Cross-species exploration of Inhibitory Control: The Role of Netrin-1/DCC System in Corticolimbic Development

Jose Maria Restrepo-Lozano

Integrated Program in Neuroscience

Faculty of Medicine

McGill University, Montreal, Canada

Submitted September, 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Table of Contents

Abstract

Identifying risk factors before psychiatric disorders manifest is essential for early interventions and has a broad societal impact, especially given that in Canada, nearly one-third of people will experience a mental health disorder in their lifetime. Despite the high prevalence, our current understanding of the biological basis and early indicators of vulnerability to these disorders is limited. Using the Research Domain Criteria (RDoC) framework as a guide, this thesis investigates possible neurobiological mechanisms affecting inhibitory control behaviors in mice and humans. These behaviors, characteristic of conditions such as attention-deficit/hyperactivity disorder (ADHD), obsessive-compulsive disorder (OCD), and substance use disorder (SUD), rely on the interaction between corticolimbic brain regions. I hypothesized that measuring and estimating molecular processes at the tissue level can provide insights, across species, into the ongoing neurodevelopmental processes within brain substrates critical for the development of inhibitory control behaviors.

The present thesis examines the role of the Netrin-1/DCC axon guidance cue signaling system in brain corticolimbic development, characterizing the net effect of variations in this highly-conserved molecular pathway on inhibitory control behaviors in both mice and humans. More specifically, my research is organized around three main objectives: 1) In mice, I examine the effects of therapeutic-like doses of amphetamine during critical neurodevelopmental periods. Our focus is on the Netrin-1/DCC signaling system, which guides brain corticolimbic development. I aim to understand how these doses affect inhibitory control behaviors in adulthood, a timely concern given the widespread use of amphetamine-based treatments; 2) Extending the insights to humans, I propose a new biological marker for inhibitory control in children. This marker is based on DCC gene co-expression networks, initially identified in mice but also evident in human brain samples during early life. I analyze its association with inhibitory control in children aged 6 and 10; and 3) I offer a thorough review of genetic risk assessment

approaches in psychiatry, emphasizing the shift from genotype-disease mapping to a more nuanced gene regulation-phenotype framework.

In the rodent study, I discovered that administering therapeutic-like doses of amphetamine during sensitive neurodevelopmental periods positively influenced adult performance in the Go/No-Go task, an assay measuring inhibitory control. This behavioral outcome was notably different from the detrimental effects observed when rodents were administered recreational-like doses. Moreover, stereological assessments revealed no significant anatomical alterations in key mesocorticolimbic dopamine targets following therapeutic psychostimulant administration. In the human study, I identified a novel polygenic signal associated with impulsivity in children, which was rooted in corticolimbic-specific DCC gene co-expression networks. This polygenic signal associated with increased motor and reflection impulsivity across three independent birth cohorts. Also, enrichment analyses supported the functional importance of these gene networks in the protracted maturation of both the prefrontal cortex and nucleus accumbens.

Together, the results of my thesis suggest that 1) therapeutic doses of amphetamine do not disrupt corticolimbic development or inhibitory control in mice; 2) in humans, lower expression-based polygenic risk scores are associated with higher impulsivity levels in children. This work offers a new method for predicting genetic risk for psychiatric conditions and enhances our understanding of the molecular networks that govern brain development and function. It serves also as a framework for understanding the potential impact of pharmacological interventions during key neurodevelopmental stages.

Résumé

L'identification de facteurs de risque en amont de troubles psychiatriques est essentielle pour une intervention précoce et a un impact sociétal ample, d'autant plus qu'au Canada, près d'un tiers des personnes connaîtront un trouble de la santé mentale au cours de leur vie. Malgré cette prévalence élevée, notre compréhension actuelle des fondements biologiques et des indicateurs précoces de vulnérabilité à ces troubles est limitée. En utilisant le schéma des Critères du Domaine de Recherche (RDoC) comme guide, cette thèse étudie les mécanismes neurobiologiques potentiels affectant le contrôle inhibiteur chez la souris et chez l'humain. Ce processus cognitif, caractéristique de troubles tels que trouble de déficit de l'attention/hyperactivité (TDAH), le trouble obsessionnel-compulsif (TOC) et le trouble lié à l'utilisation de substances psychoactives (TUS), repose sur l'interaction entre les régions corticolimbiques du cerveau. J'ai émis l'hypothèse que la mesure et l'estimation du processus moléculair au niveau des tissus du cerveau peuvent fournir des informations sur le neurodéveloppement essentiel du contrôle inhibiteur, au sein de plusieurs espèces.

Cette thèse examine le rôle du système de signalisation Netrin-1/DCC dans le développement du système corticolimbique, en caractérisant l'effet des variations de cette voie moléculaire hautement conservée sur le contrôle inhibiteur chez la souris et l'humain. Plus précisément, ma recherche s'articule autour de trois objectifs principaux : 1) Chez la souris, j'examine les effets de doses thérapeutiques d'amphétamine pendant des périodes critiques du développement neurologique. Je me concentre sur le système de signalisation Netrin-1/DCC, qui guide le développement du système corticolimbique au sein du cerveau. Je cherche à saisir comment ces doses affectent le contrôle inhibiteur à l'âge adulte, une question d'actualité étant donné l'utilisation répandue des traitements à base d'amphétamines ; 2) En étendant ces connaissances à l'humain, je propose un nouveau marqueur biologique pour le contrôle inhibiteur chez les enfants. Ce marqueur est basé sur les réseaux de co-expression des gènes autour de DCC, initialement identifiés chez la souris mais également observés dans des échantillons de cerveau

humain au cours des premières années de la vie. J'analyse son association avec le contrôle inhibiteur chez des enfants âgés de 6 et 10 ans ; et 3) je propose de revisiter les méthodes d'évaluation du risque génétique en psychiatrie, en mettant en avant la transition de l'approche basée sur la cartographie génétique des maladies vers un modèle plus nuancé qui étudie la régulation génique des phénotypes.

Dans l'étude sur les rongeurs, j'ai observé que l'administration de doses thérapeutiques d'amphétamine pendant une période critique du développement neurologique influençait positivement les performances de l'adulte dans la tâche Go/No-Go, un test mesurant le contrôle inhibiteur. Ces résultats comportementaux étaient sensiblement différents aux effets délétères observés lorsque les rongeurs recevaient des doses récréatives. En outre, les études stéréologiques n'ont révélé aucune altération anatomique significative dans les principales cibles dopaminergiques mésocorticolimbiques après l'administration de psychostimulants à des fins thérapeutiques. Dans l'étude chez l'humain, j'ai identifié un nouveau signal polygénique associé à l'impulsivité chez l'enfant, qui était ancré dans les réseaux de co-expression des gènes de DCC spécifiques au système corticolimbique. Ce signal polygénique est associé à une impulsivité motrice et cognitive accrue dans trois cohortes de naissance indépendantes. En outre, des analyses d'enrichissement ont confirmé l'importance fonctionnelle de ces réseaux de gènes dans la maturation progressive du cortex préfrontal et du noyau accumbens.

Ensemble, les résultats de ma thèse suggèrent que 1) les doses thérapeutiques d'amphétamine ne perturbent pas le développement corticolimbique ou le contrôle inhibiteur chez la souris ; 2) chez l'humain, les risques polygéniques moins élevés sont associés à des niveaux d'impulsivité plus hauts chez l'enfant. Ces travaux offrent une nouvelle méthode pour prédire le risque génétique de troubles psychiatriques; enrichissant notre compréhension des réseaux moléculaires qui régissent le développement et le fonctionnement du cerveau. Ils servent également de cadre à la compréhension de

l'impact potentiel des interventions pharmacologiques à des périodes clés du développement neurologique.

List of Figures

List of Tables

List of Abbreviations

ADHD	Attention/Deficit and Hyperactivity Disorder
ALSPAC	Avon Longitudinal Study on Parents and Children cohort
AMPH	Amphetamine
CSEA	Cell-Type Specific Expression Analysis
DCC	Deleted in Colorectal Cancer
DSM	Diagnostic and Statistical Manual of Mental Disorders
ePRS	Expression-based Polygenic Score
FUMA	Functional Mapping and Annotation of Genome-Wide Association Studies
GTEx	Genome-Tissue Expression project
GUSTO	Growing Up in Singapore Towards Healthy Outcomes cohort
GWAS	Genome-Wide Association Study
ICD	International Classification of Disorders
MAVAN	Maternal Adversity, Vulnerability and Neurodevelopment cohort
MDD	Major Depressive Disorder
NAcc	Nucleus Accumbens
NCBI	National Center for Biotechnology Information
PCR	Polymerase-chain reaction
PFC	Prefrontal Cortex
PPI	Protein-Protein Interaction
PRS	Polygenic Risk Score
PRSoS	Polygenic Risk Score on Spark pipeline
RDoC	Research Domain Criteria
SIAH	Seven in absentia homologue
SNP	Single-nucleotide polymorphism
SSRT	Stop-Signal Reaction Time
SUD	Substance Use Disorder
TH	Tyrosine Hydroxylase
VTA	Ventral Tegmental Area

Acknowledgements

This work was supported by scholarships provided by the Fonds de Recherche en Santé du Québec, Canadian Institutes of Health Research, and the McGill University Faculty of Medicine.

First and foremost, I would like to extend my deepest gratitude to my advisors, Dr. Cecilia Flores and Dr. Patricia Pelufo Silveira. Their invaluable guidance, unwavering support, and insightful critiques throughout this journey have been instrumental in shaping both this research and my growth as a scientist. Their passion for neuroscience and their commitment to advancing the field have inspired me each day.

To my committee members, Dr. Andreas Arvanitogiannis and Dr. Frederic Charron, thank you for your time, patience, and invaluable feedback. Your critical insights and thought-provoking questions have not only enriched the quality of this work but have also broadened my perspectives and deepened my understanding of the nuances of our research domain.

My appreciation also extends to the entire team at the Flores, Silveira, and Meaney Labs. Each one of you has contributed, in ways big and small, to this endeavor. I am grateful for the stimulating discussions and the camaraderie we shared.

Lastly, to my family: your unwavering faith in my capabilities, your love, and your endless encouragement have been the backbone of my journey. Thank you for being my strength, for the countless sacrifices you've made, and for reminding me of the bigger picture when the going got tough.

PREFACE TO THE THESIS

Original contributions to Knowledge

While it's widely accepted that gene disruptions linked to neural development are a common feature in neuropsychiatric disorders, there remains a significant knowledge gap regarding the molecular mechanisms by which genetic variants modify susceptibility to psychiatric phenotypes. Model organisms provide an important resource for understanding and identifying 1) biological mechanisms that mediate environmentally induced changes in neurodevelopment, and 2) tissue-specific gene co-expression networks that are relevant in this context. The main hypothesis for this doctoral thesis argues that measuring and estimating molecular processes at the tissue level can ultimately associate to ongoing neurodevelopmental events within brain substrates critical for the development of inhibitory control behaviors.

In **Chapter I**, I review background literature regarding the problem of psychiatric heterogeneity, the use of psychiatric endophenotypes to overcome this problem, and the use of such endophenotypes to map neurobiological dysfunctions to brain substrates. Next, I discuss the importance of sensitive periods during brain development and how early life experiences can permanently impact developmental trajectories. Then, I introduce inhibitory control behaviors, the main trait under study in this thesis, elaborating on its protracted maturational trajectories in mice and humans, and their striking temporal parallel with mesocorticolimbic dopamine development. Finally, I review the highly-conserved role of Netrin-1/DCC signaling system in mesocorticolimbic dopamine development, and how disruptions to its signaling during sensitive periods of development ultimately play a central role in adult corticolimbic structure and function in both mice and humans.

In **Chapter II**, I describe the results of experiments that characterize, using rodents, the molecular, behavioral, and structural consequences of administering therapeutic-like doses of amphetamine during

sensitive neurodevelopmental windows. Previous research had characterized these same molecular, behavioral, and structural effects using recreational-like doses of amphetamine, and while important insights about the mechanisms responsible for changes in the maturation of corticolimbic networks had been gathered, I looked for a translational component lacking from this line of research. First, I measure peak plasma levels achieved by an intraperitoneal injection of 0.5 or 4.0 mg/kg of amphetamine, to compare the experimental setting in this study to human pharmacokinetic studies that monitor blood profiles following exposure to clinically relevant or recreational doses of amphetamine. I show that administration of therapeutic-like doses of amphetamine before the full maturation of the brain does not induce impairments in inhibitory control behaviors in adulthood, and instead it prompts an overall increase in the performance of mice in the Go/No-Go task, several weeks after treatment cessation. These results show that developmental consequences of exposure to therapeutic- versus abused-like doses of amphetamine early in life have opposite behavioral effects and dissimilar molecular consequences, evidenced by an increase and a decrease in DCC protein in dopamine neurons, respectively.

Moving the focus from individual genes to gene co-expression networks, in **Chapter III** I show the identification and calculation of an expression-based polygenic score (ePRS) that consists of single-nucleotide polymorphisms (SNPs) within genes that are co-expressed with *DCC* in specific corticolimbic regions of the brain (the prefrontal cortex and nucleus accumbens). Results from this study provide a new functional and corticolimbic-specific marker for individual variation in impulsivity-related phenotypes in three prospective birth cohorts. Intriguingly, as increased impulsivity reflects a lack of inhibitory control, this marker (which is based on a prioritized subset of genetic variants) could reflect in and of itself an endophenotype bridging genotypes and basic behavioral assays of impulsivity.

In **Chapter IV**, I present an in-depth review of currently emerging genomic risk assessment approaches in psychiatry, placing a particular emphasis on methodologies that explore the neurobiological mechanisms by which gene networks contribute to psychiatric endophenotypes. This is meant to serve the reader as a comprehensive review of the literature on polygenic signals in psychiatry, although it also builds upon the previous chapter by discussing prospective frameworks that can be applied, other than our specific genotype-gene regulation framework- the ePRS.

In **Chapter V**, I integrate and discuss the main findings in Chapters II and III, and provide concluding remarks and future directions to the research presented in this thesis.

Statement of Originality

This thesis is presented in the manuscript-based format for a Doctoral Thesis, following the guidelines provided by the Department of Graduate and Postdoctoral Studies at McGill University. The studies described here were performed under the supervision of Dr. Cecilia Flores and Dr. Patricia Pelufo Silveira, and were discussed with my advisory committee members Dr. Frederic Charron and Dr. Andreas Arvanitogiannis. This thesis includes five chapters: Chapter I is a review of background concepts in *developmental neurobiology* relevant to this thesis. Chapters II and III are two original studies that have been published *in Addiction Biology* and *Molecular Psychiatry*, respectively. Chapter IV is an in-depth review of the emerging genomic risk assessment approaches in psychiatry published in *Biological Psychiatry Global Open Science*. Chapter V is a discussion of findings in Chapters II and III, concluding remarks, and future directions to the research presented in this thesis. Additionally, connecting statements from Chapter II to Chapter III, and from Chapter III to Chapter IV serve as a preface for each manuscript, to connect them into a single cohesive body of research presented herein.

Contribution of Authors

Chapter I

J.M.R.L wrote the general introduction in Chapter I under the supervision of Dr. C.F. and Dr. P.P.S.

Chapter II

J.M.R.L, S.C, and C.F conceived and designed the experiments. S.C and J.M.R.L performed the molecular evaluations. R.R and B.S.S quantified plasma amphetamine levels. J.M.R.L performed all behavioral experiments with help from C.P and S.C. S.H performed the neuroanatomical analyses. J.M.R.L, S.C, and C.F analyzed the results. S.I contributed to the statistical analysis. G.H contributed to the editing of the manuscript. J.M.R.L, S.C, and C.F discussed the results, and reviewed and edited the final manuscript. C.F and S.C supervised the project.

Chapter III

J.M.R.L, P.P.S, and C.F conceived and designed the project. J.M.R.L and S.P calculated the expression-based polygenic scores. J.M.R.L, Z.W, and I.P performed the statistical analyses. J.M.R.L performed the enrichment analyses. M.J.M provided scientific advice and the high-performance computational infrastructure. J.M.R.L, P.P.S, and C.F wrote the manuscript. J.M.R.L, I.P, P.P.S, and C.F discussed the results, and reviewed and edited the final manuscript. C.F and P.P.S supervised the project.

Chapter IV

Chapter IV contains a comprehensive review of functional genomics tools bridging neuroscience and psychiatry, written by J.M.R.L, C.F, and P.P.S.

Chapter V

The writing of this chapter was performed by the thesis author under the supervision of Dr. C.F. and Dr. P.P.S.

Chapter I: INTRODUCTION

1. The burden of neuropsychiatric conditions

Recent estimates from the Canadian government indicate that, annually, close to 1 in 5 Canadians are diagnosed with a mental illness, with important consequences for the general population in the form of increased disability and mortality. Additionally, current estimates for lifetime prevalence indicate that 1 in 3 Canadians will be diagnosed with a mental illness during their lifetime [1]. Despite the high prevalence and considerable burden[2], our knowledge of the underlying neurobiological processes responsible for the development of psychiatric conditions, as well as the identification of early markers of vulnerability, remain limited. Traditional diagnostic systems like the DSM-5 and ICD-10, although historically valuable for guiding clinical communication and decision making, suffer from marked heterogeneity within diagnostic categories and high comorbidity rates[3]. This is primarily due to their reliance on symptomatic criteria, which often overlook the inherent biological and social complexity of psychopathologies[4–6].

The heterogeneity problem underscores an important need for a rigorous framework that allows psychiatric researchers and clinicians to study mental disorders embracing the complexity of psychopathologies, rather than attempting to reduce it to categorical diagnoses. Developed by the National Institute of Mental Health (NIMH), the Research Domain Criteria (RDoC) framework views mental disorders as extremes along *functional dimensions of neurobiological systems*[7]. The RDoC's approach acknowledges that symptoms and psychological constructs often transcend traditional diagnostic boundaries, ultimately encouraging the discovery of *endophenotypes* that can ideally improve early detection and prevention strategies for neuropsychiatric disorders that share similarities across functional dimensions[8,9]. These dimensions integrate our knowledge in genetics, behavioral science, and neuroscience, with a constant strive to provide a biologically-based understanding of mental health and

disease[10,11]. In the context of this thesis, where the investigation of molecular networks involved in neuropsychiatric conditions takes a central place, the RDoC framework serves as an essential foundation, guiding the exploration and interpretation of the underlying endophenotypes associated with mental health disorders[5].

2. Sensitive Periods During Development

Human development, as a biological process grounded on a genetic program, unfolds through a series of overlapping stages whereby internal and external signals are constantly integrated. This process results in a gradual and exquisitely specific spatiotemporal patterning of gene expression that ultimately guides the specialization of the different tissues in the body, including the brain[12,13]. Remarkably, brain development seems to extend well into our postnatal life as neurons form and adjust synaptic connections based on functional validation[14], influenced by unique individual experiences like chronic stress[15], exercise[16] or exposure to drugs of abuse[17]. These experiences can weave themselves into our neural fabric if they occur during a period where the brain is responsive (or sensitive) to environmental influences. In sum, developmental experiences can cause permanent changes to brain structure and function, whereas adult experiences usually trigger temporary, compensatory responses that often revert to the baseline function over time[18,19].

3. Maturation of Inhibitory Control

Inhibitory control involves the ability to suppress or inhibit prepotent or automatic responses to voluntarily choose a more context-appropriate goal-directed behavior. People with strong inhibitory control can resist impulses, delay gratification, and modulate their behavior according to situational demands. On the other hand, individuals with weak inhibitory control may struggle with impulsivity, have difficulties adhering to rules or social norms, and may be more prone to engaging in risky or maladaptive behaviors[20]. While rudimentary forms of inhibitory control are present relatively early in

life, its refinement and maturation continue throughout adolescence and early adulthood, mirroring the protracted maturation and refinement of brain structural and functional networks engaged in the behavior, most notably of the prefrontal cortex (PFC)[21–23].

3.1 Inhibitory Control Impairments in Neuropsychiatric Disorders

Many studies have shown an association between inhibitory control impairments and neuropsychiatric disorders like attention-deficit/hyperactivity disorder (ADHD) [24,25], obsessive-compulsive disorder (OCD) [26], and substance use disorder (SUD) [27], among others. This behavioral impairment can lead to problems such as lack of focus, impulsive behaviors, and invasive thoughts – symptoms frequently found in these disorders. SUDs, for example, often involve clear manifestations of deficient inhibitory control as individuals may struggle with resisting the impulse to use substances despite adverse consequences. In addition, impulsivity- as well as compulsivity- are central to many theoretical models of addiction, emphasizing the shift from impulsive drug use in the early stages to compulsive use in the later stages[28]. In ADHD, the inability to inhibit impulsive behaviors is one of the primary symptoms leading to issues with attention and hyperactivity. Similarly, OCD patients often grapple with inhibiting intrusive thoughts and compulsive behaviors.

Impaired inhibitory control appears to be an important feature across several neuropsychiatric disorders, suggesting its potential consideration as an endophenotype that can help identify early in life individuals with a higher risk of developing these conditions[29]. In addition, the use of animal models has been an important tool for the identification of mechanisms involved in inhibitory control development, since the same tasks that are used for measuring these phenotypes across the human population[30,31] have been adapted for investigation using murine models, including the Go/No-Go task and the Stop-Signal task[32,33].

3.2 Neuroanatomy of Inhibitory Control

Multiple lines of research have established the central role that the brain's prefrontal and striatal regions play in our ability to inhibit impulsive or automatic responses[34,35]. The prefrontal cortex is involved in high-level cognitive processes while the striatum, a major component of the brain's basal ganglia, is critically involved in the regulation of motor and action planning, among other functions. Importantly, altered connections and communication between these brain regions have been linked to deficits in inhibitory control[23,36]. On a neurobiological level, this altered communication may involve imbalances in neurotransmitter systems that modulate neural activity within these pathways[37,38]. Aberrant signaling between the PFC and striatum can disrupt the delicate balance between inhibition and activation, making it harder for individuals to exert control over their automatic or impulsive responses[39]. Thus, understanding the nature and causes of these altered prefrontal-striatal connections is essential for elucidating the neural underpinnings of Inhibitory control.

4. The Mesocorticolimbic Dopamine System

The mesocorticolimbic dopamine system originates in the ventral tegmental area (VTA) of the midbrain, with projections reaching to various forebrain regions including the PFC and the nucleus accumbens (NAcc) in the ventral striatum[40]. Dopamine, a central neurotransmitter in this pathway, plays an essential role in a variety of functions, including motivation and the regulation of motor control[41]. In the context of inhibitory control, dysregulation of dopamine transmission in this pathway has been shown to contribute to impulsive behavior and reduced control over automatic responses[28,42,43].

Indeed, alterations in the mesocorticolimbic dopamine pathway have been implicated in several neuropsychiatric disorders that exhibit deficits in inhibitory control[44,45]. Moreover, the cognitive capacity to control and override impulsive behaviors improves gradually from childhood to early adulthood, mirroring the protracted developmental trajectory of the PFC, and its gradual quantitative and qualitative changes in dopamine innervation[14,46–48].

4.1 Segregation of mesolimbic and mesocortical dopamine axons during postnatal development

Mesocorticolimbic dopamine axons grow from the VTA to reach the striatum early in development, and while mesolimbic dopamine axons arborize and establish local connections in the ventral striatum-mainly the NAcc, mesocortical dopamine axons gradually continue to grow to the PFC throughout adolescence, subsequently modifying the structure and function of local neurons in the PFC[40,48].

The development of mesolimbic and mesocortical dopamine inputs is characterized by distinct temporal timelines[49,50], but they show a reciprocal functional connection throughout[51–54]. There's a dynamic interplay observed between these two systems during development: On one hand, it has been shown that changes in striatal dopamine regulation, such as a transient overexpression of dopamine D2 receptors during adolescence, can affect the functionality of the PFC and contribute to cognitive deficits in adulthood[55]. This suggests that the events regulating striatal dopamine maturation interact closely with those directing mesocortical dopamine axon growth. On the other hand, 1) alterations in PFC dopamine neurotransmission have been linked to corresponding changes in NAcc dopamine function[53,56], and 2) alterations in D1- or D2-expressing NAcc pathways, in turn, influence gene expression in the PFC[57]. Hence, the development of the PFC and cognitive control appears to involve the coordinated interplay of corticostriatal neuronal networks (also referred to as corticolimbic in some of our manuscripts, and frontostriatal in some areas of research), demonstrating the interconnectedness of these pathways and their collective influence on cognitive function[58,59].

5. Molecular processes guiding mesocorticolimbic dopamine development

The mesocorticolimbic dopamine system plays a pivotal role in various neurological and psychological functions, including movement control, motivation, emotion regulation, and cognition[60]. This circuitry's development stems from a complex interplay of genetic, epigenetic, and environmental

factors, including a highly orchestrated sequence of molecular events that involve several guidance-cue signaling pathways. Among them, the Netrin-1/DCC pathway has emerged as a particularly crucial mechanism regulating axonal guidance of dopaminergic neurons, as 1) ~99% of dopamine neurons in the rat VTA express DCC[65], and 2) variations in both Netrin-1 and DCC function have been associated with consequent changes in dopamine connectivity and dopamine function[43,61–64]. Indeed, there is evidence from both human studies[66] and model organisms[43,62] demonstrating the importance of Netrin-1/DCC signaling in mesocorticolimbic dopamine development and indicating that there is a high level of precision in the mechanisms guiding its development.

5.1 Netrin-1/DCC signaling in axon guidance and synaptic plasticity

Netrin-1, a laminin-related secreted protein, interacts with its receptor DCC, a member of the immunoglobulin superfamily of cell adhesion molecules, wherein upon their binding, DCC undergoes dimerization, triggering downstream signaling cascades including the activation of focal adhesion kinase (FAK) and small GTPases like Rac1 and Cdc42, all of which are instrumental in cytoskeletal rearrangements[67,68]. These downstream signaling cascades can lead to the reorganization of actin filaments and microtubules in the axonal growth cone, facilitating movement and steering. Netrin-1 has been proposed to be a bi-functional guidance cue, as its association to DCC can result in either attractive or repulsive axonal responses[69,70]. This dual function is determined by the presence of other co-receptors (such as UNC5C), allowing the recognition of final axon targets and the formation of synaptic connections within local circuitry, in addition to preventing axons from innervating and connecting with inappropriate targets. More recently, another role for Netrin-1/DCC signaling has been demonstrated in fully mature brain networks, where it acts as an effector of synaptic plasticity, suggesting that the functionality of this pathway depends on the developmental stage of the organism[71–73].

5.2 The role of the Netrin-1/DCC system across developmental stages

During the embryonic stage, the Netrin-1/DCC signaling pathway plays an essential role in guiding axonal growth by steering the growth cone of developing neurons toward their target locations. In humans, mutations in the *DCC* gene have been shown to result in dramatic neurodevelopmental alterations such as agenesis of the corpus callosum[74,75], developmental split-brain syndrome[74], and congenital mirror movements[76]. These disruptions are seen in both humans and mice[77], emphasizing the fundamental importance of DCC-dependent signaling in neurodevelopmental wiring. However, as development progresses into more mature- or adult-like states, there's a switch in the functional role of *DCC*[78]. Its expression shifts from high to low, correlating with a change from the broad organization of developing neuronal networks to the initiation of synapse assembly and the promotion of synaptic plasticity. In fact, both Netrin-1 and DCC are enriched at synapses in the mature mammalian brain[71], which corresponds to its discovered role in mature circuits[72], where DCC has been shown to be involved in synaptogenesis and synaptic plasticity, acting as a key regulator of neural complexity and connection[73]. Recent studies have also shown that many polymorphisms in *DCC*[79], as well as altered levels of gene expression[62] in the prefrontal cortex (for a review, see [78]), are related to various neuropsychiatric conditions of developmental onset characterized by deficits in PFC function and impulse control, stressing the multifaceted roles of *DCC* signaling across different phases of neural development, and highlighting its broader implications in the understanding of psychiatric conditions.

5.3 Environmentally-induced alterations to Netrin-1/DCC expression

Proper DCC-mediated Netrin-1 signaling requires a degree of robustness that extends well beyond the wiring of the mesocorticolimbic dopamine system, indicating that if the Netrin-1/DCC system is highly sensitive to *any* environmental stimulus, it would not be reliable in guiding axonal targeting and connectivity, or at least not to the degree observed across individuals and across species, and therefore its expression levels are tightly regulated by strong genetic effects. There are exceptions to this statement, however. Evidence from rodent models have shown that exposure to recreational-like doses

of amphetamine can modify the expression of *Dcc*. Critically, these effects are manifested in a developmentally-sensitive manner[80–83], as repeated exposure to recreational-like doses of amphetamine in adulthood exerts a diametrically opposite effect- compared to repeated exposure in the juvenile period- in the expression of *Dcc* in dopamine neurons. Consequently, repeated exposure to amphetamine leads to alterations in the organization of the mesocorticolimbic dopamine system, but only when the exposure occurs early in development.

6. From individual genes to tissue-specific gene co-expression networks

In principle, transcriptionally co-regulated genetic networks refer to a set of genes that are controlled together at the transcriptional level and are often associated with *similar* or *connected* biological functions[84,85]. These networks ensure that genes involved in specific molecular processes are turned on and off together, leading to a coordinated response by the cell or the specific tissue to its current *environmental* conditions.

This concept has not yet been fully adopted by the psychiatric genetics community, as the main experiment to investigate genotype-phenotype associations (the Genome-Wide Association Study-GWAS) relies on an explicit agnosticism when it comes to the biological functions implicated in disease risk, scanning in an unbiased manner the entire genomic landscape and correcting for the vast number of multiple comparisons conducted simultaneously (typically in the order of millions)[86,87]. This can be problematic because genotype-disease effects are small for most common genetic variations. However, incorporating prior knowledge from molecular and cellular biology experiments can help shift the focus from genome-wide variant identification (genotype-disease model) to quantifying the contribution of a prioritized subset of variants to an endophenotype (gene regulation-phenotype model). The fact that a large proportion of psychiatric disease risk can be explained by variants that modulate gene expression

levels is intriguing and, I suggest, may provide clues for the cellular and biological mechanisms underlying disease risk[88].

7. Rationale and Objectives

Understanding the biological and environmental factors that influence psychiatric conditions is critical for developing effective treatments and interventions. Although many studies have explored these aspects separately, combining genetic and environmental perspectives can provide a more comprehensive understanding. This thesis focuses on two primary aspects of this multidimensional puzzle: First, it delves into the consequences of environmentally-induced disruptions in Netrin-1/DCC signaling – a critical system in brain development - during sensitive neurodevelopmental periods. Specifically, it examines their impact on the organization of corticolimbic circuits, which are instrumental in regulating inhibitory control behaviors. Secondly, this thesis explores the association of co-expression based polygenic scores, grounded on corticolimbic-specific DCC-gene co-expression networks, with laboratory-based measurements of impulsivity-related phenotypes. This comprehensive approach aims to shed light on the intricate interplay between genetic and environmental factors in the development of psychiatric endophenotypes.

The first study implemented a rodent model to characterize the molecular, behavioral, and structural outcomes of disturbing a DCC-mediated signaling system within corticolimbic networks, key brain hubs orchestrating inhibitory control behaviors. I exposed mice early in development to a non-contingent regimen of a therapeutic-like dose of amphetamine, resulting in an environmentally-induced manipulation of the Netrin-1/DCC guidance cue system. This study allowed me to characterize the differential impact of therapeutic-like versus recreational-like doses of amphetamine on dopamine maturation and consequently on the development of inhibitory control behaviors. Results from this

work advanced our understanding of the consequences of psychostimulant administration during sensitive neurodevelopmental periods.

Building upon the insights gained from the rodent study, the second study involved human subjects and the exploration of a novel type of polygenic signal based on corticolimbic-specific DCC-gene co-expression networks across three ethnically diverse independent birth cohorts. This analysis aimed to create a biological marker for individual variations in impulsivity-related phenotypes, focusing particularly on our knowledge of the molecular processes involved in the maturation of corticolimbic substrates in the brain. By employing cutting-edge techniques in polygenic risk analysis and functional genomics, my research aims to shed light on potential genetic markers contributing to psychiatric susceptibility, within the complex landscape of human genetic interactions and influences that may predispose or protect against certain psychiatric conditions.

Together, these investigations aim to advance our understanding of the intricate interplay between genetic risk factors and environmental influences in susceptibility to psychiatric conditions, contributing to a nuanced and enriched understanding of these complex disorders. This introductory chapter provides a background of the association between inhibitory control impairments and psychiatric conditions, and of the cellular and molecular mechanisms involved in the development of the mesocorticolimbic dopamine system. Subsequent sections will detail the specific methodologies employed in each study, followed by a comprehensive overview of the thesis structure.

CHAPTER II: DCC-RELATED DEVELOPMENTAL EFFECTS OF ABUSED-VERSUS THERAPEUTIC-LIKE AMPHETAMINE DOSES IN ADOLESCENCE

Authors:

Santiago Cuesta[1*], José Maria Restrepo-Lozano[1,2*], Christina Popescu[1,2], Susan He[1], Lauren M. Reynolds[1,2], Sonia Israel[1], Giovanni Hernandez[1], Rana Rais[3], Barbara S. Slusher[3] and Cecilia Flores[1#]

Affiliations:

[1]Department of Psychiatry and Department of Neurology and Neurosurgery, McGill University, Douglas Mental Health University Institute, Montreal, Quebec, Canada, H4H 1R3. [2]Integrated Program in Neuroscience, McGill University, Montreal, Quebec, Canada, H3A 2B4. [3]Johns Hopkins Drug Discovery, Johns Hopkins School of Medicine, Baltimore, Maryland, USA, 21205.

***These authors contributed equally to this work**.

Published in

Addiction Biology. 2019; 25(4): e12791.

https://doi.org/10.1111/adb.12791

Reproduced with Permission

1. Abstract

The guidance cue receptor DCC controls mesocortical dopamine development in adolescence. Repeated exposure to an amphetamine regimen of 4 mg/kg during early adolescence induces, in male mice, downregulation of DCC expression in dopamine neurons by recruiting the *Dcc* microRNA repressor, miR-218. This adolescent amphetamine regimen also disrupts mesocortical dopamine connectivity and behavioral control in adulthood. Whether low doses of amphetamine in adolescence induce similar molecular and developmental effects needs to be established. Here we quantified plasma amphetamine concentrations in early adolescent mice following a 4 or 0.5 mg/kg dose and found peak levels corresponding to those seen in humans following recreational and therapeutic settings, respectively. In contrast to the high doses, the low amphetamine regimen does not alter *Dcc* mRNA or miR-218 expression, instead it *upregulates* DCC protein levels. Furthermore, high, but not low, drug doses downregulate the expression of the DCC receptor ligand, Netrin-1, in the nucleus accumbens and prefrontal cortex. Exposure to the low-dose regimen did not alter the expanse of mesocortical dopamine axons or their number/density of presynaptic sites in adulthood. Strikingly, adolescent exposure to the low-dose drug regimen does not impair behavioral inhibition in adulthood, instead it induces an overall increase in performance in a Go/No Go task. These results show that developmental consequences of exposure to therapeutic- versus abused-like doses of amphetamine in adolescence have dissimilar molecular signatures, and opposite behavioral effects. These findings have important clinical relevance since amphetamines are widely used for therapeutic purposes in youth.

2. Introduction

Adolescence represents a key neurodevelopmental stage when environmental factors can have a strong influence over ongoing structural, molecular and neurochemical changes [1]. In rodents, the adolescent growth of dopamine axons to the prefrontal cortex (PFC) is a particularly important event because it affects the structural and functional maturation of PFC circuitry and, in turn, behavioral flexibility and cognitive control in adulthood [2-5]. We have shown that the guidance cue receptor DCC, which is highly expressed by mesolimbic dopamine axons [6], promotes targeting recognition events in the nucleus accumbens (NAcc) in adolescence, preventing them from continuing to grow ectopically to the PFC [4]. Subtle changes in DCC levels in adolescence impact the extent of the dopamine innervation to the PFC, leading to substantial modifications in PFC circuitry connectivity and in cognitive processing in adulthood [4-9].

Repeated non-contingent exposure to 4 mg/kg of amphetamine in early adolescence, but not in adulthood, downregulates Dcc mRNA and protein expression in the ventral tegmental area (VTA) in male mice [10, 11]. This effect is mediated by amphetamine-induced VTA upregulation of the microRNA-218 (miR-218), which is a potent repressor of DCC in both human and rodent neurons and appears to control DCC expression in dopamine neurons across postnatal life [10, 12]. Notably, exposure to 4 mg/kg of amphetamine in early adolescence also results in disruption to the development of mesocorticolimbic dopamine connectivity, leading to an increase in the expanse of the dopamine input to the PFC, but to a significant reduction in the number of presynaptic sites of PFC dopamine axons in adulthood [5, 9]. These neuroanatomical alterations, which require downregulation of DCC receptors in dopamine neurons, are associated with deficits in PFC-dependent behaviors, including impaired behavioral inhibition [5, 9, 13-16].

To date, we have explored the effects of amphetamine in adolescence on miR-218/Dcc expression, PFC dopamine development, and behavioral inhibition using a 4 mg/kg dose. In mice, approximately the same doses have been reported to achieve plasma amphetamine levels comparable to those reached by recreational doses used in humans, ranging from 500-2500 ng/mL [17-22]. However, the doses of amphetamine that are typically used for therapeutic purposes, have been shown to reach peak plasma levels ranging from 30-140 ng/mL [23-27]. Whether exposure to equivalent low doses of amphetamine in mice downregulates miR-218/Dcc expression in the VTA, altering PFC dopamine development and cognitive control in adulthood needs to be determined. Here we quantified plasma amphetamine levels achieved by exposing early adolescent male mice to a 0.5 or 4 mg/kg dose and then determined the impact that the low dose regimen has on the Dcc-dependent maturation of mesocortical dopamine connectivity and behavioral control.

3. Materials and methods

Animals

All experiments and procedures were performed according with the guidelines of the Canadian Council of Animal Care and the McGill University/Douglas Mental Health University Institute Animal Care Committee. C57BL/6 wild-type male mice were obtained from Charles River Canada and maintained in the colony room of the Douglas Mental Health University Institute Neurophenotyping center on a 12-h light–dark cycle (light on at 0800 h) with food and water available ad libitum. All the experiments were conducted during the light part of the cycle.

Drugs

d-Amphetamine sulfate (Sigma-Aldrich, Dorset, United Kingdom) was dissolved in 0.9% saline. All amphetamine injections were administered i.p. at a volume of 0.1ml/10g and doses of 0.5 or 4 mg/kg. Different routes of administration lead to significant differences in the bioavailability of different drugs, including methylphenidate and cocaine [28, 29]. In this study we chose the i.p. route to be able to compare our findings with those reported by our lab and other groups.

Amphetamine plasma concentration

Amphetamine regimen and plasma collection: Male C57BL/6 mice received a single injection of amphetamine (0.5 or 4 mg/kg) at postnatal day (PND) 28 and trunk blood was collected after different time points (5, 15, 25 and 35 minutes) in tubes containing EDTA as anticoagulant. Blood plasma was obtained by centrifuging at 3000g for 10 minutes, at 4°C.

Bioanalysis of d-amphetamine in plasma samples: d-Amphetamine was extracted from plasma samples (25 µL) by protein precipitation with methanol (125 µL) containing 0.5 µM losartan (internal standard). Samples were vortexed and centrifuged (16,000 x g for 5 minutes) at 4°C and supernatants

were transferred to a 96 well plate. Separation of analytes was achieved on Ultimate 3000 UPLC system (Thermo Fisher Scientific, Waltham, MA, USA) with an Agilent Eclipse Plus C18 column (1.8 µm 2.1 X 100mm) (Agilent, Santa Clara, CA, USA) using a gradient run of 2.5:97.5 - 5:95 (water:acetonitrile) with 0.1% formic acid over 1.5 minutes and detected on a Thermo Q Exactive™ Focus orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). The orbitrap was run in a positive ion PRM (product reaction monitoring) mode. The ions of 136.1121 m/z and 423.1695 m/z are selected with a 1 m/z isolation window and fragmented with setting of 15 and 20 CE respectively. Transitions of 91.0563 and 119.0867 (d-amphetamine) and 207.0915 and 377.1522 (losartan) were used and quantified for concentration determinations. Calibration curves over the range of 0.03 -100 nmol/mL were constructed from the peak area ratio of the analyte to the internal standard using quadratic regression with a weighting factor of 1/(nominal concentration). Correlation coefficient >0.99 was obtained in all analytical runs for the analytes. Non-compartmental-analysis module in Phoenix WinNonlin version 7.0 (Certara USA, Inc., Princeton, NJ) was used to assess pharmacokinetic parameters. Peak plasma concentrations (C_{max}) and time to C_{max} (T_{max}) were the observed values. Area under the curve (AUC) was calculated by log–linear trapezoidal rule to the end of sample collection (AUC_{last}).

Amphetamine Regimen

Male C57BL/6 early adolescent mice were treated with saline or amphetamine injections from PND 22±1 to PND 31±1. Consistent with our previous studies, different groups of mice received one injection of amphetamine (0.5 or 4 mg/kg i.p.) or saline, every other day, for a total of 5 days. Locomotor activity was measured 15 minutes prior and 90 minutes after each saline or amphetamine injection.

In concordance with our and other studies, we define early adolescence in mice as the period between the day of weaning and PND 32 [3, 4, 8-11, 30, 31]. While this range is not an absolute margin, but an age during which mice exhibit distinct neurobehavioral characteristics, this definition seems now

to be a consensus in the rodent literature [1, 3, 32-34]. Indeed, early adolescence is the critical period when exposure to a high-dose amphetamine regimen leads to impaired behavioral inhibition, aberrant PFC dopamine connectivity, and reduced PFC dopamine function in adulthood [5, 9].

Western Blot analysis

One week after completing the saline or amphetamine treatment regimens, different cohorts of mice were rapidly decapitated and their brains were flash-frozen in 2-methylbutane (Fisher Scientific, Hampton, NH, USA) chilled with dry ice. Bilateral punches of the VTA, NAcc and PFC, were excised from 1-mm-thick coronal slices starting from sections corresponding to Plate 55 (-2.92mm, anterior/posterior relative to Bregma) and 15 (1.94mm, anterior/posterior relative to Bregma), respectively, of Paxinos and Franklin [35] and processed for western blot as before [10, 36]. Briefly, protein samples (15 µg) were separated on a 10% SDS-PAGE and transferred to a nitrocellulose membrane which was incubated overnight at 4°C with antibodies against DCC (1:1000, Cat#554223, BD Pharmingen, Mississauga, ON, Canada), Netrin-1 (1:750 dilution, Cat#NB100-1605, Novus Biologicals, Littleton, CO, USA) and β-actin (1:15000, Sigma-Aldrich, Oakville, ON, Canada).

RNA extraction and quantitative real time PCR for mouse tissue

One week after the saline or amphetamine treatment, bilateral punches of the VTA were taken from coronal sections obtained as described above. Total RNA and microRNA fractions were isolated using the miRNeasy Micro Kit protocol (Qiagen, Toronto, ON, Canada) as previously [10]. All RNA samples were determined to have 260/280 and 260/230 values ≥1.8, using the Nanodrop 1000 system (Thermo Scientific, Toronto, ON, Canada). RNA integrity was further assessed using the denaturing gel electrophoresis method proposed by Aranda et al [37]. The three well-defined 28S, 18S, and 5.8S/5S ribosomal bands were detected without any background. Reverse transcription for Dcc and Glyceraldehyde-3-phosphatedehydrogenase (Gapdh) mRNA was performed using a High-Capacity cDNA

Reverse Transcription Kit (Applied Biosystems, Foster City, CA) according to manufacturer's instructions. Real time PCR, using TaqMan assay (Applied Biosystems, Foster City, CA) was carried out with an Applied Biosystems 7900HT RT PCR system. Data for Dcc mRNA expression were analyzed by using the relative quantification standard curve method and the level of these transcript was quantified relative to the expression of the reference gene Gapdh. Reverse transcription for miR-218 was performed using the TaqMan MicroRNA Reverse Transcription Kit together with the corresponding miRNA TaqMan probes (Applied Biosystems, Foster City, CA). Expression levels were calculated using the relative quantification standard curve method. The small nucleolar RNA (snoRNA) RNU6B was used as endogenous control to normalize the expression of miR-218. In all cases, the real-time PCR was run in technical triplicates.

Neuroanatomical analysis

Perfusion. Adult mice received an intraperitoneal overdose of ketamine 50 mg/kg, xylazine 5 mg/kg, and acepromazine 1 mg/kg and were perfused intracardially with 50ml of 0.9% saline followed by 75ml of chilled fixative solution (4% paraformaldehyde in phosphate-buffered saline). Brains were dissected and placed in the fixative solution overnight at 4ºC and were then transferred to phosphate-buffered saline and stored for a maximum of 2 days. Brains were sectioned using a vibratome (35-µm-thick coronal slices).

Immunofluorescence. As we have done before, every second coronal section was processed (1:2 series) [6, 8, 36]. A rabbit polyclonal anti-TH antibody (1:500 dilution, catalog #AB152; Millipore Bioscience Research Reagents) and an Alexa Fluor 555-conjugated secondary antibody raised in goat (1:500 dilution, 1 h incubation, Invitrogen) were used.

Stereology. The TH antibody labels dopamine axons in the PFC with high specificity, and rarely labels norepinephrine axons [6, 8, 38]. As previously, and because of the lateralization of the dopamine system, we obtained counts only from the right hemisphere. To evaluate changes in mesocortical

dopamine connectivity in animals exposed to 0.5 mg/kg of amphetamine during adolescence, we performed stereological quantification of the span of TH-positive fibers in the cingulate 1, prelimbic, and infralimbic subregions of the pregenual medial PFC. The total volume of TH-positive fiber innervation (in cubic micrometers) was assessed using the Cavalieri method using Stereoinvestigator® (MicroBrightField) [4, 6, 8, 9]. Counts were performed blind. The coefficient of error was below 0.1 for all regions of interest in all sampled brains.

The medial PFC subregions were delineated according to plates spanning 14–18 (1.98mm - 1.54mm anterior/posterior relative to Bregma) of the mouse brain atlas [35]. A 5X magnification was used to trace the contours of the dense TH-positive innervation of the subregions using a Leica DM400B microscope. An unbiased counting frame (25 x 25 µm) was superimposed on each contour and counts were made at regular predetermined intervals (x=175µm, y=175µm) from a random start point. Counting of varicosities was performed at X100 magnification on 6 of the 12 sections contained within the rostrocaudal borders of our region of interest (Plates 14–18; 1:4 series). A guard zone of 5 µm was used and the optical dissector height was set to 10 µm.

Number of DA neurons. The total number of dopamine neurons in the VTA was assessed using the optical fractionator probe of Stereoinvestigator as previously [4, 7-9]. The counting scheme used a 60 x 60 µm counting frame (x=150 µm, y=150 µm intervals) with a random start point. Counting was performed at 40X magnification in a 1:4 series. A 3 µm guard zone and a probe depth of 10 µm were used.

Behavioral evaluations

Go/No-Go. We modified and optimized a Go/No-Go task for use with mice [4, 5]. Briefly, mice were food restricted for the duration of the behavioral testing to maintain a body weight of 85% of the initial weight. The task took place in operant behavioral boxes (Med Associates, Inc., St Albans, VT, USA)

equipped with a house light, 2 illuminated nose poke holes, an adjustable sonalert tone generator, and a pellet dispenser. We used chocolate flavored dustless precision pellets (BioServ, Inc., Flemington, NJ, USA) as the operant reinforcer. The experimental procedure consisted of three stages: Conditioned Reinforcement Training, Reaction Time Training, and the Go/No-Go Task. Animals were subject to one training or testing session per day.

Conditioned Reinforcement Training: At the start of each conditioned reinforcement training session, the house light turns on and remains on throughout the 20-minute session. Each trial within this session consists of the presentation of an illuminated nose poke hole for 9 seconds. If the mouse does not respond by nose-poking in the illuminated hole, the cue light is extinguished for a 10 second inter-trial interval (ITI) before the next trial/cue presentation. Thus, the purpose of the house light in this session is to signal that there is a current ongoing session.

If the mouse responds to the cue light by nose-poking in the illuminated hole, a chocolate food pellet is dispensed and the trial is counted as a "reward" trial. The location of the active (cued) nose poke hole (either Left or Right) is counterbalanced within groups and stays consistent for each individual mouse for the duration of the session and throughout each stage of the experiment. Responses to the active nose poke hole when the cue light is off, as well as responses to the non-active nose poke hole (where the cue light was never illuminated), do not result in a reward but are recorded and analyzed. Mice advanced to the next stage of training once they achieve a criterion of over 70% responses to cued trials. Mice received one 20-minute Conditioned Reinforcement training session per day.

Reaction Time Training. Once mice stably respond to the cued nose poke hole to receive the reinforcer, they are trained to 1) respond only following the illumination of the cue light, and 2) to respond within 3 seconds of the cue illumination to receive the reinforcer. To train mice to respond only when the cue light is present, the structure of the session is changed to incorporate a pretrial period

prior to each trial. In this pretrial period, the house light is illuminated for a variable amount of time (3, 6, or 9 seconds, distributed randomly) without the cue light. If the mouse nose pokes during this pretrial period, the house light is extinguished and a 10-second ITI is initiated. This is recorded as a 'Premature Response'. If the mouse does not respond during the pretrial period, the cue light is then illuminated for 3 seconds. Therefore, the purpose of the house light in this training phase is to signal the start of a pretrial period, and the purpose of this pretrial period is to train mice to respond specifically to the cue light. A response during the 3 second trial period results in the delivery of a reward pellet; if the mouse fails to respond within this window, the cue and house lights are extinguished and a 10-second ITI began. In order to advance to the Go/No-Go Task, mice have to end less than 25% of the pretrial periods with a premature response. Mice receive one 30-minute reaction time training session per day.

Go/No-Go Task. Following successful completion of both training stages, mice undergo 10 sessions of the Go/No-Go Task. This task requires mice to respond to a lighted 'Go' cue, identical to the cue used in the training sessions, or inhibit their response to this cue when presented in tandem with an auditory 'No-Go' cue. In the 'Go' trials, mice had to respond to the illuminated nose poke hole in the 3-second timeframe during which the cue light is on in order to receive a reward. This is counted as a 'Hit' in our analysis. In the 'No-Go' trials, an 85 dB tone is paired with the 3-second cue light to signal that the mouse should withhold from responding. If mice respond during the 3-second 'No-Go' trial, an ITI is initiated and no reward is dispensed. This is counted as a 'Commission Error' in our analysis. However, if mice withhold from responding for the 3 second duration of the tone/light 'No-Go' cue, a reward is dispensed. As in the Reaction Time training, a randomized, variable pretrial period of 3-9 seconds precedes each trial and the number of premature responses was recorded. The purpose of this is to mimic the setting used during the Reaction Time Training, where mice learn to withhold from nosepoking prior to the presentation of the cue light. Within each session, the number of 'Go' and 'No-

Go' trials is given in an approximately 1:1 ratio and presented in a randomized order. Each session lasts 30 minutes and consists of approximately 30-50 'Go' and 30-50 'No-Go' trials.

Statistical Analysis

All values are represented as means ± S.E.M. A significance threshold of $\alpha = 0.05$ was used in all the experiments. Statistical differences between two groups were analyzed with Student's t-tests and, when required, the significance level used to evaluate the comparisons was adjusted using the Holm–Bonferroni sequentially rejective procedure [39]. All data are normally distributed and the variance is similar between groups. Statistical differences between more than two groups were analyzed with one or two-way ANOVAs. The sample size in all the experiment varied from 4-9 animals per group.

4. Results

A low amphetamine dose leads to plasma levels seen in therapeutic settings

The concentration of d-amphetamine in plasma of early adolescent mice was measured at multiple time points post i.p. administration of 0.5 or 4 mg/kg (Fig 1a). d-Amphetamine showed quick absorption with a maximum concentration (C_{max}) of 97±21 ng/mL following administration of the 0.5 mg/kg dose (Fig 1b, 0.5 mg/kg: ANOVA with a main effect of time: $F(3, 12)=5.58$, $p=0.0124$). The 4 mg/kg dose, however, resulted in a concentration of 1300±79 ng/mL observed 5 minutes post injection (Fig 1c, 4 mg/kg: ANOVA with a main effect of time: $F(3, 12)=20.45$, $p<0.0001$). The exposures based on area under the curve (AUC_{last}) were 1928±305 for 0.5 mg/kg and 26349±1922 for 4 mg/kg. An eight-fold increase in the amphetamine dose resulted in almost a 14-fold increase in exposures. Thus, the increase in blood concentration of amphetamine was not simply proportional to the increase in the drug dose.

These results demonstrate that while the 0.5 mg/kg amphetamine dose reaches peak plasma concentrations within the range of those observed in therapeutic settings (i.e. 30-140 ng/mL) [23-27], peak plasma levels following the 4 mg/kg dose are within those seen in recreational use (i.e. 500-2500 ng/mL) [17-22].

Repeated exposure to a low dose of amphetamine in adolescence upregulates VTA DCC protein expression, without altering miR-218 or Dcc mRNA levels

The expression levels of DCC protein, Dcc mRNA, and miR-218 in the VTA are regulated by non-contingent exposure to 4 mg/kg of amphetamine during early adolescence [10]. To determine whether there is a threshold for these drug effects, we exposed early adolescent mice (PND 21±1 to PND 31±1; Fig. 2a) to 0.5 mg/kg of amphetamine. Injections of 0.5 mg/kg of amphetamine do not alter locomotor activity in comparison to saline treatment (Fig 2b, two-way ANOVA for repeated measures: significant main effect of the day, $F(4,112)=16.72$, $p<0.0001$; no significant effect of treatment, $F(1,28)=0.09$;

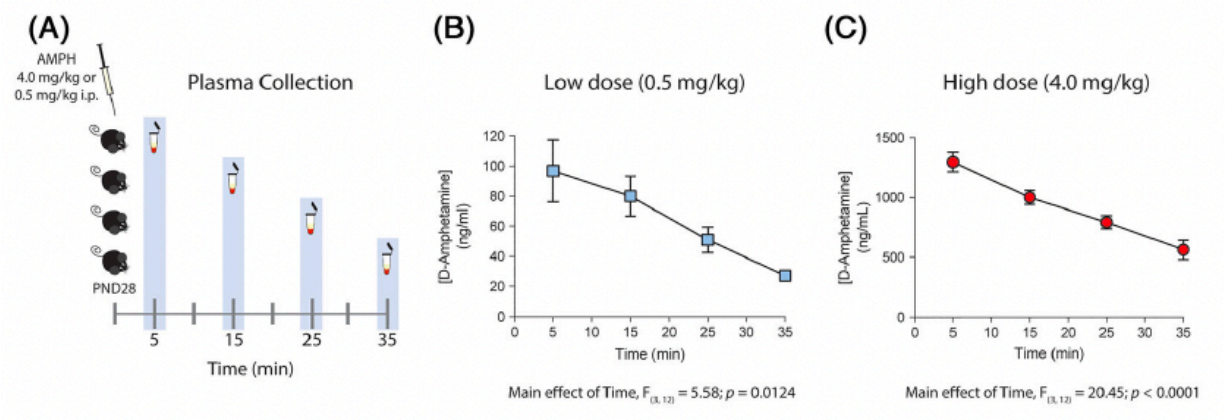

Figure 1 Peak plasma concentrations achieved by an intraperitoneal injection of 0.5 or 4.0 mg/kg of d-amphetamine (AMPH) correspond to those measured in humans in therapeutic and recreational settings, respectively. A, Diagram describing AMPH regimen and the different time points for plasma collection. B, and C, Bioanalysis of d-amphetamine in plasma observed over time after a single intraperitoneal injection of 0.5 or 4 mg/kg dose (n = 4 per time point). B, Plasma concentration of AMPH showed a maximum concentration of 97 ± 21 ng/mL 5 minutes after low-dose injection. C, Plasma concentration of AMPH showed a maximum concentration of 1300 ± 79 ng/mL 5 minutes after high-dose injection

p=0.76; or treatment × day interaction, F(4,112)=0.49, p=0.74). However, we found a significant increase in DCC protein expression in the VTA one week later in amphetamine-treated mice versus saline controls (Fig. 2c; t(14)=3.51, p=0.0035). This contrasts the DCC downregulation we previously reported following exposure to a 4 mg/kg amphetamine treatment regimen (Fig 2c inset; [10]). Furthermore, Dcc mRNA and miR-218 expression in the VTA did not differ between mice treated with 0.5 mg/kg of amphetamine or with saline in adolescence (Fig 2d Dcc mRNA: t(12)=1.20, p=0.25; Fig. 2e miR-218: t(12)=1.60, p=0.14). These findings, which are also opposite to the effects seen with the 4 mg/kg dose (Fig. 2d and e insets, [10]), indicate that the changes in DCC protein expression are posttranslational in nature.

Dose-dependent effects of amphetamine in adolescence on Netrin-1 expression

The effects of DCC receptor on axonal targeting and synapse formation are mediated by its interaction with its ligand Netrin-1 [40, 41]. Thus, we next determined the effect of high and low doses of amphetamine in adolescence on Netrin-1 expression in target regions of dopamine neurons (Fig. 2f-g). Exposure to 4 mg/kg of amphetamine significantly downregulated Netrin-1 expression in the NAcc one week later, in comparison to saline groups (Fig 2f, right panel, t(14)=0.89, p=0.019). Thus, high doses of amphetamine in adolescence not only downregulate DCC expression in mesolimbic dopamine neurons, but also reduce Netrin-1 levels in their target region. Intriguingly, this high-dose regimen also reduced Netrin-1 expression in the PFC (Fig 2g, right panel, t(14)=4.24, p=0.0008); an effect that may contribute to the reduced number of mesocortical dopamine varicosities in adulthood [5, 9].

In contrast, Netrin-1 expression in the NAcc and PFC did not differ between mice treated with the 0.5 mg/kg dose and saline (Fig. 2f, left panel, NAcc: t(14)=1.03, p=0.32; Fig. 2g, left panel, PFC: t(14)=0.16, p=0.88). The effects of amphetamine in adolescence on both DCC and Netrin-1 expression are therefore dose-dependent.

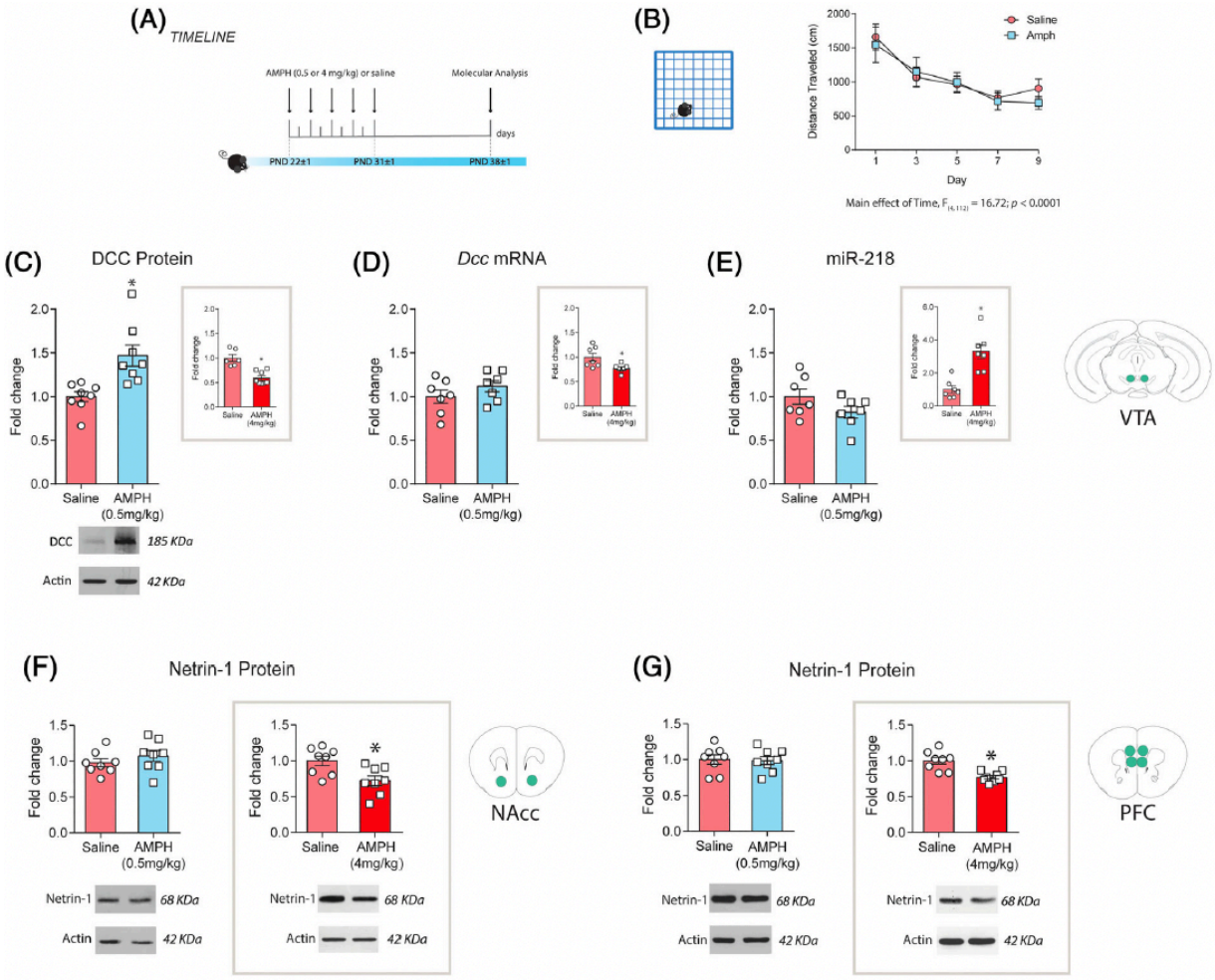

Figure 2 Exposure to a therapeutic-like dose of amphetamine in adolescence upregulates DCC protein expression without altering miR-218 or Dcc mRNA levels in the ventral tegmental area (VTA) and does not change Netrin-1 in dopamine terminal regions. A, Timeline of treatment and experimental procedures. B, Locomotor activity during the 90-minute test performed after each treatment injection. C, DCC expression is significantly increased in the VTA 1 week after the treatment with 0.5 mg/kg of amphetamine in adolescence. Inset: levels of DCC protein expression in the VTA in animals treated with 4 mg/kg of amphetamine, using the exact same schedule. In contrast to animals exposed to 4 mg/kg (D and E insets), animals treated with 0.5 mg/kg of amphetamine showed no changes in D, Dcc mRNA or E, miR-218 in the VTA compared with saline controls. F, and G, Netrin-1 expression in nucleus accumbens (NAcc) and prefrontal cortex (PFC) 1 week after termination of treatment with 0.5 or 4.0 mg/kg of amphetamine. No changes were observed in Netrin-1

expression levels in the F, NAcc or in the G, PFC in mice exposed to 0.5 mg/kg of amphetamines in comparison with saline-treated ones. Mice exposed to 4 mg/kg of amphetamine during adolescence showed a downregulation on Netrin-1 protein expression in the F, NAcc and the G, PFC compared with saline-treated controls. *Significantly different from saline group, $P < .05$. The data presented in the insets are reproduced from Cuesta et al.10 All data are shown as mean ± SEM

Exposure to low doses of amphetamine in adolescence does not alter mesocortical dopamine connectivity in adulthood

Next, we analyzed the extent of the mesocortical dopamine innervation to layers V and VI of the medial PFC in adult mice exposed to amphetamine (0.5 mg/kg) or saline during adolescence (Fig. 3a and b). We found no differences between groups in the span (i.e. volume) of TH-positive fibers (Fig 3c: two-way ANOVA, no significant main effect of treatment, $F(1, 8)=0.036$, $p=0.85$; no significant treatment × medial PFC region interaction, $F(2, 16)=0.078$, $p=0.93$; significant main effect of medial PFC region, $F(2, 16)=111.2$, $p<0.0001$). In the medial PFC nearly every dopamine varicosity forms a synapse with a dendritic spine or shaft [42], therefore, we measured the total number and the density of TH-positive varicosities. There were no group differences in the total number and density of medial PFC TH-positive varicosities in any of the three subregions analyzed (total number: Fig 3d: two-way ANOVA, no significant main effect of treatment, $F(1, 8)=0.056$, $p=0.81$; no significant treatment × medial PFC region interaction, $F(2, 16)=1.01$, $p=0.37$; significant main effect of medial PFC region, $F(2, 16)=73.92$, $p<0.0001$; density: Fig 3e: two-way ANOVA, no significant main effect of treatment, $F(1, 8)=0.11$, $p=0.75$; no significant treatment × medial PFC region interaction, $F(2, 16)=1.59$, $p=0.24$; significant main effect of medial PFC region, $F(2, 16)=0.02$, $p=0.82$). These results contrast the ones we observed in the PFC of mice exposed to 4 mg/kg of amphetamine during adolescence, namely an increase in the span of TH-positive innervation, but a reduction in the total number and density of TH-positive varicosities [9].

It is important to note that the number of TH-positive neurons in the VTA is similar between the amphetamine- and saline-treated groups (Dopamine neuron number: mean ± s.e.m.: Saline=7204±1135, AMPH=5744±1005, $t(8)=0.95$, $p=0.37$; data not shown), similar to our findings using the 4 mg/kg dose [9].

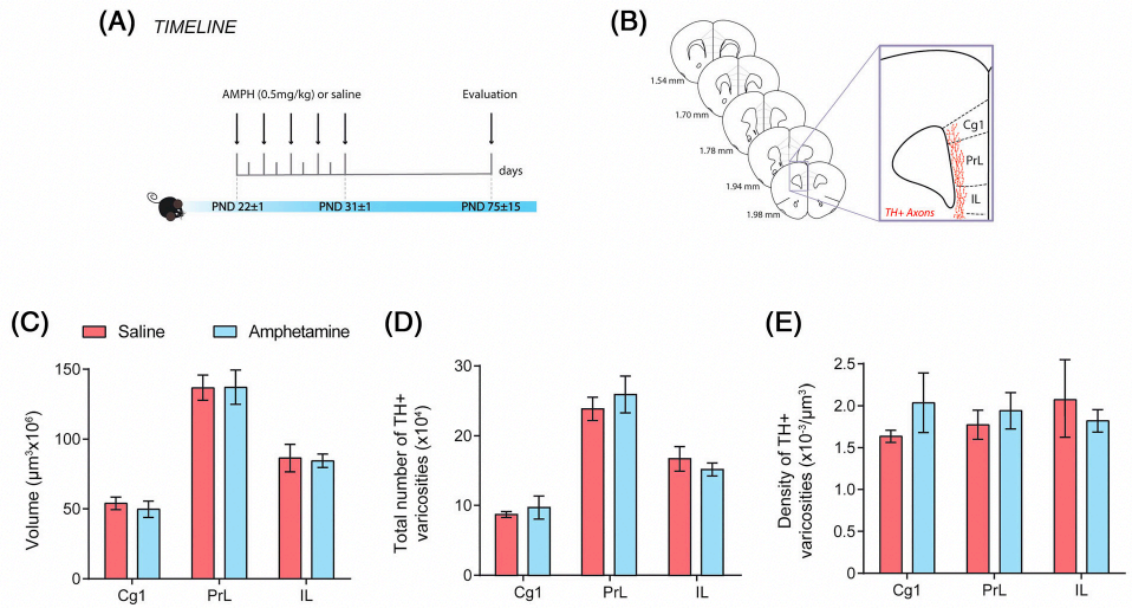

Figure 3 Low amphetamine doses in adolescence do not alter mesocortical dopamine connectivity in adulthood. A, Timeline of treatment and experimental procedures. B, Schematic representation of the regions of interest in the medial prefrontal cortex outlined according to the Mouse Brain Atlas (Paxinos and Franklin, 2008). The cingulate (Cg1), prelimbic (PrL), and infralimbic (IL) subregions of the medial prefrontal cortex were analyzed. C, Volume, D, total number, and E, density of the TH-positive fiber innervation to the inner layers of the medial prefrontal cortex. Mice treated with 0.5 mg/kg of AMPH in adolescence do not show significant differences to their saline counterparts. All data are shown as mean ± SEM (n = 4-6 per group)

Exposure to low doses of amphetamine in adolescence does not induce behavioral inhibition deficits in adulthood

We then assessed the effects of 0.5 mg/kg of amphetamine in adolescence on adult cognitive control using the Go/No-Go task [4, 5] (Fig 4a-b). There were no differences in the number of commission errors, "hits", or omission errors between amphetamine and saline groups (Fig 4c, commission errors: two-way repeated measures ANOVA, no significant effect of treatment, $F(1, 15)=0.76$, $p=0.396$ or time × treatment interaction, $F(9, 135)=0.79$, $p=0.625$, significant main effect of time, $F(9, 135)=16.69$, $p<0.0001$; Fig 4d, hits: two-way repeated measures ANOVA, no significant effect of treatment, $F(1, 15)=1.42$, $p=0.251$ or time × treatment interaction, $F(9, 135)=0.47$, $p=0.895$, significant main effect of time, $F(9, 135)=11.26$, $p<0.0001$; Fig 4e, omission errors: two-way repeated measures ANOVA, no significant effect of treatment, $F(1,15)=1.875$, $p=0.1911$ or time × treatment interaction $F(9,135)=1.351$, $p=0.2167$, significant main effect of time, $F(9,135)=16.37$, $p<0.0001$).

We also calculated the total number of correct responses across the Go/No-Go task using the following formula modified from Gubner et al [43]:

$$\frac{\frac{Go\ rewards}{Go\ trials} + \frac{NoGo\ rewards}{NoGo\ trials}}{2}$$

Interestingly, the amphetamine-treated group had a higher number of correct responses in comparison to the saline group (Fig 4f: two-way repeated measures ANOVA, significant effect of treatment, $F(1, 15)=5.78$, $p=0.0296$; no significant time × treatment interaction, $F(9, 135)=1.50$, $p=0.155$, significant main effect of time, $F(9, 135)=5.64$, $p<0.0001$). These results indicate that exposure to low doses of amphetamine in adolescence do not lead to deficits in cognitive control in adulthood, but in fact improve performance across the overall task.

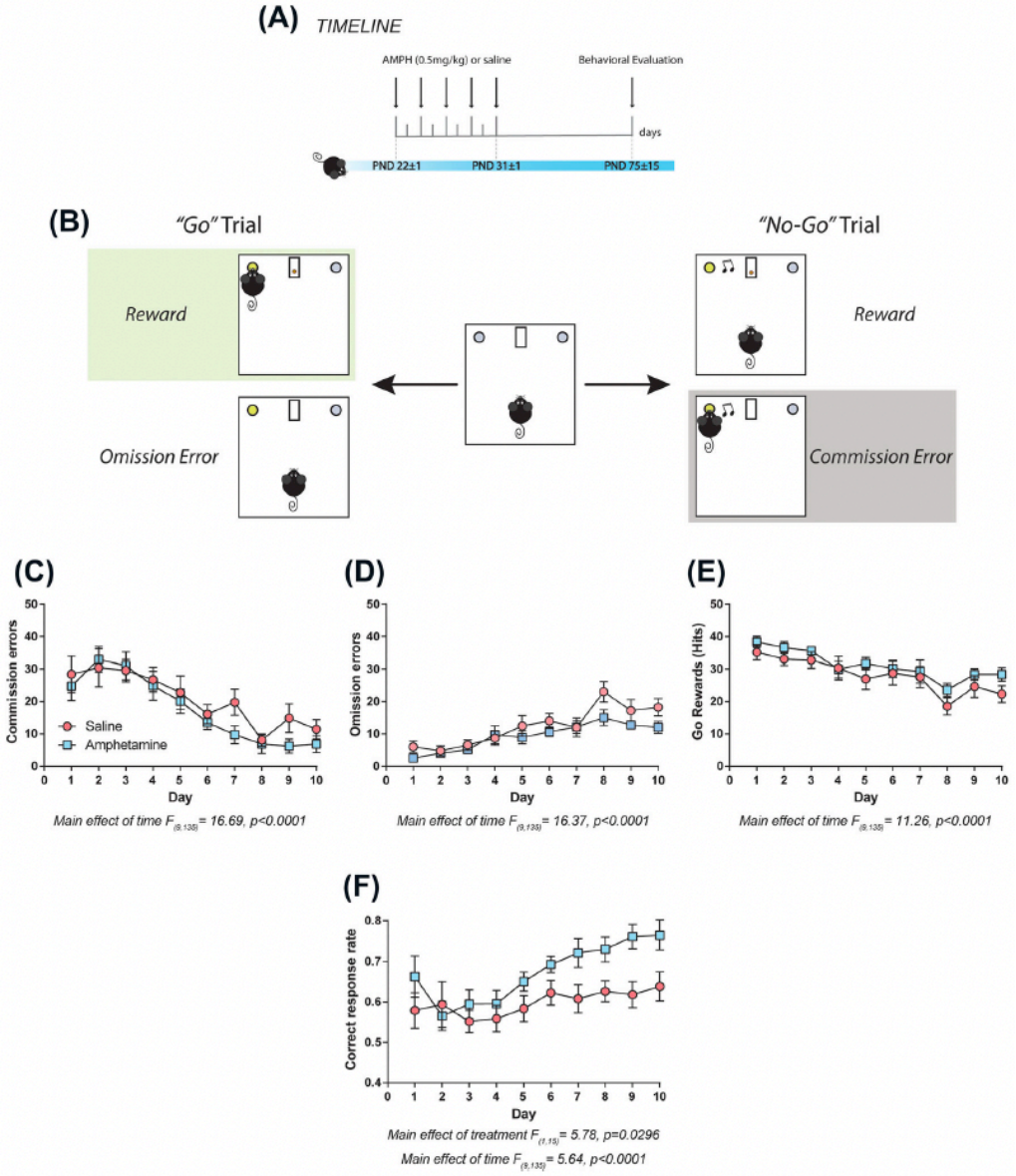

Figure 4 Low doses of amphetamine in adolescence improve overall cognitive performance in adulthood. A, Timeline of treatment and experimental procedures. B, Diagram of the go/no-go task adapted for mice. (C-E) There are no differences in the number of C, commission errors, D, omission errors, or E, "hits" between animals treated with 0.5 mg/kg or saline during adolescence. However, amphetamine exposure during adolescence induce a significant increase in the correct response rate across the test, when compared with saline-treated ones F. All data are shown as mean ± SEM (n = 8-9 per group)

5. Discussion

In this study we show that there is a dose effect of amphetamine in early adolescent male mice on miR-218/DCC/Netrin-1 expression, mesocortical dopamine development, and cognitive control. The two doses used in the present study reached plasma levels within the range of those observed in humans using the drug in therapeutic and recreational settings. For adolescent and adult mice, plasma levels are a good proxy for drug brain levels [44]; thus, it is expected that the doses used for the present studies achieved brain levels of the drug that can have relevant molecular changes.

First, while exposure to the high amphetamine dose regimen recruits miR-218 to downregulate DCC in the VTA [10, 11], the low drug dose increases DCC protein without altering *Dcc* mRNA or mir-218 levels. Second, only the high dose amphetamine regimen reduces Netrin-1 expression in the NAcc and PFC. Third, while the high dose regimen leads to a reduction in adult mesocortical dopamine connectivity and function [5, 9] inducing in turn a reduction in inhibitory control [5], the low dose regimen does not disrupt behavioral inhibition and, in fact, leads to an overall improvement in adult cognitive performance, as measured by an increase in the total number of correct responses in the Go and No-Go trials.

Exposure to 0.5 mg/kg of amphetamine increases DCC protein expression in the VTA, without altering Dcc mRNA levels or miR-218 expression in this region. This is in contrast to the downregulation of Dcc mRNA and protein expression induced by exposure to the 4 mg/kg dose regimen, which is actually mediated by drug-induced upregulation of miR-218 in the VTA [10]. Thus, DCC protein upregulation by the therapeutic-like treatment might be mediated by posttranslational mechanisms. The SIAH (seven in absentia homolog) protein family has been shown to ubiquitinate and regulate DCC protein levels in different organs, including rat brain [45, 46]. In future studies we will assess whether SIAH protein expression and/or activity is altered by amphetamine administration in adolescence.

Our previous work shows that DCC expression in mesolimbic dopamine axons determines where and when they recognize their final target and, accordingly, reduced DCC signaling in adolescence results in targeting errors and ectopic growth of mesolimbic dopamine axons from the NAcc to the PFC [3, 4, 9]. The fact that the high amphetamine regimen downregulates both DCC in dopamine neurons and Netrin-1 in the NAcc suggests that this experience potentiates the ectopic growth of mesolimbic dopamine axons to the PFC, causing disorganization of dopamine connectivity in this region. This in turn may lead to alterations in executive function [47], negatively impacting performance in tasks which require inhibition of a learned or ongoing behavior [5, 14, 15].

Recently, Netrin-1 has been demonstrated to promote synapse formation in the PFC [41] and to potentiate excitatory synaptic transmission via the insertion of GluA1 AMPA receptors in the hippocampus of adult mice [48]. Thus, the downregulation of Netrin-1 in the PFC induced by the high dose regimen may contribute to the reduction in the number and density of mesocortical dopamine synaptic sites observed in adulthood [5, 9]. Interestingly, a similar amphetamine regimen in adolescence has been shown to induce a reduction in the expression of D1 receptors in the PFC of adult male rats [49] suggesting also alterations in the electrophysiological properties at excitatory and/or inhibitory synapses in this region. In line with this idea, it has been reported that high, but not low, doses of the stimulant drug methylphenidate in adolescence reduces PFC long-term potentiation in adulthood [50]. These findings suggest that high doses of stimulant drugs in adolescence lead to functional reorganization of PFC synaptic circuitry and that local alterations in Netrin-1 signaling may contribute to this effect.

We have reported that both improvements and deficits in behavioral inhibition during the No Go task are correlated with alterations in dopamine connectivity in the adult PFC. Specifically, we have observed improved inhibitory control in mice that have increased PFC dopamine synaptic connectivity

[4]. In contrast, adult mice that have reduced PFC dopamine connectivity and turnover show deficits in inhibitory control [5]. In this study, however, we do not find changes in the span (i.e. volume) of the dopamine input or in the number/density of dopamine varicosities in the PFC of mice treated with low amphetamine doses in adolescence. We are currently setting new methodologies to assess directly dopamine neurotransmission in the PFC as well as to investigate changes in connectivity in other dopamine terminal regions or in other neurotransmitter systems.

Whether exposure to amphetamine in adolescence leads to changes in dopamine connectivity in the NAcc in adulthood remains to be addressed. The stereological methods used in this study are not sensitive enough to capture subtle changes in dopamine varicosity density and/or number in the NAcc [6, 8, 9]. Therefore, we are planning to use more sensitive methods (i.e. [4]) in future studies to address this issue.

To date, we have conducted all our studies in male mice. However, there is evidence of sex differences in the enduring behavioral effects of amphetamine exposure during adolescence [51-53], emphasizing the importance of addressing this issue in our studies. We are now beginning to assess whether the effects of amphetamine (high and low doses) in adolescence on DCC receptor signaling, PFC dopamine maturation, and behavioral control are sexually dimorphic. In addition, we also plan to examine whether adolescent exposure to the other typically prescribed psychostimulant, methylphenidate, also increases DCC protein expression one week later and leads to improved cognitive performance in adulthood. Indeed, methylphenidate administration can alter DCC expression in the VTA in adulthood [54].

To our knowledge, this is the first study to compare the effects of amphetamine doses equivalent to those used by humans for recreational versus therapeutic purposes on the expression of developmental genes coordinating the adolescent maturation of PFC dopamine circuitry. Although it is important to

keep in mind the limitation of non-contingent drug administration models and that the duration of therapeutic treatment in humans may vary and sometimes last until adulthood, these and our previous findings [9, 10] show that therapeutic versus abused doses of amphetamine in adolescence have very different long-term consequences: while abused-like doses disrupt miR-218/DCC/Netrin-1-dependent dopamine development and behavioral control, therapeutic-like doses actually increase DCC expression and improve cognitive performance in adulthood. Our findings provide insight into the critical question of whether therapeutic exposure to stimulant drugs in adolescence induces detrimental effects on ongoing neurodevelopmental events.

6. References

1. Spear LP. The adolescent brain and age-related behavioral manifestations. Neurosci Biobehav Rev. 2000;24(4):417-463.

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

3. Hoops D, Reynolds LM, Restrepo-Lozano JM, Flores C. Dopamine development in the mouse orbital prefrontal cortex is protracted and sensitive to amphetamine in adolescence. eNeuro. 2018;5(1).

4. Reynolds LM, Pokinko M, Torres-Berrio A, et al. DCC receptors drive prefrontal cortex maturation by determining dopamine axon targeting in adolescence. Biol Psychiatry. 2018a;83(2):181-192.

5. Reynolds LM, Yetnikoff L, Pokinko M, et al. Early adolescence is a critical period for the maturation of inhibitory behavior. Cereb Cortex. 2018b.

6. Manitt C, Mimee A, Eng C, et al. The netrin receptor DCC is required in the pubertal organization of mesocortical dopamine circuitry. J Neurosci. 2011;31(23):8381-8394.

7. Grant A, Hoops D, Labelle-Dumais C, et al. Netrin-1 receptor-deficient mice show enhanced mesocortical dopamine transmission and blunted behavioural responses to amphetamine. Eur J Neurosci. 2007;26(11): 3215-3228.

8. Manitt C, Eng C, Pokinko M, et al. DCC orchestrates the development of the prefrontal cortex during adolescence and is altered in psychiatric patients. Transl Psychiatry. 2013;3:e338.

9. Reynolds LM, Makowski CS, Yogendran SV, Kiessling S, Cermakian N, Flores C. Amphetamine in adolescence disrupts the development of medial prefrontal cortex dopamine connectivity in a DCC-dependent manner. Neuropsychopharmacology. 2015;40(5):1101-1112.

10. Cuesta S, Restrepo-Lozano JM, Silvestrin S, et al. Non-contingent exposure to amphetamine in adolescence recruits miR-218 to regulate DCC expression in the VTA. Neuropsychopharmacology. 2018;43(4):900-911.

11. Yetnikoff L, Almey A, Arvanitogiannis A, Flores C. Abolition of the behavioral phenotype of adult netrin-1 receptor deficient mice by exposure to amphetamine during the juvenile period. Psychopharmacology (Berl). 2011;217(4):505-514.

12. Torres-Berrío A, Lopez JP, Bagot RC, et al. DCC confers susceptibility to depression-like behaviors in humans and mice and is regulated by miR-218. Biol Psychiatry. 2017;81(4):306-315.

13. Boomhower SR, Newland MC. Effects of adolescent exposure to methylmercury and d-amphetamine on reversal learning and an extradimensional shift in male mice. Exp Clin Psychopharmacol. 2017;25(2):64-73.

14. Hammerslag LR, Waldman AJ, Gulley JM. Effects of amphetamine exposure in adolescence or young adulthood on inhibitory control in adult male and female rats. Behav Brain Res. 2014;263:22-33.

15. Hankosky ER, Kofsky NM, Gulley JM. Age of exposure-dependent effects of amphetamine on behavioral flexibility. Behav Brain Res. 2013;252:117-125.

16. Sherrill LK, Stanis JJ, Gulley JM. Age-dependent effects of repeated amphetamine exposure on working memory in rats. Behav Brain Res. 2013;242:84-94.

17. Änggård E, Gunne L-M, Jönsson L-E, Niklasson F. Pharmacokinetic and clinical studies on amphetamine dependent subjects. Eur J Clin Pharmacol. 1970;3(1):3-11.

18. Cravey RH, Jain NC. Testing for amphetamines: medico-legal hazards. Trauma. 1973;15(49).

19. Gustavsen I, Morland J, Bramness JG. Impairment related to blood amphetamine and/or methamphetamine concentrations in suspected drugged drivers. Accid Anal Prev. 2006;38(3):490-495.

20. Kramer JC, Fischman VS, Littlefield DC. Amphetamine abuse. JAMA. 1967;201:305-309.

21. Riffee WH, Ludden TM, Wilcox RE, Gerald MC. Brain and plasma concentrations of amphetamine isomers in mice. J Pharmacol Exp Ther. 1978;206(3):586-594.

22. Van Swearingen AE, Walker QD, Kuhn CM. Sex differences in novelty- and psychostimulant-induced behaviors of C57BL/6 mice. Psychopharmacology (Berl). 2013;225(3):707-718.

23. Borcherding BG, Keysor CS, Cooper TB, Rapoport JL. Differential effects of methylphenidate and dextroamphetamine on the motor activity level of hyperactive children. Neuropsychopharmacology. 1989;2(4):255-263.

24. Brown GL, Hunt RD, Ebert MH, Bunney WE Jr, Kopin IJ. Plasma levels of d-amphetamine in hyperactive children. Serial behavior and motor responses. Psychopharmacology (Berl). 1979;62(2):133-140.

25. Greenhill LL, Swanson JM, Steinhoff K, et al. A pharmacokinetic/pharmacodynamic study comparing a single morning dose of Adderall to twice-daily dosing in children with ADHD. J Am Acad Child Adolesc Psychiatry. 2003;42(10):1234-1241.

26. Kramer WG, Read SC, Tran B, Zhang Y, Tulloch SJ. Pharmacokinetics of mixed amphetamine salts extended release in adolescents with ADHD. CNS Spectr. 2005;10(S15):6-13.

27. McGough JJ, Biederman J, Greenhill LL, et al. Pharmacokinetics of SLI381 (ADDERALL XR), an extended-release formulation of Adderall. J Am Acad Child Adolesc Psychiatry. 2003;42(6):684-691.

28. Gerasimov MR, Franceschi M, Volkow ND, et al. Comparison between intraperitoneal and oral methylphenidate administration: a microdialysis and locomotor activity study. J Pharmacol Exp Ther. 2000;295(1): 51-57.

29. Pan WJ, Hedaya MA. An animal model for simultaneous pharmacokinetic/pharmacodynamic investigations: application to cocaine. J Pharmacol Toxicol Methods. 1998;39(1):1-8.

30. Adriani W, Macri S, Pacifici R, Laviola G. Peculiar vulnerability to nicotine oral self-administration in mice during early adolescence. Neuropsychopharmacology. 2002;27(2):212-224.

31. Hefner K, Holmes A. Ontogeny of fear-, anxiety- and depression-related behavior across adolescence in C57BL/6J mice. Behav Brain Res. 2007;176(2):210-215.

32. Laviola G, Macri S, Morley-Fletcher S, Adriani W. Risk-taking behavior in adolescent mice: psychobiological determinants and early epigenetic influence. Neurosci Biobehav Rev. 2003;27(1-2):19-31.

33. Schneider M. Adolescence as a vulnerable period to alter rodent behavior. Cell Tissue Res. 2013;354(1):99-106.

34. Tirelli E, Laviola G, Adriani W. Ontogenesis of behavioral sensitization and conditioned place preference induced by psychostimulants in laboratory rodents. Neurosci Biobehav Rev. 2003;27(1-2):163-178.

35. Paxinos G, Franklin KBJ. The Mouse Brain in Stereotaxic Coordinates, Compact 2nd edn. Amsterdam; Boston: Elsevier Academic Press; 2004.

36. Manitt C, Labelle-Dumais C, Eng C, et al. Peri-pubertal emergence of UNC-5 homologue expression by dopamine neurons in rodents. PLoS ONE. 2010;5(7):e11463.

37. Aranda PS, LaJoie DM, Jorcyk CL. Bleach gel: a simple agarose gel for analyzing RNA quality. Electrophoresis. 2012;33(2):366-369.

38. Miner LH, Schroeter S, Blakely RD, Sesack SR. Ultrastructural localization of the norepinephrine transporter in superficial and deep layers of the rat prelimbic prefrontal cortex and its spatial relationship to probable dopamine terminals. J Comp Neurol. 2003;466(4):478-494.

39. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat. 1979;1:65-70.

40. Deiner MS, Kennedy TE, Fazeli A, Serafini T, Tessier-Lavigne M, Sretavan DW. Netrin-1 and DCC mediate axon guidance locally at the optic disc: loss of function leads to optic nerve hypoplasia. Neuron. 1997;19(3):575-589.

41. Goldman JS, Ashour MA, Magdesian MH, et al. Netrin-1 promotes excitatory synaptogenesis between cortical neurons by initiating synapse assembly. J Neurosci. 2013;33(44):17278-17289.

42. Seguela P, Watkins KC, Descarries L. Ultrastructural features of dopamine axon terminals in the anteromedial and the suprarhinal cortex of adult rat. Brain Res. 1988;442(1):11-22.

43. Gubner NR, Wilhelm CJ, Phillips TJ, Mitchell SH. Strain differences in behavioral inhibition in a Go/No-go task demonstrated using 15 inbred mouse strains. Alcohol Clin Exp Res. 2010;34(8):1353-1362.

44. Zombeck JA, Gupta T, Rhodes JS. Evaluation of a pharmacokinetic hypothesis for reduced locomotor stimulation from methamphetamine and cocaine in adolescent versus adult male C57BL/6J mice. Psychopharmacology (Berl). 2008;201(4):589-599.

45. Hu G, Zhang S, Vidal M, La Baer J, Xu T, Fearon ER. Mammalian homologs of seven in absentia regulate DCC via the ubiquitin-proteasome pathway. Genes Dev. 1997;11(20):2701-2714.

46. Rott R, Szargel R, Haskin J, et al. Monoubiquitylation of α-synuclein by seven in absentia homolog (SIAH) promotes its aggregation in dopaminergic cells. J Biol Chem. 2008;283(6):3316-3328.

47. Robbins TW. Dissociating executive functions of the prefrontal cortex. Philos Trans R Soc Lond B Biol Sci. 1996;351(1346):1463-1470. discussion 1470-1461

48. Glasgow SD, Labrecque S, Beamish IV, et al. Activity-dependent Netrin-1 secretion drives synaptic insertion of GluA1-containing AMPA receptors in the hippocampus. Cell Rep. 2018;25(1):168-182 e166.

49. Kang S, Wu MM, Galvez R, Gulley JM. Timing of amphetamine exposure in relation to puberty onset determines its effects on anhedonia, exploratory behavior, and dopamine D1 receptor expression in young adulthood. Neuroscience. 2016;339:72-84.

50. Burgos H, Cofre C, Hernandez A, et al. Methylphenidate has long-lasting metaplastic effects in the prefrontal cortex of adolescent rats. Behav Brain Res. 2015;291:112-117.

51. Cirulli F, Laviola G. Paradoxical effects of D-amphetamine in infant and adolescent mice: role of gender and environmental risk factors. Neurosci Biobehav Rev. 2000;24(1):73-84.

52. Hankosky ER, Gulley JM. Adolescent Exposure to Amphetamines and Vulnerability to Addiction. In Neuropathology of Drug Addictions and Substance Misuse. Academic Press; 2016;292-299.

53. Shahbazi M, Moffett AM, Williams BF, Frantz KJ. Age-and sexdependent amphetamine self-administration in rats. Psychopharmacology (Berl). 2008;196(1):71-81.

54. Argento JK, Arvanitogiannis A, Flores C. Juvenile exposure to methylphenidate reduces cocaine reward and alters netrin-1 receptor expression in adulthood. Behav Brain Res. 2012;229(1):202-207.

Connecting statement to Chapter III

Chapter 2 of this thesis has provided detailed insights into the dose-dependent effects of amphetamine on the molecular and behavioral aspects of mesocorticolimbic dopamine development in early adolescent male mice. Through a multidisciplinary investigation, we have discerned the significant effects of recreational- and therapeutic-like doses of amphetamine on miR-218/DCC/Netrin-1 expression, mesocortical dopamine development, and inhibitory control. Our findings show that therapeutic-like doses increase DCC expression and enhance cognitive performance in adulthood, while recreational-like doses downregulate the expression of DCC in the tissue, disrupting axon guidance and inducing cognitive impairment.

In Chapter 3, I embarked on a broader exploration of the developmental processes that underlie psychiatric disorders characterized by deficits in inhibitory control. The underlying neurobiological events and early markers of vulnerability remain areas of interest, reflecting a broader concern with the development of the prefrontal cortex (PFC) and its connections with the nucleus accumbens (NAcc). I delve into the pivotal role of the Netrin-1/DCC guidance cue system in corticolimbic development, and the dynamic relationship between mesocortical and mesolimbic dopamine inputs. I also explore a cutting-edge systems biology approach that moves beyond traditional genotype-disease associations to construct an expression-based polygenic score (ePRS). This new methodological approach allows for the identification of early vulnerability markers for impulsivity-related phenotypes and offers a fresh perspective on the genetic underpinnings of inhibitory control.

Chapter 3 is a translational study that builds upon the molecular and behavioral insights uncovered in Chapter 2 and extends the inquiry into gene co-expression networks and developmental trajectories that shape cognitive control. By integrating novel approaches like ePRS and focusing on corticolimbic-specific *DCC* gene co-expression networks, I aim to help better understand the quantifiable relationships

between genomics, brain development, and behavior. This chapter represents a step towards understanding the multifaceted nature of psychiatric disorders of developmental origin and opens new avenues for future research.

Chapter III: CORTICOLIMBIC DCC GENE CO-EXPRESSION NETWORKS AS PREDICTORS OF IMPULSIVITY IN CHILDREN

Authors:

Jose M. Restrepo-Lozano[1,2], Irina Pokhvisneva[2,3], Zihan Wang[2,3], Sachin Patel[2,3], Michael J. Meaney[2,3,4,5,6], Patricia P. Silveira[2,3,5#], and Cecilia Flores[2,4,5#]

Affiliations:

[1]Integrated Program in Neuroscience, McGill University, Montreal, Quebec, Canada, [2]Douglas Mental Health University Institute, Montreal, Quebec, Canada, [3]Ludmer Centre for Neuroinformatics and Mental Health, Douglas Mental Health University Institute, McGill University, Montreal, Quebec, Canada, [4]Department of Neurology and Neurosurgery, Faculty of Medicine, McGill University, Douglas Mental Health University Institute, Montreal, Quebec, Canada, [5]Department of Psychiatry, Faculty of Medicine, McGill University, Montreal, Quebec, Canada, [6]Singapore Institute for Clinical Sciences, Agency for Science, Technology and Research (A*STAR), Brenner Centre for Molecular Medicine, Singapore.

Published in

Molecular Psychiatry. 2022 Jun; 27(6): 2742-2750.

https://doi.org/10.1038/s41380-022-01533-7

1. Abstract

Inhibitory control deficits are prevalent in multiple neuropsychiatric conditions. The communication- as well as the connectivity- between corticolimbic regions of the brain are fundamental for eliciting inhibitory control behaviors, but early markers of vulnerability to this behavioral trait are yet to be discovered. The gradual maturation of the prefrontal cortex (PFC), in particular of the mesocortical dopamine innervation, mirrors the protracted development of inhibitory control; both are present early on in life, but they reach full maturation by early adulthood. Evidence suggests the involvement of the Netrin-1/*DCC* signaling pathway and its associated gene networks in corticolimbic development. Here we investigated whether an expression-based polygenic score (ePRS) based on corticolimbic-specific *DCC* gene co-expression networks associates with impulsivity-related phenotypes in community samples of children. We found that lower ePRS scores associate with higher measurements of impulsive choice in 6-year-old children tested in the Information Sampling Task- and impulsive action in 6- and 10-year-old children tested in the Stop Signal Reaction Time Task. We also found the ePRS to be a better overall predictor of impulsivity when compared to a conventional PRS score comparable in size to the ePRS (4515 SNPs in our discovery cohort) derived from the latest GWAS for ADHD. We propose that the corticolimbic *DCC*-ePRS can serve as a novel type of marker for impulsivity-related phenotypes in children. By adopting a systems biology approach based on gene co-expression networks and genotype-gene expression (rather than genotype-disease) associations, these results further validate our methodology to construct polygenic scores linked to the overall biological function of tissue-specific gene networks.

2. Introduction

Several psychiatric disorders of developmental origin are characterized by deficits in cognitive control – a compromised ability to voluntarily choose a context-appropriate goal-directed response. Altered connections and communication between prefrontal and striatal regions appear to be at the core of this behavioral trait [1], but the underlying neurobiological processes, as well as early markers of vulnerability, are yet to be discovered [2-4]. The cognitive capacity to control and override impulsive behaviors improves gradually from childhood to early adulthood, mirroring the protracted developmental trajectory of the prefrontal cortex (PFC) [5-9], and its gradual quantitative and qualitative changes in dopamine innervation [10-12]. While dopamine axons establish local connections in the nucleus accumbens (NAcc) in adolescence, mesocortical dopamine axons are still growing from the NAcc to the PFC across this period [13-18]. The extent and organization of mesocortical dopamine axon growth in adolescence determines the organization of local PFC circuitry and cognitive function in adulthood [19-21].

The developmental trajectories of mesocortical and mesolimbic dopamine inputs are temporally different but have a reciprocal functional connection [21-24]. Transient postnatal developmental overexpression of dopamine D2 receptor in the striatum leads to adult mesocortical dopamine PFC dysfunction and cognitive deficits, indicating that striatal dopamine maturational events interact with those controlling mesocortical dopamine axon growth [25]. Changes in PFC dopamine neurotransmission are associated with opposite changes in NAcc dopamine function [23, 26], and alterations in D1- or D2-expressing NAcc pathways impact gene expression in the PFC [27]. Clearly, PFC and cognitive control development involve the recruitment of corticostriatal neuronal networks [28, 29].

A rapidly increasing number of studies show a strong association between genetic variability within the Netrin-1 guidance cue receptor gene, *DCC,* and several psychiatric disorders of developmental onset,

most notably those emerging in adolescence. These disorders are characterized by PFC and NAcc dysfunction and deficits in impulse control [30-33]. Early postnatal expression of the *DCC* gene network in the PFC associates with total brain volume in children [34], emphasizing that the Netrin-1/*DCC* guidance cue system is tightly linked to overall early neurodevelopment. In adolescent rodents, DCC-mediated Netrin-1 signaling organizes the maturation of dopamine networks by promoting mesolimbic dopamine axon targeting in the NAcc and controlling the growth of dopamine axons to the PFC [12, 17]. Changes in DCC receptor levels in adolescent mice lead to mistargeting of mesolimbic dopamine axons in the NAcc and to their ectopic growth to PFC, altering PFC function and cognitive control in adulthood [17, 35]. Similar anatomical and behavioral changes occur in humans that are *DCC* mutation carriers [36, 37], indicating that the Netrin-1/*DCC* pathway is part of a gene network closely involved in corticolimbic development.

To date, most human genetic studies have focused on associations between genetic variants and phenotypes, and the estimated effects of a given number of variants can be aggregated into a score that represents individual genetic risk (called polygenic risk score; PRS). This association between genetic variation and behavior/disease ultimately results in relatively few genome-wide significant variants (e.g. [38]), most of which belong to noncoding portions of the genome and whose effect is diminished by the polygenic nature of complex phenotypes [39, 40]. Several of these non-coding variants are regulatory in nature, likely affecting the expression of nearby genes [40], ultimately placing gene expression as an intermediate molecular phenotype between genetics and disease [41]. Our approach exploits the facts that genes operate within complex networks, code with remarkable tissue-specificity for precise biological functions, and the likelihood of identifying relevant biological markers increases by relying on genotype-gene expression rather than genotype-disease associations (see [42, 43]). We use a systems biology approach to create a genetic score based on genes co-expressed with a gene of interest in a specific brain region. We gather all SNPs from the co-expressed genes and assign for each SNP the effect

size estimated by the Genome-Tissue Expression (GTEx) project [44], which quantifies the influence of variants on tissue-specific gene expression. We aggregate genotypes weighted by the GTEx across all SNPs within the co-expression network into an expression-based polygenic score (ePRS), according to the individual's genotype [42, 43].

The relationship between genes and behavior is highly indirect, regardless of how strong the relationship may be. Here, we forgo direct genotype-disease associations to construct an ePRS based on genes co-expressed with *DCC* in the PFC and the NAcc. Our goal is to create a marker that captures individual variation in the processes involved in the maturation of corticolimbic substrates supporting inhibitory control. By modifying the approach to genomic profiling, we generated a biological marker that can help identify early vulnerability for impulsivity-related phenotypes. We tested the association of the biologically-informed genetic score with measurements of impulsivity in three ethnically different community samples of 6- and 10-year-old children.

3. Materials and Methods

Detailed description is provided in the Supplemental Materials and Methods.

Participants

We used genomic and phenotypic data from three prospective birth cohorts: 1) Maternal Adversity, Vulnerability and Neurodevelopment (MAVAN, [45]), 2) Growing Up in Singapore Towards Healthy Outcomes (GUSTO, [46]), and 3) Avon Longitudinal Study on Parents and Children (ALSPAC, detailed block diagram in Figure S1, [47, 48]). Informed consent was obtained from each participant, and the use of these data has been approved by: 1) McGill University, Université de Montréal, Royal Victoria Hospital, Jewish General Hospital, Centre hospitalier de l'Université de Montréal, Hôpital Maisonneuve-Rosemount, St Joseph's Hospital, and McMaster University for MAVAN; 2) The National Healthcare Group Domain Specific Review Board and the Sing Health Centralized Institutional Review Board for GUSTO; and 3) the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. See supplementary table S6 for a summary of the genotyping information for each cohort.

Identification of corticolimbic DCC gene co-expression networks and ePRS calculation

Figure 1 shows the steps involved in the identification of the gene networks and the ePRS score. The ePRS was calculated considering genes co-expressed with *DCC* in the PFC and NAcc. We aimed to capture *DCC* co-expression networks within each brain region, with the final ePRS being a joint representation of the functional co-expression networks in these two corticolimbic regions. As described previously [42, 43, 49], the score was created using the data from: 1) GeneNetwork (http://genenetwork.org), 2) BrainSpan (http://www.brainspan.org), 3) the NCBI Variation Viewer, U.S. National Library of Medicine, (NCBI) [50], 4) the GTEx project [44], and 5) genotype data in the three cohorts. We used GeneNetwork to generate a list of genes co-expressed with *DCC* in the PFC and in the

NAcc in mice (absolute value of co-expression correlation greater or equal to 0.5). We used gene expression datasets from mice (see supplementary data file) because our study is guided by our previous findings in rodents linking variations in *Dcc* expression to changes in impulse control and in mesocorticolimbic dopamine axon targeting [13, 17, 52, 52]. To retain genes that are more active when the brain is still undergoing core maturational processes in humans, we used BrainSpan to select autosomal transcripts expressed at least 1.5-fold more during the early postnatal development (0-18 months after birth) than in adulthood (20-40 years of age), with the final networks consisting of 154 genes in the PFC (see Table S4) and 72 genes in the NAcc (see Table S5). For annotations, we used GRCh37.p13 assembly of the NCBI to source chromosome and start/end position for the co-expressed genes, which, in turn, were used to gather all the gene-SNP pairs from the GTEx dataset in human PFC and NAcc (PFC: 41,053 SNPs, NAcc: 66,428 SNPs). These lists were merged with the genotyping data in each of the three cohorts, keeping only the common SNPs and subjecting the final genotyping data sets to linkage disequilibrium clumping (r^2<0.2) to eliminate highly correlated SNPs.

To calculate the ePRS, number of effect alleles at a given cis-SNP were weighted by the estimated brain-region-specific effect of the genotype on gene expression from the GTEx data. The ePRS was obtained by adding the weighted SNPs, accounting for the sign of the correlation between each gene's expression and *DCC* gene expression. The sum of the estimated effects resulted in ePRS scores for the *DCC* co-expression networks in the PFC and NAcc, which were then aggregated (by summation of the two scores) into a single global genetic score termed "corticolimbic *DCC*-ePRS".

Finally, an enrichment analysis was conducted to characterize the functional and biological properties of the gene networks that comprise the corticolimbic *DCC*-ePRS score. A description of the tools and datasets used throughout the study can be found in the supplemental material, Table S7.

Behavioral Outcomes

We tested whether the ePRS associates with two aspects of impulsivity: *(i)* impulsive choice, reflecting proneness to make risky choices, as measured by the Information Sampling Task; *(ii)* impulsive action, reflecting the ability to inhibit motor responses, as measured by the Stop-Signal Task. In both cases, the ability to self-regulate behavior is required for interrupting or inhibiting competing inputs or actions in order to accomplish a specific goal-directed response [1]. A description of the behavioral data obtained from each cohort is described in the supplemental material.

Statistical Analysis

Data were analyzed using R v3.6 [53] and Python v3.7 (https://www.python.org/), and polygenic scores were generated using the PRSoS pipeline (https://github.com/MeaneyLab/PRSoS). We considered two-tailed hypothesis tests, and significance levels for all tests were set at $\alpha < 0.05$. For each cohort we categorized the ePRS into high- and low-ePRS groups using a median split of the genetic score. Analysis of baseline characteristics was performed using Student's *t*-test for continuous data (in case of unequal variances, Welch's t-test was used) and X^2 for categorical variables. Linear regression analysis was used to examine the association of the ePRS with the behavioral outcomes, adjusting for sex and population stratification. Adjustment for multiple comparisons was applied using Bonferroni method, independently for each behavioral construct/cohort. All data were inspected to ensure that the assumptions for the tests and the linear regression analyses were met. Power analysis was conducted for linear multiple regression, considering the effect of the ePRS on the different outcomes: for $\alpha = 0.05$, sample size of 202, 398, and 4392, and a small effect size $f^2 = 0.02$, the achieved power will be .64, .88, and >.95 in MAVAN, GUSTO, and ALSPAC respectively.

4. Results

We found no differences in baseline characteristics between ePRS groups in MAVAN, GUSTO, and ALSPAC cohorts (Table 1).

Behavioral Outcomes

Lower Corticolimbic DCC-ePRS Scores Associate with Higher Measurements of Impulsivity in Children

Information Sampling Task: In MAVAN, the ePRS was associated with meanP-correct values (Figure 2A: β = -0.04, *p* = .045); the low-ePRS group had lower meanP-correct values (less information sampled, lowering the probability of making a correct choice at the point of decision) indicating higher levels of impulsive choice in comparison to the high-ePRS group.

Stop-Signal Task: In GUSTO, the ePRS was associated with the proportion of successfully inhibited responses (Figure 2B: β = -0.03, *p* = .027; Figure S3 for complete results); low-ePRS group has a lower proportion of successful stops compared to the high-ePRS group, indicating higher levels of impulsive action. In ALSPAC, the ePRS was associated with measurements of impulsive action (Figure 2C: β = -10.368, *p* = .019; Figure S4 for complete results), with the low-ePRS group showing a shorter mean reaction time in unsuccessful stop trials- indicating higher levels of impulsive action- compared to the high-ePRS group. Also in ALSPAC, there are no differences between ePRS groups when comparing the proportion of successful stops, but methodological differences (full details in supplementary methods; see Figure S2) in the way the task was conducted in CANTAB ("stop" signal delay was adjusted to subject's performance in MAVAN and GUSTO) versus ALSPAC (fixed delay of 250ms was applied irrespective of subject's performance) prevent the direct comparison of successfully inhibited responses between these cohorts. In MAVAN, we found no association between the ePRS and performance in this task (Proportion of successful stops: β = -0.03, *p* = .10; Estimated SSRT: β = -10.07, *p* = .57; see Table S2

MAVAN				
Sample description	*Total (n = 202)*	*Low ePRS (n = 96)*	*High ePRS (n = 106)*	*p-value*
Sex – male (n)	49.5% (100)	56.3% (54)	43.4% (46)	0.09
Maternal age at birth (years)	30.72 (4.90)	31.42 (5.07)	30.08 (4.67)	0.053
Gestational age (weeks)	39.03 (1.23)	38.89 (1.30)	39.16 (1.15)	0.11
Birth weight (g)	3313 (452)	3300 (450)	3224 (456)	0.71
Maternal education – University degree or above	55.7% (108)	54.3% (50)	56.9% (58)	0.47
Low family income	20.6% (50)	25.8% (23)	27.3% (27)	0.96
GUSTO				
Sample description	*Total (n = 398)*	*Low ePRS (n=202)*	*High ePRS (n=196)*	*p-value*
Sex – male (n)	53.3% (212)	49.5% (100)	57.1% (112)	0.15
Maternal age at birth (years)	31.55 (5.04)	31.81 (5.34)	31.28 (4.73)	0.37
Gestational age (weeks)	38.49 (1.28)	38.47 (1.33)	38.50 (1.23)	0.87
Birth weight (g)	3137 (416)	3127 (441)	3147 (391)	0.39
Maternal education – University degree or above	35.64% (103)	38.36% (56)	32.88% (47)	0.39
Household income < $2000 SGD per month	12.11% (35)	15.75% (23)	8.39% (12)	0.08
ALSPAC				
Sample description	*Total (n = 4392)*	*Low ePRS (n=2210)*	*High ePRS (n=2182)*	*p-value*
Sex – male (n)	49.2% (2159)	48.6% (1075)	49.7% (1084)	0.51
Maternal age at birth (years)	29.31 (4.47)	29.31 (4.47)	29.31 (4.47)	0.97
Gestational age (weeks)	39.76 (1.27)	39.79 (1.26)	39.73 (1.28)	0.11
Birth weight (g)	3499 (465)	3503 (458)	3496 (471)	0.61
Maternal education – University degree or above	18.9% (786)	19.7% (411)	18.2% (375)	0.25
Low SES [a]	35.0% (1537)	34.5% (762)	35.5% (775)	0.49

Table 1. Description of baseline characteristics of the 3 cohort samples. Continuous variables are expressed as mean (SD); categorical variables are expressed as percentage (number of subjects). [a] We used "crowding index" as a proxy measure for SES. This index was calculated by dividing the number of individuals living in the family's residence, by the number of rooms in the residence, and we considered low SES when crowding index > 0.75, and high SES when crowding index <= 0.75.

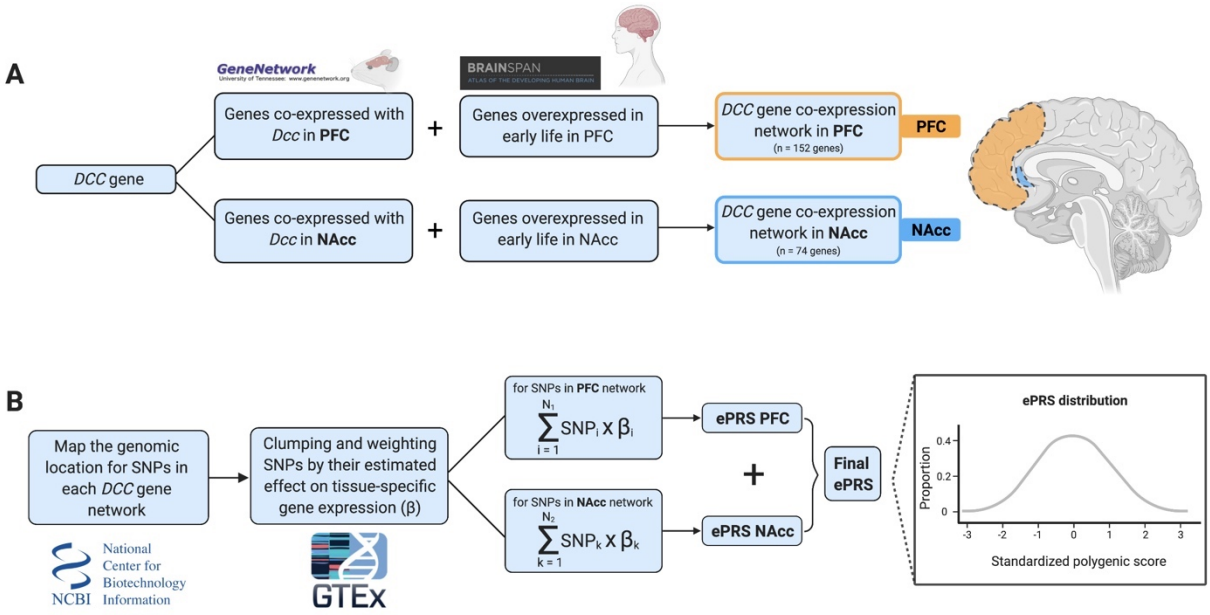

Figure 1. Flowchart depicting the steps involved in the creation of the corticolimbic DCC-ePRS score. (A) The GeneNetwork database was used to generate a Dcc gene co-expression matrix in the PFC and NAcc in mice. Genes with a correlation of co-expression $\geq |0.5|$ were retained. Brainspan was used to identify human homologous transcripts and to filter each gene list by selecting the transcripts enriched during the first 18 months of life, as compared to adulthood, defined by a differential expression ≥ 1.5, within the same brain area. Each resulting gene list comprised the DCC co-expression network for their respective brain area. (B) Based on their annotation in the NCBI library, using GRCh37.p13 assembly, common SNPs within each co-expression network, GTEx data base, and genotyping cohort were subjected to linkage disequilibrium clumping to remove highly correlated SNPs ($r^2 \geq 0.2$). Using data from the GTEx project, alleles at a given cis-SNP were weighted by the estimated brain-region-specific effect of the genotype on gene expression. The sum of these estimated effects resulted in ePRS scores for the DCC co-expression networks in the PFC and NAcc, which we aggregated into a single global ePRS score.

for complete results). After adjustments for multiple comparisons, the association between the ePRS and the proportion of successful stops in GUSTO cohort is no longer significant (p = .054).

These results show that an ePRS score reflecting variability in the expression of corticolimbic *DCC* gene co-expression networks is associated with the levels of inhibitory control in children from ethnically diverse backgrounds.

Corticolimbic DCC Gene Co-expression Networks: Enrichment Analysis

Protein-Protein Interaction (PPI)

We used STRING [54] and Cytoscape [55] to visualize catalogued PPIs in protein products of genes within each co-expression network (networks were analyzed separately; only proteins with interactions are depicted in Figure 3A). *PFC:* This network contains 152 nodes (one for each protein with at least 1 connection with another protein in the network) and 151 edges, corresponding to the mapped interactions among the nodes. The PPI enrichment (p = 0.004) indicates that this network contains more interactions than expected, compared to a network of equal size composed of a random set of proteins, and that the proteins are involved in common biological functions. *NAcc:* Contains 74 nodes and 50 edges, and the PPI enrichment (p = 5.1e-11) also suggests a strong biological connection among the proteins (see the corresponding gene networks, created using GeneMANIA [56], in Figure S5).

Tissue-Specific Gene Expression

We used FUMA [57] to visualize the expression levels of the genes from the co-expression networks across the 54 tissue-types included in GTEx. The PFC (comprising BA24 and BA9) and the NAcc are the 1st and 4th most enriched tissues for the gene networks' expression (Figure 3B).

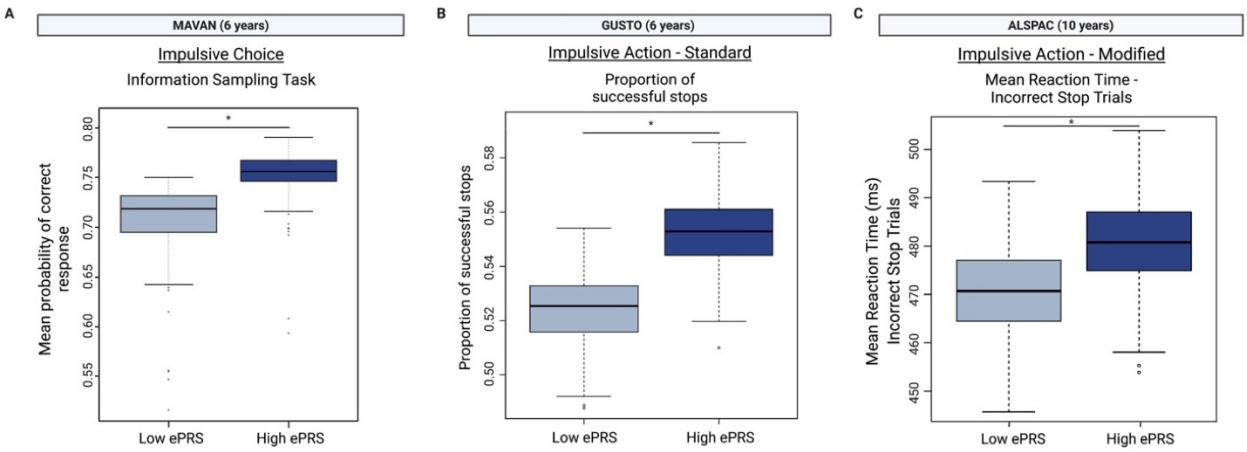

Figure 2. Association between the corticolimbic DCC-ePRS score and measurements of impulsivity, in (A) MAVAN kids (n=197) at 6 years of age (Information Sampling Task: β = -0.04, p = .045), (B) GUSTO kids (n=398) at 6 years of age (Proportion of successful stops: β = -0.03, p = .027), and (C) ALSPAC kids (n=4392) at 10 years of age (Mean RT – Incorrect stop trials: β = -10.36, p = .019). A lower DCC-ePRS score was associated with higher impulsive action and choice in the 3 ethnically-diverse cohorts. The Y-axis represents the predicted values of the measurements of impulsivity, the middle of the box is the median, the edges are the lowest and highest quartiles, and the error bars (whiskers) represent 1.5 x IQR (interquartile range). *p<.05.

Functional Ontologies

Using Metacore™, we explored the biological context in which the gene networks operate, by mapping the genes from each network onto functional categories (Figure 3C, Table S1 for more detailed results). The networks are enriched in synaptic components, predominantly in cell junction and plasma membrane regions (strongest enrichment for cell junction in PFC: p<3.3e-8, and for synapse in NAcc: p<1.3e-19). Enrichment for biological processes showed the involvement of the network in neurodevelopmental processes including neuronal differentiation and development (PFC: p<2.1e-4; NAcc: p<1e-19), neuron projection guidance (PFC: p<1e-3; NAcc: p<1e-16) and regulation of trans-synaptic signaling (PFC: p<2.9e-5; NAcc: p<3.09e-17). Enrichment for molecular functions showed a role of the networks in protein binding and cell adhesion (PFC: p<2.72e-8; NAcc: p<5.6e-6). These functions are fundamental to the establishment of brain connectivity, mainly via axon guidance (e.g. [58]) and synaptogenesis (e.g. [59]).

Developmental Gene expression

We assessed the enrichment of gene expression for each network across brain regions and developmental periods in humans, using the CSEA tool [60, 61]. Both networks are enriched across the brain during perinatal periods. Notably, the expression of the NAcc network in the PFC is enriched again during adolescence (Figure 3D: p=3.679e-04), in line with previous descriptions of the developmental trajectory mediating adolescent corticolimbic maturation [17, 25].

To explore the ability of the PFC and NAcc networks, and ultimately the ePRS itself, to capture transcriptionally co-regulated biological processes, we analyzed the networks' co-expression patterns in their corresponding brain regions during childhood and adulthood using the Brainspan dataset ([62]; Figure 4). *PFC:* In the heatmap representing correlation of gene expression during childhood, there are 3 main clusters of high co-expression, but only 1 cluster is maintained in adulthood. Finding that

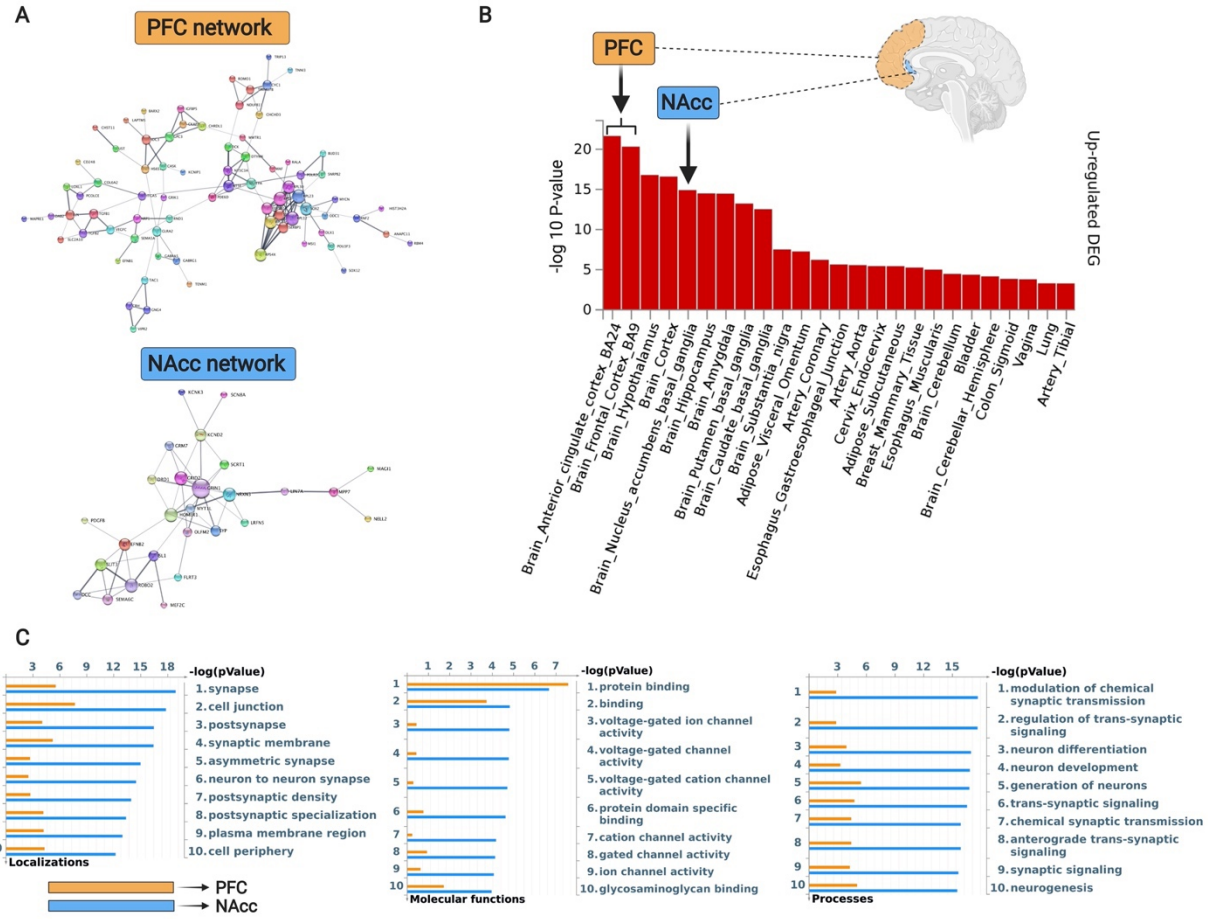
A
PFC network
NAcc network
B
PFC
NAcc
-log 10 P-value
Up-regulated DEG
Brain_Anterior_cingulate_cortex_BA24
Brain_Frontal_Cortex_BA9
Brain_Hypothalamus
Brain_Cortex
Brain_Nucleus_accumbens_basal_ganglia
Brain_Hippocampus
Brain_Amygdala
Brain_Putamen_basal_ganglia
Brain_Caudate_basal_ganglia
Brain_Substantia_nigra
Adipose_Visceral_Omentum
Artery_Coronary
Esophagus_Gastroesophageal_Junction
Artery_Aorta
Cervix_Endocervix
Adipose_Subcutaneous
Breast_Mammary_Tissue
Esophagus_Muscularis
Brain_Cerebellum
Bladder
Brain_Cerebellar_Hemisphere
Colon_Sigmoid
Vagina
Lung
Artery_Tibial
C
-log(pValue)
1. synapse
2. cell junction
3. postsynapse
4. synaptic membrane
5. asymmetric synapse
6. neuron to neuron synapse
7. postsynaptic density
8. postsynaptic specialization
9. plasma membrane region
10. cell periphery
Localizations
PFC
NAcc
-log(pValue)
1. protein binding
2. binding
3. voltage-gated ion channel activity
4. voltage-gated channel activity
5. voltage-gated cation channel activity
6. protein domain specific binding
7. cation channel activity
8. gated channel activity
9. ion channel activity
10. glycosaminoglycan binding
Molecular functions
-log(pValue)
1. modulation of chemical synaptic transmission
2. regulation of trans-synaptic signaling
3. neuron differentiation
4. neuron development
5. generation of neurons
6. trans-synaptic signaling
7. chemical synaptic transmission
8. anterograde trans-synaptic signaling
9. synaptic signaling
10. neurogenesis
Processes
D
PFC network
NAcc network
Amygdala
Cerebellum
Cortex
Hippocampus
Striatum
Thalamus
Early Fetal
Early Mid Fetal
Late Mid Fetal
Late Fetal
Neonatal Early Infancy
Late Infancy
Early Childhood
Mid_Late Childhood
Adolescence
Young Adulthood
p-values

Figure 2. Validation of the PFC and NAcc DCC co-expression networks. (A) Protein-Protein interaction (PPI) networks constructed from the gene co-expression networks in PFC and NAcc, using the Cytoscape software. The edges between the nodes indicate both functional and physical associations, and the size of the sphere is proportional to the degree of connectivity with other nodes. The protein networks represent known functional interactions between the protein products of the genes that make up the corticolimbic DCC gene networks; Significant PPI enrichment in the PFC ($p = 0.004$) and the NAcc ($p = 5.1e-11$). Tissue-specific gene expression analysis performed in FUMA (B) confirms that the genes that comprise both networks are highly upregulated in the PFC and NAcc, according to GTEx dataset v8. (C) A combined enrichment analysis for the co-expression networks performed in Metacore ™ shows enrichment for diverse neurodevelopmental processes, suggesting a common brain maturational role for the networks (full results with FDR adjusted values in Table S3). (D) Cell-type Specific Expression Analysis (CSEA) analysis reveals that the NAcc and PFC DCC co-expression networks are highly enriched throughout the brain during embryonic life and early infancy. However, the NAcc network is enriched again in the cortex during late childhood and adolescence ($p = 0.0004$ for Fisher's exact test, $p = 0.002$ after Benjamini-Hochberg correction). The hexagon levels mark the different degrees of stringency applied in the identification of selectively enriched transcripts for that brain region/developmental period. In each hexagon there are 4 levels, with the outer level representing the least stringent pSI value (0.05) and the inner-most level consisting of the most stringent pSI value (0.0001). The 2 hexagons in between represent a pSI = 0.01 and a pSI = 0.001. The size of the hexagon is proportional to the number of genes selectively enriched, and the color represents the FDR-adjusted p-values of the expected overlap between the genes in the network and the list of selectively enriched genes.

correlations of expression of genes in the PFC in childhood are not maintained in adulthood is in line with the marked developmental changes in the PFC transcriptome landscape previously described in humans and mice [63]. *NAcc:* The heatmap for childhood gene co-expression shows a large main cluster, containing several highly correlated smaller clusters. Many of the smaller NAcc clusters perdure into adulthood, indicating that the NAcc network is more stable than the PFC network.

Comparison between polygenic scores

We compared our ePRS to a traditional PRS for ADHD on the capacity to predict the same behavioral outcomes. For that, we selected the top 4515 most significant SNPs identified in the latest ADHD GWAS [64], which corresponds to the GWAS *p*-value threshold 4.912e-5, and created a score comparable in size to the ePRS in terms of number of SNPs. There was no association between the PRS and the main outcomes for MAVAN (meanP-correct: β=-0.01, *p*=0.48), GUSTO (proportion of successful inhibitions: β=-0.009, *p*=0.52; Figure S3) or ALSPAC (mean reaction time – incorrect stop trials: β=-3.248, *p*=0.46; Figure S4) cohorts. We also performed an enrichment analysis to characterize the functional/biological properties of the PRS genes and found that they are upregulated across the brain, but not as selectively- and to a lesser extent- than the genes from the ePRS. Finally, results from the CSEA show no selective spatiotemporal enrichment in the human brain (Figure S6).

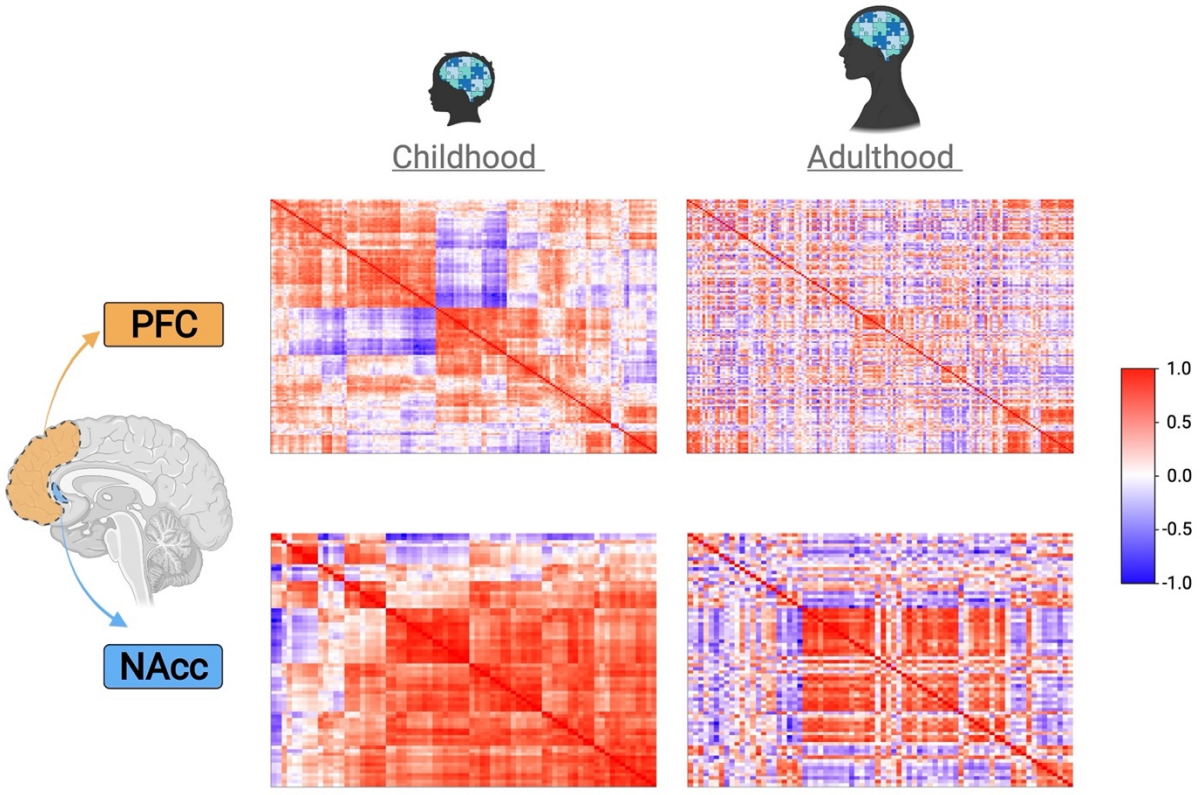

Figure 4. Heat map of the co-expression for the genes included in the corticolimbic DCC-ePRS, in human PFC (top panels) and NAcc (bottom panels). PFC: The heatmap of the co-expression in childhood (left) shows several clusters, while for the co-expression patterns in adulthood (right) most of the clusters are not maintained, suggesting that genes that are co-expressed during childhood in the PFC are rarely co-expressed in adulthood. NAcc: The heatmap in childhood (left) shows many clusters with a very high correlation of expression. Interestingly, a larger proportion of these clusters are maintained in adulthood (right) compared to the transition between childhood and adulthood in the PFC, indicating a more stable gene network. We retained the same order for the genes as in childhood, to be able to compare if the clusters that we observe in childhood are maintained in adulthood.

5. Discussion

Impulse control deficits are a common trait of numerous neurodevelopmental psychiatric disorders. Discovering their neurobiological underpinnings and early biomarkers will help identifying at risk individuals and improving/implementing early prevention and intervention strategies. Here, we generated an expression-based polygenic score (ePRS) consisting of SNPs within genes co-expressed with the axon guidance cue receptor gene, *DCC*, in the PFC and NAcc, to create a functional and corticolimbic-specific marker of vulnerability to heightened impulsivity. Our results show that the ePRS is significantly associated with different measures of impulsive behaviors in children from three ethnically diverse independent birth cohorts. Across all cohorts, the low-ePRS groups show higher impulsivity-related phenotypes. Detailed characterization of the gene networks comprising the corticolimbic *DCC*-ePRS show significant functional interactions, contribution to core neurodevelopmental processes, and enriched expression in cortical neurons, particularly from embryonic life to adolescence.

Most PRSs are characterized by a limited generalizability due to a marked disparity in prediction accuracy across different populations [65, 66]. This limitation, partially explained by the biased ancestry representation in most well-powered discovery GWASs, does not affect the ability of the ePRS to predict impulsive phenotypes across 3 independent birth cohorts from Canada, Singapore, and UK. Other studies that have implemented a similar approach to polygenic risk analysis by using the ePRS methodology reported a high predictive value of their genetic scores applied across diverse populations [42, 67]. To understand how the ePRS compares to a traditionally derived PRS, in this study we constructed a score based on the latest GWAS for ADHD and found that, even though the PRS score predicts impulsive behavior in one cohort (Figure S3), our ePRS predicts a larger number of outcomes, across all three cohorts. This is consistent with other studies that have observed a higher prediction accuracy of their ePRS compared to conventional PRSs [34, 42]. Since the ePRS methodology relies on

identifying tissue-specific gene networks and their function, instead of identifying scattered genetic variants across the genome, we are able to create a more biologically meaningful score compared to conventional PRSs. Finally, the associated phenotype to weigh the SNPs in our ePRS is gene expression, which, given the current state of technology, is a highly quantitative trait measurable with high precision, across different tissues and conditions, by high-throughput sequencing, and thus yielding a score that globally represents transcriptionally co-regulated biological processes. These results suggest that our genetic profiling approach increases the likelihood of identifying trait-relevant biological markers.

We identified co-expression networks for the guidance cue receptor, *DCC*, specifically in the NAcc and the PFC, to create a biological marker related to neurodevelopmental processes occurring in these regions, that could predict levels of impulsivity in children. In addition to establishing the ePRS' predictive power of impulsivity across 3 different cohorts, we found the co-expression networks to be highly enriched for protein-protein interactions, suggesting their involvement in common biological functions. Since DCC receptors are master organizers of neuronal circuits, and since variations in its expression in early life result not only in functional and anatomical alterations of neural pathways involved in inhibitory control, but also in alterations of inhibitory control itself [17, 36, 51], it is not surprising that *DCC* co-expression networks in these corticolimbic hubs associate with behavioral traits implicated in psychopathology. Indeed, proper establishment of neuronal circuits is essential to mental health [68]. The genes that make up the networks are highly upregulated in the PFC and the NAcc and are involved in a wide range of neurodevelopmental processes. This enrichment suggests a prominent role of the gene networks in the maturation of both PFC and NAcc circuits, validating the use of these networks as the basis for the ePRS calculation and their potential use as a functional biomarker to predict reflection and motor impulsivity in children.

Results from several studies in humans show that mutations in the *DCC* gene lead to dramatic neurodevelopmental changes, including agenesis of the corpus callosum [36, 69, 70], developmental split-brain syndrome [69], and congenital mirror movements [37, 70, 71]. Similar noticeable changes have been described in *Dcc* homozygous or haploinsufficient mice [72], highlighting the core role of *DCC* in neurodevelopmental wiring. As *DCC* expression shifts from high to low in adolescence, its functional role also shifts from broad organization of developing neuronal networks to the refinement of neuronal architecture, synaptogenesis and synaptic plasticity of established matured circuits [30, 31, 73]. Recent human studies have also shown that many polymorphisms in *DCC*, as well as altered levels of gene expression, are related to numerous neuropsychiatric conditions of developmental onset, some of which are characterized by deficits in PFC function and impulse control [30, 31]. Individual genes do not operate in isolation and cannot explain the entire spectrum of mental disorders, as it has been well established by a wealth of data from recent GWAS studies showing massive polygenicity among neuropsychiatric disorders. Therefore, DCC receptors act as a master organizer of specific synaptic circuits, as a part of a gene network, and we have shown that a PFC gene network for *DCC* is associated with overall brain structure [34]. Our functional analyses of the corticolimbic *DCC* gene networks suggest their implication in the development of the neural substrates underlying inhibitory control behaviors. The genes that comprise the networks are co-expressed in crucial brain regions (see Figure 3B, Figure 4), suggesting their spatial convergence. Furthermore, the expression of genes known to increase risk for neuropsychiatric disorders converge temporally, especially before and during the onset of the disorder [68]. Here we observed that gene expression for both networks is enriched during specific pre- and post-natal periods, including an enriched expression of the NAcc network in cortical neurons during late childhood and adolescence. As noted previously, the neurodevelopmental role that DCC plays changes as a function of developmental stage, and the fact that a DCC co-expression network is enriched again

during late childhood and adolescence suggest that alterations in its function/expression can impact the adolescent development of synaptic connectivity and function in the PFC later in life.

We propose a novel type of marker for impulsivity-related phenotypes in children. Our biologically-informed approach to polygenic risk analysis aims to capture variation in the function/expression of gene networks predominantly associated with PFC and NAcc maturation, two regions subserving inhibitory control. Whether integrating relevant SNPs associated with other forms of gene expression regulation beyond cis (e.g., transcription factors, promoter regions, and chromatin modifications) in non-coding regions changes the performance of the scores, will be investigated in future studies. Exploring the association between the ePRS and inhibitory control behaviors later in life is needed in order to investigate the possible use of this genetic marker as a probabilistic risk score for vulnerability phenotypes linked to psychopathologies of adolescent onset. Our results are an example of the utility of understanding the molecular processes that govern the development of a neural circuit, and how this knowledge can be applied to predict genetic susceptibility to endophenotypes linked to psychiatric conditions.

6. Supplementary Information

Supplemental Materials and Methods

Participants

We used genomic and phenotypic data from three prospective birth cohorts: 1) Maternal Adversity, Vulnerability and Neurodevelopment (MAVAN), 2) Growing Up in Singapore Towards Healthy Outcomes (GUSTO), and 3) Avon Longitudinal Study on Parents and Children (ALSPAC).

The MAVAN cohort consisted of children recruited from Montreal (Quebec) and Hamilton (Ontario), in Canada. Eligibility criteria for mothers specified being 18 years of age or older, with singleton pregnancies, and fluency in English or French. Approval for the MAVAN project was obtained from McGill University, Université de Montréal, Royal Victoria Hospital, Jewish General Hospital, Centre hospitalier de l'Université de Montréal, Hôpital Maisonneuve-Rosemount, St Joseph's Hospital, and McMaster University. Informed consent was obtained from each participant (for more information see []). Extensive phenotyping was carried out from birth, including measures of reflection impulsivity assessed with the Information Sampling Task (IST) at 6 years of age. After verification of complete genotypic and phenotypic data, we retained 202 children for the current study.

The GUSTO cohort involved children born at the National University Hospital (NUH) or KK Women's and Children's Hospital (KKH) in Singapore, between November 2009 and May 2011. The eligibility criteria for mothers specified being of Chinese, Malay or Indian ethnicity with homogeneous parental ethnic background, as well as being aged 18 years and above at the time of recruitment. The study was approved by the National Healthcare Group Domain Specific Review Board and the Sing Health Centralized Institutional Review Board. Informed written consent was obtained from each participant (for detailed information see [], or https://sicsdatavault.sg/gusto/). Extensive phenotyping was carried

out from birth, including measures of impulsivity in the offspring assessed with the Stop Signal Reaction Time Task (SSRT) at 6 years of age. After verification of complete genotypic and phenotypic data, we retained 398 children in the current study.

The ALSPAC cohort consisted of children with an expected delivery date between April 1 1991, and December 31 1992, from a geographically defined area in the Southwest of England. For data collected after the age of seven, the total sample size was 15,454 pregnancies, resulting in 15,589 fetuses, of whom, 14,901 were alive at 1 year of age (for complete description see []). Phenotypic data from subjects were collected in order to assess neurodevelopment, including measures of impulsivity assessed with the SSRT task at 10 years of age. After verification of complete genotypic and phenotypic data, we retained 4392 children in the current study (detailed block diagram in Figure S1). 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 (http://www.bristol.ac.uk/alspac/researchers/our-data/). Ethical 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/. Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

Genotyping

MAVAN: Autosomal SNPs were genotyped using genome-wide platforms (PsychArray/PsychChip, Illumina) according to manufacturer's guidelines, with genomic DNA derived from buccal epithelial cells. Quality control procedure was carried out using PLINK 1.951 (Purcell et al., 2007). Samples with a call rate less than 90% were removed. SNPs with a low call rate (< 95%), minor allele frequency (MAF) < 5%, and low p-values on Hardy-Weinberg Equilibrium exact test ($p < 1e\text{-}40$) were removed, which resulted in

a total of 242,211 SNPs. Imputation using the Sanger Imputation Service (McCarthy et al., 2016) and the Haplotype Reference Consortium (HRC) as the reference panel (release 1.1) was performed, resulting in 20,790,893 autosomal SNPs with an info score > 0.80 [].

*GUSTO***:** Genotyping was performed using Illumina OmniExpressExome array and split by ethnicity for quality checks. Non-autosomal SNPs, SNPs with low call rates (< 95%), MAF < 5%, and failed Hardy-Weinberg equilibrium p-value < 1e-6 were removed. Variants discordant with their respective subpopulation in the 1000 Genomes Project [] reference panel were removed (Chinese: EAS with a threshold of 0.20; Malays: EAS with a threshold of 0.30; Indian: SAS with a threshold of 0.20). Samples with call rate < 99%, cryptic relatedness and sex/ ethnic discrepancies were excluded. The resulting data were pre-phased using SHAPEIT v2.837 with family trio information. We then used Sanger Imputation Service for imputation, choosing 1000 Genomes Project Phase 3 as reference panel and imputed "with PBWT, no pre-phasing" (the Positional Burrows Wheeler Transform algorithm) as the pipeline. Imputed data that were non-monomorphic, had biallelic SNPs and an INFO score > 0.80 were retained. Imputed genotyping data that were common in all three ethnicities (5,771,259 SNPs) were used for further analyses.

ALSPAC. Children were genotyped using the Illumina HumanHap550-quad chip genotyping platforms by *23andme* subcontracting the Wellcome Trust Sanger Institute, Cambridge, UK and the Laboratory Corporation of America, Burlington, NC, US. DNA was extracted from blood, cell line, and mouthwash samples, then the resulting raw genome-wide data were subjected to standard quality control methods. Participants with inconsistent self-reported and genotyped sex, minimal or excessive heterozygosity, high levels of individual missingness (>3%) and insufficient sample replication (IBD < 0.8) were excluded. SNPs with a minor allele frequency of < 1%, a call rate of < 95% or evidence for violations of Hardy-Weinberg equilibrium (p < 5e-7) were removed. Cryptic relatedness was measured as

proportion of identity by descent (IBD > 0.1). Related subjects that passed all other quality control thresholds were retained during subsequent phasing and imputation. For all the subjects that were retained (N = 9115), a total of 500,527 SNPs passed these quality control filters, and after imputation with Impute v3 and Haplotype Reference Consortium (HRC) imputation reference panel (release 1.1), total genotyping data resulted in 38,898,739 SNPs. Consent for biological samples has been collected in accordance with the Human Tissue Act (2004). Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

See Table S6 for a summary of the genotyping information for each cohort. See Table S7 for a brief description of the tools and datasets used throughout the study.

Identification of corticolimbic DCC gene co-expression networks and ePRS calculation

The dataset for the PFC network was generated from 29 male mice, between 10-12 weeks of age. The large-scale gene expression analysis was performed across recombinant inbred (RI) strains that were derived from the C57BL6/J x DBA/2J (BXD) genetic mapping panel, resulting in the profiling of PFC tissue from 27 BXD strains and the 2 progenitor strains (N=29). The Affymetrix Mouse Genome 430 type 2.0 microarray platform was used, and the results were normalized using the robust multi-array average (RMA) expression measure. Correlations for *Dcc* gene co-expression were calculated using the Pearson correlation coefficient, using the trait ID *1440487_at*. This dataset can be downloaded from the Gene Expression Omnibus repository (GSE28515) and can be queried on GeneNetwork (GN Accession GN135).

The dataset for the nucleus accumbens network was generated from 75 mice, 54 from the BXD panel and 21 mice reported in GeneNetwork as “other” strains. The Illumina Mouse WG-6 v1.1

Expression BeadChip platform was used, and the results were normalized using the rank invariant set normalization. Correlations for gene expression were calculated using the Pearson correlation coefficient, using the trait ID*ILM100460270*. This dataset (GeneNetwork accession GN285) does not contain information to further inspect the details of the experiment.

Behavioral Outcomes

We explored whether the corticolimbic *DCC*-ePRS score associates with two different aspects of impulsivity: *(i)* impulsive choice, reflecting a proneness to make risky choices, as measured by the Information Sampling Task (IST); *(ii)* impulsive action, reflecting a compromised faculty to inhibit motor responses, as measured by the stop-signal reaction time task (SSRT). In both cases, the ability to self-regulate behavior is required for interrupting or inhibiting competing inputs or actions in order to accomplish a specific goal-directed response [] (See Figure S2).

Information Sampling Task (IST). The IST is part of the CANTAB battery of neuropsychological tests and is designed to measure impulsive decision making. For each trial, a 5x5 matrix of gray boxes is presented on a computer screen, with two additional colored boxes centered below, indicating the two possible colors hidden underneath the gray boxes. Children are told that this is a game for points and that by correctly choosing the color appearing more frequently under the gray boxes they can win points. Once a grey box is selected, it immediately opens the box to reveal which of the two colors is underneath. Subjects are told that they can open as many boxes as they want, without time limitation, before making a decision. Once the subject decides and indicates which of the 2 colors appears more frequently, the color under the remaining gray boxes is revealed, along with a message stating whether the subject chose the correct response or not. The main outcome from this task is the mean probability of being correct (meanP-correct) when the response is made. Lower scores indicate more impulsivity because they are obtained when less information is gathered before deciding which color is the most

prevalent (this score is a measure of impulsive choice). This task was performed by children in the MAVAN cohort at 72 months of age.

Stop-Signal Task (SST). The SST is designed to measure inhibitory control of a motor response, a construct of impulsive action. Participants in this task are required to respond as fast as possible to a "go" signal across many trials. On a subset of these trials, a "stop" signal is presented shortly after the "go" signal, and subjects must try to inhibit an already initiated response. We studied 2 variants of the SST:

SST – Standard: By presenting the "stop" signal with varying delays after the "go" signal, an estimate of the time required by each subject to successfully inhibit an ongoing response can be calculated. This outcome is called the Stop-signal Reaction Time or SSRT. Following each incorrect response, the subsequent presentation of the "stop" signal is delayed 50ms, while for each correctly inhibited response, the subsequent presentation of the "stop" signal is shortened by 50ms. This process serves to estimate the latency for the "stop" signal at which the participant responded correctly 50% of the time, which ultimately provides the basis for SSRT estimation. Additionally, we investigated the proportion of successfully inhibited responses when presented with the "stop" signal, which serves as another main outcome in the SST (see []). This task was performed by children in MAVAN and GUSTO cohorts at 6 years of age.

SST – Modified: In the modified version of the task, the "stop" signal is presented with a fixed delay after the "go" signal (see Figure S2). This modification prevents SSRT estimation, but the ability to actively suppress the response when presented with the "stop" signal is still required. Therefore, other components of the task serve as the main outcomes to study. Here, we investigated the mean reaction time of responses during unsuccessful stop trials and the proportion of unsuccessful stops. This task was performed by children in ALSPAC cohort at 10 years of age.

Note: While several measures can be calculated for the SST, we only studied estimated SSRT (for MAVAN and GUSTO), accuracy on stop trials (all 3 cohorts), and mean reaction time in incorrect stop trials (for ALSPAC).

Validation of the PFC and NAcc DCC co-expression networks

To characterize the functional and biological properties of the gene networks that comprise the corticolimbic *DCC*-ePRS score, we used 5 bioinformatic resources: 1) the STRING database ([] https://string-db.org/) to construct the protein-protein interaction networks and analyze functional interactions between gene products; 2) the Cytoscape software to design and visualize the final PPI networks [], https://cytoscape.org/; 3) Functional Mapping and Annotation (FUMA, [], https://fuma.ctglab.nl/) to explore the expression of the genes from the co-expression networks across the 54 tissue types reported in GTEx v8; 4) MetaCore™ (Clarivate Analytics) to perform enrichment analysis by mapping genes in the co-expression networks onto functional ontologies; and 5) the Cell-type Specific Expression Analysis (CSEA, [] http://genetics.wustl.edu/jdlab/csea-tool-2/) to analyze selective enrichment of transcripts in particular brain regions and across different developmental periods (for a complete description on how to use this resource, see [], and see CSEA explanation below in the supplemental methods). These different bioinformatic resources allowed us to explore the biological context in which the genes within the co-expression networks operate.

Finally, we used the human post-mortem brain samples BrainSpan dataset to evaluate expression levels of the genes comprising the PFC and the NAcc co-expression networks in childhood (n=6 for NAcc, n=12 for PFC, ages from 4 months to 11 years) and adulthood (n=7 for NAcc and PFC, ages from 19 to 40 years). This allows us to investigate whether the networks originally identified in mice are also observed in humans and the extent to which the pattern of co-expression in childhood is maintained across the lifespan. To simplify the visualization of the clustered genes and their consistency across development,

we kept the same order for the genes for both time points, for each brain region. This analysis was performed using the R-based package *heatmaply* [].

Cell-type Specific Expression Analysis

Results presented in Figure 3 (panel 3D) show the enrichment in gene expression in the human brain, according to region and age. The "cell-type specific expression analysis" (CSEA) tool was developed by the laboratory of Joseph Dougherty in NY (Howard Hughes Medical Institute, the Rockefeller University; see []) to investigate, among its many applications, the selective expression of genes to particular brain regions and across different developmental periods. This selective expression (termed specificity index in the tool) is calculated by comparing 60 different gene expression *profiles* – one profile for each brain region at each developmental period. For each profile, the identification of selectively enriched genes is determined at different levels of stringency (specificity index probability-pSI from 0.05 to 0.0001). A more stringent analysis results in a smaller list of enriched genes that are highly unique to that profile. Using a Fisher's exact test (with Benjamini-Hochberg correction), this CSEA tool calculates the expected overlap between an input gene list and the previously calculated enriched lists for each of the 60 *profiles*, at 4 varying degrees of stringency (pSI = {0.05, 0.01, 0.001, 0.0001}).

Comparison between the corticolimbic DCC-ePRS and other polygenic scores

We generated other polygenic scores using our accelerated pipeline (https://github.com/MeaneyLab/PRSoS, []), for each subject. To test and compare the predictive power and functional cohesiveness of similarly large networks derived from different approaches, we calculated other polygenic scores. First, we calculated a traditional polygenic score comparable in size to the ePRS in terms of number of SNPs, considering the top 4515 SNPs from the latest ADHD GWAS, which corresponds to the GWAS *p*-value threshold 4.912e-5 []. We investigated whether the PRS for ADHD would associate with impulsivity measurements in the same three cohorts, and then carried out a

functional enrichment analysis to characterize the biological properties of the resulting group of genes that, based on genomic location, were mapped from the SNPs included in the score. We also calculated a polygenic score based on a random subset of 4515 SNPs, where we matched the proportion of SNPs from each brain region to the corticolimbic DCC-ePRS (2040 SNPs in the PFC and 2475 SNPs in the NAcc) and weighted the SNPs by the corresponding brain-region-specific effect from GTEx. We tested the association of this random ePRS with impulsivity measurements.

Statistical Analysis

We generated polygenic scores for all subjects with available genotypic data, with the final genetic scores categorized into low or high PRS/ePRS using a median split. All subsequent analyses, which included the comparison of baseline characteristics and the linear regressions used to examine the association of the genetic scores with the behavioral outcomes, were ran using subjects with complete genotypic and phenotypic data. Based on our inspection for influential observations, we excluded 5 datapoints from the MAVAN dataset. The population structure of the MAVAN, GUSTO, and ALSPAC cohorts were evaluated using principal component analysis of all genotyped SNPs that passed the quality control with low allele frequency (MAF > 5%) and with the following pruning parameters: not in high linkage disequilibrium ($r^2 > 0.2$) across 50 kb regions [] and a sliding window of 5 SNPs for MAVAN and GUSTO cohorts, and not in high linkage disequilibrium across 100 kb region, increment of 5 SNPs and variance inflation factor threshold of 1.01 for ALSPAC cohort. Based on the inspection of the scree plot, the first three principal components in MAVAN and GUSTO, and the first 10 principal components in ALSPAC, were the most informative of population structure and were included in all subsequent analyses.

Supplemental Figures and Tables

Supplementary Table 1

Enrichment Analysis (Metacore™)				
	PFC Network		*NAcc Network*	
	p-value	*FDR*	*p-value*	*FDR*
Cellular Localizations				
Synapse	0.000004	0.00024	1.301e-19	4.461e-17
Cell junction	3.304e-8	0.000005	1.623e-18	2.799e-16
Postsynapse	0.00016	0.00301	3.781e-17	3.542e-15
Synaptic membrane	0.00001	0.00051	4.131e-17	3.542e-15
Asymmetric synapse	0.0039	0.034	1.231e-15	8.455e-14
Neuron to neuron synapse	0.0059	0.041	3.906e-15	2.233e-13
Postsynaptic density	0.0036	0.032	1.448e-14	7.095e-13
Postsynaptic specialization	0.00012	0.0024	5.222e-14	2.239e-12
Plasma membrane region	0.00014	0.0024	1.307e-13	4.983e-12
Cell periphery	0.000087	0.00204	8.018e-13	2.750e-11
Molecular Functions				
Protein binding	2.723e-8	0.0000056	2.210e-7	0.000092
Binding	0.000217	0.0148	0.000016	0.0015
Voltage-gated ion channel activity	0.482	0.634	0.000017	0.0015
Voltage-gated channel activity	0.487	0.643	0.000018	0.0015
Voltage-gated cation channel activity	0.675	0.781	0.000021	0.0015
Protein domain specific binding	0.224	0.403	0.000027	0.0016
Cation channel activity	0.76	0.83	0.000076	0.0038
Gated channel activity	0.15	0.32	0.000083	0.0038
Ion channel activity	0.30	0.48	0.000096	0.0040
Glycosaminoglycan binding	0.023	0.15	0.00012	0.0043
Biological Processes				
Modulation of chemical synaptic transmission	0.0025	0.016	2.199e-18	3.724e-15
Regulation of trans-synaptic signaling	0.0026	0.016	2.349e-18	3.724e-15
Neuron differentiation	0.00021	0.0028	1.171e-17	1.138e-14
Neuron development	0.00093	0.0078	1.574e-17	1.138e-14
Generation of neurons	0.0000058	0.00025	1.795e-17	1.138e-14
Trans-synaptic signaling	0.000029	0.00075	3.090e-17	1.633e-14
Chemical synaptic transmission	0.000064	0.0012	1.464e-16	5.803e-14
Anterograde trans-synaptic signaling	0.000064	0.0012	1.464e-16	5.802e-14
Synaptic signaling	0.000091	0.0016	2.627e-16	9.256e-14
Neurogenesis	0.000015	0.00048	3.390e-16	1.075e-13

Table S1. Gene ontology categories related to genes included in the *DCC* gene co-expression networks in the PFC and the NAcc.

A

MAVAN-IST	Impulsive Choice	
	MeanP-correct	
	β	*p*
DCC-ePRS	-0.04	0.045

B

MAVAN-SST	Impulsive Action - Standard			
	Successful stops		*Estimated SSRT*	
	β	*p*	*β*	*p*
DCC-ePRS	-0.03	0.10	-10.07	0.57

C

GUSTO-SST	Impulsive Action - Standard			
	Successful stops		*Estimated SSRT*	
	β	*p*	*β*	*p*
DCC-ePRS	-0.03	0.027	-19.15	0.07

D

ALSPAC-SST	Impulsive Action - Modified			
	Successful stops		*MRT - Stop trials*	
	β	*p*	*β*	*p*
DCC-ePRS	-0.005	0.299	-10.368	0.019

Table S2. Estimated effects of the ePRS on different impulsivity measures across cohorts. All models were adjusted for population stratification and sex.

Correlations between the continuous ePRS scores and the behavioral measures

	MAVAN			GUSTO		ASLPAC	
	Impulsive Choice	Impulsive Action - Standard		Impulsive Action - Standard		Impulsive Action - Modified	
	MeanP-correct	*Proportion of successful stops*	*Estimated SSRT*	*Proportion of successful stops*	*Estimated SSRT*	*Proportion of successful stops*	*MRT in Stop-Signal trials*
ePRS-DCC	0.249 **	-0.114	-0.055	0.073	0.030	0.021	0.008
PRS ADHD	0.037	0.053	0.142 *	0.049	0.052	-0.016	0.014
Random ePRS	-0.019	0.017	0.008	0.027	-0.014	-0.028	-0.007

* $p < .05$, ** $p < .001$

Table S3. Correlations between the continuous ePRS scores and the behavioral measures investigated in each cohort.

Supplementary Table 4. Co-expression network in ***PFC***

Gene name	Ensembl ID	Description
BTG3	ENSG00000281484	BTG anti-proliferation factor 3
HSPA1L	ENSG00000236251	heat shock protein family A (Hsp70) member 1 like
CXADR	ENSG00000154639	CXADR Ig-like cell adhesion molecule
SURF2	ENSG00000281024	surfeit 2
GRIK1	ENSG00000171189	glutamate ionotropic receptor kainate type subunit 1
SPRY2	ENSG00000136158	sprouty RTK signaling antagonist 2
UFM1	ENSG00000120686	ubiquitin fold modifier 1
GNG4	ENSG00000282972	G protein subunit gamma 4
SNRPB2	ENSG00000125870	small nuclear ribonucleoprotein polypeptide B2
METTL9	ENSG00000284548	methyltransferase like 9
RNF152	ENSG00000176641	ring finger protein 152
COL6A2	ENSG00000142173	collagen type VI alpha 2 chain
EEF1B2	ENSG00000283391	eukaryotic translation elongation factor 1 beta 2
TMEM47	ENSG00000147027	transmembrane protein 47
GLRA2	ENSG00000101958	glycine receptor alpha 2
SLC2A10	ENSG00000197496	solute carrier family 2 member 10
TIMM17B	ENSG00000126768	translocase of inner mitochondrial membrane 17B
BACE2	ENSG00000182240	beta-secretase 2
NXPH1	ENSG00000122584	neurexophilin 1
NRP1	ENSG00000099250	neuropilin 1
RSU1	ENSG00000148484	Ras suppressor protein 1
EFNB1	ENSG00000090776	ephrin B1
RPS4X	ENSG00000198034	ribosomal protein S4 X-linked
FJX1	ENSG00000179431	four-jointed box kinase 1
COMMD6	ENSG00000188243	COMM domain containing 6
ABHD3	ENSG00000158201	abhydrolase domain containing 3, phospholipase
GABRA5	ENSG00000186297	gamma-aminobutyric acid type A receptor subunit alpha5
PKIG	ENSG00000168734	cAMP-dependent protein kinase inhibitor gamma
FTH1	ENSG00000167996	ferritin heavy chain 1
IGSF1	ENSG00000147255	immunoglobulin superfamily member 1
HS6ST2	ENSG00000171004	heparan sulfate 6-O-sulfotransferase 2
LRRN1	ENSG00000175928	leucine rich repeat neuronal 1
GPC3	ENSG00000147257	glypican 3
TRIP13	ENSG00000071539	thyroid hormone receptor interactor 13
SEMA5A	ENSG00000112902	semaphorin 5A
MAPRE1	ENSG00000101367	microtubule associated protein RP/EB family member 1
ITPA	ENSG00000125877	inosine triphosphatase

TAC1	ENSG00000006128	tachykinin precursor 1
CYP26A1	ENSG00000095596	cytochrome P450 family 26 subfamily A member 1
LAPTM5	ENSG00000162511	lysosomal protein transmembrane 5
RPL23	ENSG00000125691	ribosomal protein L23
SDC3	ENSG00000162512	syndecan 3
CHST8	ENSG00000124302	carbohydrate sulfotransferase 8
KRT17	ENSG00000128422	keratin 17
MYCN	ENSG00000134323	MYCN proto-oncogene, bHLH transcription factor
EVL	ENSG00000196405	Enah/Vasp-like
ROMO1	ENSG00000125995	reactive oxygen species modulator 1
ETV4	ENSG00000175832	ETS variant transcription factor 4
ITGA5	ENSG00000161638	integrin subunit alpha 5
PDZRN3	ENSG00000121440	PDZ domain containing ring finger 3
TBC1D10A	ENSG00000099992	TBC1 domain family member 10A
BUD31	ENSG00000106245	BUD31 homolog
EHBP1L1	ENSG00000173442	EH domain binding protein 1 like 1
NFKBIE	ENSG00000146232	NFKB inhibitor epsilon
MAF	ENSG00000178573	MAF bZIP transcription factor
RHPN2	ENSG00000131941	rhophilin Rho GTPase binding protein 2
SLN	ENSG00000170290	sarcolipin
DAB2	ENSG00000153071	DAB adaptor protein 2
FHL3	ENSG00000183386	four and a half LIM domains 3
CHST11	ENSG00000171310	carbohydrate sulfotransferase 11
DCK	ENSG00000156136	deoxycytidine kinase
RPL30	ENSG00000156482	ribosomal protein L30
RPL12	ENSG00000197958	ribosomal protein L12
LOXL1	ENSG00000129038	lysyl oxidase like 1
NDUFB11	ENSG00000147123	NADH:ubiquinone oxidoreductase subunit B11
WFIKKN2	ENSG00000173714	WAP, follistatin/kazal, immunoglobulin, kunitz and netrin domain containing 2
SOX12	ENSG00000177732	SRY-box transcription factor 12
TSPAN6	ENSG00000000003	tetraspanin 6
BARX2	ENSG00000043039	BARX homeobox 2
COTL1	ENSG00000103187	coactosin like F-actin binding protein 1
MARCKS	ENSG00000277443	myristoylated alanine rich protein kinase C substrate
PRICKLE2	ENSG00000163637	prickle planar cell polarity protein 2
HMGCS1	ENSG00000112972	3-hydroxy-3-methylglutaryl-CoA synthase 1
UST	ENSG00000111962	uronyl 2-sulfotransferase
RGS20	ENSG00000147509	regulator of G protein signaling 20
RNF180	ENSG00000164197	ring finger protein 180
RGS9	ENSG00000108370	regulator of G protein signaling 9

SOCS2	ENSG00000120833	suppressor of cytokine signaling 2
CKAP4	ENSG00000136026	cytoskeleton associated protein 4
RPS3	ENSG00000149273	ribosomal protein S3
GABRG1	ENSG00000163285	gamma-aminobutyric acid type A receptor subunit gamma1
PPP1R14B	ENSG00000173457	protein phosphatase 1 regulatory inhibitor subunit 14B
CHRDL1	ENSG00000101938	chordin like 1
CHCHD3	ENSG00000106554	coiled-coil-helix-coiled-coil-helix domain containing 3
ZNHIT1	ENSG00000106400	zinc finger HIT-type containing 1
TNNI3	ENSG00000129991	troponin I3, cardiac type
CASK	ENSG00000147044	calcium/calmodulin dependent serine protein kinase
LRRC17	ENSG00000128606	leucine rich repeat containing 17
VIPR2	ENSG00000106018	vasoactive intestinal peptide receptor 2
SOX2	ENSG00000181449	SRY-box transcription factor 2
PRDX4	ENSG00000123131	peroxiredoxin 4
CRH	ENSG00000147571	corticotropin releasing hormone
SNX7	ENSG00000162627	sorting nexin 7
DACT1	ENSG00000165617	dishevelled binding antagonist of beta catenin 1
MSI1	ENSG00000135097	musashi RNA binding protein 1
VEGFC	ENSG00000150630	vascular endothelial growth factor C
DCAF12	ENSG00000198876	DDB1 and CUL4 associated factor 12
RALA	ENSG00000006451	RAS like proto-oncogene A
HNMT	ENSG00000150540	histamine N-methyltransferase
TTL	ENSG00000114999	tubulin tyrosine ligase
IGFBP5	ENSG00000115461	insulin like growth factor binding protein 5
CD248	ENSG00000174807	CD248 molecule
CYC1	ENSG00000179091	cytochrome c1
EZR	ENSG00000092820	ezrin
DLX1	ENSG00000144355	distal-less homeobox 1
SDHAF1	ENSG00000205138	succinate dehydrogenase complex assembly factor 1
ENDOG	ENSG00000167136	endonuclease G
ELN	ENSG00000049540	elastin
POU3F3	ENSG00000198914	POU class 3 homeobox 3
MPZL1	ENSG00000197965	myelin protein zero like 1
PCOLCE	ENSG00000106333	procollagen C-endopeptidase enhancer
DNAJB1	ENSG00000132002	DnaJ heat shock protein family (Hsp40) member B1
RNF2	ENSG00000121481	ring finger protein 2
POLR2G	ENSG00000168002	RNA polymerase II subunit G
NPNT	ENSG00000168743	nephronectin
RPL27	ENSG00000131469	ribosomal protein L27
DTYMK	ENSG00000168393	deoxythymidylate kinase

KCNIP1	ENSG00000182132	potassium voltage-gated channel interacting protein 1
WWTR1	ENSG00000018408	WW domain containing transcription regulator 1
TGFB2	ENSG00000092969	transforming growth factor beta 2
TGFB1	ENSG00000105329	transforming growth factor beta 1
DPP3	ENSG00000254986	dipeptidyl peptidase 3
WSB2	ENSG00000176871	WD repeat and SOCS box containing 2
G0S2	ENSG00000123689	G0/G1 switch 2
UNC13C	ENSG00000137766	unc-13 homolog C
GCK	ENSG00000106633	glucokinase
CD244	ENSG00000122223	CD244 molecule
PDE6D	ENSG00000156973	phosphodiesterase 6D
TLCD1	ENSG00000160606	TLC domain containing 1
MASP1	ENSG00000127241	mannan binding lectin serine peptidase 1
SSR2	ENSG00000163479	signal sequence receptor subunit 2
NT5E	ENSG00000135318	5'-nucleotidase ecto
RND1	ENSG00000172602	Rho family GTPase 1
FADS2	ENSG00000134824	fatty acid desaturase 2
FMNL1	ENSG00000184922	formin like 1
RBM4	ENSG00000173933	RNA binding motif protein 4
ANAPC11	ENSG00000141552	anaphase promoting complex subunit 11
MGLL	ENSG00000074416	monoglyceride lipase
ODC1	ENSG00000115758	ornithine decarboxylase 1
HEYL	ENSG00000163909	hes related family bHLH transcription factor with YRPW motif like
LRRTM1	ENSG00000162951	leucine rich repeat transmembrane neuronal 1
SERBP1	ENSG00000142864	SERPINE1 mRNA binding protein 1
CCDC28B	ENSG00000160050	coiled-coil domain containing 28B
TRIB2	ENSG00000071575	tribbles pseudokinase 2
LIMS2	ENSG00000072163	LIM zinc finger domain containing 2
H2AW	ENSG00000284841	H2A.W histone
TENM1	ENSG00000009694	teneurin transmembrane protein 1
TAFA1	ENSG00000183662	TAFA chemokine like family member 1
NECTIN3	ENSG00000177707	nectin cell adhesion molecule 3
SELENOW	ENSG00000178980	selenoprotein W
MACROH2A2	ENSG00000099284	macroH2A.2 histone
NT5C3A	ENSG00000122643	5'-nucleotidase, cytosolic IIIA

Table S4. Genes co-expressed with *DCC* in the PFC

Supplementary Table 5. Co-expression network in ***NAcc***

Gene name	Ensembl ID	Description
NEU4	ENSG00000277926	neuraminidase 4
PPP4R4	ENSG00000278326	protein phosphatase 4 regulatory subunit 4
MAGI1	ENSG00000282956	membrane associated guanylate kinase, WW and PDZ domain containing 1
PRKAR2B	ENSG00000284096	protein kinase cAMP-dependent type II regulatory subunit beta
FLRT3	ENSG00000125848	fibronectin leucine rich transmembrane protein 3
DCC	ENSG00000187323	DCC netrin 1 receptor
RAB27B	ENSG00000041353	RAB27B, member RAS oncogene family
SCRT1	ENSG00000284923	scratch family transcriptional repressor 1
PDGFB	ENSG00000100311	platelet derived growth factor subunit B
EFNB2	ENSG00000125266	ephrin B2
MAP1LC3A	ENSG00000101460	microtubule associated protein 1 light chain 3 alpha
KCNA6	ENSG00000151079	potassium voltage-gated channel subfamily A member 6
ARHGDIG	ENSG00000242173	Rho GDP dissociation inhibitor gamma
PGRMC1	ENSG00000101856	progesterone receptor membrane component 1
SOX4	ENSG00000124766	SRY-box transcription factor 4
KLHL13	ENSG00000003096	kelch like family member 13
GRM7	ENSG00000196277	glutamate metabotropic receptor 7
NETO2	ENSG00000171208	neuropilin and tolloid like 2
ISL1	ENSG00000016082	ISL LIM homeobox 1
ROBO2	ENSG00000185008	roundabout guidance receptor 2
KCNK3	ENSG00000171303	potassium two pore domain channel subfamily K member 3
ADAMTSL2	ENSG00000197859	ADAMTS like 2
LASP1	ENSG00000002834	LIM and SH3 protein 1
CNTN5	ENSG00000149972	contactin 5
COL5A1	ENSG00000130635	collagen type V alpha 1 chain
DPYSL4	ENSG00000151640	dihydropyrimidinase like 4
MAP3K10	ENSG00000130758	mitogen-activated protein kinase kinase kinase 10
GPR158	ENSG00000151025	G protein-coupled receptor 158
KCND2	ENSG00000184408	potassium voltage-gated channel subfamily D member 2
ADAMTS3	ENSG00000156140	ADAM metallopeptidase with thrombospondin type 1 motif 3
KCNS2	ENSG00000156486	potassium voltage-gated channel modifier subfamily S member 2
SYP	ENSG00000102003	synaptophysin
GRID2	ENSG00000152208	glutamate ionotropic receptor delta type subunit 2
BCL11A	ENSG00000119866	BAF chromatin remodeling complex subunit BCL11A
HOMER1	ENSG00000152413	homer scaffold protein 1
PODXL2	ENSG00000114631	podocalyxin like 2
BASP1	ENSG00000176788	brain abundant membrane attached signal protein 1

GRIN1	ENSG00000176884	glutamate ionotropic receptor NMDA type subunit 1
IGSF21	ENSG00000117154	immunoglobin superfamily member 21
SKIL	ENSG00000136603	SKI like proto-oncogene
SLC6A17	ENSG00000197106	solute carrier family 6 member 17
NRXN3	ENSG00000021645	neurexin 3
DRD1	ENSG00000184845	dopamine receptor D1
DHDH	ENSG00000104808	dihydrodiol dehydrogenase
TMEM163	ENSG00000152128	transmembrane protein 163
LRFN5	ENSG00000165379	leucine rich repeat and fibronectin type III domain containing 5
MEF2C	ENSG00000081189	myocyte enhancer factor 2C
SLIT3	ENSG00000184347	slit guidance ligand 3
DGKB	ENSG00000136267	diacylglycerol kinase beta
NELL2	ENSG00000184613	neural EGFL like 2
MPP7	ENSG00000150054	membrane palmitoylated protein 7
RPRM	ENSG00000177519	reprimo, TP53 dependent G2 arrest mediator homolog
CMIP	ENSG00000153815	c-Maf inducing protein
GPRIN1	ENSG00000169258	G protein regulated inducer of neurite outgrowth 1
WIZ	ENSG00000011451	WIZ zinc finger
PIP5K1C	ENSG00000186111	phosphatidylinositol-4-phosphate 5-kinase type 1 gamma
LRRC40	ENSG00000066557	leucine rich repeat containing 40
TMEM108	ENSG00000144868	transmembrane protein 108
MYT1L	ENSG00000186487	myelin transcription factor 1 like
SCN8A	ENSG00000196876	sodium voltage-gated channel alpha subunit 8
OLFM2	ENSG00000105088	olfactomedin 2
LIN7A	ENSG00000111052	lin-7 homolog A, crumbs cell polarity complex component
PCP4L1	ENSG00000248485	Purkinje cell protein 4 like 1
SERPINE2	ENSG00000135919	serpin family E member 2
TUBG1	ENSG00000131462	tubulin gamma 1
SEMA6C	ENSG00000143434	semaphorin 6C
ARHGAP20	ENSG00000137727	Rho GTPase activating protein 20
PRMT6	ENSG00000198890	protein arginine methyltransferase 6
DAB1	ENSG00000173406	DAB adaptor protein 1
RALGPS2	ENSG00000116191	Ral GEF with PH domain and SH3 binding motif 2
PLPPR5	ENSG00000117598	phospholipid phosphatase related 5
TENM3	ENSG00000218336	teneurin transmembrane protein 3
PLAAT1	ENSG00000127252	phospholipase A and acyltransferase 1
ATP5F1D	ENSG00000099624	ATP synthase F1 subunit delta

Table S5. Genes co-expressed with *DCC* in the NAcc

	MAVAN	GUSTO	ALSPAC
# of available SNPs	20,790,893	5,771,259	38,898,739
Genotyping platform	PsychArray/PsychChip, Illumina	Illumina OmniExpressExome array	Illumina HumanHap550
Reference panel	Haplotype Reference Consortium*	1000 Genomes Project (Phase 3)	Haplotype Reference Consortium*
Imputation method	Sanger Imputation Service	Sanger Imputation Service	Impute v3
# of SNPs included in ePRS	4515	2590	11473

*HRC release 1.1

Table S6. Genotyping information for the 3 cohorts

Data set/Tool name	*Description*	*Reference*
GeneNetwork	Group of linked datasets and analytic tools that allow the user to relate genetic maps and phenotypes of interest. It contains several multispecies tissue-specific datasets.	https://www.genenetwork.org
Brainspan	Online atlas of the developing human brain designed to help investigate the transcriptional mechanisms involved in human brain development. It contains several data modalities across 13 stages of development (from prenatal to adulthood, N=42), including RNA-Seq and exon array expression of 16 cortical and subcortical structures.	https://www.brainspan.org
NCBI Variation Viewer	Online tool from the National Center for Biotechnology Information designed to view, search, and navigate annotated variants from dbSNP, dbVar, and ClinVar in their genomic context, using an interactive sequence viewer. Variants can be searched based on chromosomal location, gene, or variant ID.	https://www.ncbi.nlm.nih.gov/variation/view
The Genotype-Tissue Expression (GTEx) project	Dataset that provides open access to human data including gene expression, QTLs, and histology images, offering the scientific community a resource to study human gene expression and regulation and its relationship to genetic variation (i.e., a GWAS where the trait of interest is gene expression). Samples from this dataset were collected from nearly 1000 individuals (across 54 non-diseased tissues), generating a group of densely genotyped individuals and RNA expression within individual tissues.	https://gtexportal.org/home
STRING	Database containing known and predicted protein-protein interactions, covering 24,584,628 proteins from 5,090 organisms. Protein-protein interactions include both physical interactions and functional associations.	https://string-db.org
Cytoscape	Open source software platform for visualizing networks, including molecular interactions and biological pathways.	https://cytoscape.org
Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA GWAS)	Online platform used to annotate, prioritize, visualize, and interpret GWAS results. Using the *GENE2FUNC* function the user can annotate any list of genes in a biological context.	https://fuma.ctglab.nl/
MetaCore™ (Clarivate Analytics)	Software platform for data analysis, including extensive enrichment analysis of personal datasets. Users can map gene IDs of interest onto gene IDs from a large library of ontologies (e.g., GO processes, GO Molecular Functions, and GO localizations)	https://portal.genego.com
Cell-type Specific Expression Analysis (CSEA)	Online tool that leverages profiling data from humans and mice to perform a cell-type specific expression analysis. This analysis relies on a calculated specificity index, which allows the user to input a list of genes and investigate the extent to which those genes are enriched in particular human brain regions, and/or developmental windows.	http://genetics.wustl.edu/jdlab/csea-tool-2
GeneMANIA	Software platform that allows the user to predict the function of a set of genes using a set of functional association data including genetic interactions, co-expression, co-localization. In addition, once the software has mapped the connections between the genes, it will also return an image of the network that can be imported into the Cytoscape software for further processing and analysis.	https://genemania.org/
ADHD GWAS	Latest association study of genetic variation and ADHD across common human variants.	https://www.med.unc.edu/pgc/download-results

Table S7. Description of the datasets and tools used throughout the study

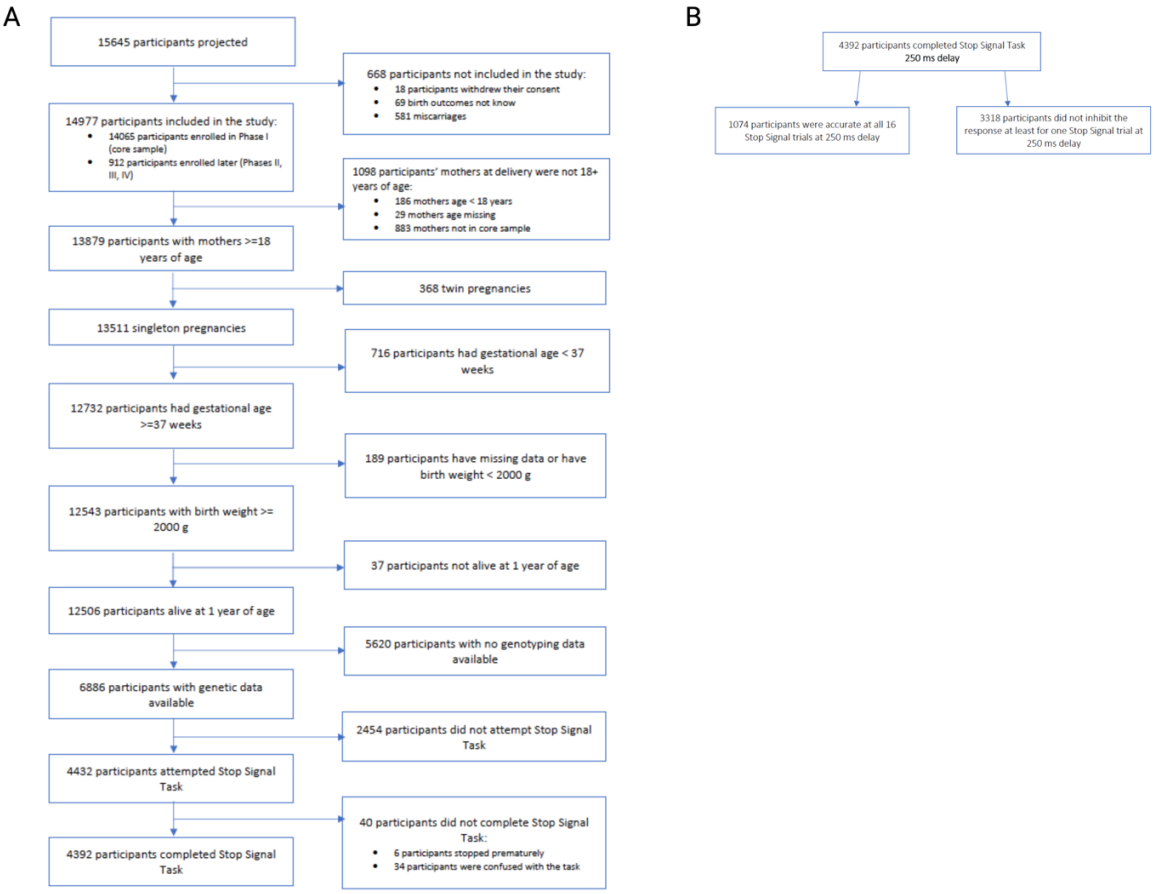

Figure S1. Block scheme depicting the steps involved in sample selection for ALSPAC cohort. Panel A shows a detailed scheme for the exclusion/inclusion criteria, together with the total number of participants, starting at the projected initial cohort (N=15,645) to the selection of participants that completed the SSRT task (n=4,392). Panel B shows the total number of participants for which we have proportion of successful stops and mean reaction time.

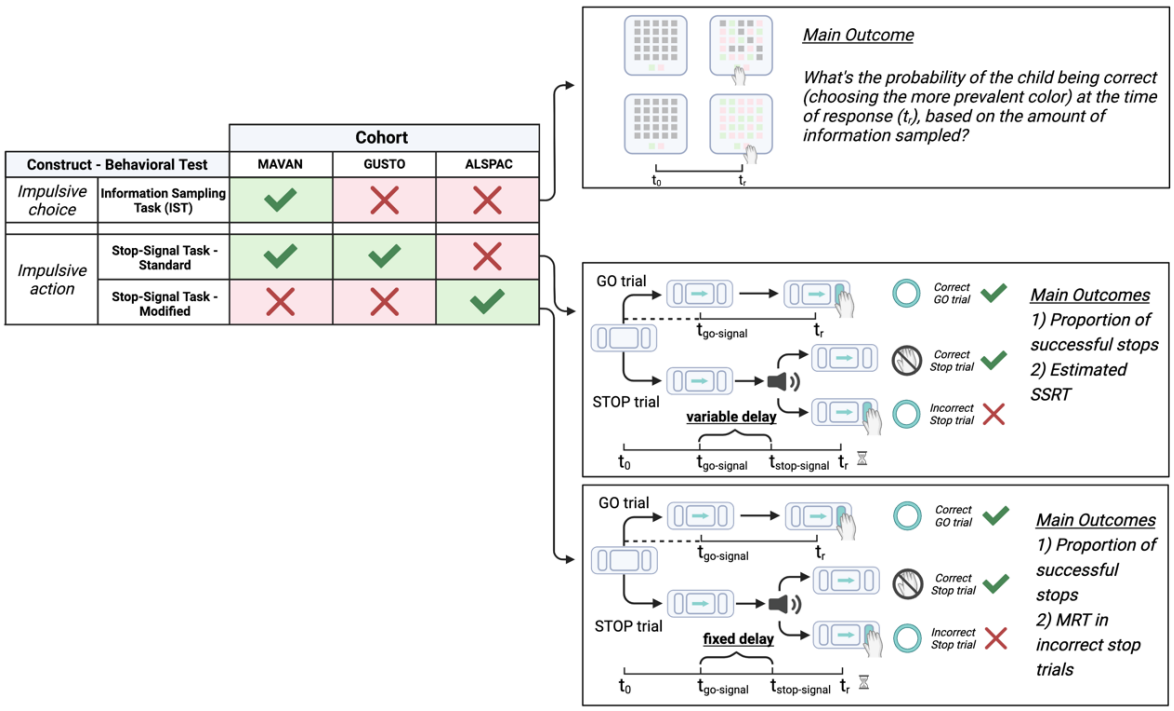

Figure S2. Depiction of the behavioral tasks conducted for each cohort and the main outcomes studied for each task.

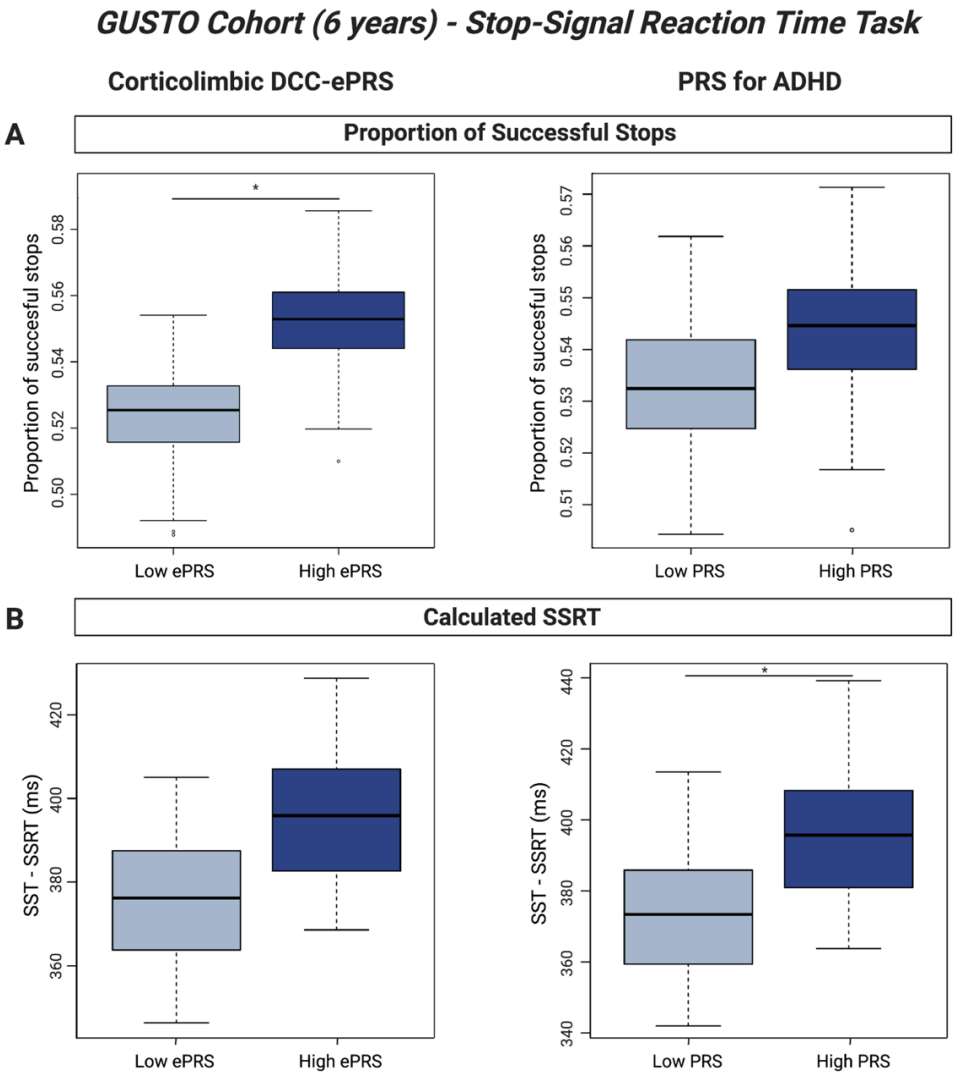

Figure S3. Associations between the computed genetic score for the GUSTO cohort and different components of the SSRT task, represented as boxplots. (A) Low-ePRS score group has a significantly lower proportion of successful stops when presented with the "go" signal, compared to the high-ePRS group (β = -0.03, p = .027). We didn't find a significant association between the PRS for ADHD and the proportion of successful stops (β = -0.009, p = .522). (B) There is no significant association between the corticolimbic DCC-ePRS (β = -19.152, p = .07) and SSRT measure, but we found that subjects with a low-PRS for ADHD had lower SSRT estimate compared to high-PRS subjects (β = -23.865, p = .026). *p<.05**.**

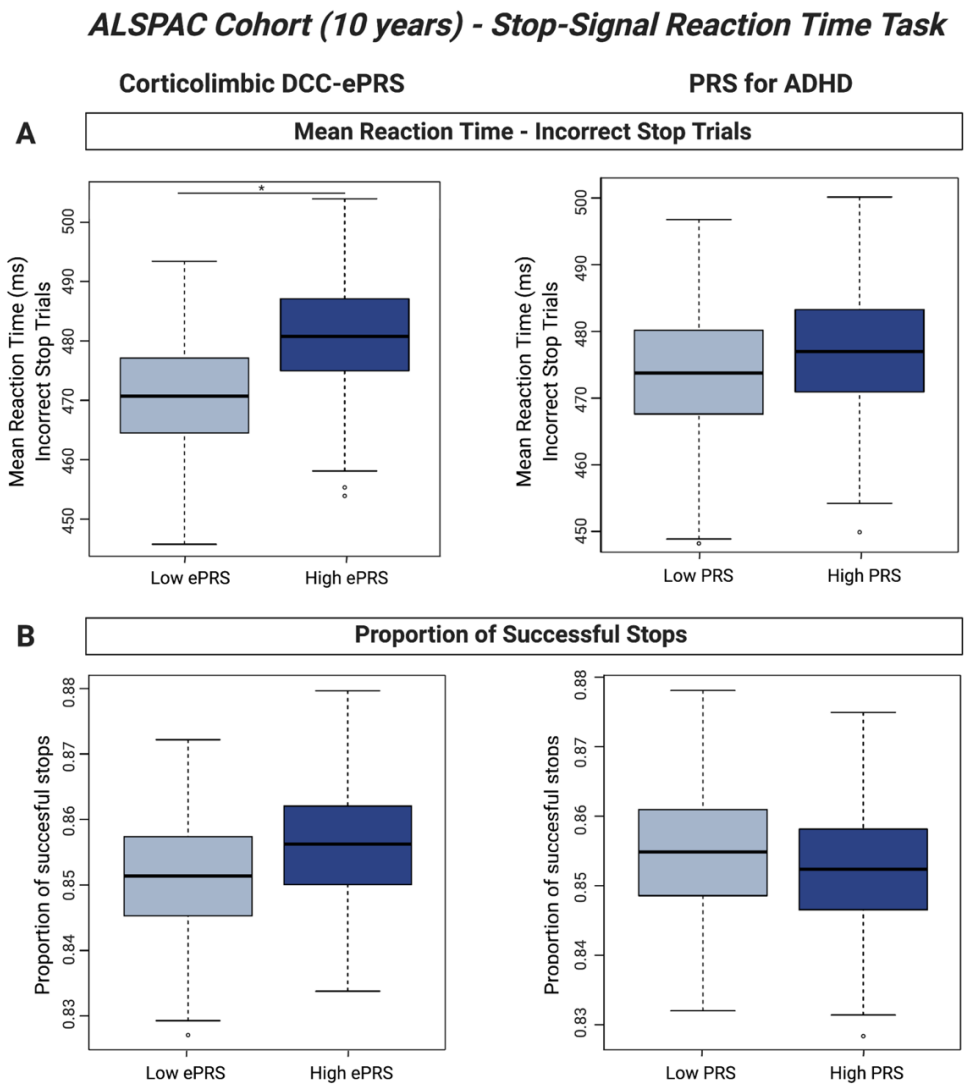

Figure S4. Associations between the computed genetic scores for the ALSPAC cohort and different measures of the SSRT task. (A) Low-ePRS score group has significantly shorter mean latency of response when presented with the "stop" signal, compared to the high-ePRS group (β = -10.368, p = .019), thus showing higher levels of impulsive action. We didn't find a significant association between the PRS for ADHD and the same outcome (β = -3.248, p = .463). (B) There is no significant association between the genetic scores and the proportion of successful stops (ePRS: β = -0.005, p = .299; PRS-ADHD: β = 0.002, p = .657). *p<.05

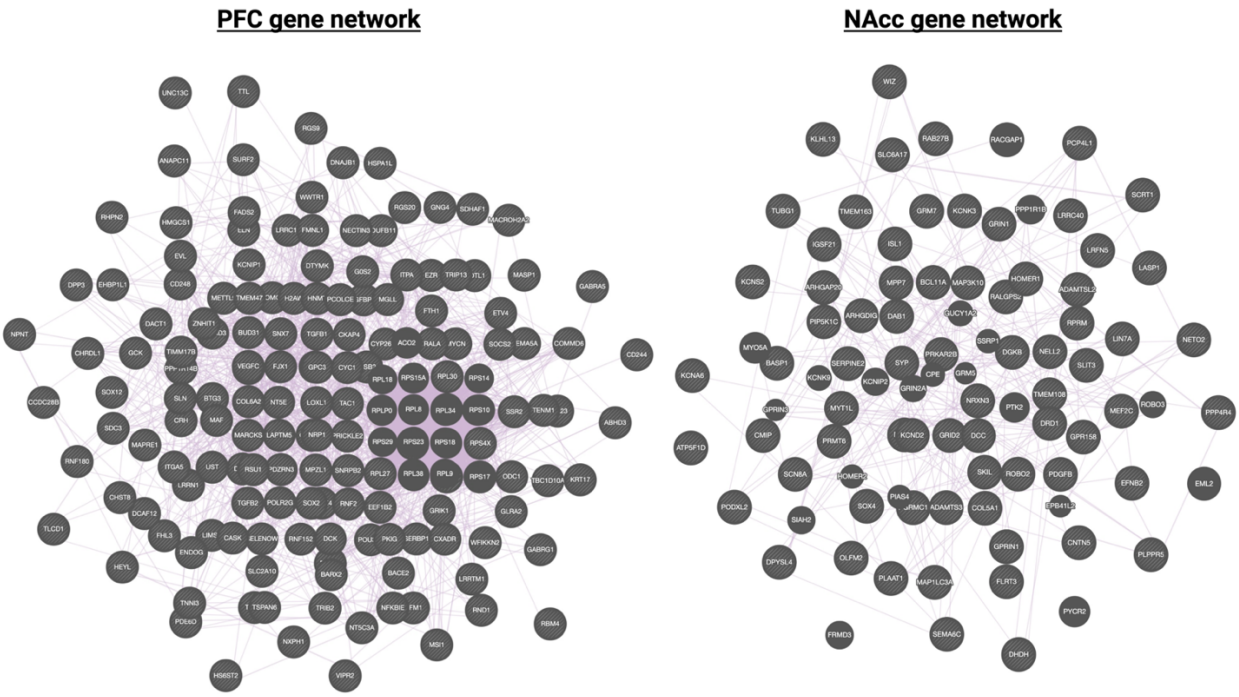

Figure S5. Gene networks for the PFC and the NAcc, created with GeneMANIA. While each gene is represented by a node, the figures clearly demonstrate the cohesiveness of both gene networks. The lines represent co-expression between genes, based on available databases on GeneMANIA.

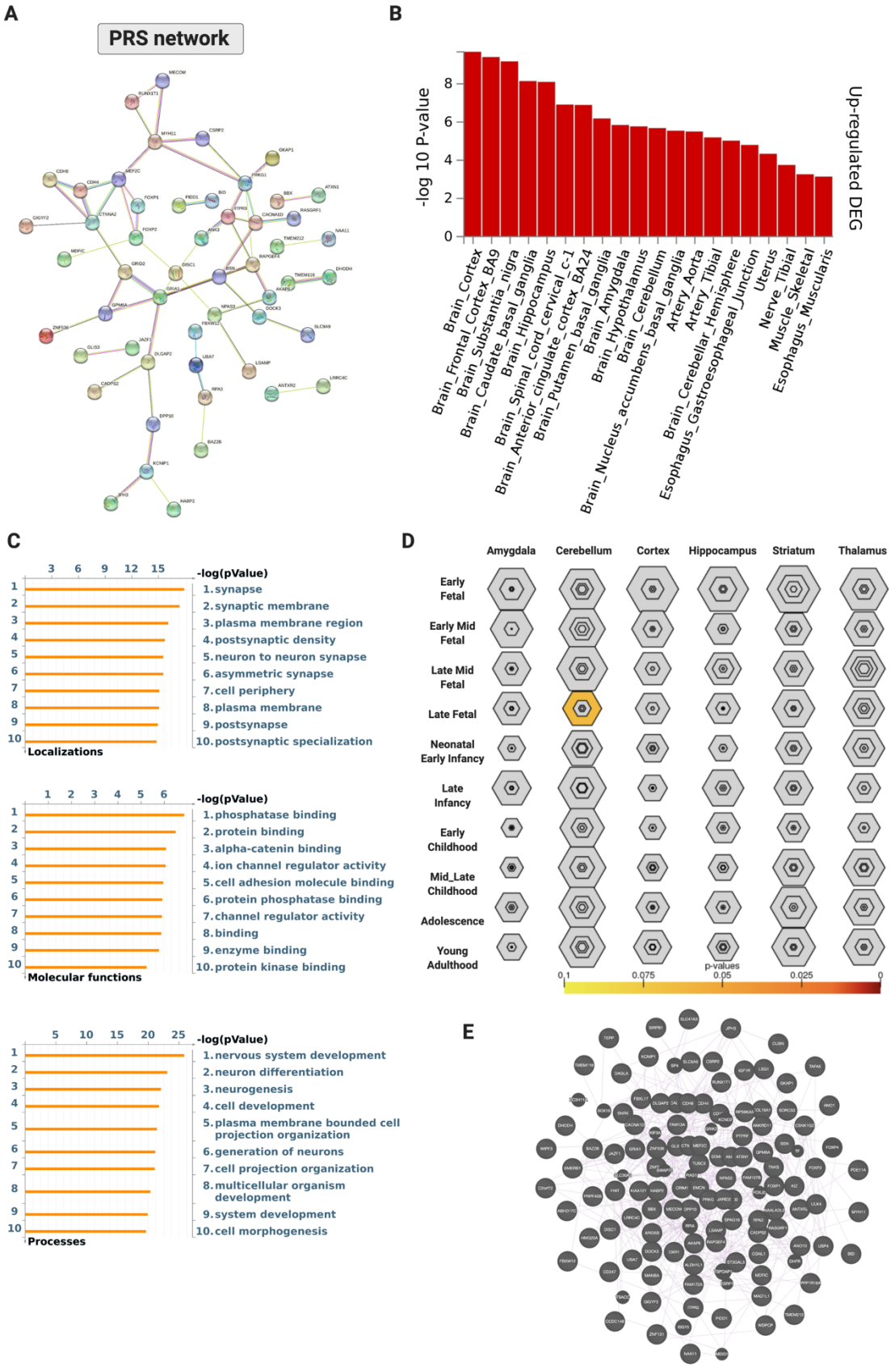
A
PRS network
B
-log 10 P-value
Up-regulated DEG
Brain_Cortex
Brain_Frontal_Cortex_BA9
Brain_Substantia_nigra
Brain_Caudate_basal_ganglia
Brain_Hippocampus
Brain_Spinal_cord_cervical_c-1
Brain_Anterior_cingulate_cortex_BA24
Brain_Putamen_basal_ganglia
Brain_Amygdala
Brain_Hypothalamus
Brain_Cerebellum
Brain_Nucleus_accumbens_basal_ganglia
Artery_Aorta
Artery_Tibial
Brain_Cerebellar_Hemisphere
Esophagus_Gastroesophageal_Junction
Uterus
Nerve_Tibial
Muscle_Skeletal
Esophagus_Muscularis
C
-log(pValue)
1. synapse
2. synaptic membrane
3. plasma membrane region
4. postsynaptic density
5. neuron to neuron synapse
6. asymmetric synapse
7. cell periphery
8. plasma membrane
9. postsynapse
10. postsynaptic specialization
Localizations
-log(pValue)
1. phosphatase binding
2. protein binding
3. alpha-catenin binding
4. ion channel regulator activity
5. cell adhesion molecule binding
6. protein phosphatase binding
7. channel regulator activity
8. binding
9. enzyme binding
10. protein kinase binding
Molecular functions
-log(pValue)
1. nervous system development
2. neuron differentiation
3. neurogenesis
4. cell development
5. plasma membrane bounded cell projection organization
6. generation of neurons
7. cell projection organization
8. multicellular organism development
9. system development
10. cell morphogenesis
Processes
D
Amygdala
Cerebellum
Cortex
Hippocampus
Striatum
Thalamus
Early Fetal
Early Mid Fetal
Late Mid Fetal
Late Fetal
Neonatal Early Infancy
Late Infancy
Early Childhood
Mid_Late Childhood
Adolescence
Young Adulthood
p-values
E

Figure S6. Enrichment analysis of the list of genes that make up the conventional PRS score comparable in size to the ePRS in terms of number of SNPs (n=4515, corresponding to the GWAS *p*-value threshold of 4.912e-5). (A) Protein-Protein interaction networks constructed from the PRS network. The protein network represents known functional interactions between the protein products of the genes (significant PPI enrichment). Tissue-specific gene expression analysis (B) shows that the genes that comprise the PRS are upregulated across the brain, albeit not as selectively and to a much lesser extent than the genes from the corticolimbic *DCC*-ePRS. (C) An enrichment analysis for the cellular localization, molecular functions, and biological processes of the genes in the PRS. Interestingly, these genes are enriched for similar categories across the 3 ontologies, compared to the genes in the corticolimbic *DCC* networks, indicating an enrichment for neurodevelopmental processes. (D) CSEA shows no selective spatiotemporal enrichment for the genes that comprise the PRS. (E) A depiction of the PRS gene network, created with GeneMANIA, where it's possible to observe a highly cohesive co-expression network.

7. References

1. Somerville LH, Casey BJ. Developmental neurobiology of cognitive control and motivational systems. Curr Opin Neurobiol. 2010;20:236–41.

2. Jentsch JD, Ashenhurst JR, Cervantes MC, Groman SM, James AS, Pennington ZT. Dissecting impulsivity and its relationships to drug addictions. Ann N. Y Acad Sci. 2014;1327:1–26.

3. Luna B, Sweeney JA. The emergence of collaborative brain function: fMRI studies of the development of response inhibition. Ann N. Y Acad Sci. 2004;1021:296–309.

4. Petrovic P, Castellanos FX. Top-down dysregulation—from ADHD to emotional instability. Front Behav Neurosci. 2016;10:1–25.

5. Casey BJ, Getz S, Galvan A. The adolescent brain. Dev Rev. 2008;28:62–77.

6. Dalley JW, Everitt BJ, Robbins TW. Impulsivity, Compulsivity, and Top-Down Cognitive Control. Vol. 69, Neuron. Cell Press; 2011. p. 680–94.

7. Galvan A, Hare T, Voss H, Glover G, Casey BJ. Risk-taking and the adolescent brain: Who is at risk? Vol. 10, Developmental Science. John Wiley & Sons, Ltd; 2007. p. 8–14.

8. Luna B, Padmanabhan A, O'Hearn K. What has fMRI told us about the Development of Cognitive Control through Adolescence? Brain Cogn. 2010;72:101–13.

9. Spear LP. The adolescent brain and age-related behavioral manifestations. Neurosci Biobehav Rev. 2000;24:417–63.

10. Caballero A, Granberg R, Tseng KY. Mechanisms contributing to prefrontal cortex maturation during adolescence. Neurosci Biobehav Rev. 2016;70:4–12.

11. Reynolds LM, Flores C. Adolescent dopamine development. In: Diagnosis, Management and Modeling of Neurodevelopmental Disorders. Elsevier; 2021. p. 295–304.

12. Hoops D, Flores C. Making dopamine connections in adolescence. Trends Neurosci. 2017;40:709–19.

13. Hoops D, Reynolds LM, Restrepo-Lozano JM, Flores C. Dopamine development I the mouse orbital prefrontal cortex is protracted and sensitive to amphetamine in adolescence. eNeuro. 2018;5:ENEURO.0372-17.2017.

14. Kalsbeek A, Voorn P, Buijs RM, Pool CW, Uylings HBM. Development of the dopaminergic innervation in the prefrontal cortex of the rat. J Comp Neurol. 1988;269:58–72.

15. Manitt C, Mimee A, Eng C, Pokinko M, Stroh T, Cooper HM, et al. The netrin receptor DCC is required in the pubertal organization of mesocortical dopamine circuitry. J Neurosci. 2011;31:8381–94.

16. Naneix F, Marchand AR, Di Scala G, Pape JR, Coutureau E. Parallel maturation of goal-directed behavior and dopaminergic systems during adolescence. J Neurosci. 2012;32:16223–32.

17. Reynolds LM, Pokinko M, Torres-Berrío A, Cuesta S, Lambert LC, Del Cid Pelliter E, et al. DCC receptors drive prefrontal cortex maturation by determining dopamine axon targeting in adolescence. Biol Psychiatry. 2018;83:181–92.

18. Willing J, Cortes LR, Brodsky JM, Kim T, Juraska JM. Innervation of the medial prefrontal cortex by tyrosine hydroxylase immunoreactive fibers during adolescence in male and female rats. Dev Psychobiol. 2017;59:583–9.

19. Manitt C, Eng C, Pokinko M, Ryan RT, Torres-Berrío A, Lopez JP, et al. Dcc orchestrates the development of the prefrontal cortex during adolescence and is altered in psychiatric patients. Transl Psychiatry. 2013;3:e338–e338.

20. Pokinko M, Moquin L, Torres-Berrío A, Gratton A, Flores C. Resilience to amphetamine in mouse models of netrin-1 haploinsufficiency: role of mesocortical dopamine. Psychopharmacol (Berl). 2015;232:3719–29.

21. Flores C. Role of netrin-1 in the organization and function of the mesocorticolimbic dopamine system. J Psychiatry Neurosci. 2011;36:296–310.

22. Grace AA, Floresco SB, Goto Y, Lodge DJ. Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. Trends Neurosci. 2007;30:220–7.

23. Jackson ME, Frost AS, Moghaddam B. Stimulation of prefrontal cortex at physiologically relevant frequencies inhibits dopamine release in the nucleus accumbens. J Neurochem. 2001;78:920–3.

24. Simpson EH, Kellendonk C, Kandel E. A possible role for the striatum in the pathogenesis of the cognitive symptoms of schizophrenia. Neuron. 2010;65:585–96.

25. Kellendonk C, Simpson EH, Polan HJ, Malleret G, Vronskaya S, Winiger V, et al. Transient and selective overexpression of dopamine D2 receptors in the striatum causes persistent abnormalities in prefrontal cortex functioning. Neuron. 2006;49:603–15.

26. Kim CK, Yang SJ, Pichamoorthy N, Young NP, Kauvar I, Jennings JH, et al. Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. Nat Methods. 2016;13:325–8.

27. Hikida T, Yao S, Macpherson T, Fukakusa A, Morita M, Kimura H, et al. Nucleus accumbens pathways control cell-specific gene expression in the medial prefrontal cortex. Sci Rep. 2020;10:1–9.

28. Dalley JW, Roiser JP. Dopamine, serotonin and impulsivity. Neuroscience. 2012;215:42–58.

29. Dalley JW, Robbins TW. Fractionating impulsivity: neuropsychiatric implications. Nat Rev Neurosci. 2017;18:158–71.

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

31. Torres-Berrío A, Hernandez G, Nestler EJ, Flores C. The Netrin-1/DCC guidance cue pathway as a molecular target in depression: translational evidence. Biol Psychiatry. 2020;88:611–24.

32. Lee PH, Anttila V, Won H, Feng YCA, Rosenthal J, Zhu Z, et al. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. Cell. 2019;179:1469–82.e11.

33. Robbins TW, Gillan CM, Smith DG, de Wit S, Ersche KD. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. Trends Cogn Sci Elsevier Curr Trends. 2012;16:81–91.

34. Morgunova A, Pokhvisneva I, Nolvi S, Entringer S, Wadhwa P, Gilmore J, et al. DCC gene network in the prefrontal cortex is associated with total brain volume in childhood. J Psychiatry Neurosci. 2021;46:E154–63.

35. Cuesta S, Nouel D, Reynolds LM, Morgunova A, Torres-Berrío A, White A, et al. Dopamine axon targeting in the nucleus accumbens in adolescence requires netrin-1. Front Cell Dev Biol. 2020;8:1–12.

36. Vosberg DE, Zhang Y, Menegaux A, Chalupa A, Manitt C, Zehntner S, et al. Mesocorticolimbic connectivity and volumetric alterations in dcc mutation carriers. J Neurosci. 2018;38:4655–65.

37. Vosberg DE, Beaulé V, Torres-Berrío A, Cooke D, Chalupa A, Jaworska N, et al. Neural function in DCC mutation carriers with and without mirror movements. Ann Neurol. 2019;85:433–42.

38. Sanchez-Roige S, Fontanillas P, Elson SL, Gray JC, De Wit H, MacKillop J, et al. Genome-wide association studies of impulsive personality traits (BIS-11 and UPPS-P) and drug experimentation in up to 22,861 adult research participants identify loci in the CACNA1I and CADM2 genes. J Neurosci. 2019;39:2562–72.

39. Lee PH, Feng YCA, Smoller JW. Pleiotropy and cross-disorder genetics among psychiatric disorders. Biol Psychiatry. 2021;89:20–31.

40. Watanabe K, Stringer S, Frei O, Umićević Mirkov M, de Leeuw C, Polderman TJC, et al. A global overview of pleiotropy and genetic architecture in complex traits. Nat Genet. 2019;51:1339–48.

41. Reynolds T, Johnson EC, Huggett SB, Bubier JA, Palmer RHC, Agrawal A, et al. Interpretation of psychiatric genome-wide association studies with multispecies heterogeneous functional genomic data integration. Neuropsychopharmacology. 2021;46:86–97.

42. Hari Dass SA, McCracken K, Pokhvisneva I, Chen LM, Garg E, Nguyen TTT, et al. A biologically-informed polygenic score identifies endophenotypes and clinical conditions associated with the insulin receptor function on specific brain regions. EBioMedicine. 2019;42:188–202.

43. Silveira PP, Pokhvisneva I, Parent C, Cai S, Rema ASS, Broekman BFP, et al. Cumulative prenatal exposure to adversity reveals associations with a broad range of neurodevelopmental outcomes

that are moderated by a novel, biologically informed polygenetic score based on the serotonin transporter solute carrier family C6, member 4. Dev Psychopathol. 2017;29:1601–17.

44. Ardlie KG, DeLuca DS, Segrè AV, Sullivan TJ, Young TR, Gelfand ET, et al. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science (80-) 2015;348:648–60.

45. O'Donnell KA, Gaudreau H, Colalillo S, Steiner M, Atkinson L, Moss E, et al. The maternal adversity, vulnerability and neurodevelopment project: Theory and methodology. Can J Psychiatry. 2014;59:497–508.

46. Soh S-E, Tint MT, Gluckman PD, Godfrey KM, Rifkin-Graboi A, Chan YH, et al. Cohort profile: growing up in Singapore towards healthy outcomes (GUSTO) birth cohort study. Int J Epidemiol. 2014;43:1401–9.

47. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort profile: The 'Children of the 90s'-The index offspring of the avon longitudinal study of parents and children. Int J Epidemiol. 2013;42:111–27.

48. Fraser A, Macdonald-wallis C, Tilling K, Boyd A, Golding J. Davey smith G, et al. Cohort profile: The avon longitudinal study of parents and children: ALSPAC mothers cohort. Int J Epidemiol. 2013;42:97–110.

49. Miguel PM, Pereira LO, Barth B, de Mendonça Filho EJ, Pokhvisneva I, Nguyen TTT, 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. 2019;86:621–30.

50. Barrett T, Beck J, Benson DA, Bollin C, Bolton E, Bourexis D, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2015;43:D6–17.

51. Reynolds LM, Yetnikoff L, Pokinko M, Wodzinski M, Epelbaum JG, Lambert LC, et al. Early adolescence is a critical period for the maturation of inhibitory behavior. Cereb Cortex. 2019;29:3676–86.

52. Cuesta S, Restrepo-Lozano JM, Popescu C, He S, Reynolds LM, Israel S, et al. DCC-related developmental effects of abused- versus therapeutic-like amphetamine doses in adolescence. Addict Biol. 2020;25:1–11.

53. Urbanek S, Bibiko H-J, Stefano ML. R: a language and environment for statistical computing. The R Foundation for Statistical Computing. 2014.

54. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47:D607–13.

55. Demchak B, Hull T, Reich M, Liefeld T, Smoot M, Ideker T, et al. Cytoscape: the network visualization tool for GenomeSpace workflows. F1000Res. 2014;2014:1–12.

56. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. 2010;38: W214–20.

57. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8:1826.

58. Ebens A, Brose K, Leonardo ED, Hanson MG, Bladt F, Birchmeier C, et al. Hepatocyte growth factor/scatter factor is an axonal chemoattractant and a neurotrophic factor for spinal motor neurons. Neuron.1996;17:1157–72.

59. Schmeisser MJ, Baumann B, Johannsen S, Vindedal GF, Jensen V, Hvalby OC, et al. IκB kinase/nuclear factor κB-dependent insulin-like growth factor 2 (Igf2) expression regulates synapse formation and spine maturation via Igf2 receptor signaling. J Neurosci. 2012;32:5688–703.

60. Xu X, Wells AB, O'Brien DR, Nehorai A, Dougherty JD. Cell type-specific expression analysis to identify putative cellular mechanisms for neurogenetic disorders. J Neurosci. 2014;34:1420–31.

61. Dougherty JD, Schmidt EF, Nakajima M, Heintz N. Analytical approaches to RNA profiling data for the identification of genes enriched in specific cells. Nucleic Acids Res. 2010;38:4218–30.

62. Miller JA, Ding SL, Sunkin SM, Smith KA, Ng L, Szafer A, et al. Transcriptional landscape of the prenatal human brain. Nature 2014;508:199–206.

63. Skene NG, Roy M, Grant SGN. A genomic lifespan program that reorganises the young adult brain is targeted in schizophrenia. Elife. 2017;6:1–30.

64. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet. 2019;51:63–75.

65. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet. 2019;51:584–91.

66. Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, et al. Human demographic history impacts genetic risk prediction across diverse populations. Am J Hum Genet. 2017;100:635–49.

67. de Lima RMS, Barth B, Arcego DM, de Mendonça Filho EJ, Clappison A, Patel S, et al. Amygdala 5-HTT gene network moderates the effects of postnatal adversity on attention problems: anatomo-functional correlation and epigenetic changes. Front Neurosci. 2020;14:1–18.

68. Shohat S, Amelan A, Shifman S. Convergence and divergence in the genetics of psychiatric disorders from pathways to developmental stages. Biol Psychiatry. 2021;89:32–40.

69. Jamuar SS, Schmitz-Abe K, D'Gama AM, Drottar M, Chan WM, Peeva M, et al. Biallelic mutations in human DCC cause developmental split-brain syndrome. Nat Genet. 2017;49:606–12.

70. Marsh APL, Heron D, Edwards TJ, Quartier A, Galea C, Nava C, et al. Mutations in DCC cause isolated agenesis of the corpus callosum with incomplete penetrance. Nat Genet. 2017;49:511–4.

71. Srour M, Rivière J-B, Pham JMT, Dubé M-P, Girard S, Morin S, et al. Mutations in DCC cause congenital mirror movements. Science. 2010;328:592.

72. Finger JH, Bronson RT, Harris B, Johnson K, Przyborski SA, Ackerman SL. The netrin 1 receptors Unc5h3 and Dcc are necessary at multiple choice points for the guidance of corticospinal tract axons. J Neurosci. 2002;22:10346–56.

73. Glasgow SD, Ruthazer ES, Kennedy TE. Guiding synaptic plasticity: novel roles for netrin-1 in synaptic plasticity and memory formation in the adult brain. J Physiol. 2021;599:493–505.

Connecting statement to Chapter IV

As detailed in Chapter 3, our exploration of gene co-expression networks, including the implementation of expression-based polygenic scores (ePRS) specific to cognitive control and impulsivity-related phenotypes, has laid the groundwork for a more nuanced understanding of psychiatric conditions. One that is much more in line with the well-known description of psychiatric conditions as being highly *polygenic*. With the advent of modern genomic profiling methodologies, the need to move beyond traditional gene-disease models has become clear. In Chapter 4, I transition from the prior focus on genotype-phenotype associations to an in-depth review of cutting-edge genotype-gene regulation frameworks. These frameworks are transformative, incorporating functional molecular phenotypes closer to genetic variation and more resilient to the multiple testing required in genome-wide association studies. This chapter will enumerate various functional genomics tools and innovative methodologies that may serve as potential biomarkers of psychiatric disease susceptibility, and how they can be used and integrated to enhance both research and clinical care. In doing so, I pave the way for generating biologically driven hypotheses and optimizing current approaches to genetic risk assessment, aligning our understanding of genetic factors with the broader biological processes involved in psychopathology.

CHAPTER IV: NOVEL FUNCTIONAL GENOMICS APPROACHES BRIDGING NEUROSCIENCE AND PSYCHIATRY – COMPREHENSIVE REVIEW OF THE LITERATURE

Authors:

Jose M. Restrepo-Lozano[1,2], Cecilia Flores[2,3,5], and Patricia P. Silveira[2,4,5]

Affiliations:

[1]Integrated Program in Neuroscience, McGill University, Montreal, Quebec, Canada, [2]Douglas Mental Health University Institute, Montreal, Quebec, Canada, [3]Departmentof Neurology and Neurosurgery, Faculty of Medicine, McGill University, Montreal, Quebec, Canada, [4]Ludmer Centre for Neuroinformatics and Mental Health, Douglas Mental Health University Institute, McGill University, Montreal, Quebec, Canada, [5]Departmentof Psychiatry, Faculty of Medicine, McGill University, Montreal, Quebec, Canada.

Published in

Biological Psychiatry Global Open Science. 2023; 3(3), 351-361.

https://doi.org/10.1016/j.bpsgos.2022.07.005

Abstract

The possibility of establishing a metric of individual genetic risk for a particular disease or trait has sparked the interest of the clinical and research communities, with many groups developing and validating genomic profiling methodologies for their potential application in clinical care. Current approaches for calculating genetic risk to specific psychiatric conditions consist of aggregating genome-wide association studies–derived estimates into polygenic risk scores, which broadly represent the number of inherited risk alleles for an individual. While the traditional approach for polygenic risk score calculation aggregates estimates of gene-disease associations, novel alternative approaches have started to consider functional molecular phenotypes that are closer to genetic variation and are less penalized by the multiple testing required in genome-wide association studies. Moving the focus from genotype-disease to genotype-gene regulation frameworks, these novel approaches incorporate prior knowledge regarding biological processes involved in disease and aggregate estimates for the association of genotypes and phenotypes using multiomics data modalities. In this review, we discuss and list different functional genomics tools that can be used and integrated to inform researchers and clinicians for a better understanding and diagnosis of psychopathology. We suggest that these novel approaches can help generate biologically driven hypotheses for polygenic signals that can ultimately serve the clinical community as potential biomarkers of psychiatric disease susceptibility.

Introduction

Establishing potential high-risk scenarios prior to the onset of neuropsychiatric conditions could profoundly improve mental health trajectories worldwide by presenting an opportunity for timely interventions, especially during sensitive neurodevelopmental windows. Although the well-established practice of inquiring about an individual's family history when diagnosing physical and psychiatric conditions is a useful tool to indirectly assess potential heritable risk (1–3), an individual's genomic profile could provide information to guide overall health management. However, the true value of genomic data relies on our understanding of the complex interaction between genes, environments, and lifestyle choices over time (4–6), and efforts to elucidate this complex interplay have the potential to help develop tools to assess disease susceptibility prior to symptom onset, informing preventive and therapeutic decisions.

Current genotyping technology allows the identification of inherited DNA differences in the order of millions, mostly in the form of single nucleotide polymorphisms (SNPs), across a given population and in a rapid and affordable manner (7). As a result, studying genotype-phenotype associations changed from interrogating a few carefully selected candidate genes at a time to unbiased genome-wide surveys, with constant increases in sample sizes leading to the identification of an increasing number of genetic loci that could modify risk for a given disease (8). Although this systematic interrogation of genomes yielded several loci reliably associated with an increased risk for psychiatric phenotypes, linking such loci to specific biological functions remains a challenge, primarily because most identified genome-wide significant associations lie in noncoding portions of the genome and require fine-mapping resolution to determine the real causal variants implicated (8–10). Establishing a neurobiological framework underlying psychiatric risk will require a multi-omics data integration approach, with the purpose of mapping the molecular processes linking genomes and disease-relevant phenotypes (11). Such frameworks may ultimately help improve models of disease risk prediction based on genomic profiles

and provide actionable insights for clinical decision making. In this review, we discuss emerging genomic risk assessment approaches in psychiatry, emphasizing methods that explore the neurobiological mechanisms by which gene networks contribute to psychiatric phenotypes.

Genome-Wide Association Studies as the Basis for Mapping Genetic Susceptibility to Psychiatric Phenotypes

To date, the most common population-based method to find genotype-phenotype associations is the performance of genome-wide association studies (GWASs) [see (12,13)], which has successfully helped identify genomic variants associated with increased risk of developing different psychiatric conditions (9,14–16). Essentially, GWASs entail the assessment of millions of variants across many individuals to detect those statistically associated with a specific phenotype. The primary outcome of GWASs typically includes a list of tested variants together with their respective effect sizes. Then, after identifying the relationship between the phenotypic variance and each genotype by means of a linear (for continuous) or logistic (for binary outcomes) regression, significant loci can be functionally annotated for post-GWAS analyses (Figure 1A). Psychiatric genomics studies for conditions such as schizophrenia (17) and depression (18) have yielded .100 robustly associated risk loci, with w43.7% and w8.9% of heritability explained by common SNPs, respectively. The remarkable collaborative effort from the Psychiatric Genomics Consortium (PGC) has helped generate important discoveries in the identification of risk-conferring variants as well as in advancing our understanding of the genetic architecture across 11 psychiatric disorders (17–27).

GWAS-derived quantified effects of common human variation have translated into different clinical applications. For example, using data derived from human genetics studies has improved the successful development of novel drugs (28,29). Another application central to this review is the calculation of polygenic risk scores (PRSs), which aim to predict the contribution of an individual's genomic profile to a

given trait or disease (16,30–35). The possibility of establishing a metric of individual genetic risk for a particular disease or trait has sparked the interest of the clinical community, with many researchers now investigating and exploiting the utility of PRS profiling in clinical care [e.g., (34) or (36)].

Aggregating GWAS-Derived Signals Into PRSs: A Proxy for Genetic Liability to Psychiatric Traits

For many years, studies in psychiatric genetics used a candidate gene approach, investigating the role of SNPs in particular phenotypes [e.g., (37), where a specific mutation in the HTR2B gene was associated with increased impulsivity]. However, this approach to study the contribution of common variants to psychiatric phenotypes required a previously defined SNP target that was arbitrarily selected, albeit with very few exceptions. Indeed, conditions such as Huntington's disease (38) are caused by large effect variants, and there is a marked increase in risk for Alzheimer's disease (AD) (although not a determinant of the disease itself) in people with the isoform e4 of the APOE gene (39). However, Huntington's disease and AD are neurologic conditions with a more defined clinical phenotype compared with psychiatric conditions such as mood disorders, where the degree of polygenicity is even more evident. The candidate gene approach is now considered outdated because it has failed to yield useful insights for psychiatry [see (8) for a perspective on how GWASs made candidate gene studies obsolete]. Current psychiatric genetics studies use an unbiased examination of the genome, as a continuously growing body of evidence established the highly polygenic architecture across disorders, with many small-effect risk loci distributed across the entire genome (40–43). As psychiatry gradually adopted a more probabilistic and risk-oriented mindset, evidence for a concept that could explain a significant proportion of heritability in independent target samples, based entirely on inherited DNA differences, began to emerge (44,45).

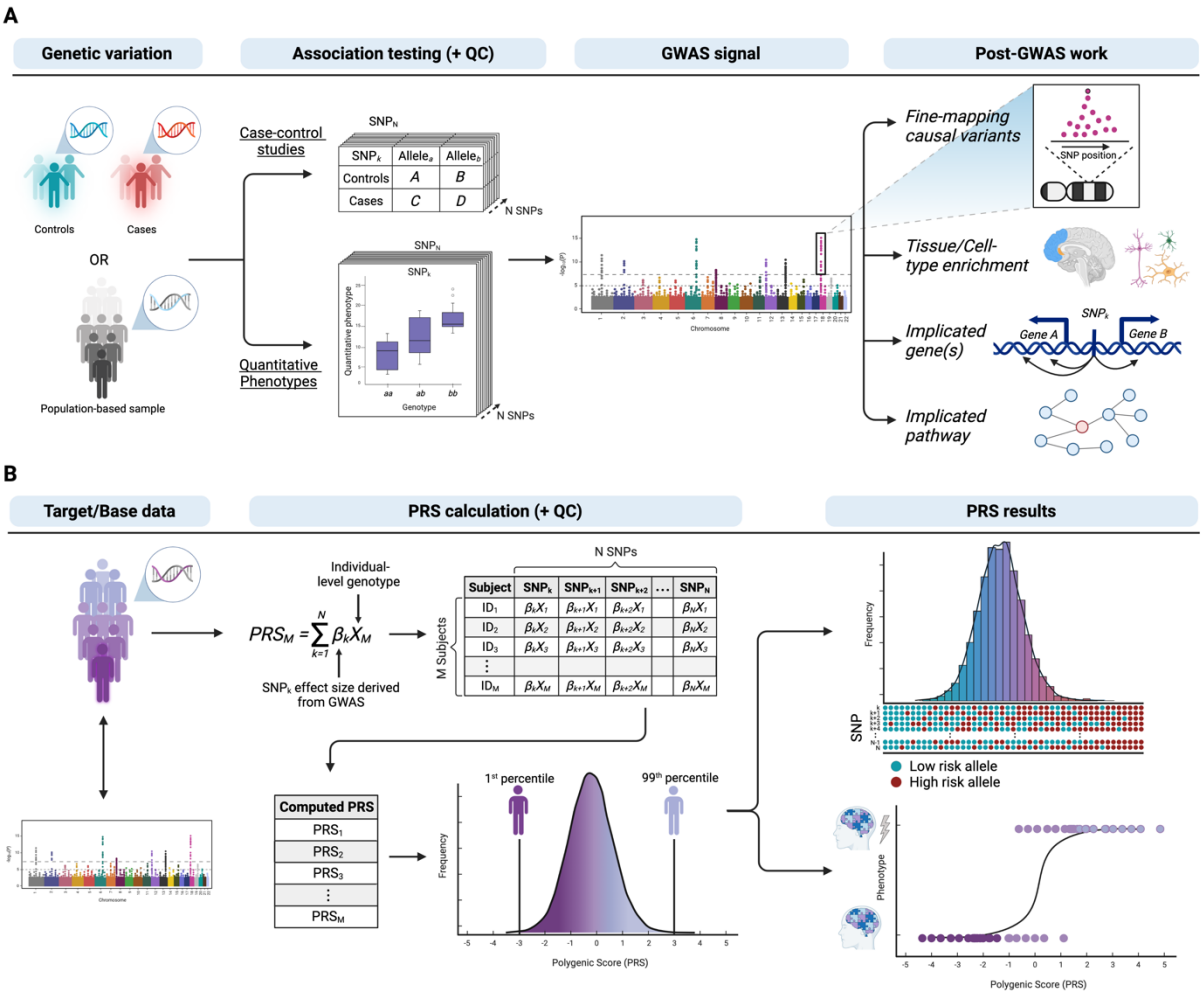

Figure 1. Overview of the steps involved when conducting a GWAS and creating a PRS. (A) Genotypic data from cases and controls or from a population-based sample are gathered to compare the proportion of specific alleles from each SNP among cases and controls or to determine the linear relationship between genotypes and a continuous trait. After proper quality control of the genotype data and determination of the underlying population structure in the sample, a statistical analysis is conducted to investigate whether the observed allele proportions (for case-control studies) or relationships (for continuous traits) deviate significantly from expected values at each SNP, correcting for the number of tests applied. When an allele is found in the cases more frequently than it would be expected by chance, it is reported as a candidate SNP for the entire haplotype block, together with its estimated effect size which quantifies the increased odds of having the disease per risk

allele count. For continuous traits, the regression coefficient will determine the effect size attributed to the "effect" allele. Ideally, the observed GWAS signal should be replicated in an independent cohort to minimize false positives and to calibrate the effect sizes attributed to all SNPs. Genome-wide signals (shown in a Manhattan plot) that have been replicated are typically further investigated during post-GWAS work, which consists of 1) fine-mapping the genomic region to find the true causal variant, 2) investigating the tissues/cell types where the variant is known to be active, 3) determining the genes that are affected by the variant, and 4) identifying the molecular pathways implicated. (B) Using a base and a target dataset, the GWAS-derived estimated effects can be applied to a target sample for which genotype data are available. The calculated PRS is an aggregated score of the individual-level genotype weighted by the SNP effect sizes described in a discovery GWAS, resulting in a normally distributed score in the target sample. The distributions depicted in panel (B) reflect raw standardized values of real PRSs, which could be associated with a particular trait of interest. GWAS, genome-wide association study; OR, odds ratio; PRS, polygenic risk score; QC, quality control; SNP, single nucleotide polymorphism.

Current Methodologies for PRS Calculation in Psychiatry and Important Considerations to Obtain Meaningful Genetic Signals

In principle, all methods of PRS calculation provide an estimate of an individual's genetic susceptibility to a trait by aggregating the GWAS-derived effect size estimates into an indexed score, as shown in Figure 1B [for a detailed PRS tutorial, see (32); for a detailed PRS review, see (46)]. The classic method of PRS calculation uses clumping or pruning and thresholding (C/P-T method) to prune out SNPs in high linkage disequilibrium and apply varying stringencies to p-value thresholds that can be higher than genome-wide significance to calibrate and maximize predictability (30,46,47). Essentially, SNPs with p values below an established threshold will keep the original estimate of their effect size, while SNPs with higher p values are excluded from the PRS, shrinking their effect sizes to 0. This process can be carried out iteratively, using a range of p-value thresholds, with the resulting PRSs tested for an association with the target trait in a test sample, determining the optimal p value in a forward selection method (48,49). Other methods for PRS calculation are based on Bayesian frameworks in which the shrinkage of all SNPs is based on a prior distribution specification [for more details, see (50,51)]. One example that seems to be particularly suited to calculate PRSs for psychiatric disorders (52) is the Bayesian multiple regression summary statistic (SBayesR) (53), which can use publicly available GWAS summary statistics while using prior distributions of alternative genetic effects and analyzing all SNPs together, accounting for their pattern of coinheritance.

Ideally, a PRS can serve as a tool to stratify the population in terms of disease risk, as this can help decide on potential follow-up actionable measures such as therapeutic interventions, more in-depth screening, or lifestyle modifications. One of the earliest examples of a successful PRS came in 2009 when the International Schizophrenia Consortium (ISC) published an aggregated polygenic signal derived from a GWAS that could predict risk for both schizophrenia and bipolar disorder (44). As the sample size for the schizophrenia GWASs increased, the phenotypic variance explained by the aggregated polygenic

signal also increased. Current estimates indicate that individuals with PRS in the top 10% and top 1% of the population have an approximate 3-fold and 6-fold increase in their risk of developing schizophrenia, respectively, compared with 1% baseline risk when selecting someone randomly from the population (17,54). Another example comes from the study of Desikan et al. (55), wherein the researchers calculated a PRS based on a large AD GWAS meta-analysis (56) to investigate the PRS predictability of age-specific risk of developing the disease. By combining epidemiological data on population-based incidence rates and PRSs, they found that individuals in the highest PRS quartile developed AD at a lower age and showed the highest yearly AD incidence rate. This finding was then replicated in other independent cohorts, where the PRS was associated with known neurodegenerative markers and with the age of disease onset (55).

It is important to note that existing GWASs are predominantly performed using individuals of European ancestry. Missing genetic effects present in other populations and genetic variants with very low frequency may dramatically decrease the accuracy of a PRS. This is especially true when the ancestry of the target sample does not match the population of the original GWAS (57,58). In addition, it has been shown that PRSs work better when considered in combination with other clinical risk factors, with a joint model improving overall disease risk calculation, the identification of individuals that can benefit from early diagnosis, and predictive accuracy (55,59–62). Prediction is a difficult task, and most GWASs necessitate many millions of individuals to allow PRSs to achieve higher discriminatory power and reach the upper bound of their predictive performance (i.e., heritability estimates) (33). Some groups have started to propose alternatives to investigate polygenic signals in psychiatry, considering phenotypes closely linked to genetic variation and therefore more directly affected by it.

Method Name	Description
MAGMA	Software tool for mapping genome-wide significant variants to genes and gene sets. Novel variations of this method (i.e., H-MAGMA and eMAGMA) are meant to refine the mapping of variants by incorporating long-range and tissue-specific interactions and the enrichment of variants across different gene modules.
LDSC	Method that leverages GWAS summary statistics and LD scores from an external panel to distinguish between inflated effect sizes and true polygenic effects. This method is commonly used for determining genetic correlation between complex traits, partitioned heritability, and stratified heritability.
SMR	Method that integrates GWAS summary statistics and data from eQTL studies, allowing the user to identify and prioritize genes whose expression levels are associated with specific complex traits.
lassosum	Method to construct a PRS in a penalized regression framework that uses GWAS summary statistics and an LD reference panel.
LD-Hub	Centralized database of GWAS summary statistics that automates LDSC analysis pipeline, allowing the user to estimate SNP heritability and genetic correlation across complex traits.
ANNOPRED	Bayesian framework for disease risk prediction that integrates genomic functional annotations using GWAS summary statistics and estimates LD from reference genotype data.
SBLUP	Method that rescales SNP effect sizes using an external LD reference panel, converting the ordinary least squares SNP estimates into approximate best linear unbiased predictions.
PRS-CS	Polygenic prediction method that infers posterior effect sizes of SNPs using GWAS summary statistics and an external LD panel. This model places a continuous shrinkage prior on SNP effect sizes.
JAMPred	Method for modeling polygenic risk using the JAM software, adjusting for local and for long-range LD. The computed polygenic risk predictions are obtained through a Bayesian variable selection framework.
SBayesR	Polygenic prediction method that adjusts SNP effect sizes based on Bayesian multiple regression model (BayesR), using GWAS summary statistics data.
LDpred-func	Probabilistic model for deriving PRS that accounts for LD and incorporates trait-specific functional priors to increase prediction accuracy. This model assumes a point-normal distribution as a prior.
LDpred2	Method for deriving polygenic scores using GWAS summary statistics and LD information from an external reference sample to infer posterior mean effect sizes of SNPs. Optimization of LD and *p*-value thresholds is achieved using a Bayesian framework for shrinkage of SNP effects. This model assumes a point-normal distribution as a prior.
PTRS	Method for calculating PTRSs that can be applied as a gene-based complement to other PRS methods, as it does not outperform other current PRS technologies. This method can help improve portability across ancestries and facilitate interpretation of underlying biological effects.

Table 1. Overview of the different Methodologies for Post-GWAS Analysis. eQTL, expression quantitative trait locus; GWAS, genome-wide association study; LD, linkage disequilibrium; LDSC, linkage disequilibrium score regression; PRS, polygenic risk score; PTRS, polygenic transcriptome risk score; SNP, single nucleotide polymorphism.

From Genetics to Functional Genomics: PRS Methodologies That Go Beyond the Link Between Genetic Variability and Psychiatric Traits by Addressing Biological Mechanisms/Functions

The PRS methodologies described so far have been useful tools for clinicians and researchers, but one common characteristic is the agnosticism when it comes to the biological functions implicated in disease risk. In the classic GWAS-PRS methods, the first step consists of identifying statistically significant genetic associations such that afterward, while conducting post-GWAS work, the biological functions implicated in those gene-disease associations can be dissected (Figure 1A) and further explored as potential therapeutic avenues. However, another way to investigate the role that genes play in disease (together with their associated transcripts, proteins, and epigenomes) is to first identify disease-relevant biological processes and functions to create PRSs that somehow capture and quantify those functions and to then test their association with disease (see Table 1). Moving the focus from genotype-disease toward genotype-gene regulation frameworks, below we review these novel methodologies and resources used by some groups to guide the selection of variants and phenotypes, emphasizing those that take into consideration 1) meaningful networks of genes coregulated (or coexpressed) with spatiotemporal specificity and 2) highly quantifiable phenotypes, such as transcriptomic or epigenomic data. We suggest that these approaches can help generate biologically driven hypotheses for polygenic signals that can ultimately serve the clinical community as potential biomarkers for disease susceptibility.

Genotype-disease effects are small for most common genetic variations, but the fact that a large proportion of disease risk can be explained by variants that modulate gene expression levels (9,63,64) is intriguing and may provide clues for the cellular and biological mechanisms underlying disease (65,66). The transcriptome-wide association study methodology was developed with the goal of detecting associations between measured or predicted levels of gene expression and particular traits (Figure 2A)

(67). For example, in the study of Girgenti et al. (68), researchers used the Million Veteran Program posttraumatic stress disorder GWAS dataset to impute gene expression and identify genes significantly associated with posttraumatic stress disorder risk and illness state, uncovering novel functional signals that confer genetic liability for posttraumatic stress disorder. This method provides key advantages with respect to GWAS. First, using a gene-based approach reduces the burden of multiple testing prevalent in other SNP-based approaches. There are approximately 20,000 genes for which one can impute transcript levels. Although large, this number is considerably smaller compared with several million SNPs in a typical GWAS that are individually tested for an association with a given trait. By incorporating functional information about the regulation of gene expression, this method can help uncover the underlying biological mechanisms affecting a trait. Another advantage of this method is that it facilitates the interpretation of the direction of the effect. A gene-based signal that includes the direction of the effect is highly amenable to systems biology approaches because if the increased (or decreased) expression of a gene is associated with a particular trait, the information can be easily incorporated into pathway or network analyses, making the interpretation of results more straightforward, especially when compared with SNP-based signals. This approach is nonetheless limited regarding tissue accessibility in study participants, in particular if the tissue of interest is a specific brain region. To this end, a more novel computational framework such as the probabilistic transcriptome-wide association study can be of help because it can predict gene expression from genotypes and investigate causal relationships between tissue- or cell-type–specific gene expression and complex traits (69).

Genotype-Based Prediction of Gene Expression in Specific Tissues

Many research groups are actively developing tools to predict the transcriptional effects of genetic variation [see (70–72) for some examples], most likely driven by similar motivations: a unidirectional effect (from genes to gene expression) that ultimately narrows the gap between genetic variation and disease. One of the most prominent examples of this type of work is the PrediXcan methodology, which

developed a machine learning algorithm to predict tissue-specific gene expression based on genomic profiles (73). Using genotype and gene expression data from the Genotype-Tissue Expression (GTEx) project (74) and other similar datasets, this method generates a database wherein, in a tissue-specific manner, transcript levels can be predicted using as input the genotypic data from any target sample (Figure 2B). PrediXcan serves to calculate an endophenotype (genetically regulated gene expression) that is known to drive biological processes to test for associations with a particular trait [for the entire data repository and the PrediXcan family of methods, see (75)]. The more novel version, MultiXcan, can help investigate the mediating role of gene expression on many complex traits, using only summary statistics from publicly available GWASs (76).

The predicted gene expression approach has been applied to existing GWASs for bipolar disorder to identify novel risk-conferring genes PTPRE and BBX, whose predicted transcript levels in whole blood and in the anterior cingulate cortex, respectively, were found to be associated with increased bipolar disorder risk (77). This study highlights the importance of gene expression to help understand the potential underlying mechanisms driving disease risk. However, this approach fails to simultaneously consider genes that are coregulated as part of common biological processes, bypassing the established polygenicity of most psychiatric phenotypes.

Gene Coexpression Networks to Inform Polygenic Metrics

As discussed previously, most identified genome-wide significant associations are devoid of a clear functional interpretation because they lie in noncoding portions of the genome, requiring fine-mapping resolution to determine the real causal variants implicated (9,10). Many of these noncoding disease-associated variants are regulatory in nature (a high proportion of these variants have been determined to be cis and/or trans–expression quantitative trait loci [eQTLs]) (9), suggesting that they likely affect the expression of their associated genes, ultimately placing gene expression as an imminent molecular

phenotype linking genetics and disease. More crucially, however, disease-associated genes do not operate in isolation, but as part of complex networks that function with an exquisite degree of spatiotemporal specificity for precise biological processes. By operating under the assumption that functional groups of genes are coregulated as part of specific molecular pathways, the identification of disease-relevant and tissue-specific gene networks provides a framework for mapping transcriptionally coregulated processes into a type of polygenic score. This approach can potentially increase the likelihood of discovering psychiatrically relevant markers of disease [see (78)].

A study that aimed to determine genetic susceptibility to cognitive disability used an unsupervised genome-wide coexpression network analysis leveraging measurements of gene expression in human hippocampal tissue, with the goal of capturing modules of covarying genes, which can ultimately provide clues for the molecular mechanisms driving the susceptibility (79). The study identified a module of 150 genes with significant enrichment for 1) genes associated with relevant cognitive phenotypes, 2) genes related to neural activity and synaptic processes, and 3) genes intolerant to mutations and that, when mutated, are associated with intellectual disability (80,81). Another group followed a similar approach, but instead of using an unsupervised analysis, they hypothesized that genes conferring risk to disease must translate into biological risk by acting as part of a coregulated gene network on a measurable molecular phenotype, which could then be associated with the disease (82). They were interested in elucidating the genetic architecture of the D2 receptor molecular pathway because genetic variation within the DRD2 gene has been linked with schizophrenia-related phenotypes, including response to treatment (17). Starting with human postmortem tissue, the authors identified a prefrontal DRD2 coexpression network using weighted gene coexpression network analysis, and then defined potential SNPs in the form of eQTLs affecting the expression of the genes within the network. Combining these regulatory SNPs into a particular PRS (referred to as polygenic coexpression index), the study captured the genetic component (eQTLs) of the expression of the network and associated the PRS with brain

activity measurements during working memory tasks. Finally, they found that individuals with a higher prefrontal cortex DRD2 coexpression PRS are predisposed to a less efficient working memory, which is a known risk-associated phenotype for schizophrenia. This study is an example of how identifying a disease or trait-relevant gene network can help generate hypotheses for novel types of PRSs based on biological frameworks.

Another innovative way to identify coregulated biological processes underlying the genetic susceptibility to psychiatric conditions leverages data from the GTEx to quantify the genotypic effect linked to gene expression across several tissues. One such example is the method eMAGMA, which integrates both genetic and transcriptomic data to identify disease-specific risk genes and test for their enrichment across different gene modules (83). This method can exploit a systems biology approach to generate polygenic signals that are essentially based on tissue-specific gene coexpression networks. A similar approach from our group has generated tissue-specific polygenic signals associated with traits or diseases (Figure 2C). We first identify coexpression networks using genome-wide gene expression data from a specific tissue, then map all SNPs within the coexpressed genes and eliminate those in linkage disequilibrium. We then assign to each SNP the weight of the association between alleles and gene expression estimated by GTEx (84), ultimately obtaining a set of SNPs that lie within a tissue-specific coexpression network, where each SNP is weighted by its estimated influence on gene expression. We can then identify all SNPs from the coexpressed genes in a test sample of subjects with available genotype data and weight the SNPs according to the GTEx. The derived expression-based polygenic signal (or expression-based PRS) reflects variation in the expression of the gene network and can be calculated in target samples with available genotype data (85–89).

In a recent study, we investigated whether an expression-based PRS based on corticolimbic-specific gene coexpression networks associates with impulsive phenotypes in children (90). We aimed at

capturing individual variation in the molecular processes involved in the maturation of corticolimbic substrates, which are known to support inhibitory control. Similar to most studies using functional polygenic signals, we compared the predictive ability of the score against a conventional PRS derived from the latest GWAS for attention-deficit/hyperactivity disorder and found the expression-based PRS to be a better overall predictor of impulsivity. This type of polygenic signal did not have generalizability problems seen in other polygenic score methodologies, as the experiment was conducted in 3 ethnically diverse cohorts, all showing similar effects. This approach exploits the fact that genes engage within complex networks for precise biological functions, and they do so with a remarkable tissue specificity. Based on knowledge of the neurobiological processes of brain development, this score aimed to predict psychiatric-relevant phenotypes.

Gene-by-Environment Interplay: Quantifying Environmental Influences and Their Interaction With Multi-omics Data

One of the biggest challenges faced by researchers studying models of disease risk prediction is to develop a methodology to accurately represent an individual's environment in a quantitative metric. Similar to how a functional PRS can represent a restricted set of phenotype-relevant biological processes, some studies have narrowed down the environment variable to a composite score made up of clearly defined constructs [see (48,49,85,86,88,89,91) for examples]. By doing so, researchers can start investigating the interplay between genes and environments while also assessing the potential ways in which genetic and environmental effects interact (Figure 2D). Although not in psychiatry, the study by Belsky et al. (92) is a good example of how an individual's environment can exert a powerful influence on his or her socioeconomic attainment. In this study, the authors tested whether a PRS based on a GWAS for educational attainment (which is currently one of the PRSs with highest predictive value) [see (35)] could predict socioeconomic mobility (i.e., any shift in a person's social class relative to that of their parents). While higher PRSs did predict more socioeconomic success than parents and siblings,

additional analyses revealed that the maternal polygenic score is associated with children's educational attainment even when adjusting for the children's own polygenic scores. This suggests an environmentally mediated genetic effect.

Some studies have started to integrate epigenetic data into genome-wide scores with the goal of identifying individuals who might be at an increased risk of psychiatric phenotypes. For example, given the association between early-life stress and behavioral and psychiatric problems later in life (93), the study of Provençal et al. (94) assessed differentially methylated sites following exposure to glucocorticoids in a human hippocampal progenitor cell line and in human blood cells. In addition, a subsequent glucocorticoid exposure induced important transcriptional changes. The overlapping differentially methylated sites were then used to calculate a weighted polyepigenetic score, which was proposed as a potential biomarker for conditions associated with prenatal glucocorticoid exposure in newborns. The calculated score was applied to newborns' cord blood DNA (n = 817), with the glucocorticoid-responsive score significantly associated with levels of maternal anxiety and depression, suggesting that early-life stress induces lasting epigenetic changes that can ultimately modify the vulnerability to stress exposure in later years.

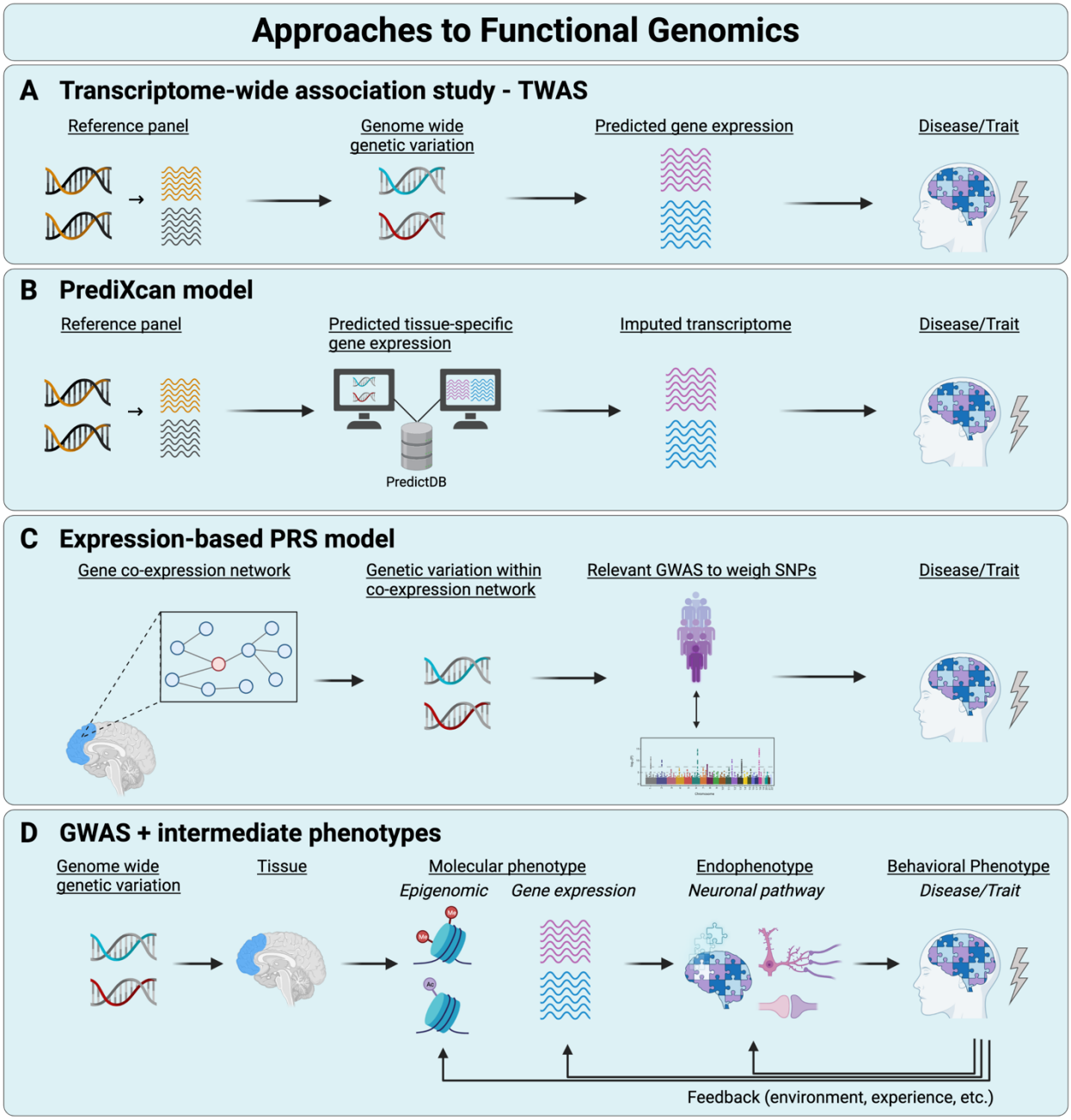

Figure 1. Novel approaches to functional genomics. (A) The TWAS consists of associating measured (or predicted) gene expression data with a disease or trait, making this a gene-based rather than a SNP-based association study, considerably reducing the number of multiple comparisons while also providing insight into the potential biological mechanisms driving disease risk. (B) Models that predict gene expression based on genotype data. These models can be tissue- or cell type–specific, allowing testing for the association of the cell-

or tissue-specific imputed transcriptome and a given disease or trait. (C) The ePRS model starts with the identification of a gene coexpression network from RNA sequencing data to identify coregulated disease-relevant biological processes in specific tissues. Once a gene network is identified, it maps the genetic variation within the coexpression network in an independent target sample to weigh the SNPs according to a GWAS (typically a GWAS of gene expression, e.g., GTEx). The resulting ePRS, which aims to capture individual variation in the expression of the gene network, can then be associated with a disease or trait, mapping the association between functional biological processes and the target phenotype. (D) A global model that incorporates all the potential intermediate phenotypes that can occur between genetic variation and disease, considering the feedback from the environment and lived experiences across all levels. This model, similar to the ones discussed earlier, generates a unidirectional effect that starts from individual genetic variability and provides alternative approaches to assess the effects of inherited DNA polymorphisms on particular traits. All these approaches can help generate biologically informed predictors of susceptibility to psychiatric-relevant phenotypes. ePRS, expression-based polygenic risk score; GTEx, Genotype-Tissue Expression project; GWAS, genome-wide association study; SNP, single nucleotide polymorphism; TWAS, transcriptome-wide association study.

Conclusions

As the field of psychiatric genomics continues to evolve, so will the models of disease risk prediction based on strong biological foundations. Advances in big data availability and complexity (e.g., longitudinal studies such as the ABCD [Adolescent Brain Cognitive Development] cohort, deep phenotyping as in the UK Biobank), mapping the developmental trajectories, and including a wealth of data in large numbers of individuals will benefit the understanding of factors that ultimately play an important role in determining mental health. Polygenic scores and polyepigenetic scores by themselves, like any other marker, have a limited capacity to predict with perfect accuracy the condition for which they were generated. It should be noted, however, that the optimal selection of genetic variants and other genomic markers and the aggregation of their associated weights are active areas of research (50,51,53). The continued improvement of the technology (increases in GWAS sample size and incorporation of different ancestries, higher genotyping resolution, etc.) entails continued revision of the guidelines for their calculation and interpretation. Owing to the recency of the appearance of several methods discussed in this review, evidence of their clinical utility is still lacking, but as the technology driving functional genomics approaches continues to improve, we expect researchers and clinicians to be encouraged to investigate or test their clinical utility in psychiatry. One can assume that some of the methods highlighted here will be replaced by newer approaches. However, incorporating functional aspects rather than being informed exclusively by data-driven approaches is our core message.

Although it is highlighted that functional PRS can focus on a particular network or system, efforts should be made to maintain a genome-wide platform for unbiased querying of the relevant signals (e.g., genome-wide RNA sequencing). Moreover, one of the advantages of the functional methods is to provide tissue-specific information, but this can be challenging for certain research questions (e.g., epigenetics markers collected from peripheral studies inferring brain mechanisms). Finally, brain gene

expression data in humans is postmortem and is limited in numbers, ancestry, and developmental stage representation. These features can influence gene expression and therefore can bias the generation of functional PRS.

Training of the clinical workforce to handle and communicate genome-wide information is an issue that is becoming more pressing with time. Commercially available and direct-to-consumer genotyping services, which allow users to download their genotype data, are already reporting PRSs for some traits, and users can upload their data into other online PRS calculators (95). It is important to clearly communicate to the public the utility and, most importantly, the limitations of PRS profiling, in particular driving away the idea that genetic testing can accurately predict every aspect of a person's health, as it has inherent limitations similar to those of other tests commonly used in clinical settings (96).

Future efforts in disease risk prediction should aim at integrating data at multiple levels, aggregating genomics, epigenomics, transcriptomics, proteomics, and metabolomics data into predictive models. Some of the examples presented in this review highlight the significant contribution from each of these data to disease liability. In addition, models should be able to assess the role of the environment at multiple levels (person, family, community) because all biological processes occurring within an individual are physically contained processes functioning together as part of society, including household, neighborhood, school, and work, and so on (4). The technological advance should occur in parallel to societal progress in overcoming the menace of racism and structural inequalities, which still are an unfortunate reality that has a major impact on mental and physical health. The ability to combine multilevel biological information with the constant changes in a person's environment for overall health risk assessment in trusted clinician-patient relationships with joint decision making can revolutionize the diagnosis and early prevention of psychopathology.

References

1. Weissman MM, Wickramaratne P, Adams P, Wolk S, Verdeli H, Olfson M (2000): Brief screening for family psychiatric history: The family history screen. Arch Gen Psychiatry 57:675–682.

2. Thompson WD, Orvaschel H, Prusoff BA, Kidd KK (1982): An evaluation of the family history method for ascertaining psychiatric disorders. Arch Gen Psychiatry 39:53–58.

3. Baselmans BML, Yengo L, van Rheenen W, Wray NR (2021): Risk in relatives, heritability, SNP-based heritability, and genetic correlations in psychiatric disorders: A review. Biol Psychiatry 89:11–19.

4. Patel CJ, Ioannidis JP (2014): Studying the elusive environment in large scale. JAMA 311:2173–2174.

5. Marigorta UM, Gibson G (2014): A simulation study of gene-by-environment interactions in GWAS implies ample hidden effects. Front Genet 5:225.

6. Idaghdour Y, Awadalla P (2013): Exploiting gene expression variation to capture gene-environment interactions for disease. Front Genet 3:228.

7. Perkel J (2008): SNP genotyping: Six technologies that keyed a revolution. Nat Methods 5:447–453.

8. Duncan LE, Ostacher M, Ballon J (2019): How genome-wide association studies (GWAS) made traditional candidate gene studies obsolete. Neuropsychopharmacology 44:1518–1523.

9. Watanabe K, Stringer S, Frei O, Umi cevi c Mirkov M, de Leeuw C, Polderman TJC, et al. (2019): A global overview of pleiotropy and genetic architecture in complex traits [published correction appears in Nat Genet. 2020 52:353]. Nat Genet 51:1339–1348.

10. Reynolds T, Johnson EC, Huggett SB, Bubier JA, Palmer RHC, Agrawal A, et al. (2021): Interpretation of psychiatric genome-wide association studies with multispecies heterogeneous functional genomic data integration. Neuropsychopharmacology 46:86–97.

11. Subramanian I, Verma S, Kumar S, Jere A, Anamika K (2020): Multiomics data integration, interpretation, and its application. Bioinform Biol Insights 14:1177932219899051.

12. Bracken MB, DeWan A, Hoh J (2016): Genome-wide association studies. In: Molecular Epidemiology: Applications in Cancer and Other Human Diseases. Boca Raton: CRC Press, 239–252.

13. Hirschhorn JN, Daly MJ (2005): Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95–108.

14. Buniello A, Macarthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. (2019): The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics. Nucleic Acids Res 47:D1005–D1012.

15. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, Yang J (2017): 10 years of GWAS discovery: Biology, function, and translation. Am J Hum Genet 101:5–22.

16. Wray NR, Lee SH, Mehta D, Vinkhuyzen AA, Dudbridge F, Middeldorp CM (2014): Research Review: Polygenic methods and their application to psychiatric traits. J Child Psychol Psychiatry 55:1068–1087.

17. Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke S, Neale BM, Corvin A, Walters JTR, Farh K-H, Holmans PA, et al. (2014): iological insights from 108 schizophrenia-associated genetic loci. Nat 511:421–427.

18. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. (2019): Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci 22:343–352.

19. Yu D, Sul JH, Tsetsos F, Nawaz MS, Huang AY, Zelaya I, et al. (2019): Interrogating the genetic determinants of Tourette's syndrome and other Tic disorders through genome-wide association studies. Am J Psychiatry 176:217–227.

20. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, et al. (2019): Identification of common genetic risk variants for autism spectrum disorder. Nat Genet 51:431–444.

21. Nievergelt CM, Maihofer AX, Klengel T, Atkinson EG, Chen CY, Choi KW, et al. (2019): International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. Nat Commun 10:4558.

22. Watson HJ, Yilmaz Z, Thornton LM, Hübel C, Coleman JRI, Gaspar HA, et al. (2019): Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. Nat Genet 51:1207–1214.

23. Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. (2019): Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk [published correction appears in Nat Genet. 2020; 52:354. Nat Genet 51:404–413.

24. Pasman JA, Verweij KJH, Gerring Z, Stringer S, Sanchez-Roige S, Treur JL, et al. (2018): GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia [published correction appears in Nat Neurosci. 2019; 22:1196]. Nat Neurosci 21:1161–1170.

25. Walters RK, Polimanti R, Johnson EC, McClintick JN, Adams MJ, Adkins AE, et al. (2018): Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. Nat Neurosci 21:1656–1669.

26. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, et al. (2019): Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet 51:63–75.

27. Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium (2018): Genomic dissection of bipolar disorder and schizophrenia, including 28 subphenotypes. Cell 173:1705–1715.e16.

28. Ochoa D, Karim M, Ghoussaini M, Hulcoop DG, McDonagh EM, Dunham I (2022): Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs. Nat Rev Drug Discov 21:551.

29. Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, et al. (2015): The support of human genetic evidence for approved drug indications. Nat Genet 47:856–860.

30. Dudbridge F (2013): Power and predictive accuracy of polygenic risk scores [published correction appears in PLoS Genet. 2013; 9. doi: 10. 1371/annotation/b91ba224-10be-409d-93f4-7423d502cba0]. PLoS Genet 9:e1003348.

31. Wray NR, Lin T, Austin J, McGrath JJ, Hickie IB, Murray GK, et al. (2021): From basic science to clinical application of polygenic risk scores: A Primer. JAMA Psychiatry 78:101–109.

32. Choi SW, Mak TSH, O'Reilly PF (2020): Tutorial: A guide to performing polygenic risk score analyses. Nat Protoc 15:2759–2772.

33. Chatterjee N, Wheeler B, Sampson J, Hartge P, Chanock SJ, Park JH (2013): Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. Nat Genet 45:400–405e1–3.

34. Torkamani A, Wineinger NE, Topol EJ (2018): The personal and clinical utility of polygenic risk scores. Nat Rev Genet 19:581–590.

35. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. (2018): Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 50:1112–1121.

36. Ikeda M, Saito T, Kanazawa T, Iwata N (2021): Polygenic risk score as clinical utility in psychiatry: A clinical viewpoint. J Hum Genet 66:53–60.

37. Bevilacqua L, Doly S, Kaprio J, Yuan Q, Tikkanen R, Paunio T, et al. (2010): A population-specific HTR2B stop codon predisposes to severe impulsivity [published correction appears in Nature. 2011;470:424]. Nature 468:1061–1066.

38. Caron NS, Wright GE, Hayden MR Huntington Disease. In: Adam MP, Mirzaa GM, Pagon RA, et al., editors. GeneReviews. Seattle (WA): University of Washington, Seattle; October 23, 1998 (updated 2020 Jun 11).

39. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, et al. (1993): Apolipoprotein E: High-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. Proc Natl Acad Sci U S A 90:1977–1981.

40. Gratten J, Wray NR, Keller MC, Visscher PM (2014): Large-scale genomics unveils the genetic architecture of psychiatric disorders. Nat Neurosci 17:782–790.

41. Boyle EA, Li YI, Pritchard JK (2017): An expanded view of complex traits: From polygenic to omnigenic. Cell 169:1177–1186.

42. Shi H, Kichaev G, Pasaniuc B (2016): Contrasting the genetic architecture of 30 complex traits from summary association data. Am J Hum Genet 99:139–153.

43. Zhang Y, Qi G, Park JH, Chatterjee N (2018): Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. Nat Genet 50:1318–1326.

44. International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, et al. (2009): Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460:748–752.

45. Wray NR, Goddard ME, Visscher PM (2008): Prediction of individual genetic risk of complex disease. Curr Opin Genet Dev 18:257–263.

46. Chatterjee N, Shi J, García-Closas M (2016): Developing and evaluating polygenic risk prediction models for stratified disease prevention. Nat Rev Genet 17:392–406.

47. Euesden J, Lewis CM, O'Reilly PF (2015): PRSice: Polygenic Risk Score software. Bioinformatics 31:1466–1468.

48. Batra A, Chen LM, Wang Z, Parent C, Pokhvisneva I, Patel S, et al. (2021): Early life adversity and polygenic risk for high fasting insulin are associated with childhood impulsivity. Front Neurosci 15:704785.

49. Chen LM, Tollenaar MS, Hari Dass SA, Bouvette-Turcot AA, Pokhvisneva I, Gaudreau H, et al. (2020): Maternal antenatal depression and child mental health: Moderation by genomic risk for attention-deficit/hyperactivity disorder. Dev Psychopathol 32:1810–1821.

50. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW (2019): Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun 10:1776.

51. Vilhjálmsson BJ, Yang J, Finucane HK, Gusev A, Lindström S, Ripke S, et al. (2015): Modeling linkage disequilibrium increases accuracy of polygenic risk scores. Am J Hum Genet 97:576–592.

52. Ni G, Zeng J, Revez JA, Wang Y, Zheng Z, Ge T, et al. (2021): A comparison of ten polygenic score methods for psychiatric disorders applied across multiple cohorts. Biol Psychiatry 90:611–620.

53. Lloyd-Jones LR, Zeng J, Sidorenko J, Yengo L, Moser G, Kemper KE, et al. (2019): Improved polygenic prediction by Bayesian multiple regression on summary statistics. Nat Commun 10:5086.

54. Ripke S, PGC SCZ WORKGROUP (2019): GWAS with over 70.000 cases and 100,000 controls. Eur Neuropsychopharmacol 29:S814.

55. Desikan RS, Fan CC, Wang Y, Schork AJ, Cabral HJ, Cupples LA, et al. (2017): Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score [published correction appears in PLoS Med. 2017 14:e1002289]. PLoS Med 14:e1002258.

56. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. (2013): Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's Disease. Nat Genet 45:1452–1458.

57. Peterson RE, Kuchenbaecker K, Walters RK, Chen CY, Popejoy AB, Periyasamy S, et al. (2019): Genome-wide association studies in ancestrally diverse populations: Opportunities, methods, pitfalls, and recommendations. Cell 179:589–603.

58. Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, et al. (2017): Human demographic history impacts genetic risk prediction across diverse populations [published correction appears in Am J Hum Genet. 2020; 107:788-789]. Am J Hum Genet 100:635–649.

59. Perkins DO, Olde Loohuis L, Barbee J, Ford J, Jeffries CD, Addington J, et al. (2020): Polygenic risk score contribution to psychosis prediction in a target population of persons at clinical high risk. Am J Psychiatry 177:155–163.

60. Mega JL, Stitziel NO, Smith JG, Chasman DI, Caulfield M, Devlin JJ, et al. (2015): Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: An analysis of primary and secondary prevention trials. Lancet 385:2264–2271.

61. Inouye M, Abraham G, Nelson CP, Wood AM, Sweeting MJ, Dudbridge F, et al. (2018): Genomic risk prediction of coronary artery disease in 480,000 adults: Implications for primary prevention. J Am Coll Cardiol 72:1883–1893.

62. Maier RM, Visscher PM, Robinson MR, Wray NR (2018): Embracing polygenicity: A review of methods and tools for psychiatric genetics research. Psychol Med 48:1055–1067.

63. Gusev A, Lee SH, Trynka G, Finucane H, Vilhjálmsson BJ, Xu H, et al. (2014): Partitioning heritability of regulatory and cell-type specific variants across 11 common diseases. Am J Hum Genet 95:535–552.

64. Powell SK, O'Shea C, Brennand KJ, Akbarian S (2021): Parsing the functional impact of noncoding genetic variants in the brain epigenome. Biol Psychiatry 89:65–75.

65. Hernandez LM, Kim M, Hoftman GD, Haney JR, de la Torre-Ubieta L, Pasaniuc B, et al. (2021): Transcriptomic insight into the polygenic mechanisms underlying psychiatric disorders. Biol Psychiatry 89:54–64.

66. Uffelmann E, Posthuma D (2021): Emerging methods and resources for biological interrogation of neuropsychiatric polygenic signal. Biol Psychiatry 89:41–53.

67. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, et al. (2016): Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet 48:245–252.

68. Girgenti MJ, Wang J, Ji D, Cruz DA, Traumatic Stress Brain Research Group, Stein MB, et al. (2021): Transcriptomic organization of the human brain in post-traumatic stress disorder. Nat Neurosci 24:24–33.

69. Zhang Y, Quick C, Yu K, Barbeira A, GTEx Consortium, Luca F, et al. (2020): PTWAS: Investigating tissue-relevant causal molecular mechanisms of complex traits using probabilistic TWAS analysis. Genome Biol 21:232.

70. Agarwal V, Shendure J (2020): Predicting mRNA abundance directly from genomic sequence using deep convolutional neural networks. Cell Rep 31:107663.

71. Zhou J, Theesfeld CL, Yao K, Chen KM, Wong AK, Troyanskaya OG (2018): Deep learning sequence-based ab initio prediction of variant effects on expression and disease risk. Nat Genet 50:1171–1179.

72. Alpay BA, Demetci P, Istrail S, Aguiar D (2020): Combinatorial and statistical prediction of gene expression from haplotype sequence. Bioinformatics 36(Suppl_1):i194–202.

73. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. (2015): A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 47:1091–1098.

74. GTEx Consortium (2020): The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science 369:1318–1330.

75. Home, Predict DB. Internet. [cited Dec 27 2021]. Available at: https://predictdb.org/. [Accessed December 27, 2021].

76. Barbeira AN, Pividori M, Zheng J, Wheeler HE, Nicolae DL, Im HK (2019): Integrating predicted transcriptome from multiple tissues improves association detection. PLoS Genet 15:e1007889.

77. Shah KP, Wheeler HE, Gamazon ER, Nicolae DL, Cox NJ, Im HK (2016): Genetic predictors of gene expression associated with risk of bipolar disorder. bioRxiv. https://doi.org/10.1101/043752.

78. Hartl CL, Ramaswami G, Pembroke WG, Muller S, Pintacuda G, Saha A, et al. (2021): Coexpression network architecture reveals the brain-wide and multiregional basis of disease susceptibility. Nat Neurosci 24:1313–1323.

79. Johnson MR, Shkura K, Langley SR, Delahaye-Duriez A, Srivastava P, Hill WD, et al. (2016): Systems genetics identifies a convergent gene network for cognition and neurodevelopmental disease. Nat Neurosci 19:223–232.

80. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB (2013): Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet 9:e1003709.

81. Zhu X, Need AC, Petrovski S, Goldstein DB (2014): One gene, many neuropsychiatric disorders: Lessons from Mendelian diseases. Nat Neurosci 17:773–781.

82. Pergola G, Di Carlo P, D'Ambrosio E, Gelao B, Fazio L, Papalino M, et al. (2017): DRD2 co-expression network and a related polygenic index predict imaging, behavioral and clinical phenotypes linked to schizophrenia. Transl Psychiatry 7:e1006.

83. Gerring ZF, Gamazon ER, Derks EM, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium (2019): A gene co-expression network-based analysis of multiple brain tissues reveals novel genes and molecular pathways underlying major depression. PLoS Genet 15:e1008245.

84. GTEx Consortium, Ardlie KG, DeLuca DS, Segrè AV, Sullivan TJ, Young TR, et al. (2015): Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science 348:648–660.

85. Dalmaz C, Barth B, Pokhvisneva I, Wang Z, Patel S, Quillfeldt JA, et al. (2021): Prefrontal cortex VAMP1 gene network moderates the effect of the early environment on cognitive flexibility in children. Neurobiol Learn Mem 185:107509.

86. de Mendonça Filho EJ, Barth B, Bandeira DR, de Lima RMS, Arcego DM, Dalmaz C, et al. (2021): Cognitive development and brain gray matter susceptibility to prenatal adversities: Moderation by the prefrontal cortex brain-derived neurotrophic factor gene coexpression network. Front Neurosci 15:744743.

87. Hari Dass SA, McCracken K, Pokhvisneva I, Chen LM, Garg E, Nguyen TTT, et al. (2019): A biologically informed polygenic score identifies endophenotypes and clinical conditions associated with the insulin receptor function on specific brain regions. EBioMedicine 42:188–202.

88. Silveira PP, Pokhvisneva I, Parent C, Cai S, Rema ASS, Broekman BFP, et al. (2017): 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.

89. Miguel PM, Pereira LO, Barth B, de Mendonça Filho EJ, Pokhvisneva I, Nguyen TTT, et al. (2019): 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.

90. Restrepo-Lozano JM, Pokhvisneva I, Wang Z, Patel S, Meaney MJ, Silveira PP, et al. (2022): Corticolimbic DCC gene co-expression networks as predictors of impulsivity in children. Mol Psychiatry 27:2742–2750.

91. McGill MG, Pokhvisneva I, Clappison AS, McEwen LM, Beijers R, Tollenaar MS, et al. (2022): Maternal prenatal anxiety and the fetal origins of epigenetic aging. Biol Psychiatry 91:303–312.

92. Belsky DW, Domingue BW, Wedow R, Arseneault L, Boardman JD, Caspi A, et al. (2018): Genetic analysis of social-class mobility in five longitudinal studies [published correction appears in Proc Natl Acad Sci U S A. 2018; 115:E10998]. Proc Natl Acad Sci U S A 115:E7275–E7284.

93. Lupien SJ, McEwen BS, Gunnar MR, Heim C (2009): Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nat Rev Neurosci 10:434–445.

94. Provençal N, Arloth J,Cattaneo A, Anacker C, CattaneN,Wiechmann T, et al. (2020): Glucocorticoid exposure during hippocampal neurogenesis primes future stress response by inducing changes in DNA methylation. Proc Natl Acad Sci U S A 117:23280–23285.

95. Folkersen L, Pain O, Ingason A, Werge T, Lewis CM, Austin J (2020): Impute.me: An open-source, non-profit tool for using data from direct-to-consumer genetic testing to calculate and interpret polygenic risk scores. Front Genet 11:578.

96. Watson J, Jones HE, Banks J, Whiting P, Salisbury C, Hamilton W (2019): Use of multiple inflammatory marker tests in primary care: Using clinical practice research datalink to evaluate accuracy. Br J Gen Pract 69:e462–e469.

97. de Leeuw CA, Mooij JM, Heskes T, Posthuma D (2015): MAGMA: Generalized gene-set analysis of GWAS data. PLoS Comput Biol 11: e1004219.

98. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium, et al. (2015): LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47:291–295.

99. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. (2016): Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 48:481–487.

100. Mak TSH, Porsch RM, Choi SW, Zhou X, Sham PC (2017): Polygenic scores via penalized regression on summary statistics. Genet Epidemiol 41:469–480.

101. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, Haycock PC, et al. (2017): LD Hub: A centralized database and web interface to perform LD score regression that maximizes the potential

of summary level GWAS data for SNP heritability and genetic correlation analysis. Bioinformatics 33:272–279.

102. Hu Y, Lu Q, Powles R, Yao X, Yang C, Fang F, et al. (2017): Leveraging functional annotations in genetic risk prediction for human complex diseases. PLoS Comput Biol 13:e1005589.

103. Robinson MR, Kleinman A, Graff M, Vinkhuyzen AAE, Couper D, Miller MB, et al. (2017): Genetic evidence of assortative mating in humans. Nat Hum Behav 1:0016.

104. Newcombe PJ, Nelson CP, Samani NJ, Dudbridge F (2019): A flexible and parallelizable approach to genome-wide polygenic risk scores. Genet Epidemiol 43:730–741.

105. Márquez-Luna C, Gazal S, Loh PR, Kim SS, Furlotte N, Auton A, et al. (2021): Incorporating functional priors improves polygenic prediction accuracy in UK Biobank and 23andMe data sets. Nat Commun 12:6052.

106. Privé F, Arbel J, Vilhjálmsson BJ (2021): LDpred2: Better, faster, stronger. Bioinformatics 36(22–23):5424–5431.

107. Liang Y, Pividori M, Manichaikul A, Palmer AA, Cox NJ, Wheeler HE, et al. (2022): Polygenic transcriptome risk scores (PTRS) can improve portability of polygenic risk scores across ancestries. Genome Biol 23:23.

Chapter V: DISCUSSION AND CONCLUSIONS

The purpose of my doctoral thesis was to venture into an exploration of molecular networks that contribute to brain corticolimbic development, and how variations in these neurodevelopmental processes can ultimately modify the development of inhibitory control behaviors in both mice and humans. I conducted two main studies: The first study, using mice, sought to dissect the molecular, behavioral, and structural consequences of environmentally-induced disturbances to the Netrin-1/DCC signaling system, which has been shown to be important for the development of inhibitory control behaviors and for brain corticolimbic maturation. In the second study, a targeted analysis of a novel functional genomic signal- based on corticolimbic-specific *DCC* gene co-expression networks- was carried out across independent and ethnically-diverse human birth cohorts, mapping the association between this novel biological marker and measurements of inhibitory control behaviors.

In the first phase of this research, by examining the long-term structural and behavioral effects of early-life exposure to therapeutic-like and recreational-like doses of amphetamine in mice, I found significant insights into how this experience can modify, in a remarkable dose-dependent fashion, mesocorticolimbic dopamine development and inhibitory control behaviors in adulthood. This study also elucidates the molecular signatures induced by the different amphetamine doses, mainly showing that while therapeutic-like doses increase DCC expression, recreational-like doses induce a downregulation of DCC and Netrin-1.

Stimulant medications, notably including amphetamines, represent one of the most prescribed pharmaceutical compounds for Canadian adolescents. Data from the Canadian Centre on Substance Use and Addiction (2019) showed that approximately 5% of Canadians aged 15-19 utilized *prescription stimulants*, while an additional 3% engaged in *recreational usage* of these substances. Given this well-known exposure, it is important to delineate- and in some sense, to better understand- what are the

long-term implications of exposure to stimulants during childhood and adolescence, on neurodevelopmental and cognitive trajectories. Notably, pharmacological treatments are currently designed with a focus to mitigate symptomatic manifestations rather than address underlying pathophysiology. This could be the results of a gap in our understanding of the early indicators and etiology of psychiatric conditions and their development. A recent investigation attempted to link polygenic signals from Genome-Wide Association Studies (GWASs) for psychiatric conditions with the corresponding molecular pathways targeted by different pharmacological interventions [89], including the use of amphetamines and other related phenethylamine-derived compounds. Employing a multi-modal analytical approach—encompassing protein-protein interactions, gene activity levels, and chromatin interactions—this study highlighted a noticeable incongruity between genetic factors implicated in psychiatric disorders and the biological pathways targeted by current drug treatments. More precisely, risk-conferring genes identified in GWASs predominantly relate to neurodevelopmental processes, in contrast to the synaptic signaling pathways that current pharmacological treatments primarily target.

In this context, mapping behavioral constructs into a set of neurobiological functions allowed me to explore and characterize the developmental consequences of psychostimulant-induced alterations to the Netrin-1/DCC system and corticolimbic function. Ironically, this molecular signaling pathway has been notoriously important in GWASs for psychiatric phenotypes [79] and perhaps it plays a critical role in mental health trajectories across the lifespan. I discovered that amphetamines have a dose-dependent effect on the Netrin-1/DCC signaling pathway in rodents, and that this effect is accompanied by a consequent behavioral effect in the Go/No-Go task, showing that therapeutic-like doses of amphetamine during sensitive neurodevelopmental periods ameliorate the overall performance of mice in the task, long after treatment cessation, when tested in adulthood. Furthermore, the analysis implemented for the description of neuroanatomical features in corticolimbic development revealed no apparent changes following early therapeutic psychostimulant treatment, measured by the volume and

the density of dopamine pre-synaptic sites observed in mesocorticolimbic dopamine targets in adulthood. The absence of evidence for structural alterations was initially puzzling, given the observed changes in inhibitory control. More refined investigations are needed to understand precisely if and how early-life amphetamine exposure changes dopamine dynamics in mesocorticolimbic dopamine targets.

Growing computational and genetic evidence underscores the role of *DCC* and Netrin-1 in the onset of adolescent psychiatric disorders, many of which are also marked by deficits in inhibitory control [61,78]. Our rodent studies demonstrate that disruptions in corticolimbic networks during critical neurodevelopmental periods associate with later-life changes in cognitive control capacity [90]. We observed that repeated, dose-sensitive adolescent exposure to psychostimulants can modulate *Dcc* gene expression in VTA dopamine neurons, influencing a cascade of neurodevelopmental processes that play an important role in adult corticolimbic structure and function. So far, VTA dopamine inputs to the PFC, the OFC, and the NAcc, have shown *sensitivity* to amphetamine-induced modulations in Netrin-1/DCC-mediated signaling, ultimately altering adult inhibitory control behaviors in mice. This empirical evidence aligns with clinical data suggesting that individuals with ADHD who have been medicated exhibit a regression to baseline risk levels for developing substance use disorders (SUDs), in contrast to their non-medicated ADHD counterparts. Importantly, it still needs to be addressed whether this effect is particular to amphetamines or if it generalizes to other psychostimulant substances like methylphenidate or cocaine.

In the second phase of my thesis, I shifted the focus to human subjects while still relying on mouse models for the generation of co-expression networks to investigate a novel type of polygenic signal. I calculated an expression-based polygenic score (ePRS) built upon corticolimbic-specific *DCC* gene co-expression networks (identified first in mice, and then enriched for human brain expression in early developmental stages), identifying unique associations with impulsivity-related phenotypes in children

This effort has not only further validated the methodology used to construct expression-based polygenic scores but also presented a pioneering approach to recognizing early markers of vulnerability in inhibitory control deficits. By going back and forth between rodents and humans, I gather mechanistic insight that corroborates the potential role of DCC in mesocorticolimbic dopamine terminals for the development of impulsivity in humans.

These results not only solidify the role of the corticolimbic DCC-gene networks for proper brain corticolimbic development and inhibitory control capacity, but also pave the way for implementing this approach in other investigations, considering that initially I rely on a highly sensitive and context-specific measurement of transcript abundance in mouse tissue. As this study was guided by the previous findings in rodents linking variations in Dcc expression to changes in impulse control and in mesocorticolimbic dopamine axon targeting, it seemed logical to begin using mice and move on to humans by applying a filter for genes that are overexpressed in humans in early life when compared to adulthood, to reflect time-sensitive maturational processes in the human brain. This process narrowed down and identified networks that contain 152 and 74 genes for the PFC and NAcc networks, respectively.

The enrichment analyses that followed allowed me to establish the biological context in which the genes within the co-expression networks operate, but it was still important to assess the co-expression of the genes in human data. I used the BrainSpan dataset containing postmortem PFC and NAcc to confirm that the genes in the PFC and NAcc networks do co-express (see Chapter III, Figure 4). These analyses, in combination with the functional enrichment analyses (shown in Chapter III, Figure 3) support the conclusion that the resulting groups of genes are in fact coherent and tissue-specific functional networks.

The temporal convergence of gene expression, particularly those genes implicated in neuropsychiatric disorders, has been shown to be especially salient before and during the onset of these disorders[91]. Our results align, to an interesting degree, with these findings, revealing distinct periods of enrichment for gene expression within both corticolimbic-specific gene networks, notably during specific pre-natal and post-natal phases. This includes the high specificity in spatiotemporal expression of the NAcc network in cortical neurons during late childhood and adolescence, which indicates that those genes (specifically: KCNS2, GRIN1, SYP, SCN8A, DGKB, PPP4R4, SLC6A17, HOMER1, ARHGDIG, and RAB27B) are expressed and enriched in the cortex during later developmental periods. As previously acknowledged, the functional role of *DCC* varies depending on the developmental stage. The re-emergence of an enriched *DCC*-related group of genes during late childhood and adolescence suggests that any perturbations in *DCC* function or expression could have significant ramifications on the development of synaptic connectivity and function within the PFC. Finally, a brief overview of the cellular components where the above-mentioned genes are found indicates enrichment in excitatory synapses, terminal boutons, and exocytic versicles, which may explain why these processes occur earlier in the NAcc compared to the PFC, as this aligns well with current knowledge about the temporal differences between mesolimbic and mesocortical development.

A limitation from this study is the lack of integration of SNPs outside the genomic regions ($\pm$ 500 KB) of the corticolimbic-specific DCC-gene co-expression networks, in particular of variants that are associated with other forms of gene expression regulation (e.g., transcription factors, promoter regions, chromatin modifications and trans-eQTLs), and it would be both interesting and informative whether the integration of such signals can change the performance of the scores. This will be investigated in future studies, and is discussed in the review paper presented in Chapter 4.

Finally, I consider important to point out that "deleted in colorectal cancer" (DCC) gets its name from its initial discovery. It was identified as a gene that is frequently absent or deleted in colorectal cancer tissues. The deletion of specific regions of DNA is a common event in the development of many cancers. When researchers identified this gene, they observed that it was often deleted in tumor samples from patients with colorectal cancer, hence they named it "deleted in colorectal cancer". However, DCC was subsequently found to encode a netrin-1 receptor and that their interaction played a crucial role in axon guidance during neural development. This discovery indicated that DCC might have functions beyond its role in colorectal cancer, at least in the nervous system. The naming convention of such genes often stems from their initial discovery or the context in which they were first identified, even if subsequent research reveals more about their broader biological roles. Making sense of these overlapping sets of molecules interacting in different tissues is a task that can only be understood via systematic probing of the molecular networks, and how disrupting the cellular milieu can have consequences observable at different levels of analysis.

Concluding remarks and future directions

Genome-wide association studies (GWAS) on human behavioral traits have become increasingly sophisticated, yielding numerous polygenic signals that not only hold substantial predictive power but, following functional post-GWAS analyses, also offer valuable biological insights. However, interpreting these genetic associations for behavioral traits is complex because, unlike physical traits which are more proximally influenced by biology, behavioral traits are more distal and subjected to multiple layers of regulation, making their genetic underpinnings less straightforward to decipher. The outcome of GWAS in the domain of human behavior is likely a composite of aggregated signals from diverse origins. As the sample sizes in these studies continue to expand, thereby increasing the power to detect weaker polygenic signals, methodological advancements such as the one presented in this work aim to delineate

the sources of these signals more precisely. Notably, the major contributory elements to the polygenic landscape include not only direct genetic influences on the trait under investigation but also genetic associations with correlated traits and environmental interactions.

In the realm of psychiatric genetics, to my knowledge and understanding, the convention of modeling SNPs through additive models has proven practical for deconstructing the complex architecture of mental disorders. While this approach has enabled important strides in identifying risk variants, it can also inadvertently obscure the potential richness of epistatic interactions—instances where the combined effect of multiple loci diverges from the sum of their individual effects. Therefore, in the current body of work, by weighing SNPs within an additive-only context, we could be inadvertently pruning the full spectrum of complex biological interactions that may contribute to inhibitory control behaviors, as the specific combination of variants would be what matter most. This methodological limitation hinders our understanding of the intricacies of gene-gene interactions and their subsequent impact on phenotypic expression. Such interactions, potentially important for better predictive models, remain largely unexplored, mainly because of the computational requirements for the deployment of such models. Consequently, a more nuanced approach that incorporates epistatic models alongside additive effects could help us further advance our current understanding of the genetic mechanisms underpinning complex traits[92].

Finally, the development of frameworks that link genotypes to cellular programs represents a paradigmatic shift in psychiatric research, potentially revolutionizing our capacity for early identification of vulnerability markers for psychiatric disorders. Traditional genomic analyses often operate in a manner that is somewhat disconnected from the underlying biological processes, offering us statistical associations but not mechanistic understanding. By contrast, a genotype-to-cellular-program framework integrates data across multiple biological layers—from genetic variants to gene expression profiles to

cellular pathways—thereby providing a more coherent and biologically informed landscape. This enriched viewpoint can help identify novel markers or sets of markers that are indicative of predisposition to psychiatric conditions at a cellular or system-level, rather than merely at the level of isolated genetic variants. Such markers could be invaluable for early interventions, allowing for timely and targeted therapeutic strategies that are rooted in a deeper understanding of disease pathophysiology. Consequently, this multi-layered approach, similar to the one presented in this work, could be instrumental in bridging the extant gap between genetic risk and the clinical manifestation of psychiatric disorders, offering a more predictive and preventative model of mental health care.

References for Non-Manuscript Chapters

1. Lim, K-L, P. Jacobs, A. Ohinmaa, D. Schopflocher, and C. S. D. Volume 28, No. 3, 2008 /Chronic Diseases in Canada. *Chronic Dis. Canada CS Dewa* **28**, 92–98 (2008).

2. Baranne, M. L. & Falissard, B. Global burden of mental disorders among children aged 5-14 years. *Child Adolesc. Psychiatry Ment. Health* **12**, 1–9 (2018).

3. Allsopp, K., Read, J., Corcoran, R. & Kinderman, P. Heterogeneity in psychiatric diagnostic classification. *Psychiatry Res.* **279**, 15–22 (2019).

4. Fanous, A. H. & Kendler, K. S. Genetic heterogeneity, modifier genes, and quantitative phenotypes in psychiatric illness: Searching for a framework. *Mol. Psychiatry* **10**, 6–13 (2005).

5. Glahn, D. C. *et al.* Arguments for the sake of endophenotypes: Examining common misconceptions about the use of endophenotypes in psychiatric genetics. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **165**, 122–130 (2014).

6. Heckers, S., Engstrom, E. J. & Kendler, K. S. "Manifestations of insanity": Kraepelin's final views on psychiatric nosology in their historical context. *Mol. Psychiatry* **27**, 328–334 (2022).

7. Cuthbert, B. N. The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry* **13**, 28–35 (2014).

8. Patrick, C. J. & Hajcak, G. RDoC: Translating promise into progress. *Psychophysiology* **53**, 415–424 (2016).

9. Morris, S. E. & Cuthbert, B. N. Research domain criteria: Cognitive systems, neural circuits, and

dimensions of behavior. *Dialogues Clin. Neurosci.* **14**, 29–37 (2012).

10. Sanislow, C. A. *et al.* Developing constructs for psychopathology research: Research domain criteria. *J. Abnorm. Psychol.* **119**, 631–639 (2010).

11. Casey, B. J., Oliveri, M. E. & Insel, T. A neurodevelopmental perspective on the research domain criteria (RDoC) framework. *Biol. Psychiatry* **76**, 350–353 (2014).

12. Zhou, Y., Song, H. & Ming, G. Genetics of human brain development. *Nat. Rev. Genet.* (2023) doi:10.1038/s41576-023-00626-5.

13. Li, M. *et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science (80-.).* **362**, (2018).

14. Yang, S. & Tseng, K. Y. Maturation of Corticolimbic Functional Connectivity During Sensitive Periods of Brain Development. *Curr. Top. Behav. Neurosci.* **53**, 37–53 (2022).

15. Watt, M. J., Weber, M. A., Davies, S. R. & Forster, G. L. Impact of juvenile chronic stress on adult cortico-accumbal function: Implications for cognition and addiction. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **79**, 136–154 (2017).

16. Kunitoki, K. *et al.* Youth Team Sports Participation Associates With Reduced Dimensional Psychopathology Through Interaction With Biological Risk Factors. *Biol. Psychiatry Glob. Open Sci.* 1–9 (2023) doi:10.1016/j.bpsgos.2023.02.001.

17. Kolb, B. & Gibb, R. Brain Plasticity and Behaviour in the Developing Brain. *J. Can. Acad. Child Adolesc. Psychiatry* **20**, 265–276 (2011).

18. Knudsen, E. I. Sensitive periods in the development of the brain and behavior. *J. Cogn. Neurosci.*

16, 1412–1425 (2004).

19. Hensch, T. K. Critical period regulation. *Annu. Rev. Neurosci.* **27**, 549–579 (2004).

20. Verdejo-Garcia, A. & Albein-Urios, N. Special issue on vulnerabilities to substance abuse impulsivity traits and neurocognitive mechanisms conferring vulnerability to substance use disorders. *Neuropharmacology* **183**, 108402 (2021).

21. Ravindranath, O., Calabro, F. J., Foran, W. & Luna, B. Pubertal development underlies optimization of inhibitory control through specialization of ventrolateral prefrontal cortex. *Dev. Cogn. Neurosci.* **58**, 101162 (2022).

22. Naneix, F., Marchand, A. R., Di Scala, G., Pape, J. R. & Coutureau, E. Parallel maturation of goal-directed behavior and dopaminergic systems during adolescence. *J. Neurosci.* **32**, 16223–16232 (2012).

23. Casey, B. J., Getz, S. & Galvan, A. The adolescent brain. *Dev. Rev.* **28**, 62–77 (2008).

24. Barkley, R. A. Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychol. Bull.* **121**, 65–94 (1997).

25. Nigg, J. T. Is ADHD a disinhibitory disorder? *Psychol. Bull.* **127**, 571–598 (2001).

26. Chamberlain, S. R., Blackwell, A. D., Fineberg, N. A., Robbins, T. W. & Sahakian, B. J. The neuropsychology of obsessive compulsive disorder: The importance of failures in cognitive and behavioural inhibition as candidate endophenotypic markers. *Neurosci. Biobehav. Rev.* **29**, 399–419 (2005).

27. Smith, J. L., Mattick, R. P., Jamadar, S. D. & Iredale, J. M. Deficits in behavioural inhibition in

substance abuse and addiction: A meta-analysis. *Drug Alcohol Depend.* **145**, 1–33 (2014).

28. Jentsch, J. D. & Taylor, J. R. Impulsivity resulting from frontostriatal dysfunction in drug abuse: Implications for the control of behavior by reward-related stimuli. *Psychopharmacology (Berl).* **146**, 373–390 (1999).

29. Iacono, W. G., Malone, S. M. & McGue, M. Behavioral disinhibition and the development of early-onset addiction: Common and specific influences. *Annu. Rev. Clin. Psychol.* **4**, 325–348 (2008).

30. Verbruggen, F. & Logan, G. D. Response inhibition in the stop-signal paradigm. *Trends Cogn. Sci.* **12**, 418–424 (2008).

31. Logan, Gordon D., and W. B. C. On the ability to inhibit thought and action: A theory of an act of control. *Psychol. Rev.* **91**, 295 (1984).

32. Eagle, D. M. & Robbins, T. W. Inhibitory Control in Rats Performing a Stop-Signal Reaction-Time Task: Effects of Lesions of the Medial Striatum and d-Amphetamine. *Behav. Neurosci.* **117**, 1302–1317 (2003).

33. Bari, A., Eagle, D. M., Mar, A. C., Robinson, E. S. J. & Robbins, T. W. Dissociable effects of noradrenaline, dopamine, and serotonin uptake blockade on stop task performance in rats. *Psychopharmacology (Berl).* **205**, 273–283 (2009).

34. Aron, A. R., Robbins, T. W. & Poldrack, R. A. Inhibition and the right inferior frontal cortex: One decade on. *Trends Cogn. Sci.* **18**, 177–185 (2014).

35. Bari, A. & Robbins, T. W. Inhibition and impulsivity: Behavioral and neural basis of response control. *Prog. Neurobiol.* **108**, 44–79 (2013).

36. Haber, S. N. Corticostriatal circuitry. *Dialogues Clin. Neurosci.* **18**, 7–21 (2016).

37. Volkow, N. D., Fowler, J. S., Wang, G. J., Baler, R. & Telang, F. Imaging dopamine's role in drug abuse and addiction. *Neuropharmacology* **56**, 3–8 (2009).

38. Sesack, S. R. & Grace, A. A. Cortico-basal ganglia reward network: Microcircuitry. *Neuropsychopharmacology* **35**, 27–47 (2010).

39. Fineberg, N. A. *et al.* Probing compulsive and impulsive behaviors, from animal models to endophenotypes: A narrative review. *Neuropsychopharmacology* **35**, 591–604 (2010).

40. Le Moal, M. & Simon, H. *Mesocorticolimbic dopaminergic network: Functional and regulatory roles*. *Physiological Reviews* vol. 71 (1991).

41. Van den Heuvel, D. M. A. & Pasterkamp, R. J. Getting connected in the dopamine system. *Prog. Neurobiol.* **85**, 75–93 (2008).

42. Dalley, J. W., Everitt, B. J. & Robbins, T. W. Impulsivity, Compulsivity, and Top-Down Cognitive Control. *Neuron* **69**, 680–694 (2011).

43. Reynolds, L. M. *et al.* DCC Receptors Drive Prefrontal Cortex Maturation by Determining Dopamine Axon Targeting in Adolescence. *Biol. Psychiatry* **83**, 181–192 (2018).

44. Volkow, N. D., Fowler, J. S., Wang, G., Ding, Y. & Gatley, S. J. R ole of dopamine in the therapeutic and reinforcing effects of methylphenidate in humans : results from imaging studies. **12**, 557–566 (2002).

45. Howes, O. D. & Kapur, S. The dopamine hypothesis of schizophrenia: Version III - The final common pathway. *Schizophr. Bull.* **35**, 549–562 (2009).

46. Rubia, K. Functional brain imaging across development. *Eur. Child Adolesc. Psychiatry* **22**, 719–731 (2013).

47. Luna, B., Marek, S., Larsen, B., Tervo-Clemmens, B. & Chahal, R. An Integrative Model of the Maturation of Cognitive Control. *Annu. Rev. Neurosci.* **38**, 151–170 (2015).

48. Reynolds, L. M. & Flores, C. Adolescent dopamine development. in *Diagnosis, Management and Modeling of Neurodevelopmental Disorders* 295–304 (Elsevier, 2021). doi:10.1016/B978-0-12-817988-8.00026-9.

49. Kalsbeek, A., Voorn, P., Buijs, R. M., Pool, C. W. & Uylings, H. B. M. Development of the dopaminergic innervation in the prefrontal cortex of the rat. *J. Comp. Neurol.* **269**, 58–72 (1988).

50. Teicher, M. H., Andersen, S. L. & Hostetter, J. C. Evidence for dopamine receptor pruning between adolescence and adulthood in striatum but not nucleus accumbens. *Dev. Brain Res.* **89**, 167–172 (1995).

51. Flores, C. Role of netrin-1 in the organization and function of the mesocorticolimbic dopamine system. *Journal of Psychiatry and Neuroscience* vol. 36 296–310 (2011).

52. Grace, A. A., Floresco, S. B., Goto, Y. & Lodge, D. J. Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends Neurosci.* **30**, 220–227 (2007).

53. Jackson, M. E., Frost, A. S. & Moghaddam, B. Stimulation of prefrontal cortex at physiologically relevant frequencies inhibits dopamine release in the nucleus accumbens. *J. Neurochem.* **78**, 920–923 (2001).

54. Simpson, E. H., Kellendonk, C. & Kandel, E. A Possible Role for the Striatum in the Pathogenesis of

the Cognitive Symptoms of Schizophrenia. *Neuron* **65**, 585–596 (2010).

55. Kellendonk, C. *et al.* Transient and selective overexpression of dopamine D2 receptors in the striatum causes persistent abnormalities in prefrontal cortex functioning. *Neuron* **49**, 603–615 (2006).

56. Kim, C. K. *et al.* Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. *Nat. Methods* **13**, 325–328 (2016).

57. Hikida, T. *et al.* Nucleus accumbens pathways control cell-specific gene expression in the medial prefrontal cortex. *Sci. Rep.* **10**, 1–9 (2020).

58. Dalley, J. W. & Robbins, T. W. Fractionating impulsivity: neuropsychiatric implications. *Nat. Rev. Neurosci.* **18**, 158–171 (2017).

59. Dalley, J. W. & Roiser, J. P. Dopamine, serotonin and impulsivity. *Neuroscience* **215**, 42–58 (2012).

60. Bromberg-Martin, E. S., Matsumoto, M. & Hikosaka, O. Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. *Neuron* **68**, 815–834 (2010).

61. Vosberg, D. E., Leyton, M. & Flores, C. The Netrin-1/DCC guidance system: dopamine pathway maturation and psychiatric disorders emerging in adolescence. *Mol. Psychiatry* **25**, 297–307 (2020).

62. Manitt, C. *et al.* Dcc orchestrates the development of the prefrontal cortex during adolescence and is altered in psychiatric patients. *Transl. Psychiatry* **3**, e338–e338 (2013).

63. Manitt, C. *et al.* The netrin receptor DCC is required in the pubertal organization of mesocortical dopamine circuitry. *J. Neurosci.* **31**, 8381–8394 (2011).

64. Cuesta, S. *et al.* Dopamine Axon Targeting in the Nucleus Accumbens in Adolescence Requires Netrin-1. *Front. Cell Dev. Biol.* **8**, 1–12 (2020).

65. Phillips, R. A. *et al.* An atlas of transcriptionally defined cell populations in the rat ventral tegmental area. *Cell Rep.* **39**, 110616 (2022).

66. Vosberg, D. E. *et al.* Mesocorticolimbic connectivity and volumetric alterations in dcc mutation carriers. *J. Neurosci.* **38**, 4655–4665 (2018).

67. Kennedy, T. E., Serafini, T., de la Torre, J. R. & Tessier-Lavigne, M. Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. *Cell* **78**, 425–435 (1994).

68. Kolodkin, A. L. & Tessier-Lavigne, M. Mechanisms and molecules of neuronal wiring: A primer. *Cold Spring Harb. Perspect. Biol.* **3**, 1–14 (2011).

69. K, H. *et al.* A ligand-gated association between cytoplasmic domains of UNC5 and DCC family recepors converts netrin-induced growth cone attraction to repulsion. *Cell* **97**, 927–941 (1999).

70. Keino-masu, K. *et al.* Deleted in Colorectal Cancer (DCC) Encodes a Netrin Receptor. **87**, 175–185 (1996).

71. Manitt, C. *et al.* Widespread expression of netrin-1 by neurons and oligodendrocytes in the adult mammalian spinal cord. *J. Neurosci.* **21**, 3911–3922 (2001).

72. Horn, K. E. *et al.* DCC Expression by Neurons Regulates Synaptic Plasticity in the Adult Brain. *Cell Rep.* **3**, 173–185 (2013).

73. Glasgow, S. D., Ruthazer, E. S. & Kennedy, T. E. Guiding synaptic plasticity: Novel roles for netrin-1 in synaptic plasticity and memory formation in the adult brain. *J. Physiol.* **599**, 493–505 (2021).

74. Jamuar, S. S. *et al.* Biallelic mutations in human DCC cause developmental split-brain syndrome. *Nat. Genet.* **49**, 606–612 (2017).

75. Marsh, A. P. L. *et al.* Mutations in DCC cause isolated agenesis of the corpus callosum with incomplete penetrance. *Nat. Genet.* **49**, 511–514 (2017).

76. Vosberg, D. E. *et al.* Neural function in DCC mutation carriers with and without mirror movements. *Ann. Neurol.* **85**, 433–442 (2019).

77. Finger, J. H. *et al.* The netrin 1 receptors Unc5h3 and Dcc are necessary at multiple choice points for the guidance of corticospinal tract axons. *J. Neurosci.* **22**, 10346–10356 (2002).

78. Torres-Berrío, A., Hernandez, G., Nestler, E. J. & Flores, C. The Netrin-1/DCC Guidance Cue Pathway as a Molecular Target in Depression: Translational Evidence. *Biol. Psychiatry* **88**, 611–624 (2020).

79. Lee, P. H. *et al.* Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **179**, 1469-1482.e11 (2019).

80. Yetnikoff, L., Eng, C., Benning, S. & Flores, C. Netrin-1 receptor in the ventral tegmental area is required for sensitization to amphetamine. **31**, 1292–1302 (2010).

81. Yetnikoff, L. & Almey, A. Abolition of the behavioral phenotype of adult netrin-1 receptor deficient mice by exposure to amphetamine during the juvenile period. 505–514 (2011) doi:10.1007/s00213-011-2312-6.

82. Yetnikoff, L. & Pokinko, M. Adolescence : a time of transition for the phenotype of dcc heterozygous mice. 1705–1714 (2014) doi:10.1007/s00213-013-3083-z.

83. Cuesta, S. *et al.* Non-Contingent Exposure to Amphetamine in Adolescence Recruits miR-218 to Regulate Dcc Expression in the VTA. *Neuropsychopharmacology* **43**, 900–911 (2018).

84. Barabási, A. L. & Oltvai, Z. N. Network biology: Understanding the cell's functional organization. *Nat. Rev. Genet.* **5**, 101–113 (2004).

85. Lee, T. I., & Young, R. A. Transcription of eukaryotic protein-coding genes. *Annu. Rev. Genet.* **34**, 77–137 (2000).

86. Sullivan, P. F., Daly, M. J. & O'Donovan, M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat. Rev. Genet. 2012 138* **13**, 537–551 (2012).

87. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *American Journal of Human Genetics* vol. 101 5–22 (2017).

88. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**, 1177–1186 (2017).

89. Fornito, A. *et al.* Linking GWAS to pharmacological treatments for psychiatric disorders. *medRxiv* 2022.05.11.22274981 (2022).

90. Reynolds, L. M. *et al.* Amphetamine disrupts dopamine axon growth in adolescence by a sex-specific mechanism. *bioRxiv* 2022.12.14.520468 (2022) doi:10.1038/s41467-023-39665-1.

91. Shohat, S., Amelan, A. & Shifman, S. Convergence and Divergence in the Genetics of Psychiatric Disorders From Pathways to Developmental Stages. *Biol. Psychiatry* **89**, 32–40 (2021).

92. Domingo, J., Baeza-Centurion, P. & Lehner, B. The Causes and Consequences of Genetic Interactions (Epistasis). *Annu. Rev. Genomics Hum. Genet.* **20**, 433–460 (2019).