A mesocorticolimbic dopamine gene network moderates the effect of early adversity on the risk for psychiatric and cardio-metabolic comorbidities

Barbara Barth, MSc[1,2,3], Danusa Mar Arcego, PhD[2,3], Euclides José de Mendonça Filho, PhD[2,3], Randriely Merscher Sobreira de Lima, MSc[4], Sachin Patel[2,3], Omar Khedr[2,3], Zihan Wang, MSc[2,3], Carine Parent, PhD[2,3], Carla Dalmaz, PhD[4], André Krumel Portella, MD, PhD [2,3] Irina Pokhvisneva, MSc[2,3], Michael J Meaney, PhD, CQ, FRSC[2,3,5,6], Patricia Pelufo Silveira, MD, PhD[2,3,5]

1. Integrated Program in Neurosciences, McGill University, Montreal, QC, Canada.
2. Douglas Mental Health University Institute, McGill University, Montreal, QC, Canada.
3. Ludmer Centre for Neuroinformatics and Mental Health, Douglas Research Centre, McGill University, Montreal, QC, Canada
4. Programa de Pós-Graduação em Neurociências, Instituto de Ciências Básicas da Saúde (ICBS), Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.
5. Department of Psychiatry, Faculty of Medicine, McGill University, Montreal, QC, Canada.
6. Translational Neuroscience Programme, Singapore Institute for Clinical Sciences, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore

*Corresponding author:
Patricia Pelufo Silveira, MD, PhD
Department of Psychiatry, Faculty of Medicine, McGill University
Douglas Research Centre, 6875 Boulevard LaSalle, Montreal, QC, H4H 1R3, Canada.
Phone: 514-761-6131 (ext.2776)
Fax: 514-761-6131
patricia.silveira@mcgill.ca

Short/Running title: Dopamine transporter gene network and early life adversity

Keywords: gene network, dopamine transporter gene, comorbidities, early life adversity, psychiatric, cardio-metabolic

Abstract

Psychiatric and cardio-metabolic conditions are commonly comorbid, being a leading cause of disability. Early adversity is a risk factor for both conditions, however, biological pathways remain unknown. The dopamine (DA) system is sensitive to early adversity and influences the development of comorbidities. We hypothesized that early life adversity would functionally link these conditions with the mesocorticolimbic dopamine system as a critical moderator pathway. We computed a co-expression based polygenic score (ePRS) reflecting variations in the function of the dopamine transporter (DAT) gene network in the prefrontal cortex and striatum, the final targets of the mesocorticolimbic pathway. We explored the interaction effects of the ePRS with a score of early life adversity on presence of psychiatric and cardio-metabolic comorbidities in adults (UK Biobank, N= 60016) and adolescents (ALSPAC, N= 910). In adults we also explored genetic and environment effects on gray matter density variations. As predicted, the mesocorticolimbic DAT1 ePRS significantly moderated the impact of early life adversity on the risk for both psychiatric (schizophrenia, neuroticism, mood and substance use disorders) and cardio-metabolic (type 2 diabetes, atherosclerosis, cardiovascular disease) comorbidities in adults and adolescents. Brain gray matter densities in the insula and prefrontal cortex were significantly associated with SNPs from the DAT1 ePRS implicating these regions as critical dopaminergic targets for psychiatric/cardio-metabolic comorbidities. These results reveal that psychiatric and cardio-metabolic comorbidities share common developmental pathways and underlying biological mechanisms.

Introduction

The co-occurrence of more than one chronic disease[1] has high prevalence in primary care settings[2], inflating health care utilization and functional disability[3]. Psychiatric and cardio-metabolic disorders, which are highly comorbid, rank amongst the leading global causes of disability-adjusted life years worldwide[4, 5]. Prospective studies show a bi-directional relationship between psychiatric and cardio-metabolic conditions[6]. Meta-analytic evidence from longitudinal studies indicates that diabetes increases the risk for depression by approximately 25% and that depression increases the risk for type 2 diabetes by 40-60%[5, 7]. The odds for depression also increases with one or more non-psychiatric coexisting chronic conditions, especially coronary artery disease, chronic arthritis, and stroke[8]. Among adult patients with schizophrenia, the prevalence of diabetes averages a prevalence of 15% and this association persists even after controlling for obesity and use of antipsychotic drugs[9].

The underlying mechanism for these co-morbidities remains unknown, but an emerging explanation is that psychiatric and cardio-metabolic disorders share common developmental pathways. Indeed early life adversity increases the risk for a range of psychiatric disorders such as depression[10], addiction[11], attention deficit hyperactivity disorder[12] and schizophrenia[13] and for metabolic diseases such as type II diabetes[14] and cardiovascular disease[15]. An obvious question concerns the biological mechanisms that underlie such a developmental trajectory.

The central dopamine system is highly sensitive to early adversity[16] and proposed as a mechanism for developmental pathways to multiple psychiatric and metabolic comorbidities[17, 18]. Dysfunction of the mesocorticolimbic dopamine system associates with a range of mood disorders and psychosis[19, 20], as well as dysregulated food intake and altered energy homeostasis[21-23]. Dopamine signaling is influenced by core hormones linked to the metabolic syndrome, such as leptin and insulin, through actions on the dopamine transporter (DAT)[21, 24].

We hypothesized that the mesocorticolimbic DAT1 network underlies the association between early life adversity and the comorbidity of psychiatric and cardio-metabolic disorders. We created a mesocorticolimbic co-expression-based polygenic score (DAT1 ePRS) and investigated if the DAT1 ePRS would moderate the association between early adversity and psychiatric/cardio-metabolic comorbidity in adults using the UK Biobank and in adolescents using the Avon Longitudinal Study of Parents and Children (ALSPAC).

Materials and Methods

Detailed description is provided in the Supplementary Materials and Methods.

Participants

We used genomic and phenotypic data from two cohorts, one of adults and one of adolescents: 1) the UK Biobank and 2) Avon Longitudinal Study on Parents and Children (ALSPAC). Informed consent was obtained from each participant, and the projects have been approved by: 1) The North West Multicentre Research 580 Ethics Committee (REC reference 11/NW/0382), the National Information Governance Board for Health and Social Care, and the Community Health Index Advisory Group for UK Biobank; 2) the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees.

Adult cohort: UK Biobank, a large population-based study from the United Kingdom, was used as the adult human cohort[25]. Participants, aged 37-73, were recruited between 2006 and 2010 resulting in 502,543 participants. After all exclusions and inclusion criteria, the number of subjects that remained for the analysis was 60,016 (mean age of 54.51, SD=7.60). For the brain imaging analysis, a subsample of these 60,016 subjects was considered (N=9,291, mean age = 54, SD = 7.371). Detailed description of each inclusion and exclusion criteria and the corresponding sample size at each step can be found in **Supplementary Figure 1.**

Adolescent cohort: To test our hypothesis in a earlier developmental time point with the aim to explore the developmental origins of psychiatric and cardiometabolic comorbidities, we used data from the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort[26-28]. This is a transgenerational prospective observational cohort that recruited 14,451 pregnant women residents in Avon County, UK. Additional recruitment (N=913) was done during later phases, bringing the total sample size to 15,454. The data from the adolescent offspring aged between 15.5 and 17.5 was used in this study. Only subjects with available phenotypic data of interest, early life adversity scores and genotyping data were considered for the analyses (N=910).

Detailed description of each inclusion and exclusion criteria and the corresponding sample size at each step can be found in **Supplementary Figure 2**.

See Supplementary Materials and Methods for details description of the genotyping procedure for each cohort.

Identification of the mesocorticolimbic DAT1co-expression gene network and ePRS calculation

Figure 1 shows the steps involved in the identification of the gene networks and the calculation of the ePRS score. The ePRS was calculated considering genes co-expressed with the DAT1 gene in the PFC and Striatum, the two brain regions representing the final targets of the mesocorticolimbic pathway structure. We aimed to identify DAT1 co-expression networks within each brain region, with the final ePRS being a joint representation of the functional co-expression networks in these two mesocorticolimbic regions. As described previously[29-33], we began by using brain region-specific RNA sequencing data from mouse available at GeneNetwork (http://genenetwork.org/)[34] to identify DAT1 gene co-expression networks (absolute value of co-expression correlation with DAT1 gene greater or equal to 0.5), which were then converted to human orthologs by using the biomaRt package[35]. GeneNetwork was used to obtain gene expression from rodents since our previous findings demonstrated an effect of early life adversity on dopaminergic mesocorticolimbic system in rodents[22, 23, 36]. Next, since we were interested in gene networks that were active during the developmental period in which adversity occurred (early life) and active when the brain is still undergoing core maturational processes in humans, we used BrainSpan to select autosomal transcripts expressed at least 1.5-fold more during fetal and child development (0–60 months after birth) in comparison to adulthood (20–40 years of age). Only common genes between the GeneNetwork human orthologs and BrainSpan derived lists were retained. These processes resulted in two lists of

DAT1 co-expressed genes (210 genes for the PFC and 51 genes for the striatum). No gene overlap was observed between the two lists. Our final list included 261 genes (**Supplementary Table 1**). We then mapped SNPs in the human ortholog genes comprising the mesocorticolimbic DAT1 gene network using the GRCh37.p13 assembly of the NCBI to source chromosome and start/ end position, and to gather all the gene-SNP pairs from the GTEx dataset in human PFC and Striatum. These lists were merged with the genotyping data from UK Biobank and ALSPAC cohort, retaining only the common SNPs and subjecting the final SNP list to linkage disequilibrium clumping ($r2 < 0.2$) within 500kb radious, which resulted in 7,940 and 1,330 independent functional SNPs retained for the PFC and striatum lists respectively in the UK Biobank cohort.

To calculate the mesocorticolimbic DAT1 ePRS, the number of effect alleles (genotype information from the study samples) at a given SNP was weighted using the estimated brain-region-specific effect of the genotype on gene expression from the GTEx data[37]. We also accounted for the direction of the co-expression of each gene with DAT1 by multiplying the weight by -1 in case the expression of a gene was negatively correlated with the expression of the DAT1 gene. The sum of the weighted values from all SNPs for each individual in the cohorts resulted in the region-specific PFC and striatum scores, that were than summed in one score. The mesocorticolimbic DAT1 expression-based polygenic score (ePRS)[30, 32, 33] reflects the variation of gene expression of the genes co-expressed with the DAT1 gene in the PFC and striatum in a combined manner. The mesocorticolimbic DAT1 ePRS score was used as a continuous variable in this study.

<u>Validation/specificity of the findings</u>: Using the same steps described above, control ePRS were created to test the validity and specificity of the mesocorticolimbic DAT1 ePRS to moderate the

effect of early life adversity on the study outcomes. The PFC DAT1 ePRS and the striatum DAT1 ePRS were used separately in the gene by environment model to test if the summed effect of these two brain regions is more informative than the individual scores. We also tested a VTA DAT1 ePRS (number of genes = 83, number of SNPs = 3.182) based on the high density of dopaminergic neurons on this region[38], to further test the specificity of the brain regions chosen for the mesocorticolimbic DAT1 ePRS. For this score, the substantia nigra gene expression information from GTEx was used to weight the ePRS score as there is no available data for the VTA brain region in GTEx.

Functional enrichment analysis: To characterize the biological functions associated with the mesocorticolimbic DAT1 co-expression gene network, functional enrichment analysis was performed using MetaCore® software from Clarivate Analytics (https://portal.genego.com). SNPs that compose the mesocorticolimbic DAT1 ePRS were used in the analysis. The significance was considered for the false discovery rate (FDR) adjusted p-value <0.05. To investigate network centrality measures, protein-protein (PPI) network interactions were identified using the STRING database (https://string-db.org)[39] and the mesocorticolimbic DAT1 ePRS genes. These elements interaction were then visualized using the Cytoscape® software[40]. The nodes are the elements of a network and edges are the connection between these elements. Bottleneck genes are defined as having high betweenness, hub genes are defined as genes with high degree (nodes with more connections). To analyze the topological properties associated with this gene network, the CentiScaPe app in Cytoscape® was used to calculate the degrees and betweenness (genes that act as ‘bridges’ between nodes in a network). We used this information to define the “hub genes” within the network, characterized as nodes with degrees higher than +1SD above the mean; and the “bottlenecks” characterized as nodes with betweenness higher than +1SD above the mean. Genes that are both bottlenecks and hubs were considered the central

nodes of the network[41]. In order to investigate if the genes that composed the mesocorticolimbic DAT1 ePRS have similar patterns of gene co-expression during different life periods in humans, we used the gene expression data from human samples extracted from the BrainSpan database[42]. Genetic correlations between the SNPs included in the score and other genome-wide association studies, accounting for linkage disequilibrium between SNPs was done using LD Hub[43] (http://ldsc.broadinstitute.org/ldhub/). We also used FUMA for GWAS enrichment analysis [44] (**Supplementary Materials and Methods)**.

Early Life Adversity Score

A cumulative measure combining multiple indicators of early life adversity according to the methods demonstrated previously[33] was calculated for UK Biobank and ALSPAC participants. (**Table 1** and **Supplementary Materials and Methods**).

Outcome measures

Adult cohort: diagnostic terms from across all participants hospital inpatient records, coded according to the International Classification of Diseases version 10 (ICD-10)[45] were used. Selected diagnostic terms from chapter V Mental and behavioural disorders, chapter IV Endocrine, nutritional and metabolic diseases and chapter IX Diseases of the circulatory system were included. The presence of at least one cardio-metabolic diagnosis and at least one psychiatric diagnosis was considered a "co-morbidity" case. Comorbidity was coded as a binary variable (1="yes" or 0="no"). T1 structural brain MRI pre-processed imaging data were generated by an image-processing pipeline developed and run on behalf of the UK Biobank[46] (**Supplementary Materials and Methods**).

Adolescent cohort: Presence of psychiatric and metabolic disorders are less common at early stages of development[47]. Since no diagnoses for described above cardio-metabolic and psychiatric disorders were available in ASLPAC cohort and to maintain consistency with the

outcome tested in UK Biobank adult cohort, we created a comorbidity risk variable by cluster analysis, considering known indicators of risk to develop psychiatric and metabolic disorders later in life. Total difficulties score measured by the Strengths and Difficulties Questionnaire, depression and anxiety scores measured by Computerized Interview Schedule – Revised (CIS-R), Homeostatic Model Assessment of Insulin Resistance (HOMA-IR), and waist circumference (cm) were considered. The resulting cluster membership, which represented comorbidity risk, was coded as a binary variable (1= "yes comorbidity" or 0= "no comorbidity"). (**Supplementary Materials and Methods, Supplementary Table 2**, **Supplementary Figure 3**).

Statistical Analysis

Statistical analysis were performed using R[48]. For the descriptive statistics, the ePRS groups were defined by median split, and a comparison between low and high ePRS groups was done using student t test for continuous variables and a chi-square test for categorical variables (**Table 2** and **Table 3**). Significance levels for all measures were set at $p < 0.05$.

The comorbidity risk variable by cluster analysis was performed using the mclust package[49]. This algorithm uses a model-based classification and density estimation of the z-standardized variables based on finite Gaussian mixture modelling. The method assumes that inputted variables can be explained by an underlying latent categorical cluster variable that represents distinct profiles within the sample, both in a qualitative and quantitative manner. A cluster size solution of two was retained, representing two distinct profiles within the sample associated with the inputted variables. All variables were z-standardized prior to entering the clustering procedure. Regression analysis was carried out to test if the means for each variable were different between the two identified cluster groups indicating two distinct profiles

(**Supplementary Table 2**). Sex and social economic status were used as covariates, because of their know influence on the variables included in the cluster analysis[50-52].

The gene by environment (GxE) interaction was explored by logistic regression analysis. The early life adversity score, the mesocorticolimbic DAT1 ePRS and the interaction term between them were used as predictors for both cohorts. In the UK Biobank analysis were adjusted by sex, age, the first forty genetic population structure principal components, genotyping array, and assessment center (**Table 2**). In the ALSPAC analysis, adjusted by sex, the first ten genetic population structure principal components and breastfeeding duration for being significant at baseline comparisons (**Table 3**). The analysis was also adjusted for zBMI at 15.5 years of age to diminish confounding factors while analyzing metabolic variables (as the ones considered in the comorbidity risk variable). In case of significant gene by environment (GxE) interactions, post hoc simple slope analysis was performed to investigate how the environment effect varies as a function of the genetic background[53].

Gray matter density from the UK Biobank participants was analyzed as an outcome in a multivariate parallel independent component analysis (pICA). This analysis was applied to identify the effect of early life adversity on the relationship between two different data modalities in a data-driven manner[54]. This analysis estimates the maximum independent components within each data modality separately while also maximizing the association between modalities using an entropy term based on information theory[54]. Each SNP that composes the mesocorticolimbic DAT1 ePRS weighted by GTEx (genotype * GTEx gene expression slope for each SNP) and whole brain voxel based gray matter density were used in the analysis. Weighted SNPs were adjusted for the most significant principal components from population stratification (ancestry). The subjects were split in two groups according to the early life adversity score (yes or no adversity present). A significant effect of age was observed on the gray matter density (r=-0.59,

$p<2.2e-16$), suggesting that as age progresses, a significant loss of gray matter density is seen in the UK Biobank participants (**Supplementary Figure 4**). To account for the age effect, the participants were separated in blocs of 5 years of age (from 40 to 70 years old), resulting in 6 blocs, as the pICA analysis does not allow inclusion of co-variates. The Fusion ICA Toolbox (http://mialab.mrn.org/software/fit/) within MATLAB® R2019 was used to run the analysis. The number of independent components was estimated using minimum description length criteria[54] for the MRI modality and SNP dimensionality inside the toolbox for the genetic modality. Components for both modalities were converted to z-scores and a threshold at $|Z| > 2.5$ was used to identify significant brain regions and SNPs that contributed the most for the component overall pattern. Significant brain regions were further analyzed in terms of frequency across the 6 groups analyzed, to narrow down important brain regions. Loading coefficients, which describe the presence of the identified component across subjects[54] were extracted for each component, modality and subject. The mean subject-specific loading coefficients of these components between subjects from high and low prenatal adversity groups was compared using Student's t-test. To identify the anatomical classification of brain areas included in the significant MRI components, Talairach coordinates were used. The significant SNPs ($|Z| > 2.5$) from the significant genetic components were analyzed using MetaCore®, to identify associated underlying biological pathways and gene ontology processes (**Figure 3**).

Results

The mesocorticolimbic DAT1 gene network is co-expressed in humans

To investigate if mouse-generated DAT1 gene network was co-expressed in humans, we explored the gene co-expression patterns of the mesocorticolimbic DAT1 gene network throughout human development using gene expression data from human postmortem samples[42]. A high co-expression was expected in childhood, as the mesocorticolimbic DAT1 gene network was enriched for genes over expressed in this life period (see **Figure 1A** and **Supplementary Materials and Methods**,). However, prominent gene co-expression clusters were also seen in adolescents and adults (**Figure 1B and 1C).** When visualizing and exploring the network properties (**Figure 1D**), 12 genes were considered central (**Figure 1E**), and NPM1 was the central node. PIK3CA was among the most important genes serving as a bridge between other nodes. Among the hub genes, HNRNPA1 was the gene with the highest degree. The main pathways associated with the network include epigenetic regulation, activation of the PI3K pathway and regulation of the mTOR transcription factor, comprising regulation of translation, cell proliferation and oxidative stress (**Supplementary Figure 5**).

Mesocorticolimbic DAT1 ePRS moderates the association of early life adversity with the risk for psychiatric/cardio-metabolic comorbidities in adults

In agreement with the literature, our early adversity score was strongly associated with the presence of co-morbidities in both UK Biobank (b= 0.442, Odds ratio (OR)= 1.556, 95% confidence interval (CI):1.492– 1.620, P <0.001) and ALSPAC (b= 0.201, OR= 1.222, 95% (CI): 1.087 –1.358, P =0.003). For the UK Biobank there were no significant main association of the ePRS with co-morbidity risk (b= 0.022, OR= 1.021, 95% (CI): 0.939 – 1.102, P= 0.603). In contrast, and consistent with our anticipated hypothesis, there was a significant interaction

between the mesocorticolimbic DAT1 ePRS and early life adversity on the presence of psychiatric/cardio-metabolic co-morbidities in UK Biobank adults (*b*= -0,069, OR = 0.933, 95% (CI): 0.868 – 0.997, *P* = 0.036, *n* = 60016) (**Figure 2A**).

In an earlier developmental time point (ALSPAC adolescents), the GxE model revealed a significant interaction between the mesocorticolimbic DAT1 ePRS and early life adversity on the presence of psychiatric/metabolic comorbidity risk (*b* = 0.180, OR= 1.198, 95% (CI): 1.060–1.335, *P*= 0.009, *n*=910) (**Figure 2B**). Similar to the findings in adults, there was no significant main association of the ePRS with the comorbidity risk (*b*= 0.105, OR= 1.110, 95% (CI): 0.945 – 1.275, *P*= 0.150). These results indicate a developmental trajectory, in which early indicators of risk to develop psychiatric and metabolic comorbidities in adulthood can be seen in adolescents as a function of the mesocorticolimbic DAT1 co-expression gene network and early life environmental conditions.

To further validate and test the specificity of our findings, the PFC DAT1 ePRS and the striatum DAT1 ePRS were used separately in the GxE model to test if the summed effect of these two brain regions is more informative than the individual scores. No significant interaction was found between early life adversity and the PFC DAT1 ePRS on the probability of having psychiatric and cardio-metabolic comorbidities in adults (*b*= -0.023, OR= 0.977, 95% (CI): 0.913 – 1.041, *P*= 0.477). The interaction model using the striatum DAT1 ePRS and early life adversity showed a non-significant trend (*b*= -0.023, OR= 0.937, 95% (CI): 0.872 – 1.002, *P*= 0.054) in predicting presence of psychiatric and cardio-metabolic comorbidities in the UK Biobank cohort. We also used the ventral tegmental area (VTA) DAT1 ePRS as another control score, as this is a brain region dense in dopaminergic neurons[38]. No significant interaction was found in adults (*b*= -0.020, OR= 0.980, 95% (CI):0.921 – 1.045, *P*= 0.549). As a whole, we conclude that the results

from the GxE model using the mesocorticolimbic DAT1 ePRS are specific to this gene network and to the combination of the PFC and striatum gene network in adulthood.

In the adolescent ALSPAC cohort, the PFC DAT1 ePRS (*b*= 0.071, OR= 1.074, 95% (CI): 0.830 – 1.111, *P*= 0.282) and the VTA DAT1 ePRS (*b*= -0.076, OR=0.926, 95% (CI): 1.190 – 0.890, *P*= 0.257) interactions with early adversity were not significant for the presence of co-morbidity risk. However, the striatal DAT1 ePRS showed a significant interaction with early adversity in predicting the presence of a psychiatric/cardio-metabolic comorbidity risk in adolescence (b = 0.165, OR = 1.179, 95% (CI): 0.597 – 1.246, *P*= 0.021). This suggests that the striatal DAT1 gene network may also have an independent role in the emergence of these co-morbidities in response to early life adversity during adolescence before the prefrontal dopaminergic connections are fully established[55].

Mesocorticolimbic DAT1 gene network SNPs are enriched in genome-wide association studies for both psychiatric and metabolic conditions

To address whether the genes included in the network were relevant for psychiatric and metabolic illnesses we compared our network with genes identified in GWAS of psychiatric and metabolic disorders. We observed that the mesocorticolimbic DAT1 ePRS was significantly enriched for MDD GWAS (ρ = 0.23, p = 0.021). When examining genetic correlations between SNPs from the mesocorticolimbic DAT1 ePRS and SNPs from GWASes from the LD Hub catalog[43], we observed genetic correlations with neuroticism (rg= 0.353; *P*= 0.0021), free cholesterol in small VLDL (rg= 0.328; *P*= 0.042) and triglycerides in very large VLDL (rg=0.248; *P*= 0.050). This analysis confirms that the genetic variation associated with the mesocorticolimbic DAT1 co-expression network is relevant for both psychiatric and metabolic phenotypes.

SNPs from the mesocorticolimbic DAT1 ePRS are related to gray matter variations in the insula and prefrontal cortex

To establish the anatomo-functional relevance of our mesocorticolimbic DAT1 gene network, we investigated correlations between the SNPs from the mesocorticolimbic DAT1 ePRS and voxel-based gray matter density in UK Biobank participants using a multivariate parallel independent component analysis (pICA)[54] (**Figure 3** and **Supplementary Materials and Methods**). Since a significant correlation of age was observed in gray matter density (**Supplementary Figure 4**), analyses were carried out in 6 groups of participants separated in blocks of 5 years of age (from 40 to 70 years old). Only the first significantly linked pair of components with higher correlation index values was selected for each age block to be further explored (**Supplementary Table 3**). A significant association with adversity was observed in components at different ages. In these cases, the relation between SNPs and gray matter were moderated by the early life adversity score (see **Supplementary Table 4** for a full list of significant SNPs per age band). These subsets of significant SNPs are related to variations in gray matter density in the insula, inferior and medial frontal gyrus, striatum (pallidum, putamen and caudate), parahippocampal gyrus and cingulate gyrus across ages (see **Supplementary Table 5** for a full list of significant brain regions from all age bands). Enrichment analysis of these significant SNPs using MetaCore® (**Supplementary Table 6**) showed that the most common pathway maps identified are involved in synaptic plasticity and transmission through the action of Ephrin-B and the cell cycle. Gene ontology enrichment revealed central nervous system and neuron development, neurogenesis, regulation of glutamate receptor signaling pathway, response to monoamine and catecholamine and regulation of insulin secretion involved in the cellular response to glucose (see **Supplementary Table 6** for a full list).

Discussion

Our study suggests that exposure to early life adversity increases the risk for later co-morbidities between cardio-metabolic and psychiatric conditions in adulthood [10-14], and our novel functional genomics approach provides evidence for the mesocorticolimbic DAT1 gene network as a salient mechanism moderating this association. This finding is aligned with the critical role of the dopaminergic system in environmental responsivity[30, 56].

Our enrichment analysis results showed that the mesocorticolimbic DAT1 gene network is highly co-expressed across childhood, adolescence and adulthood. This finding is consistent with our finding of a moderating correlation for the DAT1 ePRS in both adolescents and adults. The difference in simple slopes seen between adolescents and adults (**Figure 2**) suggests that the mesocorticolimbic DAT1 ePRS leads to a differential timing for the emergence of the co-morbidity phenotype in response to early adversity, which can be important to inform specific vulnerabilities for prevention measures.

We also identified candidate biological mechanisms underlying the moderating correlation of the mesocorticolimbic DAT1 ePRS on cardio-metabolic – psychiatric comorbidities, for example mTOR and PI3K, which are involved in intracellular insulin signaling. The mTOR (mammalian target of rapamycin) pathway is activated in insulin-induced long-term depression of DA neurons in the VTA[57]. A central gene in our network, PIK3C, operates in both the PI3K and mTOR pathways[58]. Insulin-induced activation of the PI3K/Akt/mTOR pathway is an important protective factor for cardiomyocytes apoptosis, implicated in myocardial injury associated with metabolic disturbances[59]. Insulin resistance is a risk factor for cardio-metabolic and brain-based disorders, including type II diabetes, cardiovascular disease, Alzheimer disease, and major depressive disorder[60, 61]. Insulin modulates mesocorticolimbic DA neurotransmission through different mechanisms, one of which is to increase the capacity of DA reuptake by DAT by

activating the phosphatidylinositol (PI) 3-kinase[24, 62]. Insulin also reduces DA release in rodent nucleus accumbens and medial prefrontal cortex slices[63] . Taken together, the evidence suggests that insulin modulation of dopamine function may be an underlying mechanism by which the mesocorticolimbic DAT1 gene network moderates the risk for developing psychiatric and cardio-metabolic co-morbidities in response to early life adversity.

Interestingly, the subset of SNPs from the DAT1 ePRS that was related to gray matter density in all age bands in our anatomical-functional correlation analysis is also associated with the regulation of insulin secretion and a cellular response to glucose and insulin stimulus. This subset of insulin-related SNPs from our DAT1 ePRS was correlated with gray matter density variations in the insula, medial and inferior frontal gyrus, striatum (lentiform nucleus and caudate) and parahippocampal gyrus. The insula is involved in emotional recognition and processing, cognition and motivation[64]. Altered insular gray matter volumes occur in patients with type II diabetes and are associated with increased anxiety and depression[65] as well as cognitive impairment[66]. Importantly, insulin resistance moderates the relationship between childhood abuse and insula-precuneus connectivity in depressed and overweight youth[67].. Reduced gray matter volume in the parahippocampal gyrus is a hallmark of schizophrenia[68] and functional connectivity between the parahippocampal gyrus and prefrontal areas is positively correlated with self-reported childhood adversity in adults[69], further implicating these regions as a critical site for adversity-related cardio-metabolic – psychiatric conditions.

Specificity analysis revealed that the GxE results in the UK Biobank cohort were apparent only for the mesocorticolimbic DAT1 gene network, in comparison to other important sites for dopamine transmission. However, in adolescents, the striatum DAT1 ePRS also showed a significant interaction with early adversity, suggesting an independent role in the emergence of co-morbidities in response to early life adversity. This agrees with evidence of prolonged

mesocorticolimbic dopamine axons growth (from the striatum to the PFC) during adolescence[55]. Due to this late maturation, striatal dopaminergic axons are especially vulnerable to environmental effects during development[70]. The Netrin-1 receptor, DCC, is one of the main molecules responsible for axonal guiding, determining the axons' extent of growth and final target in the PFC[71]. The DCC gene is indeed present in our mesocorticolimbic DAT1 gene network (**Supplementary Table 1**).

Given that DAT1 gene is expressed in dendrites and cell bodies of dopaminergic neurons and DAT concentration is higher in the VTA and substantia nigra[38], one might expect significant associations with a VTA DAT1 gene co-expression network. In our study, however, the terminal gene networks (PFC and striatum) were strong moderators of the association between early adversity and the co-morbidities, while the VTA co-expression network showed no significant associations. mRNA axonal transport and protein synthesis at the terminal is an important mechanism for regulation of neurotransmitter synthesis and reuptake[72], and could explain the presence of DAT1 mRNA at the terminals, consistently detected in numerous human post-mortem studies[38, 73-75]. Indeed, one of the central nodes of our network is HNRNPA1 (**Figure 1E**), a gene associated with mRNA transport and synthesis[76], suggesting that DAT1 mRNA transport to terminals may be a key mechanistic feature of our DAT1 mesocorticolimbic gene network. DAT1 protein vesicular traffic has a limited contribution to DAT1 concentration in synapses[77], hence other forms of regulation of DAT1 availability in terminals – for instance via mRNA axonal transport and terminal protein synthesis – may be in place, in agreement to our findings.

Our ePRS does not consider intronic regions, potentially ignoring important regulatory elements. Moreover, our developmental results are based on cross sectional studies, and further longitudinal data are needed to better describe this trajectory.

We observed that the association between environmental and genetic factors can place individuals at risk for adult co-morbid chronic conditions from an early age, and that dopaminergic neurotransmission has a central role in moderating the association of the early environment with the risk for these diseases. These findings open opportunities for designing more effective interventions and prevention strategies, especially with the growing support from precision medicine.

Acknowledgments

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC and UK Biobank teams, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. Funding: The UK Medical Research Council and Wellcome (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). The research using ALSPAC was specifically funded by the Wellcome Trust and MRC (Grant ref: 076467/Z/05/Z) and Wellcome Trust (Grant ref: 08426812/Z/07/Z). GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe.

This project was supported by The JPB Foundation through a grant to The JPB Research Network on Toxic Stress: A Project of the Center on the Developing Child at Harvard University. This work was also funded by the Fonds de recherche du Québec – Santé and Canadian Institutes of Health Research (CIHR, PJT-166066, PI Silveira PP). This publication is the work of the authors and Barbara Barth and Patricia P. Silveira will serve as guarantors for the contents of this paper.

Conflict of Interest

BB, DMA, EJMF, RML, SP, OK, ZW, IP, CP, AKP, CD, MJM and PPS report no biomedical financial interests or potential conflicts of interest.

References

1. Prince MJ, Wu F, Guo Y, Gutierrez Robledo LM, O'Donnell M, Sullivan R *et al.* The burden of disease in older people and implications for health policy and practice. *Lancet* 2015; **385**(9967)**:** 549-562.

2. Uijen AA, van de Lisdonk EH. Multimorbidity in primary care: prevalence and trend over the last 20 years. *Eur J Gen Pract* 2008; **14 Suppl 1:** 28-32.

3. Egede LE. Major depression in individuals with chronic medical disorders: prevalence, correlates and association with health resource utilization, lost productivity and functional disability. *Gen Hosp Psychiatry* 2007; **29**(5)**:** 409-416.

4. Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The lancet* 2016; **388**(10053)**:** 1545-1602.

5. Kyu HH, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N *et al.* Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 2018; **392**(10159)**:** 1859-1922.

6. Golden SH, Lazo M, Carnethon M, Bertoni AG, Schreiner PJ, Diez Roux AV *et al.* Examining a bidirectional association between depressive symptoms and diabetes. *JAMA* 2008; **299**(23)**:** 2751-2759.

7. Moulton CD, Pickup JC, Ismail K. The link between depression and diabetes: the search for shared mechanisms. *The Lancet Diabetes & Endocrinology* 2015; **3**(6)**:** 461-471.

8. Lotfaliany M, Bowe SJ, Kowal P, Orellana L, Berk M, Mohebbi M. Depression and chronic diseases: Co-occurrence and communality of risk factors. *J Affect Disord* 2018; **241:** 461-468.

9. Regenold WT, Thapar RK, Marano C, Gavirneni S, Kondapavuluru PV. Increased prevalence of type 2 diabetes mellitus among psychiatric inpatients with bipolar I affective and schizoaffective disorders independent of psychotropic drug use. *J Affect Disord* 2002; **70**(1)**:** 19-26.

10. Martins-Monteverde CMS, Baes CVW, Reisdorfer E, Padovan T, Tofoli SMdC, Juruena MF. Relationship Between Depression and Subtypes of Early Life Stress in Adult Psychiatric Patients. *Front Psychiatry* 2019; **10:** 19-19.

11. Enoch M-A. The role of early life stress as a predictor for alcohol and drug dependence. *Psychopharmacology* 2011; **214**(1)**:** 17-31.

12. Lahti M, Eriksson J, Heinonen K, Kajantie E, Lahti J, Wahlbeck K *et al.* Late preterm birth, post-term birth, and abnormal fetal growth as risk factors for severe mental disorders from early to late adulthood. *Psychological medicine* 2015; **45**(5)**:** 985-999.

13. Abel KM, Wicks S, Susser ES, Dalman C, Pedersen MG, Mortensen PB *et al.* Birth weight, schizophrenia, and adult mental disorder: is risk confined to the smallest babies? *Archives of general psychiatry* 2010; **67**(9)**:** 923-930.

14. Hales CN, Barker DJ. Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. *Diabetologia* 1992; **35**(7)**:** 595-601.

15. Rich-Edwards JW, Kleinman K, Michels KB, Stampfer MJ, Manson JE, Rexrode KM *et al.* Longitudinal study of birth weight and adult body mass index in predicting risk of coronary heart disease and stroke in women. *Bmj* 2005; **330**(7500)**:** 1115.

16. Gatzke-Kopp LM. The canary in the coalmine: The sensitivity of mesolimbic dopamine to environmental adversity during development. *Neuroscience & Biobehavioral Reviews* 2011; **35**(3)**:** 794-803.

17. Gragnoli C, Reeves G, Reazer J, Postolache T. Dopamine–prolactin pathway potentially contributes to the schizophrenia and type 2 diabetes comorbidity. *Translational psychiatry* 2016; **6**(4)**:** e785-e785.

18. Nousen EK, Franco JG, Sullivan EL. Unraveling the mechanisms responsible for the comorbidity between metabolic syndrome and mental health disorders. *Neuroendocrinology* 2013; **98**(4)**:** 254-266.

19. Pizzagalli DA, Berretta S, Wooten D, Goer F, Pilobello KT, Kumar P *et al.* Assessment of striatal dopamine transporter binding in individuals with major depressive disorder: in vivo positron emission tomography and postmortem evidence. *JAMA psychiatry* 2019; **76**(8)**:** 854-861.

20. Kesby J, Eyles D, McGrath J, Scott J. Dopamine, psychosis and schizophrenia: the widening gap between basic and clinical neuroscience. *Translational psychiatry* 2018; **8**(1)**:** 1-12.

21. Jones KT, Woods C, Zhen J, Antonio T, Carr KD, Reith ME. Effects of diet and insulin on dopamine transporter activity and expression in rat caudate-putamen, nucleus accumbens, and midbrain. *Journal of neurochemistry* 2017; **140**(5)**:** 728-740.

22. Laureano DP, Alves MB, Miguel PM, Machado TD, Reis AR, Mucellini AB *et al.* Intrauterine growth restriction modifies the accumbal dopaminergic response to palatable food intake. *Neuroscience* 2018.

23. Alves MB, Dalle Molle R, Desai M, Ross MG, Silveira PP. Increased palatable food intake and response to food cues in intrauterine growth-restricted rats are related to tyrosine hydroxylase content in the orbitofrontal cortex and nucleus accumbens. *Behavioural brain research* 2015; **287:** 73-81.

24. Garcia BG, Wei Y, Moron JA, Lin RZ, Javitch JA, Galli A. Akt is essential for insulin modulation of amphetamine-induced human dopamine transporter cell-surface redistribution. *Molecular pharmacology* 2005; **68**(1)**:** 102-109.

25. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015; **12**(3)**:** e1001779.

26. Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G *et al.* Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International journal of epidemiology* 2013; **42**(1)**:** 97-110.

27. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J *et al.* Cohort profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *International journal of epidemiology* 2013; **42**(1)**:** 111-127.

28. Golding J, Pembrey M, Jones R. ALSPAC--the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatric and perinatal epidemiology* 2001; **15**(1)**:** 74-87.

29. Hari Dass SA, McCracken K, Pokhvisneva I, Chen LM, Garg E, Nguyen TTT *et al.* A biologically-informed polygenic score identifies endophenotypes and clinical conditions associated with the insulin receptor function on specific brain regions. *EBioMedicine* 2019; **42:** 13.

30. Miguel PM, Pereira LO, Barth B, de Mendonça Filho EJ, Pokhvisneva I, Nguyen TT *et al.* Prefrontal cortex dopamine transporter gene network moderates the effect of perinatal hypoxic-ischemic conditions on cognitive flexibility and brain gray matter density in children. *Biological psychiatry* 2019; **86**(8)**:** 621-630.

31. Restrepo-Lozano JM, Pokhvisneva I, Wang Z, Patel S, Meaney MJ, Silveira PP *et al.* Corticolimbic DCC gene co-expression networks as predictors of impulsivity in children. *Molecular Psychiatry* 2022**:** 1-9.

32. Silveira PP, Pokhvisneva I, Parent C, Cai S, Rema ASS, Broekman BFP *et al.* Cumulative prenatal exposure to adversity reveals associations with a broad range of neurodevelopmental outcomes that are moderated by a novel, biologically informed polygenetic score based on the serotonin transporter solute carrier family C6, member 4 (SLC6A4) gene expression. *Development psychopathology* 2017; **29**(5)**:** 1601-1617.

33. de Lima RMS, Barth B, Arcego DM, de Mendonça Filho EJ, Clappison A, Patel S *et al.* Amygdala 5-HTT Gene Network Moderates the Effects of Postnatal Adversity on Attention Problems: Anatomo-Functional Correlation and Epigenetic Changes. *Front Neurosci* 2020; **14:** 198.

34. Mulligan MK, Mozhui K, Prins P, Williams RW. GeneNetwork: A Toolbox for Systems Genetics. *Methods Mol Biol* 2017; **1488:** 75-120.

35. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature protocols* 2009; **4**(8)**:** 1184.

36. Alves MB, Laureano DP, Dalle Molle R, Machado TD, de Ascenção Salvador AP, Miguel PM *et al.* Intrauterine growth restriction increases impulsive behavior and is associated with altered dopamine transmission in both medial prefrontal and orbitofrontal cortex in female rats. *Physiology & behavior* 2019; **204:** 336-346.

37. Consortium GT. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013; **45**(6)**:** 580-585.

38. Ciliax BJ, Drash GW, Staley JK, Haber S, Mobley CJ, Miller GW *et al.* Immunocytochemical localization of the dopamine transporter in human brain. *Journal of Comparative Neurology* 1999; **409**(1)**:** 38-56.

39. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* 2013; **41**(Database issue)**:** D808-815.

40. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 2003; **13**(11)**:** 2498-2504.

41. Neves de Oliveira B-H, Dalmaz C, Zeidán-Chuliá F. Network-based identification of altered stem cell pluripotency and calcium signaling pathways in metastatic melanoma. *Medical Sciences* 2018; **6**(1)**:** 23.

42. Miller JA, Ding S-L, Sunkin SM, Smith KA, Ng L, Szafer A *et al.* Transcriptional landscape of the prenatal human brain. *Nature* 2014; **508**(7495)**:** 199.

43. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, Haycock PC *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* 2017; **33**(2)**:** 272-279.

44. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017; **8**(1)**:** 1826.

45. Organization WH. ICD-10 : international statistical classification of diseases and related health problems : tenth revision. 2 edn2004.

46. Alfaro-Almagro F, Jenkinson M, Bangerter NK, Andersson JL, Griffanti L, Douaud G *et al*. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* 2018; **166:** 400-424.

47. de Ferranti SD, Gauvreau K, Ludwig DS, Neufeld EJ, Newburger JW, Rifai N. Prevalence of the metabolic syndrome in American adolescents: findings from the Third National Health and Nutrition Examination Survey. *Circulation* 2004; **110**(16)**:** 2494-2497.

48. R CT. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing: Vienna, Austria, 2014.

49. Scrucca L, Fop M, Murphy TB, Raftery AE. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R journal* 2016; **8**(1)**:** 289.

50. Tramunt B, Smati S, Grandgeorge N, Lenfant F, Arnal J-F, Montagner A *et al.* Sex differences in metabolic regulation and diabetes susceptibility. *Diabetologia* 2020; **63**(3)**:** 453-461.

51. Altemus M, Sarvaiya N, Neill Epperson C. Sex differences in anxiety and depression clinical perspectives. *Front Neuroendocrinol* 2014; **35**(3)**:** 320-330.

52. Pickett KE, Wilkinson RG. Income inequality and health: a causal review. *Social science & medicine* 2015; **128:** 316-326.

53. Bauer DJ, Curran PJ. Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate behavioral research* 2005; **40**(3)**:** 373-400.

54. Pearlson GD, Liu J, Calhoun VD. An introductory review of parallel independent component analysis (p-ICA) and a guide to applying p-ICA to genetic data and imaging phenotypes to identify disease-associated biological pathways and systems in common complex disorders. *Front Genet* 2015; **6:** 276.

55. Vosberg DE, Leyton M, Flores C. The Netrin-1/DCC guidance system: dopamine pathway maturation and psychiatric disorders emerging in adolescence. *Mol Psychiatry* 2020; **25**(2)**:** 297-307.

56. Dahoun T, Nour MM, McCutcheon RA, Adams RA, Bloomfield MA, Howes OD. The relationship between childhood trauma, dopamine release and dexamphetamine-induced positive psychotic symptoms: a [11 C]-(+)-PHNO PET study. *Translational psychiatry* 2019; **9**(1)**:** 1-12.

57. Labouebe G, Liu S, Dias C, Zou H, Wong JC, Karunakaran S *et al.* Insulin induces long-term depression of ventral tegmental area dopamine neurons via endocannabinoids. *Nat Neurosci* 2013; **16**(3)**:** 300-308.

58. Janku F, Wheler JJ, Naing A, Falchook GS, Hong DS, Stepanek VM *et al.* PIK3CA mutation H1047R is associated with response to PI3K/AKT/mTOR signaling pathway inhibitors in early-phase clinical trials. *Cancer research* 2013; **73**(1)**:** 276-284.

59. Yao H, Han X, Han X. The cardioprotection of the insulin-mediated PI3K/Akt/mTOR signaling pathway. *American Journal of Cardiovascular Drugs* 2014; **14**(6)**:** 433-442.

60. Watson KT, Simard JF, Henderson VW, Nutkiewicz L, Lamers F, Rasgon N *et al.* Association of Insulin Resistance With Depression Severity and Remission Status: Defining a Metabolic Endophenotype of Depression. *JAMA psychiatry* 2021; **78**(4)**:** 439-441.

61. Watson K, Nasca C, Aasly L, McEwen B, Rasgon N. Insulin resistance, an unmasked culprit in depressive disorders: Promises for interventions. *Neuropharmacology* 2018; **136:** 327-334.

62. Carvelli L, Morón JA, Kahlig KM, Ferrer JV, Sen N, Lechleiter JD *et al.* PI 3□kinase regulation of dopamine uptake. *Journal of neurochemistry* 2002; **81**(4)**:** 859-869.

63. Schoffelmeer AN, Drukarch B, De Vries TJ, Hogenboom F, Schetters D, Pattij T. Insulin modulates cocaine-sensitive monoamine transporter function and impulsive behavior. *Journal of Neuroscience* 2011; **31**(4)**:** 1284-1291.

64. Namkung H, Kim SH, Sawa A. The Insula: An Underestimated Brain Area in Clinical Neuroscience, Psychiatry, and Neurology: (Trends in Neuroscience 40, 200-207, 2017). *Trends Neurosci* 2018; **41**(8)**:** 551-554.

65. Roy B, Ehlert L, Mullur R, Freeby MJ, Woo MA, Kumar R *et al.* Regional brain gray matter changes in patients with type 2 diabetes mellitus. *Scientific reports* 2020; **10**(1)**:** 1-13.

66. Zhang D, Lei Y, Gao J, Qi F, Yan X, Ai K *et al.* Right Frontoinsular Cortex: A Potential Imaging Biomarker to Evaluate T2DM-Induced Cognitive Impairment. *Frontiers in Aging Neuroscience* 2021; **13:** 255.

67. Sun KL, Watson KT, Angal S, Bakkila BF, Gorelik AJ, Leslie SM *et al.* Neural and endocrine correlates of early life abuse in youth with depression and obesity. *Front Psychiatry* 2018; **9:** 721.

68. Glahn DC, Laird AR, Ellison-Wright I, Thelen SM, Robinson JL, Lancaster JL *et al.* Meta-analysis of gray matter anomalies in schizophrenia: application of anatomic likelihood estimation and network analysis. *Biological psychiatry* 2008; **64**(9)**:** 774-781.

69. Wade-Bohleber L, Boeker H, Grimm S, Gärtner M, Ernst J, Recher D *et al.* Depression is associated with hyperconnectivity of an introspective socio-affective network during the recall of formative relationship episodes. *Journal of Affective Disorders* 2020; **274:** 522-534.

70. Hoops D, Flores C. Making dopamine connections in adolescence. *Trends in neurosciences* 2017; **40**(12)**:** 709-719.

71. Reynolds LM, Pokinko M, Torres-Berrío A, Cuesta S, Lambert LC, Pellitero EDC *et al.* DCC receptors drive prefrontal cortex maturation by determining dopamine axon targeting in adolescence. *Biological psychiatry* 2018; **83**(2)**:** 181-192.

72. Dalla Costa I, Buchanan CN, Zdradzinski MD, Sahoo PK, Smith TP, Thames E *et al.* The functional organization of axonal mRNA transport and translation. *Nature Reviews Neuroscience* 2020**:** 1-15.

73. Choi CS, Hong M, Kim KC, Kim J-W, Yang SM, Seung H *et al.* Effects of atomoxetine on hyper-locomotive activity of the prenatally valproate-exposed rat offspring. *Biomolecules & therapeutics* 2014; **22**(5)**:** 406.

74. Jijun L, Zaiwang L, Anyuan L, Shuzhen W, Fanghua Q, Lin Z *et al.* Abnormal expression of dopamine and serotonin transporters associated with the pathophysiologic mechanism of Tourette syndrome. *Neurology India* 2010; **58**(4)**:** 523.

75. Sun Z, Ma Y, Xie L, Huang J, Duan S, Guo R *et al.* Behavioral changes and neuronal damage in rhesus monkeys after 10 weeks of ketamine administration involve prefrontal cortex dopamine D2 receptor and dopamine transporter. *Neuroscience* 2019; **415:** 97-106.

76. Dreyfuss G, Kim VN, Kataoka N. Messenger-RNA-binding proteins and the messages they carry. *Nature reviews Molecular cell biology* 2002; **3**(3)**:** 195-205.

77. Bagalkot TR, Block ER, Bucchin K, Balcita-Pedicino JJ, Calderon M, Sesack SR *et al.* Dopamine transporter localization in medial forebrain bundle axons indicates its long-range transport primarily by membrane diffusion with a limited contribution of vesicular traffic on retromer-positive compartments. *Journal of Neuroscience* 2021; **41**(2)**:** 234-250.

Figure 1.

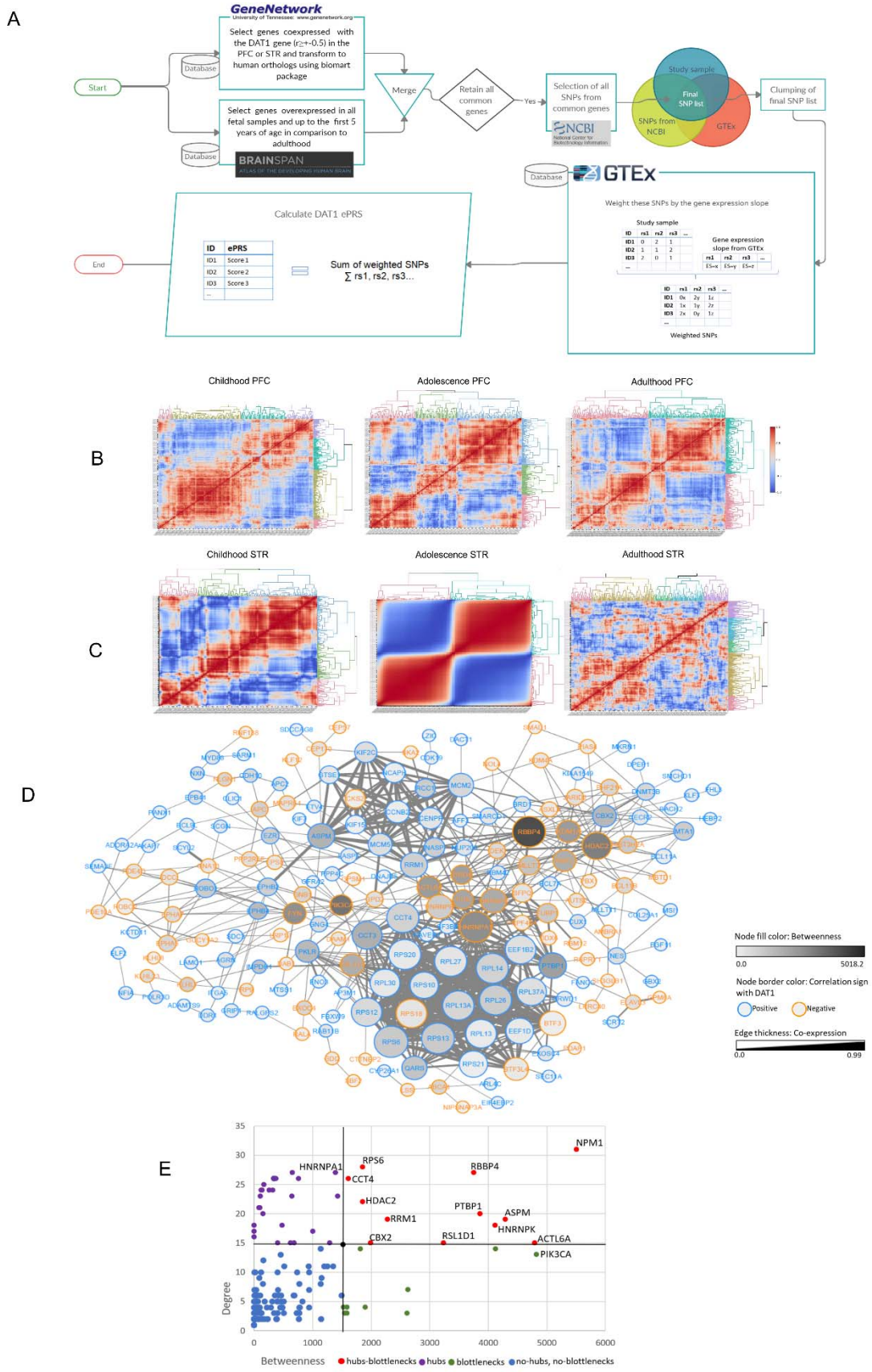
A
GeneNetwork
Select genes coexpressed with the DAT1 gene (r≥+-0.5) in the PFC or STR and transform to human orthologs using biomart package
Select genes overexpressed in all fetal samples and up to the first 5 years of age in comparison to adulthood
BRAINSPAN
Start
Database
Merge
Retain all common genes
Yes
Selection of all SNPs from common genes
NCBI
Study sample
Final SNP list
SNPs from NCBI
GTEx
Clumping of final SNP list
GTEx
Weight these SNPs by the gene expression slope
Study sample
Gene expression slope from GTEx
Weighted SNPs
Calculate DAT1 ePRS
ID ePRS
Sum of weighted SNPs
∑ rs1, rs2, rs3...
End
B
Childhood PFC
Adolescence PFC
Adulthood PFC
C
Childhood STR
Adolescence STR
Adulthood STR
D
Node fill color: Betweenness
0.0
5018.2
Node border color: Correlation sign with DAT1
Positive
Negative
Edge thickness: Co-expression
0.0
0.99
E
NPM1
HNRNPA1
RPS6
CCT4
RBBP4
HDAC2
PTBP1
ASPM
RRM1
HNRNPK
CBX2
RSL1D1
ACTL6A
PIK3CA
Degree
Betweenness
hubs-blottlenecks
hubs
blottlenecks
no-hubs, no-blottlenecks

Figure 1. Construction and characterization of the mesocorticolimbic DAT1 gene network

A, calculation of the expression based polygenic score (ePRS) from genes co-expressed with the DAT1 gene in the PFC and striatum (mesocorticolimbic DAT1 ePRS). Steps were followed separately for the two brain regions. GeneNetwork was used to generate the list of genes co-expressed with the DAT1 in the PFC or striatum in mice, which were then converted to human ortologs. BrainSpan was used to identify genes overexpressed within these same brain regions in all fetal samples and up to 5 years of age in comparison to adult samples. Common genes between GeneNetwork and BrainSpan lists were retained.

All the SNPs from these genes, common between the study sample, NCBI and GTEx databases were retained and included in the final list of SNPs. This final list was subjected to linkage disequilibrium clumping, with removal of highly correlated SNPs. Next, for each SNP, a number of alleles at a given SNP (rs1, rs2…) was multiplied by the estimated effect of the genotype-gene expression association from GTEx. The sum of these values over all SNPs provides the PFC or striatum DAT1 ePRS. These two ePRS were summed to compose the mesocorticolimbic DAT1 ePRS. **B**, co-expression of genes included in the mesocorticolimbic DAT1 gene network in the PFC and **C**, striatum in humans at different ages according to BrainSpan. **D,** protein-protein interactions of the mesocorticolimbic DAT1 co-expression network. Nodes represent genes and edges represent protein-protein interactions and co-expression according to STRING. The color of the node represents the correlation sign with the DAT1 gene according to GeneNetwork co-expression matrix (orange represents negative and blue positive correlation). Size of the node demonstrates the number of connections (total degree). Node's filling intensity depicts betweenness (number of times a node acts as a bridge between nodes). **E**, topological properties of the mesocorticolimbic DAT1 gene network, showing hubs (with degrees higher than +1SD above mean), bottlenecks (betweenness higher than +1SD above the mean), and hub-bottlenecks.

Lines in black indicate mean + 1 SD for degrees and betweenness. The first 3 genes with higher degree in the network (NPM1, RBBP4, ASPM) have a role in cell proliferation, DNA replication and neuron migration respectively. The gene with highest betweenness was PIK3CA important for insulin signaling. Among the hub genes, HNRNPA1 was the gene with the highest degree, involved in the packaging of pre-mRNA into particles and transport from the nucleus to the cytoplasm, as well as splicing.

Figure 2.

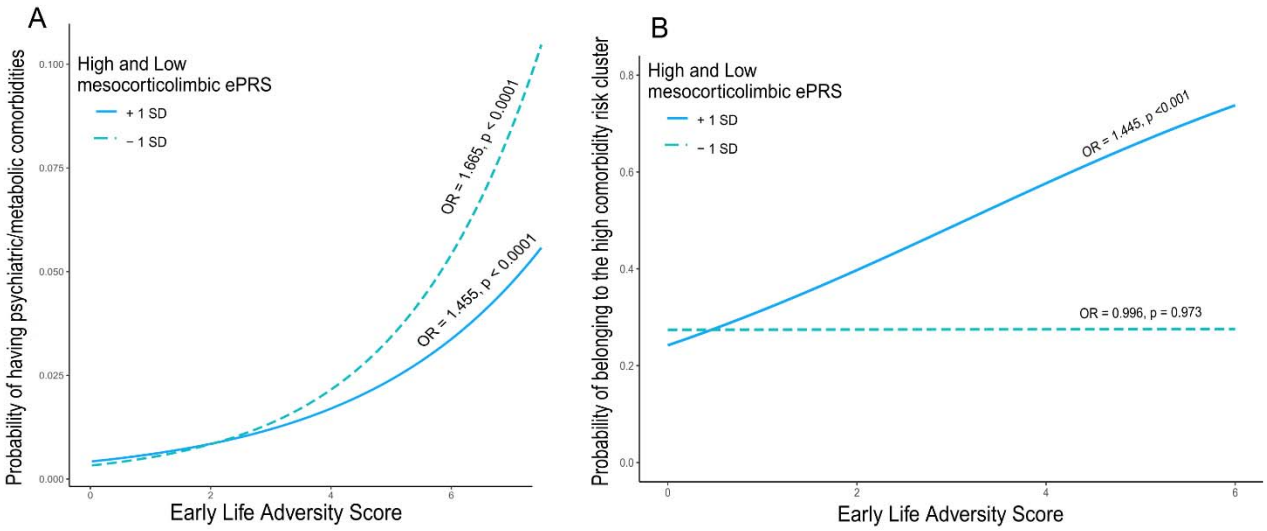

Figure 2. Mesocorticolimbic DAT1 ePRS moderates the effect of early life adversity on the risk for psychiatric and cardio-metabolic comorbidities. Probability of having comorbidities in individuals with high and low mesocorticolimbic DAT1 ePRS as a function of Early Life Adversity Score. **A**, UK Biobank cohort, N=60016. The risk for comorbidity increases as early life adversity increases, especially at lower ePRS values (low: *b*=0.509, $OR_{Low\ ePRS}$= 1.66, 95% CI:1.576 – 1.753, *P*= 1.42E-29; high: *b*= 0.374, $OR_{High\ ePRS}$= 1.45, 95% CI:1.358 – 1.542, *P*= 1.03E-15). **B**, ALSPAC cohort, N= 910. The risk for comorbidity increases as early life adversity increases, especially at higher ePRS values (low: *b*= -0.003, $OR_{Low\ ePRS}$= 0.996, 95% CI: 0.788 – 1.204, *P*= 0.973; high: b= 0.368, $OR_{High\ ePRS}$= 1.455, 95% CI: 1.258 – 1.633, *P*< 0.001).

Figure 3.

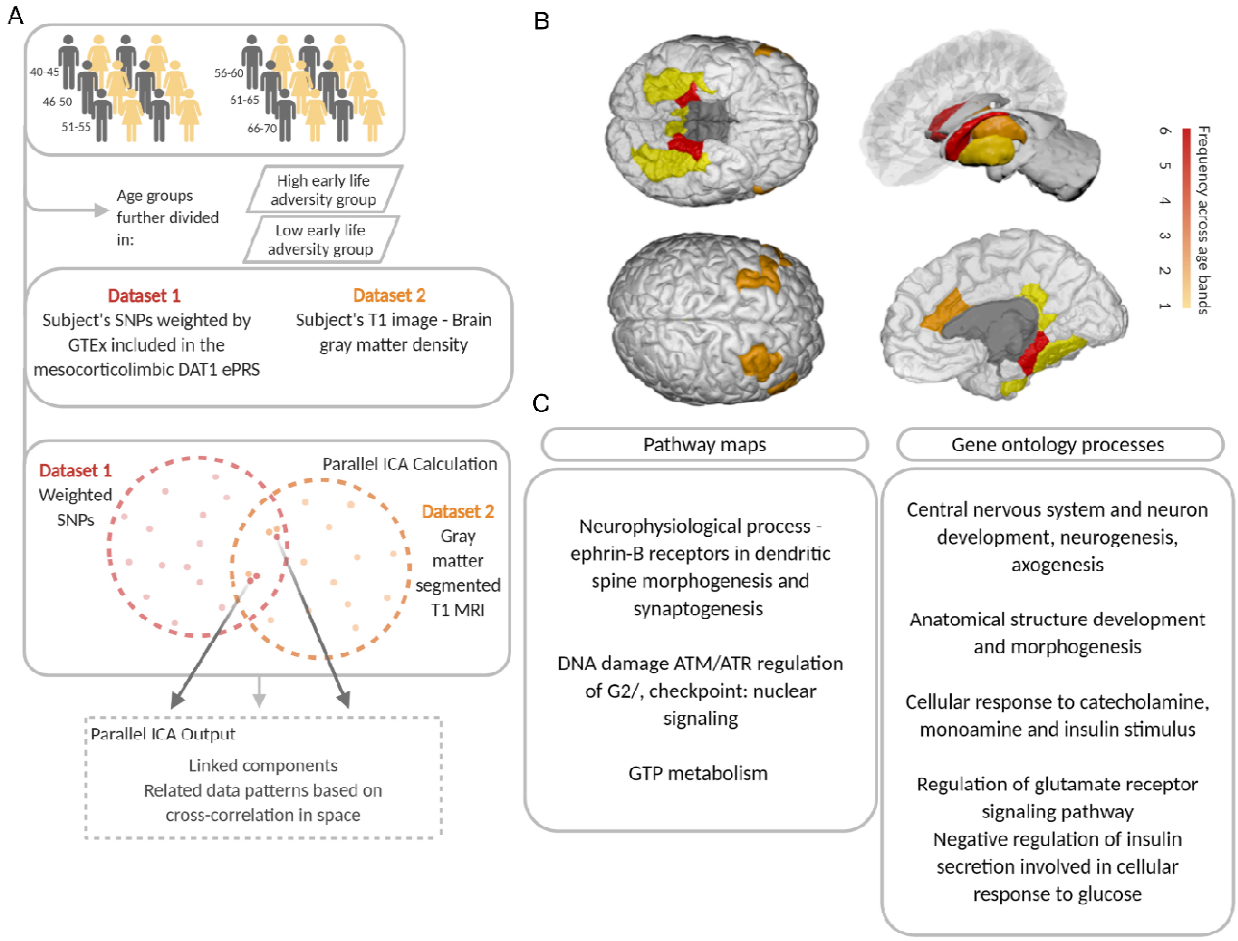

Figure 3. Parallel ICA analysis. A, schematic representation of parallel ICA method. Two different data modalities (SNPs and voxel-based gray matter) were used to establish anatomical-functional correlations between the mesocorticolimbic DAT1 ePRS and brain features from UK Biobank participants (N=9,291). Analyses were carried out in 6 groups of participants separated in blocks of 5 years of age (from 40 to 70 years old). These groups were further separated into high and low early life adversity scores. The analysis estimates the maximum independent components within each data modality separately while also maximizing the association between modalities using an entropy term based on information theory. **B**, significant brain regions associated with SNPs from the mesocorticolimbic DAT1 ePRS across all age bands included insula, inferior and medial frontal gyrus, striatum (pallidum, putamen and caudate),

parahippocampal gyrus and cingulate gyrus. Color scheme represents the frequency that each brain region appeared as significant across the analyzed age groups. **C**, summary of significant pathway maps and gene ontology processes related to SNPs from the mesocorticolimbic DAT1 ePRS associated with gray matter across all age bands.

Tables

Table 1. Variables used to create early life adversity scores. Presence of each component (described in each line) yielded 1 point, and the scores represent the summation of points.

	UK Biobank	**ALSPAC**
Early Life Adversity Indicators	Born small or large for gestational age	Born small or large for gestational age
	Smoking during pregnancy	Smoking during pregnancy
		Hospitalization in the first 6 months of life
		Financial difficulties
	Domestic violence	Domestic violence
	Nobody to take to doctor when needed	Poor family function
		Poor maternal mental health
	Felt hated during childhood	Poor maternal sensitivity
	Not felt loved during childhood	
	Maximum score: 6 Minimum score: 0	Maximum score: 8 Minimum score: 0

Table 2. Description of the baseline characteristics in UK Biobank sample for high and low mesocorticolimbic DAT1 ePRS groups, which were defined by median split of the ePRS score.

UK Biobank (n = 60016)					
Characteristics	Low ePRS (n =30008)		High ePRS (n = 30008)		P-value
	Mean or percentage	SD or N	Mean or percentage	SD or N	
Sex - Male	37.5%	11265	37.4%	11230	0.77
Birth weight (grams)	3347	614	3340	609	0.18
Breastfed as a baby	71.9%	19785	71.8%	19875	0.92
Maternal smoking around birth	27.7%	8317	27.4%	8230	0.43
Completed full-time education at 14-years of age or younger	0.6%	97	0.6%	93	0.75
Age at recruitment (years)	54.51	7.70	54.57	7.69	0.29
Townsend deprivation index at recruitment	-1.71	2.84	-1.75	2.82	0.15
Current or previous smoking	40.4%	12090	39.8%	11934	0.43
BMI at recruitment	26.63	4.63	26.63	4.64	0.99
ICD10 F10-F19 Mental and behavioural disorders due to psychoactive substance use	2.2%	663	2.1%	633	0.41
ICD10 F20-F29 Schizophrenia, schizotypal and delusional disorders	0.1%	32	0.1%	32	1.00
ICD10 F30-F39 Mood [affective] disorders	2.7%	804	2.5%	741	0.11
ICD10 F40-F48 Neurotic, stress-related and somatoform disorders	1.8%	549	1.6%	481	0.03*
ICD10 E11 Non-insulin-dependent diabetes	2.8%	833	2.5%	763	0.08
ICD10 I70 Atherosclerosis	0.1%	34	0.2%	49	0.12
ICD10 I63 Cerebral infarction	0.4%	123	0.4%	128	0.80
ICD10 I20-I25 Ischaemic heart diseases	4.8%	1428	4.8%	1428	1.00

Townsend deprivation index in UK Biobank: A measure of the level of social deprivation that a person lives in. The index was calculated based on previous national census. Participant is given a score reflecting the output area in which their postal code is located. Four key aspects are considered in the index: the percentage of unemployment, overcrowded households, households without a car and non-home ownership.

Table 3. Description of the baseline characteristics of the ALSPAC sample used according to high and low mesocorticolimbic DAT1 ePRS. Groups were divided by the median of the ePRS score.

ALSPAC (n = 910)					
Characteristics	Low ePRS (n =455)		High ePRS (n =455)		P-value
	Mean or percentage	SD or N	Mean or percentage	SD or N	
Gender - Male	48.57%	221	47.47%	216	0.79
Maternal age at birth (years)	30.23	4.16	29.98	4.15	0.35
Full weeks of gestation	39.74	1.28	39.75	1.30	0.87
Smoking during pregnancy	11.86%	54	11.20%	51	0.83
Birth weight (grams)	3523.73	461.82	3507.56	432.08	0.58
Low SES	35.26%	158	36.24%	162	0.98
Breastfeeding for at least 3 months	64.83 %	295	57.80 %	263	0.03*
Waist circumference (cm) at 15.5 years of age	76.42	8.25	76.05	8.16	0.49
SDQ Total difficulties score at 16.5 years of age	5.55	4.34	5.20	4.47	0.22
CIS-R Depression score at 17.5 years of age	0.24	0.69	0.32	0.77	0.12
CIS-R Anxiety score at 17.5 years of age	0.20	0.59	0.27	0.71	0.11
HOMA2-IR at 17.5 years of age	211.49	117.62	212.91	119.73	0.55
zBMI at 15.5 years of age	0.26	0.97	0.26	0.97	0.94

Low socioeconomic status (SES) in ALSPAC: crowding index higher than 0.75 at 2-year-and-9-month time point was considered as low SES. Crowding index was calculated by dividing the number of individuals living in the family dwelling by the number of rooms in the family dwelling and was used as a proxy measure of socioeconomic status.