Interplay between Genomics and the Early Environment influences the risk for Psychiatric and Cardiometabolic comorbidities

Barbara Barth

Integrated Program in Neuroscience

Faculty of Medicine

McGill University, Montreal

December 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Neuroscience.

Table of Contents

Table of Contents **1**

Acknowledgments **5**

Preface to the thesis **8**

Contribution to original knowledge 8

Contribution of authors 10

Abstract **13**

Résumé **15**

List of tables **17**

List of figures **19**

List of abbreviations **21**

Chapter I. Introduction **23**

Chapter II. Comprehensive review of the literature **37**

Introduction 38

The Dopaminergic System 41

The Development of Dopaminergic Pathways 43

Early Life Adversity: Beyond Low Birth Weight 45

Genetic Studies Inform About Biological Functions 46

Diathesis Stress Versus Differential Susceptibility 50

Evidence of Dopamine as the Biological Bases of Early Life Programming 52

Conclusions and Final Overview 58

References 60

Connecting statement to chapter III **74**

Chapter III. Genetically predicted gene expression of prefrontal DRD4 gene and the differential susceptibility to childhood emotional eating in response to positive environment 76

Abstract 78
Introduction 79
Methods 82
Results 93
Discussion 96
References 99

Connecting statement to chapter IV 107

Chapter IV. Genome-wide main effect and environment interaction-associated SNPs are distinct 108

Abstract 109
Introduction 109
Results 110
Discussion 115
Methods 116
References 119

Connecting statement to chapter V 122

Chapter V. Expression-based polygenic scores - A gene network perspective to capture individual differences in biological processes 124

Abstract 126
Introduction 127
Results 129
Discussion 148
Supplementary Information 151
Supplementary Figures 151
Supplementary Methods 157

References 162
Connection statement to chapter VI 171
Chapter. VI. Striatum DAT1 ePGS moderates the effect of early adversity on the risk for adult psychiatric and cardiometabolic comorbidity 174
Abstract 176
Introduction 176
Methods 179
Results 191
Discussion 200
Supplementary Information 207
Supplementary Figures 207
Supplementary Tables 215
Supplementary Methods 224
References 231
Chapter VII. Discussion 244
Chapter VIII. Conclusion 254
Reference list for Non-Manuscript Chapters (Contributions to original knowledge, connecting statements and Chapters I, VII and VIII) 255

I aspire not to be impervious to life's challenges but, instead, to inhabit a world where compassion and understanding render such fortitude irrelevant.

Akin to an orchid child's heightened reactivity and susceptibility to their surroundings, I seek resilience that is attuned and responsive, rather than impenetrable.

I perceive resilience not as a shield against difficulties, but as a guiding compass directing me toward decisions aligned with my values. It stands as a force, ensuring my integrity remains unwavering in challenging times—an anchor steadfast amid the storms of uncertainties, empowering me to navigate with purpose in a world full of traps, where some are woven from the threads of vanity, and others from the fabric of inequality.

Some may say I'm too tender for this world; I say this world is too empty for my purposes.

Acknowledgments

To my Ph.D. supervisor **Dr. Patricia Pelufo Silveira**, thank you for your support over the years, for your patience in teaching and guiding me through a new field, in which I had no previous experience on, and for seeing potential in me. I admire your enthusiasm for science. It was a joy to participate in many important moments from the Silveira Lab. I will cherish these experiences and carry the lessons learned through out my career. My committee members, **Dr. Cecília Flores** and **Dr. Marco Leyton**, thank you for enthusiastically pushing my work to its best version over our committee meetings. **Dr. Micheal Meaney** and **Josie Diorio**, thank you for your fundamental support and many opportunities offered over these years. The interactions we had made me feel seen and appreciated. **Irina Pokhvisneva,** thank you for being patient with the many requests and doubts over the many projects we worked on together. I admire your work ethic and your diligent way of conducting research. Thank you **Dr. André Krumel Portella** for the advice and availability to discuss and improve my work. **Sachin Patel, Zihan Wang, Guillaume Elgbeili,** and **Nick O'Toole**, thank you for your amazing technical support, it was a pleasure to count with your support, but most importantly with your kindness! **Xianglan Wen, Tie Yuan Zhang** and **Anne-Marie Arel-Dubeau**, thank you for your kindness in teaching so many things I didn't know. It was a pleasure learning from you. Thank you for the experienced advice, support, and discussions **Dr. Bonnie Alberry**, **Dr. Patricia Maidana Miguel**, **Dr. Eamon Fitzgerald**, **Dr. Marcio Bonesso Alves** and **Dr. João Paulo Maires Hoppe**. Thank you, **Dr. Roberta Dalle Molle**, for your support. It was a joy to work with you after following your work for many years before. Thank you, **Kelly Craig**, for sharing with the lab your excellent humor and for eagerly supporting our work. To my fellow student lab mates who made the student experience easier

by sharing perspectives, good and bad times, **Lawrence Chan**, **Ranjani Nadarajan**, **Ameyalli Gómez-Ilescas**. To my lab mate and friend **Aashita Batra**, thank you for your friendship, for being so generous with me, sharing your points of view and giving me support. Thank you to my former supervisor, **Dr. Lisiane Bizarro**, for providing endless recommendation letters, every time with the same enthusiasm. It was a joy being able to work with you again during your visiting professor time at McGill. Your enthusiasm for science is always inspiring. To **Dr. Carla Dalmaz**, thank you for generously sharing your knowledge, providing support, and offering guidance. Your astute approach to exploring problems or ideas is truly inspirational. I am proud to have learned from and worked with you.

I consider myself fortunate to have shared the Ph.D. journey with two individuals who were not only my lab mates but, most importantly, my friends and my family away from home. **Dr. Danusa Mar Arcego** and **Dr. Randriely Merscher Sobreira de Lima**, thank you for teaching me so much, for eagerly engaging in countless discussions about my work, and celebrating with me every accomplishment along this challenging journey. Your support was fundamental, and I will be forever grateful. Most importantly, thank you for the undoubtable source of support through hard times and for reminding me of my value.

Special thanks to my husband, **Dr. Euclides José de Mendonça Filho**, who played a dual role as both a lab mate and my number one supporter. To my lab mate, I am grateful for sharing your knowledge, for being readily available with a kind smile for every question I had, and for the many hours devoted to discussing my work. To my husband, I wouldn't have been able to accomplish this without you. I'm so proud of the team we are, no challenge is big enough when we tackle it as a team. Thank you for believing in me during times when I couldn't, for

recognizing strength and value in me when I felt most vulnerable. I admire and love you enormously. Special mention and thanks to my feline companions, **Lisbela** and **Neno**, your existence brings pure joy to my life. Your unshakable confidence and self-esteem are truly inspirational for a mere human like me. Your instincts for showing me love when I needed it most are something to be studied. I'm grateful for being your human tutor (servant).

To my family, who understood all my absence during these five years. To my aunts and cousin, **Ivete**, **Nivene**, **Eliane** and **Bruna** and my grandmother, **Maria Cordélia**, thank you for your unshakable believe and love for me. To **my family from Salvador**, having you in my life brings me joy and a sense of belongingness, thank you for sharing Keno with me. To my brother, **Paulinho**, for being my most discrete yet fierce admirer. To my father, **Paulo**, for being my first science teacher. You instigated in me the need to search for answers to my questions, and I always had I ton; thank you for your patience. Thank you for believing in my academic career aspirations since the very beginning. To my mother, **Eloisa**, for being my most important critical thinker teacher. You taught me to doubt everything and think for myself, this is my most valuable quality as a scientist. Thank you for all your support over this Ph.D. journey, your visits lifted my spirit when I needed it the most. Amo vocês com todo meu coração!

Preface to the thesis

Contribution to original knowledge

The interest in understanding the neurobiological mechanisms underlying the effects of early life adversity has grown in recent years, yet these mechanisms remain to be fully defined. This thesis contributes to original knowledge by exploring the role of individual variation in gene expression related to brain dopaminergic neurotransmission in linking ELA exposure to the development of psychiatric and cardiometabolic disorders. The results on chapter VI are the first to show that a novel expression-based polygenic score, based on the striatal dopamine transporter gene co-expression network moderates the association between birth weight and psychiatric and cardiometabolic disorders in both adults and adolescents. These results can inform future molecular studies aiming to causally relate dopaminergic gene expression variations and the development of chronic disorders in the context of ELA exposure.

In chapter III, we utilized a predicted prefrontal DRD4 gene expression measure calculated using the PrediXcan technique[1]. This enabled us to investigate the relationship between individual variations in whole genome tissue-specific predicted expression of DRD4 and eating behavior, representing a significant innovation and advancement compared to previous studies that instead focused on specific polymorphisms within this gene and their relationship to eating behavior[2]. The use of this novel and more comprehensive functional genomics measure has already inspired two other studies addressing the connection between predicted DRD4 expression and eating behaviour[3,4]. To the best of our knowledge, this thesis work is the first to show that a positive environment measure interacts with prefrontal DRD4-predicted gene expression variation to influence emotional eating in children. We demonstrated that a high

prefrontal DRD4-predicted gene expression impacts emotional eating phenotype differently depending on the early life environment exposure, highlighting the protective factor of a positive early life environment in individuals susceptible to the impacts of environment exposure.

In Chapter IV we demonstrated for the first time that variants detected in GWAS and GWEIS results do not overlap when calculated in the same population and using the same outcomes. We contributed to the field of GxE studies by concluding that PRS derived from GWAS have a limited application to study GxE effects.

In Chapter V, we demonstrated that a genetic score integrating tissue-specific information, emphasizing the concept of the functional genetic unit as a network rather than isolated genes, and incorporating functional genomics through gene expression while maintaining a genome-wide perspective, offers advantages in representing biological information over traditional polygenic risk scores. We also demonstrated that the ePGS have greater trans-ancestry portability, which can facilitate the use of this type of genetic scores across diverse populations.

Contribution of authors

The present thesis was written by Barbara Barth with editing from Dr. Patricia Silveira. Details on manuscript chapters can be found below.

Chapter I

Barbara Barth wrote the introduction under the supervision of Dr. Patricia Silveira.

Chapter II

Comprehensive review of the literature published as a book chapter entitled " The Interplay Between Dopamine and Environment as the Biological Basis for the Early Origins of Mental Health". The book chapter was written by Barbara Barth with supervision and editing from Dr. Patricia Silveira and Dr. André Krumel Portella and editing by Dr. Michael Meaney and Dr. Laurette Dubé.

Chapter III

The published manuscript entitled "Genetically predicted gene expression of prefrontal DRD4 gene and the differential susceptibility to childhood emotional eating in response to supportive environment" was written by Barbara Barth with editing from Dr. Patricia Silveira, Dr. Lisiane Bizarro, Dr. Patricia Maidana Miguel, Dr. Kieran O'Donnel and Dr. Michael Meaney. Barbara Barth and Dr. Patricia Silveira conceived and designed the experiments. Statistical analyses were supervised by Dr. Patricia Silveira and conducted by Barbara Barth. The entire project was planned and supported with Dr. Patricia Silveira, Dr. Laurette Dubé, Dr. Robert Levitan and Dr. Michael Meaney.

Chapter IV

The submitted manuscript entitled “Genome-wide main effect and environment interaction-associated SNPs are distinct” was written by Barbara Barth with editing from Dr. Patricia Silveira, Irina Pokhvisneva, Dr. Patricia Maidana Miguel, Dr. André Krumel Portella and Dr. Michael Meaney. Barbara Barth, Dr. Patricia Silveira, Irina Pokhvisneva, Dr. Patricia Maidana Miguel and Dr. André Krumel Portella conceived and designed the experiments. Statistical analyses were supervised by Dr. Patricia Silveira and conducted by Irina Pokhvisneva and Barbara Barth. The entire project was planned and supported by Dr. Michael Meaney and Dr. Patricia Silveira.

Chapter V

The manuscript to be submitted entitled “Expression-based polygenic scores - A gene network perspective to capture individual differences in biological processes” was written by Barbara Barth with editing from Irina Pokhvisneva, Dr. Michael Meaney and Dr. Patricia Silveira. Barbara Barth and Patricia Silveira conceived and designed the experiments. Experiments and statistical analysis were supervised by Dr. Patricia Silveira and Irina Pokhvisneva and conducted by Barbara Barth, Euclides José de Mendonça Filho, and Danusa Mar Arcego. The entire project was planned and supported by Dr. Michael Meaney and Dr. Patricia Silveira.

Chapter VI

The manuscript to be submitted entitled “Striatal dopamine gene network moderates the effect of early adversity on the risk for adult psychiatric and cardiometabolic comorbidity” was written by Barbara Barth with editing from Dr. André Krumel Portella, Irina Pokhvisneva, Dr. Michael Meaney and Dr. Patricia Silveira. Human cohorts’ data managing was conducted by Irina

Pokhvisneva and Carine Parent. Barbara Barth and Patricia Silveira conceived and designed the experiments. Experiments and statistical analysis were supervised by Dr. Patricia Silveira and Irina Pokhvisneva and conducted by Barbara Barth, Danusa Mar Arcego, Euclides José de Mendonça Filho, Randriely Merscher Sobreira de Lima, and Carla Dalmaz. The entire project was planned and supported by Dr. Michael Meaney and Dr. Patricia Silveira.

Chapter VII

Barbara Barth wrote the discussion under the supervision of Dr. Patricia Silveira.

Chapter VIII

Barbara Barth wrote the conclusion under the supervision of Dr. Patricia Silveira.

Abstract

Cardiometabolic and psychiatric disorders frequently occur in the same patient, and an important shared risk factor among them is exposure to early life adversity (ELA). Nevertheless, the specific biological mechanisms connecting ELA to the co-occurrence of cardiometabolic and psychiatric conditions are not yet understood. In fact, not all individuals exposed to ELA will develop deleterious consequences of this exposure, suggesting an individual variability in the susceptibility to ELA effects. Dopamine (DA) neurotransmission is responsive to ELA and plays a role in shaping the development of these two categories of disorders. Here we demonstrated that individual genetic variations associated with dopaminergic genes modulate the risk for psychiatric and cardiometabolic comorbidities and altered eating behaviour in adults, adolescents, and children. We began by showing that individual variation in the genetically predicted expression of the dopamine receptor D4 (DRD4) gene in the prefrontal cortex moderates the effect of early environment on child emotional eating. We then discussed ways of capturing gene by environmental (GxE) effects in humans while also integrating functional genomics and a notion of complex systems in biology to elucidate potential shared biological mechanisms underlying psychiatric and cardiometabolic comorbidities. We showed that polygenic risk scores (PRS) derived from Genome-wide association studies (GWAS) were not suited to capture GxE effects, by demonstrating that identified variants from GWAS do not significantly overlap with variants from Genome-wide by environment interaction studies (GWEIS) in the same outcomes and population. With the aim to capture biological information through a complex system approach in biology while also maintaining a genome-wide perspective, we adopted the use of expression-based polygenic risk scores (ePGS). We then

compared the ePGS performance against the traditional PRS through several experiments and ancestry-specific analysis. We were able to demonstrate ePGS' advantage in representing biological information and portability across ancestries. Lastly, we applied the ePGS technique and demonstrated the role of the striatal dopamine transporter (SLC6A3) base gene network in moderating the impact of ELA on psychiatric and cardiometabolic disorders in adults and adolescents. Enrichment analysis from this study points to possible molecular mechanisms associated with insulin signaling disturbances. This work may contribute to increased awareness of the differential impact of ELA in individuals, the role of GxE in the development of chronic disorders, the role of dopamine-related brain gene expression in modulating the effects of ELA exposure, and the need to include functional genomics into genetic prediction scores.

Résumé

Les troubles cardiométaboliques et psychiatriques surviennent fréquemment chez le même patient et un facteur de risque commun important est l'exposition à l'adversité au début de la vie (ELA). Néanmoins, les mécanismes biologiques spécifiques reliant l'ELA à la cooccurrence des troubles cardiométaboliques et psychiatriques ne sont pas encore compris. En fait, tous les individus exposés à l'ELA ne développeront pas nécessairement des conséquences néfastes de cette exposition, suggérant une variabilité individuelle dans la susceptibilité aux effets de l'ELA. La neurotransmission de la dopamine (DA) répond à l'ELA et joue un rôle dans le développement de ces deux catégories de troubles. Ici, nous avons démontré que les variations génétiques individuelles associées aux gènes dopaminergiques modulent le risque de comorbidités psychiatriques et cardiométaboliques et de comportements alimentaires altérés chez les adultes, les adolescents et les enfants. Nous avons commencé par montrer que la variation individuelle dans l'expression génétiquement prédite du gène du récepteur de la dopamine de type 4 (DRD4) dans le cortex préfrontal modère l'effet de l'environnement précoce sur l'alimentation émotionnelle de l'enfant. Nous avons ensuite discuté des moyens de capturer les effets gène-environnement (GxE) chez les humains tout en intégrant la génomique fonctionnelle et une notion de système complexe en biologie pour élucider les éventuels mécanismes biologiques partagés sous-jacents aux comorbidités psychiatriques et cardiométaboliques. Nous avons montré que les scores de risque polygénique (PRS) dérivés d'études d'association pangénomique (GWAS) ne convenaient pas pour capturer les effets GxE en démontrant que les variants identifiés dans les GWAS ne chevauchent pas significativement les variants des études d'interaction Génome-Environnement (GWEIS) dans les mêmes

phénotypes et la même population. Dans le but de capturer des informations biologiques grâce à une approche de système complexe en biologie tout en maintenant une perspective pangénomique, nous avons adopté l'utilisation de scores de risque polygénique basés sur l'expression (ePGS). Nous avons ensuite comparé les performances de l'ePGS par rapport au PRS traditionnel à travers plusieurs expériences et des analyses spécifiques à l'ascendance. Nous avons pu démontrer l'avantage de l'ePGS dans la représentation d'informations biologiques et la portabilité entre les ascendants. Enfin, nous avons appliqué la technique de l'ePGS et démontré le rôle du réseau génique de base du transporteur de la dopamine (SLC6A3) dans la modulation de l'impact de l'ELA sur les troubles psychiatriques et cardiométaboliques chez les adultes et les adolescents. L'analyse d'enrichissement génique de cette étude pointe vers des mécanismes moléculaires possibles associés à des perturbations de la signalisation de l'insuline. Ce travail peut contribuer à une meilleure prise de conscience de l'impact différencié de l'ELA chez les individus, du rôle du GxE dans le développement de troubles chroniques, du rôle de l'expression génique cérébrale liée à la dopamine dans la modulation des effets de l'exposition à l'ELA et de la nécessité d'inclure la génomique fonctionnelle dans les scores de prédiction génétique.

List of tables

Table 1. Sample description and differences between High and Low DRD4 predicted gene expression groups in MAVAN. 90

Table 2. Sample description and differences between High and Low DRD4 predicted gene expression groups in GUSTO. 91

Table 3. SNPs and genes included in ePRS and PRS scores..... 132

Table 4. Description of the baseline characteristics in UK Biobank sample for high and low striatum *SLC6A3* ePGS groups, which were defined by median split of the ePGS score..... 181

Table 5. Description of the baseline characteristics of the ALSPAC sample used according to high and low striatum *SLC6A3* ePGS. Groups were divided by the median of the ePGS score..... 182

Supplementary Table 1. Genes co-expressed with the SLC6A3 in mice striatum data, used to calculate striatum *SLC6A3* ePGS genetic score. 215

Supplementary Table 2. Means of variables used in the cluster analysis for the ALSPAC cohort..... 219

Supplementary Table 3. Main effect of PRS on psychiatric and cardio-metabolic comorbidity..... 220

Supplementary Table 4. Interaction effect between PRS and birth weight on psychiatric and cardio-metabolic comorbidity. 220

Supplementary Table 5. Characteristics comparison between UK Biobank low vs randomly-selected non-low birth weight groups (n=1906) used in the parallel ICA analysis..... 221

Supplementary Table 6. Characteristics comparison between randomly-selected non-low birth weight group used in the parallel ICA analysis vs full sample of non-low birth weight individuals with MRI data available in the UK Biobank (n=10214) 222

Supplementary Table 7. Subset of significant SNPs related to gray matter density variations according to pICA analysis. 223

List of figures

Figure 1. Theoretical framework on the interplay between environmental adversity, altered metabolism and DA altered function across development. 43

Figure 2. Variables and cut-offs used to create the Positive postnatal environmental score in MAVAN 84

Figure 3. Scheme for the generation of predicted DRD4 gene expression on discovery and replication cohorts 86

Figure 4. Variables and cut-offs used to create the positive postnatal environmental score in GUSTO 88

Figure 5. Evidence of differential susceptibility - Interaction between positive postnatal environmental score and predicted DRD4 gene expression on Emotional over-eating at 48 months of age. MAVAN Cohort 94

Figure 6. Interaction between positive postnatal environmental score and predicted DRD4 gene expression on Desire to Drink at 60 months of age. Gusto Cohort 95

Figure 7. Investigation of the overlap between significant SNPs identified in the GWAS for different chronic diseases and those identified in the GWEIS (SNP-by-early adversity model) for the same conditions 112

Figure 8. GWAS may not capture SNPs for which the effect is modified by the environment 115

Figure 9. Schematic figure representing the key steps to calculate the ePGS 129

Figure 10. Network visualization comparison of SLC6A3 derived ePGS and comparable size PRSes 134

Figure 11. Unique elements for 'neuron differentiation', a common gene ontology enrichment analysis term for both ePGS and PRS 138

Figure 12. Correlation matrix of gene expression for ePGS gene networks and PRS gene networks based on BrainSpan human post-mortem brain tissue (from embryonic to adulthood, N=42) 140

Figure 13. Manhattan plot for Howard (2019) broad depression GWAS results and SLC6A3 PFC ePGS SNPs 143

Figure 14. Venn diagrams of gene overlap for ePGSes and PRSes calculated based on the ePGS and PRS in the 1000 Genomes Project dataset 146

Supplementary Figure 1. Schematic overview of RNA sequencing process: from tissue collection to data analysis. 152

Supplementary Figure 2. Network visualization comparison of DRD2 derived ePGS and comparable size PRSes 154

Supplementary Figure 3. Visualization for the SLC6A3 Striatum ePGS gene network. . 155

Supplementary Figure 4. ePGS and PRS score distribution across 5 different ancestries 156

Figure 19. Construction and characterization of the striatum SLC6A3 gene network. .. 194

Figure 20. Striatum SLC6A3 ePGS moderates the effect of early life adversity on the risk for mental health disorder and cardiometabolic comorbidity 197

Figure 21. Parallel ICA analysis 200

Supplementary Figure 5. UK Biobank Sample size block scheme for each inclusion and exclusion criteria applied 208

Supplementary Figure 6. ALSPAC Sample size block scheme for each inclusion and exclusion criteria applied 210

Supplementary Figure 7. Comorbidity risk clusters – ALSPAC 213

Supplementary Figure 8. Striatal SLC6A3 gene network enrichment analysis 214

List of abbreviations

Abbreviation	full form
AADC	Aromatic amino acid decarboxylase
ACE	Adverse Childhood Experiences
ADHD	Attention deficit hyperactivity disorder
BDNF	Brain-derived neurotrophic factor
BMI	Body mass index
CAD	Coronary artery disease
CEBQ	Child Eating Behavior Questionnaire
CNS	Central nervous system
DA	Dopamine; 3-hydroxytyramine
DAT	Dopamine transporter
DOHaD	Developmental origins of health and disease
DRD4	Dopamine receptor D4
DLPFC	Dorsolateral Prefrontal Cortex
DNA	Deoxyribonucleic acid
ELA	Early life adversity
EPDS	Edinburgh Postnatal Depression Scale
ePGS or ePRS	Expression-based polygenic risk score
FAD	Family Assessment Device
GCs	Glucocorticoids
GUSTO	Growing Up in Singapore Towards Healthy Outcomes
GWAS	Genome-wide association study
GWEIS	Genome-wide by environment interaction studies
GABA	Gamma-aminobutyric acid
GxE	Gene by environment interaction
GTEx	Genotype-Tissue Expression

HPA	Hypothalamic-Pituitary-Adrenal
HBT	Human Brain Transcriptome
IUGR	Intrauterine growth restriction
LBW	Low birth weight
L-DOPA	L-3,4-dihydroxyphenylalanine
MAVAN	Maternal Adversity, Vulnerability and Neurodevelopment
NAc	Nucleus accumbens
NGFIA	Nerve growth factor-induced clone A
OFC	Orbitofrontal cortex
PFC	Prefrontal cortex
PRS	Polygenic risk score
PSRP	Preschool Separation – Reunion Procedure
SES	Socioeconomic status
SGA	Born small for gestational age
SLC6A4	Solute carrier family 6 member 4
SNP	Single nucleotide polymorphisms
STAI	State-Trait Anxiety Inventory
pTH	Phospho-tyrosine hydroxylase
TH	Tyrosine hydroxylase
VMAT	Vesicular monoamine transporter
VTA	Ventral tegmental area
7R	7-repeat allele

Chapter I. Introduction

In recent years, the significance of the early life environment in shaping human development has gained growing recognition. However, the historical trail of evidence underscoring this crucial connection dates back much further, laying a foundation for our understanding of its paramount importance. In early 1930 researchers concluded that life expectancy from a British and Swedish sample was apparently determined by early life conditions[5]. Later, in 1976 results from the tragic Dutch Famine that happened during the second world war, revealed an association between early in-utero deprivation of food and later obesity during adulthood[6]. This association was further explored in 1989, when researchers demonstrated the positive relationship between low birth weight and death by ischaemic heart disease[7]. In 1998, The Adverse Childhood Experiences (ACE) study showed the positive relationship between exposure to adverse experiences during childhood and presence of health risk factors in adulthood[8]. Such evidence and many others culminated in the concept of developmental origins of health and disease[9], adopted by many researchers that are dedicated to elucidate the associations and mechanisms linking the early life environment to later health outcomes. Today, it is known that the exposure to early life adversity (ELA) can have profound implications in the human development[8,10-18]. Variations in children's health outcomes are influenced by the developmental context and the exposure to ELA is involved in health disparities in childhood[19,20]. Some studies suggest a dose-dependent effect between the number of exposures to adverse childhood experiences and adverse health outcomes[21,22] while others emphasize the increased burden of exposure to multiple forms of ELA[10,23]. Considering the growing evidence supporting this relationship, increased attention is being given to the

importance of the biology of adversity and resilience and its possible utility in improving pediatric clinical practice[24-26]. In fact, the timing of adversity exposure plays a pivotal part in determining adverse related outcomes[11,27]. The first years of life are marked by critical periods of brain development[28,29], thus granting vulnerability to the effects of ELA exposure[11,27]. Known forms of ELA are related to threat and deprivation in both physical and psychosocial domains[30-32] (e.g. malnutrition, chronic infections, physical and emotional abuse and neglect), and confer an increase risk for negative outcomes, although the degree of risk associated with particular forms of ELA may vary to some extent[33-37].

ELA affects the development of different systems by causing biological disruptions in the stress response [38-41], the developing brain[42], cardiovascular[43,44], immune[45,46 47] and metabolic[48,49] systems. The Hypothalamic-Pituitary-Adrenal (HPA)-axis controls the hormonal and subsequential behavioral and physiological responses to stress[50]. In response to stress exposure, corticotropin-releasing hormone is released by specific hypothalamic neurons, stimulating the secretion of adrenocorticotropic hormone from the pituitary, which stimulates the adrenal cortex production of glucocorticoids (GCs)[50,51]. ELA exposure is associated with hyper-responsive HPA-axis which leads to greater GCs production, both at baseline and in response to subsequent stress[38,39]. Evidence points to a causal relationship between ELA exposure, DNA methylation and GC receptor expression[40]. Studies using animal models, evidenced the positive relationship between poor maternal care and increased release of adrenocorticotropic hormone in response to stress by failure in the negative feedback of GCs by the HPA-axis[52-54]. This impaired negative feedback was shown to be related to a diminished quantity of GC receptors in the hippocampus, which in turn was causally related to methylation changes that modified the

access of GC receptor transcription factor NGFIA to the promoter region of the GC gene, thus modulating the gene expression of the GC receptor[40,55].

The over production of GC, by a hyper-responsive HPA-axis, affects the development and function of cells in the body, including those of the immune system[38,39]. GCs are important suppressors of chronic inflammation[56] and their over production can lead to resistance, damaging the anti-inflammatory action and compromising the immune system[45]. In fact, inflammatory markers are found to be elevated in individuals exposed to ELA and who developed psychiatric disorders[57-62], suggesting greater immune activation associated with ELA exposure. The interplay between stress and immune signaling pathways may be a possible candidate linking ELA exposure to endothelial damage and alterations in the cardiovascular system[43], although further studies are needed [44]. Activation of GC receptors can inhibit the signaling of an immune messenger (e.g. cytokine interleukin-4) that is important for cardiomyocytes development, which are specialized muscle cells found in the heart[43].

Overproduction of GCs also affects the central nervous system (CNS)[38,39], since they cross the blood-brain barrier and influence brain function via binding to GC and mineralocorticoid receptors [50,63]. ELA exposure alters brain neuronal plasticity, as exposure to stress is a negative regulator of neurogenesis. This happens especially in the hippocampus, a region that has a large expression of GC receptors and is involved in adult neurogenesis and in the pathophysiology of mood disorders[64-70]. GCs activity in the hippocampus is considered a way by which stress exposure disrupts neurogenesis and subsequently brain neuronal plasticity[66,68,71]. The brain-derived neurotrophic factor (BDNF) is an essential player in neuronal survival and growth and

evidence points to the interplay between GCs and BDNF as a candidate mechanism affecting neuronal plasticity[72].

There is also a relationship between GCs and alterations in neuroendocrine appetite signaling with differential effects of acute and chronic stress on eating behavior[73-77]. In fact, ELA exposure appears to impact ghrelin and leptin signaling[78-80] as well as the hypothalamic neuropeptide Y and agouti-related peptide expression[81], important appetite and energy balance regulators, with consequent impact on metabolic function. Another important hormone linked to glucose metabolism and regulation of appetite is insulin[82,83]. Insulin resistance, a condition that impairs the proper cellular response to insulin and the regulation of glucose levels in the body, is associated with elevated GC levels via their interference with various points of the insulin signaling cascade[74,84], explaining the relationship between high GCs and glucose intolerance [85,86]. A pro-inflammatory state, which is associated with ELA exposure and higher GCs[57-62], is also associated with insulin resistance[87,88]. This evidence demonstrates the connection between ELA exposure, HPA-axis alterations and impaired insulin signaling.

In fact, ELA exposure is associated with elevated risk for type 2 diabetes in adulthood, a disorder defined by insulin resistance[89-91]. Being born low birth weight (LBW) is a prevalent type of early life adversity associated with adverse health outcomes[92-96], including insulin resistance and later in life type 2 diabetes[97,98]. The metabolic challenges imposed by intrauterine growth restriction (IUGR) are the basis for an augmented insulin sensitivity seen in LBW and a possible mechanism is due to the diminished number of pancreatic β-cells in individuals born with LBW[99,100-102], which reduces insulin production, leading to higher cellular sensitivity of insulin receptors[103,104]. A hallmark of LBW is the subsequent catch-up growth, a compensatory mechanism for LBW[105-

[107], an insulin-dependent process that boosts the uptake of glucose by muscles and fat cells, promoting rapid growth[103,104,108,109]. As these events happen in a critical period, "metabolic programming" happens and propitiates the continuation of this process through adulthood, contributing to the development of later in life type 2 diabetes[97,110].

Altered insulin signaling followed by ELA exposure also impacts the CNS, as insulin receptor distribution in the brain overlaps with neurotransmitter systems implicated in neuronal communication[111-113]. In fact, insulin signaling modulates brain functions related to both metabolic and cognitive outcomes[111,114] and some studies suggest that insulin sensitivity of the CNS may underly the pathology of metabolic and cognitive dysfunctions[114,115].

These alterations, which affect both physical and mental health, are the basis for the long-term effects of early life adversity exposure[116-118]. For example, early life adversity is associated with increased risk for both psychiatric[12-16,22,31,33,34,36,37,119] and cardiometabolic disorders[93 94,120-126], which are often comorbid[127,128]. These categories of disorders are both prominent contributors to the worldwide burden of disability-adjusted life years globally[129,130] and important contributors for inflated health care utilization[131,132]. Although ELA is associated with increased risk for both mental and metabolic disease at the population level, not all individuals exposed to ELA will develop these diseases. Measures of early life adversity alone have poor accuracy in predicting adult disease at the individual level[133-135], indicating the existence of individual variability in the susceptibility to ELA exposure effects. Evidence suggests that this variability is in part attributed to the genetic background[136,137].

Genetics can directly contribute to the likelihood of developing a phenotype as in the case of monogenic diseases caused by a single rare mutation[138,139] and shown by studies comparing the differential genetic disease risk in monozygotic and dizygotic twins[138-141], which are relatively rare. The genetic etiology of most diseases is polygenic, or the additive result of many genetic variants (single nucleotide polymorphisms, SNPs)[142].

Genome-Wide Association Studies (GWAS) are designed to identify the isolated individual influence of many SNPs on a specific phenotype[143] by using a case-control study design and stringent corrections for multiple comparisons at the genome-wide level[137,142,144,145]. While these findings have enhanced our comprehension of the genetic basis of numerous diseases, it is worth noting that, even when considered in an additive manner, they account for only a small portion of the variability in complex traits[146-148]. This is because complex traits such as chronic disorders are influenced by heritability, environmental exposures, and also the interaction between them[149-153] (GxE).

GxE interaction refers to the joint effects of genetics and environment in influencing an individual's phenotype. These factors are operationally considered to interact if the effect of one depends on the level of the other factor[151]. The genetic risk for a certain phenotype, in this perspective, might be modified (attenuated or increased) as a result from different environmental exposures. Inversely, genetic factors can also modulate or modify the impact of environmental factors on a particular trait, meaning that the impact of environmental exposures can be different according to a person's genotype[148,154].

This view resonates with an evolutionary-inspired biological reasoning, in which researchers propose that some individuals are genetically more responsive to environmental variations than others. The genetic plasticity to environmental variations is set as a bet-hedging against an uncertain future to avoid a costly mismatch between the individual's ability to face the environmental conditions and the actual challenges that the environment could impose[155-157]. Genetic plasticity maps onto phenotypic flexibility, enabling the individual to adjust and mature in accordance with cues from the present environmental conditions[54,158]. ELA exposure interacts with the genetic background and defines the risk for later disease, although mechanisms are still unknown. Studies addressing the genetic moderation of environmental exposure effects on phenotypes such as depression[159,160], executive function[161-163], eating behaviour[1,164] and obesity[2,165,166] corroborate this view.

The objective of this thesis is to explore the GxE interaction effect involved in the susceptibility to develop psychiatric and cardiometabolic comorbidities. We aim at elucidating possible biological mechanisms underlying these comorbidities and the effect of early life environment-genetics interplay.

Neurotransmitter systems involved in environmental responsivity may therefore serve as the neurobiological basis to understand the differential effect of ELA exposure on the development of psychiatric and cardiometabolic disorders. Evidence suggests that specific brain regions are particularly affected by ELA exposure. Parts of the limbic system (hypothalamus, amygdala and hippocampus) and their connections to the prefrontal cortex (PFC), striatal circuits including the nucleus accumbens (NAc), as well as the ventral tegmental area (VTA), have all been implicated in ELA effects[18,167-174].

The serotonin action on the brain has been linked to the differential effects of ELA exposure. Special attention was given to the serotonin transporter gene (SLC6A4) and its polymorphisms[175-182], although not supported by replication studies or metanalyis[183,184]. The serotonergic system, originating in the raphe nuclei, projects to the brain regions mentioned earlier and is implicated in several behavioral, neuropsychological, and physiological processes, including mood, perception, reward, cardiovascular function, digestion, and energy balance[185,186]. Serotonin function in early life regulates developmental processes linked to neural circuit formation[187-190]. During brain formation, thalamocortical axons are essential to convey sensory information to cortical areas[191,192] and modifying serotonin availability to the embryonic brain disrupts the proper formation of sensory maps by thalamocortical axons[193], possibly by interfering with the netrin-1 axonal guidance cue[194]. Serotonin dysregulation during early life, induced by pharmacological blocking of the serotonin transporter, is related to the development of depression and anxiety-like behaviours[195-197].

Gamma-Aminobutyric acid (GABA) dysregulation has been associated with exposure to ELA, suggesting a possible mechanism by which ELA negatively impacts behaviour and brain development[198-200]. GABA is the main inhibitory neurotransmitter of the human CNS, being important for the typical development of the CNS and implicated in various psychiatric disorders[201,202]. GABAergic projections are present in regions involved in mood regulation and reward and that are implicated in ELA exposure effects, such as NAc, PFC and VTA[203]. ELA exposure also impacts the main excitatory neurotransmitter of the CNS, glutamate, by altering its cerebral content and neurotransmission[204-212] in areas such as the hippocampus and the prefrontal cortex[208-210,212]. A study found that genes related to glutamate receptor activity

moderate the impact of ELA exposure on fetal brain development[213]. Alterations in the glutamatergic system can impact synaptic plasticity, learning, and memory and are implicated in the neuropathology of schizophrenia, autism, and depression[214-217].

ELA can disrupt the development of endocannabinoid system leading to deficits in its function, especially in the hippocampus[218-220]. The endocannabinoid system has receptors localized in the hippocampus, a region heavily impacted by ELA exposure[221], and has a role in regulating stress responsivity and mood[222-224]. ELA can also affect behavioral processes associated to opioid system functioning, such as augmented vulnerability to develop opioid use disorder[225-227] through alterations in the opioid system function[228-231]. Opioids act through three receptors, mu delta and kappa, being related to modulation of pain perception and reward among other functions[232-234].

The dopaminergic system is equally impacted by ELA exposure[235-241] and is considered especially sensitive to its effects[158]. ELA exposure can affect the amount of midbrain dopaminergic neurons[237,240], the amount of dopamine (DA) receptors in the nucleus accumbens[236,241], orbitofrontal and medial prefrontal cortex[239], the expression of dopaminergic genes in the midbrain[235] and alter DA transmission in response to reward[239,242] among other alterations[243,244].

Involved in the synthesis and secretion of DA, the dopaminergic system originates in the midbrain VTA and projects to subcortical and cortical regions[245]. In the mesolimbic pathway dopaminergic neurons project from the VTA to the ventral striatum, including the NAc and in the mesocortical pathway they project to the PFC. VTA, NAc and PFC are implicated in the ELA

exposure effects as mentioned earlier. Together these pathways are called the mesocorticolimbic dopaminergic pathway which is involved in reward and motivation[246,247].

ELA exposure is known to affect value attribution and predicted reward errors[248-251]. For example, individuals exposed to ELA showed altered processing of reward and loss in adulthood in a reward-processing task[248] and deficits in functions attributed to the ventral striatum such as reward responsiveness and approach motivation[249]. ELA exposure also affects brain regions implicated in dopaminergic pathways, such as altered connectivity between the ventral striatum and PFC[252], between VTA and hippocampus[251] and between orbitofrontal cortex (OFC) and PFC[173].

DA is operationally related to reward and motivation, influencing various aspects of behavior, including reward prediction errors and reinforcement learning[247,253], the level at which rewards trigger responses, the speed of learning through associations, and the ability to adapt when actions are unsuccessful[158]. This indicates that DA plays a role in how individuals perceive their environment, especially in computing the value attribution of environmental cues. Evidence shows that both hyper and hypo-responsive DA system could result from different ELA exposures[158]. This may imply that the DA system indeed adapts in a flexible manner under conditions of exposure to ELA, thereby conferring phenotypic versatility. This perspective aligns with the idea that DA-related genes are plasticity genes, involved in the differential susceptibility to environmental influence[155] and with evidence showing that the dopaminergic system has a prolonged maturation, making it especially vulnerable to environmental effects[18,254-256] (this topic is further explored in chapter II).

ELA exposure impacts on the dopaminergic system can lead to alterations in encoding of reward value and thus to fundamental changes in reinforcement learning and consequentially behaviour. This vision is corroborated by studies showing that poor fetal growth or LBW [92,96] is related to increased intake of palatable foods, rich in fat and sugar[2,238,241,257,258] and altered hedonic responses to sweet taste[231,259]. In animal models, this altered feeding behavior is related to altered dopaminergic signaling in the NAc and medial PFC measured by chronoamperometry recordings[238,239] and differential tyrosine hydroxylase (TH) content in response to sweet food intake in the OFC and NAc[242]. Even though LBW appears to increase appetitive behaviour towards palatable foods, evidence shows that it can also diminish reinforcement learning, demonstrated by decreased conditioned place preference to a palatable diet[241] and lower monetary-related expenditure to buy a snack in an ecological food choice test[173]. These changes are subtle but persist over the life span, likely contributing to increased adiposity, metabolic disarrangements and the development of chronic diseases later in life, such as psychiatric and cardiometabolic[110,260].

Dopaminergic alterations have been implicated in co-occurring psychiatric and cardiometabolic disorders[261,262]. Dopamine is influenced by core metabolic signals such as insulin, which modulates dopaminergic signaling in the PFC[263] and insulin receptor distribution overlaps with dopaminergic projections[111]. Leptin also is considered to play a role in DA reward system affecting the regulation of feeding behaviour[264]. The dopaminergic system also regulates inflammation through dopamine receptors influence on T cells[265,266] which are lymphocytes and part of the immune system[267]. Dopamine is known to be involved in the etiology and

pathophysiology of psychiatric disorders[268] and related to the action of antidepressant and attention-deficit/hyperactivity disorder medication[269,270].

Considering the dopaminergic system's late maturation, its role in environmental responsivity, reward processing and decision-making, and its connection to metabolic and psychiatric disorders, DA emerges as an important neurobiological target for the investigation of GxE interaction effects involved in the development and maintenance of cardiometabolic and psychiatric comorbidities, and the biological individual differences in response to ELA exposure.

As a starting point for investigating the contribution of DA to the biological mechanism involved in the programming by early life environmental conditions, we first conducted a comprehensive literature review (Chapter II), published as a book chapter in 2019, on the role of the interplay between environment and DA as the basis for the early origins of menta health[260]. In this chapter, we delved into the characteristics of the dopaminergic system, its development, the concept that the dopaminergic system could play a role in individual variation in susceptibility to environmental influences, and methods for capturing this role using genetic studies.

Next, we empirically explored the role of a genetic score based on the predicted gene expression of a dopaminergic gene in modulating ELA impact on chapter III. Published in the journal “Appetite” in 2020, this study explored the interplay between positive early-life environmental factors and genomics in children's emotional eating behavior, with a specific focus on the role of the prefrontal DRD4 predicted gene expression. Our findings indicate that children exposed to less positive early environments had higher emotional overeating, a non-adaptive behavior linked to obesity. More positive environments were associated with a

decrease in emotional overeating, but this was especially true for children characterized by high levels of predicted prefrontal DRD4 gene expression, confirming that variations in the genetic background linked to the expression of DRD4 influence the individual responsivity to ELA effects on emotional eating.

Chapters IV and V delved deeper into the discussion about methods to capture GxE interaction effects in large populations of humans using genome-wide approaches. In chapter IV, we empirically demonstrated that signals derived from GWAS (main genetic effects) do not significantly overlap with variants from GWEIS (GxE interaction effects) in the same outcomes and population. This explains why traditional PRS derived from GWAS are not suited to explore GxE effects. In line with our goal to elucidate possible underlying biological mechanisms involved in the relationship between ELA, the genetic background and psychiatric/cardiometabolic comorbidity, we adopted a genetic tool (ePGS) based on a complex system in biology approach. The ePGS leverages information from basic science research while also maintaining a genome-wide perspective. The performance of ePGS was compared to PRS in Chapter V across several experiments, demonstrating that the ePGS is better suited to represent individual variations in biological functions than the standard PRS. These chapters were fundamental to advance and better equip my final empirical study in Chapter VI to explore the main objective of the thesis.

Chapter VI details the investigation of the interplay between ELA exposure and variations in a dopamine-related striatal gene network on psychiatric and cardiometabolic disorders. We applied the ePGS technique and generated a score based on the expression of the dopamine

transporter gene (SLC6A3) network in the striatum. We observed that the SLC6A3 gene network-based score moderated the effect of ELA on psychiatric and cardiometabolic disorders in adults and adolescents. We also described brain regions involved in this relationship and the possible involvement of insulin signaling disturbances as a mechanism involved in these associations.

Chapter II. Comprehensive review of the literature

Publication information: Barth, B., Portella, A. K., Dubé, L., Meaney, M. J. & Silveira, P. P. (2019). The Interplay Between Dopamine and Environment as the Biological Basis for the Early Origins of Mental Health. in *Early Life Origins of Ageing and Longevity,* 121-140 (Springer).

The Interplay Between Dopamine and Environment as the Biological Basis for the Early Origins of Mental Health

Barbara Barth, André K. Portella, Laurette Dubé, Michael J. Meaney and Patricia Pelufo Silveira

P. P. Silveira (Corresponding author)
e-mail: patricia.silveira@mcgill.ca

B. Barth, M. J. Meaney, P. P. Silveira
Integrated Program in Neuroscience, McGill University, Montreal, QC, Canada

B. Barth, P. P. Silveira
Department of Psychiatry, Faculty of Medicine and Ludmer Center for Neuroinformatics and Mental Health, McGill University, Montreal, QC, Canada

A. K. Portella · L. Dubé
Desautels Faculty of Management, McGill Center for the Convergence of Health and Economics, McGill University, Montreal, QC, Canada

M. J. Meaney, P. P. Silveira
Douglas Hospital Research Centre, 6875 Boulevard LaSalle, Montreal H4H 1R3, QC, Canada

M. J. Meaney
Singapore Institute for Clinical Sciences, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore

A. Vaiserman (ed.), Early Life Origins of Ageing and Longevity,

Healthy Ageing and Longevity 9, https://doi.org/10.1007/978-3-030-24958-8_7

Keywords Perinatal stress · Dopamine neurons · Mental health · Attention deficit hyperactivity disorder · Depression · Metabolic dysregulation

Introduction

The knowledge that conditions existent during the perinatal period have persistent effects on individuals' functioning, health and disease patterns across the lifetime is well established in the literature[1,2]. This is evident as the decline in neonatal mortality, especially since mid-1960's, took place[3], resulting from the improvement of perinatal medical care and subsequent reduction in mortality rates[4]. One large study from Finland, that covered 96% of all children born in the region in 1966 and followed them up until 14 years of age, is one of the first consistent retrospective studies to demonstrate that perinatal conditions have persistent effects on children's neurodevelopment. Among the 12,058 participants, 411 children were considered as having low birth weight (< 2500 g) and this group had a higher incidence of cerebral palsy, epilepsy, severe hearing defects and educational problems[5]. A Danish retrospective cohort study involving 4300 participants born between 1973 and 1975 demonstrates a relationship between low birth weight and poor cognitive performance[2]. Still along those lines, a study using data from 357,768 Swedish military conscripts born between 1973 and 1981 describes that being born small for gestational age (SGA) increases the risk for subnormal intellectual and psychological performance in males[6], even after adjusting for maternal and socioeconomic factors[7]. Amore recent example comes from a 240,351-sample size Western Australian population-based cohort study reporting a U-shaped association between fetal growth and the risk for intellectual disability[8].

Besides the early evidence related to cognitive performance, other studies have identified that adversity such as exposure to stress, infections or malnutrition in early life compromises the quality of growth and development, and increases the risk for chronic, non-transmittable diseases in the offspring in the long term. For instance, increased rates of coronary heart disease and cerebrovascular disease later in life have been observed in subjects born with poor fetal growth[9]. Poor fetal growth has also been associated with glucose intolerance, less capacity to secrete insulin and increased risk for type II diabetes[10,11]. Early adverse conditions exposure also increases the individuals' risk for developing attention deficit hyperactivity disorder (ADHD), schizophrenia and major depression[12,13].

It is thought that these effects could be connected to the concept of developmental plasticity, defined by a critical window during development when a system is plastic and sensitive to the nutritional, hormonal and metabolic environment. For most organs and systems, the critical period occurs in utero and early postnatal life, and may give rise to a range of different physiological or morphological states in response to a variety of conditions existent during development[9]. In the presence of adversity, the fetus/newborn responds through specific adaptations, increasing allocation of energy to favor the brain, heart and adrenal glands development, but reducing the blood flow to other organs and producing lifelong changes in blood pressure and metabolism[14]. Variability and plasticity of physiological and behavioral responses help the growing organism to adapt effectively to the uncertainty of later environmental conditions. This plasticity allows the emergence of phenotypes that are better suited to their surrounding conditions, being more efficiently adapted than it would be possible if the exact same phenotypes were to be produced for all environments. But this process is

done at the expense of the individuals' health[15]. Indeed, healthy subjects born at term with poor fetal growth show impaired vascular endothelial function already in infancy[16,17], disturbed vascular regulation, premature stiffening of the carotid artery[18], and left ventricular hypertrophy[19]. Early life adversities impose a change in the individual's developmental trajectory, altering its metabolism and susceptibility to several outcomes.

Although studies on the influences of environmental variation during the perinatal period on development have mainly focused on being born small for a given gestational age (SGA) or being exposed to intrauterine growth restriction (IUGR) as indicators of prenatal adversity, this concept of early life adversity can be broader, and not necessarily involving differences in fetal growth, also impacting an infant's development and wellbeing[20]. The developmental origins of health and disease (DOHaD) hypothesis suggests that intrauterine signals affect the individual predisposition to specific health outcomes, thus shaping individual differences in the risk for chronic illnesses across the lifespan[21].

Exposure to early life adversity increases the risk for non-transmittable diseases like type II diabetes, but also psychiatric conditions like ADHD or depression, and these common developmental risk factors suggest overlapping underlying mechanisms. Although the co-morbidity between metabolic disease, poor cognitive performance and psychiatric disorders is well-established, the mechanisms are poorly understood. In this chapter we will explore the idea that the co-morbidity between conditions such as psychiatric outcomes and metabolic dysregulation occurs, in part at least, because of an influence of metabolic neuroendocrine signals on the development of the mescorticolimbic dopamine (DA) pathway. The neurons in this pathway regulate cognitive-emotional states, notably impulsivity, and responsivity to

environmental challenges through reward-based decision-making, thus defining behavioral phenotypes that in turn can contribute to promote metabolic disorders.

The Dopaminergic System

Dopamine (3-hydroxytyramine; DA) is a catecholamine neurotransmitter that is synthesized in the brain, as DA does not pass through the blood brain barrier. Instead, its precursor amino acid L-3,4-dihydroxyphenylalanine (L-DOPA) crosses the blood brain barrier and is converted into DA. Tyrosine hydroxylase converts tyrosine to L-DOPA, that in turn is converted into DA by aromatic amino acid decarboxylase (AADC). DA is then transported by VMAT (vesicular monoamine transporter) to inside the neurotransmitter vesicles. Vesicles containing DA move towards the presynaptic membrane as an electrical impulse arrives at the terminal and the vesicle fuses with the presynaptic membrane, releasing the neurotransmitter into the synaptic cleft. There, DA can bind to specific proteins called dopamine receptors on the membrane of the postsynaptic neuron. Dopamine transporter (DAT), which is a membrane-spanning protein, pumps DA out of the synaptic cleft back into the presynaptic cytosol and vesicle. Dopamine reuptake by DAT provides the primary mechanism through which DA concentration in the synaptic cleft is balanced. Moreover, DA receptor D2 acts as a presynaptic auto receptor and also plays a role in regulating the dopaminergic system by providing feedback inhibition. This controls cell firing, and the synthesis, release, and uptake of DA[22-24].

There are four distinct pathways of DA signaling. The tuberoinfundibular pathway refers to a group of DA neurons in the arcuate nucleus of the hypothalamus that projects to the median eminence. There, DA is released into the portal vessels, acting to inhibit the secretion of

prolactin from the anterior pituitary. The nigrostriatal pathway originates in the substantia nigra, and projects to the dorsal striatum. Degeneration of these projections has been shown to cause Parkinson's Disease, impairing planning, initiation, and control of movements, and for that reason this area is thought to be implicated in motor activity[25]. The mesocortical pathway, and the mesolimbic pathway also referred to the mesocorticolimbic system projects from the midbrain to the striatum, limbic and frontal cortical regions. Particularly the mesocortical pathway projects from the ventral tegmental area (VTA) to the frontal and temporal cortices, especially the anterior cingulate, entorhinal, and prefrontal cortices. The mesolimbic pathway also originates in the VTA but instead innervates the ventral striatum, including the nucleus accumbens. This mesocorticolimbic system is involved in cognitive-emotional states, namely reward-based decision-making, and the experience of pleasure, impulsivity, concentration and executive functions[26,27] and for that reason it is the focus of this chapter.

Alterations on the dopaminergic pathways can lead to increased sensitivity to reward and impulsivity[26] and consequently to poor decision-making processes, prompting non-adaptive behaviors such as addiction and altered eating behavior[28-30]. Evidence deriving from the fact that drugs including amphetamine and methylphenidate, known for being dopamine enhancers, improve behavioral symptoms of most children with ADHD, suggests that the DA systems plays a role on the onset and maintenance of this condition. In fact, DA system genes were described as candidate genes for the well-established heritability of ADHD[31]. In the field of schizophrenia research, it is understood that an abnormal neurochemistry related to presynaptic striatal hyper dopaminergia is the common pathway that explains the disease symptoms[32]. Finally, DA is thought to play a role, at least to some extent, in major depression symptoms, since

impairments in motivation, psychomotor speed, concentration and anhedonia are all related to the disorder, and also regulated in part by the dopaminergic systems[27]. Thus, DA dysregulation seems to be the basis of several neurological and behavioral disorders.

Besides that, there is evidence that both prenatal and post-natal adversity are linked to alterations in the dopaminergic pathways, in humans as well as in animal models[32]. DA pathways seem to play a role on the interplay between the influence of the environment and the development of non-communicable diseases during the life-course (see **Figure 1**).

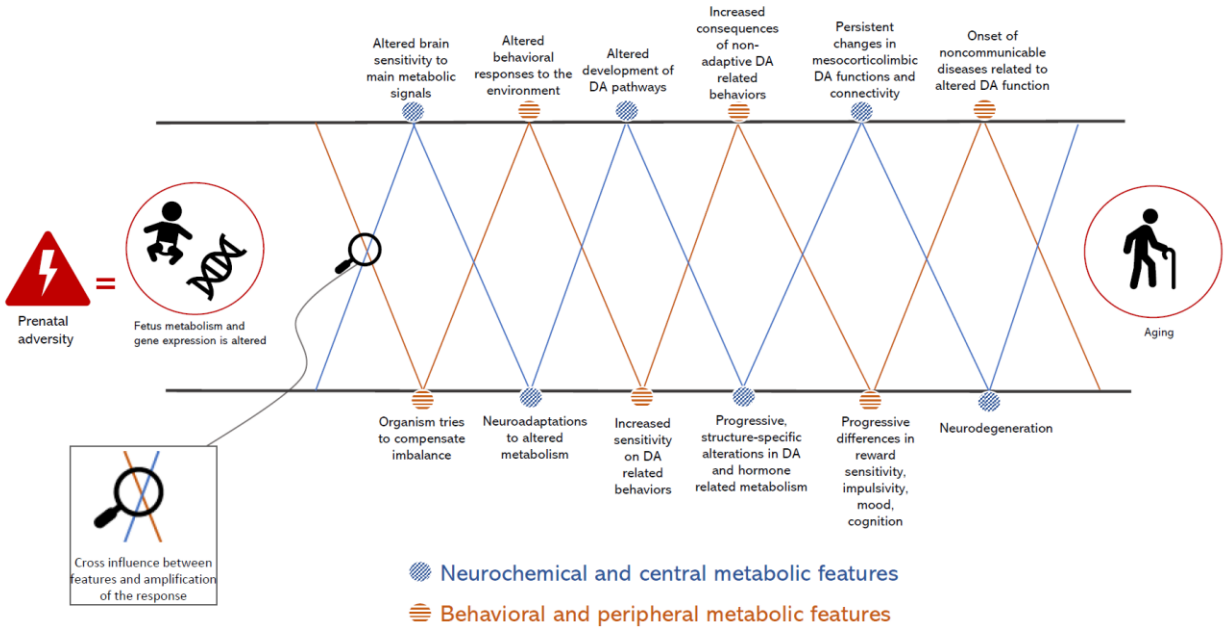


Figure 1. *Theoretical framework on the interplay between environmental adversity, altered metabolism and DA altered function across development.*

The Development of Dopaminergic Pathways

There is evidence that the development of DA pathways is prolonged, reaching maturation at around early adulthood[33] alongside the maturation of the pre-frontal cortex (PFC). DA is one of

the most important PFC neuromodulators, given its role on reward-based decision making. Studies performed in animal models have quantified the density of dopaminergic afferents to the pre-frontal cortex (PFC) during development, describing that it begins as early as the embryonic development, starting with axon extensions from the VTA and increasing during neonatal, juvenile and adult periods[34].

In humans, mesocorticolimbic DA axons continue to grow during adolescence towards the PFC[35] and there is evidence that netrin-1 receptor DCC is responsible for coordinating this maturation. Work by Reynolds et al. (2018)[33] in animal models has demonstrated that DCC acts as a guidance cue receptor, controlling the extent of growth by determining the axons' final target in the PFC. Changes in this growth trajectory can significantly modify PFC structural and functional development[33]. Interestingly, these axons are especially vulnerable to environmental effects[35], increasing the individuals' susceptibility to develop several disturbances when exposed to stress or adversity during development.

In conclusion, DA pathways, especially the mesocorticolimbic pathway, finish their development later in life when compared to other neurotransmitter systems, suggesting that this pathway is susceptible to the influence of the environment for a longer period of time, being an obvious candidate for a biological mechanism involved in the programming by prenatal and postnatal adverse conditions. This idea strongly corroborates with the concept of DA genes being considered as "plasticity genes" (see below) and points the importance of studying this system when investigating the impact of environmental adversities on neurodevelopment.

Early Life Adversity: Beyond Low Birth Weight

As mentioned above, environmental conditions existent during early developmental stages have a dramatic influence on the health/disease patterns of the individual over the life course. The measure of low birth weight has been used in the literature as a marker of exposure to fetal adversity, although it is known that fetal adversity not necessarily impacts birth weight. For instance, the presence of multiple psychosocial stressors during pregnancy is associated with higher systolic and diastolic blood pressure in children aged 5–7 years[36]. Higher prenatal stress during the first 20 weeks of pregnancy is a predictor, among adults, of mood dysregulation, lower overall gray-matter volume, and lower gray-matter volume in mid-dorsolateral frontal cortex, anterior cingulate cortex, and precuneus[37]. Social adversity during the prenatal period is a risk factor for elevated inflammation in adulthood, independently of the exposure to adversities during childhood[38]. Fetal exposure tomaternal depression during pregnancy also has persistent effects on immune function of the young adult offspring[39]. Despite the large amount of evidence, studies exploring the relationship between prenatal adversity and risk for disease in the offspring focus either exclusively on the prenatal social environment[38], maternal mental health[40,41] or biological risk[12,42,43], but these conditions are highly inter-correlated in the lives of children, and have yet to be considered in a cumulative manner.

To tackle the issues described above, Silveira et al. (2017)[20] conducted a study in two community birth cohorts and created a cumulative prenatal adversity score, accounting for information on health during pregnancy, birthweight, gestational age, income, domestic violence/sexual abuse, marital strains, as well as maternal smoking, anxiety, and depression. As

a result of the effort to improve the representation of the early adversity environment, they were able to demonstrate that the cumulative adversity score was a better predictor of neurodevelopmental outcomes than any single factor in isolation. The knowledge about the interplay between components that represent an adverse environment is crucial to design better and more precise preventive measures and interventions to improve neurodevelopment and physical health in the general population.

Genetic Studies Inform About Biological Functions

Genetic studies contribute vastly to our knowledge of the etiology and mechanisms underlying different diseases. Although association studies between variation in the environment, human behavior and specific outcomes can inform about risk factors, they do not apprise which biological processes are implicated. Genetic studies enlighten about how gene expression can modulate these associations and contribute to the understanding of the link between the functionality of a system and a specific outcome.

Earlier studies using genes to better understand the etiology and mechanism of a particular disease involved candidate genes, that focused on testing the association between a specific variant (e.g. a single nucleotide polymorphisms, SNP) and a given disease[44]. Those studies made a significant impact in the scientific community, but also raised questions on the extend of the contribution of this method to complex questions, including the onset of noncommunicable diseases such as obesity. Considering the millions of existing SNPs on the human genome, it is very unlikely that a single SNP could explain much of the variation on the risk to develop a specific condition, or on a behavioral trait. It is now clear that complex phenotypes such as

obesity and mental conditions arise from the cumulative influence of multiple genetic variants[45,46].

This shift in the understanding of disease risk, together with advances in genomic technologies, has led to larger and better powered genome-wide association studies (GWAS) that have permitted large scale analyses of common markers, by associating SNPs inside a gene with a specific outcome. For example, in a large sample size study conducted to find susceptibility loci for coronary artery disease (CAD), SNPs significantly associated with this outcome were able to explain approximately 10.6% of CAD heritability[47]. The use of methods of genomic risk profiling is consistent with the idea that the genetic contribution to a certain condition is derived from a combination of small effects from many genetic variants. To consider the effects of many SNPs, a new concept of polygenic risk score (PRS) was introduced. The PRS summarizes individual's genetic risk for a specific condition[48], characterizes subject's response to therapy[49], or describes variation in specific measures associated with a disease. A polygenic risk score is calculated for each subject in the target sample as a sum of the risk alleles count, weighted by the effect size described in a discovery GWAS[50,51]. The specific alleles that compose the PRS also come from GWAS studies, like the CAD study mentioned above, that have remained significant after multiple comparison corrections. As an example, Scott et al. (2012)[52] performed a GWAS for fasting insulin, adjusting for age, sex and body mass index (BMI), compiling data from 108,557 individuals from 56 studies. They demonstrated that loci associated with fasting insulin concentrations also show an association with lipid levels and fat distribution, suggesting a relationship between genetic variation in these loci and cardio-metabolic risk. GWAS studies have identified several variants associated with complex traits, although still leaving questions

open, especially regarding the mechanisms involved in these associations. This approach neglects the fact that genes operate in networks and code for precise biological functions in specific tissues. Besides that, since the significant SNPs come from GWAS studies, only the ones that "survive" calculations to adjust for multiple comparisons will end up being recognized as significant, resulting in strong genetic main effects and leaving no space for the potential environmental variability contained in the sample.

More recently, new methods are accounting for mechanisms that can further explain variations in biological processes associated with unfavorable outcomes. For example, gene expression is composed by three main constituents: a genetically determined component, a trait-related component, and a component determined by the remaining factors, including the environment[53]. To represent the genetically determined component of gene expression, these authors created an algorithm called Predixcan, proposing a gene-based association method that directly tests the molecular mechanisms through which genetic variation affects phenotype. This is done by estimating the amount of gene expression that is determined by an individual's genetic profile and correlating this biologically imputed gene expression with the trait of interest. More specifically, genotype information from a sample of interest is compared to a reference dataset that has both genotype and gene expression information, then a tissue-specific prediction model involving a machine learning algorithm is used to estimate the genetically determined component of gene expression from the subjects of the target sample[53]. One example of the use of this method can be found on the work by Huckins et al. (2017)[54] in which data from 40,299 schizophrenia cases and 65,264 47 matched controls were used to predict gene expression levels in 12 brain regions. They found 413 genes associated with

schizophrenia across 12 brain regions, being the Dorsolateral Prefrontal Cortex (DLPFC) the one with higher number of associated genes.

Although the focus on the function of a single gene is interesting, it is known that genes operate in networks, and code for precise biological functions in specific tissues. We recently developed a novel approach to genomic profiling, informed by biological function, and characterizing gene networks based on the levels of coexpression with a determined gene in a specific tissue. This genetic score is called ePRS[20,55,56]. The principle of gene networks considers that gene expression is co-regulated by other genes, and consequently genes involved in the same network are expected to have similar expression profiles[57]. Analyzing genomic data through gene sets defined by functional pathways represents a potentially powerful and biologically oriented link between genotypes and phenotypes[58]. Recent advances in this field comprise the assessment of spatial and temporal transcriptomes in the human brain by the BrainSpan[59] and HBT (Human Brain Transcriptome) datasets[60] and the integration of genetic variation with gene expression in the brain by the Genotype-Tissue Expression (GTEx) project[61]. Using these online databases, matrices of co-expression can be generated, and the genetic variation and association with gene expression in the genes that compose the network can be used to reflect the function of the machinery involved in that biological process. An example of the application of such method is the work by Silveira et al. (2017)[20]. A tissue-specific (hippocampal) ePRS score was created for the SLC6A4 or serotonin transporter gene, which has been related to the responsivity to environmental adversity and effects on psychopathology across the life span. Using a cumulative environmental score of prenatal adversity, they searched for interactions between the SLC6A4 ePRS and environmental quality on neurodevelopmental and socio-

emotional outcomes. The ePRS reveled significant interaction effects that were not identified with the use of candidate polymorphism 5-HTTPLR variant.

These new genomic approaches integrate information from molecular neurobiology with GWAS technology to develop biologically-informed polygenic scores based on gene co-expression or genetically predicted gene expression in specific brain regions, creating novel measures to identify vulnerability for childhood behavioral phenotypes that predict later neuropsychiatric conditions in community-based samples, and gene network by environment interactions.

Diathesis Stress Versus Differential Susceptibility

As shown on the previous section, genes modulate the cellular response to environmental variation. Although discrete and potentially differential gene by environment interactions are difficult to be detected using simple association studies, some theoretical paradigms guide the understanding of these relationships. The dominant paradigm on gene by environment interaction studies is based on the diathesis-stress hypothesis, which states that some individuals are more vulnerable than others to the negative effects of the environment (e.g., insensitive parenting, childhood maltreatment, poverty). However, this theoretical framework does not consider variations in resilience, for instance, raising an intriguing question on why would natural selection craft an individual to be more susceptible only to the negative effects of the context?

The alternative differential susceptibility hypothesis[62,63], firstly observed in psychiatric-genetic research[64], suggests that individuals vary both in relation to how much they are negatively affected by environmental adverse events[65,66] and how much they are positively influenced by

the provision of resources and support[67] or the simple absence of contextual adversity[68]. In fact, it has been proposed that individuals should vary in their susceptibility to environmental influences[62] based on an evolutionary perspective, where the future is uncertain. In order to maximize the probability of survival/reproduction, natural selection would favor systems/genes that are able to respond to both poor as well as rich environmental conditions, so that the offspring would be "hedging their bets" against an unclear future.

This theoretical framework has advantages since it considers a broader spectrum of environmental influences. One applicability of this concept can be seen on the proposed idea of "plasticity genes", in which DA is one of the main "plasticity systems" that may have been set up as a form of preparation of the individual to vary its responses according to diverse environmental conditions[68]. This enhanced sensitivity to the environmental context, therefore, increases the range of phenotypic possibilities, and from the research stand point, moves the interest not only to specific vulnerabilities, but to identify both the several patterns of environmental sensitivity, and the significant factors involved in the manifestation of these patterns[69].

A study from our group showed, for instance, that girls carrying the 7-repeat allele of the DRD4 gene and living under adverse socioeconomic conditions have higher fat intake, while those carrying the same gene variant but living in a healthy environment have lower fat intake when compared to non-carriers[70]. This suggests that the previously considered obesity "risk allele" (DRD4 7-repeat)[71] in fact determines openness to environmental modification and/or intervention. A metanalysis conducted with data from 15 studies also revealed results that corroborate the idea of dopamine genes functioning as "plasticity genes". Bakermans-

Kranenburg and Van Ijzendoorn (2011)[72] described that children with a larger number of less efficient dopamine-related variants performed worse in negative environments but also profited the most from positive conditions, in comparison with children with a lower number of these variants.

Shifting from a "vulnerability" to a "differential susceptibility" paradigm not only enables the study of the full range of negative and positive G versus E interactions, but also has the potential to bring more impactful and targeted interventions to improve health outcomes of the individuals who are also the most vulnerable.

Evidence of Dopamine as the Biological Bases of Early Life Programming

As mentioned above, early life adversity could include many different types of adversities happening pre or postnatally. In the past, we have studied the effects of intrauterine growth restriction on behaviors involving sensitivity to reward and impulsivity in different human cohorts. We have shown, for instance, that poor fetal growth is linked to alterations in the hedonic responses to sucrose as early as the first day of life in human newborns[73], a finding corroborated by Rotstein et al. (2015)[74] and Laureano et al. (2016)[75]. We also demonstrated that 3-year old girls born SGA are more impulsive towards a sweet reward using the Snack Delay Task[76], a behavioral feature that is associated with fat preferences and higher body mass index later in childhood[76]. Similarly, another study demonstrated that SGA children at 10 years of age have significantly higher percent energy intake derived from fat when compared to controls, which is associated with higher waist circumference, insulin and HOMA-IR levels[77]. Our group showed that adult women born with severe growth restriction have a higher intake of

carbohydrates, and increased carbohydrate to protein ratio in their diets[78]. The association between being born low birthweight and having specific food preferences later in life, initially described by our group [78], was later replicated in several different human cohorts around the world[79-83].

Although consistent, the studies in humans are correlational. “Bedside-to-Bench” translational approaches with relevant animal models hold the promise of (1) establishing causal relations and (2) identifying underlying mechanisms. The latter is critical for developing objective measures of risk at the level of the individual child. We have explored the long-term effects of poor fetal growth using a rat model based on caloric restriction of 50% initiated at gestational day 10[84]. Pups from food restricted and controls dams are fostered by control dams within 24 hours of life, which ensures growth restriction only during the fetal period[85-87]. The experiments reveal a higher preference for palatable foods in food restricted (FR) animals with a choice between standard and palatable chow (diet with higher contents of sugar and fat)[85,87]. The behavioral phenotype of increased palatable food consumption in FR rats is comparable to that described in humans[73,75,76,78,88-92]. Following our findings of altered taste reactivity to sucrose in human newborns[73], we also saw that FR rat newborns demonstrate more persistent hedonic responses to sucrose when compared to control pups[93] already in the first day of life. FR animals also have reduced conditioned place preference to sweet food when compared to controls[87], which combined with the behavioral profile described above strongly suggests that FR affects the functioning of the dopaminergic mesocorticolimbic pathway, closely associated with appetite regulation and eating behaviours.

Indeed, we described robust alterations in tyrosine hydroxylase (TH), an enzyme involved in DA synthesis, as well as phospho-tyrosine hydroxylase (pTH) levels in the NAcc of FR rats[87], and in the orbitofrontal cortex (OFC) in response to sweet food intake[85]. FR rats have reduced levels of dopamine type 2 (D2) receptors in the NAcc when compared to controls[87], which explains FR rats' inability to condition their preference to a place paired with palatable food[94].

In humans, much of the evidence suggesting DA as a biological basis for early life programming was generated in gene by environment interaction studies. For instance, we recently saw an interaction between a multilocus score reflecting DA signaling capacity and poor fetal growth on spontaneous sugar intake in 48-month children. Using five polymorphisms to create a composite score, the hypofunctional variants (TaqIA-A1 allele, DRD2-141C Ins/Ins, DRD4 7-repeat, DAT1-10-repeat, Met/Met-COMT) received the lowest scores. While in IUGR children there was a correlation between the genetic score and the consumption of sugar, no association was found in non-IUGR children[95].

Levitan et al. (2017)[96] showed in two birth cohorts (one from Canada and another from Netherlands) a significant interaction between maternal sensitivity and the presence of the 7-repeat allele (7R) of DRD4, predicting higher body mass indices (BMI) and/or obesity risk. When exposed to poor maternal sensitivity, 7R carriers have a higher chance of being obese or overweight, especially in Canadian girls or in Dutch boys. The presence of 7R is also associated with higher body mass index (BMI) in women who had seasonal affective disorder and were born in the spring[97], suggesting a fetal programming effect.

As mentioned above, we have shown that variations in this specific mutation of the DRD4 gene interacts with socioeconomic status (SES) according to the differential susceptibility framework,

to predict fat intake in girls at 4 years of age. In other words, the same individuals who are genetically more prone to develop obesogenic behaviorswhen raised in low SES conditions, are also more predisposed to eat less fat when raised in a supportive, high SES environment[70]. As a follow up study, we used the entire genotype information in the same cohort to calculate the genetically predicted gene expression of DRD4 in the prefrontal cortex, evaluating the differential responsivity to positive scenarios on eating outcomes. There was a significant interaction between the exposure to positive environments and the predicted prefrontal DRD4 gene expression on emotional over-eating measured by the Children Eating Behavior Questionnaire applied at 48 months. This interaction also followed the differential susceptibility framework, in which the children that have high predicted DRD4 gene expression and show elevated emotional eating in a less positive environment, have less emotional eating symptoms in more positive environments[98]. This highlights the idea of dopamine genes acting as "plasticity genes", as noted by Bakermans-Kranenburg and Van Ijzendoorn (2011)[72] meta-analysis, proposing the dopamine-related genes as markers of differential susceptibility, and other studies showing that variations in genes that code proteins implicated in the dopamine pathways are sensible to environmental variation[99,100].

In the previously mentioned study from our group, reporting significant interactions between a score for cumulative prenatal exposure to adversity and the ePRS based on the serotonin transporter (SLC6A4) on neurodevelopmental outcomes[20], an enrichment analysis of the genes represented in the polygenic score also suggests the involvement of DA in these fetal programming effects. The most significant biological process enriched in the score was the

dopaminergic neuron differentiation, which can be explained by the common source for monoaminergic progenitors during neurodevelopment[101,102].

It is known that early adverse conditions increase the risk for many different psychopathologies, including ADHD[12,13]. Based on that, Neuman et al. (2007)[103] described that children carrying specific variations of the DAT1 and DRD4 genes, when exposed to prenatal smoking, are more likely to be either diagnosed or have symptoms of ADHD in comparison with non-exposed, non-carrier children. This once more suggests a role of the dopaminergic pathways on modulating the relationship between early environmental exposure to adversity
and the development of psychopathology.

Following the idea that genes operate in networks, Miguel et al. (2019)[56] created an expression-based polygenic score that reflects variations in the function of the dopamine transporter DAT1 gene network (ePRS-DAT1) in the prefrontal cortex (PFC). Using data from two prospective birth cohorts (Canadian based Maternal Adversity, Vulnerability and Neurodevelopment—MAVAN; and GUSTO Growing Up in Singapore Towards Healthy Outcomes), they evaluated differences in cognitive flexibility according to the exposure to hypoxic-ischemic conditions at birth. More intense exposure to hypoxic-ischemic conditions was associated with longer latency to respond and lower accuracy in the attentional set shifting paradigms—measures related to attentional flexibility—but only in the high DAT1 ePRS group. In addition, the relationship between these genes involved in the machinery associated with prefrontal DAT1 function and the PFC and thalamic gray matter volumes was different between children exposed or not to hypoxic-ischemic conditions. These findings indicate that variations in the function of the DAT1 gene

network seem to be important for attention flexibility and its deviances, especially in the context of early life adversity like poor oxygenation levels at birth.

Being born small for gestational age (SGA), a proxy of a poor intrauterine environment, is associated with higher risk of hospitalization for all mental disorders, higher risk of anxiety and adjustment, personality and psychotic disorders[104]. At least until young adulthood, individuals born SGA are at increased risk of severe mental disorders such as schizophrenia[105] and suicide attempts and completeness[106], independently of their gestational age at birth. Schizophrenia is linked to altered dopamine signaling[32], suggesting once more a possible link between early life adversity, altered DA signaling and risk for later mental health disturbances.

All the above findings indicate that exposure to early life adversities could modulate the behavioral phenotype over the life course, inducing neurochemical responses and adaptations that are, again, reflected in altered behaviors. This could, in the long term, lead to unfavorable outcomes. **Figure 1** outlies the proposed theoretical framework on how early adversity affects an individual's development. Exposure to prenatal adversity leads to fetal altered metabolism and gene expression[21]. This prompts the organism to compensate, and neuroadaptations occurring in response to the altered metabolism take place as a trade-off (represented by the orange dots and lines). Because of these neuroadaptations, very specific behavioral features appear already early in life, for instance altered hedonic responses to sucrose[73,75]. These small, persistent alterations in the behavioral response help shaping the development of the DA mesocorticolimbic pathways, affecting neurotransmission and the connectivity between the striatal and prefrontal region[107]. These in turn trigger altered behaviors such as increased impulsivity[76,92,108] altered reward sensitivity[26] and preferences for palatable foods[77-79,82,109]. All

these progressively contribute to generate consequences of non-adaptive DA related behaviors, and these in turn lead to structure-specific alterations in DA and hormone-related metabolism. The continuation of this process, perpetuated by the behavioral features (e.g. chronic increased intake of high fat, high sugar foods), leads to systemic overload and exhaustion, loss of homeostasis, chronic diseases such as type II diabetes[107,110], cardiovascular disease[14], atherosclerosis[111], mood disorders[104,106], Alzheimer's disease [112,113] and neurodegeneration[114].

Conclusions and Final Overview

We gathered evidence of the role of dopamine pathways on modulating the relationship between environmental adversities and later health outcomes. There is an intricate cascade of successive metabolic and behavioral events that feed each other forward and seem to contribute to both metabolic disarrangements and psychiatric disease/neurodegeneration in the long term. This life course approach represents a major change in how we think and study human lives, adding new dimensions of recurring interactions between context, and sensitivity/resilience (plasticity) in biological processes. Beyond understanding the biological factors involved in dopaminergic plasticity effects, this approach also highlights the importance of understanding the key environmental factors that trigger such plastic adaptations. Our focus on dopamine genes as "plasticity genes", and the view of differential susceptibility hypothesis as an important player in this phenomenon, opens up the possibility of both identifying vulnerability but also opportunities for prevention.

Prenatal adversity is associated with higher risk of hospitalization for mental disorders, higher risk of anxiety and adjustment disorders, personality and psychotic disorders[104],

schizophrenia[105], suicide attempts and completeness[106], type II diabetes[11,115], hypertension[36], cardiovascular disease[116] and atherosclerosis[18]. The cost to society should focus not only on health outcomes, but also poor academic achievement and reduced human capital[117-119]. Our studies of the biological basis for developmentally-determined co-morbid metabolic and psychiatric conditions represent a novel approach to understanding the pathophysiology of common, chronic illnesses. These studies have direct implications for targeted preventive measures against metabolic and mental health diseases associated with the exposure to early life adversity.

References

1. Silveira, P.P. & Manfro, G. Retrospective Studies. in *Perinatal Programming of Neurodevelopment. Advances in Neurobiology*, Vol. 10 (ed. (eds), A.M.) (Springer, New York, NY, 2015).
2. Sørensen, H.T. *et al.* Birth weight and cognitive function in young adult life: historical cohort study. *BMJ* **315**, 401-403 (1997).
3. David, R.J. & Siegel, E. Decline in neonatal mortality, 1968 to 1977: better babies or better care? *Pediatrics* **71**, 531-540 (1983).
4. Kessel, S.S., Villar, J., Berendes, H.W. & Nugent, R.P. The changing pattern of low birth weight in the United States 1970 to 1980. *JAMA* **251**, 1978-1982 (1984).
5. Rantakallio, P. & von Wendt, L. Prognosis for low-birthweight infants up to the age of 14: a population study. *Developmental Medicine & Child Neurology* **27**, 655-663 (1985).
6. Lundgren, M., Cnattingius, S., Jonsson, B. & Tuvemo, T. Intellectual performance in young adult males born small for gestational age. *Growth Hormone & IGF Research* **14**, 7-8 (2004).
7. Bergvall, N., Iliadou, A., Tuvemo, T. & Cnattingius, S. Birth characteristics and risk of low intellectual performance in early adulthood: are the associations confounded by socioeconomic factors in adolescence or familial effects? *Pediatrics* **117**, 714-721 (2006).
8. Leonard, H. *et al.* Relation between intrauterine growth and subsequent intellectual disability in a ten-year population cohort of children in Western Australia. *American journal of epidemiology* **167**, 103-111 (2007).

9. Barker, D., Osmond, C., Winter, P.D., Margetts, B. & Simmonds, S. Weight in infancy and death from ischaemic heart disease. *The Lancet* **334**, 577-580 (1989).

10. Barker, D.J. *et al.* Type 2 (non-insulin-dependent) diabetes mellitus, hypertension and hyperlipidaemia (syndrome X): relation to reduced fetal growth. *Diabetologia* **36**, 62-7 (1993).

11. Phipps, K. *et al.* Fetal growth and impaired glucose tolerance in men and women. *Diabetologia* **36**, 225-228 (1993).

12. Lahti, M. *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* **45**, 985-999 (2015).

13. Murphy, S.K. *et al.* Maternal infection and stress during pregnancy and depressive symptoms in adolescent offspring. *Psychiatry research* **257**, 102-110 (2017).

14. Visentin, S. *et al.* Early origins of adult disease: low birth weight and vascular remodeling. *Atherosclerosis* **237**, 391-399 (2014).

15. Barker, D.J. The developmental origins of well-being. *Philos Trans R Soc Lond B Biol Sci* **359**, 1359-66 (2004).

16. Rich-Edwards, J.W. *et al.* Longitudinal study of birth weight and adult body mass index in predicting risk of coronary heart disease and stroke in women. *Bmj* **330**, 1115 (2005).

17. Davis, G.K., Newsome, A.D., Ojeda, N.B. & Alexander, B.T. Effects of Intrauterine Growth Restriction and Female Sex on Future Blood Pressure and Cardiovascular Disease. *Curr Hypertens Rep* **19**, 13 (2017).

18. Martin, H., Hu, J., Gennser, G. & Norman, M. Impaired endothelial function and increased carotid stiffness in 9-year-old children with low birthweight. *Circulation* **102**, 2739-2744 (2000).

19. Wang, K.C. *et al.* Akt signaling as a mediator of cardiac adaptation to low birth weight. *J Endocrinol* **233**, R81-r94 (2017).

20. Silveira, P.P. *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* **29**, 1601-1617 (2017).

21. O'Donnell, K.J. & Meaney, M.J. Fetal Origins of Mental Health: The Developmental Origins of Health and Disease Hypothesis. *Am J Psychiatry* **174**, 319-328 (2017).

22. Ford, C. The role of D2-autoreceptors in regulating dopamine neuron activity and transmission. *Neuroscience* **282**, 13-22 (2014).

23. Beaulieu, J.-M. & Gainetdinov, R.R. The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacological reviews* **70**, 110.002642 (2011).

24. Brady, S., Siegel, G., Albers, R.W. & Price, D. *Basic neurochemistry: principles of molecular, cellular, and medical neurobiology*, (Academic press, 2011).

25. Hopes, L. *et al.* Magnetic resonance imaging features of the nigrostriatal system: biomarkers of Parkinson's disease stages? *PloS one* **11**(2016).

26. Bear, M.F., Connors, B.W. & Paradiso, M.A. *Neuroscience exploring the brain*, (Wolters Kluwer, 2016).

27. Dunlop, B.W. & Nemeroff, C.B. The role of dopamine in the pathophysiology of depression. *Archives of general psychiatry* **64**, 327-337 (2007).

28. Wilson, G.T. Eating disorders, obesity and addiction. *European Eating Disorders Review* **18**, 341-351 (2010).

29. Loxton, N.J. & Tipman, R.J. Reward sensitivity and food addiction in women. *Appetite* **115**, 28-35 (2017).

30. Robbins, T.W. & Clark, L. Behavioral addictions. *Current Opinion in Neurobiology* **30**, 66-72 (2015).

31. Li, D., Sham, P.C., Owen, M.J. & He, L. Meta-analysis shows significant association between dopamine system genes and attention deficit hyperactivity disorder (ADHD). *Human molecular genetics* **15**, 2276-2284 (2006).

32. Howes, O.D. & Kapur, S. The dopamine hypothesis of schizophrenia: version III—the final common pathway. *Schizophrenia bulletin* **35**, 549-562 (2009).

33. Reynolds, L.M. *et al.* DCC receptors drive prefrontal cortex maturation by determining dopamine axon targeting in adolescence. *Biological psychiatry* **83**, 181-192 (2018).

34. Willing, J., Cortes, L.R., Brodsky, J.M., Kim, T. & Juraska, J.M. Innervation of the medial prefrontal cortex by tyrosine hydroxylase immunoreactive fibers during adolescence in male and female rats. *Developmental psychobiology* **59**, 583-589 (2017).

35. Hoops, D. & Flores, C. Making dopamine connections in adolescence. *Trends in neurosciences* **40**, 709-719 (2017).

36. Van Dijk, A.E., Van Eijsden, M., Stronks, K., Gemke, R.J. & Vrijkotte, T.G. The association between prenatal psychosocial stress and blood pressure in the child at age 5–7 years. *PLoS One* **7**(2012).

37. Marecková, K. *et al.* Prenatal Stress, Mood, and Gray Matter Volume in Young Adulthood. *Cereb Cortex* **29**, 1244-1250 (2019).

38. Slopen, N. *et al.* Early origins of inflammation: An examination of prenatal and childhood social adversity in a prospective cohort study. *Psychoneuroendocrinology* **51**, 403-13 (2015).

39. Plant, D.T., Pawlby, S., Sharp, D., Zunszain, P.A. & Pariante, C.M. Prenatal maternal depression is associated with offspring inflammation at 25 years: a prospective longitudinal cohort study. *Transl Psychiatry* **6**, e936 (2016).

40. O'Donnell, K.J., Glover, V., Barker, E.D. & O'Connor, T.G. The persisting effect of maternal mood in pregnancy on childhood psychopathology. *Dev Psychopathol* **26**, 393-403 (2014).

41. Pearson, R.M. *et al.* Maternal depression during pregnancy and the postnatal period: risks and possible mechanisms for offspring depression at age 18 years. *JAMA Psychiatry* **70**, 1312-9 (2013).

42. Räikkönen, K. *et al.* Depression in young adults with very low birth weight: the Helsinki study of very low-birth-weight adults. *Arch Gen Psychiatry* **65**, 290-6 (2008).

43. Laursen, T.M., Munk-Olsen, T., Nordentoft, M. & Bo Mortensen, P. A comparison of selected risk factors for unipolar depressive disorder, bipolar affective disorder,

schizoaffective disorder, and schizophrenia from a danish population-based cohort. *J Clin Psychiatry* **68**, 1673-81 (2007).

44. Auer, P.L. & Stitziel, N.O. Genetic association studies in cardiovascular diseases: do we have enough power? *Trends in cardiovascular medicine* **27**, 397-404 (2017).

45. Lee, S.H., Wray, N.R., Goddard, M.E. & Visscher, P.M. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* **88**, 294-305 (2011).

46. Purcell, S.M. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748-52 (2009).

47. Deloukas, P. *et al.* Large-scale association analysis identifies new risk loci for coronary artery disease. *Nature genetics* **45**, 25-33 (2013).

48. Hachiya, T. *et al.* Genetic predisposition to ischemic stroke: a polygenic risk score. *Stroke* **48**, 253-258 (2017).

49. Natarajan, P. *et al.* Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. *Circulation* **135**, 2091-2101 (2017).

50. Wray, N.R. *et al.* Research review: polygenic methods and their application to psychiatric traits. *Journal of Child Psychology and Psychiatry* **55**, 1068-1087 (2014).

51. Dudbridge, F. Power and predictive accuracy of polygenic risk scores. *PLoS genetics* **9**, e1003348 (2013).

52. Scott, R.A. *et al.* Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet* **44**, 991-1005 (2012).

53. Gamazon, E.R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nature genetics* **47**, 1091 (2015).

54. Huckins, L.M. *et al.* Gene expression imputation across multiple brain regions reveals schizophrenia risk throughout development. *Preprint at bioRxiv*, 222596 (2017).

55. Hari Dass, S.A. *et al.* A biologically-informed polygenic score identifies endophenotypes and clinical conditions associated with the insulin receptor function on specific brain regions. *EBioMedicine* **42**, 13 (2019).

56. Miguel, P.M. *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* **86**, 621-630 (2019).

57. Gaiteri, C., Ding, Y., French, B., Tseng, G.C. & Sibille, E. Beyond modules and hubs: the potential of gene coexpression networks for investigating molecular mechanisms of complex brain disorders. *Genes, Brain, Behavior* **13**, 13-24 (2014).

58. Ramanan, V.K. *et al.* Genome-wide pathway analysis of memory impairment in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort implicates gene candidates, canonical pathways, and networks. *Brain imaging behavior* **6**, 634-648 (2012).

59. Miller, J.A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199 (2014).

60. Kang, R., Zeh, H.J., Lotze, M.T. & Tang, D. The Beclin 1 network regulates autophagy and apoptosis. *Cell death differentiation* **18**, 571 (2011).

61. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nature Genetics* **45**, 580 (2013).

62. Belsky, J. Variation in susceptibility to environmental influence: An evolutionary argument. *Psychological inquiry* **8**, 182-186 (1997).

63. Boyce, W.T. & Ellis, B.J. Biological sensitivity to context: I. An evolutionary–developmental theory of the origins and functions of stress reactivity. *Development and psychopathology* **17**, 271-301 (2005).

64. Belsky, J. & Pluess, M. Genetic moderation of early child-care effects on social functioning across childhood: a developmental analysis. *Child Dev* **84**, 1209-25 (2013).

65. Caspi, A. *et al.* Role of genotype in the cycle of violence in maltreated children. *Science* **297**, 851-854 (2002).

66. Caspi, A. *et al.* Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301**, 386-389 (2003).

67. Blair, C. Early intervention for low birth weight, preterm infants: The role of negative emotionality in the specification of effects. *Development and psychopathology* **14**, 311-332 (2002).

68. Belsky, J. *et al.* Vulnerability genes or plasticity genes? *Molecular psychiatry* **14**, 746 (2009).

69. Moore, S.R. & Depue, R.A. Neurobehavioral foundation of environmental reactivity. *Psychological bulletin* **142**, 107 (2016).

70. Silveira, P.P. *et al.* Genetic differential susceptibility to socioeconomic status and childhood obesogenic behavior: why targeted prevention may be the best societal investment. *JAMA pediatrics* **170**, 359-364 (2016).

71. Levitan, R.D. *et al.* The dopamine-4 receptor gene associated with binge eating and weight gain in women with seasonal affective disorder: an evolutionary perspective. *Biological Psychiatry* **56**, 665-669 (2004).

72. Bakermans-Kranenburg, M.J. & Van Ijzendoorn, M.H. Differential susceptibility to rearing environment depending on dopamine-related genes: New evidence and a meta-analysis. *Development psychopathology* **23**, 39-52 (2011).

73. Ayres, C. *et al.* Intrauterine growth restriction and the fetal programming of the hedonic response to sweet taste in newborn infants. *International journal of pediatrics* **2012**(2012).

74. Rotstein, D.L., Healy, B.C., Malik, M.T., Chitnis, T. & Weiner, H.L. Evaluation of no evidence of disease activity in a 7-year longitudinal multiple sclerosis cohort. *JAMA neurology* **72**, 152-158 (2015).

75. Laureano, D.P., Molle, R.D., Portella, A.K. & Silveira, P.P. Facial Expressions in Small for Gestational Age Newborns. *J Child Neurol* **31**, 398-9 (2016).

76. Silveira, P.P. *et al.* Preliminary evidence for an impulsivity-based thrifty eating phenotype. *Pediatric research* **71**, 293-298 (2012).

77. Crume, T.L. *et al.* The Long-term impact of intrauterine growth restriction in a diverse US cohort of children: The EPOCH study. *Obesity* **22**, 608-615 (2014).

78. Barbieri, M.A. *et al.* Severe intrauterine growth restriction is associated with higher spontaneous carbohydrate intake in young women. *Pediatric research* **65**, 215-220 (2009).

79. Kaseva, N. *et al.* Diet and nutrient intake in young adults born preterm at very low birth weight. *The Journal of Pediatrics* **163**, 43-48 (2013).

80. Stein, A.D., Rundle, A., Wada, N., Goldbohm, R.A. & Lumey, L.H. Associations of gestational exposure to famine with energy balance and macronutrient density of the diet at age 58 years differ according to the reference population used. *J Nutr* **139**, 1555-61 (2009).

81. Lussana, F. *et al.* Prenatal exposure to the Dutch famine is associated with a preference for fatty foods and a more atherogenic lipid profile. *Am J Clin Nutr* **88**, 1648-52 (2008).

82. Perälä, M.-M. *et al.* Body size at birth is associated with food and nutrient intake in adulthood. *PLoS One* **7**, e46139 (2012).

83. Kampmann, F.B. *et al.* Being born small-for-gestational-age is associated with an unfavourable dietary intake in Danish adolescent girls: findings from the Danish National Birth Cohort. *Journal of Developmental Origins of Health and Disease* **10**, 488-496 (2019).

84. Desai, M., Gayle, D., Babu, J. & Ross, M.G. Programmed obesity in intrauterine growth-restricted newborns: modulation by newborn nutrition. *Am J Physiol Regul Integr Comp Physiol* **288**, R91-6 (2005).

85. Alves, M.B., Dalle Molle, R., Desai, M., Ross, M.G. & Silveira, P.P. 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* **287**, 73-81 (2015).

86. Cunha, F.d.S. *et al.* Both food restriction and high-fat diet during gestation induce low birth weight and altered physical activity in adult rat offspring: the “Similarities in the Inequalities” model. *PLoS One* **10**, e0118586 (2015).

87. Dalle Molle, R. *et al.* Intrauterine growth restriction increases the preference for palatable foods and affects sensitivity to food rewards in male and female adult rats. *brain research* **1618**, 41-49 (2015).

88. Portella, A.K. *et al.* Effects of in utero conditions on adult feeding preferences. *J Dev Orig Health Dis* **3**, 140-52 (2012).

89. Portella, A.K. & Silveira, P.P. Neurobehavioral determinants of nutritional security in fetal growth-restricted individuals. *Ann N Y Acad Sci* **1331**, 15-33 (2014).

90. Silveira, P.P. Playing with food: the fetal programming of food preferences. *Obesity (Silver Spring)* **22**, 1210 (2014).

91. Dalle Molle, R. & Silveira, P.P. Small for gestational age children have specific food preferences. *J Pediatr* **166**, 1547 (2015).

92. Reis, R.S. *et al.* Impulsivity-based thrifty eating phenotype and the protective role of n-3 PUFAs intake in adolescents. *Translational psychiatry* **6**, e755-e755 (2016).

93. Laureano, D. *et al.* Intrauterine growth restriction modifies the hedonic response to sweet taste in newborn pups–Role of the accumbal μ-opioid receptors. *Neuroscience* **322**, 500-508 (2016).

94. Smith, J.W., Fetsko, L.A., Xu, R. & Wang, Y. Dopamine D2L receptor knockout mice display deficits in positive and negative reinforcing properties of morphine and in avoidance learning. *Neuroscience* **113**, 755-65 (2002).

95. Silveira, P.P. *et al.* Fetal growth interacts with multilocus genetic score reflecting dopamine signaling capacity to predict spontaneous sugar intake in children. *Appetite* **120**, 596-601 (2018).

96. Levitan, R.D. *et al.* A DRD 4 gene by maternal sensitivity interaction predicts risk for overweight or obesity in two independent cohorts of preschool children. *Journal of Child Psychology Psychiatry* **58**, 180-188 (2017).

97. Levitan, R.D. *et al.* A birth-season/DRD4 gene interaction predicts weight gain and obesity in women with seasonal affective disorder: a seasonal thrifty phenotype hypothesis. *Neuropsychopharmacology* **31**, 2498 (2006).

98. Barth, B. *et al.* Genetically predicted gene expression of prefrontal DRD4 gene and the differential susceptibility to childhood emotional eating in response to positive environment. *Appetite* **148**, 104594 (2020).

99. Nikolova, Y.S., Ferrell, R.E., Manuck, S.B. & Hariri, A.R. Multilocus genetic profile for dopamine signaling predicts ventral striatum reactivity. *Neuropsychopharmacology* **36**, 1940-7 (2011).

100. Stice, E., Yokum, S., Burger, K., Epstein, L. & Smolen, A. Multilocus genetic composite reflecting dopamine signaling capacity predicts reward circuitry responsivity. *J Neurosci* **32**, 10093-100 (2012).

101. Abeliovich, A. & Hammond, R. Midbrain dopamine neuron differentiation: Factors and fates. *Developmental Biology* **304**, 447-454 (2007).

102. Cheng, A. *et al.* Monoamine oxidases regulate telencephalic neural progenitors in late embryonic and early postnatal development. *J Neurosci* **30**, 10752-62 (2010).

103. Neuman, R.J. *et al.* Prenatal smoking exposure and dopaminergic genotypes interact to cause a severe ADHD subtype. *Biol Psychiatry* **61**, 1320-8 (2007).

104. Monfils Gustafsson, W., Josefsson, A., Ekholm Selling, K. & Sydsjö, G. Preterm birth or foetal growth impairment and psychiatric hospitalization in adolescence and early adulthood in a Swedish population-based birth cohort. *Acta Psychiatrica Scandinavica* **119**, 54-61 (2009).

105. Abel, K.M. *et al.* Birth weight, schizophrenia, and adult mental disorder: is risk confined to the smallest babies? *Archives of general psychiatry* **67**, 923-930 (2010).

106. Niederkrotenthaler, T., Rasmussen, F. & Mittendorfer-Rutz, E. Perinatal conditions and parental age at birth as risk markers for subsequent suicide attempt and suicide: a population based case–control study. *European journal of epidemiology* **27**, 729-738 (2012).

107. Hajnal, A. & Norgren, R. Accumbens dopamine mechanisms in sucrose intake. *Brain research* **904**, 76-84 (2001).

108. Reis, R.S. *et al.* Poor infant inhibitory control predicts food fussiness in childhood–a possible protective role of n-3 PUFAs for vulnerable children. *Prostaglandins, Leukotrienes Essential Fatty Acids* **97**, 21-25 (2015).

109. Migraine, A. *et al.* Effect of preterm birth and birth weight on eating behavior at 2 y of age–. *The American journal of clinical nutrition* **97**, 1270-1277 (2013).

110. Ravelli, A.C. *et al.* Glucose tolerance in adults after prenatal exposure to famine. *The Lancet* **351**, 173-177 (1998).

111. Giussani, M. *et al.* Cardiovascular risk assessment in children: role of physical activity, family history and parental smoking on BMI and blood pressure. *Journal of hypertension* **31**, 983-992 (2013).

112. Cai, Z., Xiao, M., Chang, L. & Yan, L.-J. Role of insulin resistance in Alzheimer's disease. *Metabolic brain disease* **30**, 839-851 (2015).

113. Rani, V., Deshmukh, R., Jaswal, P., Kumar, P. & Bariwal, J. Alzheimer's disease: Is this a brain specific diabetic condition? *Physiology behavior* **164**, 259-267 (2016).

114. Verdile, G., Fuller, S.J. & Martins, R.N. The role of type 2 diabetes in neurodegeneration. *Neurobiology of disease* **84**, 22-38 (2015).

115. Hales, C.N. & Barker, D.J. Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. *Diabetologia* **35**, 595-601 (1992).

116. Rich-Edwards, J.W. *et al.* Birth weight and risk of cardiovascular disease in a cohort of women followed up since 1976. *Bmj* **315**, 396-400 (1997).

117. Currie, J. Health disparities and gaps in school readiness. *Future Child* **15**, 117-38 (2005).

118. Reichman, N.E. Low birth weight and school readiness. *Future Child* **15**, 91-116 (2005).

119. Strauss, R.S. Adult functional outcome of those born small for gestational age: twenty-six-year follow-up of the 1970 British Birth Cohort. *Jama* **283**, 625-32 (2000).

Connecting statement to chapter III

In Chapter II we discussed in more depth the idea that the dopaminergic system might play a role in individual variation in susceptibility to environmental influences. As a starting point to empirically address this idea, in Chapter III, we explored the role of prefrontal DRD4-predicted gene expression in moderating the impact of early life environment to predict emotional eating in children.

The DRD4 gene encodes the D4 subtype of the dopamine receptor (D4R) and is most expressed in the prefrontal cortex[271,272] corroborating the idea that the DRD4 plays a role in complex behaviours influenced by cortical DA transmission, such as executive control[273]. DA receptors have DA as their primary endogenous ligand and are a class of G-protein-coupled receptors expressed in many types of cells in the CNS but also peripheral organs[272]. The D4R is part of the D2-like receptor family that induces inhibition of adenylyl cyclase[267] and is thought to be expressed postsynaptically on dopamine target cells[272].

The function of the DRD4 gene has been linked to eating disorders[274]. The DRD4 48-base-pair variable number of tandem repeats (VNTR) in exon III 7-repeat allele polymorphism is thought to modify the functionality of the D4R activity and has been shown to moderate the association between exposure to early life environment conditions and altered eating behaviors[275,276]. In 2016 Silveira et al reported that girls living under adverse socioeconomic conditions and carrying this 7-repeat allele polymorphism of DRD4 consume more calories derived from fat compared to non-carriers; however, the same individuals consumed fewer calories derived from fat when living in a privileged economic and social stratum, when compared to non-carriers [277]. We decided to expand this previous work conducted by the Silveira lab, but instead of

investigating a single polymorphism on this gene, we used a more sophisticated genomics methodology, that predicts gene expression of the entire DRD4 gene in tissue specific manner. We also expanded previous literature by accessing the contribution of a positive environment measure in the GxE model, contrasting with the usual focus on negative aspects of the environment observed in the literature.

Altered eating behavior early in life, especially the ones related to pro-intake features, may serve as an endophenotype for later metabolic disorders, including obesity. Emotional eating has a psychological component related to emotional regulation[278] and might also act as an endophenotype for later psychopathology.

Chapter III. Genetically predicted gene expression of prefrontal DRD4 gene and the differential susceptibility to childhood emotional eating in response to positive environment

Publication information: Barth, B., Bizarro, L., Miguel, P. M., Dubé, L., Levitan, R., O'Donnell, K., Meaney, M. J., & Silveira, P. P. (2020). Genetically predicted gene expression of prefrontal DRD4 gene and the differential susceptibility to childhood emotional eating in response to positive environment. Appetite, 148, 104594. https://doi.org/10.1016/j.appet.2020.104594.

Genetically predicted gene expression of prefrontal DRD4 gene and the differential susceptibility to childhood emotional eating in response to positive environment

Barbara Barth[1,2], Lisiane Bizarro[2], Patrícia Maidana Miguel[3], Laurette Dubé[4], Robert Levitan[5], Kieran O'Donnell[1,6], Michael J. Meaney[1,6,7], Patricia Pelufo Silveira[1,3,5,*]

[1] Integrated Program in Neurosciences, McGill University, 3801 University Street, Montreal, Quebec, H3A 2B4, Canada
[2] Programa de Pós-Graduação em Psicologia, Instituto de Psicologia, Universidade Federal do Rio Grande do Sul, 2600 Ramiro Barcelos, Porto Alegre, Rio Grande do Sul, 90035003, Brazil
[3] Programa de Pós-Graduação em Neurociências, Instituto de Ciências Básicas da Saúde (ICBS), Universidade Federal do Rio Grande do Sul, 500 Sarmento Leite, Porto Alegre, Rio Grande do Sul, 90.046-900, Brazil
[4] Desautels Faculty of Management, McGill Center for the Convergence of Health and Economics, McGill University, 1001 Sherbrooke Street West, Montreal, Quebec, H3A 1G5, Canada
[5] Department of Psychiatry, University of Toronto and Centre for Addiction and Mental Health, 250 College St, Toronto, Ontario, M5T 1R8, Canada
[6] Department of Psychiatry, Faculty of Medicine and Ludmer Centre for Neuroinformatics and Mental Health, Douglas Hospital Research Centre, McGill University, 6875 Boulevard Lasalle, Montréal, Quebec, H4H 1R3, Canada
[7] Singapore Institute for Clinical Sciences, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore
[*] **Corresponding author**. Department of Psychiatry, Faculty of Medicine, McGill University, Douglas Hospital Research Centre, 6875 Boulevard LaSalle, Montreal, QC, H4H 1R3, Canada.

E-mail address: patricia.silveira@mcgill.ca (P.P. Silveira).

Funding: This work was supported by Toxic Stress Research network of the JPB Foundation and Ludmer Centre for Neuroinformatics and Mental Health.

Declaration of competing interest: None.

Acknowledgements: This work 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, and by the Ludmer Centre for Neuroinformatics and Mental Health, McGill University. Dr. Levitan acknowledges support from the Cameron Holcombe Wilson Chair in Depression studies, CAMH and university of Toronto.

Abstract

Genetic differential susceptibility states that individuals may vary both by exhibiting poor responses when exposed to adverse environments, and disproportionally benefiting from positive settings. The dopamine D4 receptor gene (DRD4) may be particularly implicated in these effects, including disturbed eating behaviors that might lead to obesity. Here, we explore differential susceptibility to positive environments according to the predicted genetically regulated gene expression of prefrontal cortex DRD4 gene. Using MAVAN as the discovery cohort (Maternal Adversity, Vulnerability and Neurodevelopment) and GUSTO as the replication cohort (Growing Up in Singapore Towards Healthy Outcomes), we analyzed the interaction between a) a Positive postnatal environmental score, that accounts for positive outcomes in the postnatal period and b) the genetically regulated gene expression of prefrontal DRD4, computed using a machine learning prediction method (PrediXcan). The outcome measures were the pro-intake domains (Emotional over-eating, Food Responsiveness, Food Enjoyment and Desire to Drink) from the Child Eating Behavior Questionnaire at 48 months of age (MAVAN) and 60 months of age (GUSTO). The interaction between the positive environment and the predicted prefrontal DRD4 gene expression was significant for emotional over-eating in MAVAN ($\beta=-0.403$, $p < 0.02$), in which the high gene expression group had more or less emotional eating according to the exposure to lower or higher positive environment respectively, showing evidence of differential susceptibility criteria. In the replication cohort, a similar result was found with the pro-intake domain Desire to drink ($\beta = -0.583$, $p < 0.05$). These results provide further evidence for the genetic differential susceptibility, accounting for the benefit of positive environments.

Keywords: Emotional eating, Gene expression, Differential susceptibility, DRD4

Introduction

Genes can modulate the cellular and behavioral responses to environmental variation and some theoretical paradigms guide the understanding of these relationships. The Diathesis-Stress paradigm states that some individuals are more vulnerable than others to the negative effects of the environment[1]. However, it does not consider variations in positive aspects. The genetic differential susceptibility states that individuals may vary both by exhibiting poor responses when exposed to adverse environments, and disproportionally benefiting from positive settings (including the simple absence of adversity). These would occur to guarantee survival in different contexts. This idea is aligned with evolutionary analysis of human development, in which plasticity to environmental variations is set as a bet hedging against an uncertain future, and to avoid a costly mismatch between the individual's ability to face the environmental conditions and the actual challenges that the environment could impose[2-4]. This framework has advantages since it considers a broader spectrum of environmental influences, also shedding light on positive aspects of the environment and its consequences on development. This theoretical concept can also be seen on the proposed idea of “plasticity genes”, in which dopamine seems to have a central role[2,5]. In this sense, individuals that are highly responsive to the environment, in a differential susceptibility perspective, while being more vulnerable to the damaging effects of an exposure to environmental adversity, can also benefit more from positive environmental conditions than the nonresponsive individuals. This is equivalent to the ‘orchid’ children described by Boyce and Ellis[6], in a theory called biological sensitivity to context.

This is corroborated by evidence showing that the mesocorticolimbic pathway finishes its development later in life, compared to other neurotransmitter systems. This pathway therefore is susceptible to the influence of the environment for a much longer period of time, being an obvious candidate for a biological mechanism involved in the programming by environmental conditions. This enhanced sensitivity to the environmental context, associated with specific dopamine signaling, increases the range of phenotypic possibilities, not limited to vulnerabilities, but also involving better outcomes in particular environmental settings[7].

Phenotypes known to be affected by these gene by environment (GxE) interactions include disturbed eating behaviors that can lead to obesity[8]. In fact, alterations on the dopaminergic pathways can lead to increased sensitivity to reward and impulsivity[9]. For example, drugs such as amphetamine and methylphenidate, known for being dopamine enhancers, improve behavioral symptoms of most children with attention deficit hyperactivity disorder (ADHD), suggesting that dopamine signaling plays a role on the onset and maintenance of this condition related to impulsivity and other executive functions impairments[10]. Similarly dopamine function is thought to play a role in major depression symptoms, since impairments in motivation and anhedonia are all related to the disorder, and also regulated in part by the DA neurotransmission systems[11]. These dopamine signaling alterations can lead to poor decision-making processes, prompting non-adaptive behaviors such as addiction and altered eating behavior[12-14].

The dopamine D4 receptor gene (DRD4) exon III VNTR polymorphism has been particularly implicated in these effects. In 2016, Silveira et al. described that variations in this specific mutation interacted with socioeconomic status (SES) according to the differential susceptibility

framework, influencing fat preferences of girls at 4 years of age[15]. The same girls who are genetically more prone to develop obesogenic behaviors (increased fat intake) when raised in low SES conditions, are also less likely to develop obesogenic behaviors when raised in a positive, high SES environment. Similarly van Strien (2015)[16] found that hypofunctional variants of the DRD4 were associated with higher emotional eating in females. However, single polymorphism approaches may not capture the whole complexity of the function of a gene. Novel genomics approaches using machine learning algorithms to predict gene expression in tissue specific regions are available[17], and these are likely able to provide a more comprehensive view of the role of a specific gene in modulating an individual response to environmental variations.

Even though the differential susceptibility hypothesis accounts for both extremes of the environmental influence (positive and negative, including the simple absence of adversity)[4], few studies have used measures that account for positive aspects of the environment[18]. Work is needed to improve empirical evidence on the responsivity to positive or supporting conditions, showing that this theoretical framework is in fact relevant to understand effectiveness of interventions.

Here, we propose to expand previous work done by our laboratory[8,15,19,20] by using an innovative and more comprehensive genomics approach to evaluate differential susceptibility to obesogenic behaviors in children. If the framework is indeed applicable, variations in the predicted DRD4 gene expression in the prefrontal cortex (where D4 receptors are predominantly localized) would be associated with differential responsiveness to positive

circumstances, here represented by measures associated with supporting conditions in the postnatal period.

Methods

Subjects

The sample was derived from the prospective birth cohort MAVAN[21] (Maternal Adversity, Vulnerability and Neurodevelopment) which followed up children at different time points in the first years of life in Montreal (Quebec) and Hamilton (Ontario), Canada. Exclusion criteria were severe maternal chronic illness, placenta previa, and history of incompetent cervix, impending delivery, or a fetus/infant affected by a major anomaly or born at a gestational age less than 37 weeks. Ethical approvals were obtained from obstetricians performing deliveries at the study hospitals and by the ethics committees and university affiliates (McGill University and Université de Montréal, the Royal Victoria Hospital, Jewish General Hospital, Centre hospitalier de l'Université de Montréal, Hôpital Maisonneuve-Rosemont, St Joseph's Hospital and McMaster University, Hamilton, Ontario, Canada). The study was conducted in accordance with the rules and regulations of the university ethics committees and informed consent was obtained from all participants.

Procedure

Information collected at birth as well as at 48 months of age was used. A total of 132 out of 630 participants had data available for all the measures relevant for this study (birth records, genotype, the Child Eating Behavior Questionnaire at 4 years of age and positive postnatal

environmental score). Children and mothers came to the laboratory for testing and to complete the scales (CEBQ, see details below). Birth records were obtained directly from the birthing units.

Predictors

Positive postnatal environmental score: This score accounts for positive environmental conditions on the postnatal period of life. **Figure 2** shows which variables and cut-offs were used to compute this score. Presence of each component established by its cut-off point yield one point. The total score is represented by the summation of points. The score was built in a cumulative index manner[18], accounting for stablished predictors of child health and development[8]. Birth weight percentiles and household gross income were calculated using the local reference[22,23]. Maternal mental health information was extracted from different questionnaires: Beck Depression Inventory, a 21-question multiple-choice self-report inventory[24]; Edinburgh Postnatal Depression Scale (EPDS), a 10-item self-report scale designed to screen for postpartum depression[25] and State-Trait Anxiety Inventory (STAI), a two versions 20 item each self-report scaling to measure state and trait anxiety[26]. To measure types of attachment styles in preschool-aged children the Preschool Separation – Reunion Procedure (PSRP) was used[27,28], having a baseline interaction followed by two separation and reunion episodes lasting 5 min video recorded and scored (reliability k = 0.83). The Family Assessment Device (FAD), a 60-item self-report instrument, was used to assess different domains of family functioning[29]. The Marital Strain Scale of Pearlin and Schooler was used to assess chronic stress with the romantic partner[30]. Lastly, a self-report breastfeeding questionnaire[31] was used to inquire the age at which the baby (in weeks) was fed for the first time with something other

than breast milk, and the age of the baby (in weeks) when mothers stopped nursing (or giving breast milk).

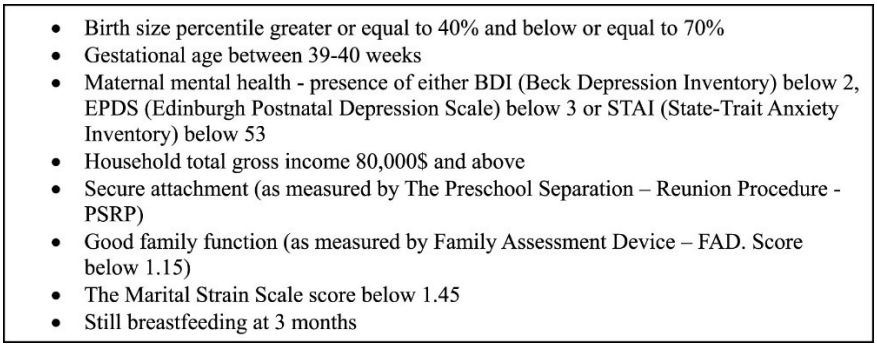


Figure 2. *Variables and cut-offs used to create the Positive postnatal environmental score in MAVAN*. Presence of each component (described in each bullet) yielded 1 point, and the scores represent the summation of points.

The rationale behind including these variables that represent both phenotype measures (e.g. birth size, attachment) and family environment measures (e.g. maternal mental health, marital strain) together into the same score was based on the literature of early life dversity/protection and their long-term effects on child neurodevelopment and behavior. Phenotypes such as birth size have been extensively shown in the literature to have programming" effects on the individual's metabolism, altering the response to the environment and subsequently increasing the likelihood of developing non-communicable diseases such as obesity. For example, a well-known effect of poor fetal growth is the programming of food preferences, widely explored by our lab[20,32-34], and confirmed by others[35-38]. These long-lasting "programming" effects work as if they were a first or immediate "layer" of the environment, dictated by the individual's current metabolic features that result from a past exposure. The inclusion of attachment style is aligned

with the same idea. Evidence has shown its effects on development of several socioemotional characteristics[39,40], having a programming effect on socioemotional development[41,42]. We discussed extensively about these environmental "layers" in a review[8].

Genetically regulated expression of prefrontal DRD4 gene: The genetically regulated expression of prefrontal DRD4 gene is computed using a machine learning prediction method (PrediXcan)[17]. This algorithm was built using a reference dataset from human brain donors (postmortem), being therefore tissue-specific. This reference dataset is composed by data from GTEx project[43], GEUVADIS[44] and DGN[45] containing both genotype and gene expression levels. The PrediXcan prediction model, proposed by Gamazon & et al (2015)[17], uses a machine learning approach to generate algorithms to estimate the genetically determined component of gene expression in specific brain regions from the subject's genotype in the target sample, in this case MAVAN cohort. For the genetic score used in this study, we applied this algorithm to our two samples, and were able to calculate a predicted DRD4 PFC gene expression using the genotype information available in the children from our birth cohorts (**Figure 3**).

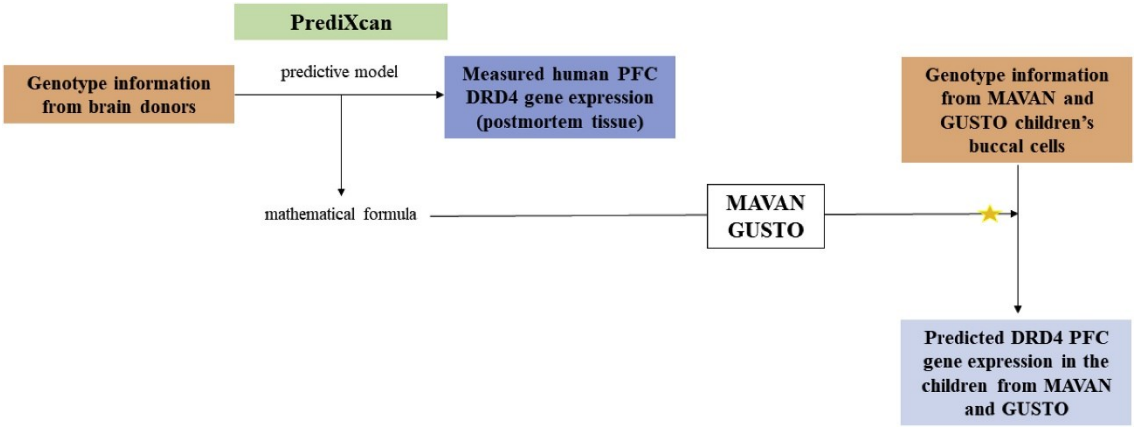


Figure 3. *Scheme for the generation of predicted DRD4 gene expression on discovery and replication cohorts*. PrediXcan prediction model is applied to PFC gene expression data from human brain donors, that also had genotype data. The gene expression information was translated into a model that uses only the genotype information from our sample (in this case, MAVAN or GUSTO) to estimate the gene expression of a given gene (in this case, DRD4).

In MAVAN, we genotyped 242,211 autosomal SNPs using genome-wide platforms (PsychArray/PsychChip, Illumina) from 200 ng of genomic DNA derived from the buccal epithelial cells. After quality control procedures and imputation, 20,790,893 SNPs with an info score >0.80 and posterior genotype probabilities >0.90 were available to be used in PrediXcan.

Outcome

The Child Eating Behavior Questionnaire (CEBQ)[46] is designed to assess children's eating styles that have been hypothesized to contribute both to underweight and overweight. Having domains that reflect behaviors of food pro-intake (positive inclinations for eating or food approach) and anti-intake (or food avoidance). It is a parent-report measure comprised of 35 items, each rated on a five-point Likert scale that ranges from never to always. The instrument is ideal for use in research investigating the early precursors of eating disorders or obesity. The psychometric properties of the instrument have been evaluated and show robust factor structure, good internal and test-retest reliability[46]. A more recent study also shows validity of the questionnaire against behavioral measures of eating[47]. The outcome measures used were the four domains from the questionnaire that reflect pro-intake behaviors[48]: Enjoyment of Food, Food Responsiveness, Desire to Drink and Emotional over-eating. Overall these items describe

pro-intake behaviors either by enjoyment of food, being responsive to food, having a high desire to drink or over-eating in response to negative emotions.

Replication cohort

Subjects

The sample included children from the prospective birth cohort GUSTO (Growing Up in Singapore Towards Healthy Outcomes)[49]. Pregnant women aged 18 years and above were recruited at the National University Hospital (NUH) and KK Women's and Children's Hospital (KKH) in Singapore, being of Chinese, Malay or Indian ethnicity with homogeneous parental ethnic background. Mothers receiving chemotherapy, psychotropic drugs or who had type I diabetes mellitus were excluded. Besides that, for the sake of comparison with the MAVAN cohort, only non-preterm children (born above 37 weeks of gestation) were considered. The study was approved by the National Healthcare Group Domain Specific Review Board (NHG DSRB) and the Sing Health Centralized Institutional Review Board (CIRB). Informed written consent was obtained from each participant. A descriptive paper details other aspects of the cohort[49].

Procedures

We used information collected at birth as well as at 5 years of age. A total of 443 participants out of 1173 had data available for all the measures relevant for this study (birth records, genotype, CEBQ at 60 months of age and positive postnatal environmental score). Children and mothers came to the laboratory for testing and to complete scales. Birth records were obtained directly from the birthing units.

Predictors

Positive postnatal environmental score: Was defined and calculated as described in the MAVAN cohort above, except for attachment style and marital relationship quality that were not available in this cohort. Differences can be seen on **Figure 4** that shows variables and cut-offs used in the GUSTO cohort, chosen to best match the score created in the discovery cohort.

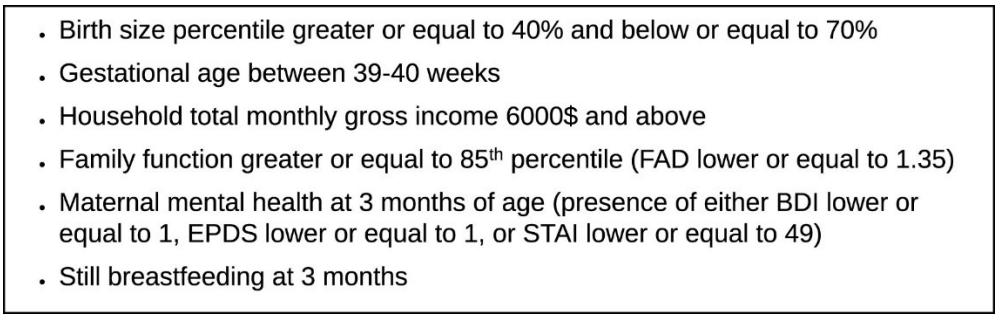


Figure 4. *Variables and cut-offs used to create the positive postnatal environmental score in GUSTO.* Presence of each component (described in each bullet) yielded 1 point, and the scores represent the summation of points.

Genetically regulated expression of prefrontal DRD4 gene: It was computed using the same machine learning prediction method (PrediXcan)[17] and brain region as described in the MAVAN cohort. Genomic DNA in GUSTO was extracted from frozen umbilical cord specimens. Samples were genotyped on Illumina Omni express arrays and on Illumina Exome arrays, following the manufacturer's instructions (Expression Analysis Inc). Further quality control on the genotyping calls were previously described[50]. SNPs were verified for a genotyping rate ≥95% and no deviation from Hardy–Weinberg equilibrium ($P < 0.001$), and minor allele frequency ≥0.05, using PLINK[51,52].

Outcome

The outcome measures were the same used in the MAVAN cohort from the Child Eating Behavior Questionnaire[46], with the four domains that reflect pro-intake behaviors: Enjoyment of Food, Food Responsiveness, Desire to Drink and Emotional over-eating.

Statistical analysis

Statistical analysis of the participants' baseline characteristics was performed using Student's T test for continuous data and chi-square tests for categorical variables (**Table 1** and **Table 2**). For the baseline comparisons, a median split was used to define the high and low DRD4 predicted gene expression groups. For the main analysis, linear regression models using continuous DRD4 predicted gene expression values on the PFC, positive postnatal environmental and the interaction term between these two variables were performed for the four domains of the CEBQ considered in this study (Enjoyment of Food, Food Responsiveness, Desire to Drink and Emotional Over-Eating). Regression analysis were corrected for multiple comparisons. The replication analysis considered statistically significant results using one-tailed P-value thresholds. We considered the analysis done in the discovery cohort (MAVAN) to be exploratory and in this case, we used two-tailed P-value thresholds, since the direction of the forthcoming results were not anticipated. For the analysis done in the replication cohort (GUSTO) we anticipated results direction based on what we found in the discovery cohort. A one-tailed test is appropriate if the estimated value may depart from a reference value in only one direction. For that reason, the one-tailed P value thresholds were considered appropriated to confirm the results direction we saw in the discovery cohort. Preliminary analysis adjusted by sex showed no

main effect or interaction with sex, therefore in the main analysis boys and girls were analyzed together. To verify if the gene by environment interaction finding was aligned with the differential susceptibility model, we followed criteria developed by Roisman et al. (2012)[53]. Three measures were considered; if regions of significance were inside the range of the environmental variation; if the markers PA (proportion affected) and PoI (proportion of interaction) were consistent with differential susceptibility; and if there was absence of nonlinear terms X2 and ZX2.

Table 1. Sample description and differences between High and Low DRD4 predicted gene expression groups in MAVAN.

Sample Description

	Total sample (n=132)		**Low DRD4 (n=67)**		**High DRD4 (n=65)**		
Variable	**Mean or n**	**SD or %**	**Mean or n**	**SD or %**	**Mean or n**	**SD or %**	***p***
Birth weight (g)	3320.95	455.38	3322.79	450.15	3318.93	464.66	0.96
Gestational age (weeks)	39.18	1.21	39.15	1.06	39.2	1.36	0.82
Maternal age at birth (years)	30.81	4.75	31.13	4.04	30.46	5.09	0.4
Montreal site	76	57%	42	31.8%	27	20.4%	0.42
Female sex	68	51%	37	28%	31	23.4%	0.61
Income below	56	44%	45	34%	22	16.6%	0.36

Can$80,000							
Maternal education high school or less	2	1.5%	2	1.5%	0	0.0%	0.47
Positive postnatal environmental score	3.4	1.52	3.57	1.51	3.34	1.52	0.38
Food Responsiveness	2.27	0.8	2.14	0.81	2.39	0.77	0.09
Food enjoyment	3.58	0.75	3.46	0.8	3.72	0.68	0.06
Desire to drink	3	1.07	3.02	1.11	2.98	1.04	0.82
Emotional over-eating	1.61	0.6	1.62	0.6	1.61	0.6	0.9
PrediXCan DRD4 PFC	-0.13	0.22	-0.32	0.15	0.05	0.06	-

MAVAN participants' characteristics by prefrontal DRD4 predicted gene expression group. Data are expressed as means (standard deviations) or number of participants (percentages).

Table 2. Sample description and differences between High and Low DRD4 predicted gene expression groups in GUSTO.

	Sample Description						
	Total sample (n=428)		**Low DRD4 (n=223)**		**High DRD4 (n=205)**		
Variable	**Mean or n**	**SD or %**	**Mean or n**	**SD or %**	**Mean or n**	**SD or %**	**p**
Birth weight (g)	3122.42	427.06	3151.0	422.8	3091.	430.4	0.15
Gestational age (weeks)	38.46	1.28	38.56	1.2	38.35	1.36	0.1
Maternal age at birth (years)	31.31	5.08	31.18	4.91	31.45	5.26	0.58

	Sample Description						
Female sex	203	47.4%	114	51.1%	89	43.4%	0.11
Income below $6,000	302	70,6%	156	70%	146	71.2%	0.77
Maternal education high school or less	277	64.7%	144	64.9%	133	65.5%	0.89
Positive postnatal environmental score	2.11	1.24	2.09	1.27	2.14	1.2	0.67
Food responsiveness	2.4	0.69	2.41	0.69	2.39	0.69	0.83
Food enjoyment	3.5	0.79	3.51	0.82	3.5	0.76	0.89
Desire to drink	2.74	0.9	2.84	0.94	2.62	0.84	0.01*
Emotional over-eating	2.79	0.79	2.76	0.77	2.82	0.82	0.44
PrediXCan DRD4 PFC	-0.01	0.11	-0.10	0.11	0.06	0.04	-

GUSTO participants' characteristics by prefrontal DRD4 predicted gene expression group. Data are expressed as means (standard deviations) or number of participants (percentages).

We examined population structure (i.e. presence of a systematic difference in allele frequencies between subpopulations in a population, possibly due to different ancestry) and the models were adjusted by principal components that reflect population stratification[54,55]. By adding the principal components, we aim to adjust for false results due to ancestry differences. For that, first we pruned our datasets to common variants (MAF>0.05) that were not in linkage disequilibrium (r2<0.20) with a sliding window (50 kilobases) approach that examined linkage disequilibrium in increments of 5 SNPs using PLINK 1.9[56]. We performed a principal component analysis using SMARTPCA on this pruned dataset and generated a scree plot (see Hari Dass et al., 2019[57]) for scree plot for the MAVAN cohort). Based on the inspection of the scree plot, the first three principal components were the most informative of population structure in both cohorts and were included in all analyses. No other co-variates were used in the regression

analysis. Data were analyzed using the Statistical Package for the Social Sciences (SPSS) version 20.0 software (SPSS Inc., Chicago, IL, USA) and R software[58,59]. Significance levels for all measures were set at $p < 0.05$.

Results

Baseline comparisons between predicted gene expression groups can be seen in **Table 1**, **Table 2**. No differences were found between the two groups (high and low predicted prefrontal DRD4 gene expression) in relation to the main confounding variables in both cohorts.

In MAVAN, we observed a statistically significant interaction effect between the positive environment score and the predicted prefrontal DRD4 gene expression on emotional over-eating ($\beta = -0.403$, $p = 0.0159$). A simple slope analysis revealed that a more positive environment is associated with lower emotional over-eating as the DRD4 predicted gene expression increases. On **Figure 5**, groups are divided by plus and minus one standard deviation for the sake of visualization. We confirmed that the interaction is aligned with the differential susceptibility model according to Roisman et al. (2012) method[53], since the regions of significance were inside the range of the environmental variation; moreover, the markers PA = 0.54 and PoI = 0.52 were consistent with differential susceptibility, as well as the absence of nonlinear terms X2 and ZX2. This means that the same genetic profile associated with increased benefit from a more positive environment, is also more affected by a less positive environment, showing more emotional over-eating. After adjusting by multiple comparison this result remains significant.

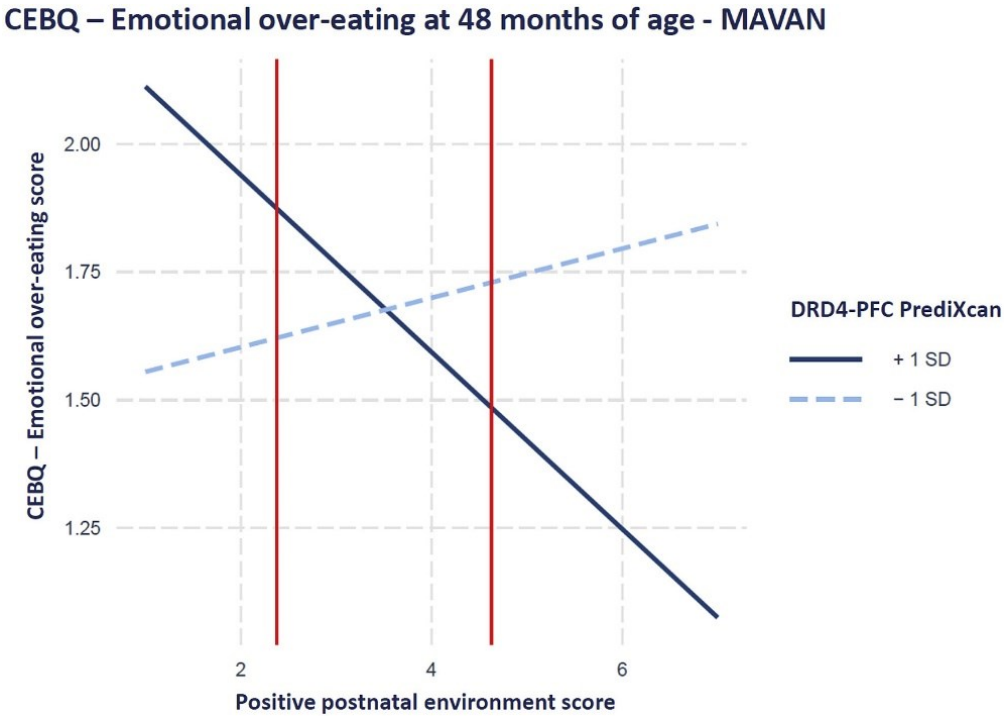


Figure 5. *Evidence of differential susceptibility - Interaction between positive postnatal environmental score and predicted DRD4 gene expression on Emotional over-eating at 48 months of age.* MAVAN Cohort. The vertical lines depict the regions of significance.

On the same analysis, the predicted prefrontal DRD4 expression had an independent effect on emotional over-eating (β = 1.388. p = 0.0240) as did the positive postnatal environmental score (β = −0.098, p = 0.0129). The same association was not found for the other domains in the CEBQ: Desire to drink (β = −0.142, p = 0.62051); Food Enjoyment (β = -0.088, p = 0.660) and Food Responsiveness (β = -0.047, p = 0.968).

In the replication cohort, similar results were found with another pro-intake domain from the CEBQ. The interaction between the positive environment and the predicted prefrontal DRD4 gene expression was statistically significant on the domain desire to drink (β = −0.579,

p = 0.01455). Simple slope analysis revealed that as the score for the positive environment increases and the gene expression score also increases, there is a decrease in the desire to drink score. For the sake of visualization of the results, on **Figure 6** the participants are divided in plus and minus one standard deviation from the mean. After adjusting for multiple comparisons this result was found marginally significant (p = 0.0582). No other association was found for the other CEBQ pro intake domains: Emotional over-eating (β = −0.046, p = 0.3903), Food Enjoyment (β = −0.357, p = 0.08866), Food Responsiveness (β = -0.375, p = 0.0660); no evidence for differential susceptibility was detected in this cohort.

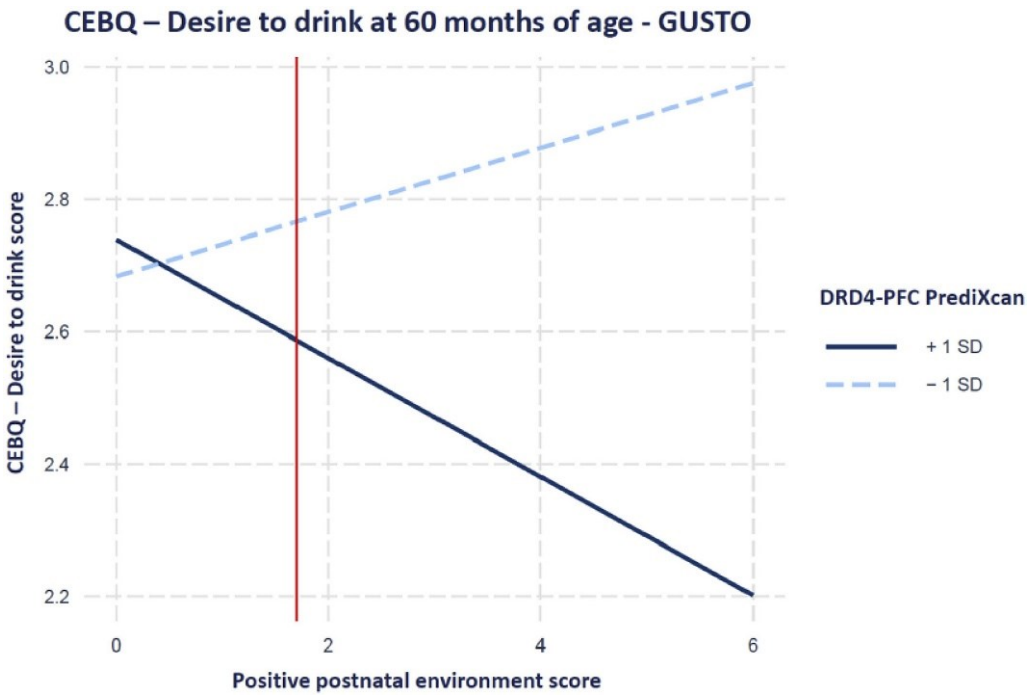


Figure 6. *Interaction between positive postnatal environmental score and predicted DRD4 gene expression on Desire to Drink at 60 months of age.* Gusto Cohort. The vertical lines depict the regions of significance.

Discussion

In this study, we demonstrated on both cohorts that environment and genetics were associated with some obesogenic behaviors in children. In MAVAN, a high predicted prefrontal DRD4 gene expression decreases the risk for the development of behaviors associated with emotional over-eating in children as young as 4 years old that are raised in a more positive environment. Since we found evidence of differential susceptibility, the opposite relationship is also true, in which these same children, if raised in a less positive environment are in a higher risk to develop obesogenic behaviors as measured by the CEBQ instrument. In fact, emotional over-eating has been linked to difficulties in weight loss among adults that underwent treatment for obesity[60] being a stronger predictor of weight gain than life style factors such as little physical activity and consumption of fruits and vegetables[61]. Emotional over-eating seems to be a risk factor not only for the development of obesity but for its maintenance as well.

In the GUSTO cohort, a high predicted prefrontal DRD4 gene expression decreases the risk for the development of behaviors associated with the domain desire to drink in children as young as 5 years old that are raised in a more positive environment. Although the domains desire to drink and emotional over-eating are known to be weakly correlated[46], it is also known that both have a relationship with the onset of obesogenic behaviors[60-62]. Besides that the domain desire to drink is also considered pro-intake, and is associated with the consumption of high sugar-sweetened beverages[63]. In fact, the overconsumption of sugary drinks[64] and the desire to drink domain have been related to obesity and overweight in children[62]. Although this result did not survive correction for multiple comparisons, it could be seen as valid since it emerged from an a

priori hypothesis and previous published results[15,19,65] characterizing this analysis as non-exploratory.

Despite the difference in the significant domains between the two cohorts, we were able to demonstrate the effect of the interaction between positive environmental conditions and the predicted prefrontal DRD4 gene expression on eating behaviors associated with obesity and overweight. Explanations for the dissimilar results between the cohorts may involve cultural or behavioral aspects associated with eating styles. The lack of evidence for differential susceptibility in GUSTO could be explained by the fact that the positive environment score in this cohort does not include an evaluation of attachment styles as does MAVAN, due to the lack of this data in GUSTO.

Evidence from the literature showing the relationship between pro intake behaviors and the function of the DRD4 gene variants[15,16,66,67] and also between dopamine related genes and susceptibility for environment influences[2], corroborates the relationship seen on this work. It is important to emphasize that we used a novel genomic approach to predict gene expression in a tissue specific manner[17], being able to provide a more comprehensive view of the role of a specific gene in modulating an individual response to environmental variations. It seems that individual variation on the function of dopaminergic pathways, here represented by the variations of the predicted prefrontal DRD4 gene expression, could be one of the underlying biological process that explain the relationship between variations in a positive environment and reduced probability to develop obesogenic behaviors. This could be happening by altering the subjects' reward sensitivity and decision-making behaviors at critical time points during development.

Insights from neuroscience and GxE studies are crucial to understand the biological processes underlying children's behavior and susceptibility to negative/positive outcomes. This has implications for understanding the development of several important health outcomes, including growth and its deviations, as well as metabolic alterations.

These results provide further evidence for the genetic differential susceptibility[2], that accounts not only for how vulnerable an individual is to adversity, but also how much they will benefit from positive environments. It is known that children vary according to their susceptibility to the environmental variations, but this framework brings a biological explanation for this observed phenomenon, and accounts for a better characterization of the adverse as well as the positive environment. Indeed, this is demonstrated here, being the characterization of the environment in terms of positive circumstances one of the innovative aspects of this study. It gives strong support for the theoretical framework used, since most of the studies in the area focus on measures characterizing the environment in terms of adversity only[18]. Here we show that even when the starting point is a positive characterization of the environment, a moderation effect in agreement with the genetic differential susceptibility framework can be detected, in this case in relation to eating behavior. Applying this novel approach to the developmental neuropsychology and developmental origins of health and disease agenda guides the elaboration of more efficacious and cost-effective interventions, targeting individuals that would benefit the most from interventions. Furthermore, this broadens the scope of scientific evidence for interventions that focus on promotion of health rather than preventing diseases.

References

1. Patten, S.B. Major depression epidemiology from a diathesis-stress conceptualization. *BMC psychiatry* **13**, 19 (2013).
2. Belsky, J. *et al.* Vulnerability genes or plasticity genes? *Molecular psychiatry* **14**, 746 (2009).
3. Belsky, J. & Hartman, S. Gene-environment interaction in evolutionary perspective: differential susceptibility to environmental influences. *World Psychiatry* **13**, 87-9 (2014).
4. Ellis, B.J., Boyce, W.T., Belsky, J., Bakermans-Kranenburg, M.J. & van Ijzendoorn, M.H. Differential susceptibility to the environment: an evolutionary--neurodevelopmental theory. *Dev Psychopathol* **23**, 7-28 (2011).
5. Barth, B., Portella, A.K., Dubé, L., Meaney, M.J. & Silveira, P.P. The Interplay Between Dopamine and Environment as the Biological Basis for the Early Origins of Mental Health. in *Early Life Origins of Ageing and Longevity* 121-140 (Springer, 2019).
6. Boyce, W.T. & Ellis, B.J. Biological sensitivity to context: I. An evolutionary–developmental theory of the origins and functions of stress reactivity. *Development and psychopathology* **17**, 271-301 (2005).
7. Moore, S.R. & Depue, R.A. Neurobehavioral foundation of environmental reactivity. *Psychological bulletin* **142**, 107 (2016).
8. Dalle Molle, R. *et al.* Gene and environment interaction: Is the differential susceptibility hypothesis relevant for obesity? *Neurosci Biobehav Rev* **73**, 326-339 (2017).
9. Bear, M.F., Connors, B.W. & Paradiso, M.A. *Neuroscience exploring the brain*, (Wolters Kluwer, 2016).

10. Li, D., Sham, P.C., Owen, M.J. & He, L. Meta-analysis shows significant association between dopamine system genes and attention deficit hyperactivity disorder (ADHD). *Human molecular genetics* **15**, 2276-2284 (2006).

11. Dunlop, B.W. & Nemeroff, C.B. The role of dopamine in the pathophysiology of depression. *Archives of general psychiatry* **64**, 327-337 (2007).

12. Loxton, N.J. & Tipman, R.J. Reward sensitivity and food addiction in women. *Appetite* **115**, 28-35 (2017).

13. Robbins, T.W. & Clark, L. Behavioral addictions. *Current Opinion in Neurobiology* **30**, 66-72 (2015).

14. Wilson, G.T. Eating disorders, obesity and addiction. *European Eating Disorders Review* **18**, 341-351 (2010).

15. Silveira, P.P. *et al.* Genetic differential susceptibility to socioeconomic status and childhood obesogenic behavior: why targeted prevention may be the best societal investment. *JAMA pediatrics* **170**, 359-364 (2016).

16. van Strien, T., Levitan, R.D., Engels, R.C. & Homberg, J.R. Season of birth, the dopamine D4 receptor gene and emotional eating in males and females. Evidence of a genetic plasticity factor? *Appetite* **90**, 51-57 (2015).

17. Gamazon, E.R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nature genetics* **47**, 1091 (2015).

18. Silveira, P.P. *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* **29**, 1601-1617 (2017).

19. Silveira, P.P. *et al.* Association between the seven-repeat allele of the dopamine-4 receptor gene (DRD4) and spontaneous food intake in pre-school children. *Appetite* **73**, 15-22 (2014).

20. Silveira, P.P. *et al.* Fetal growth interacts with multilocus genetic score reflecting dopamine signaling capacity to predict spontaneous sugar intake in children. *Appetite* **120**, 596-601 (2018).

21. O'Donnell, K.A. *et al.* The maternal adversity, vulnerability and neurodevelopment project: theory and methodology. *The Canadian Journal of Psychiatry* **59**, 497-508 (2014).

22. Kramer, M.S. *et al.* A new and improved population-based Canadian reference for birth weight for gestational age. *Journal of Pediatrics* **108**, e35-e35 (2001).

23. Québec, I.d.l.s.d. *Enquête générale sur la santé et le bien-être de la population 1998 : questionnaires (QAA et QRI). In D.S. Québec (Ed.) (p. 79). Quebec: Groupe Léger & Léger inc.*, (Quebec: Groupe Léger & Léger inc., 1998).

24. Beck, A.T., Ward, C.H., Mendelson, M., Mock, J. & Erbaugh, J. An inventory for measuring depression. *Archives of general psychiatry* **4**, 561-571 (1961).

25. Cox, J.L., Holden, J.M. & Sagovsky, R. Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. *The British journal of psychiatry* **150**, 782-786 (1987).

26. Spielberger, C.D. & Gorsuch, R. *State-trait anxiety inventory (form Y)*, (Consulting Psychologists Press, 1983).

27. Cassidy, J., Parke, R.D., Butkovsky, L. & Braungart, J.M. Family-peer connections: The roles of emotional expressiveness within the family and children's understanding of emotions. *Child development* **63**, 603-618 (1992).

28. Cassidy, J. & Marvin, R. Attachment organization in 2 1/2 to 4 1/2 year olds. Coding manual.(with the McArthur Working Group on Attachment). *Unpublished coding manual, University of Virginia* (1992).

29. Epstein, N.B., Baldwin, L.M. & Bishop, D.S. The McMaster family assessment device. *Journal of marital family therapy* **9**, 171-180 (1983).

30. Pearlin, L.I. & Schooler, C. The structure of coping. *Journal of health and social behavior*, 2-21 (1978).

31. Adedinsewo, D.A. *et al.* Maternal anxiety and breastfeeding: findings from the MAVAN (Maternal Adversity, Vulnerability and Neurodevelopment) Study. *Journal of Human Lactation* **30**, 102-109 (2014).

32. Ayres, C. *et al.* Intrauterine growth restriction and the fetal programming of the hedonic response to sweet taste in newborn infants. *International journal of pediatrics* **2012**(2012).

33. Barbieri, M.A. *et al.* Severe intrauterine growth restriction is associated with higher spontaneous carbohydrate intake in young women. *Pediatric research* **65**, 215-220 (2009).

34. Bischoff, A.R. *et al.* Low birth weight is associated with increased fat intake in school-aged boys. *Br J Nutr* **119**, 1295-1302 (2018).

35. Crume, T.L. *et al.* The Long-term impact of intrauterine growth restriction in a diverse US cohort of children: The EPOCH study. *Obesity* **22**, 608-615 (2014).

36. Lussana, F. *et al.* Prenatal exposure to the Dutch famine is associated with a preference for fatty foods and a more atherogenic lipid profile. *Am J Clin Nutr* **88**, 1648-52 (2008).

37. Migraine, A. *et al.* Effect of preterm birth and birth weight on eating behavior at 2 y of age–. *The American journal of clinical nutrition* **97**, 1270-1277 (2013).

38. Perälä, M.-M. *et al.* Body size at birth is associated with food and nutrient intake in adulthood. *PLoS One* **7**, e46139 (2012).

39. Cooke, J.E., Kochendorfer, L.B., Stuart-Parrigon, K.L., Koehn, A.J. & Kerns, K.A. Parent–child attachment and children's experience and regulation of emotion: A meta-analytic review. *Emotion* **19**, 1103 (2019).

40. Groh, A.M. *et al.* Attachment and Temperament in the Early Life Course: A Meta-Analytic Review. *Child Dev* **88**, 770-795 (2017).

41. Farrell, A.K. *et al.* Early maternal sensitivity, attachment security in young adulthood, and cardiometabolic risk at midlife. *Attach Hum Dev* **21**, 70-86 (2019).

42. Nonnenmacher, N., Noe, D., Ehrenthal, J.C. & Reck, C. Postpartum bonding: the impact of maternal depression and adult attachment style. *Arch Womens Ment Health* **19**, 927-35 (2016).

43. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nature Genetics* **45**, 580 (2013).

44. Lappalainen, T. *et al.* Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* **501**, 506-511 (2013).

45. Battle, A. *et al.* Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. *Genome Res* **24**, 14-24 (2014).

46. Wardle, J., Guthrie, C.A., Sanderson, S. & Rapoport, L. Development of the Children's Eating Behaviour Questionnaire. *J Child Psychol Psychiatry* **42**, 963-70 (2001).

47. Carnell, S. & Wardle, J. Measuring behavioural susceptibility to obesity: validation of the child eating behaviour questionnaire. *Appetite* **48**, 104-113 (2007).

48. Vandeweghe, L., Vervoort, L., Verbeken, S., Moens, E. & Braet, C. Food Approach and Food Avoidance in Young Children: Relation with Reward Sensitivity and Punishment Sensitivity. *Front Psychol* **7**, 928 (2016).

49. Soh, S.-E. *et al.* Cohort profile: Growing Up in Singapore Towards healthy Outcomes (GUSTO) birth cohort study. *International journal of epidemiology* **43**, 1401-1409 (2013).

50. Qiu, A. *et al.* Prenatal maternal depression alters amygdala functional connectivity in 6-month-old infants. **5**, e508 (2015).

51. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).

52. Qiu, A. *et al.* Effects of antenatal maternal depressive symptoms and socio-economic status on neonatal brain development are modulated by genetic risk. **27**, 3080-3092 (2017).

53. Roisman, G.I. *et al.* Distinguishing differential susceptibility from diathesis-stress: recommendations for evaluating interaction effects. *Dev Psychopathol* **24**, 389-409 (2012).

54. Patterson, N., Price, A.L. & Reich, D. Population structure and eigenanalysis. *PLoS genetics* **2**(2006).

55. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics* **38**, 904-909 (2006).

56. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, s13742-015-0047-8 (2015).

57. Hari Dass, S.A. *et al.* A biologically-informed polygenic score identifies endophenotypes and clinical conditions associated with the insulin receptor function on specific brain regions. *EBioMedicine* **42**, 13 (2019).

58. J.A., L. _jtools: Analysis and Presentation of Social Scientific Data_. R package version 1.1.1, <URL: https://cran.r-project.org/package=jtools>. (2018).

59. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis.*, (New York, 2016).

60. López-Guimerà, G. *et al.* CLOCK 3111 T/C SNP interacts with emotional eating behavior for weight-loss in a Mediterranean population. *PLoS One* **9**, e99152 (2014).

61. Koenders, P.G. & van Strien, T. Emotional eating, rather than lifestyle behavior, drives weight gain in a prospective study in 1562 employees. *J Occup Environ Med* **53**, 1287-93 (2011).

62. dos Passos, D.R., Gigante, D.P., Maciel, F.V. & Matijasevich, A. [Children's eating behaviour: comparison between normal and overweight children from a school in Pelotas, Rio Grande do Sul, Brazil]. *Rev Paul Pediatr* **33**, 42-9 (2015).

63. Cantoral, A. *et al.* Early introduction and cumulative consumption of sugar-sweetened beverages during the pre-school period and risk of obesity at 8-14 years of age. *Pediatr Obes* **11**, 68-74 (2016).

64. van Grieken, A., Renders, C.M., van de Gaar, V.M., Hirasing, R.A. & Raat, H. Associations between the home environment and children's sweet beverage consumption at 2-year follow-up: the 'Be active, eat right' study. *Pediatr Obes* **10**, 126-33 (2015).

65. Levitan, R.D. *et al.* A DRD 4 gene by maternal sensitivity interaction predicts risk for overweight or obesity in two independent cohorts of preschool children. *Journal of Child Psychology Psychiatry* **58**, 180-188 (2017).

66. Fontana, C. *et al.* DRD4 and SLC6A3 gene polymorphisms are associated with food intake and nutritional status in children in early stages of development. *The journal of nutritional biochemistry* **26**, 1607-1612 (2015).

67. Levitan, R.D. *et al.* The dopamine-4 receptor gene associated with binge eating and weight gain in women with seasonal affective disorder: an evolutionary perspective. *Biological Psychiatry* **56**, 665-669 (2004).

Connecting statement to chapter IV

In Chapter III we showed the contribution of predicted dopaminergic gene expression variation on moderating the impact of early environment on eating behavior. A limitation of this study is the use of a single gene. Complex disorders, such as psychiatric and cardiometabolic, are the product of many factors, including multiple genetic signals. Most chronic diseases are polygenic, resulting from the effect of many SNPs[142], that can be mapped to many different genes. GWAS are designed to find multiple SNPs associated with a phenotype in a genome-wide level. A question that remained at this point was if PRSs derived from GWAS were suited to study GxE interaction effects, especially considering the limited success of studies using PRS to investigate GxE interaction effects in mental and physical disorders[279-281]. We then asked, in the next study described in chapter IV, "how do the genetic variants that reveal statistical associations with common disorders or traits compare with those that moderate the impact of environmental conditions?".

Chapter IV. Genome-wide main effect and environment interaction-associated SNPs are distinct

Publication information: This manuscript, in form of a brief communication, is in submission phase, currently being submitted to the Communications Medicine Journal.

Genome-wide main effect and environment interaction-associated SNPs are distinct.

Barbara Barth[1,2,3], Irina Pokhvisneva[2], Patricia Maidana Miguel[4], André Krumel Portella[2], Michael J Meaney[2,4,5,6], Patricia Pelufo Silveira[2,3,4,5]

[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] Department of Psychiatry, Faculty of Medicine, McGill University, Montreal, QC, Canada.
[5] Department of Paediatrics, Yong Lin Loon School of Medicine, National University of Singapore, Republic of Singapore.
[6] Singapore Institute for Clinical Sciences, Agency for Science, Technology and Research, Republic of 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

Abstract

Predisposition to common adult chronic physical and mental diseases is influenced by genetics, environments, and the interaction between these factors. We compared single nucleotide polymorphisms (SNPs) associated with common chronic disorders to those that moderate the association with environmental risk factors. Our findings reveal that genome-wide variants with main effects on common chronic diseases (GWAS) show limited overlap with those underlying gene-environment interaction effects (GWEIS) on the same diseases and in the same population.

Introduction

Genome-Wide Association Studies (GWAS) revolutionized our understanding of the genetic basis of various human conditions by identifying markers associated with diseases and traits[1]. The GWAS Catalog (https://www.ebi.ac.uk/gwas/) is a public database summarizing information from genetic associations for numerous traits including chronic, common adult diseases such as type II diabetes, cardiovascular disease and psychopathology[2]. It provides evidence for the influence of genetic background on the development of these diseases. Recent analyses of GWAS data sets revealed both the pluripotency of genetic influences and the polygenicity of common disorders.

Exposure to adverse conditions (e.g., trauma, emotional abuse, neglect, poverty) over the life course also influences the risk of chronic adult diseases, increasing the risk for a wide-range of outcomes such as major depression and other psychopathologies, ischemic heart conditions, and type II diabetes[3-6]. While these associations apply to the general population, there are

individual differences in susceptibility: not all individuals exposed to adversity develop any form of chronic adult disease[7]. Indeed, measures of exposure to adversity alone cannot distinguish between individuals with or without later problems of physical or mental health[7]. There is now considerable evidence for the importance of heritable genetic variation as a major source of susceptibility to environmental exposure[8].

These findings reflect the complex interplay between genetic and environmental influences in defining individual differences in health, well-being, and productivity over the life span. These findings also raise an important yet unresolved question: how do the genetic variants that reveal statistical associations with common disorders or traits compare with those that moderate the impact of environmental conditions?

Results

To address this issue, we conducted genome-wide association studies (GWAS) on common adult chronic physical and mental diseases using individuals from the UK Biobank. We aimed to identify the genetic background associated with the risk for these conditions. Then, in the exact same population, we performed genome-wide environment interaction association studies (GWEIS) for the same conditions. This analysis aimed to characterize the genetic variants associated with disease risk in response to exposure to early adversity. Finally, we investigated the overlap between single nucleotide polymorphisms (SNPs) identified in the GWAS and GWEIS.

We considered a sample of 97,583 participants from the UK Biobank that had information available on genotype, early life adversity, and targeted health outcomes. Of these, 11,682

(11.9%) reported exposures to early adversity (reported feeling hated or not loved during childhood) and 85,901 (88.1%) reported no such exposures. The outcomes used in the GWAS and GWEIS were cardiovascular disorders (5,713 cases and 91,870 controls), non-insulin dependent diabetes (3,071 cases and 94,512 controls), mood disorders (2,508 cases and 95,075 controls) and neurotic disorders (1,679 cases and 95,904 controls). *See Methods section for a detailed description of inclusion criteria and different variables used in the study.*

Figure 7 demonstrates the overlap between statistically significant SNPs in the GWAS and GWEIS for the different health conditions at the $p<10^{-3}$ threshold. Only a few significant SNPs are commonly identified in both GWAS and GWEIS for each condition. Fisher exact tests are not statistically significant in each instance, meaning that the limited overlap between the two approaches is likely due to chance, considering the entire number of SNPs available for GWAS and GWEIS analyses. In sum, the genetic background associated with risk for these chronic conditions as a main effect (genetic risk identified in the GWAS) is distinct from the genetic background that modifies the association between early adversity and chronic disease (SNPs that respond to environmental variation identified in the GWEIS).

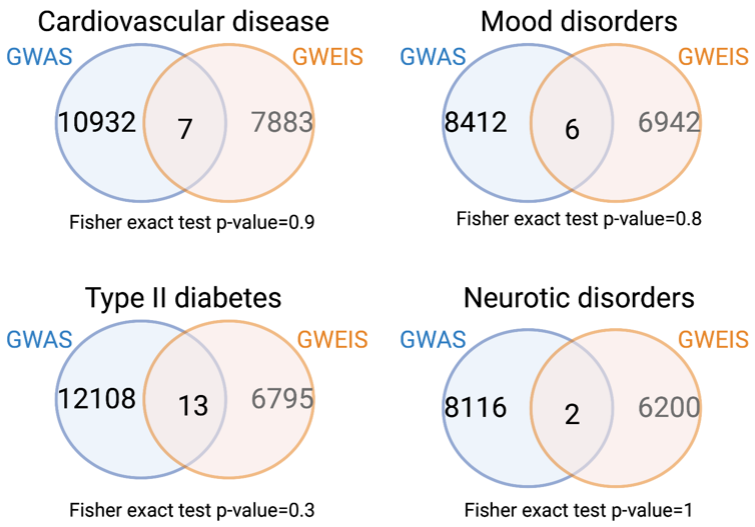
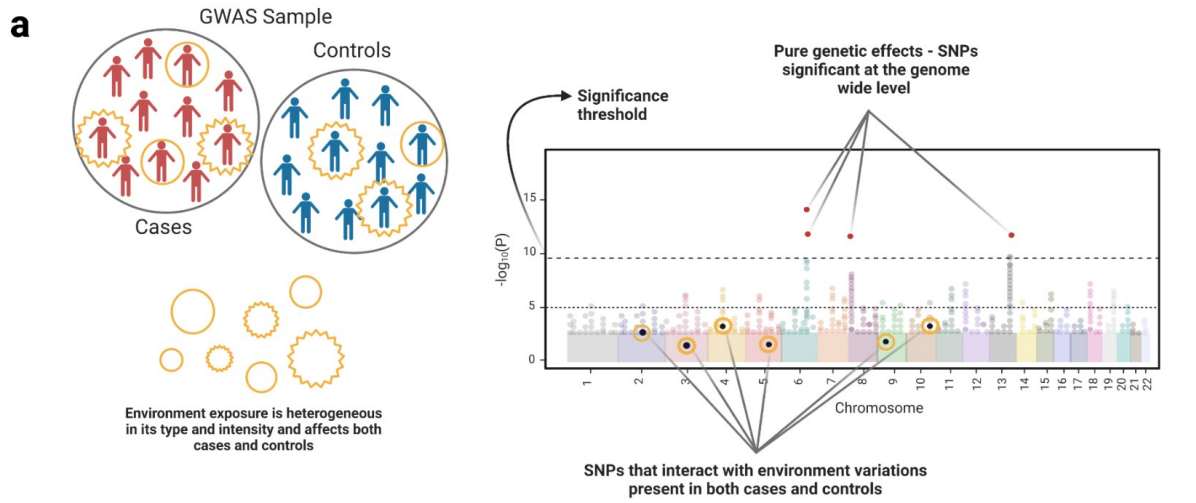


Figure 7. *Investigation of the overlap between significant SNPs identified in the GWAS for different chronic diseases and those identified in the GWEIS (SNP-by-early adversity model) for the same conditions.* There is no significant overlap in the SNPs identified by the two types of genome-wide study in the same population (Fisher exact test is non-significant in the four examples).

A Manhattan plot of significance in the Y axis against chromosomal location in the X axis is a common approach for visualization of GWAS results. Each dot represents a genetic locus, and genome-wide significant loci lay at the top of the plot[9]. GWAS is based on heavily penalized p-values from multiple testing[10] and can identify statistically significant SNPs even in the presence of environmental heterogeneity within the sample. The main effect assumption implies that highly significant GWAS hits will emerge mostly *despite* the large variation in environmental exposures of the participants. These exposures are distributed in the entire sample of the GWAS, both in cases (individuals with the disease) and controls (individuals without the

disease). Moreover, the environmental exposure is randomly distributed among individuals carrying different genotypes (AA, AB or BB) at every SNP. The existence of a significant GxE interaction effect at a given SNP can modify the strength of the association between the number of effect alleles and the outcome, influencing the likelihood of that SNP being highly significant in the GWAS. This is because environmental heterogeneity, present in both cases and controls of the GWAS, interferes with the association between each SNP and the outcome in the entire sample. A statistically significant association for a SNP with the target outcome is less likely to emerge if the association is dependent upon an environmental condition. In Figure 2a, we frame the hypothesis that variants driving GxE interactions (statistically significant SNPs in a GWEIS) will therefore likely be located on non-statistically significant portions of the Manhattan plot from a GWAS. In **Figure 8b**, we confirmed this hypothesis for the four chronic diseases analyzed in this study.


a
GWAS Sample
Controls
Cases
Environment exposure is heterogeneous in its type and intensity and affects both cases and controls
Significance threshold
Pure genetic effects - SNPs significant at the genome wide level
-log10(P)
Chromosome
SNPs that interact with environment variations present in both cases and controls

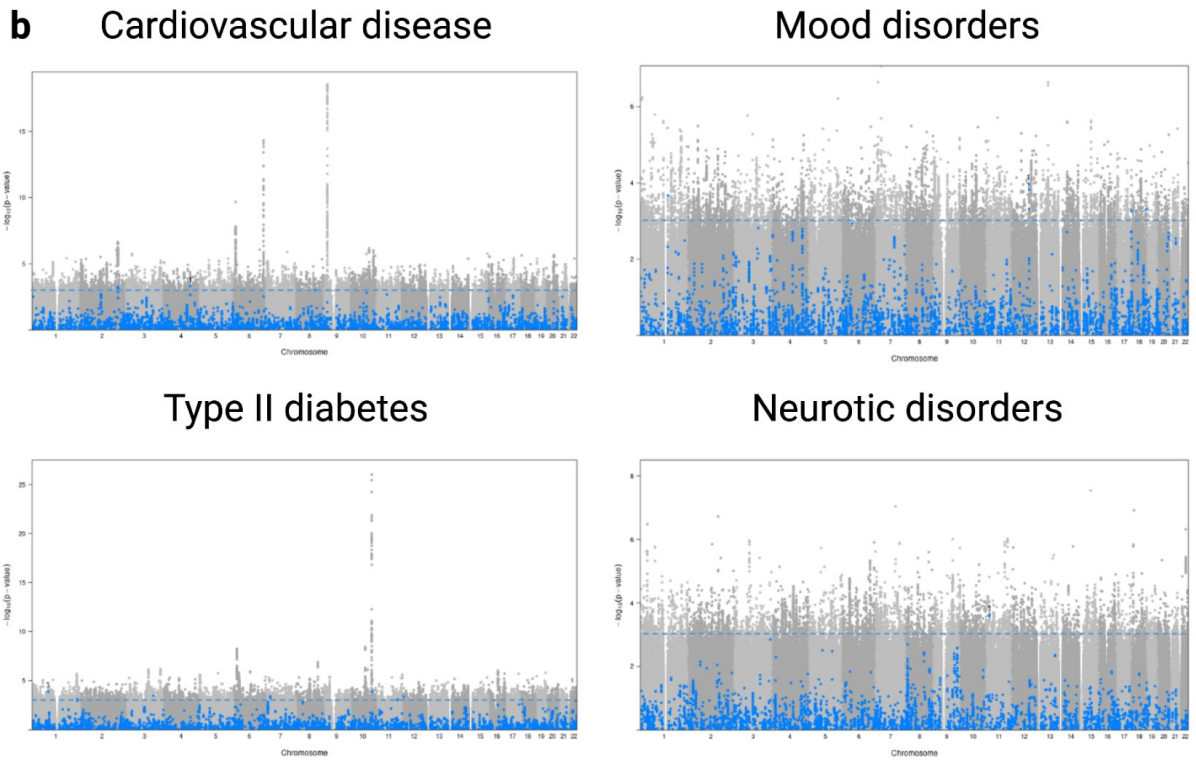

b
Cardiovascular disease
Mood disorders
Type II diabetes
Neurotic disorders
Chromosome

Figure 8. *GWAS may not capture SNPs for which the effect is modified by the environment.* **a**, GWAS design is based on a sample of cases and controls (individuals with or without the target outcome). Both groups may or may not have been exposed to environmental adversity in early life, and this exposure varies in type and intensity. The Manhattan plot orders the genetic variants in a GWAS from the most statistically significant (lower p-value) at the top, to the least statistically significant ones at the bottom. SNPs that remain statistically significant after the stringent adjustments for multiple comparisons do so despite the environmental heterogeneity of the sample. Therefore, SNPs for which the effect is modified by the environment, and thus more likely to be statistically significant in interaction models or GWEIS, are unlikely to reach the upper regions of the GWAS Manhattan plot, reserved for variants with lower p-values. **b**, Manhattan plots for SNPs associated with cardiovascular disease, type II diabetes, mood disorders and neurotic disorders in the GWAS in the same group of 97,583 UK Biobank participants are depicted in gray. SNPs that were statistically significant at the $p<10^{-3}$ threshold in the GWEIS were in the GWAS Manhattan plot and identified in blue. As hypothesized in a, SNPs that are statistically significant in the GWEIS analysis do not reach the top portions of the GWAS Manhattan plot**.**

Discussion

To summarize, we observed that genetic variants linked to main genetic associations with chronic adult diseases, as identified in the GWAS, differ greatly from those for which the association is modified by environmental exposures, or that amplify/diminish the effects of early adversity on adult disease (identified in the GWEIS). These two types of analyses identify different risk components and should be considered accordingly. The results from this study help elucidate the limited success of studies using GWAS-derived polygenic scores to investigate gene-environment interactions in mental and physical disorders[11-13]. The use of GWEIS, though technically challenging and requiring large samples, is an alternative to identify genetic

moderation effects. Moreover, innovative approaches and polygenic metrics that combine functional features of the genome[14-16] or are based on specific pathways or systems responsive to environmental variation[17] are needed.

Methods

Study population

The UK Biobank cohort is a population-based cohort consisting of 502,543 individuals aged 37-73 recruited at 23 centers across the United Kingdom between 2006 and 2010. Participants provided both phenodata and genodata. Genotyping data was available for 487,409 subjects. We excluded participants who withdrew their consent, with inconsistencies in genetic and reported sex, as well as outliers for heterozygosity. Next, we excluded 147,605 participants with shared relatedness of up to the third degree (kinship coefficients >0.044 calculated using the KING software). Also, we retained only those subjects who identified themselves as "white British" (ID21000) and had data available on Adversity (ID20487, ID20489). We also removed variants with minor allele frequency <0.01, an imputation accuracy Info score < 0.1 as well as duplicated and ambiguous SNPs. After applying the above-mentioned criteria, there were 97,583 individuals and 7,351,435 variants in the data set. This research was conducted using the UK Biobank Resource under Application Number 41975. Approval for the UK Biobank was obtained by the North West Multicentre Research 580 Ethics Committee (REC reference 11/NW/0382; www.ukbiobank.ac.uk/ethics/), the National Information Governance Board for Health and Social Care and the Community Health Index Advisory Group. Local Ethics approval was obtained from the Research Ethics Board – Mental Health and Neuroscience subcommittee based at the Centre intégré universitaire de santé et de services sociaux de l'Ouest-de-l'Île-de-

Montréal Research Ethics Board under the number 2022-371, IUSMD 21-73. Information about the genotyping protocol, QC and imputation is found elsewhere [18]. To determine population structure, principal components were computed on UK Biobank individuals. The population structure of the UK Biobank cohort was evaluated using fastPCA [19] algorithm for principal component analysis. To account for population stratification, the first forty principal components were included in the UK Biobank analysis.

Reported Early adversity

The Mental Health online follow-up included questions from the Childhood Trauma Screener CTS-5[20-22]. We considered that participants were exposed to adversity if they reported feeling hated during childhood by a family member sometimes, often or very often ("When I was growing up... I felt that someone in my family hated me", ID20487) or if they reported never or rarely feeling loved during childhood ("When I was growing up... I felt loved", ID20489).

Phenotypes

The phenotypes used in the GWAS and GWEIS were obtained from the diagnoses made during hospital inpatient admissions according to the International Classification of Disease version 10 (ICD-10) (ID 41270): cardiovascular disorders (I20-25 + I70 + I67.2), non-insulin dependent diabetes (E11), mood disorders (F30-F39) and neurotic disorders (F40-F48).

Genetic Association analysis (GWAS and GWEIS)

We applied linear regression analysis using SNPTEST v2.5.4-beta1 to explore the main effect of each SNP (in the GWAS) and the interaction effect between each SNP and the adversity (in the

GWEIS) on the phenotypes, adjusting for genetic sex (ID22001) and age (ID21022) at recruitment, genotype array (ID22000), and first ten genetic principal components (ID22009).

GWAS model: Phenotype ~ (genotype) + genetic_sex + age + genotype_array + PC1 + PC2 + PC3 + PC4 + PC5 + PC6 + PC7 + PC8 + PC9 + PC10

GWEIS model: Phenotype ~ (genotype) * Adversity + genetic_sex_f22001_0_0 + age_f21022 + genotype_array + PC1 + PC2 + PC3 + PC4 + PC5 + PC6 + PC7 + PC8 + PC9 + PC10

Statistical analysis

To investigate the overlap between the significant SNPs in GWAS and GWEIS for each of the four outcomes, we selected top SNPs (p-value < 10-3) in GWAS and GWEIS and explored their overlap using Fisher's exact test.[23]

Data availability

The raw genetic and phenotypic data that support the findings of this study are available from UK Biobank but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Summary statistics of the GWAS and GWEIS are available at https://github.com/SilveiraLab and the UK Biobank website.

References

1. Deloukas, P. *et al.* Large-scale association analysis identifies new risk loci for coronary artery disease. *Nature genetics* **45**, 25-33 (2013).
2. MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Research* **45**, D896-D901 (2017).
3. Bjorkenstam, E., Burstrom, B., Vinnerljung, B. & Kosidou, K. Childhood adversity and psychiatric disorder in young adulthood: An analysis of 107,704 Swedes. *J Psychiatr Res* **77**, 67-75 (2016).
4. Chapman, D.P. *et al.* Adverse childhood experiences and the risk of depressive disorders in adulthood. *J Affect Disord* **82**, 217-25 (2004).
5. Schussler-Fiorenza Rose, S.M., Snyder, M.P. & Slavich, G.M. Adverse childhood experiences, diabetes and associated conditions, preventive care practices and health care access: A population-based study. *Prev Med* **160**, 107044 (2022).
6. Dong, M. *et al.* Insights into causal pathways for ischemic heart disease: adverse childhood experiences study. *Circulation* **110**, 1761-6 (2004).
7. Baldwin, J.R. *et al.* Population vs Individual Prediction of Poor Health From Results of Adverse Childhood Experiences Screening. *JAMA Pediatrics* **175**(2021).
8. Silveira, P.P. & Meaney, M.J. Examining the biological mechanisms of human mental disorders resulting from gene-environment interdependence using novel functional genomic approaches. *Neurobiology of Disease* **178**, 106008 (2023).
9. Gibson, G. Hints of hidden heritability in GWAS. *Nature genetics* **42**, 558-560 (2010).

10. Pe'er, I., Yelensky, R., Altshuler, D. & Daly, M.J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society* **32**, 381-385 (2008).

11. Mullins, N. *et al.* Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. *Psychol Med* **46**, 759-70 (2016).

12. Peyrot, W.J. *et al.* Does Childhood Trauma Moderate Polygenic Risk for Depression? A Meta-analysis of 5765 Subjects From the Psychiatric Genomics Consortium. *Biol Psychiatry* **84**, 138-147 (2018).

13. Trotta, A. *et al.* Interplay between Schizophrenia Polygenic Risk Score and Childhood Adversity in First-Presentation Psychotic Disorder: A Pilot Study. *PLoS One* **11**, e0163319 (2016).

14. de Lima, R.M.S. *et al.* Leptin receptor co-expression gene network moderates the effect of early life adversity on eating behavior in children. *Commun Biol* **5**, 1092 (2022).

15. de Mendonça Filho, E.J., Barth, B., Bandeir, D. R., de Lima, R. M. S., Arcego, D. M. Dalmaz, C., Pokhvisneva, I., Sassi, R. B., Hall, G. B. C., Meaney, M. J., Silveira, P. P. . Cognitive Development and Brain Gray Matter Susceptibility to Prenatal Adversities: Moderation by the Prefrontal Cortex Brain-Derived Neurotrophic Factor Gene Co-expression. *Frontiers in Neuroscience* (2021).

16. Dalmaz, C. *et al.* Prefrontal cortex VAMP1 gene network moderates the effect of the early environment on cognitive flexibility in children. *Neurobiol Learn Mem* **185**, 107509 (2021).

17. Silveira, P.P. & Meaney, M.J. Examining the biological mechanisms of human mental disorders resulting from gene-environment interdependence using novel functional genomic approaches. *Neurobiol Dis* **178**, 106008 (2023).

18. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-209 (2018).

19. Galinsky, K.J. *et al.* Fast Principal-Component Analysis Reveals Convergent Evolution of ADH1B in Europe and East Asia. *Am J Hum Genet* **98**, 456-472 (2016).

20. Glaesmer, H. *et al.* Der Childhood Trauma Screener (CTS) - Entwicklung und Validierung von Schwellenwerten zur Klassifikation. *Psychiatrische Praxis* **40**, 220-226 (2013).

21. De Bellis, M.D. & Zisk, A. The Biological Effects of Childhood Trauma. *Child and Adolescent Psychiatric Clinics of North America* **23**, 185-222 (2014).

22. Initial reliability and validity of a new retrospective measure of child abuse and neglect. *American Journal of Psychiatry* **151**, 1132-1136 (1994).

23. RCoreTeam. *R: A language and environment for statistical computing*, (Vienna, Austria, 2021).

Connecting statement to chapter V

In Chapter IV we concluded that SNPs derived from GWAS do not represent genetic signals that moderate the effect of environment on an outcome, explaining the limitation of PRS derived from GWAS to detect GxE interaction effects. GWEIS would be a much more appropriate approach to investigate GxE effects. However, both signals (GWEIS or GWAS) are potentially limited in terms of representation of biological information, as in both cases the identification of variants is based solely on isolated statistical associations (the relationship between each SNP and the outcome is tested in isolation). In a complex system in biology approach, the structure, function, and complex interactions between the elements of a system through time, including macro-scale objects such as environment and genome are considered[282]. This view is aligned with the notion that genes do not operate in isolation, but rather in networks, with regulatory functions among them, representing associated molecular functions and coding for precise biological functions in specific tissues[283].

For the purposes of identifying underlying biological mechanisms involved in ELA exposure and development of psychiatric and cardiometabolic disorders, we adopted a genomic measure capable of representing biological information while also maintaining a genome-wide perspective. The ePGS technique is a mid ground between genome-wide associated signals and a hypothesis-driven genetic score tailored to represent biological information. Using genome-wide RNA sequencing we identify genes functionally related to with our gene of interest in a specific tissue. This functional relationship is defined by coexpression, since genes that are co-expressed are considered to be part of the same biological process[283]. This information is then

combined with a tissue specific GWAS for gene expression called GTEx. The ePGS scores reflect individual variation in the expression of a network of genes in a specific tissue.

We explored, in Chapter VI, the comparison between ePGS and PRS scores in representing variations in a biological process, as well as their portability across different ancestries in several enrichment analysis experiments. Conclusion from this work will address the question regarding the most suited way to capture GxE interaction effects for the exploration of the biological mechanisms involved in the association between ELA exposure and psychiatric and cardiometabolic comorbidities.

Chapter V. Expression-based polygenic scores - A gene network perspective to capture individual differences in biological processes

Publication information: This manuscript is in submission phase, currently being submitted to the Nature Methods Journal.

Expression-based polygenic scores - A gene network perspective to capture individual differences in biological processes.

Barbara Barth[1,2,3], Euclides José de Mendonça Filho[2,3], Danusa Mar Arcego[2,3], Irina Pokhvisneva [2,3], Michael J. Meaney[2,3,4,5], Patrícia Pelufo Silveira[2,3,4]

[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] Department of Psychiatry, Faculty of Medicine, McGill University, Montreal, QC, Canada.
[5] Translational Neuroscience Program, 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

Acknowledgments

We thank the 1000 Genomes Project for the data availability. Sachin Patel for genetic score calculations. This research was supported by the Canadian Institutes of Health Research (CIHR, PJT-166066 and PJT-173237, PPS), Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN-2018-05063, PPS) and Fonds de recherche du Québec – Santé (FRQS to BB and PPS), awarded for the project: Le rôle de l'expression du réseau de gènes du

transporteur de la dopamine sur le cerveau dans la modulation des réponses aux facteurs environnementaux au début de la vie

Author contributions

BB and PPS designed the study, BB, EJMF, DMA and IP conducted data analysis, BB, EJMF, DMA and IP generated the figures and BB and PPS wrote the manuscript with input from all authors. PPS and MJM supervised the research. All authors read and approved the final manuscript.

Competing interest declaration

The authors declare no competing interests.

Abstract

Incorporating functional aspects to polygenic scores may accelerate early diagnosis and the discovery of therapeutic targets. Yet, existing polygenic scores simply summarize information from genome-wide statistical associations between SNPs and phenotypes. We developed the novel biologically informed, expression-based polygenic scores (ePRS or ePGS). The method characterizes tissue specific gene co-expression networks from genome-wide RNA sequencing data and incorporates this information into polygenic scores. Performance and characteristics of the ePGS were compared to traditional polygenic risk score (PRS). We observed that ePGS differs from PRS for aggregating information on; i. the relation between different genes (co-expression); ii. the levels of tissue-specific gene expression; iii. the genetic variation of the target sample; iv. the tissue-specific effect size of the association between genotyping and gene expression; v. the portability across different ancestries. Variations in the ePGS represent individual variations in the expression of a tissue-specific gene co-expression network, and this methodology may profoundly influence the way we study human disease biology.

Introduction

Genome-wide association studies (GWAS) are used to identify genetic variants statistically associated with a disease or trait[1], by generally comparing cases and controls in thousands of single nucleotide polymorphisms (SNPs) across the genome. GWAS results can be applied in target samples for the calculation of polygenic risk scores (PRS). PRS aggregates the GWAS information by summing the risk alleles count weighted by the effect size for each SNP presented in the GWAS[2,3]. PRS combines the isolated small effects of multiple genetic variants in a single score which represents the genetic risk for a disease or a trait. However, complex diseases are polygenic, involving the function of diverse genes and molecules that interact with each other in cellular networks[4]. Genes do not operate in isolation but conjointly in tissue-specific networks representing associated molecular functions, and code for precise biological functions in specific tissues[5]. These biological intricacies and functional relationships are not captured by traditional polygenic risk scores.

We have created an innovative approach to genomic profiling, informed by biological function, that characterizes gene networks based on the levels of co-expression within a specific tissue[6-17]. A gene network involves a number of genes co-expressed within a specific tissue or brain region that exert a concerted effect on a target biological process. The co-expression-based polygenic score is devised using gene co-expression data in a specific brain region (or other tissue), defining the gene network that then serves as the basis for gene selection for the polygenic score. SNPs from these genes are functionally annotated and subjected to linkage disequilibrium clumping for removal of highly correlated SNPs. We then use a count function of the number of effect alleles at a given SNP, weighted by the effect size of the association between

the individual SNP and gene expression in a specified tissue (GTEx[18]). The sum of these values from the total number of SNPs provides the expression-based polygenic score (ePGS) (**Figure 9, Supplementary Figure 1**) (**Supplementary information - Methods**).

The ePGS combines information on: i. the relation between different genes (co-expression); ii. the levels of tissue-specific gene expression (bulk or single-cell genome-wide RNA-sequencing); iii. the genetic variation of the target sample (genotyping data); iv. the tissue-specific effect size of the association between variants and gene expression (GTEx). Therefore, variations in the ePGS represent individual variations in the expression of the tissue-specific gene co-expression network. Here we present the method of calculating and advantages of ePGSs, demonstrating that the ePGS represents cohesive and tissue-specific gene co-expression networks, and have higher trans-ancestry portability in comparison to traditional polygenic scores (PRS).

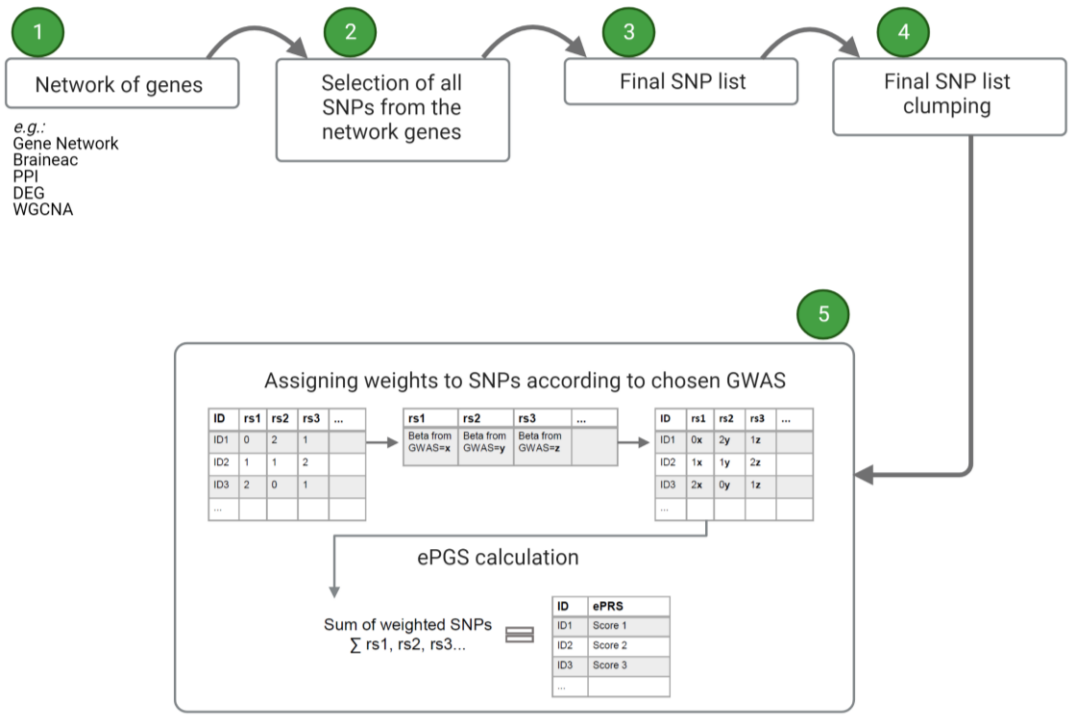


Figure 9. *Schematic figure representing the key steps to calculate the ePGS***.** 1) Construction of a network of genes that is defined by a set of genes that interact in a biologically meaningful way. Some examples are co-expression of transcripts from animal models, as used in the current study, and different expression analysis. Additionally, it can be defined by protein-protein interaction (PPI), co-expression of transcripts from human samples (Braineac) and by weighted gene co-expression network analysis (WGCNA). At this step, tissue specificity can be defined by selecting transcript data from specific tissues of interest. The list of genes can also be filtered by a specific developmental time point, for example, by using publicly available databases such as the BrainSpan[19]. Furthermore, the list of genes can be filtered by other conditions and interests. 2) Selection of all existing SNPs from the gene network was done using biomaRt package. From this list we retained common SNPs with a) SNPs from the study sample genotyping data and b) SNPs present in GTEx (which is the GWAS chosen to weight the selected SNPs in the examples provided here). The common SNPs represent the final SNP list that is subjected to linkage disequilibrium clumping (r^2>0.2). 5) Weight the SNPs: the number of effect alleles (genotype information from the study sample) at a given SNP is multiplied by the effect size of the association between SNPs and the gene expression (GTEx). The sum of all weighted SNPs for each individual corresponds to the individual ePGS.

Results

Expression-based polygenic scores (ePGS) calculation

The expression-based polygenic scores are constructed in several steps (**Figure 9**): i) identification of a gene co-expression network using genome-wide RNA sequencing data. The following are some examples of publicly available RNA sequencing databases in both rodents (e.g. GeneNetwork [20]) and humans (e.g. BrainEAC [21]) which can be used to identify gene networks. Researchers can use both co-expression[6-16] and differential expression[17] data, from publicly available or their own datasets, see **Supplementary Figure 1**. The most common

approach for identification of the gene networks is to anchor the gene network around a specific target. In this case the gene network is composed of the genes co-expressed with the target gene in a specific brain region or tissue. In principal, gene networks can be identified in many ways, and the pipeline described in this work can incorporate any form of gene co-expression network, e.g. obtained from Weighted gene co-expression network analysis (WGCNA[22]). Since the expression of gene networks varies from region to region, the resulting approach informs on the relevance of both the gene network and the brain region or tissue, obtaining gene networks that are tissue-specific. ii) Using biomaRT R package[23,24] genes are aligned to humans homologous genes, and all the existing SNPs from these genes are gathered; iii) this SNP list is combined with a list of SNPs available in the target human study sample, thus generating a final list of common SNPs; iv) the final SNP list is subjected to linkage disequilibrium clumping to inform the removal of highly correlated SNPs (r^2>0.2) across 500kb regions; v) A count function of the number of effect allele at a given SNP is weighted by the effect size of the association between the SNP and a gene expression in a specific tissue (data from GTEx). The sum of these values from the total number of SNPs provides the ePGS score.

In the examples provided here, the expression-based polygenic scores were created considering genes co-expressed with two important regulators of dopamine neurotransmission in the brain, the dopamine transporter gene (SLC6A3) and the dopamine receptor D2 gene (DRD2). We considered co-expression levels with these genes in the prefrontal cortex (PFC), one of the final targets of the dopaminergic mesocorticolimbic pathway (See **Table 3,** please access https://github.com/SilveiraLab/Expression-based-polygenic-scores.-Supplemental-Table-1 to download the table). The scores were calculated according to the protocol previously

proposed by our lab[9,25,26] (see **Figure 9** for schematic representation and **Supplementary Figure 1** for gene co-expression rationale). The calculations were done separately for each gene network of interest. The GeneNetwork (http://genenetwork.org) database was used to generate a co-expression matrix of genes with SLC6A3 or DRD2 genes in the PFC from RNA sequencing data from mice. The genes with an absolute value of co-expression correlation ≥ 0.5 with the gene of interest (SLC6A3 or DRD2) were retained. Using the biomaRt package each mouse gene was converted to a human ortholog and all the existing SNPs were gathered from these genes. Common SNPs were selected between the three sources (the SNPs gathered from the gene networks of interest, the SNPs from the GTEx project[18] data in human PFC, and with the SNPs from the study sample (1000 Genomes Project[27])) and were subjected to linkage disequilibrium clumping (r^2<0.2) within 500kb radius. Next, the number of effect alleles at a given SNP is weighted using the estimated effect of the tissue specific genotype-gene expression association from the GTEx project[18]. We also accounted for the direction of the co-expression of each gene with SLC6A3 or DRD2 by multiplying the weight by -1 in case the expression of a gene was negatively correlated with the expression of the SLC6A3 or DRD2 genes. The sum of the weighted values from all SNPs, divided by the number of SNPs, provided the region-specific ePGS scores. The scores were calculated separately for each ancestry in the 1000 Genomes Project, which includes African (N=661), American (N=347), East Asian (N=504), European (N=503) and South Asian (N=489). Since the majority of donors in the GTEx project were of European ancestry[28] (see donor information at: https://gtexportal.org/home/tissueSummaryPage), analyses that do not intent to show performance of scores across ancestries was done using 1000 Genomes Project European

sample, for both ePGS and PRS scores (see Supplement). The SLC6A3 network for European ancestry included 262 genes and 15387 SNPs, and the DRD2 network for European ancestry had 281 genes and 12595 SNPs.

Table 3. SNPs and genes included in ePRS and PRS scores. Detail description of genes and SNPs included in all scores described in the study. Please access https://github.com/SilveiraLab/Expression-based-polygenic-scores.-Supplemental-Table-1 to download the table.)

ePGSs reflect cohesive, biologically meaningful gene networks

We then compared the gene network structure represented by same size ePGS and PRS. To achieve that, we mined gene co-expression information from GeneMANIA[29,30] (http://genemania.org) to identify and quantify connections between the genes from each score. We also used the Centiscape tool[31] in Cytoscape®[32], to estimate two centrality measures of the networks: degree, which is the number of connections between each node (each gene) and betweenesss, that estimates the number of times a node lies on the shortest path between other nodes. **Figure 10a** depicts the gene network for SLC6A3 PFC ePGS (number of genes = 262), with a dense connection pattern between genes. Similar-sized PRSs for broad depression resulted in less dense networks, depicted in **Figure 10b** (number of genes = 281). When comparing the total degree between genes in the different scores using a one-way ANOVA, results show that the SLC6A3 PFC ePGS derived gene network has significantly more total connections than the broad depression PRS (**Figure 10c**). The same results were found for the

DRD2 PFC ePGS (265 genes, **Supplementary Figure 2a**) and its comparable size broad depression PRS (**Supplementary Figure 2b and 2c**).

Figure 10. *Network visualization comparison of SLC6A3 derived ePGS and comparable size PRSes*. **a.** SLC6A3 PFC ePGS gene network; **b.** Broad depression PRS gene network comparable size with SLC6A3 PFC ePGS; **c.** One-way ANOVA of total connectivity (total degree values) for ePGS and PRS comparable size. Gene co-expression interactions were obtained from GeneMANIA (http://genemania.org) and used to generate the networks with Cytoscape® application, which specifies amount of interactions between pairs of genes based on their co-expression, represented by the number of edges (gray lines) in the networks. The Centiscape plug-in in Cytoscape® was used to calculate the centrality of the genes in each network, defining the degree (number of connections with other nodes, represented by node size, in which bigger nodes indicates more connections with other nodes) and betweenness (number of times a node lies on the shortest path between other nodes, represented by node's color in which darker colors indicate higher betweenness in the networks) for the components of the networks.

It is important to highlight main conceptual differences between ePGS and PRS that can explain dissimilarities in total connectivity. PRS are built selecting SNPs from a GWAS based on their genome-wide significance level, and for that reason both intron and exon DNA sequences are considered. Introns are non-coding DNA sequences within the genome, and therefore are not mapped to genes. Introns embody 25% of the human genome and are 4 to 5 times the size of exons[33]. In fact, a large number of significant SNPs from GWAS are in intronic and intergenic regions[34,35]. On the other hand, the ePGS is built from gene co-expression information, and therefore considers only protein-coding DNA sequences, the exons, resulting in every SNPs being mapped to a gene. Moreover, the significantly lower number of connections between genes in PRS compared to ePGS suggests that the traditional PRS method has a less consistent relationship to a biological process. ePGS, on the other hand, maps into a cohesive and dense group of genes that interact with each other in a cellular network, representing associated molecular functions as described below.

ePGS and PRS represent different biological mechanisms

Because of the differences in SNP selection between ePGS (from a gene co-expression network identified in RNA sequencing data) and PRS (most statistically significant SNPs from a GWAS), it is expected that the two types of scores will differ in the biological mechanisms that they represent. We compared PRS and ePGS enrichment analyses using MetaCore™ (Clarivate Analytics, version 21.4) (https://portal.genego.com) and the function "compare experiments". We identified a significant common gene ontology (GO) term and exported unique elements from each network

that are significantly associated to that GO term (FDR < 0.05) for comparison purposes. Networks were constructed for direct interactions between selected objects and filtered for brain tissue and human species.

It is noteworthy that “neuron differentiation (FDR<0.001)” was a common GO process associated with genes from both PRSs and ePGS genes. However, this finding was due to specific element networks in each score (**Figure 11**). In ePGS, “neuron differentiation” was mapped to elements such as “Nestin”, which is present in neural stem and progenitor cells and directly involved in differentiation process[36]. In PRS, “neuron differentiation” was mapped to elements such as “olfactory receptor” and less connections are seen between elements. Taken together, the findings depicted in **Figure 11** suggest that while both ePGS and PRSs are linked to processes related to neuron projection development, these relationships occur via unique and specific mechanisms. The unique elements related to the ePGS score are richer and more connected, suggesting that variations in the ePGS score represent variation on these specific biological processes. On the other hand, the PRS score is less meaningfully linked to biological mechanisms.

Common Gene Ontology process: neuron differentiation (FDR<0.001)

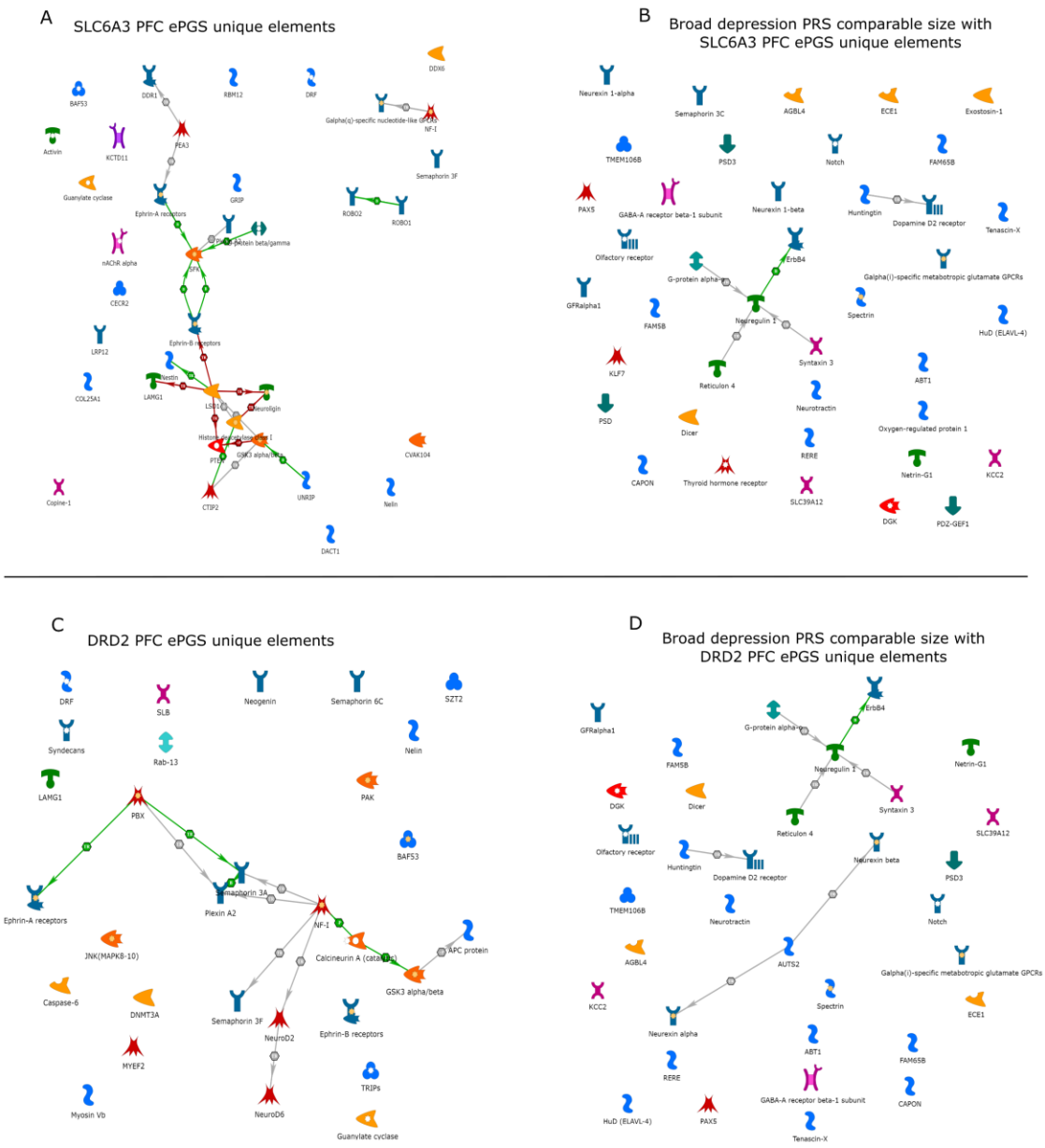


Figure 11. *Unique elements for 'neuron differentiation', a common gene ontology enrichment analysis term for both ePGS and PRS.* Gene ontology (GO) enrichment analysis was performed using Metacore®. The function "compare experiments" was used to obtain common significant (FDR <0.05) GO terms between the gene networks while also identifying the unique elements from each network that are significantly associated to the GO term. Networks were plotted in MetaCore® using the unique elements of each network for the GO enrichment term selected. Figures **a**, **b**, **c**, and **d** show visual comparisons of the different contributions of ePGS and PRS to the GO term. The details of the legends of the network's figures can be found in https://portal.genego.com/legends/MetaCoreQuickReferenceGuide.pdf

ePGS genes represent co-expression networks that are preserved across species

Since our example ePGSs were originally informed by co-expression networks identified in mice (**Supplementary information - Methods**), we aimed at exploring if ePGS genes would also represent co-expression networks in humans, and compare brain co-expression patterns between ePGS genes and traditional PRS genes. To achieve that, we used PFC gene expression data in human post-mortem brain tissue from the BrainSpan database (from embryonic to adulthood, N= 42)[19] and analyzed the correlation between the expression levels on the PFC for the ePGS and PRS gene lists. Our results show that ePGS gene networks have a greater PFC gene co-expression percentage in humans in comparison to PRS gene lists (**Figure 12**). For the SLC6A3 PFC ePGS, 40% of the gene pairs had an absolute expression correlation r>=0.5, and 80% of the correlations were significant at P<0.05. However, when using the genes of a traditional PRS for broad depression, a much lower percentage of co-expression was observed (17% of the gene pairs had an absolute expression correlation r>= 0.5, and only 62% of the correlations were significant at P<0.05). The same comparisons were done for the DRD2 PFC ePGS and its

respective comparable size broad depression PRS, and more robust co-expression patterns were consistently observed in ePGS in comparison to PRSs for broad depression. The results from these examples indicate that ePGSs informed by mice RNA-sequencing data represent brain gene co-expression networks also in humans, and these gene networks are much more tightly connected than those represented by genes that constitute the traditional PRS. This demonstrates a successful cross species translation of genome functional annotation into the ePGS scores.

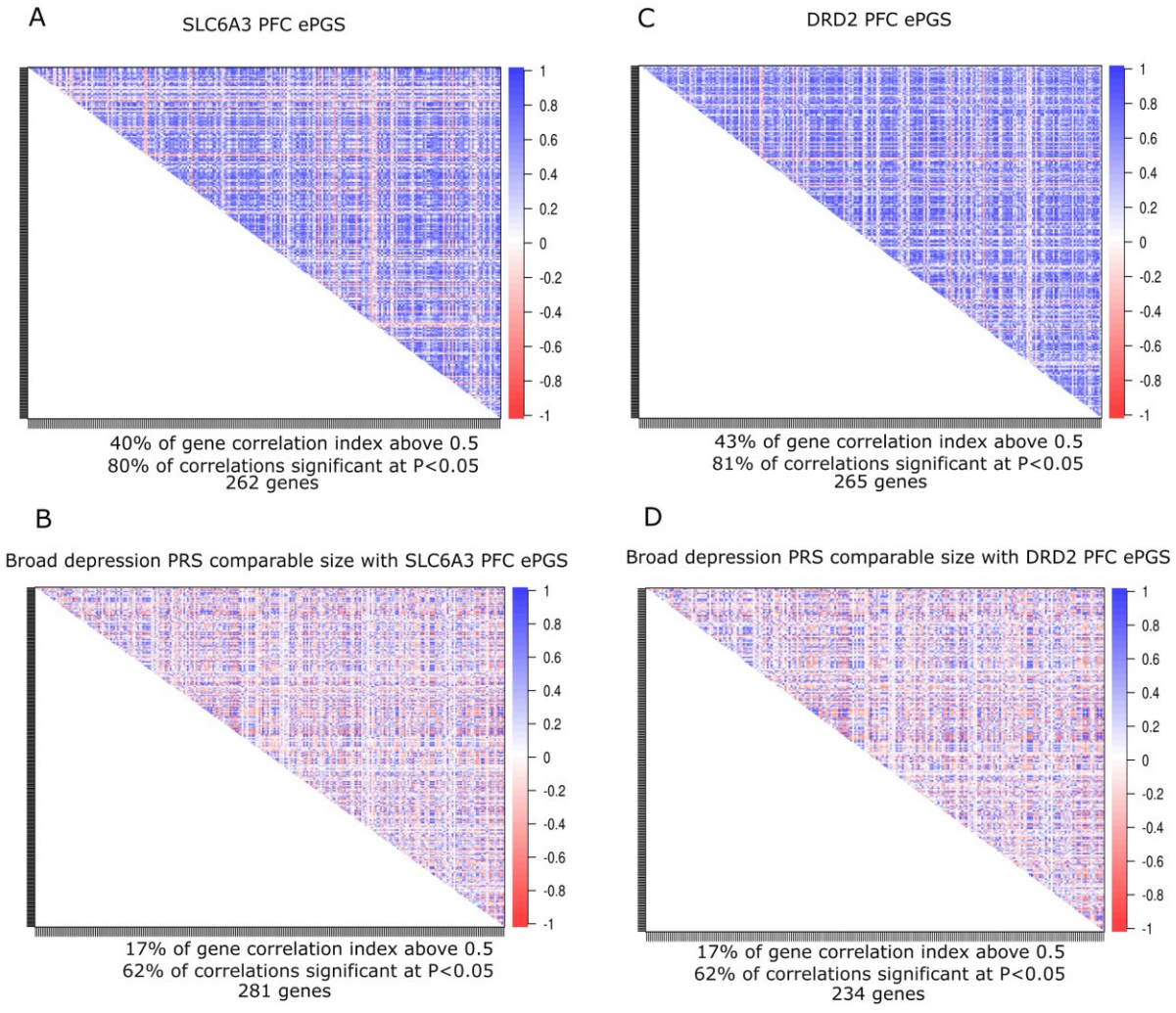


Figure 12. *Correlation matrix of gene expression for ePGS gene networks and PRS gene networks based on BrainSpan human post-mortem brain tissue (from embryonic to adulthood, N=42).* **a.** SLC6A3 PFC ePGS gene network: 40% of the gene correlations was above 0.5 and 80% of the correlations are significant at P<0.05; **b.** Broad depression PRS gene network comparable size with the SLC6A3 PFC ePGS: 17% of the gene correlations above 0.5; 62% correlations significant at P<0.05; **c.** DRD2 PFC ePGS gene network: 43% of gene correlations above 0.5 and 81% of correlations significant at P<0.05; **d.** Broad depression PRS gene network comparable size with the DRD2 PFC ePGS: 17% of the gene correlations above 0.5; 62% correlations significant at P<0.05.

ePGS reflects tissue specific co-expression networks

The ePGS calculation is informed by RNA sequencing data, which quantifies genome-wide tissue-specific gene expression levels (**Supplementary Figure 1**). Therefore, the ePGS considers: a) the gene co-expression data in a specific tissue and uses it to identify the gene network that will serve as the base for the ePGS calculation; b) the tissue specific genotype-gene expression association from GTEx to weight the ePGS SNPs. These features assign layers of information to the ePGS that the PRS does not have, namely a) the relationships between the genes included in the score and b) the degree of association between SNPs and gene expression. PRSs are purely based on the genotype information, which is the same across different cells and tissue types. While both PRS and ePGS summarize the small effects of multiple SNPs using the genotype information, the use of tissue-specific gene expression data in the ePGS technique transforms the polygenic score into a functional genomic tissue-specific measure.

To exemplify this characteristic of the ePGS, we compared two gene networks built on the same gene as the initial anchor, SLC6A3, in two different brain regions: PFC and striatum.

Please note the differences in visualization of the SLC6A3 PFC (total number of genes = 262) and SLC6A3 Striatum (total number of genes = 346) networks (**Figure 10a** and **Supplementary Figure 3a** respectively). Both networks are highly connected representing a cohesive gene network. By computing gene overlap we identified 53 genes in common between the networks (**Supplementary Figure 3b**), which represents a small percentage of the total number of genes from both regions (21% for SLC6A3 PFC ePGS and 15% for SLC6A3 Striatum ePGS). This highlights that the networks are indeed tissue/region specific, even when based on the same initial gene as the anchor, which demonstrates the ability of the ePGS to represent tissue-specific information.

ePGS interacts with environmental variation

In a GWAS, the most significant SNPs represent genetic variants that are strongly associated with a condition or trait, independent or despite the environmental variation existent in the GWAS sample. It is therefore no surprise that investigations of gene-environment interactions using polygenic scores derived from GWASs show modest success [37 38 39]. SNPs that moderate the effect of the environment will hardly be significant as main effects in a GWAS, considering the rigorous GWAS-level of statistical significance for main effects.

Intriguingly, studies published by our laboratory have demonstrated an enhanced capacity of the ePGS to identify gene-environment interactions, using multiple measures of environmental quality and exposure in different cohorts. For example, De Lima et al (2022) described that an prefrontal-based leptin receptor ePGS moderated the effect of postnatal adversity on child eating

behaviour[40]. Arcego et al (2023) demonstrated that an ePGS based on macaques' hippocampal RNA sequencing (gene networks responsive to glucocorticoid injection) moderates the impact of early life adversity on mental health outcomes in adults[41]. Dalmaz et al (2021) showed that a network of genes co-expressed with the synaptic protein VAMP1 gene in the PFC moderates the influence of the early environment on cognitive function in children[42]. Miguel et al (2019) found a significant association between history of exposure to perinatal hypoxic ischemic conditions and children's cognitive flexibility, but this was moderated by the PFC SLC6A3 ePGS[26]. More recently, we identified an environmentally responsive (EnvResp) gene network by selecting genes commonly regulated in the ventral dentate gyrus of mice exposed to stressful or supportive environments, and this informed the calculation of a polygenic score in different human cohorts. Interaction between the EnvResp polygenic score and history of early adversity exposure on anxiety/depression outcomes were statistically significant in more than 60,000 human participants at different ages. As adversity increased, there was a higher risk for psychopathology only among individuals with a high EnvResp score. However, no main effect of the score was detected, confirming the ability of the ePGS to identify gene-environment interactions and individual variation in mental health outcomes in response to environmental change[17].

This enhanced capacity of the ePGS to detect situations of gene-environment interplay is likely due to the fact that the gene networks queried for identification of co-expression are highly sensitive or responsive to the environment, as in the examples above. Following our rationale that SNPs that moderate the effect of the environment will hardly be significant as main effects in a GWAS, we hypothesized that SNPs composing the ePGS would be positioned at the bottom of a Manhattan plot of a GWAS, and therefore not statistically significant when

mapped into this existing GWAS. Indeed, **Figure 13** shows a Manhattan plot for the broad depression GWAS[43]. SNPs in green are the ones included in the SLC6A3 PFC ePGS, confirming that the variants included in the ePGS are below the GWAS significance level. This may explain why the ePGS may be more suited to identify GxE interaction effects[44].

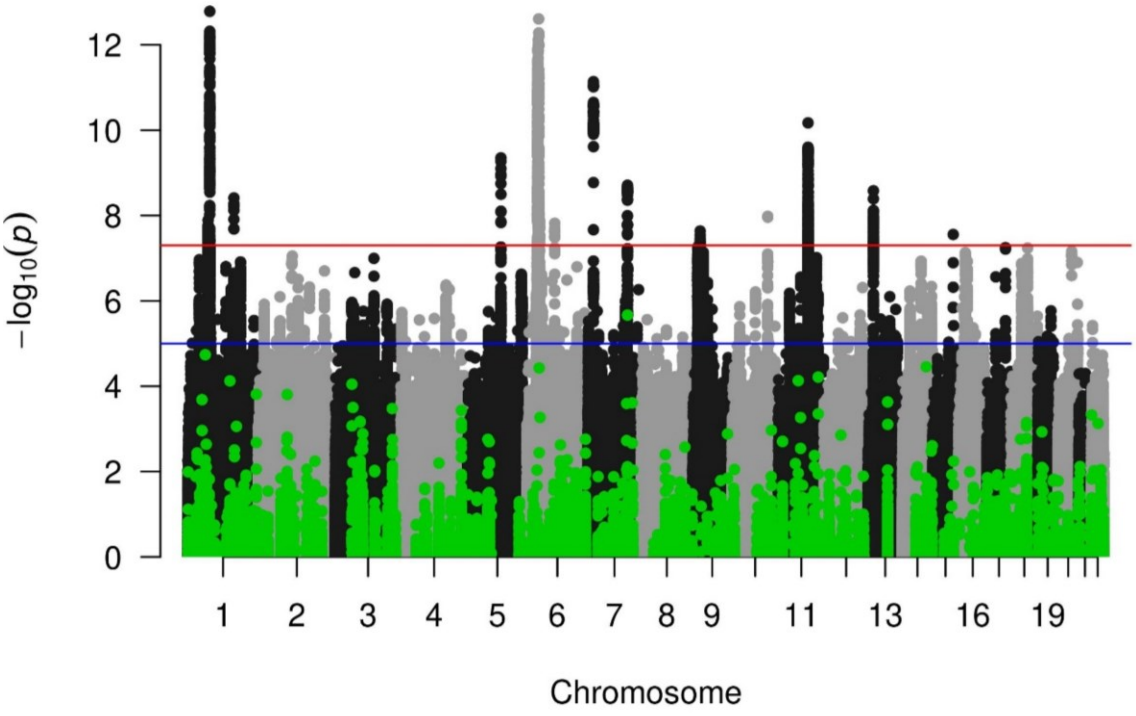


Figure 13. *Manhattan plot for Howard (2019) broad depression GWAS results and SLC6A3 PFC ePGS SNPs.* Gray and black dots represent -log10(p) from the broad depression GWAS. Green dots represent -log10(p) from GTEx for the SNPs included in SLC6A3 PFC ePGS. It demonstrates that all SNPs from the ePGS are not statistically significant at the genome-wide level.

ePGS has high trans-ancestry portability of genetic data

Allele frequency varies across ancestries[45] and the lack of proper diverse populations representation in current genetic association studies hampers the translation of findings into

clinical applications[46]. Efforts are being made to identify genetic variations common and unique to different populations, such as the 1000 Genomes Project that identified novel SNPs[47] and the HapMap consortium[48], but the same level of precision currently available for European ancestry is still not uniformly available for other ancestries[49]. In PRS, the SNP list is derived from the GWAS and the same variants are included in the calculation of the polygenic score in diverse populations, which makes PRS trans-ancestry portability extremely unreliable[46,50,51]. On the other hand, as the ePGS calculation always starts from the gene list, the SNPs included in the same ePGS may differ across ancestries but will still represent the same gene list and the same coexpression network.

The use of genetic scores that perform functional annotation or that consider genes as the first level of information, instead of SNPs, may have advantages for trans-ancestry portability of genetic data[52,53], as is the case of the ePGS method. Indeed, we have seen high trans-ancestry portability and replicability of findings using ePGS in previous studies from our laboratory[9,15-17,26,42]. To illustrate the differences between the traditional PRS and the ePGS in terms of score composition and trans-ancestry portability, we calculated PRSs of comparable size to ePGS (SLC6A3 or DRD2) in the 1000 Genomes Project dataset. The scores were calculated separately for each ancestry in order to account for ancestry-specific allele frequencies and linkage disequilibrium. Ancestries include African, American, East Asian, European and South Asian (**Supplementary information - Methods**). The same number of SNPs present in each ePGS for each ancestry was selected from the most significant variants described in the reference GWAS (broad depression[43]), and subjected to linkage disequilibrium clumping (r^2<0.2) for calculation of PRS separately in each ancestry. Next, the SNPs derived from the calculated PRSs

for each ancestry were assigned to genes, and were compared with ePGSs gene list. **Figure 14** shows the gene overlap between the five different ancestries for each ePGS and their respective comparable size PRS. The ePGS has many more common genes between different ancestries in comparison to PRS scores. These results could explain the good performance in terms of replication seen in studies using the ePGS method[9,15-17,26,42] since ePGS preserves more information (number of genes) across ancestries in comparison to PRS. We also compared the score distribution density across ancestries (**Supplementary Figure 4**). Overall, the ePGS has a greater density overlap between ancestries than the PRS.

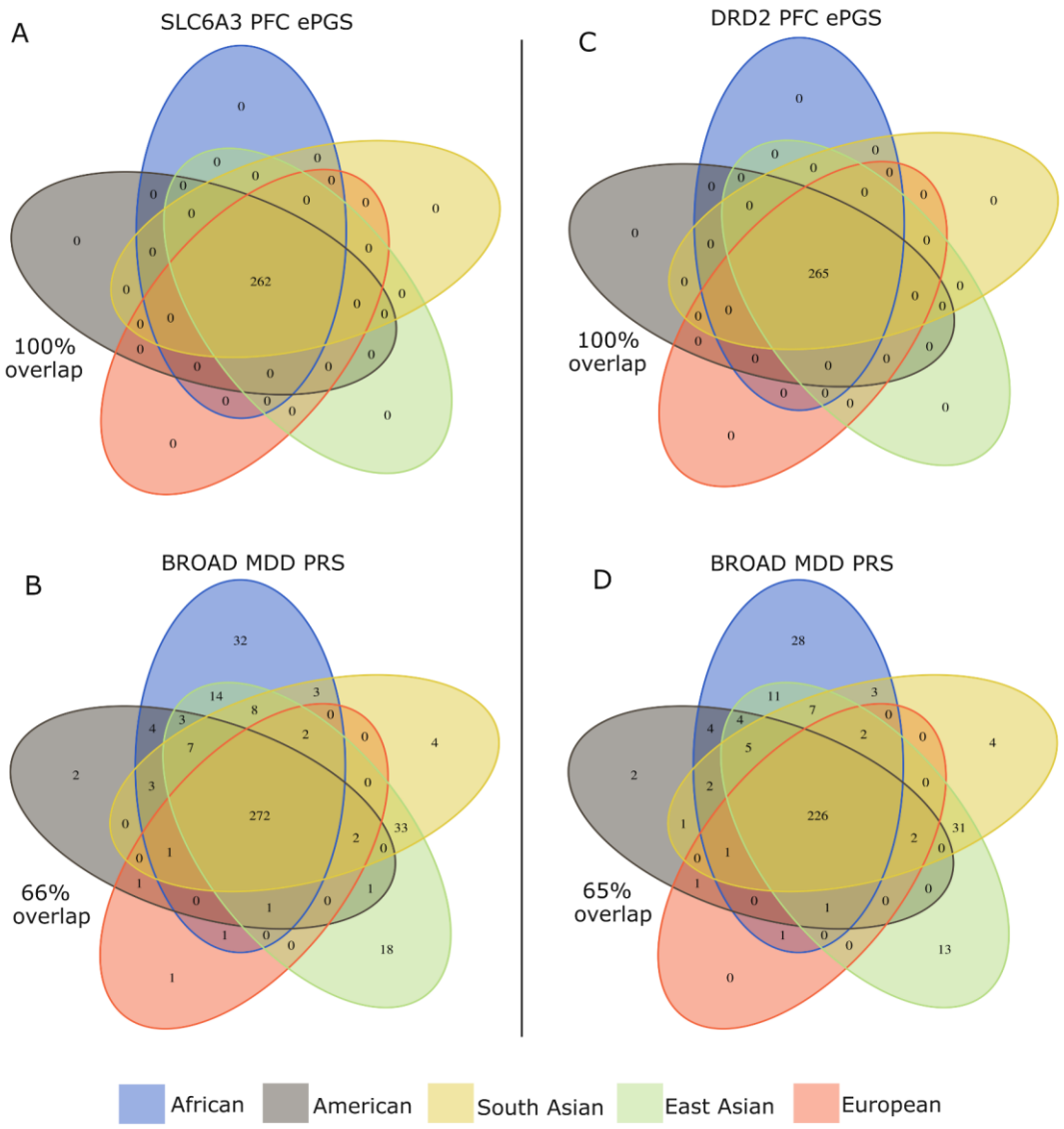


Figure 14. *Venn diagrams of gene overlap for ePGSes and PRSes were calculated based on the ePGS and PRS in the 1000 Genomes Project dataset.* Gene overlaps between the five different ancestries for SLC6A3 and DRD2 ePGS and their respective comparable size PRS. It demonstrates that the ePGS have more common genes between different ancestries in comparison to PRS scores.

Future steps and perspectives in ePGS research

The ePGS calculation is initiated by the definition of a biologically relevant gene network, and this can be done in multiple ways. The examples provided here utilized co-expression data from mice anchored in specific genes for the identification of co-expression networks (SLC6A3 or DRD2). However, other types of data and levels of information can also be used to inform the calculation of ePGS, such as protein-protein interactions, DNA methylation data, or differently expressed gene lists[17]. A promising venue currently being used in our lab consist of utilizing weighted gene correlation network analyses (WGCNA)[54] in RNA sequencing data to identify co-expression gene networks significantly associated with an exposure or condition in controlled animal model experiments or in postmortem human tissue, in a data driven manner, thus completely abandoning the hypothesis-driven approach. This perspective is well aligned with the complex system in biology paradigm, and it is an anticipated improvement of the method. Arcego et al (2023) is a demonstration of this improvement as the authors used WGCNA to identify a hippocampal network of genes responsive to glucocorticoid treatment in macaques and then calculated an ePGS in humans based on this identified gene network[41].

After the selection of the gene network, the list of genes can be filtered by diverse parameters. Adding filters allows the integration of additional information such as the developmental period, by filtering the gene selection for genes upregulated during a certain stage using Brainspan[9,19,55]. Chromosome conformation information can also be added[56], by using data from high-throughput sequencing (Hi-C) and assigning noncoding SNPs to their cognate genes based on long-range interactions using H-MAGMA[57] input files that describe gene–SNP pairs based on brain Hi-C data[58]. FIMO[59] can also be used to include variants affecting transcription factor binding motifs from the genes of the network. Finally, candidate regulatory

variants can be added by mapping available SNPs on promoter regions (up to 4kb upstream of the transcription start site) of the genes that compose the network. Lastly, the weight attributed to each SNP in the ePGS calculation can be derived from different GWASs. In the current examples, a GWAS for gene expression (GTEx[18]) was used, thus reflecting individual variations in gene expression of the network in the specific brain region. All these parameters can be accommodated to contemplate different research questions. Finally, adaptation of the ePGS technique for the use of single-cell and spatial transcriptomics will add still increased resolution and specificity to the polygenic scores.

Discussion

Aligned with the idea of incorporating functional genomics information to PRS technology, we have developed the expression-based polygenic score (ePGS). The ePGS reflects the combined biological function of gene networks. Here we demonstrated the consequences of rethinking SNP selection and incorporating other levels of information to polygenic scores, such as gene expression and tissue-specific data. Our ePGS reflects cohesive gene networks in comparison to PRS-based gene lists, demonstrating a higher level of co-expression between the genes in ePGS versus PRS. This difference is mainly explained by ePGS considering only exon DNA sequences and being built from gene co-expression information. It is important to highlight that since genes do not work in isolation, but rather in networks[5], the use of a gene network perspective has the potential to better reflect biological functions associated with these genes. We demonstrated that the ePGS and PRS reflect different biological processes, when comparing unique elements that are related to a common gene ontology term. The ePGS unique elements are richer and more connected in comparison to PRS unique elements, suggesting that

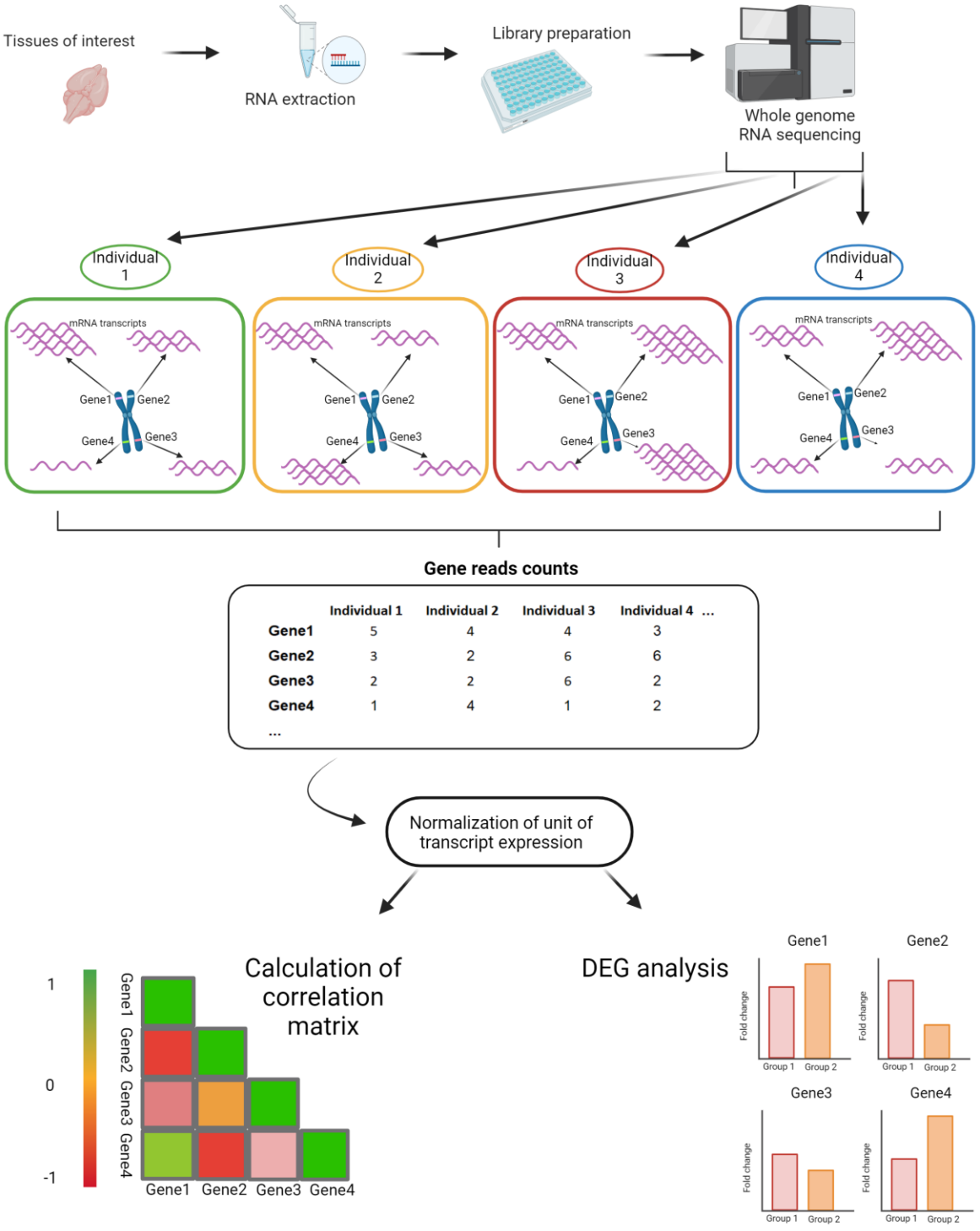
variations in the ePGS score will represent variation on a specific biological process. We also demonstrated that ePGS-based gene networks represent tissue-specific co-expression networks in humans, confirming the biological relevance of our ePGS. The possibility of reflecting functional genomics information in a tissue-specific manner is one of the strengths of the ePGS, demonstrated here by the uniqueness of the SLC6A3 PFC gene network in comparison to the SLC6A3 Striatum gene network. As a consequence of these above-mentioned features, the ePGS is suited to test gene by environment effects, evidenced by previous papers published by our lab [9,16,26,41,42]. The content of ePGS on different ancestries seems consistent when comparing the ePGS and PRS score gene overlap. This is expected since the use of genome functional annotation has the power to improve the prediction of complex traits within and between ancestries[60] and the incorporation of functional markers, such as gene expression, improves trans-ancestry portability of genomic data[61]. The ePGS uses genome functional annotation in two steps of its calculation; in the co-expression basis and by weighing the SNPs using GTEx genotype-gene expression association.

An advantage of using a gene network approach like the ePGS is the possibility of integrating other data modalities also represented by networks or with high dimensionality. For example, the integration of genetic and neuroimage information by parallel independent component analysis (pICA), which 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 [62]. Studies using pICA and the ePGS have found interesting results linking both data modalities and informing on the neuroanatomical basis of the effects of variations in the gene network expression. [9,26,42,63]

In conclusion, the ePGS method is purely based on biological, co-expression data and no information on association with outcomes of interest (e.g. GWAS for diseases) is used. When compared to conventional PRSs, ePGSs represent more cohesive and tissue-specific biological processes. The ePGS performs much better than PRS in gene by environment interaction models and across ancestries, suggesting that our method is indeed superior in capturing individual biological variation in response to environmental changes [7,17], and may profoundly influence the way we study human disease biology.

Supplementary Information

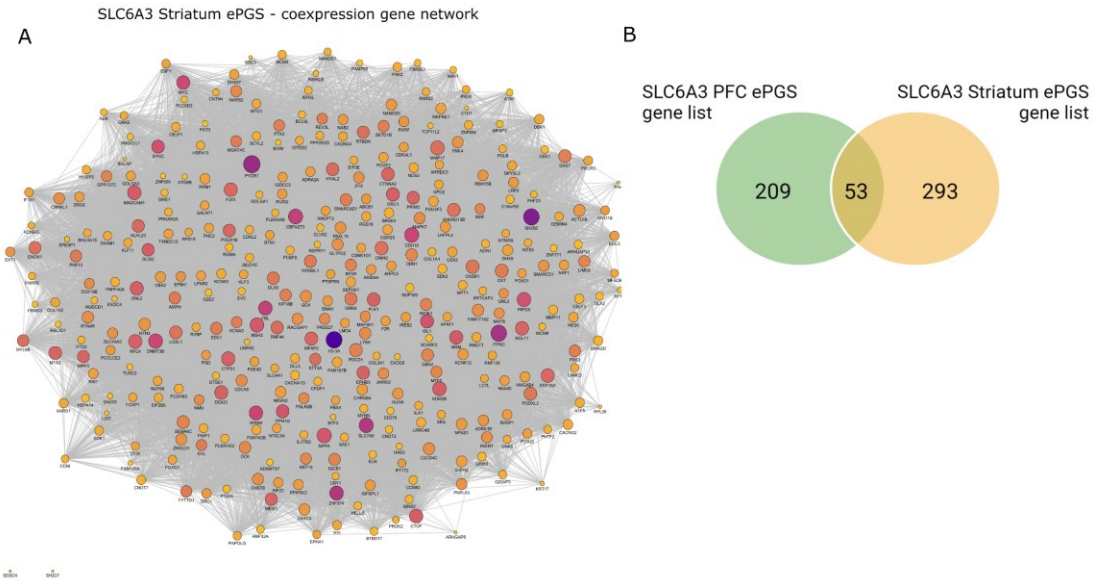
Supplementary Figures



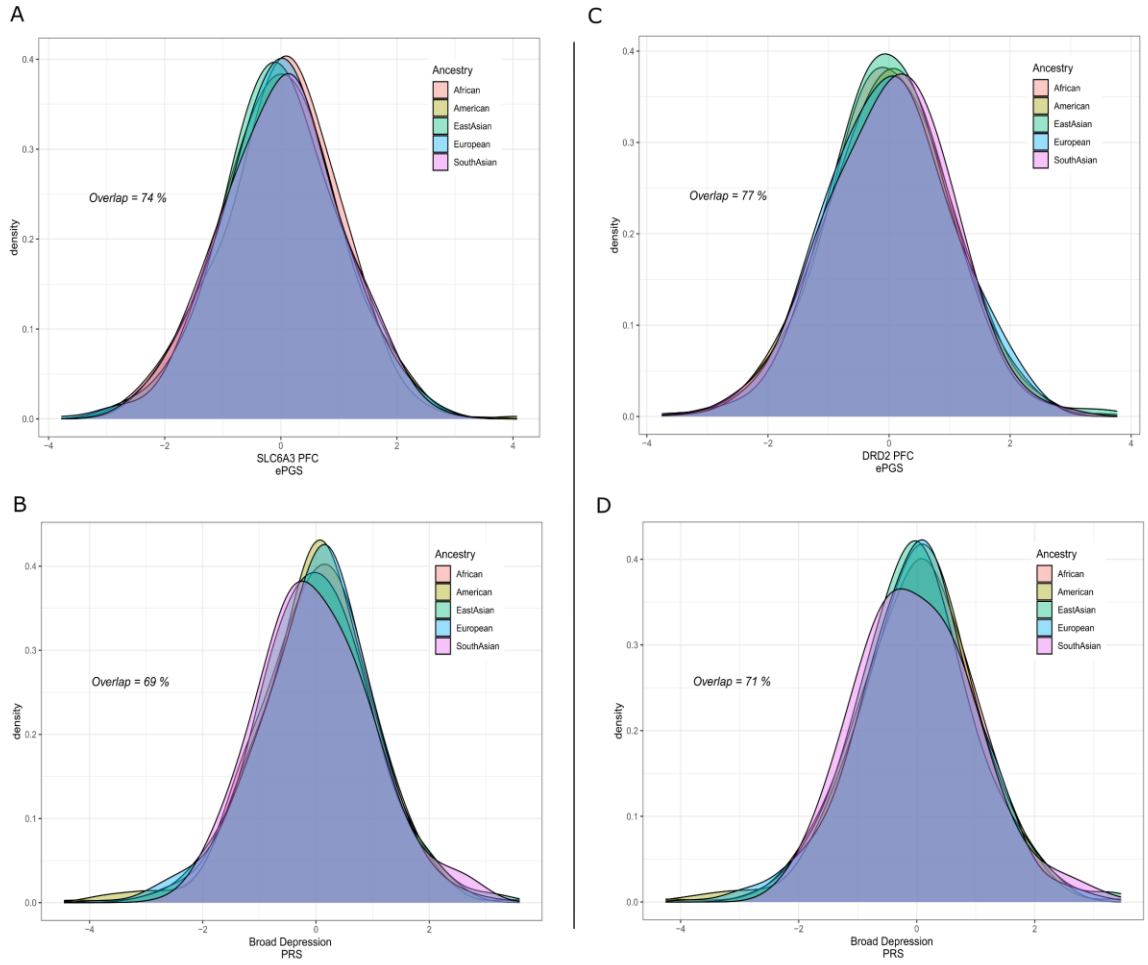
Supplementary Figure 1. *Schematic overview of RNA sequencing process: from tissue collection to data analysis***.** Upper panel - a tissue of interest in chosen and collected, RNA extraction, library preparation and then whole genome RNA sequencing are performed. The process of whole genome RNA sequencing consists of identifying the amount of mRNA in each mapped gene, in each chromosome, in the whole genome. As the result, the gene read counts are obtained for each sample and for all genes mapped in the whole genome sequencing. These units of transcript expression are then normalized, by transforming in reads per kilo base per million mapped reads (RPKM) to account for variations in genes' length. RPKM can be analyzed in several ways. Here are represented two possible analyses. On the right is the gene differential expression analysis which compares expression of a single gene between the groups to identify the genes that are significantly different between the groups in terms of expression (differentially expressed genes, DEG). On the left is a correlation matrix between all genes that were sequenced, which will inform which genes are co-expressed together, meaning that the expression of these genes is varying together, both in a positive or negative way. Co-expression, in the present study, is used as an indicative of how much these genes are working together, thus possibly in the same biological process.


A
DRD2 PFC ePGS - coexpression gene network
B
Broad depression PRS comparable size with DRD2 PFC ePGS
coexpression gene network
C
p<2e-12
Total degree
80
60
40
20
0
Mean=70.91, SD=23.89
N genes=265
Mean=50.81, SD=24.18
N genes=234
DRD2 PFC ePGS
Broad depression PRS comparable size
with DRD2 PFC ePGS

Supplementary Figure 2. *Network visualization comparison of DRD2 derived ePGS and comparable size PRSes.* **a.** DRD2 PFC ePGS gene network; **b.** broad depression PRS gene network comparable size with DRD2 PFC ePGS; **c.** one-way ANOVA results of total connectivity comparison (total degree values) for DRD2 ePGS and PRS comparable size. Gene co-expression interactions were obtained from GeneMANIA (http://genemania.org) and used to generate the networks with Cytoscape® application, which specifies amount of interactions between pairs of genes based on their co-expression, represented by the number of edges (gray lines) in the networks. The Centiscape plug-in in Cytoscape® was used to calculate the centrality of the genes in each network, defining the degree (number of connections with other nodes, represented by node size, in which bigger nodes indicates more connections with other nodes) and betweenness (number of times a node lies on the shortest path between other nodes, represented by node's color in which darker colors indicate higher betweenness in the networks) for the components of the networks.



Supplementary Figure 3. *Visualization for the SLC6A3 Striatum ePGS gene network.* **a.** SLC6A3 Striatum ePGS coexpresion gene network. Gene co-expression interactions were obtained from GeneMANIA (http://genemania.org) and used to generate the networks with Cytoscape® application, which specifies the amount of interactions between pairs of genes based on their co-expression, represented by the number of edges (gray lines) in the networks. The CentiScaPe plug-in[31] in Cytoscape® was used to calculate the centrality of the genes in each network, defining the degree (number of connections with other nodes, represented by node size, in which bigger nodes indicates more connections with other nodes) and betweenness (number of times a node lies on the shortest path between other nodes, represented by node's color, in which darker colors indicate higher betweenness in the networks) for the components of the networks; **b.** Gene overlap between SLC6A3 PFC ePGS gene list and SLC6A3 Striatum ePGS gene list.



Supplementary Figure 4. *ePGS and PRS score distribution across 5 different ancestries.* Scores for each ancestry were z-transformed and distributions were plotted and then compared. **a.** ePGS SLC6A3 PFC score; **b.** Broad depression PRS of comparable size with ePGS SLC6A3 PFC score; **c.** ePGS DRD2 PFC score; and **d.** broad depression PRS of comparable size with ePGS DRD2 PFC score.

Supplementary Methods

1000 Genomes Project genotyping

1000 Genomes Project[27] includes the data for 2504 unrelated participants with ancestry from 26 populations in Africa, Asia, Europe, South Asia and the Americas. All individuals were sequenced using both whole-genome sequencing (mean depth = 7.4×) and targeted exome sequencing (mean depth = 65.7×), and also genotyped using high-density SNP microarrays. A detailed description of genotyping and processing pipeline can be found in [The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 526, 68–74 (2015). https://doi.org/10.1038/nature15393].

PRS calculation – 1000 Genomes Project

We generated polygenic scores using our accelerated pipeline PRSoS (https://github.com/MeaneyLab/PRSoS) [64], for each individual in each ancestry of the 1000 Genomes Project. The PRS were generated using broad depression GWAS[43] considering the same number of top SNPs as in the corresponding ePGS PFC SLC6A3 (12147 SNPs) or ePGS PFC DRD2 (10481 SNPs). We used the PRS for broad depression since dysfunction of the dopamine system is associated with depression symptoms[65]. First, we selected top SNPs from broad GWAS (12147 or 10481). After selection, SNPs were subjected to linkage disequilibrium clumping (r^2>0.2 within 500 kb radius from an index variant) separately for each ancestry. Lastly, the PRS was calculated in each ancestry as a weighted sum of the SNPs remained after the clumping procedure. The comparisons then were made between the ePGS and the PRS for broad

depression of comparable size to corresponding ePGS: 1) ePGS PFC SLC6A3 was compared to broad depression PRS that was calculated based on top 12147 SNPs; 2) ePGS PFC DRD2 was compared to broad depression PRS calculated based on top 10481 SNPs. The number of SNPs to select in GWAS was chosen based on the number of SNPs included in ePGS scores (SLC6A3 and DRD2) calculated for all ancestries (**Table 3,** Please access https://github.com/SilveiraLab/Expression-based-polygenic-scores.-Supplemental-Table-1 to download the table.).

ePGS and PRS networks properties

The topological network structures for the genes that composed the SLC6A3 PFC ePGS, DRD2 PFC ePGS, PRS MDD comparable size with SLC6A3 PFC ePGS and PRS MDD comparable size with DRD2 PFC ePGS were visualized using Cytoscape® software[66]. To obtain gene by gene co-expression quantification we utilized GeneMANIA app (http://genemania.org)[29], selecting only connections based on co-expression. The GeneMANIA database (Application version: 3.6.0) includes curated information on genes coexpression from published papers. Cytoscape uses this information to plot the networks, giving a visual representation of the quantification of connections between the genes based in coexpression from GeneMANIA. It is important to notice that studies sourced by GeneMANIA are not tissue-specific. Also, the sources used by GeneNetwork and GeneMANIA might not be identical, leading to dissimilarities in gene coexpression. Then, using the Centiscape app, inside Cytoscape, we calculated the total degree centrality measure for each gene, which reflects the number of connections a gene has with

other genes. One-way ANOVA was used to verify if the difference in mean total connectivity between the networks (based on ePGS and PRS) is significant (**Figure 11c** and **Figure 12c**). Centiscape app was also used to calculate another centrality measure, betweenness (number of times a node lies on the shortest path between other nodes). Both centrality measures were incorporated in the network visualization, the degree is represented by node size, in which bigger nodes indicate a higher degree in the network visualization and betweenness is represented by node color, in which darker colors indicate higher betweenness in the networks. A control gene network was constructed using the same steps described above to visually compare the network properties with the ones described in this study in **Figure 11a** and **Figure 12a**. The control gene network was based on a list of genes co-expressed with SLC6A3 gene in the striatum (number of genes = 346).

Biological functions associated with ePGS and PRS networks

To evaluate if the ePGS and PRS share common biological processes, we performed functional enrichment analysis for the genes that compose each network, considering a false discovery rate (FDR) adjusted p-value <0.05 to select significant gene ontology (GO) terms. The genes that compose each network were uploaded into MetaCore® software from Clarivate Analytics (https://portal.genego.com). The function “compare experiments” in MetaCore® was utilized to obtain common and unique enrichment terms between the two gene networks. We selected a significant common GO term present in both SLC6A3 and DRD2 comparisons, the term neuron differentiation. Then we built direct interaction networks inside MetaCore® using the unique

elements of each network for the GO enrichment term selected to visually compare the different contributions of ePGS and PRS to the term.

Assessment of gene expression patterns across human development for the ePGS and PRS associated genes

We tested if the genes from ePGS and PRS have a notable pattern of gene expression in humans, especially for the ePGS since the co-expression data was generated from animal models. For that, we used human post-mortem gene expression samples from BrainSpan (http://brainspan.org) [67] and selected gene expression data from the genes comprising our ePGS or PRS networks. We then analyzed the correlation between expression levels for these genes in the prefrontal cortex (including the ventrolateral prefrontal cortex, orbital frontal cortex, medial prefrontal cortex, dorsolateral prefrontal cortex) from BrainSpan (from embryonic to adulthood, N= 42). Visualization of co-expression correlation matrix for each gene list was computed in R software (https://www.r-project.org) using the "cor" function and plotting it as a correlation matrix. Next, we computed the percentage of pairs of genes with an absolute value of expression correlation higher than 0.5 (considered a high correlation for gene expression) and the correlations significant at p-value <0.05, to numeric see the differences of co-expression across the different gene lists.

Comparison between ancestries

To compare the traditional PRS and the ePGS in terms of scores' composition and trans-ancestry portability we calculated the scores split by ancestry in the 1000 Genomes Project dataset (see Methods sections "Expression-based polygenic scores (ePGS or ePGS) calculation" and "PRS calculation – 1000 Genomes Project" and Results section) and compared the PRS and ePGS for each ancestry group in two ways. 1) SNP included in PRS and ePGS were converted to genes using the biomaRt R package[23,24], and we compared the sets of genes between PRS and ePGS using the VennDiagram R package[68]. The percentage of overlap between the PRS and ePGS distributions was calculated as the number of genes in the overlap divided by the total number of unique genes included in PRS and ePGS. 2) We also compared the score distributions across ancestries. For that, we approximated distributions of the PRS and ePGS with kernel density and calculated the overlap as the proportion of the area that is overlapped to the total area.

References

1 Uffelmann, E. *et al.* Genome-wide association studies. *Nature Reviews Methods Primers* **1**, doi:10.1038/s43586-021-00056-9 (2021).

2 Wray, N. R. *et al.* Research review: Polygenic methods and their application to psychiatric traits. *J Child Psychol Psychiatry* **55**, 1068-1087, doi:10.1111/jcpp.12295 (2014).

3 Dudbridge, F. Power and predictive accuracy of polygenic risk scores. *PLoS genetics* **9**, e1003348, doi:10.1371/journal.pgen.1003348 (2013).

4 Schadt, E. E. Molecular networks as sensors and drivers of common human diseases. *Nature* **461**, 218-223, doi:10.1038/nature08454 (2009).

5 Gaiteri, C., Ding, Y., French, B., Tseng, G. C. & Sibille, E. Beyond modules and hubs: the potential of gene coexpression networks for investigating molecular mechanisms of complex brain disorders. *Genes, Brain, Behavior* **13**, 13-24 (2014).

6 Silveira, P. P. *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. *Dev Psychopathol* **29**, 1601-1617, doi:10.1017/S0954579417001262 (2017).

7 Hari Dass, S. A. *et al.* A biologically-informed polygenic score identifies endophenotypes and clinical conditions associated with the insulin receptor function on specific brain regions. *EBioMedicine* **42**, 188-202, doi:10.1016/j.ebiom.2019.03.051 (2019).

8 Miguel, P. M. *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. *Biol Psychiatry* **86**, 621-630, doi:10.1016/j.biopsych.2019.03.983 (2019).

9 de Lima, R. M. 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* **14**, 198, doi:10.3389/fnins.2020.00198 (2020).

10 Morgunova, A. *et al.* DCC gene network in the prefrontal cortex is associated with total brain volume in childhood. *Journal of psychiatry & neuroscience : JPN* **46**, E154-E163, doi:10.1503/jpn.200081 (2020).

11 Potter-Dickey, A. *et al.* Associations Among Parental Caregiving Quality, Cannabinoid Receptor 1 Expression-Based Polygenic Scores, and Infant-Parent Attachment: Evidence for Differential Genetic Susceptibility? *Frontiers in neuroscience* **15**, 704392, doi:10.3389/fnins.2021.704392 (2021).

12 Dalmaz, C. *et al.* Prefrontal cortex VAMP1 gene network moderates the effect of the early environment on cognitive flexibility in children. *Neurobiol Learn Mem* **185**, 107509, doi:10.1016/j.nlm.2021.107509 (2021).

13 Selenius, J. S. *et al.* The relationship between health-related quality of life and melancholic depressive symptoms is modified by brain insulin receptor gene network. *Sci Rep* **11**, 21588, doi:10.1038/s41598-021-00631-w (2021).

14 de Mendonca Filho, E. J. *et al.* Cognitive Development and Brain Gray Matter Susceptibility to Prenatal Adversities: Moderation by the Prefrontal Cortex Brain-Derived Neurotrophic Factor Gene Co-expression Network. *Frontiers in neuroscience* **15**, 744743, doi:10.3389/fnins.2021.744743 (2021).

15 Restrepo-Lozano, J. M. *et al.* Corticolimbic DCC gene co-expression networks as predictors of impulsivity in children. *Molecular psychiatry* **27**, 2742-2750, doi:10.1038/s41380-022-01533-7 (2022).

16 de Lima, R. M. S., Barth, B., Arcego, D. M., de Mendonça Filho, E. J., Patel, S., Wang, Z., Pokhvisneva, I., Parent, C. Levitan, R.D., Kobor, M.S., Bittencourt, A.P.S.V, Meaney, M.J. Dalmaz, C., Silveira, P.P. Leptin receptor co-expression gene network moderates the effect of early life adversity on eating behavior in children. *Communications Biology* **Accepted** (2022).

17 Latsko, M. S., Wang, Z., Zhang, T.Y., Parent, C., O'Toole, N., Pokhvisneva I., Kee, M. Z. L., Wen, X., Craig, K., Boyce, W.T., Meaney, M. J., Silveira, P. P. A translational polygenic score of biological sensitivity to context. *Am J Psychiat* **Accepted** (2022).

18 Consortium, G. T. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-585, doi:10.1038/ng.2653 (2013).

19 Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199 (2014).

20 Mulligan, M. K., Mozhui, K., Prins, P. & Williams, R. W. GeneNetwork: A Toolbox for Systems Genetics. *Methods Mol Biol* **1488**, 75-120, doi:10.1007/978-1-4939-6427-7_4 (2017).

21 Ramasamy, A. *et al.* Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci* **17**, 1418-1428, doi:10.1038/nn.3801 (2014).

22 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559, doi:10.1186/1471-2105-9-559 (2008).

23 Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature protocols* **4**, 1184 (2009).

24 Durinck, S. *et al.* BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**, 3439-3440, doi:10.1093/bioinformatics/bti525 (2005).

25 Silveira, P. P. *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* **29**, 1601-1617 (2017).

26 Miguel, P. M. *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* **86**, 621-630 (2019).

27 Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74, doi:10.1038/nature15393 (2015).

28 Gay, N. R. *et al.* Impact of admixture and ancestry on eQTL analysis and GWAS colocalization in GTEx. *Genome Biology* **21**, 233, doi:10.1186/s13059-020-02113-0 (2020).

29 Warde-Farley, D. *et al.* The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic acids research* **38**, W214-W220 (2010).

30 Montojo, J. *et al.* GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics* **26**, 2927-2928, doi:10.1093/bioinformatics/btq562 (2010).

31 Scardoni, G., Petterlini, M. & Laudanna, C. Analyzing biological network parameters with CentiScaPe. *Bioinformatics* **25**, 2857-2859, doi:10.1093/bioinformatics/btp517 (2009).

32 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).

33 Sakharkar, M. K., Chow, V. T. & Kangueane, P. Distributions of exons and introns in the human genome. *In silico biology* **4**, 387-393 (2004).

34 Bartonicek, N. *et al.* Intergenic disease-associated regions are abundant in novel transcripts. *Genome Biology* **18**, 241, doi:10.1186/s13059-017-1363-3 (2017).

35 Zhang, F. & Lupski, J. R. Non-coding genetic variants in human disease. *Human molecular genetics* **24**, R102-R110 (2015).

36 Bernal, A. & Arranz, L. Nestin-expressing progenitor cells: function, identity and therapeutic implications. *Cell Mol Life Sci* **75**, 2177-2195, doi:10.1007/s00018-018-2794-z (2018).

37 Mullins, N. *et al.* Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. *Psychol Med* **46**, 759-770, doi:10.1017/S0033291715002172 (2016).

38 Peyrot, W. J. *et al.* Does Childhood Trauma Moderate Polygenic Risk for Depression? A Meta-analysis of 5765 Subjects From the Psychiatric Genomics Consortium. *Biol Psychiatry* **84**, 138-147, doi:10.1016/j.biopsych.2017.09.009 (2018).

39 Trotta, A. *et al.* Interplay between Schizophrenia Polygenic Risk Score and Childhood Adversity in First-Presentation Psychotic Disorder: A Pilot Study. *PloS one* **11**, e0163319, doi:10.1371/journal.pone.0163319 (2016).

40 de Lima, R. M. S. *et al.* Leptin receptor co-expression gene network moderates the effect of early life adversity on eating behavior in children. *Communications Biology* **5**, 1092 (2022).

41 Arcego, D. M. *et al.* A Glucocorticoid-Sensitive Hippocampal Gene Network Moderates the Impact of Early-Life Adversity on Mental Health Outcomes. *Biological Psychiatry*, doi:https://doi.org/10.1016/j.biopsych.2023.06.028 (2023).

42 Dalmaz, C. *et al.* Prefrontal cortex VAMP1 gene network moderates the effect of the early environment on cognitive flexibility in children. *Neurobiology of Learning and Memory*, 107509 (2021).

43 Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience* **22**, 343-352, doi:10.1038/s41593-018-0326-7 (2019).

44 Silveira, P. P. & Meaney, M. J. Examining the biological mechanisms of human mental disorders resulting from gene-environment interdependence using novel functional genomic approaches. *Neurobiol Dis* **178**, 106008, doi:10.1016/j.nbd.2023.106008 (2023).

45 Choudhury, A. *et al.* Population-specific common SNPs reflect demographic histories and highlight regions of genomic plasticity with functional relevance. *BMC genomics* **15**, 1-20 (2014).

46 Sirugo, G., Williams, S. M. & Tishkoff, S. A. The missing diversity in human genetic studies. *Cell* **177**, 26-31 (2019).

47 McVean, G. A. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65, doi:10.1038/nature11632 (2012).

48 Consortium, I. H. A haplotype map of the human genome. *Nature* **437**, 1299 (2005).

49 Fitipaldi, H. & Franks, P. W. Ethnic, gender and other sociodemographic biases in genome-wide association studies for the most burdensome non-communicable diseases: 2005-2022. *Hum Mol Genet* **32**, 520-532, doi:10.1093/hmg/ddac245 (2023).

50 Martin, A. R. *et al.* Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature genetics* **51**, 584-591, doi:10.1038/s41588-019-0379-x (2019).

51 Lam, M. *et al.* Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nature genetics* **51**, 1670-1678, doi:10.1038/s41588-019-0512-x (2019).

52 Liang, Y. *et al.* Polygenic transcriptome risk scores (PTRS) can improve portability of polygenic risk scores across ancestries. *Genome biology* **23**, 23, doi:10.1186/s13059-021-02591-w (2022).

53 Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature genetics* **47**, 1228-1235, doi:10.1038/ng.3404 (2015).

54 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics* **9**, 559 (2008).

55 de Mendonça Filho, E. J., Barth, B., Bandeir, D. R., de Lima, R. M. S., Arcego, D. M. Dalmaz, C., Pokhvisneva, I., Sassi, R. B., Hall, G. B. C., Meaney, M. J., Silveira, P. P. . Cognitive Development and Brain Gray Matter Susceptibility to Prenatal Adversities: Moderation by the Prefrontal Cortex Brain-Derived Neurotrophic Factor Gene Co-expression. *Frontiers in neuroscience*, doi:10.3389/fnins.2021.744743 (2021).

56 Dalmaz, C., Pokhvisneva, I., Wang, Z., Barth, B., Patel, S., de Lima, R.M.S, de Mendonça Filho, E.J., Arcego, D. M., Kobor, M. S., O'Donnell, K. J., Meaney, M. J., Silveira, P. P. . Syntaxin-1A gene network moderates the vulnerability/resilience to early life trauma-induced depressive symptoms in women. (Submitted).

57 Sey, N. Y. A. *et al.* A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. *Nat Neurosci* **23**, 583-593, doi:10.1038/s41593-020-0603-0 (2020).

58 Wang, D. *et al.* Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, doi:10.1126/science.aat8464 (2018).

59 Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME Suite. *Nucleic acids research* **43**, W39-W49, doi:10.1093/nar/gkv416 (2015).

60 Zheng, Z. *et al.* Leveraging functional genomic annotations and genome coverage to improve polygenic prediction of complex traits within and between ancestries. *bioRxiv*, 2022.2010.2012.510418, doi:10.1101/2022.10.12.510418 (2022).

61 Amariuta, T. *et al.* Improving the trans-ancestry portability of polygenic risk scores by prioritizing variants in predicted cell-type-specific regulatory elements. *Nature Genetics* **52**, 1346-1354, doi:10.1038/s41588-020-00740-8 (2020).

62 Pearlson, G. D., Liu, J. & Calhoun, V. D. 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* **6**, 276, doi:10.3389/fgene.2015.00276 (2015).

63 de Mendonça Filho, E. J. *et al.* Cognitive Development and Brain Gray Matter Susceptibility to Prenatal Adversities: Moderation by the Prefrontal Cortex Brain-Derived Neurotrophic Factor Gene Co-expression Network. *Frontiers in Neuroscience* **15**, doi:10.3389/fnins.2021.744743 (2021).

64 Chen, L. M. *et al.* PRS-on-Spark (PRSoS): a novel, efficient and flexible approach for generating polygenic risk scores. *BMC Bioinformatics* **19**, 295, doi:10.1186/s12859-018-2289-9 (2018).

65 Belujon, P. & Grace, A. A. Dopamine System Dysregulation in Major Depressive Disorders. *Int J Neuropsychopharmacol* **20**, 1036-1046, doi:10.1093/ijnp/pyx056 (2017).

66 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **13**, 2498-2504 (2003).

67 Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199-206, doi:10.1038/nature13185 (2014).

68 Chen, H. & Boutros, P. C. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**, 35, doi:10.1186/1471-2105-12-35 (2011).

Connection statement to chapter VI

In Chapter V we suggest that ePGS is better suited to represent meaningful biological information and possibly GxE interaction effects in comparison to traditional PRS. In Chapter VI, we applied the ePGS technique to investigate the main objective of the thesis; to explore the GxE interaction effect involved in the susceptibility to develop psychiatric and cardiometabolic comorbidities and to elucidate possible biological mechanisms underlying these comorbidities.

The hypothesis proposing a common underlying biological mechanism between psychiatric and cardiometabolic disorders is based on the compilation of various pieces of evidence. First, psychiatric and cardiometabolic disorders share a common risk factor: ELA exposure. There is a well-known connection between early rearing conditions and the development of psychiatric and cardiometabolic disorders later in life. This phenomena is aligned with the developmental origins of health and disease (DOHaD) hypothesis, which suggests that perinatal signals affect the individuals' predisposition to specific health outcomes, thus shaping individual differences in the risk for chronic illnesses across the lifespan[284,285]. Second, there is a high incidence of comorbidity between psychiatric and cardiometabolic disorders, and they are suggested to be bidirectionally linked, in that the cardiometabolic disorders increase the risk of psychiatric disorders and vice versa[286,287]. Third, pleiotropy – defined by one gene affecting multiple phenotypes and suggestive of shared genetic architecture[288] – is abundant among many complex disorders[289], including complex psychiatric diseases and metabolic traits [290]. The biological mechanisms linking these two types of disorders are still unknown, but dopaminergic alterations have been suggested as a potential link[261,262].

Aiming at exploring gene expression variation in dopaminergic-related processes and its contribution to the GxE interplay involved in the development and maintenance of psychiatric and cardiometabolic disorders, we identified a network of genes coexpressed with the SLC6A3 gene, an important player in the regulation of DA levels in the synapse. The SLC6A3 gene encodes the dopamine transporter (DAT) protein that mediates sodium- and chloride-dependent dopamine reuptake. DAT is responsible for transporting DA out of the synaptic cleft back into the presynaptic neuron (cytosol). DA reuptake mediated by DAT is the main mechanism of DA clearance from the synapse, thus playing a fundamental role in controlling spatial and temporal dynamics of DA neurotransmission[291]. DAT is found in areas in which there are dopaminergic innervations, including the mesocorticolimbic pathway[292]. The striatum is involved in functions related to decision-making and reward-processing and is part of the corticobasal ganglia circuitry[293]. DA transmission in the striatum has been linked to addiction[293] a disorder marked by altered reward processing. Striatal dopaminergic DAT acts on DA axons to spatially and temporally regulate DA signaling in the striatum[294].

The environmental exposure used in Chapter VI is birth weight since previous evidence from the Silveira lab has shown that IUGR and subsequent LBW are related to altered dopaminergic signaling in the ventral striatum[241,242]. IUGR is known to be associated with metabolic alterations such as higher adult total cholesterol[295] and increased adiposity[296] and the catch-up growth commonly associated with LBW is a risk for cardiovascular disease[297,298]. IUGR and LBW are also implicated in behavioural alterations known to predict psychopathologies, such as increased impulsivity[239] and anxiety[299,300].

Birth weight and a striatal network of genes coexpressed with the DAT1 are therefore interesting factors to investigate GxE processes involved in the risk for psychiatric and cardiometabolic disorders.

Chapter. VI. Striatum DAT1 ePGS moderates the effect of early adversity on the risk for adult psychiatric and cardiometabolic comorbidity

Publication information: This manuscript is in submission phase, currently being submitted to the Communications Biology Journal.

Striatal dopamine gene network moderates the effect of early adversity on the risk for adult psychiatric and cardiometabolic comorbidity

Barbara Barth[1,2,3], Danusa Mar Arcego[2,3], Euclides José de Mendonça Filho[2,3,5] Randriely Merscher Sobreira de Lima[2,3,5], Carine Parent[2], Carla Dalmaz[4], André Krumel Portella[2], Irina Pokhvisneva[2], Michael J Meaney[2,5,6,7], Patricia Pelufo Silveira[2,3,5,6*]

[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 and Health Sciences, McGill University, Montreal, QC, Canada.
[6] Department of Paediatrics, Yong Lin Loon School of Medicine, National University of Singapore, Singapore.

[7]Translational Neuroscience Programme, Singapore Institute for Clinical Sciences, Agency for Science, Technology & Research, Singapore

* Corresponding author:
Email: patricia.silveira@mcgill.ca

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. The UK Medical Research Council and

Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and Patricia Pelufo Silveira and Barbara Barth will serve as guarantors for the contents of this paper. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). This 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). 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).

Abstract

Cardiometabolic and psychiatric disorders often co-exist and share common early life risk factors, such as low birth weight. However, the biological pathways linking early adversity to adult cardiometabolic/psychiatric comorbidity remain unknown. Dopamine (DA) neurotransmission in the striatum is sensitive to early adversity and influences the development of both cardiometabolic and psychiatric diseases. Here we show that a co-expression-based polygenic score (ePGS) reflecting individual variations in the expression of the striatal dopamine transporter gene (*SLC6A3*) network significantly interacts with birth weight to predict psychiatric and cardiometabolic comorbidities in both adults (UK Biobank, N= 225,972) and adolescents (ALSPAC, N= 1188). Decreased birth weight is associated with an increased risk for psychiatric and cardiometabolic comorbidities, but the effect is dependent on a striatal *SLC6A3* ePGS, that reflects individual variation in gene expression of genes coexpressed with the SLC6A3 gene in the striatum. Neuroanatomical analyses revealed that SNPs from the striatum *SLC6A3* ePGS were significantly associated with prefrontal cortex gray matter density, suggesting a neuroanatomical basis for the link between early adversity and psychiatric and cardiometabolic comorbidity. Our study reveals that psychiatric and cardiometabolic diseases share common developmental pathways and underlying neurobiological mechanisms that includes dopamine signaling in the prefrontal cortex.

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

cardiometabolic disorders, which are highly comorbid[4,5], rank amongst the leading global causes of disability-adjusted life years worldwide[6,7]. Prospective studies show a bi-directional relationship between psychiatric and cardiometabolic conditions[8]. 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%[7,9]. The odds for depression also increase with one or more non-psychiatric coexisting chronic conditions, especially coronary artery disease, chronic arthritis, and stroke[10]. Anxiety is also associated with 41% increased risk of developing cardiovascular disease[11]. Among adult patients with schizophrenia, the prevalence of diabetes averages of 15%, which is higher than the 10% prevalence of diabetes in the general population. This association persists even after controlling for obesity and use of antipsychotic drugs[12].

The underlying mechanism for these comorbidities remains unknown, but an emerging explanation is that psychiatric and cardiometabolic disorders share common developmental pathways. For example, low birth weight broadly reflects an unhealthy fetal environment and is considered a prevalent form of early life adversity[13] associated with increased morbimortality throughout the life course[14] and is specifically linked to cardiometabolic[14-18] and psychiatric disorders[19-24]. An obvious question concerns the biological mechanisms that underlie such a developmental trajectory involved in the development of cardiometabolic and psychiatric comorbidities.

The brain dopamine system is highly sensitive to early adversity[25] and proposed as a mechanism underlying developmental pathways to multiple psychiatric and metabolic comorbidities[26,27]. Early life adversity such as fetal growth restriction that leads to low birth weight alters

dopaminergic signaling [28-30]. Dysfunction of dopamine neurotransmission in both the ventral and dorsal striatum associates with depression[31], as well as dysregulated food intake and altered energy homeostasis[29,30,32]. Striatal dopamine signaling also appears to regulate systemic glucose metabolism in humans.[33] The striatum harbors dopaminergic neurons[34] and the striatal dopamine transporter (DAT) is a critical regulator of striatal dopamine release and reuptake.[35] Dopamine signaling is influenced by core metabolic hormones such as leptin and insulin, through their actions on the expression and function of DAT [32,36], which is encoded by the *SLC6A3* (solute carrier family 6 member 3) gene.

Based on the large evidence supporting the relation between metabolism, mental health and striatal dopaminergic neurotransmission, as well as the effects of early adversity on striatal dopamine function, we hypothesized that the striatal *SLC6A3* gene network underlies the association between early life adversity and the comorbidity between psychiatric and cardiometabolic disorders in humans. We therefore aimed to test if individual differences in the function of a striatal *SLC6A3* gene network might moderate the effects of early life adversity on psychiatric and cardiometabolic comorbidities in adults and adolescents. To achieve this, we created a *SLC6A3* striatal co-expression-based polygenic score (striatum *SLC6A3* ePGS) reflecting the genetic capacity for expression of the striatal DAT1 gene network (possibly influencing dopamine signaling) and analyzed the effect of its interaction with birth weight on the comorbidity of psychiatric and cardiometabolic conditions in adults (UK Biobank) and adolescents (Avon Longitudinal Study of Parents and Children, ALSPAC).

Methods

Participants

We used genomic and phenotypic data from two cohorts, one from adults (Uk Biobank), and one from adolescents (Avon Longitudinal Study on Parents and Children, ALSPAC).

Adult cohort: The UK Biobank is a large population-based study from the United Kingdom[37]. Participants, aged 37-73, were recruited between 2006 and 2010 resulting in 502,543 subjects. Detailed description of the inclusion/exclusion criteria for the current analysis and the corresponding sample size at each step can be found in *Supplementary information* (**Supplementary Figure 5**). After all exclusions and inclusion criteria, the number of subjects that remained for the analysis was 225,972 (mean age = 55.22, SD = 8.08) (**Table 4**). We used all the data available for the brain imaging analysis considering the inclusion/exclusion criteria (**Supplementary Figure 5,** N=11,167, mean age = 53.86, SD = 7.39).

Adolescent cohort: To explore our findings in an earlier developmental time point we used data from the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort[38-40]. This is a transgenerational prospective observational cohort that recruited 14, 541 pregnant women residents in Avon County, UK. Additional recruitment (N=913) was performed later during Phases II, III and IV respectively, bringing the total sample size of prospective mother-child dyads to 15,658. For more information on ALSPAC variables, please see http://www.bristol.ac.uk/alspac/researchers/our-data/. Data from the adolescent offspring aged between 15.5 and 17.5 were used in this study. Only subjects with available phenotypic

data of interest, early life adversity measure, in this case birth weight and genotyping data were considered for the analyses (N=1,188) (**Table 5**). Detailed description of the inclusion/exclusion criteria and the corresponding sample size at each step can be found in *Supplementary information* (**Supplementary Figure 6**).

See **Supplementary Information - Methods** for a detailed description of the genotyping procedure for each cohort.

Table 4. Description of the baseline characteristics in UK Biobank sample and associations with striatum *SLC6A3* ePGS.

UK Biobank (n = 225972)				
Characteristics	Mean / N	SD / %	ePGS correlation/mean difference	p_value
Sex - Male	88939	39.358%	0.007	0.109
Birth weight (grams)	3317.492	658.366	0.008	<0.001
Completed full-time education at 14-years of age or younger	1667	1.159%	-0.025	0.308
Age at recruitment (years)	55.218	8.079	0.013	<0.001
Townsend deprivation index at recruitment	-1.482	2.983	-0.027	<0.001
BMI at recruitment	27.265	4.866	-0.001	0.570
ICD10 F10-F19 Mental and behavioural disorders due to psychoactive substance use	9278	4.106%	-0.009	0.363
ICD10 F20-F29 Schizophrenia, schizotypal and delusional disorders	556	0.246%	-0.022	0.613
ICD10 F30-F39 Mood [affective] disorders	556	0.246%	-0.022	0.613
ICD10 F40-F48 Neurotic, stress-related and somatoform disorders	5541	2.452%	-0.015	0.272
ICD10 E11 Non-insulin-dependent diabetes	10269	4.544%	-0.037	<0.001
ICD10 I70 Atherosclerosis	632	0.28%	0.001	0.978
ICD10 I63 Cerebral infarction	1758	0.778%	-0.006	0.809
ICD10 I20-I25 Ischaemic heart diseases	15389	6.81%	-0.002	0.815

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 5. Description of the baseline characteristics in ALSPAC sample and associations with striatum *SLC6A3* ePGS.

ALSPAC (n = 1188)				
Characteristics	Mean / N	SD / %	ePGS correlation/mean difference	p_value
Sex - male	565	47.559%	0.086	0.144
Gestational Age (Weeks)	39.731	1.289	0.007	0.822
Birth weight (grams)	3508.933	450.425	0.034	0.239
SES (Crowding index above 1)	27	2.314%	0.080	0.665
Waist circumference (cm) at 15.5 years of age	76.237	8.262	0.028	0.329
SDQ Total difficulties score at 16.5 years of age	5.477	4.372	-0.047	0.109
CIS-R Depression score at 17.5 years of age	0.283	0.747	0.014	0.640
CIS-R Anxiety score at 17.5 years of age	0.231	0.667	0.021	0.467
HOMA2-IR at 17.5 years of age	0.879	0.6	0.006	0.843
zBMI at 15.5 years of age	0.271	0.989	0.037	0.209

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.

Ethics approval and consent

UK Biobank: Informed consent was obtained from each participant, and the project has been approved by 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. Consenting participants provided baseline information, answered questions, had measurements and biological samples collected.

This research has been conducted using the UK Biobank Resource under application number 41975.

ALSPAC: Participants provided informed written consent to participate in the study. Ethics approval for the study was obtained from the ALSPAC Ethics and Law Committee and the local research ethics committees (a full list of the ethics committees that approved different aspects of the ALSPAC studies is available at http://www.bristol.ac.uk/alspac/researchers/research-ethics/). Consent for biological samples has been collected in accordance with the Human Tissue Act (2004).

Consent for publication was obtained from UK Biobank and ALSPAC management teams. The use of these datasets was locally approved by the Centre intégré universitaire de santé et de services sociaux de l'Ouest-de-l'Île-de-Montréal Research Ethics Board under application number IUSMD-21-73.

Identification of the striatal *SLC6A3* co-expression gene network and ePGS calculation

Figure 19a shows the steps involved in the identification of the gene co-expression networks and the calculation of the ePGS score. The ePGS was calculated considering genes co-expressed with the *SLC6A3* gene in the striatum. As described previously[41-48], we began by using brain region-specific RNA sequencing data from mice available at GeneNetwork (http://genenetwork.org/, HBP Rosen Striatum M430V2 (Apr05) RMA Clean)[49] to identify *SLC6A3* co-expressed genes (absolute value of co-expression correlation with *SLC6A3* gene

greater or equal to r= 0.5). GeneNetwork was used to obtain gene expression from rodents since our previous findings demonstrated multiple effects of early life adversities, especially poor fetal growth, on dopaminergic mesocorticolimbic system in rodents[28-30,50-54]. We then converted *SLC6A3* co-expressed genes to human orthologs by using the biomaRt package[55]. Since we were interested in gene networks that were active during the early developmental period in which adversity occurred and 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 childhood periods (0–60 months after birth) in comparison to adulthood (20–40 years of age). This process resulted in a list of striatal *SLC6A3* co-expressed genes. We then mapped all the existent SNPs in the human ortholog genes comprising the striatum *SLC6A3* gene network using biomaRt package[55] in R and gathered all gene-SNP pairs from the GTEx dataset in human striatum. These lists were merged with the genotyping data from UK Biobank and ALSPAC cohorts, respectively, retaining only common SNPs and subjecting the final SNP lists to linkage disequilibrium clumping ($r2 < 0.2$) within 500kb radious to eliminate redundant SNPs. The process resulted in 1532 independent functional SNPs retained in UK Biobank and 1663 SNPs in ALSPAC. The final score included 67 genes in our discovery sample (UKB) (**Supplementary Table 1**).

To calculate the striatal *SLC6A3* ePGS, 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 SNP on gene expression from the GTEx data[56]. We also accounted for the direction of the co-expression of each gene with *SLC6A3* in the network, by multiplying the weight by -1 in case the expression of a gene was negatively correlated with the expression of the *SLC6A3* gene in

the network – therefore, the higher the score, the higher the expression of the genes that compose the network. The sum of the weighted values from all SNPs for each individual in the cohorts resulted in the region-specific striatal scores. The striatal *SLC6A3* expression-based polygenic score (ePGS) is a continuous measure that reflects variation of gene expression of the genes co-expressed with the *SLC6A3* gene in the striatum.

Comparison between Polygenic risk scores and ePGS

To compare the results obtained with the striatum *SLC6A3* ePGS, we generated traditional polygenic risk scores (PRS) using our accelerated pipeline (https://github.com/MeaneyLab/PRSoS)[57]. A traditional PRS is a cumulative score calculated based on a relevant GWAS that represents a risk for a certain health outcome or trait[58]. The sum of the allele count weighted by the effect size across all SNPs in GWAS at a specified threshold was used to calculate type 2 diabetes[59] and major depression disorder[60] PRSs. The number of SNPs included was defined based on the number of SNPs present in our striatum *SLC6A3* ePGS calculated in the discovery cohort. For MDD PRS we used the GWAS results that were obtained without UK Biobank or 23andMe subjects.

Functional enrichment analysis

Enrichment analysis was performed using MetaCore® software from Clarivate Analytics (https://portal.genego.com) to characterize the putative biological functions associated with the

striatal *SLC6A3* co-expression gene network. Genes that comprise the striatal *SLC6A3* ePGS were used in the analysis and the whole genome was used as a background. The significance was considered for the false discovery rate (FDR) adjusted p-value <0.05. To investigate network centrality measures, co-expression patterns were mined from geneMANIA[61].The gene interactions were then visualized using the Cytoscape® software[62]. The nodes are the elements of a network (genes) and edges are the connection between these elements. Bottleneck genes are defined as those having a high betweenness (the extent to which genes act as 'bridges' between other genes in a network), hub genes are defined as those with a high degree (genes with more connections to other genes). To analyze the topological properties associated with this gene network, the CentiScaPe app in Cytoscape® was used to calculate the degree and betweenness of each gene. 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. A gene that is both bottleneck and hub was considered as a central node of the network[63]. We also mined protein-protein (PPI) network interactions using the STRING database (https://string-db.org)[64] and the striatum *SLC6A3* ePGS genes, with the objective of querying the physical interactions of the genes that compose our genetic score. Although we mapped the mice co-expressed gene list to human orthologs, not necessarily the co-expression features would be recapitulated in humans. In order to confirm if the genes that comprised the striatal *SLC6A3* ePGS are co-expressed in humans and to analyze their patterns of co-expression during different life periods in humans, we used the gene expression data from human postmortem samples from the BrainSpan database[65] (see **Supplementary information - Methods**).

Outcome measures

Adult cohort: Psychiatric disorder diagnosis was defined based on the primary or secondary diagnosis of a mental, mood, schizophrenia and neurotic disorders according to participants hospital inpatient records, coded according to the International Classification of Diseases version 10 (ICD-10)[66] (UK Biobank field 41270; ICD10 codes: F10-F19 Mental and behavioural disorders due to psychoactive substance use, F20-F29 Schizophrenia, schizotypal and delusional disorders, F30-F39 Mood [affective] disorders, F40-F48 Neurotic, stress-related and somatoform disorders). Cardiometabolic disorders diagnosis was defined by the ICD-10 codes from chapter IV Endocrine, nutritional and metabolic diseases and chapter IX Diseases of the circulatory system (UK Biobank fields: 41270; ICD10 codes: E11-Non-insulin-dependent diabetes, I70-Atherosclerosis, I63-Cerebral infarction, I20-I25 Ischaemic heart diseases). The presence of at least one mental disorder diagnosis and at least one cardiometabolic diagnosis was considered a comorbidity case. Comorbidity variable 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[67] (**Supplementary information - Methods**).

Adolescent cohort: No diagnoses for the psychiatric and cardiometabolic disorders noted above were available in ALSPAC. As recommended by the American Academy of Pediatrics (AAP)[68], we defined disease risk in adolescents using continuous measures of 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) (**Supplementary information - Methods**, **Supplementary Table 2**, **Supplementary Figure 7**). We then characterized two groups of children: low and high cardiometabolic and psychiatric comorbidity risk (see Statistical Analysis).

Statistical Analysis

Statistical analysis were performed using R[69]. For the descriptive statistics, the ePGS groups were defined by median split, and a comparison between low and high ePGS groups was done using Student t test for continuous variables and a chi-square test for categorical variables (**Table 4** and **Table 5**). Significance levels for all tests were set at $p < 0.05$.

We performed cluster analysis using the mclust package to construct the comorbidity risk variable in ALSPAC adolescent sample [70]. This algorithm applies a model-based classification and density estimation of the z-standardized variables based on finite Gaussian mixture modelling. The method assumes that predictors can be explained by an underlying latent categorical variable (cluster) that represents distinct profiles within the sample, both in a qualitative and quantitative manner. We defined a priori a cluster size solution of two (lower and higher risk for comorbidity). All predictors were z-transformed and adjusted for sex prior to entering the clustering procedure. Regression analysis was carried out to demonstrate the difference between the two clusters in the means for each variable used in the cluster analysis

(**Supplementary Table 2**). The resulting cluster membership, which represented comorbidity risk, was coded as a binary variable (1= “yes comorbidity” or 0= “no comorbidity”).

The gene by environment (GxE) interaction effect on binary outcomes was explored by logistic regression analysis. Birth weight as a continuous variable was used as a proxy for early life environment exposure in UK Biobank (variable ID20022) and ALSPAC. Early life adversity (E), striatal *SLC6A3* ePGS (G) and the interaction term between them were included in the model as main predictors for both cohorts. UK Biobank analyses were also adjusted by sex, age, the first forty genetic principal components, genotyping array, and assessment center, and ALSPAC analyses were adjusted by sex and the first ten genetic principal components. In case of a significant GxE interaction effect, post hoc simple slope analysis was performed to investigate how the environment effect varies as a function of the genetic background[71]. The directionality of the GxE effect was explored in the UK Biobank, our discovery cohort, using a two-tailed P-value threshold. The directionality of the GxE effect on ALSPAC was anticipated based on the finding from UK Biobank, thus a one-tailed P-value threshold was considered.

The relation between early life adversity, ePGS and gray matter density in UK Biobank was analyzed in a multivariate parallel independent component analysis (pICA). This analysis was applied to identify the effect of early life adversity on the relation between two different data modalities (genetic and gray matter density) in a data-driven manner[72]. This analysis separately estimates the maximum independent components within each data modality while also maximizing the association between modalities using an entropy term based on information theory[72]. Each SNP that composes the striatal *SLC6A3* ePGS weighted by striatal GTEx data (genotype * GTEx striatum 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 genetic principal components (ancestry). The subjects were split in two groups according to the birth weight (low birth weight group: subjects with birth weight <= 2.5kg, n = 953) and a randomly selected group of non-low birth weight individuals (subjects with birth weight > 2.5kg, n = 953, please see https://www.who.int/data/nutrition/nlis/info/low-birth-weight), since there was a large discrepancy between cases and controls sample size within the subsample of individuals with T1 structural brain MRI available. Comparison between low birth weight and randomly selected non-low birth weight individuals on main descriptive variables can be seen on Supplementary information (**Supplementary Table 5**). Comparison between the randomly selected group and the full sample of non-low birth weight individuals with MRI available can be seen on Supplementary information (**Supplementary Table 6**). T1 structural brain MRI pre-processed images were adjusted by age and sex (See **Supplementary information - Methods)**. 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[72] 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[72]. Loading coefficients, which describe the presence of the identified component across subjects[72], were extracted for each component, modality, and subject. The mean subject-specific loading coefficients of these components from low birth weight and non-low birth weight groups were compared using Student's t-test. Talairach coordinates were used to identify the anatomical classification of brain areas included in the

identified MRI component[73]. The significant SNPs (|Z| > 2.5) from the identified genetic component were analyzed using MetaCore®, to identify associated gene ontology processes terms (See **Figure 21a** for graphical representation of pICA analysis).

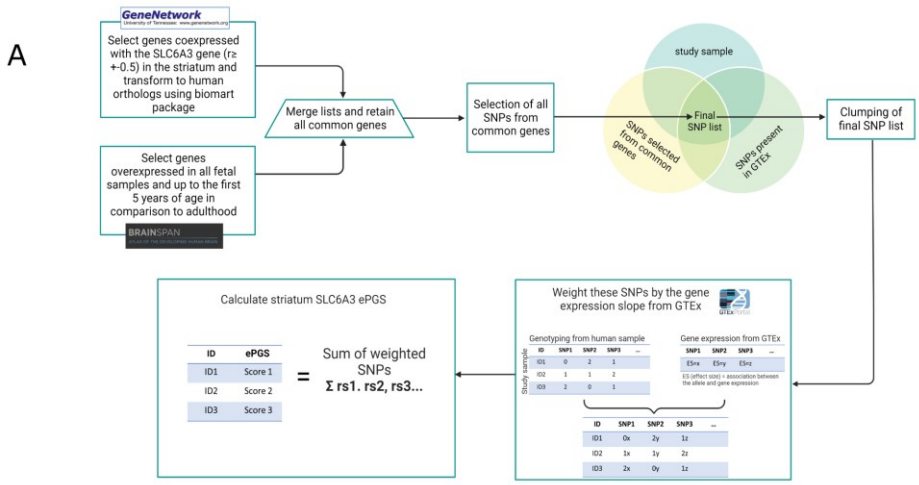
Results

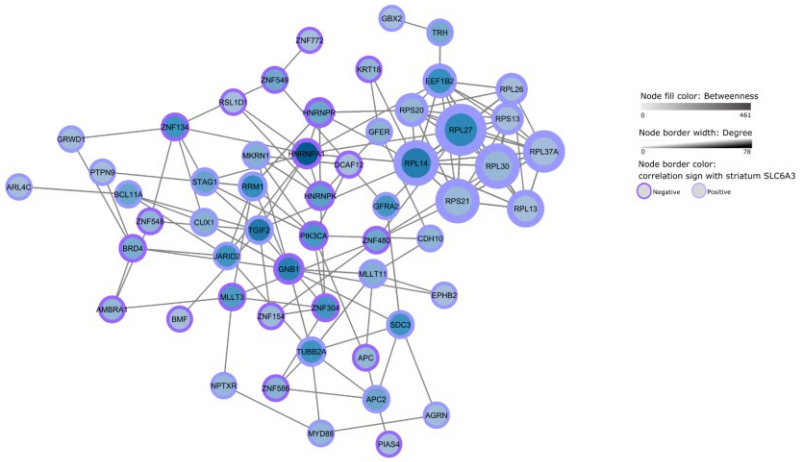
Characteristics of the striatal *SLC6A3* gene network

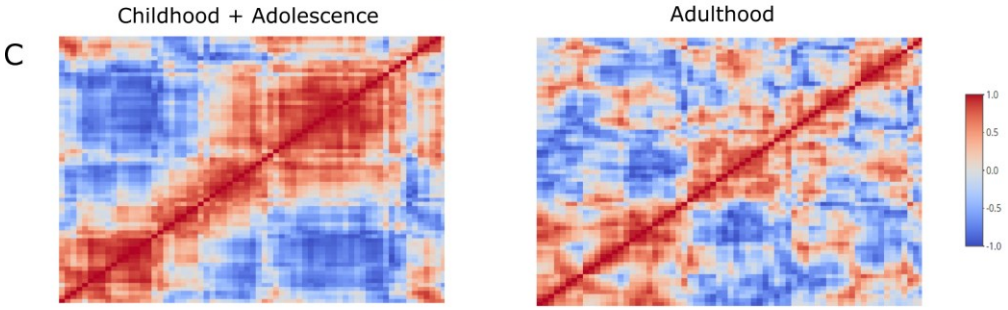
We developed a polygenic score to explore the genetic moderation of early life adversity on psychiatric and cardiometabolic comorbidities focusing on a specific gene network (**Figure 19a and 19b**). We first used brain region-specific RNA sequencing data from mice available at GeneNetwork (http://genenetwork.org/)[49] to identify genes co-expressed with the *SLC6A3* gene in the striatum. These genes were then converted to human orthologs (**Supplementary Table 1** and **Figure 19b**). This list was used to inform the calculation of the expression-based polygenic score (Striatum *SLC6A3* ePGS) in UK Biobank and ALSPAC participants as described in the Methods.

To investigate if mouse-generated *SLC6A3* gene network was co-expressed in humans, we queried the gene co-expression patterns of the striatum *SLC6A3* gene network throughout human development using gene expression data from human postmortem samples[65]. A high co-expression was expected in childhood/adolescence, as the striatum *SLC6A3* gene network was enriched for genes overexpressed in this period of life (see **Figure 19c** and **Supplementary information - Methods**). Prominent gene co-expression clusters were also seen in adults (**Figure 19c**). These findings confirm that the striatal *SLC6A3* gene network, originally from murine data,

is also co-expressed in humans, and that co-expression is observed at different ages. When visualizing and exploring the network properties (**Figure 19d**), we observed that the central gene (hub) and the hub-bottleneck genes are related to ribosomal structure. Among the bottleneck genes, which are important connectors between groups of genes, we observed HNRNPA1, which is involved in the packaging of pre-mRNA into particles and transportation from the nucleus to the cytoplasm, as well as splicing. We also observed SDC3 gene, which plays a role in cell shape organization and has been associated with obesity[74]. Protein-protein interactions of the striatum *SLC6A3* co-expression network mined from STRING revealed that the network has significantly more interactions than expected by chance (P<1.0e-16), suggesting that a significant number of the genes in this co-expression network also have physical interactions at the protein level. The main gene ontology processes terms associated with the network (**Supplementary Figure 8**) include: insulin signaling and response terms (Insulin receptor signaling pathway via phosphatidylinositol 3-kinase; Insulin receptor signaling pathway; Cellular response to insulin stimulus; Response to insulin), ribosome production related terms (Ribosome biogenesis; Ribosomal large subunit biogenesis), dopamine receptor signaling pathway (Adenylate cyclase-activating dopamine receptor signaling pathway) and inflammatory response related terms (Regulation of cytokine production involved in inflammatory response; Negative regulation of cytokine production involved in inflammatory response).


A
GeneNetwork
Select genes coexpressed with the SLC6A3 gene (r= +-0.5) in the striatum 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
Merge lists and retain all common genes
Selection of all SNPs from common genes
study sample
SNPs selected from common genes
Final SNP list
SNPs present in GTEx
Clumping of final SNP list
Weight these SNPs by the gene expression slope from GTEx
Genotyping from human sample
Gene expression from GTEx
Calculate striatum SLC6A3 ePGS
ID
ePGS
ID1
Score 1
ID2
Score 2
ID3
Score 3
=
Sum of weighted SNPs
Σ rs1. rs2, rs3...


B
Node fill color: Betweenness
Node border width: Degree
Node border color:
correlation sign with striatum SLC6A3
Negative
Positive


C
Childhood + Adolescence
Adulthood
1.0
0.5
0.0
-0.5
-1.0

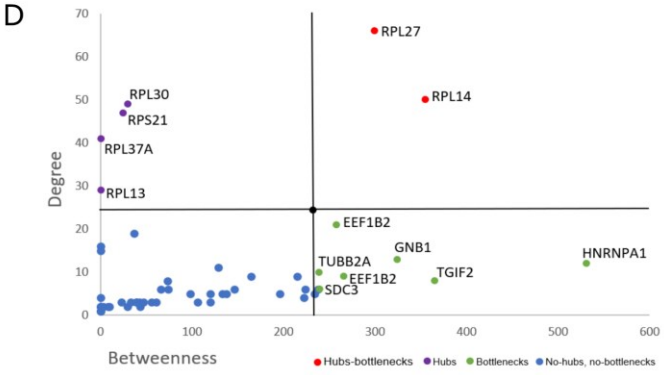

D
Degree
Betweenness
RPL27
RPL14
RPL30
RPS21
RPL37A
RPL13
EEF1B2
GNB1
HNRNPA1
TUBB2A
TGIF2
SDC3
Hubs-bottlenecks
Hubs
Bottlenecks
No-hubs, no-bottlenecks

Figure 19. *Construction and characterization of the striatum SLC6A3 gene network*. **a**. calculation of the expression-based polygenic score (ePGS) from the genes co-expressed with the *SLC6A3* gene in striatum. GeneNetwork was used to generate a list of genes co-expressed with *SLC6A3* in striatum in mice, which were then converted to human orthologs. BrainSpan was used to identify genes overexpressed within striatum in fetal samples and up to 5 years of age in comparison to adult samples. All SNPs from these genes, common between the study sample 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 striatum *SLC6A3* ePGS. **b.** striatum *SLC6A3* ePGS co-expression gene network. Co-expression pattern was mined from GeneMANIA[61]. The color of the node border represents the correlation sign with the *SLC6A3* gene according to GeneNetwork co-expression matrix (dark purple represents negative and light purple positive correlation). Node color intensity represents betweenness (number of times a node acts as a bridge between nodes). Node border width represents the number of connections a node has with other nodes (total degree). **c.** co-expression of genes included in the striatum *SLC6A3* gene network in humans at different ages according to BrainSpan. **d.** topological properties of the striatum *SLC6A3* 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. Hub and hub-bottleneck genes are related to ribosomal structure. Among the bottleneck genes, HNRNPA1 is involved in the packaging of pre-mRNA into particles and transport from the nucleus to the cytoplasm, as well as splicing. The SDC3 gene may play a role in cell shape organization and has been associated with obesity[74].

Striatum *SLC6A3* ePGS moderates the association between birth weight and the risk for psychiatric and cardiometabolic comorbidities in adults

Lower birth weight was associated with the presence of comorbidities in UK Biobank (b= -0.206, Odds ratio (OR)= 0.814, 95% confidence interval (95% CI): 0.781 – 0.847, P <0.001). However, in ALSPAC this association was not significant (b= -0.091, OR= 0.913, 95% CI:0.680 – 1.228, P = 0.548).

For the UK Biobank, there was no significant main association of the ePGS with comorbidity (b= 0.006, OR= 1.006, 95% CI: 0.977 – 1.036, P=0.678). In contrast, and consistent with our anticipated hypothesis, there was a significant interaction effect between the striatum *SLC6A3* ePGS and birth weight on the presence of psychiatric and cardiometabolic comorbidities in UK Biobank adults (b= 0.042, OR = 1.043, 95% CI: 1.001 – 1.086, P = 0.044). The risk for comorbidity increased as birth weight decreased, especially at lower ePGS values (Low ePGS: b = -0.247, OR = 0.781, P< 0.001, 95% CI 0.738 – 0.826; High ePGS: b = -0.166, OR = 0.847, P< 0.001, 95% CI 0.800 – 0.897) (**Figure 20a**). As we considered birth weight as a continuous variable ranging from low to high values and comorbidity as a dichotomous variable, being the presence of comorbidity computed as 1 and absence as 0, the odds ratio represents the negative association between birth weight and the probability of having comorbidity. Results are presented in the low birth weight perspective.

In ALSPAC adolescents the GxE model revealed a significant interaction effect between the striatum *SLC6A3* ePGS and birth weight on the probability of belonging to the high comorbidity risk cluster (b = 0.271, OR= 1.311, 95% CI: 1.015 – Inf, P = 0.041, n=1,188). The risk of belonging to the high comorbidity risk cluster increased as birth weight decreased, especially at lower ePGS values (Low ePGS: b = -0.373, OR = 0.688, 95% CI 0.447 – 1.066, P=0.090; High ePGS: b = 0.176, OR = 1.192, P = 0.419, 95% CI 0.779 – 1.825) (**Figure 20b**). Similar to the findings in adults, there

was no significant main effect association of the striatum *SLC6A3* ePGS on the comorbidity risk (b= 0.086, OR= 1.090, 95% CI: 0.957 – 1.240, P= 0.195). 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 interaction of the striatum *SLC6A3* co-expression gene network and birth weight.

To benchmark our method against the classical polygenic risk score derived from a GWAS, we performed the same GxE interaction analysis using birth weight and PRSs based on GWAS for major depressive disorder[60] and type 2 diabetes[59]. These are phenotypes related to our main outcome, psychiatric and cardiometabolic comorbidity. We found significant main effects of Type 2 diabetes and MDD PRSs on the comorbidity outcome in UK Biobank, but not in ALSPAC (**Supplementary Table 3**) and no significant GxE interaction on comorbidity using these PRSs in the UK Biobank or ALSPAC (**Supplementary Table 4**).

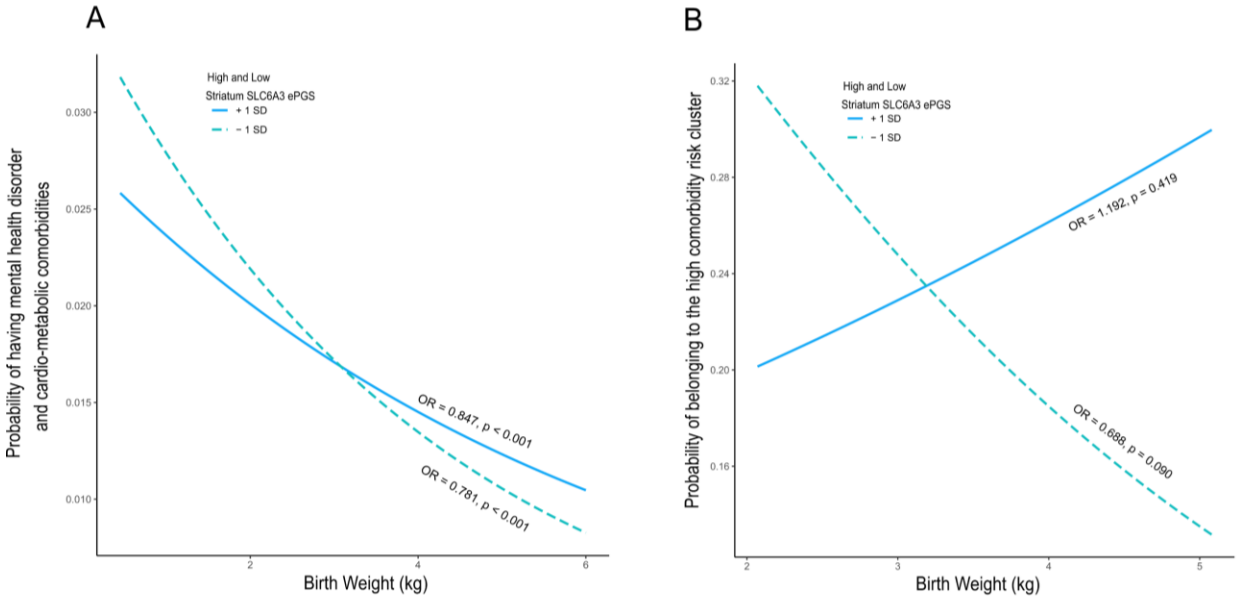


Figure 20. *Striatum SLC6A3 ePGS moderates the effect of early life adversity on the risk for mental health disorder and cardiometabolic comorbidity.* Probability of having comorbidity in individuals with high and low striatum *SLC6A3* ePGS as a function of birth weight. **a.** UK Biobank cohort, N=225,972. The risk for comorbidity increases as birth weight decreases, especially at lower ePGS values (Low ePGS: b = -0.247, OR = 0.781, P< 0.001, 95% CI 0.738 – 0.826; High ePGS: b = -0.166, OR = 0.847, P< 0.001, 95% CI 0.800 – 0.897). **b.** ALSPAC cohort, N= 1,188. The risk for comorbidity increases as birth weight decreases, especially at lower ePGS values (Low ePGS: b = -0.373, OR = 0.688, 95% CI 0.447 – 1.066, P=0.090; High ePGS: b = 0.176, OR = 1.192, P = 0.419, 95% CI 0.779 – 1.825).

SNPs from the striatum *SLC6A3* ePGS are related to gray matter variations in the frontal cortex

We then explored the neuroanatomo-functional relevance of the relation between the striatal *SLC6A3* gene network and early adversity. Functional refers to the variation in gene expression represented by the weight attributed to the SNPs that compose the striatum *SLC6A3* ePGS. We used a multivariate parallel independent component analysis (pICA)[72] (**Figure 21a** and **Supplementary information - Methods**) and investigated correlations between the SNPs from the striatum *SLC6A3* ePGS and voxel-based gray matter density in UK Biobank participants from low birth weight and non-low birth weight groups. This analysis identifies independent components within each data modality separately (SNPs and MRI) while also maximizing the association between these two modalities. The estimated number of components for the MRI modality was 28 and for the genetic modality was 34. Only the most significantly linked pair of components that resulted from the multivariate analysis with a higher correlation index value was selected to be further explored: the pair combining the genetic component 13 and MRI component 18 ($r=-0.201$, $p=6.779e^{-19}$). A statistically significant difference between birth weight

groups was observed for both the genetic component 13 (t=2,214, p=0.026) as well as the MRI component 18 (t=-3,318, p<0.001). These differences between the adversity groups suggest that the relations between data pattern variations (i.e., the relationships between SNPs and gray matter) within this pair of components are significantly different between the two birth weight groups. We then explored the content of genetic component 13 and MRI component 18. The subset of significant SNPs within component 13 is related to variations in gray matter density in the frontal cortex, including the prefrontal cortex, and also more specifically the orbitofrontal cortex, part of the prefrontal cortex, cingulate cortex and temporal cortex (**Figure 21b**). Enrichment analysis of this subset of significant SNPs (**Supplementary Table 7**) using MetaCore® (FDR<0.05) showed that the most significant gene ontology enrichment terms are related to regulation of dendrite development, regulation of neuron remodeling, positive regulation of nervous system development, pyruvate biosynthetic process, ATP metabolic process and response to epinephrine (**Figure 21c**).

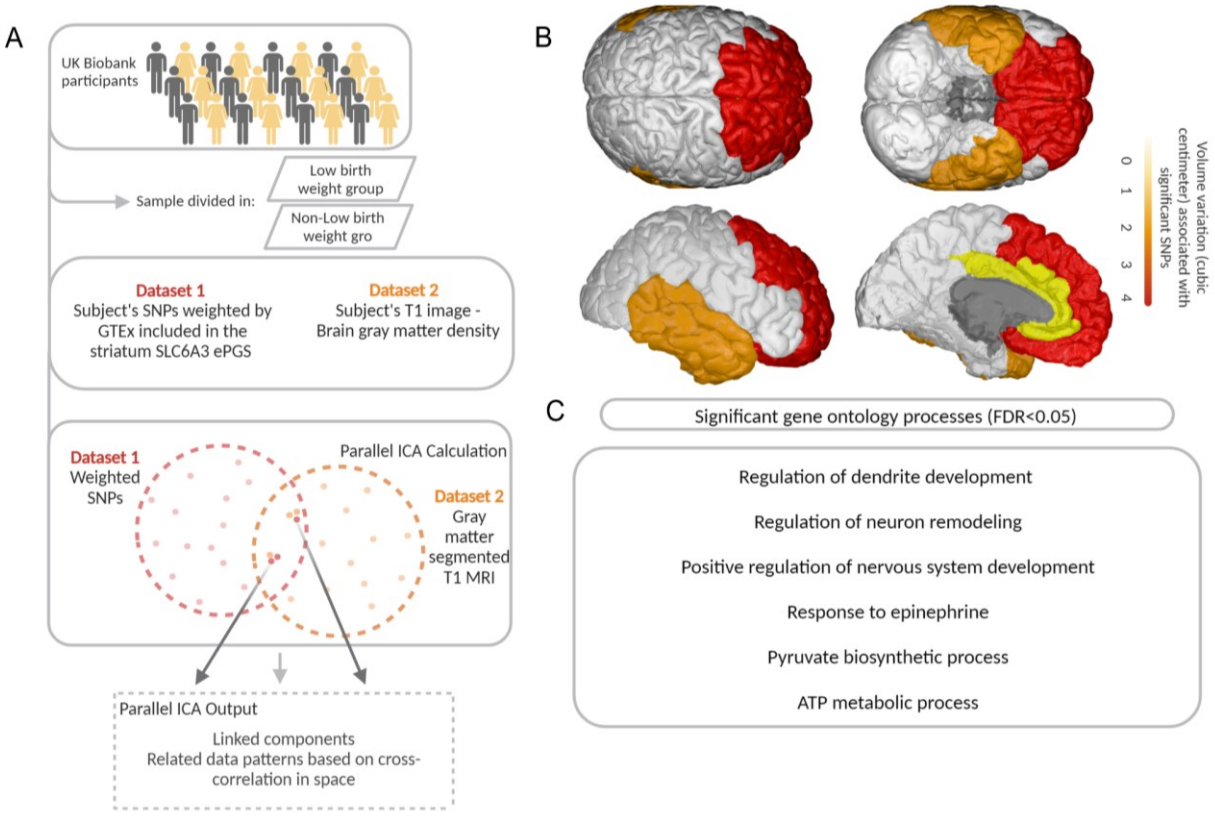

A
UK Biobank participants
Sample divided in:
Low birth weight group
Non-Low birth weight gro
Dataset 1
Subject's SNPs weighted by GTEx included in the striatum SLC6A3 ePGS
Dataset 2
Subject's T1 image - Brain gray matter density
Dataset 1
Weighted SNPs
Parallel ICA Calculation
Dataset 2
Gray matter segmented T1 MRI
Parallel ICA Output
Linked components
Related data patterns based on cross-correlation in space
B
Volume variation (cubic centimeter) associated with significant SNPs
0 1 2 3 4
C
Significant gene ontology processes (FDR<0.05)
Regulation of dendrite development
Regulation of neuron remodeling
Positive regulation of nervous system development
Response to epinephrine
Pyruvate biosynthetic process
ATP metabolic process

Figure 21. *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 striatum *SLC6A3* ePGS and brain features from UK Biobank participants (N=11,167). Participants were separated into low birth weight and normal birth weight groups. 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 striatum *SLC6A3* were frontal cortex, including the orbitofrontal and prefrontal cortex, cingulate cortex and temporal cortex. Color scheme represents the amount of volume variation (cubic centimeter) significantly associated with the subset of SNPs. **c.** summary of significant gene ontology processes related to SNPs from the striatum *SLC6A3* ePGS associated with gray matter.

Discussion

Our study suggests that being born with lower birth weight increases the risk for later comorbidities between cardiometabolic and psychiatric conditions in adulthood. In fact, being born at low birth weight, which reflects prenatal adversity[13], independently associates with increased risk for developing both cardiometabolic[15-18] and psychiatric disorders[19-24] corroborating our findings. Our functional genomics approach provides evidence for the striatal *SLC6A3* co-expression gene network as a salient mechanism moderating this association. This finding is aligned with the critical role of the dopaminergic system in environmental responsivity[42,75].

Although lower birth weight is associated with an increased risk for co-morbidity in both ePGS groups in the UK Biobank, low ePGS participants have significantly more risk than high ePGS individuals. In the low ePGS group in adolescence, there is a suggestion of increased risk of

being part of a high comorbidity risk as birth weight decreases (p=0.09). Although the simple slope for the high ePGS group in adolescence shows a positive inclination between birth weight and risk for comorbidity, this slope is not significant therefore the risk for comorbidity does not vary according to birth weight in the high ePGS group. No information on gestational age was available in the UK Biobank cohort, to maintain consistency, birth weight as a continuous variable and not corrected for gestational age was used in both cohorts. The lack of information about gestational age in our study may be especially affecting the adolescent analysis and this may potentially explain why the simple slope for the low ePRS group does not reach statistical significance, although the interaction between ePRS and birth weight is statistically significant in ALSPAC.

Our enrichment analysis showed that the striatal *SLC6A3* gene network is co-expressed in humans across childhood/adolescence and adulthood (**Figure 19c**), which is aligned with our interaction between striatum *SLC6A3* ePGS and birth weight observed in both adolescents and adults. Our results therefore demonstrate that striatum *SLC6A3* ePGS is able to detect individual differences in response to early adversity at both ages.

Based on the comparisons observed in this study, only the striatum *SLC6A3* ePGS was capable of capturing gene by environment interaction effects, while different PRSs did not significantly interact with early life adversity to predict the main outcome. GWAS-derived PRS reflect main genetic effects and thus are unlikely to capture individual differences in response to environmental variation. Indeed we found significant main effects of the PRSs of Major depressive disorder[60] and Type 2 diabetes [59] on comorbidity in UK Biobank. Overall, these results align with the well-known capacity of PRS to detect main genetic effects, as well as

demonstrate the ability of our ePGS technique in identifying responsivity to environmental change as compared to traditional GWAS-based PRS[76]. PRS main effects were not observed in adolescents from ALSPAC probably due to the specificity of the GWAS to the features of the original discovery sample; the majority of GWASs are generated based on adult samples[77], thus limiting the extrapolation of the effects in different ages.

We also identified putative biological mechanisms underlying the moderating effect of the striatum *SLC6A3* ePGS on the association between early adversity and cardiometabolic–psychiatric comorbidities (**Figure 19d** and **Supplementary Figure 8**). Central genes of the striatum *SLC6A3* network are related to ribosomal structure and the entire gene network is significantly enriched for ribosome production related terms (Ribosome biogenesis; Ribosomal large subunit biogenesis) as seen in our gene-ontology analysis. Ribosome functioning is highly related to cell growth, proliferation, and protein synthesis in all cells. The ribosome biogenesis is a critical process to form mature ribosomes[78]. Dysfunction of ribosomal gene expression has been seen in animals' models of depression[79] and the use of anti-ribosomal P antibodies, that targets phosphorylated protein (P) components of ribosomes, was able to induce depression-like behavior in mice[80]. Interestingly, the same antibody is used to detect systemic lupus erythematosus in humans[80]. This autoimmune disease is related to high levels of systemic inflammation. Our *SLC6A3* gene network is significantly enriched for inflammatory response related terms, especially cytokine production. One of the central genes of our network, *GNB1*, has been implicated in the regulation of inflammasomes, multiple protein complexes responsible for activation of inflammatory responses[81]. As a matter of fact, many patients with

major depressive disorder have elevated levels of inflammatory cytokines [82] and there is evidence linking cardiometabolic syndrome to higher levels of circulating cytokines[83].

Not surprisingly, our *SLC6A3* gene network is also significantly enriched for the dopamine receptor signaling pathway. *HNRNPA1*, a central node of our network, is involved in mRNA transport and synthesis[84]. mRNA axonal transport and protein synthesis at the terminal is an important mechanism for regulation of neurotransmitter synthesis and reuptake[85]. Although *SLC6A3* gene expression occurs at the level of the ventral tegmental area, having *HNRNPA1* as a central node of our network could explain the presence of *SLC6A3* mRNA at the striatal terminal, consistent with numerous human post-mortem studies[86-89]. *SLC6A3* mRNA transport to terminals may be a key mechanistic feature of our *SLC6A3* striatum gene network. It has been shown that *SLC6A3* protein vesicular traffic has a limited contribution to *SLC6A3* concentration in synapses[90], hence other forms of regulation of *SLC6A3* availability in terminals – for instance via mRNA axonal transport and terminal protein synthesis – are likely in place, in agreement to our findings.

Another important gene in the network is the *SDC3*, which may play a role in cell shape organization and has been associated with obesity[74]. Obesity is related to increased risk for cardiovascular disease[91], high levels of inflammation, and insulin resistance[92]. Our striatum *SLC6A3* gene network is significantly enriched for insulin signaling and response related terms. Being born with low birth weight is associated with insulin resistance in children and adolescents[93] and insulin resistance is a risk factor for cardiometabolic and brain-based disorders, including type II diabetes, cardiovascular disease, Alzheimer's disease, and major depressive disorder[94,95]. Metformin, a medication to treat insulin resistance, has shown

beneficial psychotropic effects in psychiatric conditions[96,97]. Evidence shows that insulin has a role in modulating mesocorticolimbic DA neurotransmission through different mechanisms, one of which is increasing DA reuptake by activating the phosphatidylinositol (PI) 3-kinase[36,98]. Insulin also reduces DA release in rodent nucleus accumbens and medial prefrontal cortex slices[99]. Our significant gene by environment results, using birth weight as our environmental proxy, corroborate with the literature showing elevated risk for developing psychiatric and cardiometabolic disorders among individuals born with low birth weight. Our genetic enrichment analysis results indicate that insulin signaling disturbances may be a potential mechanism involved in the interaction effect between birth weight and the striatum *SLC6A3* co-expression gene network on the risk for cardiometabolic and psychiatric comorbidities. This is aligned with many other studies suggesting that altered insulin function is an important mechanism linking early adversity to later disease[100,101].

The subset of SNPs from the striatum *SLC6A3* ePGS that are related to gray matter density variations in our neuroanatomical-functional correlation analysis is associated with regulation of dendrite morphogenesis, neuron remodeling, and positive regulation of nervous system development. These are important processes linked to the prolonged maturation of the mesocorticolimbic dopamine system during the life-course[102], which makes striatal dopaminergic axons especially vulnerable to environmental effects during development[103]. According to these findings, we recently showed that both rodent poor fetal growth and insulin treatment affect the expression of the Netrin-1/DCC axonal guidance cue system, which is involved in the maturation of the mesocorticolimbic DA circuitry[104]. Pyruvate biosynthetic process and ATP metabolic process also emerged as significant enrichment terms for the subset

of significant SNPs related to gray matter density. Both processes have connections with insulin secretion: regulation of insulin secretion in pancreatic β cells is modulated by ATP synthesis and release in mitochondria[105] and by pyruvate transport through mitochondrial pyruvate carriers[106].

The frontal, prefrontal and orbitofrontal cortices were related to the significant subset of SNPs identified by the pICA analysis. This is aligned with evidence demonstrating that resting state functional connectivity between the orbitofrontal cortex and dorsolateral prefrontal cortex is altered in human individuals born small for gestational age, at different ages during development[107].

The cingulate and temporal cortices also emerged as significant brain regions in the pICA analysis. The anterior cingulate has been implicated in affective abnormalities in mood disorders and volume reduction in patients with major depressive disorder[108]. Abnormal posterior cingulate functional connectivity has also been reported in major depression[109]. Temporal lobe alterations are related to insulin resistance pathophysiology in different imaging modalities[110].

Interestingly, the relationship between genetic and MRI components was significantly different between the two early life adversity groups, suggesting that the biological mechanisms represented by the genetic component and the brain regions highlighted by the MRI component are relevant for the effects of early adversity on adult disease.

Taken together, the evidence suggests that ribosomal function, inflammation, and insulin modulation of dopamine function may be underlying mechanisms by which the striatum *SLC6A3* gene network moderates the risk for developing psychiatric and cardiometabolic comorbidities

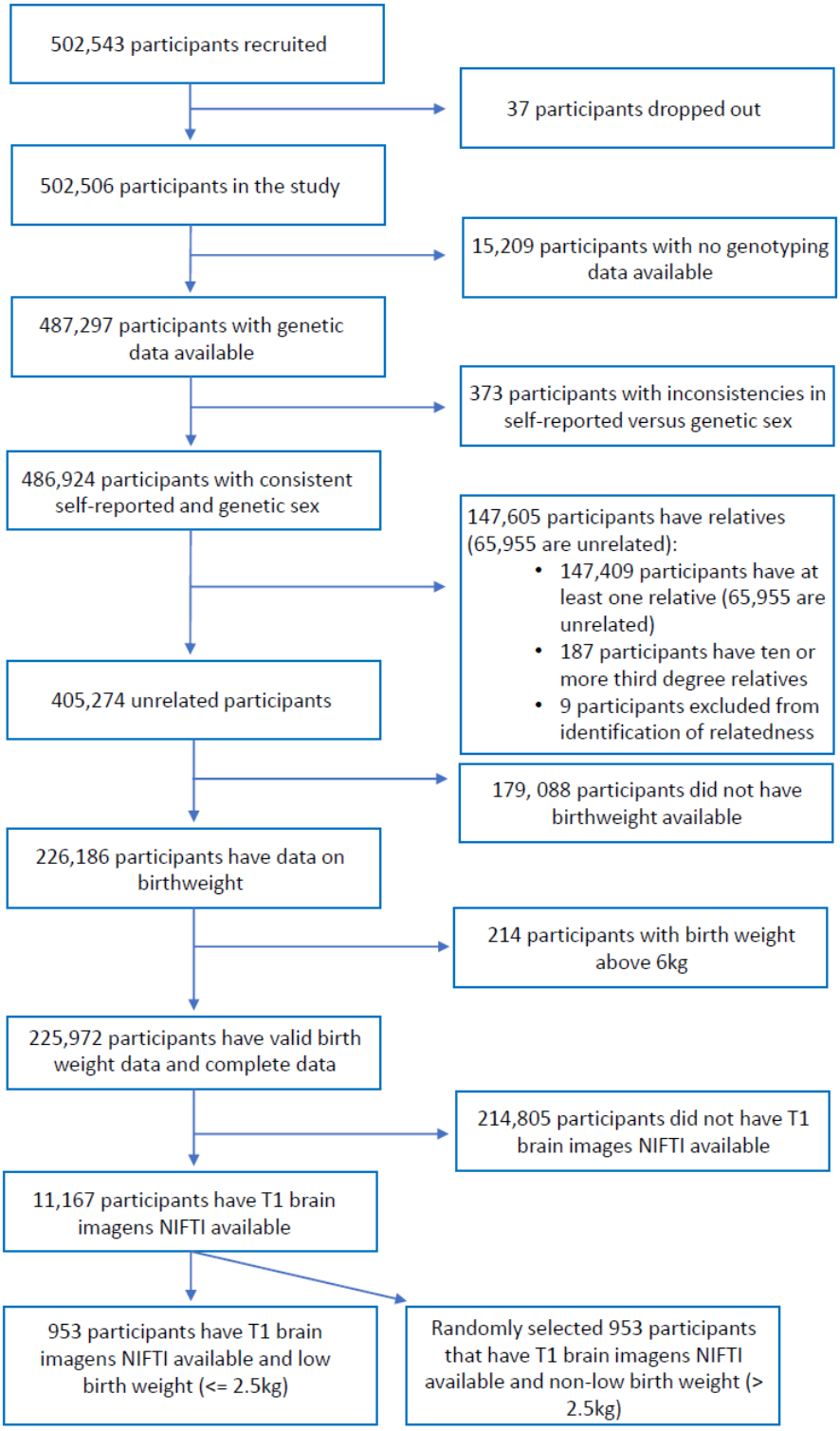
in response to early life adversity. These mechanisms might be especially important in brain areas involving the prefrontal and orbitofrontal cortices, cingulate and temporal cortices.

Our study is limited by the fact that the ePGS do not consider intronic regions, potentially ignoring other important regulatory elements. Moreover, our developmental results are based on cross-sectional studies, and further longitudinal data are needed to better describe this trajectory.

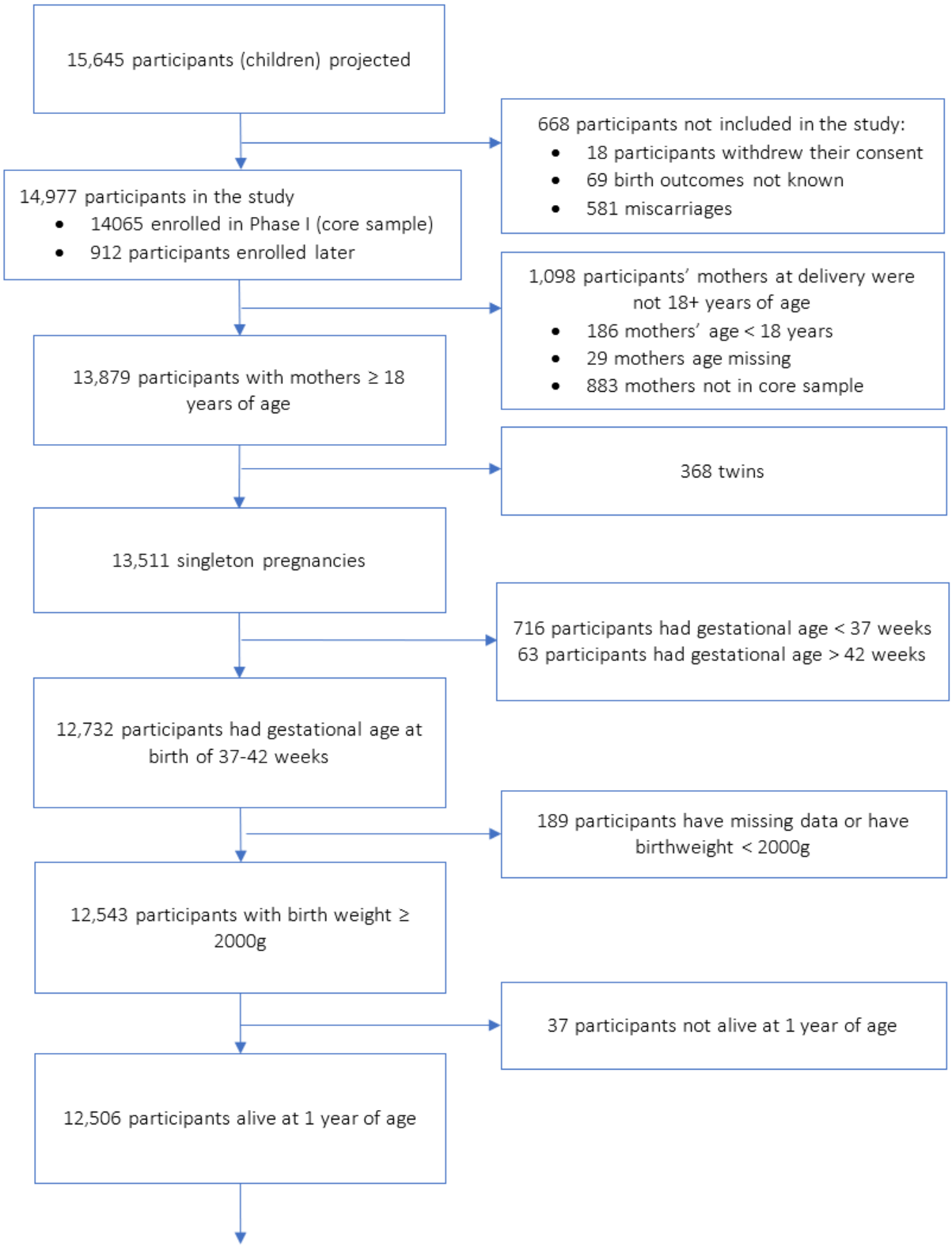
In sum, we observed that the association between environmental and genetic factors can place individuals at risk for adult comorbid chronic conditions from an early age, and that a striatal dopamine transporter gene network expression has a central role in moderating the association of the early environment with the risk for these diseases. These findings open opportunities for the exploration of the understudied field of precision prevention in pediatrics, and the potential design of more effective interventions and primary care strategies.

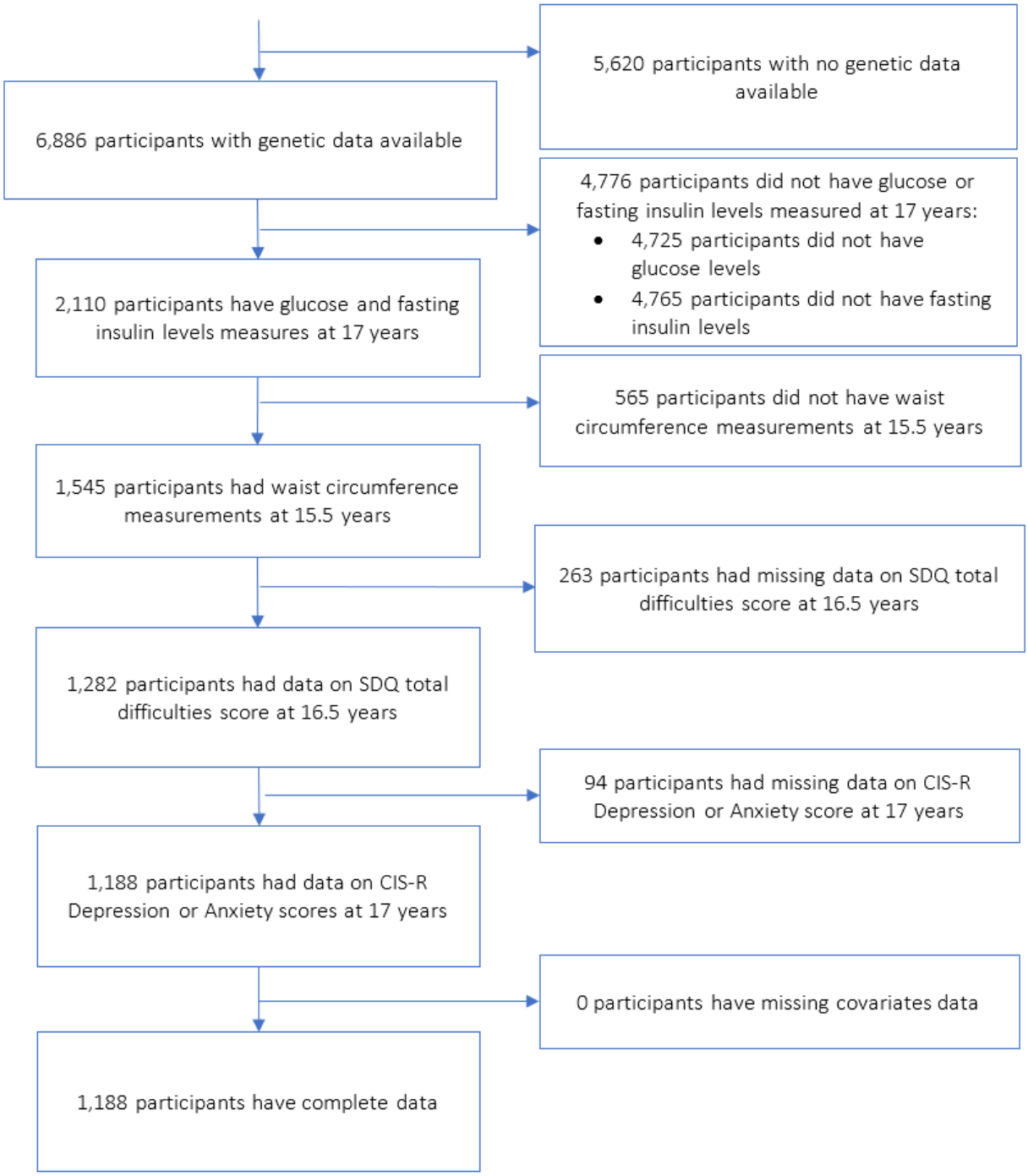
Supplementary Information

Supplementary Figures

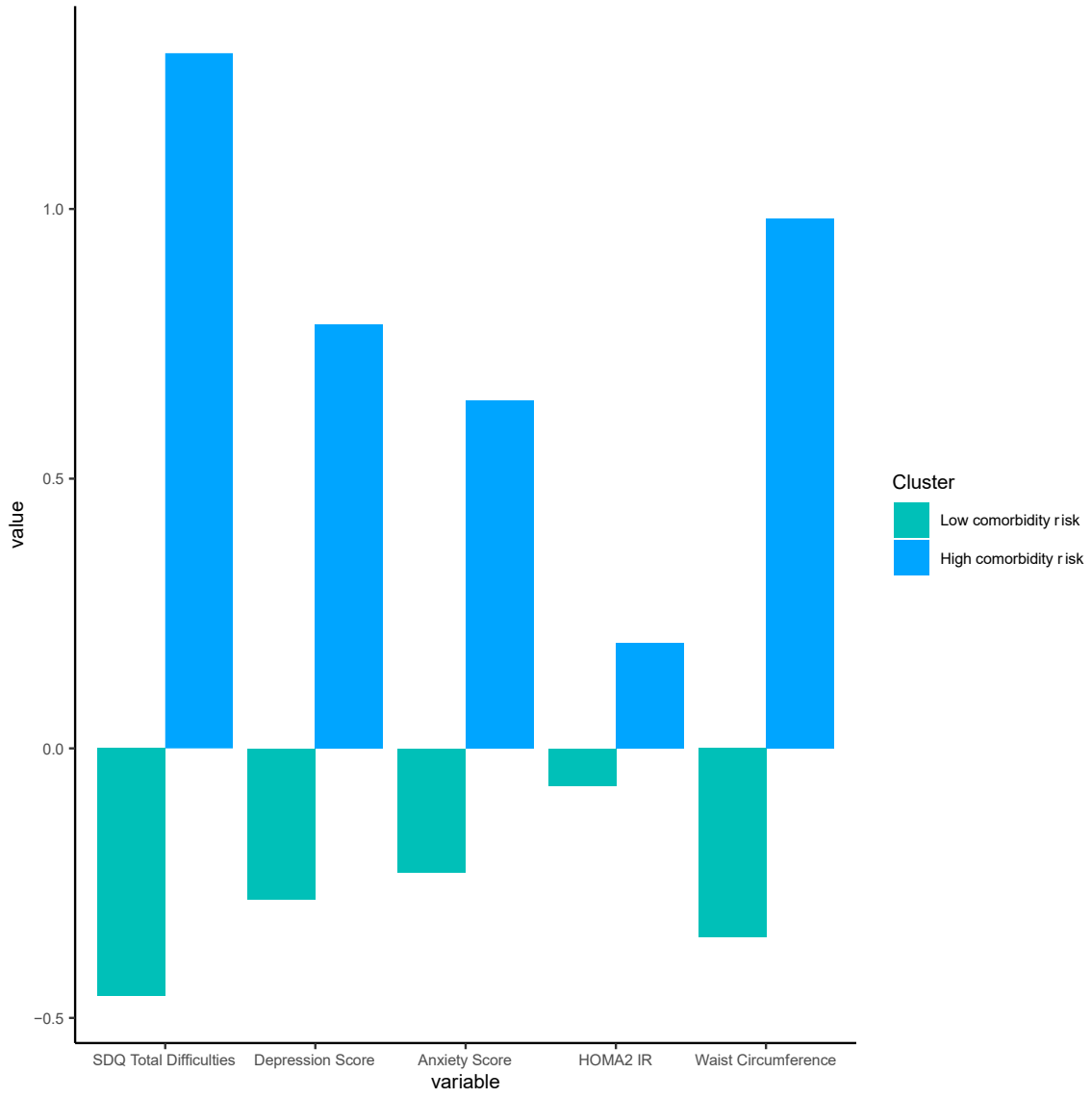


Supplementary Figure 5. *UK Biobank Sample size block scheme for each inclusion and exclusion criteria applied.* Block scheme depicting each step of inclusion and exclusion criteria applied to the original sample (N=502,543) until reaching final sample size for main hypothesis testing (N=225,972) and final sample size for parallel ICA analysis group of low birth weight (N=953) and non-low birth weight (N=953).

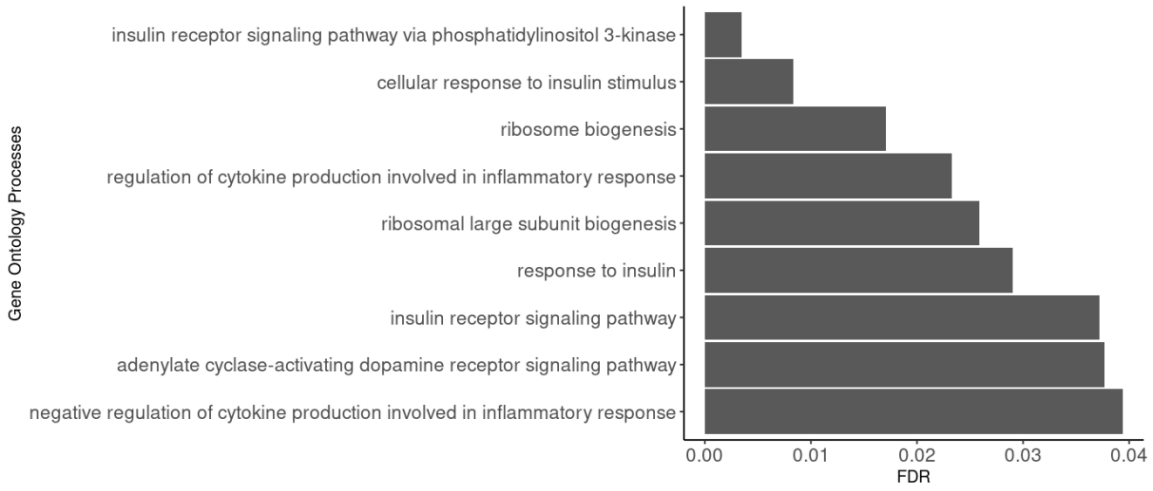

15,645 participants (children) projected
668 participants not included in the study:
• 18 participants withdrew their consent
• 69 birth outcomes not known
• 581 miscarriages
14,977 participants in the study
• 14065 enrolled in Phase I (core sample)
• 912 participants enrolled later
1,098 participants' mothers at delivery were not 18+ years of age
• 186 mothers' age < 18 years
• 29 mothers age missing
• 883 mothers not in core sample
13,879 participants with mothers ≥ 18 years of age
368 twins
13,511 singleton pregnancies
716 participants had gestational age < 37 weeks
63 participants had gestational age > 42 weeks
12,732 participants had gestational age at birth of 37-42 weeks
189 participants have missing data or have birthweight < 2000g
12,543 participants with birth weight ≥ 2000g
37 participants not alive at 1 year of age
12,506 participants alive at 1 year of age



Supplementary Figure 6. *ALSPAC Sample size block scheme for each inclusion and exclusion criteria applied.* Block scheme depicting each step of inclusion and exclusion criteria applied to the original sample (N=15,645) until reaching final sample size for main hypothesis testing (N=1,188).


1.0
0.5
0.0
−0.5
value
SDQ Total Difficulties
Depression Score
Anxiety Score
HOMA2 IR
Waist Circumference
variable
Cluster
Low comorbidity risk
High comorbidity risk

Supplementary Figure 7. *Comorbidity risk clusters – ALSPAC.* Cluster 1 – Low Comorbidity Risk (N=876) and Cluster 2 – High Comorbidity Risk (N=312) mean z score values for the variables considered in the cluster analysis. Cluster high comorbidity risk represents the psychiatric and metabolic comorbidity risk profile in adolescents. SDQ Total Difficulties is a score from the Strengths & Difficulties Questionnaire (SDQ). Depression and anxiety scores were computed from the Computerized Interview Schedule – Revised (CIS-R). Homeostatic Model Assessment of Insulin Resistance (HOMA-IR) was computed using the HOMA2 calculation tool. Waist circumference was measured in centimeters. Clustering of these variables was performed using mclust package in R.



Supplementary Figure 8. *Striatal SLC6A3 gene network enrichment analysis.* Gene ontology processes related to genes included in striatal SLC6A3 gene network. Enrichment was performed using MetaCore®. The significance was considered for the false discovery rate (FDR) adjusted p-value <0.05.

Supplementary Tables

Supplementary Table 1. Genes co-expressed with the SLC6A3 in mice striatum data, used to calculate striatum *SLC6A3* ePGS genetic score.

Mouse gene	Mouse Ensembl gene ID	Human gene	Human Ensembl gene ID	Description
Rnf165	ENSMUSG00000025427	RNF165	ENSG00000141622	ring finger protein 165 [Source:HGNC Symbol;Acc:HGNC:31696]
Rps21	ENSMUSG00000039001	RPS21	ENSG00000171858	ribosomal protein S21 [Source:HGNC Symbol;Acc:HGNC:10409]
Ipmk	ENSMUSG00000060733	IPMK	ENSG00000151151	inositol polyphosphate multikinase [Source:HGNC Symbol;Acc:HGNC:20739]
Mllt3	ENSMUSG00000028496	MLLT3	ENSG00000171843	MLLT3 super elongation complex subunit [Source:HGNC Symbol;Acc:HGNC:7136]
Hnrnpk	ENSMUSG00000021546	HNRNPK	ENSG00000165119	heterogeneous nuclear ribonucleoprotein K [Source:HGNC Symbol;Acc:HGNC:5044]
Rsl1d1	ENSMUSG00000005846	RSL1D1	ENSG00000171490	ribosomal L1 domain containing 1 [Source:HGNC Symbol;Acc:HGNC:24534]
Gfra2	ENSMUSG00000022103	GFRA2	ENSG00000168546	GDNF family receptor alpha 2 [Source:HGNC Symbol;Acc:HGNC:4244]
Nptxr	ENSMUSG00000022421	NPTXR	ENSG00000221890	neuronal pentraxin receptor [Source:HGNC Symbol;Acc:HGNC:7954]
Tgif2	ENSMUSG00000062175	TGIF2	ENSG00000118707	TGFB induced factor homeobox 2 [Source:HGNC Symbol;Acc:HGNC:15764]
Jarid2	ENSMUSG00000038518	JARID2	ENSG00000008083	jumonji and AT-rich interaction domain containing 2 [Source:HGNC Symbol;Acc:HGNC:6196]
Bmf	ENSMUSG00000040093	BMF	ENSG00000104081	Bcl2 modifying factor [Source:HGNC Symbol;Acc:HGNC:24132]
6820408C15Rik	ENSMUSG00000032680	C20orf96	ENSG00000196476	chromosome 20 open reading frame 96 [Source:HGNC Symbol;Acc:HGNC:16227]
Rpl14	ENSMUSG00000025794	RPL14	ENSG00000188846	ribosomal protein L14 [Source:HGNC Symbol;Acc:HGNC:10305]
Stk32a	ENSMUSG00000039954	STK32A	ENSG00000169302	serine/threonine kinase 32A [Source:HGNC Symbol;Acc:HGNC:28317]
Cdh10	ENSMUSG00000022321	CDH10	ENSG00000040731	cadherin 10 [Source:HGNC Symbol;Acc:HGNC:1749]
Sdc3	ENSMUSG00000025743	SDC3	ENSG00000162512	syndecan 3 [Source:HGNC Symbol;Acc:HGNC:10660]
Mkrn1	ENSMUSG00000029922	MKRN1	ENSG00000133606	makorin ring finger protein 1 [Source:HGNC Symbol;Acc:HGNC:7112]

Bbc3	ENSMUSG00000002083	BBC3	ENSG00000105327	BCL2 binding component 3 [Source:HGNC Symbol;Acc:HGNC:17868]
Tubb2a	ENSMUSG00000058672	TUBB2A	ENSG00000137267	tubulin beta 2A class IIa [Source:HGNC Symbol;Acc:HGNC:12412]
Bcl11a	ENSMUSG00000000861	BCL11A	ENSG00000119866	BAF chromatin remodeling complex subunit BCL11A [Source:HGNC Symbol;Acc:HGNC:13221]
Rpl37a	ENSMUSG00000046330	RPL37A	ENSG00000197756	ribosomal protein L37a [Source:HGNC Symbol;Acc:HGNC:10348]
Cotl1	ENSMUSG00000031827	COTL1	ENSG00000103187	coactosin like F-actin binding protein 1 [Source:HGNC Symbol;Acc:HGNC:18304]
Arl4c	ENSMUSG00000049866	ARL4C	ENSG00000188042	ADP ribosylation factor like GTPase 4C [Source:HGNC Symbol;Acc:HGNC:698]
Rpl27	ENSMUSG00000063316	RPL27	ENSG00000131469	ribosomal protein L27 [Source:HGNC Symbol;Acc:HGNC:10328]
Rrm1	ENSMUSG00000030978	RRM1	ENSG00000167325	ribonucleotide reductase catalytic subunit M1 [Source:HGNC Symbol;Acc:HGNC:10451]
Rpl30	ENSMUSG00000058600	RPL30	ENSG00000156482	ribosomal protein L30 [Source:HGNC Symbol;Acc:HGNC:10333]
Gbx2	ENSMUSG00000034486	GBX2	ENSG00000168505	gastrulation brain homeobox 2 [Source:HGNC Symbol;Acc:HGNC:4186]
Cux1	ENSMUSG00000029705	CUX1	ENSG00000257923	cut like homeobox 1 [Source:HGNC Symbol;Acc:HGNC:2557]
Dcaf12	ENSMUSG00000028436	DCAF12	ENSG00000198876	DDB1 and CUL4 associated factor 12 [Source:HGNC Symbol;Acc:HGNC:19911]
Klhl8	ENSMUSG00000029312	KLHL8	ENSG00000145332	kelch like family member 8 [Source:HGNC Symbol;Acc:HGNC:18644]
Zfy1	ENSMUSG00000053211	ZNF586	ENSG00000083828	zinc finger protein 586 [Source:HGNC Symbol;Acc:HGNC:25949]
Zfy1	ENSMUSG00000053211	ZNF480	ENSG00000198464	zinc finger protein 480 [Source:HGNC Symbol;Acc:HGNC:23305]
Zfy1	ENSMUSG00000053211	ZNF548	ENSG00000188785	zinc finger protein 548 [Source:HGNC Symbol;Acc:HGNC:26561]
Stag1	ENSMUSG00000037286	STAG1	ENSG00000118007	stromal antigen 1 [Source:HGNC Symbol;Acc:HGNC:11354]
Apc	ENSMUSG00000005871	APC	ENSG00000134982	APC regulator of WNT signaling pathway [Source:HGNC Symbol;Acc:HGNC:583]
Zfy1	ENSMUSG00000053211	ZNF549	ENSG00000121406	zinc finger protein 549 [Source:HGNC Symbol;Acc:HGNC:26632]
Zfy1	ENSMUSG00000053211	ZNF134	ENSG00000213762	zinc finger protein 134 [Source:HGNC Symbol;Acc:HGNC:12918]
Zfy1	ENSMUSG00000053211	ZNF304	ENSG00000131845	zinc finger protein 304 [Source:HGNC Symbol;Acc:HGNC:13505]

Zfy1	ENSMUSG00000053211	ZNF154	ENSG00000179909	zinc finger protein 154 [Source:HGNC Symbol;Acc:HGNC:12939]
Zfy1	ENSMUSG00000053211	ZNF793	ENSG00000188227	zinc finger protein 793 [Source:HGNC Symbol;Acc:HGNC:33115]
Zfy1	ENSMUSG00000053211	ZNF772	ENSG00000197128	zinc finger protein 772 [Source:HGNC Symbol;Acc:HGNC:33106]
Ptpn9	ENSMUSG00000032290	PTPN9	ENSG00000169410	protein tyrosine phosphatase non-receptor type 9 [Source:HGNC Symbol;Acc:HGNC:9661]
Hnrnpa1	ENSMUSG00000046434	HNRNPA1	ENSG00000135486	heterogeneous nuclear ribonucleoprotein A1 [Source:HGNC Symbol;Acc:HGNC:5031]
Myd88	ENSMUSG00000032508	MYD88	ENSG00000172936	MYD88 innate immune signal transduction adaptor [Source:HGNC Symbol;Acc:HGNC:7562]
Brd4	ENSMUSG00000024002	BRD4	ENSG00000141867	bromodomain containing 4 [Source:HGNC Symbol;Acc:HGNC:13575]
Ube2i	ENSMUSG00000015120	UBE2I	ENSG00000103275	ubiquitin conjugating enzyme E2 I [Source:HGNC Symbol;Acc:HGNC:12485]
Apc2	ENSMUSG00000020135	APC2	ENSG00000115266	APC regulator of WNT signaling pathway 2 [Source:HGNC Symbol;Acc:HGNC:24036]
Trh	ENSMUSG00000005892	TRH	ENSG00000170893	thyrotropin releasing hormone [Source:HGNC Symbol;Acc:HGNC:12298]
Pklr	ENSMUSG00000041237	PKLR	ENSG00000143627	pyruvate kinase L/R [Source:HGNC Symbol;Acc:HGNC:9020]
Krt18	ENSMUSG00000023043	KRT18	ENSG00000111057	keratin 18 [Source:HGNC Symbol;Acc:HGNC:6430]
Hnrnpr	ENSMUSG00000066037	HNRNPR	ENSG00000125944	heterogeneous nuclear ribonucleoprotein R [Source:HGNC Symbol;Acc:HGNC:5047]
Agrn	ENSMUSG00000041936	AGRN	ENSG00000188157	agrin [Source:HGNC Symbol;Acc:HGNC:329]
Cct3	ENSMUSG00000001416	CCT3	ENSG00000163468	chaperonin containing TCP1 subunit 3 [Source:HGNC Symbol;Acc:HGNC:1616]
Rpl26	ENSMUSG00000060938	RPL26	ENSG00000161970	ribosomal protein L26 [Source:HGNC Symbol;Acc:HGNC:10327]
Eef1b2	ENSMUSG00000025967	EEF1B2	ENSG00000114942	eukaryotic translation elongation factor 1 beta 2 [Source:HGNC Symbol;Acc:HGNC:3208]
Gfer	ENSMUSG00000040888	GFER	ENSG00000127554	growth factor, augmenter of liver regeneration [Source:HGNC Symbol;Acc:HGNC:4236]
Ephb2	ENSMUSG00000028664	EPHB2	ENSG00000133216	EPH receptor B2 [Source:HGNC Symbol;Acc:HGNC:3393]
Zfp454	ENSMUSG00000048728	ZNF454	ENSG00000178187	zinc finger protein 454 [Source:HGNC Symbol;Acc:HGNC:21200]

Rps20	ENSMUSG00000028234	RPS20	ENSG00000008988	ribosomal protein S20 [Source:HGNC Symbol;Acc:HGNC:10405]
Mllt11	ENSMUSG00000053192	MLLT11	ENSG00000213190	MLLT11 transcription factor 7 cofactor [Source:HGNC Symbol;Acc:HGNC:16997]
Pias4	ENSMUSG00000004934	PIAS4	ENSG00000105229	protein inhibitor of activated STAT 4 [Source:HGNC Symbol;Acc:HGNC:17002]
Rps13	ENSMUSG00000090862	RPS13	ENSG00000110700	ribosomal protein S13 [Source:HGNC Symbol;Acc:HGNC:10386]
Grwd1	ENSMUSG00000053801	GRWD1	ENSG00000105447	glutamate rich WD repeat containing 1 [Source:HGNC Symbol;Acc:HGNC:21270]
Ambra1	ENSMUSG00000040506	AMBRA1	ENSG00000110497	autophagy and beclin 1 regulator 1 [Source:HGNC Symbol;Acc:HGNC:25990]
Pik3ca	ENSMUSG00000027665	PIK3CA	ENSG00000121879	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha [Source:HGNC Symbol;Acc:HGNC:8975]
Rpl13	ENSMUSG00000000740	RPL13	ENSG00000167526	ribosomal protein L13 [Source:HGNC Symbol;Acc:HGNC:10303]
Gnb1	ENSMUSG00000029064	GNB1	ENSG00000078369	G protein subunit beta 1 [Source:HGNC Symbol;Acc:HGNC:4396]

558 unique mouse genes with absolute correlation index **>=** 0.5 were retained from GeneNetwork. These genes were converted to 381 unique human genes. The 381 unique human genes were filtered using BrainSpan data (overly expressed striatum genes in all prenatal vs adult samples) resulting in 72 genes. Of these 72 genes, 67 were found in GTEx and our study sample and were used to calculate striatum *SLC6A3 ePGS.*

Supplementary Table 2. Means of variables used in the cluster analysis for the ALSPAC cohort.

Variables	Cluster - Low comorbidity risk		Cluster – High comorbidity risk		P-value
	Mean / %	SE / N	Mean / %	SE / N	
SDQ Total Difficulties	5.02	0.14	6.77	0.26	<0.001
Depression Score	0	0	1.08	0.06	<0.001
Anxiety Score	0	0	0.88	0.06	<0.001
HOMA2-IR	0.81	0.01	1.08	0.05	<0.001
Waist Circumference	75.9	0.25	77.17	0.57	0.04
Sex -Male	49.5%	434	42%	131	0.03
zBMI	.26	0.03	0.31	0.06	0.45
Birth weight (kg)	3.515	0.015	3.49	0.025	0.44
	N=876		N=312		

Supplementary Table 3. Main effect of PRS on psychiatric and cardio-metabolic comorbidity.

		Main effect		
	PRS scores	*p*	β	OR
UK Biobank	Type 2 diabetes PRS_EPIC	< 0.001	0.073	1.076
	Major depression disorders PRS	< 0.001	0.065	1.068
	PRS scores	*p*	β	OR
ALSPAC	Type 2 diabetes PRS_EPIC	0.253	-0.076	0.927
	Major depression disorders PRS	0.882	0.020	1.020

Supplementary Table 4. Interaction effect between PRS and birth weight on psychiatric and cardio-metabolic comorbidity.

		Interaction effect		
	PRS scores	*p*	β	OR
UK Biobank	Type 2 diabetes PRS_EPIC	0.490	0.014	1.014
	Major depression disorders PRS	0.608	-0.011	0.989
	PRS scores	*p*	β	OR
ALSPAC	Type 2 diabetes PRS_EPIC	0.461	0.112	1.119
	Major depression disorders PRS	0.798	-0.078	0.925

Supplementary Table 5. Characteristics comparison between UK Biobank low vs randomly-selected non-low birth weight groups (n=1906) used in the parallel ICA analysis

Characteristics	Low Birth Weight (n =953)		Non-low Birth Weight (n = 953)		P-value
	Mean or %	SD or N	Mean or %	SD or N	
Sex - Male	34.7%	331	44.7%	426	P<0.001
Birth weight (grams)	2129	376	3464	493	P<0.001
Completed full-time education at 14-years of age or younger	0.7%	4	0.9%	5	0.97
Age at recruitment (years)	54.95	7.37	53.8	7.45	P<0.001
Townsend deprivation index at recruitment	-1.79	2.75	-2.02	2.69	0.06
BMI at recruitment	26.83	4.45	26.52	4.24	0.11
DAT1 STR ePGS	0.02	0.95	0.07	0.96	0.29

Supplementary Table 6. Characteristics comparison between randomly-selected non-low birth weight group used in the parallel ICA analysis vs full sample of non-low birth weight individuals with MRI data available in the UK Biobank (n=10214)

Characteristics	Randomly-selected Non-low Birth Weight (n =953)		Full sample of Non-low Birth Weight (n = 9261)		P-value
	Mean or %	SD or N	Mean or %	SD or N	
Sex - Male	44.7%	426	42.3%	3920	0.17
Birth weight (grams)	3464	493	3460	495	0.79
Completed full-time education at 14-years of age or younger	0.9%	5	0.5%	27	0.4
Age at recruitment (years)	53.8	7.45	53.75	7.38	0.85
Townsend deprivation index at recruitment	-2.02	2.69	-2.07	2.59	0.617
BMI at recruitment	26.52	4.24	26.51	4.34	0.956
DAT1 STR ePGS	0.07	0.96	0.05	0.96	0.607

Supplementary Table 7. Subset of significant SNPs related to gray matter density variations according to pICA analysis.

Significant SNPs	
SNP	Z Score
rs34605159	12.99878
rs7259029	-11.3745
rs139012629	11.11315
rs115385848	-9.25124
rs116962884	8.970151
rs11704200	8.799921
rs184236146	-7.28849
rs6737722	7.180836
rs112664980	5.614825
rs56988084	5.071326
rs2556378	-5.00336
rs76380377	-4.92952
rs67621412	4.895654
rs72795692	4.635065
rs6947066	-4.32333
rs1169557	4.131825
rs3762272	-3.79013
rs372339543	-3.7461
rs58399787	-3.63446
rs2966318	-3.50433
rs2967848	3.478973
rs140310740	3.390835
rs70953651	3.35623
rs149133487	-3.20932
rs71526493	-3.17741
rs3128100	3.123754
rs2967871	-3.06511
rs117055030	2.941269
rs1536096	2.915037
rs11689362	2.813534
rs3121575	-2.7558
rs57843539	-2.61199
rs11707190	2.509964
rs2048008	2.500896

Supplementary Methods

Participants

UK Biobank – Adult cohort: UK Biobank contains information on participants' lifestyle and health data at baseline or follow-up, which were collected through questionnaires, physical measurements, and biological samples. For the purpose of this project, only unrelated subjects were considered in the analysis. Exclusion criteria were 1) participants who withdrew their consent from the study, 2) no genotyping data, 3) related participants (genetic kinship to other participants > 0.04), 4) inconsistencies in genetic and reported sex and 5) outliers for heterozygosity. All subjects selected for the analysis satisfied the following criteria: (1) have genotyping data available and (2) have both the diagnosis outcome and birth weight available. Detailed description of sample selection process can be found in **Supplementary Figure 1**.

ALSPAC – Adolescent cohort: Data were collected during clinic visits or with postal questionnaires. Please note that the study website contains details of all the data that is available through a fully searchable data dictionary and variable search tool at http://www.bristol.ac.uk/alspac/researchers/our-data/. The following inclusion criteria were applied: unrelated individuals, gestational age between 37 and 42 weeks inclusively, maternal age at delivery >= 18 years old, birth weight of at least 2000 grams, singleton pregnancies. Detailed description of sample selection process can be found in **Supplementary Figure 2**. The participant attrition block scheme differs slightly in originally projected ALSPAC participant numbers as we are using an earlier data file from 2019 to complete these analyses.

Genotyping

UK Biobank – Adult cohort: Blood samples from the UK Biobank were genotyped at the Affymetrix Research Services Laboratory in Santa Clara, California, USA. Genotyping was conducted using a bespoke BiLEVE Axiom array for 50,000 participants and the remaining 450,000 participants were genotyped using the Affymetrix UK Biobank Axiom array. The two SNP arrays are very similar with over 95% common marker content. Axiom Array plates were processed on the Affymetrix GeneTitan® Multi-Channel (MC) Instrument. Genotypes were then called from the resulting intensities in batches of ~4,700 samples (~4,800 including the controls) using the Affymetrix Power Tools software and the Affymetrix Best Practices Workflow. Individuals with the same genotype at any given SNP will cluster together in a two-dimensional intensity space (one dimension for each targeted allele). For the interim data release, Affymetrix performed further rounds of genotype calling using algorithms customized for the UK Biobank project. These algorithms targeted very rare SNPs with 6 or fewer minor alleles in a batch, and a subset of SNPs for which the generic calling algorithm did not perform optimally. After genotype calling, Affymetrix performed quality control in each batch separately, to exclude SNPs with poor cluster properties. If a SNP did not meet the Affymetrix prescribed QC thresholds in a given batch, it was set to missing in all individuals from that batch. Hardy-Weinberg equilibrium was performed for each batch. Affymetrix also checked sample quality (such as DNA concentration) and genotype calls were provided only for samples with sufficient DNA metrics. For SNP-based QC metrics, only individuals with similar ancestry and the population structure were characterized by computing principal components using only UK Biobank individuals. The array also includes coding variants across a range of minor allele frequencies (MAFs), including rare markers (<1% MAF); and markers that provide good genome-wide coverage for imputation in

European populations in the common (>5%) and low frequency (1–5%) MAF ranges. More information about the genotyping protocol, QC and imputation could be found in [111]. The population structure of the UK Biobank cohort was evaluated using fastPCA algorithm for principal component analysis[112]. To account for population stratification, the first forty principal components were included in the UK Biobank analysis.

ALSPAC – Adolescent cohort: Subjects in ALSPAC cohort were genotyped using the Illumina HumanHap550 quad genome-wide SNP genotyping platform by the Welcome Trust Sanger Institute (Cambridge, UK) and the Laboratory Corporation of America (Burlington, NC, US)[113]. The following quality control procedure was applied: participants with inconsistencies in self-reported and genotyped sex, minimal or extreme heterozygosity, high levels of individual missingness (>3%), and insufficient sample replication (IBD < 0.8) were excluded. SNPs with MAF <1%, call rate <95%, or those not in HWE ($p < 5 \times 10-7$) were removed. Imputation was conducted using Impute v3 and Haplotype Reference Consortium (HRC) imputation reference panel (release 1.1). The resulting data set consisted of 8,365 individuals and 38,898,739 SNPs available for analysis. The population structure of the ALSPAC cohort was described using principal component analysis[114,115], which was conducted on the genotyped autosomal SNPs with MAF > 5% with the following pruning parameters for linkage disequilibrium: 100-SNP sliding window, an increment of 5 SNPs, and variance inflation factor threshold of 1.01. To account for population stratification, the first ten PCs were included in the analysis. Processing of the genotyping data was done using PLINK 1.9 [116](authors: Shaun Purcell, Christopher Chang; www.cog-genomics.org/plink/1.9/).

Gene expression levels at different developmental stages

In order to confirm if the genes that composed the striatum SLC6A3 ePGS are co-expressed in humans and investigate their patterns of gene co-expression during different life periods, we used the human postmortem striatal gene expression data from the BrainSpan database[65]. The co-expression patterns were analyzed during two different stages of development: childhood/adolescence (0 to 19 years old, N=7) and adulthood (20 to 40 years old, N=6). The analyses were carried out in R (https://www.r-project.org)[69] using the heatmaply package[117].

Outcome measures

UK Biobank – Adult cohort: To access the presence of psychiatric and cardio-metabolic disorders within the UK Biobank cohort, we utilized diagnostic terms from across all participants hospital inpatient records, coded according to the International Classification of Diseases version 10 (ICD-10)[66]. The presence of at least one cardio-metabolic diagnosis and at least one psychiatric diagnosis was considered a comorbidity case.

Psychiatric disorder diagnosis definition:

Cases: From Hospital Episodes Data from UK bodies (English HES Data, Scottish Morbidity Register, Patient Episode Data) (Fields 41270):

Any primary or secondary diagnosis of ICD-10 Codes for:

o ICD10 F10-F19 Mental and behavioural disorders due to psychoactive substance use'

o ICD10 F20-F29 Schizophrenia, schizotypal and delusional disorders

o ICD10 F30-F39 Mood [affective] disorders

o ICD10 F40-F48 Neurotic, stress-related and somatoform disorders

Controls: No primary or secondary diagnosis of ICD-10 codes for:

o ICD10 F10-F19 Mental and behavioural disorders due to psychoactive substance use'

o ICD10 F20-F29 Schizophrenia, schizotypal and delusional disorders

o ICD10 F30-F39 Mood [affective] disorders

o ICD10 F40-F48 Neurotic, stress-related and somatoform disorders

Cardio-metabolic disorder diagnosis definition:

Cases: From Hospital Episodes Data from UK bodies (English HES Data, Scottish Morbidity Register, Patient Episode Data) (Fields 41270):

Any primary or secondary diagnosis of ICD-10 Codes for:

o ICD10 E11 Non-insulin-dependent diabetes

o ICD10 I70 Atherosclerosis

o ICD10 I63 Cerebral infarction

o ICD10 I20-I25 Ischaemic heart diseases

Controls: No primary or secondary diagnosis of ICD-10 codes for:

o ICD10 E11 Non-insulin-dependent diabetes

o ICD10 I70 Atherosclerosis

o ICD10 I63 Cerebral infarction

o ICD10 I20-I25 Ischaemic heart diseases

ALSPAC – Adolescent cohort: The indicators of risk to develop psychiatric disorders comprised: a) total difficulties score as measured by the Strengths and Difficulties Questionnaire[118] filled out by primary caregivers of 16.6 year old participants. This instrument evaluates behavioral problems, and a total difficulties score is computed by adding four domains of the scale (emotional symptoms, conduct problems, hyperactivity/inattention and peer relationship problems) that represent negative behaviors; b) depression score and c) anxiety score as measured by the Computerized Interview Schedule – Revised (CIS-R) that establishes the nature and severity of neurotic symptoms[119], applied to 17.5 year old participants. The indicators of risk to develop metabolic disorders involved: d) insulin resistance as measured by the Homeostatic Model Assessment of Insulin Resistance (HOMA-IR), calculated using plasma fasting glucose (mmol/l) and insulin (pmol/l) levels collected at 17.5 years of age. The calculation followed the updated version of the HOMA-IR index developed by Wallace et al (2004)[120] and was computed using the HOMA2 calculation tool (http://www.dtu.ox.ac.uk/homacalculator/); and e) waist circumference (cm) measured at 15.5 years of age. Similar to UK Biobank, we created comorbidity risk variable based on the indicators of risk to develop psychiatric and metabolic disorders. Precisely, to construct comorbidity risk variable we performed cluster analysis on five risk indicators: Total difficulties, depression and anxiety scores, HOMA-IR and waist circumference. All predictors were z-transformed prior to entering the clustering procedure and adjusted by sex. We defined a cluster solution of two clusters, representing lower and higher risk for comorbidity. Regression

analysis was carried out to demonstrate the difference between the means for each variable included in the cluster analysis (*Supplementary information*, **Supplementary Table 2, Supplementary Figure 3**).

Gray matter density in UK Biobank participants: T1 structural brain MRI pre-processed imaging data were generated by an image-processing pipeline developed and run on behalf of the UK Biobank[67]. High-resolution T1-structural images for the whole brain were acquired with straight sagittal orientation using a Siemens Skyra 3T running VD13A SP4, with a standard Siemens 32-channel RF receive head coil. The following parameters were used: resolution 1x1x1mm; field-of-view 208x256x256 matrix; 5 minutes duration; 1 mm isotropic resolution using 3D MPRAGE acquisition; in-plane acceleration iPAT=2; prescan-normalize. Full 3D gradient distortion correction (GDC) was applied to the original T1 image and the field of view (FOV) was cut down to reduce the amount of non-brain tissue. Tools used to achieve this include BET (Brain Extraction Tool), FLIRT (FMRIB's Linear Image Registration Tool), and the MNI152 “nonlinear 6th generation” standard-space T1 template. A non-linear registration to MNI152 space was used with FNIRT (FMRIB's Nonlinear Image Registration Tool). Using the inverse of the MNI152 alignment warp, a standard-space brain mask was transformed into the native T1 space and applied to the T1 image to generate a brain-extracted T1. Tissue-type segmentation was applied to T1 weighted images using FSL/FAST (FMRIB's Automated Segmentation Tool). Then, the T1-weighted gray matter images were selected and adjusted for age and sex. For each voxel separately, we applied linear regression analysis to regress the intensity on age and sex and used the residuals in p-ICA analysis.

References

1. Prince MJ, Wu F, Guo Y, et al. The burden of disease in older people and implications for health policy and practice. *Lancet* 2015; **385**(9967): 549-62.

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): 409-16.

4. Tian YE, Di Biase MA, Mosley PE, et al. Evaluation of Brain-Body Health in Individuals With Common Neuropsychiatric Disorders. *JAMA psychiatry* 2023.

5. Souama C, Lamers F, Milaneschi Y, et al. Depression, cardiometabolic disease, and their co-occurrence after childhood maltreatment: an individual participant data meta-analysis including over 200,000 participants. *BMC medicine* 2023; **21**(1): 93.

6. Vos T, Allen C, Arora M, 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): 1545-602.

7. Kyu HH, Abate D, Abate KH, 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): 1859-922.

8. Golden SH, Lazo M, Carnethon M, et al. Examining a bidirectional association between depressive symptoms and diabetes. *JAMA* 2008; **299**(23): 2751-9.

9. Moulton CD, Pickup JC, Ismail K. The link between depression and diabetes: the search for shared mechanisms. *The Lancet Diabetes & Endocrinology* 2015; **3**(6): 461-71.

10. 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**: 461-8.

11. De Hert M, Detraux J, Vancampfort D. The intriguing relationship between coronary heart disease and mental disorders. *Dialogues in Clinical Neuroscience* 2018; **20**(1): 31-40.

12. 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.

13. Lawn JE, Ohuma EO, Bradley E, et al. Small babies, big risks: global estimates of prevalence and mortality for vulnerable newborns to accelerate change and improve counting. *The Lancet* 2023; **401**(10389): 1707-19.

14. Hanson Ma, Gluckman P. Early developmental conditioning of later health and disease: physiology or pathophysiology? *Physiological reviews* 2014.

15. de Mendonça E, de Lima Macêna M, Bueno NB, de Oliveira ACM, Mello CS. Premature birth, low birth weight, small for gestational age and chronic non-communicable diseases in adult life: A systematic review with meta-analysis. *Early Hum Dev* 2020; **149**: 105154.

16. Wang SF, Shu L, Sheng J, et al. Birth weight and risk of coronary heart disease in adults: a meta-analysis of prospective cohort studies. *J Dev Orig Health Dis* 2014; **5**(6): 408-19.

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

18. Rich-Edwards JW, Kleinman K, Michels KB, 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.

19. Betts KS, Williams GM, Najman JM, Scott J, Alati R. The association between lower birth weight and comorbid generalised anxiety and major depressive disorder. *J Affect Disord* 2013; **146**(2): 231-7.

20. Wiles NJ, Peters TJ, Leon DA, Lewis G. Birth weight and psychological distress at age 45-51 years: results from the Aberdeen Children of the 1950s cohort study. *Br J Psychiatry* 2005; **187**: 21-8.

21. Lahti M, Eriksson J, Heinonen 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): 985-99.

22. Abel KM, Wicks S, Susser ES, et al. Birth weight, schizophrenia, and adult mental disorder: is risk confined to the smallest babies? *Archives of general psychiatry* 2010; **67**(9): 923-30.

23. Loret de Mola C, de França GV, Quevedo Lde A, Horta BL. Low birth weight, preterm birth and small for gestational age association with adult depression: systematic review and meta-analysis. *Br J Psychiatry* 2014; **205**(5): 340-7.

24. Colman I, Ataullahjan A, Naicker K, Van Lieshout RJ. Birth weight, stress, and symptoms of depression in adolescence: evidence of fetal programming in a national Canadian cohort. *Can J Psychiatry* 2012; **57**(7): 422-8.

25. 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.

26. 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-e.

27. Nousen EK, Franco JG, Sullivan EL. Unraveling the mechanisms responsible for the comorbidity between metabolic syndrome and mental health disorders. *Neuroendocrinology* 2013; **98**(4): 254-66.

28. Alves MB, Laureano DP, Dalle Molle R, 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**: 336-46.

29. 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.

30. Laureano DP, Alves MB, Miguel PM, et al. Intrauterine growth restriction modifies the accumbal dopaminergic response to palatable food intake. *Neuroscience* 2018.

31. Pizzagalli DA, Berretta S, Wooten D, 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): 854-61.

32. 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): 728-40.

33. ter Horst KW, Lammers NM, Trinko R, et al. Striatal dopamine regulates systemic glucose metabolism in humans and mice. *Science Translational Medicine* 2018; **10**(442): eaar3752.

34. Huot P, Parent A. Dopaminergic neurons intrinsic to the striatum. *Journal of neurochemistry* 2007; **101**(6): 1441-7.

35. Threlfell S, Mohammadi AS, Ryan BJ, et al. Striatal Dopamine Transporter Function Is Facilitated by Converging Biology of α-Synuclein and Cholesterol. *Front Cell Neurosci* 2021; **15**: 658244.

36. 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): 102-9.

37. Sudlow C, Gallacher J, Allen N, 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.

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

39. Boyd A, Golding J, Macleod 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): 111-27.

40. 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.

41. Hari Dass SA, McCracken K, Pokhvisneva I, 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.

42. Miguel PM, Pereira LO, Barth B, 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): 621-30.

43. Restrepo-Lozano JM, Pokhvisneva I, Wang Z, et al. Corticolimbic DCC gene co-expression networks as predictors of impulsivity in children. *Molecular Psychiatry* 2022: 1-9.

44. Silveira PP, Pokhvisneva I, Parent C, 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): 1601-17.

45. de Lima RMS, Barth B, Arcego DM, 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.

46. de Lima RMS, Barth B, Mar Arcego D, et al. Leptin receptor co-expression gene network moderates the effect of early life adversity on eating behavior in children. *Communications Biology* 2022; **5**(1): 1092.

47. de Mendonça Filho EJ, Barth B, Bandeira DR, et al. Cognitive development and brain gray matter susceptibility to prenatal adversities: Moderation by the prefrontal cortex brain-derived neurotrophic factor gene co-expression network. *Frontiers in Neuroscience* 2021; **15**: 744743.

48. Dalmaz C, Barth B, Pokhvisneva I, et al. Prefrontal cortex VAMP1 gene network moderates the effect of the early environment on cognitive flexibility in children. *Neurobiology of Learning and Memory* 2021: 107509.

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

50. Miguel PM, Deniz BF, Deckmann I, et al. Prefrontal cortex dysfunction in hypoxic-ischaemic encephalopathy contributes to executive function impairments in rats: potential contribution for attention-deficit/hyperactivity disorder. *The World Journal of Biological Psychiatry* 2018; **19**(7): 547-60.

51. Dalle Molle R, Laureano DP, Alves MB, et al. Intrauterine growth restriction increases the preference for palatable foods and affects sensitivity to food rewards in male and female adult rats. *brain research* 2015; **1618**: 41-9.

52. Cunha FdS, Dalle Molle R, Portella AK, et al. Both food restriction and high-fat diet during gestation induce low birth weight and altered physical activity in adult rat offspring: the "Similarities in the Inequalities" model. *PLoS One* 2015; **10**(3): e0118586.

53. Portella AK, Silveira PP, Laureano DP, et al. Litter size reduction alters insulin signaling in the ventral tegmental area and influences dopamine-related behaviors in adult rats. *Behavioural brain research* 2015; **278**: 66-73.

54. Silveira PP, Portella AK, Assis SACdN, et al. Early life experience alters behavioral responses to sweet food and accumbal dopamine metabolism. *International journal of developmental neuroscience* 2010; **28**(1): 111-8.

55. 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.

56. Consortium GT. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013; **45**(6): 580-5.

57. Chen LM, Yao N, Garg E, et al. PRS-on-Spark (PRSoS): a novel, efficient and flexible approach for generating polygenic risk scores. *BMC Bioinformatics* 2018; **19**(1): 295.

58. Wray NR, Goddard ME, Visscher PM. Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Res* 2007; **17**(10): 1520-8.

59. Cai L, Wheeler E, Kerrison ND, et al. Genome-wide association analysis of type 2 diabetes in the EPIC-InterAct study. *Scientific Data* 2020; **7**(1): 393.

60. Wray NR, Ripke S, Mattheisen M, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature genetics* 2018; **50**(5): 668.

61. Warde-Farley D, Donaldson SL, Comes O, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research* 2010; **38**(suppl_2): W214-W20.

62. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 2003; **13**(11): 2498-504.

63. 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.

64. Franceschini A, Szklarczyk D, Frankild S, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* 2013; **41**(Database issue): D808-15.

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

66. Organization WH. ICD-10 : international statistical classification of diseases and related health problems : tenth revision. 2 ed; 2004.

67. Alfaro-Almagro F, Jenkinson M, Bangerter NK, et al. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* 2018; **166**: 400-24.

68. Magge SN, Goodman E, Armstrong SC, et al. The metabolic syndrome in children and adolescents: shifting the focus to cardiometabolic risk factor clustering. *Pediatrics* 2017; **140**(2).

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

70. 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.

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

72. 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.

73. Talairach J, Tournoux P. A coplanar stereotaxic atlas of the human brain: An approach to medical cerebral imaging. Thieme Medical Publishers, Inc., New York.; 1988.

74. Schüring AN, Lutz F, Tüttelmann F, Gromoll J, Kiesel L, Götte M. Role of syndecan-3 polymorphisms in obesity and female hyperandrogenism. *J Mol Med (Berl)* 2009; **87**(12): 1241-50.

75. 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.

76. Silveira PP, Meaney MJ. Examining the biological mechanisms of human mental disorders resulting from gene-environment interdependence using novel functional genomic approaches. *Neurobiology of Disease* 2023; **178**: 106008.

77. Sirugo G, Williams SM, Tishkoff SA. The missing diversity in human genetic studies. *Cell* 2019; **177**(1): 26-31.

78. Kang J, Brajanovski N, Chan KT, Xuan J, Pearson RB, Sanij E. Ribosomal proteins and human diseases: molecular mechanisms and targeted therapy. *Signal Transduct Target Ther* 2021; **6**(1): 323.

79. Smagin DA, Kovalenko IL, Galyamina AG, Bragin AO, Orlov YL, Kudryavtseva NN. Dysfunction in Ribosomal Gene Expression in the Hypothalamus and Hippocampus following Chronic Social Defeat Stress in Male Mice as Revealed by RNA-Seq. *Neural Plast* 2016; **2016**: 3289187.

80. Katzav A, Solodeev I, Brodsky O, et al. Induction of autoimmune depression in mice by anti-ribosomal P antibodies via the limbic system. *Arthritis Rheum* 2007; **56**(3): 938-48.

81. Murakami T, Ruengsinpinya L, Nakamura E, et al. Cutting Edge: G Protein Subunit β 1 Negatively Regulates NLRP3 Inflammasome Activation. *J Immunol* 2019; **202**(7): 1942-7.

82. Roohi E, Jaafari N, Hashemian F. On inflammatory hypothesis of depression: what is the role of IL-6 in the middle of the chaos? *Journal of Neuroinflammation* 2021; **18**(1): 45.

83. Karhunen V, Gill D, Huang J, et al. The interplay between inflammatory cytokines and cardiometabolic disease: bi-directional mendelian randomisation study. *BMJ Medicine* 2023; **2**(1): e000157.

84. 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.

85. Dalla Costa I, Buchanan CN, Zdradzinski MD, et al. The functional organization of axonal mRNA transport and translation. *Nature Reviews Neuroscience* 2020: 1-15.

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

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

88. Sun Z, Ma Y, Xie L, 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.

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

90. Bagalkot TR, Block ER, Bucchin K, 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): 234-50.

91. Mathieu P, Lemieux I, Després JP. Obesity, inflammation, and cardiovascular risk. *Clinical pharmacology & therapeutics* 2010; **87**(4): 407-16.

92. Shoelson SE, Herrero L, Naaz A. Obesity, inflammation, and insulin resistance. *Gastroenterology* 2007; **132**(6): 2169-80.

93. Martín-Calvo N, Goni L, Tur JA, Martínez JA. Low birth weight and small for gestational age are associated with complications of childhood and adolescence obesity: Systematic review and meta-analysis. *Obes Rev* 2022; **23 Suppl 1**: e13380.

94. Watson KT, Simard JF, Henderson VW, et al. Association of Insulin Resistance With Depression Severity and Remission Status: Defining a Metabolic Endophenotype of Depression. *JAMA psychiatry* 2021; **78**(4): 439-41.

95. Watson K, Nasca C, Aasly L, McEwen B, Rasgon N. Insulin resistance, an unmasked culprit in depressive disorders: Promises for interventions. *Neuropharmacology* 2018; **136**: 327-34.

96. Calkin CV, Chengappa KNR, Cairns K, et al. Treating Insulin Resistance With Metformin as a Strategy to Improve Clinical Outcomes in Treatment-Resistant Bipolar Depression (the TRIO-BD Study): A Randomized, Quadruple-Masked, Placebo-Controlled Clinical Trial. *J Clin Psychiatry* 2022; **83**(2).

97. Erensoy H, Niafar M, Ghafarzadeh S, Aghamohammadzadeh N, Nader ND. A pilot trial of metformin for insulin resistance and mood disturbances in adolescent and adult women with polycystic ovary syndrome. *Gynecol Endocrinol* 2019; **35**(1): 72-5.

98. Carvelli L, Morón JA, Kahlig KM, et al. PI 3-kinase regulation of dopamine uptake. *Journal of neurochemistry* 2002; **81**(4): 859-69.

99. 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): 1284-91.

100. Batra A, Latsko M, Portella AK, Silveira PP. Early adversity and insulin: neuroendocrine programming beyond glucocorticoids. *Trends Endocrinol Metab* 2021; **32**(12): 1031-43.

101. Alberry B, Silveira PP. Brain insulin signaling as a potential mediator of early life adversity effects on physical and mental health. *Neuroscience & Biobehavioral Reviews* 2023; **153**: 105350.

102. 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.

103. Hoops D, Flores C. Making dopamine connections in adolescence. *Trends in neurosciences* 2017; **40**(12): 709-19.

104. Batra A, Cuesta S, Alves MB, et al. Relationship between insulin and Netrin-1/DCC guidance cue pathway regulation in the prefrontal cortex of rodents exposed to prenatal dietary restriction. *Journal of Developmental Origins of Health and Disease* 2023; **14**(4): 501-7.

105. Wang C, Geng B, Cui Q, Guan Y, Yang J. Intracellular and extracellular adenosine triphosphate in regulation of insulin secretion from pancreatic β cells (β). *J Diabetes* 2014; **6**(2): 113-9.

106. Patterson JN, Cousteils K, Lou JW, Fox JEM, MacDonald PE, Joseph JW. Mitochondrial metabolism of pyruvate is essential for regulating glucose-stimulated insulin secretion. *Journal of Biological Chemistry* 2014; **289**(19): 13335-46.

107. Dalle Molle R, de Mendonça Filho EJ, Minuzzi L, et al. Thrifty-Eating Behavior Phenotype at the Food Court - Programming Goes Beyond Food Preferences. *Front Endocrinol (Lausanne)* 2022; **13**: 882532.

108. Yucel K, McKinnon MC, Chahal R, et al. Anterior Cingulate Volumes in Never-Treated Patients with Major Depressive Disorder. *Neuropsychopharmacology* 2008; **33**(13): 3157-63.

109. Leech R, Sharp DJ. The role of the posterior cingulate cortex in cognition and disease. *Brain* 2013; **137**(1): 12-32.

110. Cui Y, Tang TY, Lu CQ, Ju S. Insulin resistance and cognitive impairment: evidence from neuroimaging. *Journal of Magnetic Resonance Imaging* 2022; **56**(6): 1621-49.

111. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018; **562**(7726): 203.

112. Galinsky KJ, Bhatia G, Loh PR, et al. Fast Principal-Component Analysis Reveals Convergent Evolution of ADH1B in Europe and East Asia. *Am J Hum Genet* 2016; **98**(3): 456-72.

113. Richmond RC, Timpson NJ, Felix JF, et al. Using genetic variation to explore the causal effect of maternal pregnancy adiposity on future offspring adiposity: a Mendelian randomisation study. *PLoS medicine* 2017; **14**(1).

114. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS genetics* 2006; **2**(12).

115. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics* 2006; **38**(8): 904-9.

116. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015; **4**(1): s13742-015-0047-8.

117. Galili T, O'Callaghan A, Sidi J, Sievert C. heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* 2018; **34**(9): 1600-2.

118. Goodman R. The Strengths and Difficulties Questionnaire: a research note. *Journal of child psychology and psychiatry* 1997; **38**(5): 581-6.

119. Lewis G. Assessing psychiatric disorder with a human interviewer or a computer. *Journal of Epidemiology & Community Health* 1994; **48**(2): 207-10.

120. Wallace TM, Levy JC, Matthews DR. Use and abuse of HOMA modeling. *Diabetes care* 2004; **27**(6): 1487-95.

Chapter VII. Discussion

The high prevalence and comorbidity of psychiatric and cardiometabolic disorders present a complex puzzle in clinical research that is not easily understood if we consider diseases as static entities. This thesis shows that the elucidation of the mechanisms involved in these two types of disorders can be contextualized in a developmental perspective. We started by showing that DRD4 predicted gene expression variation significantly interacted with a measure of early life environment to predict emotional eating in children. As discussed earlier, this altered eating behavior, especially early in life and related to a pro-intake eating behavior, may serve as an endophenotype linked to the development of metabolic disorders, including obesity. The same reasoning is applied to the emotional domain of this behavior, which is related to emotional regulation[278] and might also act as an endophenotype for later psychopathology. Subsequently, we showed that a genetic score based in a network of genes coexpressed with the SLC6A3 gene in the striatum significantly interacts with birth weight, which represents the quality of the uterine environment, to predict cardiometabolic and psychiatric comorbidity in adults. In the same study we showed that in adolescents, a profile indicative of high risk for the development of psychiatric and metabolic comorbidity is also predicted by the same GxE interplay seen in adults. The profile indicative of high risk to develop these disorders comprised metabolic markers, including high insulin resistance score and higher waist circumference, and psychological markers, including high depression and anxiety scores. These results together highlight the role of behavioral and metabolic endophenotypes as precursors of a possible cascade of events that lead to the development of chronic disorders in adulthood. These cascades of events are thought to start with ELA exposure, that leads to early in life

behavioral[2,231,238] and metabolic alterations[49,99,102], but also at the dopaminergic system level[235,238,239], which consequently may further impact metabolism and behavior, especially the ones related to value processing. These alterations in behavior, metabolism, and neurotransmission, even if small, are persistent over the life course and interact with several layers of the environment, both at the macro and individual level, leading to the magnification and/or further alterations in these domains. The findings described in chapters III and VI are aligned with this view, and a conceptual framework of this cascade of events is discussed in chapter I, as well as in Alberry & Silveira (2023)[263], which details the role of insulin signaling alterations followed by ELA exposure and Silveira et al (2018)[110] that discussed the impacts followed by LBW in this context. The view that disease development is dynamic is aligned with other important conceptual frameworks, such as the ecological system theory, which states that human development is shaped by the complex interplay between the individual and the many layers of its environment. In this view, there is a contextual and individual variability that shapes human developmental processes[301]. Likewise, the developmental origins of health and disease hypothesis also considers the development of chronic disorders in a dynamic fashion, either by conferring increased sensitivity to environmental influences or by programming development through alterations led by ELA exposure[17,285].

In Chapter VI we found that the score based on the SLC6A3 striatum gene network, which significantly interacts with birth weight, is enriched for gene ontology process related to insulin signaling. Early life alteration in insulin signaling, such as augmented insulin sensitivity seen in LBW, is thought to be related to altered eating behaviors seen in these individuals, such as increased intake of palatable foods, rich in fat and sugar[2,238,241,257,258] and altered hedonic

responses to sweet taste[231,259]. These altered eating behaviours, related to the processing of reward information, have been linked to alterations in dopaminergic signaling[238,239,241,242]. Insulin signaling in the brain is increasingly being recognized as a modulator of dopaminergic activity[111,302,303]. Gruber et al (2023)[302] suggest that the actions of insulin on dopaminergic reward processing are a possible pathophysiological mechanism linking major depressive disorder and type 2 diabetes comorbidity. Similarly, Sullivan et al (2023)[303] discuss that insulin signaling in the brain plays an important role in behavioural relevant CNS functions, including behavioral flexibility. DA, especially in the prefrontocortical area, is also related to behavioral flexibility[304] and, as discussed earlier in chapter I, plays a crucial role in adaptive behavior. Both hyper and hypo-responsive DA systems could result from different ELA exposures[158], implicating DA in phenotypic versatility. Thus, both insulin and DA signaling alterations and their interplay are involved in the development and manifestation of altered behavioral endophenotypes seen early in life followed by ELA exposure, especially in the case of LBW.

It is important to mention that sensitivity analyses are crucial tools when evaluating the differential effects of composite genomic scores, such as the ePGS. The calculation of the ePGS facilitates tailoring the score according to the gene network of interest and the expression of this network of genes in a specific tissue. A common inquiry is to what extent the results, both at the GxE interaction level and in the investigation of putative gene ontology processes, are specific to the score used in the study. Sensitivity analyses were performed for the results seen in Chapters V and VI (data not shown), testing the same GxE interaction effect and investigating gene ontology processes. These preliminary data have shown that the results are, in fact, specific and distinct for each tailored ePGS, underscoring the capability of the ePGS to reflect

specific gene networks. This is an advantage when investigating underlying biological mechanisms at the human translational level. The ePGS, although stablished and used in several publications including high impact journals[162,305], is still considered a novel technique and a difficulty is observed in the scientific community to absorb the full scope and potential of the ePGS. Earlier submissions of the manuscript from Chapter VI lead us to consider that the ePGS performance would be better benchmarked in terms of clarity if compared to traditional PRS, rather than alternative ePGS featuring different gene networks and based on different brain regions.

The interpretation of GxE interaction effects implies that the risk for a certain phenotype can be either attenuated or increased by an environmental exposure, or that genetic factors can modify the impact of environmental factors. These interpretations take into account the nature of moderation analysis, in which the genetic component can modify the strength of the relationship between independent and dependent variables. An important distinction when investigating GxE interaction effects is the examination of the gene-environment correlation phenomenon, wherein genes can influence individual variations in exposure to different types of environments[306]. If GxE correlation exists, the interpretation of GxE interaction effects is hindered. In Chapter VI, our environmental predictor, birth weight, significantly correlates with our genomic predictor, striatum SLC6A3 ePGS, in the UK Biobank cohort (See **Table 4**), which should be considered a possible confounder for interpreting the results of GxE interactions. However, the relationship between large sample sizes, as in the case of the UK Biobank, and the identification of statistically significant minuscule effects exists and, in some cases, leads to the identification of misleading associations[307,308]. The correlation coefficient between birth weight

and striatum SLC6A3 ePGS on the UK Biobank is extremely small, being hundreds of times smaller than the small correlation threshold of 0.3[309]. In Chapter VI, if we examine the same gene-environment correlation on the ALSPAC cohort, a much smaller sample size in comparison to the UK Biobank reveals no significant correlation (See **Table 5**). The absence of gene-environment correlation in ALSPAC, together with the extremely small correlation coefficient in the UK Biobank, leads us to assume that, in Chapter VI, no gene-environment correlation phenomenon exists. Thus, this guides our interpretation of the results in the direction of GxE interactions.

The work presented in this thesis contributes to the growing knowledge on the long-term effects of ELA exposure, which underscores the imperative for further efforts to address and alleviate the consequences associated with ELA exposure. The meritocratic ideology within societies that prioritize individuality and personal gain becomes disconcerting when considering the research findings that link ELA exposure to numerous developmental challenges and subsequent risks for psychiatric and cardiometabolic diseases in adulthood. The field dedicated to studying the effects of ELA exposure highlights the injustices concealed by such a simplistic worldview, emphasizing the urgent need to formulate public policies that prioritize health equity. The concept of equity acknowledges the diverse circumstances of each individual, necessitating the fair distribution of resources and opportunities to achieve an equal outcome. It is also defined by the absence of avoidable or remediable differences among groups of people[310]. In this view, a public health system lacking prevention strategies is unlikely to provide equitable care to its population. Thus, preventive measures are crucial to mitigate the impacts of ELA exposure, especially the ones concerning the enhancement of prenatal care and

implementation of postnatal support policies for mothers and caregivers, domains that have a direct impact in the child's early life environment. Efforts in these areas are essential to address the disparities resulting from ELA exposure and promote a more equitable and just society. Medicine has revolutionized prenatal and neonatal care, significantly decreasing mortality rates for both mother[311,312] and infants under the age of five[313] in the last decades. Similar efforts should be put in place to mitigate other types of ELA exposure, including socioeconomic and psychological challenges, ideally by prevention rather than remediation, since contributing factors are known and potentially preventable. Our results showing a differential effect of enviroment exposure according to our genomic predictors, possibly indicates why some children benefit more than others from early childhood intervations to improve multiple domains[314]. In fact, the impact of these interventions has been relatively modest[315], suggesting that children do not profit equally. This variation in response to enviromental interventions leads to opportunities to better understand and effectivelly evaluate variations in intervention effectivenness and also indicates the need to plan for flexible implementation of interventions, aiming to increase the impact for those who may benefit the least.

This thesis work brings a translational perspective to explore the joint contribution of genetics and environment in disease development and manifestation. Leveraging from extensive knowledge from basic science we elaborated hypothesis-driven explorations of the neurobiological mechanisms involved in the association between ELA exposure and altered eating behavior and psychiatric and cardiometabolic comorbidities. Our hypothesis about individual differences in the expression of a dopaminergic gene network was tested by computing genetic scores that leveraged basic science data in animals and humans. We then

used this information to explore GxE effects in humans at different developmental stages. The results from chapters III and VI are contextualized at the human association level and may inform future molecular studies preoccupied with exploring and causally testing ELA exposure effects. Back translation from the findings of this work is fundamental for the effective progression of the objectives addressed here. The continuous translation of information from basic science to its application at the human association level, or through more robust research designs like randomized controlled trials, and the subsequent backtranslation to bench work, is a crucial process that should remain at the forefront while pursuing a research agenda over time. Instead of marginalizing diverse research methodologies, scientific progress is better achieved by prioritizing the integration of information and the seamless translation and back-translation from micro to macro aspects of a phenomenon.

In chapters IV and V, we discussed ways of capturing GxE effects. We concluded that GWAS is not suited to capture GxE effects and discussed the idea that utilization of genome-wide isolated signals may be detrimental when exploring possible underlying biological mechanisms involved in complex diseases, since these genome-wide isolated signals are not aggregated by biological function but by statistical association. We propose the incorporation of a complex system in biology approach, that considers genes as part of a co-regulatory network, instead of isolated signals. The complex system approach not only favors the utilization of gene networks instead of isolated genetic signals but also the incorporation of the interaction between the elements of a system, including macro-scale objects such as environment and genome. The complex system in biology is a shift from a reductionist paradigm often adopted in biological sciences, preoccupied with knowing and understanding the contribution of small parts of a

system with a cross-sectional view of time, usually resulting in moving away from the larger scenario. Medicine also traditionally follows this view, and approaches disorders as separate entities, with treatment following the same reasoning[316]. The reductionist paradigm is detrimental when complex diseases are being investigated, such as psychiatric and cardiometabolic conditions and their comorbidity, especially facing emerging evidence that pleiotropy – defined by one gene affecting multiple phenotypes and suggestive of shared genetic architecture[288] – is abundant among many complex disorders[289] and that genes, environment, and the interplay between them play a role in the etiology of these diseases[149-153]. Alongside the reductionist paradigm, the deterministic view in biology is still present in modern science[317]. However, it is known that not all individuals exposed to ELA will develop chronic diseases in the long term[133-135]. There is individual variability, which is partly attributed to the genetic background[136,137]. A complex system in biology approach preoccupied with abandoning reductionist and deterministic paradigms is potentially more philosophically aligned with the nature of the complex interplay between environment and genetics resulting from ELA exposure.

In chapter V we compared a technique that is informed by the gene networks perspective (ePGS) with the standard polygenic risk score calculations which considers isolated genetic signals (PRS). Through the examples given in chapter V we explored the promising aspects of ePGS in representing biological information, made possible by incorporating information into the ePGS that are not present in PRS. This additional information includes the coexpression between genes and the tissue-specific effect size of the association between genotyping and gene expression. Some authors suggest that the translation of findings using genetic risk scores

into actual improvements in healthcare requires the advancement of techniques that allow moving from the discovery of simple association signals to their functional interpretation. This could be achieved by incorporating other levels of data such as transcriptomics and proteomics, accelerating the translation of GWAS findings to therapeutic interventions[318]. This view is aligned with the conceptual framework of the ePGS technique.

This thesis has limitations that should be considered. Results from chapters III and VI explore the idea that genetic individual variations of dopaminergic genes are related to the differential response to ELA exposure effects, and they do so at the human association level. Even though we have demonstrated the value of such results, they are limited to the association level and molecular studies are needed to causally link dopaminergic gene expression variations to the differential effects of ELA exposure. Results from chapters III and VI are interpreted as a potential developmental cascade of events that links ELA exposure to the development of later in life psychiatric and cardiometabolic comorbidities, but the data used in these studies are cross-sectional. Future studies addressing this idea from a longitudinal perspective are imperative to more robustly address the interpretation given here. In chapter VI, insulin signaling is suggested as an underlying biological mechanism involved in the interaction effect between LBW and SLC6A3 striatal gene network expression on psychiatric and cardiometabolic comorbidities. This result is from putative information, gathered through gene enrichment analysis, that systematically combines literature information to provide insight into genetic results. Thus, further molecular studies are needed to link insulin signaling, LBW, SLC6A3 striatal gene network expression and psychiatric and cardiometabolic comorbidities, ideally by animal models where confounding variables can be more rigorously addressed. Our genetic scores are

largely based on GWAS from European ancestries and further studies following the same steps but with appropriate ancestry specific GWAS are needed, in a future where such GWAS will, hopefully, be available. We also did not consider sex-specific analysis in the studies presented in this thesis. Psychiatric and cardiometabolic disorders have potential sex-specific biological pathways that should be addressed in future studies to better understand the shared biological mechanisms between these comorbid disorders and to effectively translate findings to human therapeutics.

Chapter VIII. Conclusion

This thesis provides evidence that variations in dopaminergic gene expression in brain-specific regions significantly interact with environmental factors to predict emotional eating behavior in children, the probability of belonging to a high comorbidity risk profile in adolescents and the risk to developing psychiatric and cardiometabolic comorbidities in adulthood. Through the work presented in this thesis, we contribute to the growing body of knowledge on the potential neurobiological mechanisms connecting ELA exposure to the development of psychiatric and cardiometabolic disorders.

Our findings suggest that GWAS may not effectively capture GxE effects, and we showed that a genetic score based on a gene network perspective, incorporating biological information into the calculation of these scores, may better capture GxE effects and provide insights into biological mechanisms. This work contributes to an increased awareness of the differential impact of ELA on individuals and emphasizes the necessity of fostering preventive interventions to mitigate the effects of ELA exposure. It also sheds light on the role of GxE in the development of chronic disorders and the influence of variations in dopamine-related brain gene expression in modulating the effects of ELA exposure. Finally, it underscores the importance of incorporating functional genomics into genetic prediction scores.

Reference list for Non-Manuscript Chapters (Contributions to original knowledge, connecting statements and Chapters I, VII and VIII)

1. Barth, B. *et al.* Genetically predicted gene expression of prefrontal DRD4 gene and the differential susceptibility to childhood emotional eating in response to positive environment. *Appetite* **148**, 104594 (2020).
2. Silveira, P.P. *et al.* Genetic differential susceptibility to socioeconomic status and childhood obesogenic behavior: why targeted prevention may be the best societal investment. *JAMA pediatrics* **170**, 359-364 (2016).
3. Portella, A.K. *et al.* Genetically-predicted prefrontal DRD4 gene expression modulates differentiated brain responses to food cues in adolescent girls and boys. *Scientific Reports* **11**, 24094 (2021).
4. Portella, A.K. *et al.* Predicted DRD4 prefrontal gene expression moderates snack intake and stress perception in response to the environment in adolescents. *PLOS ONE* **15**, e0234601 (2020).
5. Kermack, W.O., McKendrick, A.G. & McKinlay, P.L. Death-Rates in Great Britain and Sweden. Some General Regularities and their Significance. *Lancet*, 698-703 (1934).
6. Ravelli, G.-P., Stein, Z.A. & Susser, M.W. Obesity in young men after famine exposure in utero and early infancy. *New England Journal of Medicine* **295**, 349-353 (1976).
7. Barker, D., Osmond, C., Winter, P.D., Margetts, B. & Simmonds, S. Weight in infancy and death from ischaemic heart disease. *The Lancet* **334**, 577-580 (1989).
8. Felitti, V.J. *et al.* Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) Study. *American journal of preventive medicine* **14**, 245-258 (1998).
9. Gluckman, P.D. & Hanson, M.A. Living with the past: evolution, development, and patterns of disease. *Science* **305**, 1733-1736 (2004).
10. Hughes, K. *et al.* The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. *Lancet Public Health* **2**, e356-e366 (2017).
11. Nelson, C.A. *et al.* Adversity in childhood is linked to mental and physical health throughout life. *Bmj* **371**, m3048 (2020).
12. Green, J.G. *et al.* Childhood Adversities and Adult Psychiatric Disorders in the National Comorbidity Survey Replication I: Associations With First Onset of DSM-IV Disorders. *Archives of General Psychiatry* **67**, 113-123 (2010).
13. Edwards, V.J., Holden, G.W., Felitti, V.J. & Anda, R.F. Relationship between multiple forms of childhood maltreatment and adult mental health in community respondents: results from the adverse childhood experiences study. *Am J Psychiatry* **160**, 1453-60 (2003).
14. Dube, S.R. *et al.* Childhood abuse, household dysfunction, and the risk of attempted suicide throughout the life span: findings from the Adverse Childhood Experiences Study. *Jama* **286**, 3089-96 (2001).
15. Dube, S.R., Anda, R.F., Felitti, V.J., Edwards, V.J. & Croft, J.B. Adverse childhood experiences and personal alcohol abuse as an adult. *Addict Behav* **27**, 713-25 (2002).
16. Dube, S.R. *et al.* Childhood abuse, neglect, and household dysfunction and the risk of illicit drug use: the adverse childhood experiences study. *Pediatrics* **111**, 564-72 (2003).
17. O'Donnell, K.J. & Meaney, M.J. Fetal Origins of Mental Health: The Developmental Origins of Health and Disease Hypothesis. *Am J Psychiatry* **174**, 319-328 (2017).
18. Smith, K.E. & Pollak, S.D. Early life stress and development: potential mechanisms for adverse outcomes. *Journal of Neurodevelopmental Disorders* **12**, 34 (2020).

19. Victorino, C.C. & Gauthier, A.H. The social determinants of child health: variations across health outcomes – a population-based cross-sectional analysis. *BMC Pediatrics* **9**, 53 (2009).
20. Shonkoff, J.P., Boyce, W.T. & McEwen, B.S. Neuroscience, Molecular Biology, and the Childhood Roots of Health Disparities: Building a New Framework for Health Promotion and Disease Prevention. *JAMA* **301**, 2252-2259 (2009).
21. Pietrek, C., Elbert, T., Weierstall, R., Müller, O. & Rockstroh, B. Childhood adversities in relation to psychiatric disorders. *Psychiatry Research* **206**, 103-110 (2013).
22. Weber, K. *et al.* Stress load during childhood affects psychopathology in psychiatric patients. *BMC Psychiatry* **8**, 63 (2008).
23. McEwen, B.S. Mood disorders and allostatic load. *Biol Psychiatry* **54**, 200-7 (2003).
24. Garner, A., Yogman, M., Child, C.o.P.A.o. & Health, F. Preventing childhood toxic stress: partnering with families and communities to promote relational health. *Pediatrics* **148**(2021).
25. Shonkoff, J.P. *et al.* Translating the Biology of Adversity and Resilience Into New Measures for Pediatric Practice. *Pediatrics* **149**(2022).
26. Shonkoff, J.P., Boyce, W.T., Levitt, P., Martinez, F.D. & McEwen, B. Leveraging the Biology of Adversity and Resilience to Transform Pediatric Practice. *Pediatrics* **147**(2021).
27. Nelson, C.A., 3rd & Gabard-Durnam, L.J. Early Adversity and Critical Periods: Neurodevelopmental Consequences of Violating the Expectable Environment. *Trends Neurosci* **43**, 133-143 (2020).
28. Knudsen, E.I. Sensitive periods in the development of the brain and behavior. *J Cogn Neurosci* **16**, 1412-25 (2004).
29. Nelson, C.A., 3rd, Zeanah, C.H. & Fox, N.A. How Early Experience Shapes Human Development: The Case of Psychosocial Deprivation. *Neural Plast* **2019**, 1676285 (2019).
30. Wade, M., Wright, L. & Finegold, K.E. The effects of early life adversity on children's mental health and cognitive functioning. *Translational Psychiatry* **12**, 244 (2022).
31. Kessler, R.C. *et al.* Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *Br J Psychiatry* **197**, 378-85 (2010).
32. Sheridan, M.A. & McLaughlin, K.A. Dimensions of early experience and neural development: deprivation and threat. *Trends Cogn Sci* **18**, 580-585 (2014).
33. Gardner, M.J., Thomas, H.J. & Erskine, H.E. The association between five forms of child maltreatment and depressive and anxiety disorders: A systematic review and meta-analysis. *Child Abuse Negl* **96**, 104082 (2019).
34. Humphreys, K.L. *et al.* Child maltreatment and depression: A meta-analysis of studies using the Childhood Trauma Questionnaire. *Child Abuse Negl* **102**, 104361 (2020).
35. Liu, R.T., Scopelliti, K.M., Pittman, S.K. & Zamora, A.S. Childhood maltreatment and non-suicidal self-injury: a systematic review and meta-analysis. *Lancet Psychiatry* **5**, 51-64 (2018).
36. Struck, N. *et al.* Childhood maltreatment and adult mental disorders - the prevalence of different types of maltreatment and associations with age of onset and severity of symptoms. *Psychiatry Res* **293**, 113398 (2020).
37. Angelakis, I., Austin, J.L. & Gooding, P. Association of Childhood Maltreatment With Suicide Behaviors Among Young People: A Systematic Review and Meta-analysis. *JAMA Netw Open* **3**, e2012563 (2020).
38. van Bodegom, M., Homberg, J.R. & Henckens, M. Modulation of the Hypothalamic-Pituitary-Adrenal Axis by Early Life Stress Exposure. *Front Cell Neurosci* **11**, 87 (2017).
39. Maniam, J., Antoniadis, C. & Morris, M.J. Early-Life Stress, HPA Axis Adaptation, and Mechanisms Contributing to Later Health Outcomes. *Front Endocrinol (Lausanne)* **5**, 73 (2014).
40. Weaver, I.C.G. *et al.* Epigenetic programming by maternal behavior. *Nature neuroscience* **7**, 847 (2004).

41. McGowan, P.O. *et al.* Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nature Neuroscience* **12**, 342-348 (2009).
42. Nakama, N., Usui, N., Doi, M. & Shimada, S. Early life stress impairs brain and mental development during childhood increasing the risk of developing psychiatric disorders. *Prog Neuropsychopharmacol Biol Psychiatry* **126**, 110783 (2023).
43. Apaydin, D.C. *et al.* Early-Life Stress Regulates Cardiac Development through an IL-4-Glucocorticoid Signaling Balance. *Cell Rep* **33**, 108404 (2020).
44. Murphy, M.O., Cohn, D.M. & Loria, A.S. Developmental origins of cardiovascular disease: Impact of early life stress in humans and rodents. *Neuroscience & Biobehavioral Reviews* **74**, 453-465 (2017).
45. Mondelli, V. & Vernon, A.C. From early adversities to immune activation in psychiatric disorders: the role of the sympathetic nervous system. *Clinical and Experimental Immunology* **197**, 319-328 (2019).
46. Elwenspoek, M.M.C., Kuehn, A., Muller, C.P. & Turner, J.D. The effects of early life adversity on the immune system. *Psychoneuroendocrinology* **82**, 140-154 (2017).
47. de Mendonça Filho, E.J. *et al.* Linking specific biological signatures to different childhood adversities: findings from the HERO project. *Pediatr Res* **94**, 564-574 (2023).
48. Kaseva, N. *et al.* Cardiometabolic health in adults born with very low birth weight—a sibling study. *Pediatric Research*, 1-9 (2023).
49. Malik, S. & Spencer, S.J. Early life stress and metabolism. *Current Opinion in Behavioral Sciences* **28**, 25-30 (2019).
50. McEwen, B.S. Physiology and Neurobiology of Stress and Adaptation: Central Role of the Brain. *Physiological Reviews* **87**, 873-904 (2007).
51. Chrousos, G.P. & Gold, P.W. The concepts of stress and stress system disorders. Overview of physical and behavioral homeostasis. *Jama* **267**, 1244-52 (1992).
52. Liu, D. *et al.* Maternal care, hippocampal glucocorticoid receptors, and hypothalamic-pituitary-adrenal responses to stress. *Science* **277**, 1659-62 (1997).
53. FRANCIS, D. *et al.* The Role of Early Environmental Events in Regulating Neuroendocrine Development: Moms, Pups, Stress, and Glucocorticoid Receptors. *Annals of the New York Academy of Sciences* **794**, 136-152 (1996).
54. Francis, D., Diorio, J., Liu, D. & Meaney, M.J. Nongenomic transmission across generations of maternal behavior and stress responses in the rat. *Science* **286**, 1155-8 (1999).
55. Meaney, M.J. & Szyf, M. Environmental programming of stress responses through DNA methylation: life at the interface between a dynamic environment and a fixed genome. *Dialogues Clin Neurosci* **7**, 103-23 (2005).
56. Barnes, P.J. Anti-inflammatory actions of glucocorticoids: molecular mechanisms. *Clin Sci (Lond)* **94**, 557-72 (1998).
57. Aas, M. *et al.* Childhood maltreatment severity is associated with elevated C-reactive protein and body mass index in adults with schizophrenia and bipolar diagnoses. *Brain Behav Immun* **65**, 342-349 (2017).
58. Hepgul, N. *et al.* Childhood maltreatment is associated with increased body mass index and increased C-reactive protein levels in first-episode psychosis patients. *Psychological medicine* **42**, 1893-1901 (2012).
59. Grosse, L. *et al.* Cytokine levels in major depression are related to childhood trauma but not to recent stressors. *Psychoneuroendocrinology* **73**, 24-31 (2016).
60. Munjiza, A. *et al.* Higher concentration of interleukin 6 - A possible link between major depressive disorder and childhood abuse. *Psychiatry Res* **264**, 26-30 (2018).

61. Baumeister, D., Akhtar, R., Ciufolini, S., Pariante, C.M. & Mondelli, V. Childhood trauma and adulthood inflammation: a meta-analysis of peripheral C-reactive protein, interleukin-6 and tumour necrosis factor-α. *Mol Psychiatry* **21**, 642-9 (2016).
62. Rasmussen, L.J.H. *et al.* Cumulative childhood risk is associated with a new measure of chronic inflammation in adulthood. *J Child Psychol Psychiatry* **60**, 199-208 (2019).
63. Reul, J.M. & de Kloet, E.R. Two receptor systems for corticosterone in rat brain: microdistribution and differential occupation. *Endocrinology* **117**, 2505-11 (1985).
64. Kempermann, G., Song, H. & Gage, F.H. Neurogenesis in the Adult Hippocampus. *Cold Spring Harb Perspect Biol* **7**, a018812 (2015).
65. Warner-Schmidt, J.L. & Duman, R.S. Hippocampal neurogenesis: opposing effects of stress and antidepressant treatment. *Hippocampus* **16**, 239-49 (2006).
66. McEwen, B.S. Stress and hippocampal plasticity. *Annu Rev Neurosci* **22**, 105-22 (1999).
67. Czéh, B. *et al.* Stress-induced changes in cerebral metabolites, hippocampal volume, and cell proliferation are prevented by antidepressant treatment with tianeptine. *Proc Natl Acad Sci U S A* **98**, 12796-801 (2001).
68. Sapolsky, R.M. Depression, antidepressants, and the shrinking hippocampus. *Proc Natl Acad Sci U S A* **98**, 12320-2 (2001).
69. Derks, N.A., Krugers, H.J., Hoogenraad, C.C., Joëls, M. & Sarabdjitsingh, R.A. Effects of Early Life Stress on Synaptic Plasticity in the Developing Hippocampus of Male and Female Rats. *PLoS One* **11**, e0164551 (2016).
70. Abbink, M.R., Naninck, E.F.G., Lucassen, P.J. & Korosi, A. Early-life stress diminishes the increase in neurogenesis after exercise in adult female mice. *Hippocampus* **27**, 839-844 (2017).
71. Schoenfeld, T.J. & Gould, E. Differential effects of stress and glucocorticoids on adult neurogenesis. *Curr Top Behav Neurosci* **15**, 139-64 (2013).
72. Arango-Lievano, M. *et al.* Neurotrophic-priming of glucocorticoid receptor signaling is essential for neuronal plasticity to stress and antidepressant treatment. *Proceedings of the National Academy of Sciences* **112**, 15737-15742 (2015).
73. Dallman, M.F. Stress-induced obesity and the emotional nervous system. *Trends Endocrinol Metab* **21**, 159-65 (2010).
74. Adam, T.C. & Epel, E.S. Stress, eating and the reward system. *Physiol Behav* **91**, 449-58 (2007).
75. Gibson, E.L. Emotional influences on food choice: sensory, physiological and psychological pathways. *Physiol Behav* **89**, 53-61 (2006).
76. Epel, E., Lapidus, R., McEwen, B. & Brownell, K. Stress may add bite to appetite in women: a laboratory study of stress-induced cortisol and eating behavior. *Psychoneuroendocrinology* **26**, 37-49 (2001).
77. Maniam, J. & Morris, M.J. The link between stress and feeding behaviour. *Neuropharmacology* **63**, 97-110 (2012).
78. Joung, K.E. *et al.* Early life adversity is associated with elevated levels of circulating leptin, irisin, and decreased levels of adiponectin in midlife adults. *J Clin Endocrinol Metab* **99**, E1055-60 (2014).
79. Shin, S. *et al.* Early adversity promotes binge-like eating habits by remodeling a leptin-responsive lateral hypothalamus–brainstem pathway. *Nature Neuroscience* **26**, 79-91 (2023).
80. Daniels, T.E. *et al.* Associations of early life stress with leptin and ghrelin in healthy young adults. *Psychoneuroendocrinology* **149**, 106007 (2023).
81. Yam, K.Y. *et al.* Ghrelin and hypothalamic NPY/AgRP expression in mice are affected by chronic early-life stress exposure in a sex-specific manner. *Psychoneuroendocrinology* **86**, 73-77 (2017).
82. Kullmann, S. *et al.* Central nervous pathways of insulin action in the control of metabolism and food intake. *Lancet Diabetes Endocrinol* **8**, 524-534 (2020).

83. Murray, S., Tulloch, A., Gold, M.S. & Avena, N.M. Hormonal and neural mechanisms of food reward, eating behaviour and obesity. *Nature Reviews Endocrinology* **10**, 540-552 (2014).
84. van der Valk, E.S., Savas, M. & van Rossum, E.F.C. Stress and Obesity: Are There More Susceptible Individuals? *Curr Obes Rep* **7**, 193-203 (2018).
85. Kivimäki, M., Bartolomucci, A. & Kawachi, I. The multiple roles of life stress in metabolic disorders. *Nature Reviews Endocrinology* **19**, 10-27 (2023).
86. Russell, G. & Lightman, S. The human stress response. *Nature Reviews Endocrinology* **15**, 525-534 (2019).
87. Olefsky, J.M. & Glass, C.K. Macrophages, inflammation, and insulin resistance. *Annu Rev Physiol* **72**, 219-46 (2010).
88. Rehman, K. & Akash, M.S.H. Mechanisms of inflammatory responses and development of insulin resistance: how are they interlinked? *Journal of Biomedical Science* **23**, 87 (2016).
89. Huang, H. *et al.* Adverse childhood experiences and risk of type 2 diabetes: A systematic review and meta-analysis. *Metabolism* **64**, 1408-18 (2015).
90. Ilchmann-Diounou, H. *et al.* Early life stress induces type 2 diabetes-like features in ageing mice. *Brain, behavior, and immunity* **80**, 452-463 (2019).
91. Portha, B., Chavey, A. & Movassat, J. Early-life origins of type 2 diabetes: fetal programming of the beta-cell mass. *Exp Diabetes Res* **2011**, 105076 (2011).
92. Ashorn, P. *et al.* Small vulnerable newborns-big potential for impact. *Lancet* **401**, 1692-1706 (2023).
93. de Mendonça, E., de Lima Macêna, M., Bueno, N.B., de Oliveira, A.C.M. & Mello, C.S. Premature birth, low birth weight, small for gestational age and chronic non-communicable diseases in adult life: A systematic review with meta-analysis. *Early Hum Dev* **149**, 105154 (2020).
94. Wang, S.F. *et al.* Birth weight and risk of coronary heart disease in adults: a meta-analysis of prospective cohort studies. *J Dev Orig Health Dis* **5**, 408-19 (2014).
95. Mathewson, K.J. *et al.* Mental health of extremely low birth weight survivors: A systematic review and meta-analysis. *Psychol Bull* **143**, 347-383 (2017).
96. Lawn, J.E. *et al.* Small babies, big risks: global estimates of prevalence and mortality for vulnerable newborns to accelerate change and improve counting. *The Lancet* **401**, 1707-1719 (2023).
97. Hales, C.N. & Barker, D.J. Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. *Diabetologia* **35**, 595-601 (1992).
98. Mi, D., Fang, H., Zhao, Y. & Zhong, L. Birth weight and type 2 diabetes: A meta-analysis. *Exp Ther Med* **14**, 5313-5320 (2017).
99. Garofano, A., Czernichow, P. & Bréant, B. In utero undernutrition impairs rat beta-cell development. *Diabetologia* **40**, 1231-4 (1997).
100. Gatford, K.L. *et al.* Impaired β-Cell Function and Inadequate Compensatory Increases in β-Cell Mass after Intrauterine Growth Restriction in Sheep. *Endocrinology* **149**, 5118-5127 (2008).
101. Cook, J.T. *et al.* Association of low birth weight with beta cell function in the adult first degree relatives of non-insulin dependent diabetic subjects. *Bmj* **306**, 302-6 (1993).
102. Jabary, M. *et al.* Fetal growth restriction followed by early catch-up growth impairs pancreatic islet morphology in male rats. *Scientific Reports* **13**, 2732 (2023).
103. Iñiguez, G. *et al.* Longitudinal changes in insulin-like growth factor-I, insulin sensitivity, and secretion from birth to age three years in small-for-gestational-age children. *J Clin Endocrinol Metab* **91**, 4645-9 (2006).
104. Mericq, V. *et al.* Longitudinal changes in insulin sensitivity and secretion from birth to age three years in small-and appropriate-for-gestational-age children. *Diabetologia* **48**, 2609-2614 (2005).

105. Nobili, V., Alisi, A., Panera, N. & Agostoni, C. Low birth weight and catch-up-growth associated with metabolic syndrome: a ten year systematic review. *Pediatr Endocrinol Rev* **6**, 241-7 (2008).
106. Hack, M., Merkatz, I.R., McGrath, S.K., Jones, P.K. & Fanaroff, A.A. Catch-up Growth in Very-Low-Birth-Weight Infants: Clinical Correlates. *American Journal of Diseases of Children* **138**, 370-375 (1984).
107. Cianfarani, S., Germani, D. & Branca, F. Low birthweight and adult insulin resistance: the "catch-up growth" hypothesis. *Archives of Disease in Childhood - Fetal and Neonatal Edition* **81**, F71-F73 (1999).
108. Soto, N.s. *et al.* Insulin sensitivity and secretion are related to catch-up growth in small-for-gestational-age infants at age 1 year: results from a prospective cohort. *The Journal of Clinical Endocrinology Metabolism* **88**, 3645-3650 (2003).
109. Li, C., McDonald, T.J., Wu, G., Nijland, M.J. & Nathanielsz, P.W. Intrauterine growth restriction alters term fetal baboon hypothalamic appetitive peptide balance. *J Endocrinol* **217**, 275-82 (2013).
110. Silveira, P.P. *et al.* Birth weight and catch up growth are associated with childhood impulsivity in two independent cohorts. *Scientific reports* **8**(2018).
111. Batra, A., Latsko, M., Portella, A.K. & Silveira, P.P. Early adversity and insulin: neuroendocrine programming beyond glucocorticoids. *Trends Endocrinol Metab* **32**, 1031-1043 (2021).
112. Pomytkin, I. *et al.* Insulin receptor in the brain: Mechanisms of activation and the role in the CNS pathology and treatment. *CNS neuroscience & therapeutics* **24**, 763-774 (2018).
113. Hill, J.M., Lesniak, M.A., Pert, C.B. & Roth, J. Autoradiographic localization of insulin receptors in rat brain: prominence in olfactory and limbic areas. *Neuroscience* **17**, 1127-38 (1986).
114. Kullmann, S. *et al.* Brain Insulin Resistance at the Crossroads of Metabolic and Cognitive Disorders in Humans. *Physiol Rev* **96**, 1169-209 (2016).
115. Ghasemi, R. *et al.* Brain insulin dysregulation: implication for neurological and neuropsychiatric disorders. *Mol Neurobiol* **47**, 1045-65 (2013).
116. Shonkoff, J.P. *et al.* The Lifelong Effects of Early Childhood Adversity and Toxic Stress. *Pediatrics* **129**, e232-e246 (2012).
117. Bellis, M.A. *et al.* Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: a systematic review and meta-analysis. *Lancet Public Health* **4**, e517-e528 (2019).
118. Nurius, P.S., Green, S., Logan-Greene, P. & Borja, S. Life course pathways of adverse childhood experiences toward adult psychological well-being: A stress process analysis. *Child Abuse Negl* **45**, 143-53 (2015).
119. Loret de Mola, C., de França, G.V., Quevedo Lde, A. & Horta, B.L. Low birth weight, preterm birth and small for gestational age association with adult depression: systematic review and meta-analysis. *Br J Psychiatry* **205**, 340-7 (2014).
120. Alastalo, H. *et al.* Early life stress and blood pressure levels in late adulthood. *J Hum Hypertens* **27**, 90-4 (2013).
121. Van Dijk, A.E., Van Eijsden, M., Stronks, K., Gemke, R.J. & Vrijkotte, T.G. The association between prenatal psychosocial stress and blood pressure in the child at age 5–7 years. *PLoS One* **7**(2012).
122. Rich-Edwards, J.W. *et al.* Longitudinal study of birth weight and adult body mass index in predicting risk of coronary heart disease and stroke in women. *Bmj* **330**, 1115 (2005).
123. Reid, B.M. *et al.* Early Life Adversity with Height Stunting Is Associated with Cardiometabolic Risk in Adolescents Independent of Body Mass Index. *J Pediatr* **202**, 143-149 (2018).
124. Lee, T.K., Wickrama, K.A.S. & O'Neal, C.W. How Early Stressful Life Experiences Combine With Adolescents' Conjoint Health Risk Trajectories to Influence Cardiometabolic Disease Risk in Young Adulthood. *J Youth Adolesc* **50**, 1234-1253 (2021).

125. Winning, A., Glymour, M.M., McCormick, M.C., Gilsanz, P. & Kubzansky, L.D. Childhood Psychological Distress as a Mediator in the Relationship Between Early-Life Social Disadvantage and Adult Cardiometabolic Risk: Evidence From the 1958 British Birth Cohort. *Psychosom Med* **78**, 1019-1030 (2016).
126. Miller, N.E. & Lacey, R.E. Childhood adversity and cardiometabolic biomarkers in mid-adulthood in the 1958 British birth cohort. *SSM Popul Health* **19**, 101260 (2022).
127. Tian, Y.E. *et al.* Evaluation of Brain-Body Health in Individuals With Common Neuropsychiatric Disorders. *JAMA psychiatry* (2023).
128. Souama, C. *et al.* Depression, cardiometabolic disease, and their co-occurrence after childhood maltreatment: an individual participant data meta-analysis including over 200,000 participants. *BMC medicine* **21**, 93 (2023).
129. Vos, T. *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* **388**, 1545-1602 (2016).
130. Kyu, H.H. *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* **392**, 1859-1922 (2018).
131. Egede, L.E. Major depression in individuals with chronic medical disorders: prevalence, correlates and association with health resource utilization, lost productivity and functional disability. *Gen Hosp Psychiatry* **29**, 409-16 (2007).
132. Griffith, L.E. *et al.* Insights on multimorbidity and associated health service use and costs from three population-based studies of older adults in Ontario with diabetes, dementia and stroke. *BMC Health Services Research* **19**, 313 (2019).
133. Baldwin, J.R. *et al.* Population vs Individual Prediction of Poor Health From Results of Adverse Childhood Experiences Screening. *JAMA Pediatrics* **175**, 385-393 (2021).
134. Danese, A. & Lewis, S.J. New directions in research on childhood adversity. *The British Journal of Psychiatry* **220**, 107-108 (2022).
135. Meehan, A.J., Baldwin, J.R., Lewis, S.J., MacLeod, J.G. & Danese, A. Poor Individual Risk Classification From Adverse Childhood Experiences Screening. *Am J Prev Med* **62**, 427-432 (2022).
136. Belsky, J. & Pluess, M. Beyond diathesis stress: differential susceptibility to environmental influences. *Psychol Bull* **135**, 885-908 (2009).
137. Silveira, P.P. & Meaney, M.J. Examining the biological mechanisms of human mental disorders resulting from gene-environment interdependence using novel functional genomic approaches. *Neurobiology of Disease* **178**, 106008 (2023).
138. Drumm, M.L. *et al.* Chloride conductance expressed by delta F508 and other mutant CFTRs in Xenopus oocytes. *Science* **254**, 1797-9 (1991).
139. Dalemans, W. *et al.* Altered chloride ion channel kinetics associated with the delta F508 cystic fibrosis mutation. *Nature* **354**, 526-8 (1991).
140. Sahu, M. & Prasuna, J.G. Twin Studies: A Unique Epidemiological Tool. *Indian J Community Med* **41**, 177-82 (2016).
141. Boomsma, D., Busjahn, A. & Peltonen, L. Classical twin studies and beyond. *Nature Reviews Genetics* **3**, 872-882 (2002).
142. Wray, N.R., Goddard, M.E. & Visscher, P.M. Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Res* **17**, 1520-8 (2007).
143. Wu, M.C. *et al.* Powerful SNP-set analysis for case-control genome-wide association studies. *The American Journal of Human Genetics* **86**, 929-942 (2010).

144. Bush, W.S. & Moore, J.H. Genome-wide association studies. *PLoS Comput Biol* **8**, e1002822 (2012).
145. Tam, V. *et al.* Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics* **20**, 467-484 (2019).
146. Janssens, A.C.J., Gwinn, M., Subramonia-Iyer, S. & Khoury, M.J. Does genetic testing really improve the prediction of future type 2 diabetes? *PLoS Med* **3**, e114 (2006).
147. Manolio, T.A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747-753 (2009).
148. Manuck, S.B. & McCaffery, J.M. Gene-Environment Interaction. *Annual Review of Psychology* **65**, 41-70 (2014).
149. Geus, E.J.C.d. Mendelian Randomization Supports a Causal Effect of Depression on Cardiovascular Disease as the Main Source of Their Comorbidity. *Journal of the American Heart Association* **10**, e019861 (2021).
150. Bookman, E.B. *et al.* Gene-environment interplay in common complex diseases: forging an integrative model—recommendations from an NIH workshop. *Genetic Epidemiology* **35**, 217-225 (2011).
151. Dempfle, A. *et al.* Gene–environment interactions for complex traits: definitions, methodological requirements and challenges. *European Journal of Human Genetics* **16**, 1164-1172 (2008).
152. Virolainen, S.J., VonHandorf, A., Viel, K.C.M.F., Weirauch, M.T. & Kottyan, L.C. Gene–environment interactions and their impact on human health. *Genes & Immunity* **24**, 1-11 (2023).
153. Meaney, M.J. Epigenetics and the biological definition of gene× environment interactions. *Child development* **81**, 41-79 (2010).
154. Ottman, R. Gene-environment interaction: definitions and study designs. *Prev Med* **25**, 764-70 (1996).
155. Belsky, J. *et al.* Vulnerability genes or plasticity genes? *Molecular psychiatry* **14**, 746 (2009).
156. Belsky, J. & Hartman, S. Gene-environment interaction in evolutionary perspective: differential susceptibility to environmental influences. *World Psychiatry* **13**, 87 (2014).
157. Ellis, B.J., Boyce, W.T., Belsky, J., Bakermans-Kranenburg, M.J. & Van IJzendoorn, M.H. Differential susceptibility to the environment: An evolutionary–neurodevelopmental theory. *Development and psychopathology* **23**, 7-28 (2011).
158. Gatzke-Kopp, L.M. The canary in the coalmine: The sensitivity of mesolimbic dopamine to environmental adversity during development. *Neuroscience & Biobehavioral Reviews* **35**, 794-803 (2011).
159. Hosang, G.M., Shiles, C., Tansey, K.E., McGuffin, P. & Uher, R. Interaction between stress and the BDNF Val66Met polymorphism in depression: a systematic review and meta-analysis. *BMC Med* **12**, 7 (2014).
160. Appel, K. *et al.* Moderation of adult depression by a polymorphism in the FKBP5 gene and childhood physical abuse in the general population. *Neuropsychopharmacology* **36**, 1982-91 (2011).
161. Dalmaz, C. *et al.* Prefrontal cortex VAMP1 gene network moderates the effect of the early environment on cognitive flexibility in children. *Neurobiology of Learning and Memory*, 107509 (2021).
162. Miguel, P.M. *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* **86**, 621-630 (2019).
163. de Mendonça Filho, E.J. *et al.* Cognitive development and brain gray matter susceptibility to prenatal adversities: Moderation by the prefrontal cortex brain-derived neurotrophic factor gene co-expression network. *Frontiers in Neuroscience* **15**, 744743 (2021).

164. de Lima, R.M.S. *et al.* Leptin receptor co-expression gene network moderates the effect of early life adversity on eating behavior in children. *Communications Biology* **5**, 1092 (2022).
165. Levitan, R.D. *et al.* A DRD 4 gene by maternal sensitivity interaction predicts risk for overweight or obesity in two independent cohorts of preschool children. *Journal of Child Psychology Psychiatry* **58**, 180-188 (2017).
166. Walter, S., Mejía-Guevara, I., Estrada, K., Liu, S.Y. & Glymour, M.M. Association of a Genetic Risk Score With Body Mass Index Across Different Birth Cohorts. *Jama* **316**, 63-9 (2016).
167. Miguel, P.M., Pereira, L.O., Silveira, P.P. & Meaney, M.J. Early environmental influences on the development of children's brain structure and function. *Developmental Medicine & Child Neurology* **61**, 1127-1133 (2019).
168. Tottenham, N. *et al.* Prolonged institutional rearing is associated with atypically large amygdala volume and difficulties in emotion regulation. *Developmental science* **13**, 46-61 (2010).
169. Tottenham, N. & Sheridan, M.A. A review of adversity, the amygdala and the hippocampus: a consideration of developmental timing. *Frontiers in human neuroscience* **3**, 1019 (2010).
170. Peña, C.J. *et al.* Early life stress confers lifelong stress susceptibility in mice via ventral tegmental area OTX2. *Science* **356**, 1185-1188 (2017).
171. Evans, G.W. *et al.* Childhood Cumulative Risk Exposure and Adult Amygdala Volume and Function. *J Neurosci Res* **94**, 535-43 (2016).
172. Hanson, J.L. *et al.* Behavioral problems after early life stress: contributions of the hippocampus and amygdala. *Biol Psychiatry* **77**, 314-23 (2015).
173. Dalle Molle, R. *et al.* Thrifty-Eating Behavior Phenotype at the Food Court - Programming Goes Beyond Food Preferences. *Front Endocrinol (Lausanne)* **13**, 882532 (2022).
174. Mucellini, A.B. *et al.* Diminished insulin sensitivity is associated with altered brain activation to food cues and with risk for obesity - Implications for individuals born small for gestational age. *Appetite* **169**, 105799 (2022).
175. Parade, S.H. *et al.* Stress exposure and psychopathology alter methylation of the serotonin receptor 2A (HTR2A) gene in preschoolers. *Dev Psychopathol* **29**, 1619-1626 (2017).
176. Soga, T., Teo, C.H. & Parhar, I. Genetic and Epigenetic Consequence of Early-Life Social Stress on Depression: Role of Serotonin-Associated Genes. *Front Genet* **11**, 601868 (2020).
177. Caspi, A. *et al.* Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301**, 386-389 (2003).
178. Caspi, A., Hariri, A.R., Holmes, A., Uher, R. & Moffitt, T.E. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *American journal of Psychiatry* **167**, 509-527 (2010).
179. Belsky, J., Zhang, X. & Sayler, K. Differential susceptibility 2.0: Are the same children affected by different experiences and exposures? *Development and Psychopathology* **34**, 1025-1033 (2022).
180. Haberstick, B.C. *et al.* Depression, Stressful Life Events, and the Impact of Variation in the Serotonin Transporter: Findings from the National Longitudinal Study of Adolescent to Adult Health (Add Health). *PLoS One* **11**, e0148373 (2016).
181. Miller, J.M. *et al.* Reported childhood abuse is associated with low serotonin transporter binding in vivo in major depressive disorder. *Synapse* **63**, 565-73 (2009).
182. Ichise, M. *et al.* Effects of early life stress on [11C]DASB positron emission tomography imaging of serotonin transporters in adolescent peer- and mother-reared rhesus monkeys. *J Neurosci* **26**, 4638-43 (2006).
183. Culverhouse, R.C. *et al.* Collaborative meta-analysis finds no evidence of a strong interaction between stress and 5-HTTLPR genotype contributing to the development of depression. *Molecular psychiatry* **23**, 133-142 (2018).

184. Richard Border, M.A. , *et al.* No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression Across Multiple Large Samples. *American Journal of Psychiatry* **176**, 376-387 (2019).
185. Berger, M., Gray, J.A. & Roth, B.L. The expanded biology of serotonin. *Annu Rev Med* **60**, 355-66 (2009).
186. Charnay, Y. & Léger, L. Brain serotonergic circuitries. *Dialogues Clin Neurosci* **12**, 471-87 (2010).
187. Dayer, A. Serotonin-related pathways and developmental plasticity: relevance for psychiatric disorders. *Dialogues Clin Neurosci* **16**, 29-41 (2014).
188. Cases, O. *et al.* Lack of barrels in the somatosensory cortex of monoamine oxidase A-deficient mice: role of a serotonin excess during the critical period. *Neuron* **16**, 297-307 (1996).
189. Persico, A.M. *et al.* Barrel pattern formation requires serotonin uptake by thalamocortical afferents, and not vesicular monoamine release. *J Neurosci* **21**, 6862-73 (2001).
190. Salichon, N. *et al.* Excessive activation of serotonin (5-HT) 1B receptors disrupts the formation of sensory maps in monoamine oxidase a and 5-ht transporter knock-out mice. *J Neurosci* **21**, 884-96 (2001).
191. Lokmane, L. *et al.* Sensory map transfer to the neocortex relies on pretarget ordering of thalamic axons. *Curr Biol* **23**, 810-6 (2013).
192. Lokmane, L. & Garel, S. Map transfer from the thalamus to the neocortex: Inputs from the barrel field. *Seminars in Cell & Developmental Biology* **35**, 147-155 (2014).
193. Sinclair-Wilson, A. *et al.* Plasticity of thalamocortical axons is regulated by serotonin levels modulated by preterm birth. *Proceedings of the National Academy of Sciences* **120**, e2301644120 (2023).
194. Bonnin, A., Torii, M., Wang, L., Rakic, P. & Levitt, P. Serotonin modulates the response of embryonic thalamocortical axons to netrin-1. *Nat Neurosci* **10**, 588-97 (2007).
195. Smit-Rigter, L.A. *et al.* Prenatal fluoxetine exposure induces life-long serotonin 5-HT_3 receptor-dependent cortical abnormalities and anxiety-like behaviour. *Neuropharmacology* **62**, 865-70 (2012).
196. Popa, D., Léna, C., Alexandre, C. & Adrien, J. Lasting syndrome of depression produced by reduction in serotonin uptake during postnatal development: evidence from sleep, stress, and behavior. *J Neurosci* **28**, 3546-54 (2008).
197. Ansorge, M.S., Morelli, E. & Gingrich, J.A. Inhibition of serotonin but not norepinephrine transport during development produces delayed, persistent perturbations of emotional behaviors in mice. *J Neurosci* **28**, 199-207 (2008).
198. Mitchell, S.J. *et al.* Early-life adversity selectively impairs α2-GABA(A) receptor expression in the mouse nucleus accumbens and influences the behavioral effects of cocaine. *Neuropharmacology* **141**, 98-112 (2018).
199. Englund, J. *et al.* Downregulation of kainate receptors regulating GABAergic transmission in amygdala after early life stress is associated with anxiety-like behavior in rodents. *Translational Psychiatry* **11**, 538 (2021).
200. Karst, H. *et al.* Acceleration of GABA-switch after early life stress changes mouse prefrontal glutamatergic transmission. *Neuropharmacology* **234**, 109543 (2023).
201. Bettler, B., Kaupmann, K., Mosbacher, J. & Gassmann, M. Molecular structure and physiological functions of GABA(B) receptors. *Physiol Rev* **84**, 835-67 (2004).
202. Ganguly, K., Schinder, A.F., Wong, S.T. & Poo, M. GABA itself promotes the developmental switch of neuronal GABAergic responses from excitation to inhibition. *Cell* **105**, 521-32 (2001).
203. Sarawagi, A., Soni, N.D. & Patel, A.B. Glutamate and GABA Homeostasis and Neurometabolism in Major Depressive Disorder. *Front Psychiatry* **12**, 637863 (2021).

204. Reiner, A. & Levitz, J. Glutamatergic Signaling in the Central Nervous System: Ionotropic and Metabotropic Receptors in Concert. *Neuron* **98**, 1080-1098 (2018).
205. Averill, L.A. *et al.* Early life stress and glutamate neurotransmission in major depressive disorder. *Eur Neuropsychopharmacol* **35**, 71-80 (2020).
206. Coplan, J.D. *et al.* Maternal Cerebrospinal Fluid Glutamate in Response to Variable Foraging Demand: Relationship to Cerebrospinal Fluid Serotonin Metabolites in Grown Offspring. *Chronic Stress (Thousand Oaks)* **2**(2018).
207. Mathew, S.J. *et al.* A magnetic resonance spectroscopic imaging study of adult nonhuman primates exposed to early-life stressors. *Biol Psychiatry* **54**, 727-35 (2003).
208. Chen, Y. *et al.* Early Life Stress Induces Different Behaviors in Adolescence and Adulthood May Related With Abnormal Medial Prefrontal Cortex Excitation/Inhibition Balance. *Front Neurosci* **15**, 720286 (2021).
209. Marrocco, J. *et al.* Anxiety-Like Behavior of Prenatally Stressed Rats Is Associated with a Selective Reduction of Glutamate Release in the Ventral Hippocampus. *The Journal of Neuroscience* **32**, 17143-17154 (2012).
210. Adrover, E. *et al.* Glutamate neurotransmission is affected in prenatally stressed offspring. *Neurochem Int* **88**, 73-87 (2015).
211. Corcoran, M. *et al.* Are working memory and glutamate concentrations involved in early-life stress and severity of psychosis? *Brain Behav* **10**, e01616 (2020).
212. Guan, J. *et al.* Early Life Stress Increases Brain Glutamate and Induces Neurobehavioral Manifestations in Rats. *ACS Chem Neurosci* **11**, 4169-4178 (2020).
213. Qiu, A. *et al.* Effects of antenatal maternal depressive symptoms and socio-economic status on neonatal brain development are modulated by genetic risk. **27**, 3080-3092 (2017).
214. Eltokhi, A., Santuy, A., Merchan-Perez, A. & Sprengel, R. Glutamatergic Dysfunction and Synaptic Ultrastructural Alterations in Schizophrenia and Autism Spectrum Disorder: Evidence from Human and Rodent Studies. *Int J Mol Sci* **22**(2020).
215. Riedel, G., Platt, B. & Micheau, J. Glutamate receptor function in learning and memory. *Behav Brain Res* **140**, 1-47 (2003).
216. Konradi, C. & Heckers, S. Molecular aspects of glutamate dysregulation: implications for schizophrenia and its treatment. *Pharmacol Ther* **97**, 153-79 (2003).
217. Lee, P.H. *et al.* Multi-locus genome-wide association analysis supports the role of glutamatergic synaptic transmission in the etiology of major depressive disorder. *Transl Psychiatry* **2**, e184 (2012).
218. Goldstein Ferber, S., Trezza, V. & Weller, A. Early life stress and development of the endocannabinoid system: A bidirectional process in programming future coping. *Developmental Psychobiology* **63**, 143-152 (2021).
219. Suárez, J. *et al.* Early maternal deprivation induces gender-dependent changes on the expression of hippocampal CB(1) and CB(2) cannabinoid receptors of neonatal rats. *Hippocampus* **19**, 623-32 (2009).
220. Ceci, C. *et al.* Prenatal corticosterone and adolescent URB597 administration modulate emotionality and CB1 receptor expression in mice. *Psychopharmacology (Berl)* **231**, 2131-44 (2014).
221. Mackie, K. Distribution of cannabinoid receptors in the central and peripheral nervous system. *Handb Exp Pharmacol*, 299-325 (2005).
222. Hill, M.N. *et al.* Endogenous cannabinoid signaling is essential for stress adaptation. *Proc Natl Acad Sci U S A* **107**, 9406-11 (2010).
223. Hill, M.N. *et al.* Functional interactions between stress and the endocannabinoid system: from synaptic signaling to behavioral output. *J Neurosci* **30**, 14980-6 (2010).

224. Valverde, O. & Torrens, M. CB1 receptor-deficient mice as a model for depression. *Neuroscience* **204**, 193-206 (2012).
225. Anda, R.F. *et al.* The enduring effects of abuse and related adverse experiences in childhood. A convergence of evidence from neurobiology and epidemiology. *Eur Arch Psychiatry Clin Neurosci* **256**, 174-86 (2006).
226. Quinn, K. *et al.* The relationships of childhood trauma and adulthood prescription pain reliever misuse and injection drug use. *Drug Alcohol Depend* **169**, 190-198 (2016).
227. Leyrer-Jackson, J.M., Overby, P.F., Nagy, E.K. & Olive, M.F. Early Life Stress Promotes Heroin Seeking But Does Not Alter the Excitability of Insular Pyramidal Cells Targeting the Nucleus Accumbens. *Front Behav Neurosci* **15**, 777826 (2021).
228. Ordoñes Sanchez, E. *et al.* Early life adversity promotes resilience to opioid addiction-related phenotypes in male rats and sex-specific transcriptional changes. *Proc Natl Acad Sci U S A* **118**(2021).
229. Kalinichev, M., Easterling, K.W. & Holtzman, S.G. Early neonatal experience of Long-Evans rats results in long-lasting changes in reactivity to a novel environment and morphine-induced sensitization and tolerance. *Neuropsychopharmacology* **27**, 518-33 (2002).
230. Michaels, C.C. & Holtzman, S.G. Early postnatal stress alters place conditioning to both μ-and κ-opioid agonists. *Journal of pharmacology and experimental therapeutics* **325**, 313-318 (2008).
231. Laureano, D. *et al.* Intrauterine growth restriction modifies the hedonic response to sweet taste in newborn pups–Role of the accumbal μ-opioid receptors. *Neuroscience* **322**, 500-508 (2016).
232. van Steenbergen, H., Eikemo, M. & Leknes, S. The role of the opioid system in decision making and cognitive control: A review. *Cogn Affect Behav Neurosci* **19**, 435-458 (2019).
233. Eduardo, E.B. Endogenous opioid systems. *Neurology* **79**, 807 (2012).
234. Le Merrer, J., Becker, J.A., Befort, K. & Kieffer, B.L. Reward processing by the opioid system in the brain. *Physiol Rev* **89**, 1379-412 (2009).
235. Son, G.H. *et al.* Hyperactivity and alteration of the midbrain dopaminergic system in maternally stressed male mice offspring. *Biochemical and Biophysical Research Communications* **352**, 823-829 (2007).
236. Meehan, C. *et al.* Effects of immune activation during early or late gestation on schizophrenia-related behaviour in adult rat offspring. *Brain, Behavior, and Immunity* **63**, 8-20 (2017).
237. Meyer, U., Engler, A., Weber, L., Schedlowski, M. & Feldon, J. Preliminary evidence for a modulation of fetal dopaminergic development by maternal immune activation during pregnancy. *Neuroscience* **154**, 701-709 (2008).
238. Laureano, D.P. *et al.* Intrauterine growth restriction modifies the accumbal dopaminergic response to palatable food intake. *Neuroscience* (2018).
239. Alves, M.B. *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* **204**, 336-346 (2019).
240. Chocyk, A. *et al.* The impact of maternal separation on the number of tyrosine hydroxylase-expressing midbrain neurons during different stages of ontogenesis. *Neuroscience* **182**, 43-61 (2011).
241. Dalle Molle, R. *et al.* Intrauterine growth restriction increases the preference for palatable foods and affects sensitivity to food rewards in male and female adult rats. *brain research* **1618**, 41-49 (2015).
242. Alves, M.B., Dalle Molle, R., Desai, M., Ross, M.G. & Silveira, P.P. 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* **287**, 73-81 (2015).

243. Rodrigues, A.J., Leão, P., Carvalho, M., Almeida, O.F. & Sousa, N. Potential programming of dopaminergic circuits by early life stress. *Psychopharmacology (Berl)* **214**, 107-20 (2011).
244. Levis, S.C., Baram, T.Z. & Mahler, S.V. Neurodevelopmental origins of substance use disorders: Evidence from animal models of early-life adversity and addiction. *Eur J Neurosci* **55**, 2170-2195 (2022).
245. Hoops, D. & Flores, C. Making dopamine connections in adolescence. *Trends in neurosciences* **40**, 709-719 (2017).
246. Kelley, A.E. & Berridge, K.C. The neuroscience of natural rewards: relevance to addictive drugs. *J Neurosci* **22**, 3306-11 (2002).
247. Schultz, W. Predictive Reward Signal of Dopamine Neurons. *Journal of Neurophysiology* **80**, 1-27 (1998).
248. Birn, R.M., Roeber, B.J. & Pollak, S.D. Early childhood stress exposure, reward pathways, and adult decision making. *Proceedings of the National Academy of Sciences* **114**, 13549-13554 (2017).
249. Novick, A.M. *et al.* The effects of early life stress on reward processing. *Journal of Psychiatric Research* **101**, 80-103 (2018).
250. Herzberg, M.P. & Gunnar, M.R. Early life stress and brain function: Activity and connectivity associated with processing emotion and reward. *NeuroImage* **209**, 116493 (2020).
251. Marusak, H.A., Hatfield, J.R., Thomason, M.E. & Rabinak, C.A. Reduced ventral tegmental area–hippocampal connectivity in children and adolescents exposed to early threat. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **2**, 130-137 (2017).
252. Marshall, N.A. *et al.* Socioeconomic disadvantage and altered corticostriatal circuitry in urban youth. *Human brain mapping* **39**, 1982-1994 (2018).
253. Lee, D., Seo, H. & Jung, M.W. Neural basis of reinforcement learning and decision making. *Annu Rev Neurosci* **35**, 287-308 (2012).
254. Reynolds, L.M. *et al.* DCC receptors drive prefrontal cortex maturation by determining dopamine axon targeting in adolescence. *Biological psychiatry* **83**, 181-192 (2018).
255. Willing, J., Cortes, L.R., Brodsky, J.M., Kim, T. & Juraska, J.M. Innervation of the medial prefrontal cortex by tyrosine hydroxylase immunoreactive fibers during adolescence in male and female rats. *Developmental psychobiology* **59**, 583-589 (2017).
256. Pechtel, P. & Pizzagalli, D.A. Effects of early life stress on cognitive and affective function: an integrated review of human literature. *Psychopharmacology (Berl)* **214**, 55-70 (2011).
257. Bischoff, A.R. *et al.* Low birth weight is associated with increased fat intake in school-aged boys. *British Journal of Nutrition* **119**, 1295-1302 (2018).
258. Silveira, P.P. *et al.* Early life experience alters behavioral responses to sweet food and accumbal dopamine metabolism. *International journal of developmental neuroscience* **28**, 111-118 (2010).
259. Ayres, C. *et al.* Intrauterine growth restriction and the fetal programming of the hedonic response to sweet taste in newborn infants. *International journal of pediatrics* **2012**(2012).
260. Barth, B., Portella, A.K., Dubé, L., Meaney, M.J. & Silveira, P.P. The Interplay Between Dopamine and Environment as the Biological Basis for the Early Origins of Mental Health. in *Early Life Origins of Ageing and Longevity* 121-140 (Springer, 2019).
261. Gragnoli, C., Reeves, G., Reazer, J. & Postolache, T. Dopamine–prolactin pathway potentially contributes to the schizophrenia and type 2 diabetes comorbidity. *Translational psychiatry* **6**, e785-e785 (2016).
262. Nousen, E.K., Franco, J.G. & Sullivan, E.L. Unraveling the mechanisms responsible for the comorbidity between metabolic syndrome and mental health disorders. *Neuroendocrinology* **98**, 254-266 (2013).

263. Alberry, B. & Silveira, P.P. Brain insulin signaling as a potential mediator of early life adversity effects on physical and mental health. *Neuroscience & Biobehavioral Reviews* **153**, 105350 (2023).
264. Khanh, D.V., Choi, Y.H., Moh, S.H., Kinyua, A.W. & Kim, K.W. Leptin and insulin signaling in dopaminergic neurons: relationship between energy balance and reward system. *Front Psychol* **5**, 846 (2014).
265. Ma, P. & Ou, Y. Correlation between the dopaminergic system and inflammation disease: a review. *Molecular Biology Reports* **50**, 7043-7053 (2023).
266. Feng, Y. & Lu, Y. Immunomodulatory Effects of Dopamine in Inflammatory Diseases. *Front Immunol* **12**, 663102 (2021).
267. Levite, M. Dopamine and T cells: dopamine receptors and potent effects on T cells, dopamine production in T cells, and abnormalities in the dopaminergic system in T cells in autoimmune, neurological and psychiatric diseases. *Acta Physiol (Oxf)* **216**, 42-89 (2016).
268. Grace, A.A. Dysregulation of the dopamine system in the pathophysiology of schizophrenia and depression. *Nature Reviews Neuroscience* **17**, 524 (2016).
269. Stahl, S.M. *et al.* A Review of the Neuropharmacology of Bupropion, a Dual Norepinephrine and Dopamine Reuptake Inhibitor. *Prim Care Companion J Clin Psychiatry* **6**, 159-166 (2004).
270. Faraone, S.V. The pharmacology of amphetamine and methylphenidate: Relevance to the neurobiology of attention-deficit/hyperactivity disorder and other psychiatric comorbidities. *Neurosci Biobehav Rev* **87**, 255-270 (2018).
271. Noaín, D. *et al.* Identification of brain neurons expressing the dopamine D4 receptor gene using BAC transgenic mice. *Eur J Neurosci* **24**, 2429-38 (2006).
272. Beaulieu, J.-M. & Gainetdinov, R.R. The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacological reviews* **70**, 110.002642 (2011).
273. Swanson, J.M. *et al.* Dopamine genes and ADHD. *Neurosci Biobehav Rev* **24**, 21-5 (2000).
274. Botticelli, L. *et al.* Underlying Susceptibility to Eating Disorders and Drug Abuse: Genetic and Pharmacological Aspects of Dopamine D4 Receptors. *Nutrients* **12**(2020).
275. Fontana, C. *et al.* DRD4 and SLC6A3 gene polymorphisms are associated with food intake and nutritional status in children in early stages of development. *The journal of nutritional biochemistry* **26**, 1607-1612 (2015).
276. Levitan, R.D. *et al.* A birth-season/DRD4 gene interaction predicts weight gain and obesity in women with seasonal affective disorder: a seasonal thrifty phenotype hypothesis. *Neuropsychopharmacology* **31**, 2498 (2006).
277. Silveira, P.P. *et al.* Genetic Differential Susceptibility to Socioeconomic Status and Childhood Obesogenic Behavior. *JAMA Pediatrics* (2016).
278. Evers, C., Marijn Stok, F. & de Ridder, D.T.D. Feeding Your Feelings: Emotion Regulation Strategies and Emotional Eating. *Personality and Social Psychology Bulletin* **36**, 792-804 (2010).
279. Mullins, N. *et al.* Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. *Psychol Med* **46**, 759-70 (2016).
280. Peyrot, W.J. *et al.* Does Childhood Trauma Moderate Polygenic Risk for Depression? A Meta-analysis of 5765 Subjects From the Psychiatric Genomics Consortium. *Biol Psychiatry* **84**, 138-147 (2018).
281. Trotta, A. *et al.* Interplay between Schizophrenia Polygenic Risk Score and Childhood Adversity in First-Presentation Psychotic Disorder: A Pilot Study. *PLoS One* **11**, e0163319 (2016).
282. Ma'ayan, A. Complex systems biology. *Journal of the Royal Society Interface* **14**, 20170391 (2017).

283. Gaiteri, C., Ding, Y., French, B., Tseng, G.C. & Sibille, E. Beyond modules and hubs: the potential of gene coexpression networks for investigating molecular mechanisms of complex brain disorders. *Genes, Brain, Behavior* **13**, 13-24 (2014).
284. O'Donnell, K.J. & Meaney, M.J. Fetal origins of mental health: the developmental origins of health and disease hypothesis. *American Journal of Psychiatry* **174**, 319-328 (2016).
285. Silveira, P.P., Portella, A.K., Goldani, M.Z. & Barbieri, M.A. Developmental origins of health and disease (DOHaD). *Jornal de pediatria* **83**, 494-504 (2007).
286. Gold, S.M. *et al.* Comorbid depression in medical diseases. *Nature Reviews Disease Primers* **6**, 69 (2020).
287. Penninx, B.W.J.H., Milaneschi, Y., Lamers, F. & Vogelzangs, N. Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile. *BMC Medicine* **11**, 129 (2013).
288. Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M. & Smoller, J.W. Pleiotropy in complex traits: challenges and strategies. *Nature Reviews Genetics* **14**, 483-495 (2013).
289. Sivakumaran, S. *et al.* Abundant pleiotropy in human complex diseases and traits. *The American Journal of Human Genetics* **89**, 607-618 (2011).
290. Kember, R.L. *et al.* Genetic pleiotropy between mood disorders, metabolic, and endocrine traits in a multigenerational pedigree. *Translational psychiatry* **8**, 1-12 (2018).
291. Vaughan, R.A. & Foster, J.D. Mechanisms of dopamine transporter regulation in normal and disease states. *Trends Pharmacol Sci* **34**, 489-96 (2013).
292. Ciliax, B.J. *et al.* Immunocytochemical localization of the dopamine transporter in human brain. *Journal of Comparative Neurology* **409**, 38-56 (1999).
293. Yager, L.M., Garcia, A.F., Wunsch, A.M. & Ferguson, S.M. The ins and outs of the striatum: role in drug addiction. *Neuroscience* **301**, 529-541 (2015).
294. Threlfell, S. *et al.* Striatal Dopamine Transporter Function Is Facilitated by Converging Biology of α-Synuclein and Cholesterol. *Front Cell Neurosci* **15**, 658244 (2021).
295. Davies, A.A., Smith, G.D., Ben-Shlomo, Y. & Litchfield, P. Low birth weight is associated with higher adult total cholesterol concentration in men: findings from an occupational cohort of 25,843 employees. *Circulation* **110**, 1258-62 (2004).
296. Ravelli, A.C., van Der Meulen, J.H., Osmond, C., Barker, D.J. & Bleker, O.P. Obesity at the age of 50 y in men and women exposed to famine prenatally. *Am J Clin Nutr* **70**, 811-6 (1999).
297. Eriksson, J.G. *et al.* Catch-up growth in childhood and death from coronary heart disease: longitudinal study. *Bmj* **318**, 427-31 (1999).
298. Barker, D.J., Osmond, C., Forsén, T.J., Kajantie, E. & Eriksson, J.G. Trajectories of growth among children who have coronary events as adults. *N Engl J Med* **353**, 1802-9 (2005).
299. Vasiliadis, H.M., Buka, S.L., Martin, L.T. & Gilman, S.E. Fetal growth and the lifetime risk of generalized anxiety disorder. *Depress Anxiety* **27**, 1066-72 (2010).
300. Furuta, M. *et al.* Exposure to social defeat stress in adolescence improves the working memory and anxiety-like behavior of adult female rats with intrauterine growth restriction, independently of hippocampal neurogenesis. *Hormones and Behavior* **70**, 30-37 (2015).
301. Bronfenbrenner, U. Toward an experimental ecology of human development. *American psychologist* **32**, 513 (1977).
302. Gruber, J. *et al.* Impact of insulin and insulin resistance on brain dopamine signalling and reward processing - An underexplored mechanism in the pathophysiology of depression? *Neurosci Biobehav Rev* **149**, 105179 (2023).
303. Sullivan, M. *et al.* Insulin and disorders of behavioural flexibility. *Neurosci Biobehav Rev* **150**, 105169 (2023).

304. Floresco, S. Prefrontal dopamine and behavioral flexibility: shifting from an “inverted-U” toward a family of functions. *Frontiers in Neuroscience* **7**(2013).
305. Arcego, D.M. *et al.* A Glucocorticoid-Sensitive Hippocampal Gene Network Moderates the Impact of Early-Life Adversity on Mental Health Outcomes. *Biological Psychiatry* (2023).
306. Wilkinson, P.O., Trzaskowski, M., Haworth, C.M. & Eley, T.C. The role of gene-environment correlations and interactions in middle childhood depressive symptoms. *Dev Psychopathol* **25**, 93-104 (2013).
307. Ziliak, S. P values and the search for significance. *Nature methods* **14**, 3-4 (2017).
308. Gómez-de-Mariscal, E. *et al.* Use of the p-values as a size-dependent function to address practical differences when analyzing large datasets. *Scientific Reports* **11**, 20942 (2021).
309. Cohen, J. Statistical power analysis. *Current directions in psychological science* **1**, 98-101 (1992).
310. Solar, O. & Irwin, A. A conceptual framework for action on the social determinants of health. (WHO Document Production Services, 2010).
311. Hogan, M.C. *et al.* Maternal mortality for 181 countries, 1980–2008: a systematic analysis of progress towards Millennium Development Goal 5. *The lancet* **375**, 1609-1623 (2010).
312. Wilmoth, J., Mathers, C., Say, L. & Mills, S. Maternal deaths drop by one-third from 1990 to 2008: a United Nations analysis. *Bulletin of the World Health Organization* **88**, 718-718a (2010).
313. Sharrow, D. *et al.* Global, regional, and national trends in under-5 mortality between 1990 and 2019 with scenario-based projections until 2030: a systematic analysis by the UN Inter-agency Group for Child Mortality Estimation. *Lancet Glob Health* **10**, e195-e206 (2022).
314. Shonkoff, J.P. Rethinking the Definition of Evidence-Based Interventions to Promote Early Childhood Development. *Pediatrics* **140**(2017).
315. Shonkoff, J.P. & Fisher, P.A. Rethinking evidence-based practice and two-generation programs to create the future of early childhood policy. *Dev Psychopathol* **25**, 1635-53 (2013).
316. Goldstein, J.M., Hale, T., Foster, S.L., Tobet, S.A. & Handa, R.J. Sex differences in major depression and comorbidity of cardiometabolic disorders: impact of prenatal stress and immune exposures. *Neuropsychopharmacology* **44**, 59-70 (2019).
317. Graves Jr, J.L. Great is their sin: Biological determinism in the age of genomics. *The Annals of the American Academy of Political and Social Science* **661**, 24-50 (2015).
318. Zeggini, E., Gloyn, A.L., Barton, A.C. & Wain, L.V. Translational genomics and precision medicine: Moving from the lab to the clinic. *Science* **365**, 1409-1413 (2019).