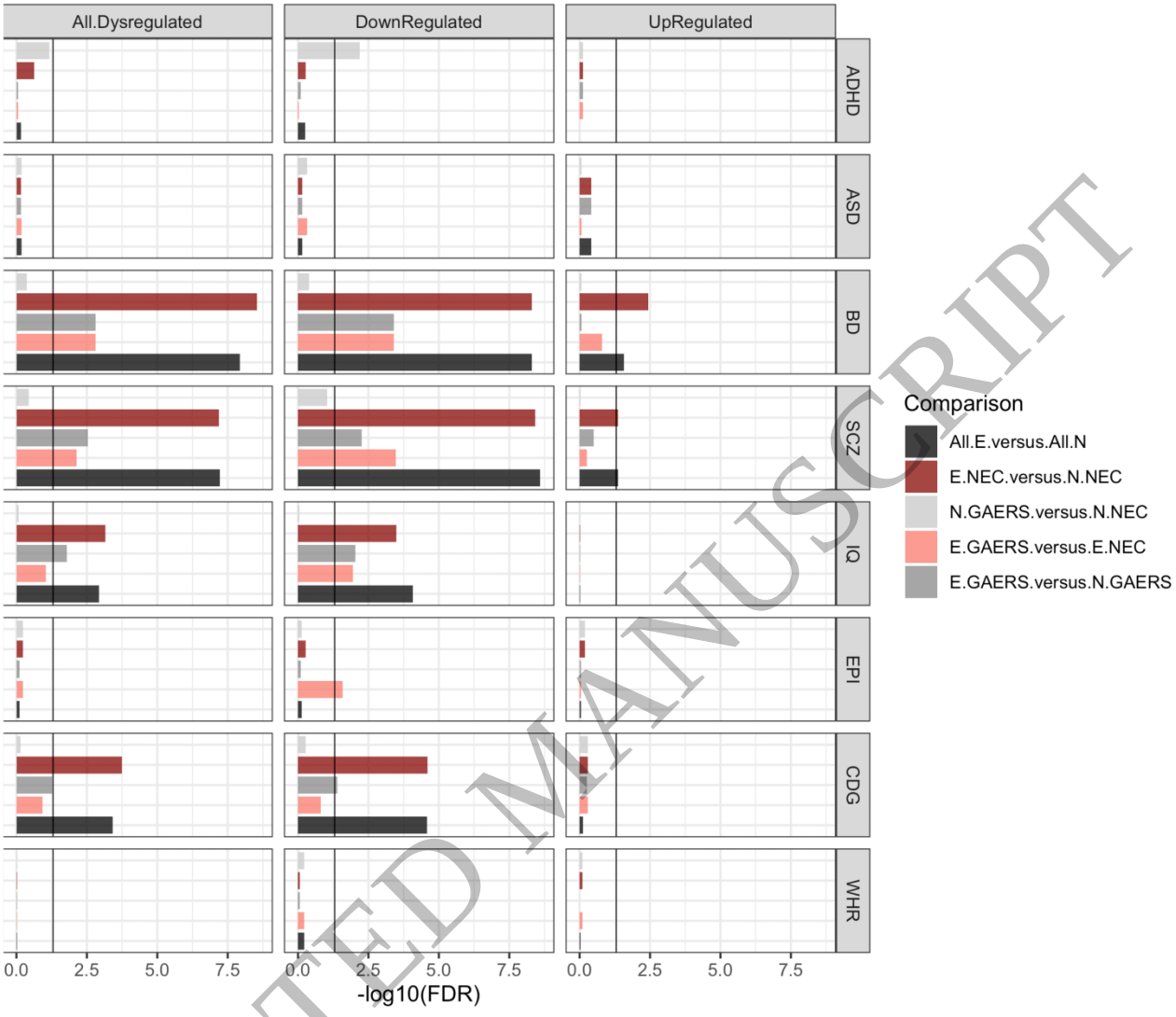

Figure 4
349x334 mm (x DPI)

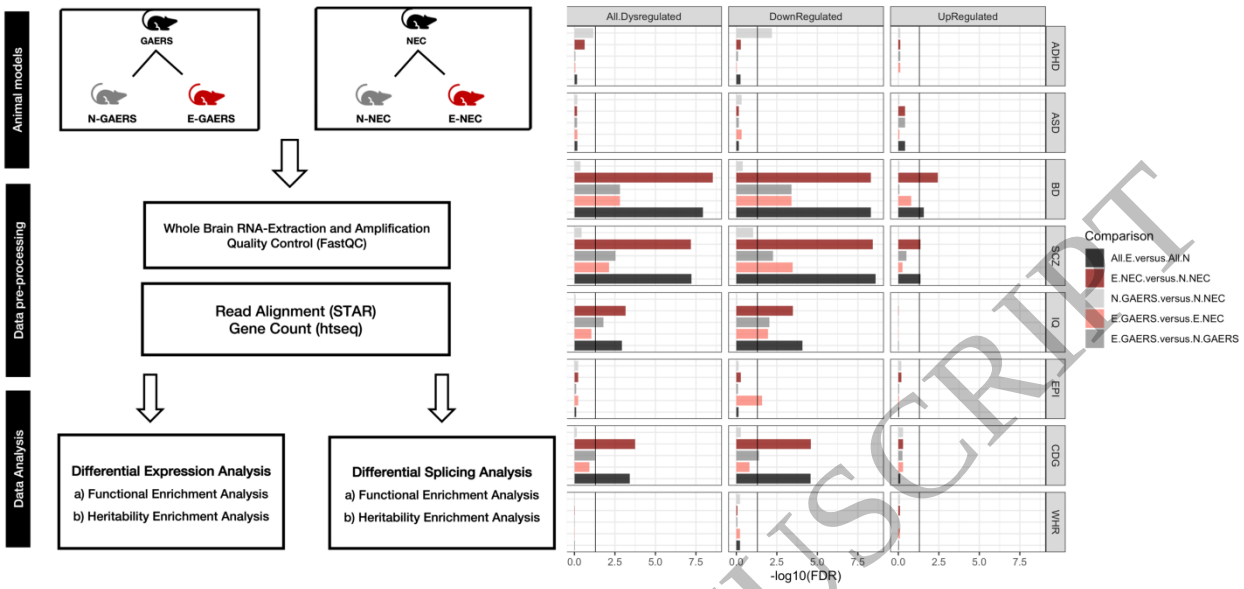

Figure 5
428x348 mm (x DPI)