bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Title: A biologically-informed polygenic score identifies endophenotypes and clinical conditions associated with the insulin receptor function on specific brain regions

Authors: Shantala A. Hari Dass[3], Kathryn McCracken[6,7], Irina Pokhvisneva[3], Lawrence M. Chen[3], Elika Garg[3], Thao T. T. Nguyen[3],Zihan Wang[3], Barbara Barth[7], Moein Yaqubi[7], Lisa M. McEwen[5], Julie L. MacIsaac[5], Josie Diorio[3], Michael S. Kobor[5], Kieran J. O'Donnell[1,2,3], Michael J. Meaney[1,2,3,4], Patricia P. Silveira[1,2,3]*.

Affiliations:

[1]Department of Psychiatry, Faculty of Medicine, McGill University, QC, Canada.

[2]Ludmer Centre for Neuroinformatics and Mental Health, Douglas Mental Health University Institute, McGill University, QC, Canada.

[3]Sackler Institute for Epigenetics & Psychobiology at McGill University, QC, Canada.

[4]Singapore Institute for Clinical Sciences, Agency for Science, Technology and Research (A*STAR)

[5]Centre for Molecular Medicine and Therapeutics, BC Children's Hospital Research Institute, Department of Medical Genetics, The University of British Columbia, BC, Canada.

[6]John Abbott College, Sainte-Anne-de-Bellevue, QC, Canada

[7]McGill University, QC, Canada.

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

Abstract: Conventional polygenic scores derived from genome-wide association studies do not reflect gene networks that code for biological functions. We present an alternative approach creating a biologically-informed polygenic score based on the insulin receptor (IR) gene networks in the mesocorticolimbic system and hippocampus that regulate reward sensitivity/inhibitory control and memory, respectively. Across multiple samples (n = 4300) our biologically-informed IR-PRS score showed better prediction of child impulsivity and cognitive performance, as well as risk for early addiction onset and Alzheimer's disease in comparison to conventional polygenic scores for ADHD, addiction and dementia. This novel, biologically-informed approach enables the use of genomic datasets to probe relevant biological processes involved in neural function and disorders.

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

One Sentence Summary: A polygenic score based on genes co-expressed with the insulin receptor predicts childhood behavior and adult disease.

Main Text:

The co-morbidity between metabolic and neuropsychiatric disorders is well-established, but poorly understood. The high co-occurrence of several psychiatric conditions (e.g. major depression, bipolar disorder, dementia) with insulin resistance suggests a common underlying mechanism (Anderson, Freedland et al. 2001). Insulin receptors (IR) are expressed throughout the brain, in areas including the ventral tegmental area, striatum, prefrontal cortex and hippocampus. Insulin is actively transported across the blood-brain barrier, and its action on mesocorticolimbic receptors modulates synaptic plasticity in dopaminergic neurons, affecting dopamine-related behaviors such as response to reward, impulsivity and decision-making (Kullmann, Heni et al. 2016). IR location on hippocampal glutamatergic synapses suggests a role of insulin in neurotransmission, synaptic plasticity and modulation of learning and memory, while its inhibition is described in Alzheimer's disease and related animal models (Bomfim, Forny-Germano et al. 2012).

Genetic studies can be a pertinent tool to investigate the neurobiological mechanisms that explain the co-morbidity between metabolic and neuropsychiatric conditions. Genome-wide association studies (GWAS) provide the basis for cumulative variants that associate with health outcomes and reflect genetic predispositions to common disorders where individual variants carry small effects. The cumulative polygenetic risk of the individual can be used to estimate risk through polygenic risk scores (PRS), by multiplying the measured number of risk alleles at a locus by the effect size of the association between a particular genotype and the outcome from the relevant GWAS study, and summing across the genome (Wray and Goddard 2010). GWAS and PRS methodologies are focused on statistically significant candidate associations between scattered loci and a certain condition or trait, not accounting for the fact that genes operate in networks, and code for precise biological functions in specific tissues. Although recent studies explore the ability of conventional polygenic risk scores (PRS) to detect vulnerability endophenotypes, the conservative approach does not provide information about the underlying network in which these genes operate, nor the mechanism of the disease. This makes the biological underpinnings of the PRS complicated to disentangle, restricting their clinical scope/implications. Since the PRS are conventionally based on statistical thresholds, while they might associate with the final disease state, they are not informative of the various clinical manifestations of the disease, and endophenotypes that may precede the disease development. A score that captures a genetic risk to a clinical manifestation- potentially an early onset manifestation- would be able to mold subsequent clinical and therapeutic endeavors. With the reduction in cost of genotyping strategies we are poised to expand the implications of a multitude of GWA studies into healthcare. Hence, we aimed to address this need by developing a novel genomics approach that provides a biologically-informed genetic score, based on genes co-expressed with the IR in specific brain regions. Our new methodology enabled a hypothesis-driven, neurobiological analysis of phenotypes linked to brain disorders across developmental stages. It allowed us to identify differences in dopamine-related behavioral outcomes, namely impulsivity and risk for addiction, when focused on the mesocorticolimbic system. By comparison, the hippocampal IR gene network permitted the distinction of differences in childhood cognitive performance, and the presence of Alzheimer's disease in adults.

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Establishment of the mesocorticolimbic biologically-informed genetic score based on genes co-expressed with the insulin receptor (ePRS-IR)

The steps for the development of the IR expression-based polygenic risk score (ePRS-IR) are depicted in Figure 1 (Fig. 1A). We (a) extracted brain-region specific IR co-expression data from publicly available RNA sequencing data sets in mice (Mulligan, Mozhui et al. 2017); (b) generated a co-expression matrix with IR in the ventral striatum and in the prefrontal cortex using transcriptomic data (i.e., the mesolimbic IR matrix), only including genes with a correlation coefficient >|0.5|; (c) identified consensus autosomal transcripts from this gene list, performing a developmental enrichment within the same brain areas, by selecting transcripts differentially expressed at ≥1.5 fold during fetal and child development as compared to adult samples (Miller, Ding et al. 2014); the final list included 281 genes; (d) based on their functional annotation, we gathered all the existing SNPs from these genes (total = 11,068); (e) subjected this list of SNPs to linkage disequilibrium clumping (Wray, Lee et al. 2014), to inform removal of highly correlated SNPs, resulting in 1897 independent functional SNPs and (f) compiled these SNPs in a polygenic score(Wray and Goddard 2010) using the association betas described in a conventional ADHD GWAS (Neale, Medland et al. 2010) (Table S1).

This biologically-informed method for selecting SNPs is designed to capture risk associated with a functional unit of a gene network, and hence it can be expected that none of the individual SNPs included in the score are statistically significantly associated with ADHD diagnosis in the GWAS (Fig. 1B), but instead, the cumulative ePRS-IR score associates with the relevant endophenotypes. A gene network analysis was performed comparing the list of the 281 genes included in the ePRS-IR with (a) the genes associated with a random selection of 11,068 SNPs from the ADHD GWAS (Neale, Medland et al. 2010) and (b) the top 281 genes associated with the most significant SNPs from the ADHD GWAS (Neale, Medland et al. 2010) (Fig. 1C). The analysis shows that the number of interactions between the genes that comprise the ePRS-IR is significantly higher than the control random list, or the genes from the ADHD GWAS (Fig. 1D), suggesting that the biologically-informed score represents a much more cohesive gene network. We examined the co-expression patterns of the genes included in the mesocorticolimbic ePRS-IR over the life-course using data from BrainSpan (Miller, Ding et al. 2014). Several genes are highly co-expressed in childhood, as expected from the methodological approach used to select the genes included in the score. This co-expression pattern is not maintained into adulthood, although other clusters of co-expression emerge (Fig. 1E). As a validation of our mesocorticolimbic ePRS-IR score, we created 30 random ePRS and repeated linear regression analyses for each of the random ePRSs. The results of the regression analyses involving different ePRS scores were pooled together to obtain an estimate of the effect of the ePRS interaction with sex on the IST performance at 72 months. The results showed a non-significant interaction term (fixed condition: =0.38, p =0.81; decreasing condition =0.14, p =0.92). When analyzing the shared heritability between our ePRS-IR and other complex traits using a LD score regression (Bulik-Sullivan, Loh et al. 2015), there was a significant genetic correlation between the GWAS for cross-disorder (Cross-Disorder Group of the Psychiatric Genomics 2013) and the mesocorticolimbic ePRS-IR (rg=1.1, p=0.004), which was expected given that the ADHD GWAS is part of the cross-disorder genome-wide analysis. However, we observe the brain specificity of the mesocorticolimbic ePRS-IR in relation to peripheral insulin levels or sensitivity, as there was no significant shared heritability between the ePRS-IR and the fasting

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

insulin GWAS(Manning, Hivert et al. 2012) (rg= 0.53, p=0.25) or HOMA-IR GWAS (Dupuis, Langenberg et al. 2010)(rg=0.62, p=0.16). Brain specificity is also demonstrated when comparing the co-expression patterns of the genes included in our score and an equivalent number of genes associated with the most significant SNPs from the fasting insulin GWAS (Scott, Lagou et al. 2012) in the PFC and striatal regions (Fig. 1F). While the mesocorticolimbic ePRS-IR has highly co-expressed genes in these brain areas, genes from the fasting pro-insulin GWAS are less consistently co-expressed in these regions. This confirms the ability of the ePRS methodology to generate a genetic score that denotes tissue-specific gene networks, and highlights the meaningful difference between the information contained in the ePRS-IR score and peripheral levels of metabolic hormones such as insulin.

Fig.1.The mesocorticolimbic ePRS-IR

Endophenotypic differences predicted by the mesocorticolimbic ePRS-IR

We then analyzed whether the ePRS-IR would predict impulsivity, a highly sex-specific trait, in 6-year old children (O'Donnell 2014) tested using a computer-based, Information Sampling task (IST). There was a significant interaction effect between ePRS-IR and sex on the Information Sampling Task (IST) (fixed condition = 2.385, p = 0.016; decreasing condition = 2.479, p = 0.006) applied at 72 months (Fig. 2A and B). While a simple slopes analysis showed no relationship between the ePRS-IR score and mean P Correct values in girls (fixed condition = −0.636, p = 0.36; decreasing condition = −0.91, p = 0.15), a lower ePRS-IR was significantly related to lower mean P Correct (less certainty when coming up to a decision or higher reflexive impulsivity) in boys (fixed condition = 1.75, p = 0.01; decreasing condition = 1.57, p = 0.01). We next examined whether a “random” PRS or other conventional PRSs (for ADHD (Neale, Medland et al. 2010) or addiction (Tobacco and Genetics 2010)) of comparable size in terms of SNP number could also predict impulsivity in these children. None of these control scores was associated with the interaction effect on reflection impulsivity in this sample (Fig. 2C). We then hypothesized that a lower ePRS-IR, associated with childhood impulsivity in boys as shown above, would predict risk for addiction in men, considering the clinical overlap between impulsive phenotypes and risk for addiction (Fatseas, Hurmic et al. 2016). We therefore investigated if the ePRS-IR was associated with the risk for addiction phenotypes using data from the Study of Addiction: Genetics and Environment (SAGE) repository(Bierut, Saccone et al. 2002). The analysis revealed a significant interaction between the ePRS-IR score and sex for the number of addiction comorbidities (interaction effect; = −11.33, p = 0.01) and probability of alcohol dependence (interaction effect; = −13.71, p = 0.02). Further simple slopes analysis confirmed that a lower ePRS-IR was significantly associated with number of addiction comorbidities only in males (males, simple slope = −7.30, p = 0.04; females, simple slope = 4.59, p = 0.14); simple slopes analysis for alcohol dependence was marginally significant in males (= −7.99, p = 0.06, OR=3.38) with no effect in females (= 6.40, p = 0.10) (Fig. 2D and E) (Table S2).

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. 2. Phenotypic differences predicted by the mesocorticolombic ePRS-IR

Comparison between mesocorticolimbic ePRS-IR using two versions of the ADHD GWAS - 2010 (Neale, Medland et al. 2010) and 2019 (Demontis, Walters et al. 2019)

Considering the theoretical premise of the ePRS-IR (biological polygenic score based on gene co-expression), we hypothesized that the choice of the GWAS for providing effect sizes for weighing the SNPs would not have a major influence on the prediction ability of the expression based polygenic score. We then created simultaneously two mesocorticolimbic ePRS-IR scores at the point of the SNP selection, weighing one using the 2010 ADHD GWAS (Neale, Medland et al. 2010) and the second one using the recently published 2019 ADHD GWAS (Demontis, Walters et al. 2019) . Confirming our hypothesis, both scores were associated with a significant interaction effect with sex on impulsivity phenotype measured by the IST task in MAVAN (2010: = 1.13, p = 0.05; 2018: = 128.8, p = 0.05) (Figure S1). Considering that several studies described associations between polygenic scores using the 2010 ADHD GWAS and endophenotypes associated with attention-impulsivity problems in community samples (rather than ADHD cases vs. controls, similar to our main cohort MAVAN)(Groen-Blokhuis, Middeldorp et al. 2014, Martin, Hamshere et al. 2014), we opted for our main analysis to be focused on the 2010 ADHD GWAS.

The hippocampal ePRS-IR

To further validate the brain-specificity of the biologically-informed polygenic score, we followed the same approach depicted in Fig 1A and created the hippocampal ePRS-IR by changing the brain region of interest in the process, and compiling the SNPs using the association betas described in the Alzheimer's GWAS (Lambert, Ibrahim-Verbaas et al. 2013) (Table S3). Similar to the mesocorticolimbic ePRS-IR, the selected SNPs are not statistically significantly associated with Alzheimer's in the GWAS (Fig. 3A). The gene network analysis demonstrates that the hippocampal ePRS-IR also represents a highly interconnected network (Fig. 3B). Analysis of the co-expression pattern of the genes included in the hippocampal ePRS-IR revealed that the high co-expression clusters from childhood are not maintained in adult life (Fig. 3C). Informed by the role of the IR in the hippocampus on cognition, we explored this endophenotype in children by evaluating cognitive performance on the Number Knowledge Task applied at 48 months (O'Donnell 2014). There was a main effect of the hippocampal ePRS-IR in predicting task performance (p = 0.03, = −1350.1), in which children with higher ePRS-IR scores have lower performance on the cognitive test (Fig. 3D). We found no significant interaction effect between the hippocampal ePRS-IR and sex on the task at this age. The hippocampal ePRS-IR was also able to distinguish the presence of Alzheimer's disease diagnosis in the GenADA study (Li, Wetten et al. 2008) (p = 6.82e-09, = 972.40), in which there are more cases in individuals with higher ePRS-IR scores (Fig. 3E). No significant interaction between ePRS and sex was found in the adult cohort.

Fig. 3. The hippocampal ePRS-IR

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Comparison between mesocorticolimbic and hippocampal ePRS-IR scores

To examine the isolated contribution of unique SNPs to the above described effects, we performed a SNP by sex association analysis for the mesocorticolimbic ePRS-IR (investigating the interaction with sex) and a genome wide association analysis for the hippocampal ePRS-IR (investigating the main effects of ePRS-IR) using linear regressions. There was no unique SNP included in the mesocorticolimbic ePRS-IR responsible for the SNP by sex interaction on impulsivity in children, or early addiction onset in adults. Similarly, no isolated SNP included in the hippocampal ePRS-IR was associated with poorer cognitive performance in children, or Alzheimer's disease, confirming that variations in the biological function of the gene networks represented in the ePRS-IR scores predict their respective outcomes as a global score, with small contributions from all included SNPs, but not by any isolated mutations (Fig. 4A). To further validate the brain region specificity of the scores, we analyzed the distribution of children in each quartile of the mesocorticolimbic and hippocampal ePRS-IR in MAVAN (Fig. 4B) and verified that the scores are independent ($r=0.062$, $p=0.416$). In addition, the mesocorticolimbic ePRS-IR did not predict cognitive performance ($p = 0.08$, $= -12.36$ for main effect, $p = 0.62$, $= -6.93$ for the interaction with sex). Likewise, the hippocampal ePRS-IR was not associated with IST scores ($p = 0.11$, $= 0.02$ for main effect, $p = 0.29$, $= -0.020$ for the interaction with sex). Enrichment analysis of the SNPs comprising the mesocorticolimbic ePRS-IR using Metacore® (Thomson Reuters) showed statistically significant enrichment for pathways involved in cell cycle regulation (FDR q = 1.18e-02). Process networks were significant for neurogenesis/axonal guidance (FDR q = 6.667e-03), transcription and translation (FDR q = 1.39e-02). The hippocampal ePRS-IR is enriched for pathways involved in development (hedgehog signaling, FDR q=1.19e-05) and transcription (FDR q=1.42e-04). Process networks involve cell division (Fig. 4C). Both ePRS-IRs were significantly associated with similar transcription factors including CREB1, C-Myc, SP1 and ESF1 (Fig. 4D).

Fig. 4. ePRS-IR scores are brain-region specific

Discussion

There has been a marked increase in the number of GWAS available. We now have access to association studies addressing a range of health related- disease and otherwise-outcomes. Conventional polygenic risk scores derived from existing GWAS's are intended to reflect genetic vulnerability. However, the PRS approach is limited by the small explained variance in GWAS of common disorders and a lack of insight into the phenotypic constructs of the disease. We propose an alternative approach that integrates the power of GWAS datasets to form an ePRS informed by research in neuroscience, as well as the strength of transcriptomic datasets. The resulting brain-region specific ePRS reflects an integrated gene network and molecular pathway that allows the use of genotyping data to test candidate biological mechanisms (Fig. 5).

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The conventional polygenic profiling method is driven by a statistical approach. Additional levels of biological complexity, such as biomarkers and neuroanatomic loci of relevance - from experimental or theoretical databases - can be superimposed and inform the use of this technology, as we demonstrate here. The need to consider gene networks is further exemplified by the observation that the power to detect a causal SNP is drastically reduced when the trait/disorder is caused by two or more pathways (Wray and Maier 2014). We showed that a biologically-informed polygenic risk score based on genes co-expressed with the central IR represents highly cohesive and relevant gene networks. The mesocorticolimbic-specific ePRS-IR is more strongly associated with impulsivity and the risk for early onset of substance dependence in males than is a conventional PRS for either ADHD or addiction. The sex-specificity of our findings was expected, given the increased prevalence of ADHD and behavioral alterations associated with this condition (such as impulsivity) in boys compared to girls (Silveira, Agranonik et al. 2012, Vasiliadis, Diallo et al. 2017). Likewise the hippocampal-specific ePRS-IR was associated both with cognitive performance in children and with Alzheimer's diagnosis in later ages, demonstrating that the biologically-informed framework is brain region and endophenotype-specific.

Our network analysis shows that the ePRS-IR represents a cohesive gene network with significantly more connections than the list of genes extracted from the GWAS for ADHD, or random lists SNPs. This robust approach therefore goes beyond describing associations between single gene variants and the outcomes, but captures information about the whole gene network, and its function, in specific brain regions. One of the critiques of GWAS is that they are 'a static global' measure, unable to discern spatial or temporal differences, as compared to techniques such as RNAseq or WGBS, which are fundamental to the study of any trait/disorder. Our approach integrates a static global measure-which is easy to acquire in your dataset of interest-with a temporal and spatial information -which are, usually, not feasible to acquire in all human datasets - to compute a score that is adept at informing spatially relevant endophenotypes.

Conduct disorder and impulsivity are the foremost risk traits for alcohol use disorder among the 80 personality disorder criteria of DSM-IV (Rosenstrom, Torvik et al. 2017). There is a relationship between childhood ADHD and the risk for developing drug addiction later in life (Levy, Katusic et al. 2014), especially considering the impulsivity component, rather than inattention (Sihvola, Rose et al. 2011), in agreement to the findings described here. Insulin function is associated with the risk for drug addiction (O'Dell and Nazarian 2016). Diminished insulin sensitivity is related to less endogenous dopamine at D2/3 receptors in the ventral striatum(Caravaggio, Borlido et al. 2015), reinforcing the idea that metabolic processes are involved in dysfunctions of the mesocorticolimbic system, such as drug dependence. The enrichment analysis of our mesocorticolimbic ePRS-IR is consistent with the known role of insulin in cell division (Perdereau, Cailliau et al. 2015), neurogenesis (Ronaghi, Zibaii et al. 2019), signal transduction (Pearson-Leary, Jahagirdar et al. 2018), and axon guidance (Gupta, Yadav et al. 2018).

The hippocampal formation is the region with the higher IR gene expression in the brain (De Felice and Benedict 2015). Lower expression of IR and altered levels of different components of the insulin signaling pathways were described in hippocampi from Alzheimer's patients (Talbot, Wang et al. 2012). A down-regulation in insulin signaling causes memory impairment through inhibition of serine phosphorylation of the insulin receptor via activation of stress kinases like c-Jun-n-terminal kinase (JNK)(Bomfim, Forny-Germano et al. 2012). JNK enhances glycogen synthase kinase 3β (GSK3β) activity that leads to tau phosphorylation

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

followed by accumulation of neurofibrillary tangles – a hallmark of Alzheimer's disease - in the brain (Bedse, Di Domenico et al. 2015). Others have shown that impaired insulin sensitivity is linked to cognitive deficits in the elderly (Benedict, Brooks et al. 2012), and that measures of cognitive function in youth predict later risk for Alzheimer's disease (Snowdon, Greiner et al. 2000, Bloss, Delis et al. 2008), in alignment to our current results.

Our genomic approach integrates information from molecular neurobiology with GWAS technology to develop a biologically-informed polygenic score based on gene co-expression data from specific brain regions. This approach creates a novel genomic measure to identify genetic vulnerability for childhood behavioral phenotypes that predict later neuropsychiatric conditions in community-based samples, highlighting possible targets for drug development.#add a closing clinical statement here#.

Fig. 5. Theoretical framework and applications

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgments:

Funding: This work was funded by the Toxic Stress Research network of the JPB Foundation. The MAVAN Cohort was funded by the Canadian Institutes for Health Research, the Ludmer Family Foundation, the Norlien Foundation (Calgary, Canada), the WOCO Foundation (London, Canada), the Blema & Arnold Steinberg Family Foundation, and the Faculty of Medicine of McGill University. Additional funding was provided by the Jacobs Foundation (Switzerland). LMC was funded by the Brain Canada and Kids Brain Health Network Developmental Neurosciences Research Training Award; **Author contributions:** Conceptualization PPS, MJM, SHD; data curation KM, IP, LMC, TTTN, EG, LMM, JLM; formal analysis IP, SHD, PPS; funding acquisition MJM, MK; investigation PPS, KM, SHD; methodology PPS, SHD, KOD, MY; project administration JD; resources LMM, JLM, EG, TTTN; supervision MJM, PPS; validation SHD, LMC, MY; visualization SHD, KM, LMC; writing – original draft PPS, SHD, MJM; writing – review & editing all authors; **Competing interests:** Authors declare no competing interests; and **Data and materials availability:** Data and materials are available upon request.

Supplementary Materials:

Materials and Methods

Figure S1

Tables S1 to S3

Figure legends:

Fig. 1. The mesocorticolimbic ePRS-IR: A) Flowchart depicting the steps involved in creating the expression-based and mesocorticolimbic-specific polygenic risk score based on genes co-expressed with the insulin receptor (ePRS-IR) using gene co-expression databases. i) GeneNetwork was used to generate a co-expression matrix with insulin receptor (IR) in the ventral striatum and in the prefrontal cortex in mice (absolute value of the co-expression correlation $r \geq 0.5$); ii) BrainSpan was then used to identify consensus autosomal transcripts from this list with developmental enrichment within the same brain areas, selecting transcripts differentially expressed at ≥ 1.5 fold during child and fetal development as compared to adult samples. The final list included 281 genes; iii) Based on their functional annotation in the National Center for Biotechnology Information, U.S. National Library of Medicine using GRCh37.p13, we gathered all the existing SNPs from these genes (total = 11,068) and subjected this list of SNPs to linkage disequilibrium clumping, to inform removal of highly correlated SNPs ($r2>0.2$) across 500kb regions, resulting in 1897 independent functional SNPs based on the children's genotype data from MAVAN (Study Sample ids); iv) We used a count function of the number of alleles at a given SNP (rs1, rs2…) weighted by the effect size (ES) of the association between the individual SNP and ADHD. The sum of these values from the total number of SNPs provides the ePRS-IR score. References; 1. Mulligan, Mozhui et al. 2017, 2. Miller, Ding et al. 2014, 3. Neale, Medland et al. 2010, 4. Wray, Lee et al. 2014, 5. Chen, Yao et al. 2018. **B)** Single nucleotide polymorphisms (SNPs) included in the mesocorticolimbic ePRS-IR, shown in

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

green, in relation to the Manhattan plot of the ADHD GWAS(Neale, Medland et al. 2010). The biologically-informed method for selecting SNPs is designed to capture signals associated with a functional gene network, and hence it can be expected that none of the individual SNPs included in the score are statistically significantly associated with the disease (in this case, ADHD). **C)** Gene Network analysis of the mesocorticolimbic ePRS-IR (larger panel), random (top) and ADHD 2010 (bottom small panel) 281 genes. The random gene list was created selecting a random group of SNPs of the same number of SNPs included in the ePRS-IR. The ADHD 2010 gene list was created selecting genes associated with the most significant SNPs from the ADHD GWAS (Neale, Medland et al. 2010). The picture demonstrates that the ePRS-IR represents a network with significantly higher connectivity than a list of genes associated with a random list of SNPs, and also in comparison to the top genes from the ADHD GWAS. **D)** Comparison of the number of connections between the genes in each network (mesocorticolimbic ePRS-IR, random list and top genes in the ADHD GWAS). Genes included in the mesocorticolimbic ePRS-IR have significantly more interactions than the random list, and the top genes from the ADHD GWAS (One-Way ANOVA, $p<0.05$). This suggests a more cohesive gene network of biological relevance. **E)** Coexpression of the genes included in the mesocorticolimbic ePRS-IR, in childhood and adulthood, combining gene expression data from PFC and striatal regions. Top panels: The heatmap of the genes' co-expression in childhood (left) shows several clusters of highly co-expressed genes. Although the clusters are not maintained in the retained order heatmap for adulthood (right), some other clusters of co-expressed genes are observed. Bottom panels: values of the gene expression correlation coefficients in childhood (left) and adulthood (right). Each vertical line represents correlation with a unique gene. The red line is drawn at a correlation score of zero. Especially in childhood, most genes included in the score have highly correlated gene expression values. Data for this analyses were extracted from BrainSpan (Miller, Ding et al. 2014). **F)** Co-expression of the genes included in the mesocorticolimbic ePRS-IR (top panel) and a comparable number of genes associated with the most significant SNPs from the fasting insulin GWAS (Manning, Hivert et al. 2012) (bottom panel), from birth to 11 years of age in the PFC and striatal regions. The heatmaps demonstrate that while the mesocorticolimbic ePRS-IR includes genes that are highly co-expressed in striatum and PFC, genes from the fasting insulin GWAS are less consistently co-expressed in these brain regions. This suggests that the ePRS methodology captures gene networks that are cohesive in specific brain regions, which may not be represented even when analyzing a genetically correlated GWAS (fasting insulin).

Fig. 2. Phenotypic differences predicted by the mesocorticolombic ePRS-IR: A) and B) Performance in the Information Sampling Task (IST, CANTAB) at 72 months according to sex and ePRS-IR in fixed (A) and decreasing (B) conditions. There is a significant interaction between the genetic score and sex on IST outcome, in which boys with lower ePRS-IR sample less information before taking the decision, being more impulsive than boys with higher ePRS-IR. Boys are depicted in blue and girls in red. N=207. **C)** Validation of the findings by using different polygenic risk scores to investigate the interaction with sex on the Information Sampling Task in children. The mesocorticolimbic ePRS-IR significantly interacted with sex as shown in Figure 2A, predicting impulsivity in boys. A PRS with a random SNP selection, a classic PRS from two different GWASes for ADHD (Neale, Medland et al. 2010, Demontis, Walters et al. 2019) or a PRS for addiction (i.e. smoking) (Tobacco and Genetics 2010) were unable to predict the behavioral phenotype. **D)** Number of addiction co-morbidities according to sex and mesocorticolimbic ePRS-IR in the SAGE cohort. There is a significant interaction between the genetic score and sex on the number of co-morbidities, in which ePRS-IR has a

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

negative effect on the number of co-morbidities in men but not women. **E)** Probability of alcohol abuse in the SAGE cohort. A significant interaction between the genetic score and sex was found on the probability of alcohol dependence, in which ePRS-IR has a negative effect on the probability of alcohol dependence in men but not women.

Fig. 3. The hippocampal ePRS-IR: A) Single nucleotide polymorphisms (SNPs) included in the hippocampal ePRS_IR, shown in green, in relation to the Manhattan plot of the Alzheimer's GWAS (Lambert, Ibrahim-Verbaas et al. 2013). The SNPs captured in the score do not have genome-wide significance. **B)** Gene network analysis in the hippocampal ePRS-IR 543 genes. The picture demonstrates that the hippocampal ePRS-IR also represents a highly cohesive gene network. **C)** Coexpression of the genes included in the hippocampal ePRS-IR, in childhood and adulthood. Top panels: The heatmap of the genes' co-expression in childhood (left) shows several clusters of highly co-expressed genes. The clusters are not maintained in the retained order heatmap for adulthood (right) suggesting that these gene networks are relatedly connected in childhood, but not anymore in adulthood. Bottom panels: values of the correlation coefficients of the gene expression values in childhood (left) and adulthood (right). Each vertical line represents correlation with a unique gene. The red line is drawn at a correlation score of zero. Genes included in the score have highly correlated gene expression in childhood, but this is not seen in adult samples. **D)** Main effect of hippocampal ePRS-IR on Number Knowledge Task performance, applied at 48 months, in which children with higher ePRS-IR scores have lower performance on the cognitive test. No significant interaction between ePRS and sex was found at this age. **E)** Presence of Alzheimer's disease diagnosis according to hippocampal ePRS-IR scores in the GenADA study. There was a main effect of the hippocampal ePRS-IR in predicting Alzheimer's cases in which there are more cases in individuals with higher ePRS-IR scores. N=1565.

Fig. 4. ePRS-IR scores are brain-region specific: A) Manhattan plots investigating the contribution of isolated SNPs into the reported findings. There was no unique SNP included in the mesocorticolimbic ePRS-IR responsible for the SNP by sex interaction on impulsivity in children (top left) or alcohol dependence in adults (top right). Similarly, no isolated SNP included in the hippocampal ePRS-IR was associated with poorer cognitive performance in children (bottom left), or Alzheimer's disease (bottom right). These findings confirm that variations in the biological function of the gene networks represented in the mesocorticolimbic ePRS-IR and hippocampal ePRS-IR can predict these outcomes as a global score, with small contributions from all included SNPs, but not by any isolated mutations. **B)** Pearson's product-moment correlation demonstrating the distribution of children in each quartile of the mesocorticolimbic and hippocampal ePRS-IRs. There is no correlation between the quartiles of the two ePRSs ($r=0.062$, $p=0.416$). This shows that a subject with a high mescorticolimbic ePRS-IR would not necessarily have a high score in the hippocampal ePRS-IR, and vice-versa, demonstrating their brain specificity and independence. **C)** Enrichment analysis of mesocorticolimbic (upper panel) and hippocampal (bottom panel) ePRS-IR. **D)** Transcription factor analysis of the two brain-region specific ePRS-IRs.

Fig. 5. Theoretical framework and applications: The expression based, brain region specific ePRS-IR represents a developmentally relevant gene network with enrichment for specific processes like translation and transcription. Variation in this score predicts particular endophenotypes in childhood – for example, impulsivity and cognitive performance. These

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

behaviors map onto risk for diseases later in life, such as early addiction onset and Alzheimer's disease, and the eRPS-IR is equally associated with these outcomes. Psychopathology related to poor inhibitory control or decision-making encompass a wide range of conditions (ADHD, addiction, eating disorders, gambling, suicide, etc.). Similarly, there is an increasing incidence of Alzheimer's disease and other dementias. The neurobiological understanding of these conditions contributes to the development of tools for early identification and primary prevention.

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References:

American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders. Washington, D.C.

Anderson, R. J., K. E. Freedland, R. E. Clouse and P. J. Lustman (2001). "The prevalence of comorbid depression in adults with diabetes: a meta-analysis." Diabetes Care **24**(6): 1069-1078.

Bedse, G., F. Di Domenico, G. Serviddio and T. Cassano (2015). "Aberrant insulin signaling in Alzheimer's disease: current knowledge." Front Neurosci **9**: 204.

Benedict, C., S. J. Brooks, J. Kullberg, J. Burgos, M. J. Kempton, R. Nordenskjold, R. Nylander, L. Kilander, S. Craft, E. M. Larsson, L. Johansson, H. Ahlstrom, L. Lind and H. B. Schioth (2012). "Impaired insulin sensitivity as indexed by the HOMA score is associated with deficits in verbal fluency and temporal lobe gray matter volume in the elderly." Diabetes Care **35**(3): 488-494.

Bennett, D., S. Oldham, A. Dawson, L. Parkes, C. Murawski and M. Yu Cel (2016). "Systematic Overestimation of Reflection Impulsivity in the Information Sampling Task." Biol Psychiatry.

Bierut, L. J. (2007). "Genetic variation that contributes to nicotine dependence." Pharmacogenomics **8**(8): 881-883.

Bierut, L. J., P. A. Madden, N. Breslau, E. O. Johnson, D. Hatsukami, O. F. Pomerleau, G. E. Swan, J. Rutter, S. Bertelsen, L. Fox, D. Fugman, A. M. Goate, A. L. Hinrichs, K. Konvicka, N. G. Martin, G. W. Montgomery, N. L. Saccone, S. F. Saccone, J. C. Wang, G. A. Chase, J. P. Rice and D. G. Ballinger (2007). "Novel genes identified in a high-density genome wide association study for nicotine dependence." Hum Mol Genet **16**(1): 24-35.

Bierut, L. J., N. L. Saccone, J. P. Rice, A. Goate, T. Foroud, H. Edenberg, L. Almasy, P. M. Conneally, R. Crowe, V. Hesselbrock, T. K. Li, J. Nurnberger, Jr., B. Porjesz, M. A. Schuckit, J. Tischfield, H. Begleiter and T. Reich (2002). "Defining alcohol-related phenotypes in humans. The Collaborative Study on the Genetics of Alcoholism." Alcohol Res Health **26**(3): 208-213.

Bierut, L. J., J. R. Strickland, J. R. Thompson, S. E. Afful and L. B. Cottler (2008). "Drug use and dependence in cocaine dependent subjects, community-based individuals, and their siblings." Drug Alcohol Depend **95**(1-2): 14-22.

Bloss, C. S., D. C. Delis, D. P. Salmon and M. W. Bondi (2008). "Decreased cognition in children with risk factors for Alzheimer's disease." Biol Psychiatry **64**(10): 904-906.

Bomfim, T. R., L. Forny-Germano, L. B. Sathler, J. Brito-Moreira, J. C. Houzel, H. Decker, M. A. Silverman, H. Kazi, H. M. Melo, P. L. McClean, C. Holscher, S. E. Arnold, K. Talbot, W. L. Klein, D. P. Munoz, S. T. Ferreira and F. G. De Felice (2012). "An anti-diabetes agent protects the mouse brain from defective insulin signaling caused by Alzheimer's disease- associated Abeta oligomers." J Clin Invest **122**(4): 1339-1353.

Bucholz, K. K., R. Cadoret, C. R. Cloninger, S. H. Dinwiddie, V. M. Hesselbrock, J. I. Nurnberger, Jr., T. Reich, I. Schmidt and M. A. Schuckit (1994). "A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA." J Stud Alcohol **55**(2): 149-158.

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bulik-Sullivan, B., H. K. Finucane, V. Anttila, A. Gusev, F. R. Day, P. R. Loh, C. ReproGen, C. Psychiatric Genomics, C. Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control, L. Duncan, J. R. Perry, N. Patterson, E. B. Robinson, M. J. Daly, A. L. Price and B. M. Neale (2015). "An atlas of genetic correlations across human diseases and traits." Nat Genet **47**(11): 1236-1241.

Bulik-Sullivan, B. K., P. R. Loh, H. K. Finucane, S. Ripke, J. Yang, C. Schizophrenia Working Group of the Psychiatric Genomics, N. Patterson, M. J. Daly, A. L. Price and B. M. Neale (2015). "LD Score regression distinguishes confounding from polygenicity in genome-wide association studies." Nat Genet **47**(3): 291-295.

Canada, S. (2005). Low Income Cut Offs for 2005 and Low Income measures for 2004. Catalogue no. 75F0002MIE, Statistics Canada. **4**.

Caravaggio, F., C. Borlido, M. Hahn, Z. Feng, G. Fervaha, P. Gerretsen, S. Nakajima, E. Plitman, J. K. Chung, Y. Iwata, A. Wilson, G. Remington and A. Graff-Guerrero (2015). "Reduced insulin sensitivity is related to less endogenous dopamine at D2/3 receptors in the ventral striatum of healthy nonobese humans." Int J Neuropsychopharmacol **18**(7): pyv014.

Chen, L. M., N. Yao, E. Garg, Y. Zhu, T. T. T. Nguyen, I. Pokhvisneva, S. A. Hari Dass, E. Unternaehrer, H. Gaudreau, M. Forest, L. M. McEwen, J. L. MacIsaac, M. S. Kobor, C. M. T. Greenwood, P. P. Silveira, M. J. Meaney and K. J. O'Donnell (2018). "PRS-on-Spark (PRSoS): a novel, efficient and flexible approach for generating polygenic risk scores." BMC Bioinformatics **19**(1): 295.

Clark, L. and T. W. Robbins (2016). "Reply to: Systematic Overestimation of Reflection Impulsivity in the Information Sampling Task." Biol Psychiatry.

Cross-Disorder Group of the Psychiatric Genomics, C. (2013). "Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis." Lancet **381**(9875): 1371-1379.

De Felice, F. G. and C. Benedict (2015). "A Key Role of Insulin Receptors in Memory." Diabetes **64**(11): 3653-3655.

Demontis, D., R. K. Walters, J. Martin, M. Mattheisen, T. D. Als, E. Agerbo, G. Baldursson, R. Belliveau, J. Bybjerg-Grauholm, M. Baekvad-Hansen, F. Cerrato, K. Chambert, C. Churchhouse, A. Dumont, N. Eriksson, M. Gandal, J. I. Goldstein, K. L. Grasby, J. Grove, O. O. Gudmundsson, C. S. Hansen, M. E. Hauberg, M. V. Hollegaard, D. P. Howrigan, H. Huang, J. B. Maller, A. R. Martin, N. G. Martin, J. Moran, J. Pallesen, D. S. Palmer, C. B. Pedersen, M. G. Pedersen, T. Poterba, J. B. Poulsen, S. Ripke, E. B. Robinson, F. K. Satterstrom, H. Stefansson, C. Stevens, P. Turley, G. B. Walters, H. Won, M. J. Wright, A. W. G. o. t. P. G. Consortium, L. Early, C. Genetic Epidemiology, T. andMe Research, O. A. Andreassen, P. Asherson, C. L. Burton, D. I. Boomsma, B. Cormand, S. Dalsgaard, B. Franke, J. Gelernter, D. Geschwind, H. Hakonarson, J. Haavik, H. R. Kranzler, J. Kuntsi, K. Langley, K. P. Lesch, C. Middeldorp, A. Reif, L. A. Rohde, P. Roussos, R. Schachar, P. Sklar, E. J. S. Sonuga-Barke, P. F. Sullivan, A. Thapar, J. Y. Tung, I. D. Waldman, S. E. Medland, K. Stefansson, M. Nordentoft, D. M. Hougaard, T. Werge, O. Mors, P. B. Mortensen, M. J. Daly, S. V. Faraone, A. D. Borglum and B. M. Neale (2019). "Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder." Nat Genet **51**(1): 63-75.

Dupuis, J., C. Langenberg, I. Prokopenko, R. Saxena, N. Soranzo, A. U. Jackson, E. Wheeler, N. L. Glazer, N. Bouatia-Naji, A. L. Gloyn, C. M. Lindgren, R. Magi, A. P. Morris, J. Randall, T. Johnson, P. Elliott, D. Rybin, G. Thorleifsson, V. Steinthorsdottir, P. Henneman, H. Grallert, A. Dehghan, J. J. Hottenga, C. S. Franklin, P. Navarro, K. Song, A. Goel, J. R. Perry, J. M. Egan, T.

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Lajunen, N. Grarup, T. Sparso, A. Doney, B. F. Voight, H. M. Stringham, M. Li, S. Kanoni, P. Shrader, C. Cavalcanti-Proenca, M. Kumari, L. Qi, N. J. Timpson, C. Gieger, C. Zabena, G. Rocheleau, E. Ingelsson, P. An, J. O'Connell, J. Luan, A. Elliott, S. A. McCarroll, F. Payne, R. M. Roccasecca, F. Pattou, P. Sethupathy, K. Ardlie, Y. Ariyurek, B. Balkau, P. Barter, J. P. Beilby, Y. Ben-Shlomo, R. Benediktsson, A. J. Bennett, S. Bergmann, M. Bochud, E. Boerwinkle, A. Bonnefond, L. L. Bonnycastle, K. Borch-Johnsen, Y. Bottcher, E. Brunner, S. J. Bumpstead, G. Charpentier, Y. D. Chen, P. Chines, R. Clarke, L. J. Coin, M. N. Cooper, M. Cornelis, G. Crawford, L. Crisponi, I. N. Day, E. J. de Geus, J. Delplanque, C. Dina, M. R. Erdos, A. C. Fedson, A. Fischer-Rosinsky, N. G. Forouhi, C. S. Fox, R. Frants, M. G. Franzosi, P. Galan, M. O. Goodarzi, J. Graessler, C. J. Groves, S. Grundy, R. Gwilliam, U. Gyllensten, S. Hadjadj, G. Hallmans, N. Hammond, X. Han, A. L. Hartikainen, N. Hassanali, C. Hayward, S. C. Heath, S. Hercberg, C. Herder, A. A. Hicks, D. R. Hillman, A. D. Hingorani, A. Hofman, J. Hui, J. Hung, B. Isomaa, P. R. Johnson, T. Jorgensen, A. Jula, M. Kaakinen, J. Kaprio, Y. A. Kesaniemi, M. Kivimaki, B. Knight, S. Koskinen, P. Kovacs, K. O. Kyvik, G. M. Lathrop, D. A. Lawlor, O. Le Bacquer, C. Lecoeur, Y. Li, V. Lyssenko, R. Mahley, M. Mangino, A. K. Manning, M. T. Martinez-Larrad, J. B. McAteer, L. J. McCulloch, R. McPherson, C. Meisinger, D. Melzer, D. Meyre, B. D. Mitchell, M. A. Morken, S. Mukherjee, S. Naitza, N. Narisu, M. J. Neville, B. A. Oostra, M. Orru, R. Pakyz, C. N. Palmer, G. Paolisso, C. Pattaro, D. Pearson, J. F. Peden, N. L. Pedersen, M. Perola, A. F. Pfeiffer, I. Pichler, O. Polasek, D. Posthuma, S. C. Potter, A. Pouta, M. A. Province, B. M. Psaty, W. Rathmann, N. W. Rayner, K. Rice, S. Ripatti, F. Rivadeneira, M. Roden, O. Rolandsson, A. Sandbaek, M. Sandhu, S. Sanna, A. A. Sayer, P. Scheet, L. J. Scott, U. Seedorf, S. J. Sharp, B. Shields, G. Sigurethsson, E. J. Sijbrands, A. Silveira, L. Simpson, A. Singleton, N. L. Smith, U. Sovio, A. Swift, H. Syddall, A. C. Syvanen, T. Tanaka, B. Thorand, J. Tichet, A. Tonjes, T. Tuomi, A. G. Uitterlinden, K. W. van Dijk, M. van Hoek, D. Varma, S. Visvikis-Siest, V. Vitart, N. Vogelzangs, G. Waeber, P. J. Wagner, A. Walley, G. B. Walters, K. L. Ward, H. Watkins, M. N. Weedon, S. H. Wild, G. Willemsen, J. C. Witteman, J. W. Yarnell, E. Zeggini, D. Zelenika, B. Zethelius, G. Zhai, J. H. Zhao, M. C. Zillikens, D. Consortium, G. Consortium, B. C. Global, I. B. Borecki, R. J. Loos, P. Meneton, P. K. Magnusson, D. M. Nathan, G. H. Williams, A. T. Hattersley, K. Silander, V. Salomaa, G. D. Smith, S. R. Bornstein, P. Schwarz, J. Spranger, F. Karpe, A. R. Shuldiner, C. Cooper, G. V. Dedoussis, M. Serrano-Rios, A. D. Morris, L. Lind, L. J. Palmer, F. B. Hu, P. W. Franks, S. Ebrahim, M. Marmot, W. H. Kao, J. S. Pankow, M. J. Sampson, J. Kuusisto, M. Laakso, T. Hansen, O. Pedersen, P. P. Pramstaller, H. E. Wichmann, T. Illig, I. Rudan, A. F. Wright, M. Stumvoll, H. Campbell, J. F. Wilson, C. Anders Hamsten on behalf of Procardis, M. investigators, R. N. Bergman, T. A. Buchanan, F. S. Collins, K. L. Mohlke, J. Tuomilehto, T. T. Valle, D. Altshuler, J. I. Rotter, D. S. Siscovick, B. W. Penninx, D. I. Boomsma, P. Deloukas, T. D. Spector, T. M. Frayling, L. Ferrucci, A. Kong, U. Thorsteinsdottir, K. Stefansson, C. M. van Duijn, Y. S. Aulchenko, A. Cao, A. Scuteri, D. Schlessinger, M. Uda, A. Ruokonen, M. R. Jarvelin, D. M. Waterworth, P. Vollenweider, L. Peltonen, V. Mooser, G. R. Abecasis, N. J. Wareham, R. Sladek, P. Froguel, R. M. Watanabe, J. B. Meigs, L. Groop, M. Boehnke, M. I. McCarthy, J. C. Florez and I. Barroso (2010). "New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk." Nat Genet **42**(2): 105-116.

Edenberg, H. J. (2002). "The collaborative study on the genetics of alcoholism: an update." Alcohol Res Health **26**(3): 214-218.

Edenberg, H. J., X. Xuei, H. J. Chen, H. Tian, L. F. Wetherill, D. M. Dick, L. Almasy, L. Bierut, K. K. Bucholz, A. Goate, V. Hesselbrock, S. Kuperman, J. Nurnberger, B. Porjesz, J. Rice, M.

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Schuckit, J. Tischfield, H. Begleiter and T. Foroud (2006). "Association of alcohol dehydrogenase genes with alcohol dependence: a comprehensive analysis." Hum Mol Genet **15**(9): 1539-1549.
Fatseas, M., H. Hurmic, F. Serre, R. Debrabant, J. P. Daulouede, C. Denis and M. Auriacombe (2016). "Addiction severity pattern associated with adult and childhood Attention Deficit Hyperactivity Disorder (ADHD) in patients with addictions." Psychiatry Res **246**: 656-662.
Filippini, N., A. Rao, S. Wetten, R. A. Gibson, M. Borrie, D. Guzman, A. Kertesz, I. Loy-English, J. Williams, T. Nichols, B. Whitcher and P. M. Matthews (2009). "Anatomically-distinct genetic associations of APOE epsilon4 allele load with regional cortical atrophy in Alzheimer's disease." Neuroimage **44**(3): 724-728.
Franceschini, A., D. Szklarczyk, S. Frankild, M. Kuhn, M. Simonovic, A. Roth, J. Lin, P. Minguez, P. Bork, C. von Mering and L. J. Jensen (2013). "STRING v9.1: protein-protein interaction networks, with increased coverage and integration." Nucleic Acids Res **41**(Database issue): D808-815.
Groen-Blokhuis, M. M., C. M. Middeldorp, K.-J. Kan, A. Abdellaoui, C. E. M. van Beijsterveldt, E. A. Ehli, G. E. Davies, P. A. Scheet, X. Xiao, J. J. Hudziak, J.-J. Hottenga, B. M. Neale and D. I. Boomsma (2014). "Attention-Deficit/Hyperactivity Disorder Polygenic Risk Scores Predict Attention Problems in a Population-Based Sample of Children." Journal of the American Academy of Child & Adolescent Psychiatry **53**(10): 1123-1129.e1126.
Gupta, S., K. Yadav, S. S. Mantri, N. K. Singhal, S. Ganesh and R. Sandhir (2018). "Evidence for Compromised Insulin Signaling and Neuronal Vulnerability in Experimental Model of Sporadic Alzheimer's Disease." Mol Neurobiol **55**(12): 8916-8935.
Kullmann, S., M. Heni, M. Hallschmid, A. Fritsche, H. Preissl and H. U. Haring (2016). "Brain Insulin Resistance at the Crossroads of Metabolic and Cognitive Disorders in Humans." Physiol Rev **96**(4): 1169-1209.
Lambert, J. C., C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, C. Bellenguez, A. L. DeStafano, J. C. Bis, G. W. Beecham, B. Grenier-Boley, G. Russo, T. A. Thorton-Wells, N. Jones, A. V. Smith, V. Chouraki, C. Thomas, M. A. Ikram, D. Zelenika, B. N. Vardarajan, Y. Kamatani, C. F. Lin, A. Gerrish, H. Schmidt, B. Kunkle, M. L. Dunstan, A. Ruiz, M. T. Bihoreau, S. H. Choi, C. Reitz, F. Pasquier, C. Cruchaga, D. Craig, N. Amin, C. Berr, O. L. Lopez, P. L. De Jager, V. Deramecourt, J. A. Johnston, D. Evans, S. Lovestone, L. Letenneur, F. J. Moron, D. C. Rubinsztein, G. Eiriksdottir, K. Sleegers, A. M. Goate, N. Fievet, M. W. Huentelman, M. Gill, K. Brown, M. I. Kamboh, L. Keller, P. Barberger-Gateau, B. McGuiness, E. B. Larson, R. Green, A. J. Myers, C. Dufouil, S. Todd, D. Wallon, S. Love, E. Rogaeva, J. Gallacher, P. St George-Hyslop, J. Clarimon, A. Lleo, A. Bayer, D. W. Tsuang, L. Yu, M. Tsolaki, P. Bossu, G. Spalletta, P. Proitsi, J. Collinge, S. Sorbi, F. Sanchez-Garcia, N. C. Fox, J. Hardy, M. C. Deniz Naranjo, P. Bosco, R. Clarke, C. Brayne, D. Galimberti, M. Mancuso, F. Matthews, I. European Alzheimer's Disease, Genetic, D. Environmental Risk in Alzheimer's, C. Alzheimer's Disease Genetic, H. Cohorts for, E. Aging Research in Genomic, S. Moebus, P. Mecocci, M. Del Zompo, W. Maier, H. Hampel, A. Pilotto, M. Bullido, F. Panza, P. Caffarra, B. Nacmias, J. R. Gilbert, M. Mayhaus, L. Lannefelt, H. Hakonarson, S. Pichler, M. M. Carrasquillo, M. Ingelsson, D. Beekly, V. Alvarez, F. Zou, O. Valladares, S. G. Younkin, E. Coto, K. L. Hamilton-Nelson, W. Gu, C. Razquin, P. Pastor, I. Mateo, M. J. Owen, K. M. Faber, P. V. Jonsson, O. Combarros, M. C. O'Donovan, L. B. Cantwell, H. Soininen, D. Blacker, S. Mead, T. H. Mosley, Jr., D. A. Bennett, T. B. Harris, L. Fratiglioni, C. Holmes, R. F. de Bruijn, P. Passmore, T. J. Montine, K. Bettens, J. I. Rotter, A. Brice, K. Morgan, T. M. Foroud, W. A.

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Kukull, D. Hannequin, J. F. Powell, M. A. Nalls, K. Ritchie, K. L. Lunetta, J. S. Kauwe, E. Boerwinkle, M. Riemenschneider, M. Boada, M. Hiltuenen, E. R. Martin, R. Schmidt, D. Rujescu, L. S. Wang, J. F. Dartigues, R. Mayeux, C. Tzourio, A. Hofman, M. M. Nothen, C. Graff, B. M. Psaty, L. Jones, J. L. Haines, P. A. Holmans, M. Lathrop, M. A. Pericak-Vance, L. J. Launer, L. A. Farrer, C. M. van Duijn, C. Van Broeckhoven, V. Moskvina, S. Seshadri, J. Williams, G. D. Schellenberg and P. Amouyel (2013). "Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease." Nat Genet **45**(12): 1452-1458.
Lemelin, J. P., M. Boivin, N. Forget-Dubois, G. Dionne, J. R. Seguin, M. Brendgen, F. Vitaro, R. E. Tremblay and D. Perusse (2007). "The genetic-environmental etiology of cognitive school readiness and later academic achievement in early childhood." Child Dev **78**(6): 1855-1869.
Levy, S., S. K. Katusic, R. C. Colligan, A. L. Weaver, J. M. Killian, R. G. Voigt and W. J. Barbaresi (2014). "Childhood ADHD and risk for substance dependence in adulthood: a longitudinal, population-based study." PLoS One **9**(8): e105640.
Li, H., S. Wetten, L. Li, P. L. St Jean, R. Upmanyu, L. Surh, D. Hosford, M. R. Barnes, J. D. Briley, M. Borrie, N. Coletta, R. Delisle, D. Dhalla, M. G. Ehm, H. H. Feldman, L. Fornazzari, S. Gauthier, N. Goodgame, D. Guzman, S. Hammond, P. Hollingworth, G. Y. Hsiung, J. Johnson, D. D. Kelly, R. Keren, A. Kertesz, K. S. King, S. Lovestone, I. Loy-English, P. M. Matthews, M. J. Owen, M. Plumpton, W. Pryse-Phillips, R. K. Prinjha, J. C. Richardson, A. Saunders, A. J. Slater, P. H. St George-Hyslop, S. W. Stinnett, J. E. Swartz, R. L. Taylor, J. Wherrett, J. Williams, D. P. Yarnall, R. A. Gibson, M. C. Irizarry, L. T. Middleton and A. D. Roses (2008). "Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease." Arch Neurol **65**(1): 45-53.
Luciana, M. and C. A. Nelson (2002). "Assessment of neuropsychological function through use of the Cambridge Neuropsychological Testing Automated Battery: performance in 4- to 12-year-old children." Dev Neuropsychol **22**(3): 595-624.
Manning, A. K., M. F. Hivert, R. A. Scott, J. L. Grimsby, N. Bouatia-Naji, H. Chen, D. Rybin, C. T. Liu, L. F. Bielak, I. Prokopenko, N. Amin, D. Barnes, G. Cadby, J. J. Hottenga, E. Ingelsson, A. U. Jackson, T. Johnson, S. Kanoni, C. Ladenvall, V. Lagou, J. Lahti, C. Lecoeur, Y. Liu, M. T. Martinez-Larrad, M. E. Montasser, P. Navarro, J. R. Perry, L. J. Rasmussen-Torvik, P. Salo, N. Sattar, D. Shungin, R. J. Strawbridge, T. Tanaka, C. M. van Duijn, P. An, M. de Andrade, J. S. Andrews, T. Aspelund, M. Atalay, Y. Aulchenko, B. Balkau, S. Bandinelli, J. S. Beckmann, J. P. Beilby, C. Bellis, R. N. Bergman, J. Blangero, M. Boban, M. Boehnke, E. Boerwinkle, L. L. Bonnycastle, D. I. Boomsma, I. B. Borecki, Y. Bottcher, C. Bouchard, E. Brunner, D. Budimir, H. Campbell, O. Carlson, P. S. Chines, R. Clarke, F. S. Collins, A. Corbaton-Anchuelo, D. Couper, U. de Faire, G. V. Dedoussis, P. Deloukas, M. Dimitriou, J. M. Egan, G. Eiriksdottir, M. R. Erdos, J. G. Eriksson, E. Eury, L. Ferrucci, I. Ford, N. G. Forouhi, C. S. Fox, M. G. Franzosi, P. W. Franks, T. M. Frayling, P. Froguel, P. Galan, E. de Geus, B. Gigante, N. L. Glazer, A. Goel, L. Groop, V. Gudnason, G. Hallmans, A. Hamsten, O. Hansson, T. B. Harris, C. Hayward, S. Heath, S. Hercberg, A. A. Hicks, A. Hingorani, A. Hofman, J. Hui, J. Hung, M. R. Jarvelin, M. A. Jhun, P. C. Johnson, J. W. Jukema, A. Jula, W. H. Kao, J. Kaprio, S. L. Kardia, S. Keinanen-Kiukaanniemi, M. Kivimaki, I. Kolcic, P. Kovacs, M. Kumari, J. Kuusisto, K. O. Kyvik, M. Laakso, T. Lakka, L. Lannfelt, G. M. Lathrop, L. J. Launer, K. Leander, G. Li, L. Lind, J. Lindstrom, S. Lobbens, R. J. Loos, J. Luan, V. Lyssenko, R. Magi, P. K. Magnusson, M. Marmot, P. Meneton, K. L. Mohlke, V. Mooser, M. A. Morken, I. Miljkovic, N. Narisu, J. O'Connell, K. K. Ong, B. A. Oostra, L. J. Palmer, A. Palotie, J. S. Pankow, J. F. Peden, N. L. Pedersen, M. Pehlic, L. Peltonen, B. Penninx, M. Pericic, M. Perola, L. Perusse, P.

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A. Peyser, O. Polasek, P. P. Pramstaller, M. A. Province, K. Raikkonen, R. Rauramaa, E. Rehnberg, K. Rice, J. I. Rotter, I. Rudan, A. Ruokonen, T. Saaristo, M. Sabater-Lleal, V. Salomaa, D. B. Savage, R. Saxena, P. Schwarz, U. Seedorf, B. Sennblad, M. Serrano-Rios, A. R. Shuldiner, E. J. Sijbrands, D. S. Siscovick, J. H. Smit, K. S. Small, N. L. Smith, A. V. Smith, A. Stancakova, K. Stirrups, M. Stumvoll, Y. V. Sun, A. J. Swift, A. Tonjes, J. Tuomilehto, S. Trompet, A. G. Uitterlinden, M. Uusitupa, M. Vikstrom, V. Vitart, M. C. Vohl, B. F. Voight, P. Vollenweider, G. Waeber, D. M. Waterworth, H. Watkins, E. Wheeler, E. Widen, S. H. Wild, S. M. Willems, G. Willemsen, J. F. Wilson, J. C. Witteman, A. F. Wright, H. Yaghootkar, D. Zelenika, T. Zemunik, L. Zgaga, D. I. G. Replication, C. Meta-analysis, C. Multiple Tissue Human Expression Resource, N. J. Wareham, M. I. McCarthy, I. Barroso, R. M. Watanabe, J. C. Florez, J. Dupuis, J. B. Meigs and C. Langenberg (2012). "A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance." Nat Genet **44**(6): 659-669.

Martin, J., M. L. Hamshere, E. Stergiakouli, M. C. O'Donovan and A. Thapar (2014). "Genetic Risk for Attention-Deficit/Hyperactivity Disorder Contributes to Neurodevelopmental Traits in the General Population." Biological Psychiatry **76**(8): 664-671.

McCarthy, S., S. Das, W. Kretzschmar, O. Delaneau, A. R. Wood, A. Teumer, H. M. Kang, C. Fuchsberger, P. Danecek, K. Sharp, Y. Luo, C. Sidore, A. Kwong, N. Timpson, S. Koskinen, S. Vrieze, L. J. Scott, H. Zhang, A. Mahajan, J. Veldink, U. Peters, C. Pato, C. M. van Duijn, C. E. Gillies, I. Gandin, M. Mezzavilla, A. Gilly, M. Cocca, M. Traglia, A. Angius, J. C. Barrett, D. Boomsma, K. Branham, G. Breen, C. M. Brummett, F. Busonero, H. Campbell, A. Chan, S. Chen, E. Chew, F. S. Collins, L. J. Corbin, G. D. Smith, G. Dedoussis, M. Dorr, A. E. Farmaki, L. Ferrucci, L. Forer, R. M. Fraser, S. Gabriel, S. Levy, L. Groop, T. Harrison, A. Hattersley, O. L. Holmen, K. Hveem, M. Kretzler, J. C. Lee, M. McGue, T. Meitinger, D. Melzer, J. L. Min, K. L. Mohlke, J. B. Vincent, M. Nauck, D. Nickerson, A. Palotie, M. Pato, N. Pirastu, M. McInnis, J. B. Richards, C. Sala, V. Salomaa, D. Schlessinger, S. Schoenherr, P. E. Slagboom, K. Small, T. Spector, D. Stambolian, M. Tuke, J. Tuomilehto, L. H. Van den Berg, W. Van Rheenen, U. Volker, C. Wijmenga, D. Toniolo, E. Zeggini, P. Gasparini, M. G. Sampson, J. F. Wilson, T. Frayling, P. I. de Bakker, M. A. Swertz, S. McCarroll, C. Kooperberg, A. Dekker, D. Altshuler, C. Willer, W. Iacono, S. Ripatti, N. Soranzo, K. Walter, A. Swaroop, F. Cucca, C. A. Anderson, R. M. Myers, M. Boehnke, M. I. McCarthy, R. Durbin and C. Haplotype Reference (2016). "A reference panel of 64,976 haplotypes for genotype imputation." Nat Genet **48**(10): 1279-1283.

Miller, J. A., S. L. Ding, S. M. Sunkin, K. A. Smith, L. Ng, A. Szafer, A. Ebbert, Z. L. Riley, J. J. Royall, K. Aiona, J. M. Arnold, C. Bennet, D. Bertagnolli, K. Brouner, S. Butler, S. Caldejon, A. Carey, C. Cuhaciyan, R. A. Dalley, N. Dee, T. A. Dolbeare, B. A. Facer, D. Feng, T. P. Fliss, G. Gee, J. Goldy, L. Gourley, B. W. Gregor, G. Gu, R. E. Howard, J. M. Jochim, C. L. Kuan, C. Lau, C. K. Lee, F. Lee, T. A. Lemon, P. Lesnar, B. McMurray, N. Mastan, N. Mosqueda, T. Naluai-Cecchini, N. K. Ngo, J. Nyhus, A. Oldre, E. Olson, J. Parente, P. D. Parker, S. E. Parry, A. Stevens, M. Pletikos, M. Reding, K. Roll, D. Sandman, M. Sarreal, S. Shapouri, N. V. Shapovalova, E. H. Shen, N. Sjoquist, C. R. Slaughterbeck, M. Smith, A. J. Sodt, D. Williams, L. Zollei, B. Fischl, M. B. Gerstein, D. H. Geschwind, I. A. Glass, M. J. Hawrylycz, R. F. Hevner, H. Huang, A. R. Jones, J. A. Knowles, P. Levitt, J. W. Phillips, N. Sestan, P. Wohnoutka, C. Dang, A. Bernard, J. G. Hohmann and E. S. Lein (2014). "Transcriptional landscape of the prenatal human brain." Nature **508**(7495): 199-206.

Mulligan, M. K., K. Mozhui, P. Prins and R. W. Williams (2017). "GeneNetwork: A Toolbox for Systems Genetics." Methods Mol Biol **1488**: 75-120.

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Neale, B. M., S. E. Medland, S. Ripke, P. Asherson, B. Franke, K. P. Lesch, S. V. Faraone, T. T. Nguyen, H. Schafer, P. Holmans, M. Daly, H. C. Steinhausen, C. Freitag, A. Reif, T. J. Renner, M. Romanos, J. Romanos, S. Walitza, A. Warnke, J. Meyer, H. Palmason, J. Buitelaar, A. A. Vasquez, N. Lambregts-Rommelse, M. Gill, R. J. Anney, K. Langely, M. O'Donovan, N. Williams, M. Owen, A. Thapar, L. Kent, J. Sergeant, H. Roeyers, E. Mick, J. Biederman, A. Doyle, S. Smalley, S. Loo, H. Hakonarson, J. Elia, A. Todorov, A. Miranda, F. Mulas, R. P. Ebstein, A. Rothenberger, T. Banaschewski, R. D. Oades, E. Sonuga-Barke, J. McGough, L. Nisenbaum, F. Middleton, X. Hu, S. Nelson and G. C. A. S. Psychiatric (2010). "Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder." J Am Acad Child Adolesc Psychiatry **49**(9): 884-897.

O'Dell, L. E. and A. Nazarian (2016). "Enhanced vulnerability to tobacco use in persons with diabetes: A behavioral and neurobiological framework." Prog Neuropsychopharmacol Biol Psychiatry **65**: 288-296.

O'Donnell, K. G., H.; Colalillo, S.; Steiner, M.; Atkinson, L.; Moss, E.; Karama, S.; Matthews, S.; Lydon, J.; Silveira, P.P. Wazana, A.; Levitan, R. Sokolowski, M.; Kennedy J.; Fleming, A. Meaney, M. (2014). "The Maternal Adversity Vulnerability and Neurodevelopment (MAVAN) Project: Theory and methodology." Can J Psychiatry **59**(9): 497-508.

Okamoto, Y. and R. Case (1996). "Exploring the microstructure of children's central conceptual structures in the domain of number." Monogr Soc Res Child Dev **61**(1-2): 27-58.

Pearson-Leary, J., V. Jahagirdar, J. Sage and E. C. McNay (2018). "Insulin modulates hippocampally-mediated spatial working memory via glucose transporter-4." Behav Brain Res **338**: 32-39.

Perdereau, D., K. Cailliau, E. Browaeys-Poly, A. Lescuyer, N. Carre, F. Benhamed, D. Goenaga and A. F. Burnol (2015). "Insulin-induced cell division is controlled by the adaptor Grb14 in a Chfr-dependent manner." Cell Signal **27**(4): 798-806.

Pokhvisneva, I., É. Léger, M. J. Meaney and P. P. Silveira (2018). "Systematic Overestimation of Reflection Impulsivity in the Information Sampling Task: Age Dependency in Children." Biological Psychiatry **83**(2): e33-e34.

Price, A. L., N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick and D. Reich (2006). "Principal components analysis corrects for stratification in genome-wide association studies." Nat Genet **38**(8): 904-909.

Ronaghi, A., M. I. Zibaii, S. Pandamooz, N. Nourzei, F. Motamedi, A. Ahmadiani and L. Dargahi (2019). "Entorhinal cortex stimulation induces dentate gyrus neurogenesis through insulin receptor signaling." Brain Res Bull **144**: 75-84.

Rosenstrom, T., F. A. Torvik, E. Ystrom, N. O. Czajkowski, N. A. Gillespie, S. H. Aggen, R. F. Krueger, K. S. Kendler and T. Reichborn-Kjennerud (2017). "Prediction of alcohol use disorder using personality disorder traits: a twin study." Addiction.

Saccone, S. F., A. L. Hinrichs, N. L. Saccone, G. A. Chase, K. Konvicka, P. A. Madden, N. Breslau, E. O. Johnson, D. Hatsukami, O. Pomerleau, G. E. Swan, A. M. Goate, J. Rutter, S. Bertelsen, L. Fox, D. Fugman, N. G. Martin, G. W. Montgomery, J. C. Wang, D. G. Ballinger, J. P. Rice and L. J. Bierut (2007). "Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs." Hum Mol Genet **16**(1): 36-49.

Saito, R., M. E. Smoot, K. Ono, J. Ruscheinski, P. L. Wang, S. Lotia, A. R. Pico, G. D. Bader and T. Ideker (2012). "A travel guide to Cytoscape plugins." Nat Methods **9**(11): 1069-1076.

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Scott, R. A., V. Lagou, R. P. Welch, E. Wheeler, M. E. Montasser, J. Luan, R. Magi, R. J. Strawbridge, E. Rehnberg, S. Gustafsson, S. Kanoni, L. J. Rasmussen-Torvik, L. Yengo, C. Lecoeur, D. Shungin, S. Sanna, C. Sidore, P. C. Johnson, J. W. Jukema, T. Johnson, A. Mahajan, N. Verweij, G. Thorleifsson, J. J. Hottenga, S. Shah, A. V. Smith, B. Sennblad, C. Gieger, P. Salo, M. Perola, N. J. Timpson, D. M. Evans, B. S. Pourcain, Y. Wu, J. S. Andrews, J. Hui, L. F. Bielak, W. Zhao, M. Horikoshi, P. Navarro, A. Isaacs, J. R. O'Connell, K. Stirrups, V. Vitart, C. Hayward, T. Esko, E. Mihailov, R. M. Fraser, T. Fall, B. F. Voight, S. Raychaudhuri, H. Chen, C. M. Lindgren, A. P. Morris, N. W. Rayner, N. Robertson, D. Rybin, C. T. Liu, J. S. Beckmann, S. M. Willems, P. S. Chines, A. U. Jackson, H. M. Kang, H. M. Stringham, K. Song, T. Tanaka, J. F. Peden, A. Goel, A. A. Hicks, P. An, M. Muller-Nurasyid, A. Franco-Cereceda, L. Folkersen, L. Marullo, H. Jansen, A. J. Oldehinkel, M. Bruinenberg, J. S. Pankow, K. E. North, N. G. Forouhi, R. J. Loos, S. Edkins, T. V. Varga, G. Hallmans, H. Oksa, M. Antonella, R. Nagaraja, S. Trompet, I. Ford, S. J. Bakker, A. Kong, M. Kumari, B. Gigante, C. Herder, P. B. Munroe, M. Caulfield, J. Antti, M. Mangino, K. Small, I. Miljkovic, Y. Liu, M. Atalay, W. Kiess, A. L. James, F. Rivadeneira, A. G. Uitterlinden, C. N. Palmer, A. S. Doney, G. Willemsen, J. H. Smit, S. Campbell, O. Polasek, L. L. Bonnycastle, S. Hercberg, M. Dimitriou, J. L. Bolton, G. R. Fowkes, P. Kovacs, J. Lindstrom, T. Zemunik, S. Bandinelli, S. H. Wild, H. V. Basart, W. Rathmann, H. Grallert, D. I. G. Replication, C. Meta-analysis, W. Maerz, M. E. Kleber, B. O. Boehm, A. Peters, P. P. Pramstaller, M. A. Province, I. B. Borecki, N. D. Hastie, I. Rudan, H. Campbell, H. Watkins, M. Farrall, M. Stumvoll, L. Ferrucci, D. M. Waterworth, R. N. Bergman, F. S. Collins, J. Tuomilehto, R. M. Watanabe, E. J. de Geus, B. W. Penninx, A. Hofman, B. A. Oostra, B. M. Psaty, P. Vollenweider, J. F. Wilson, A. F. Wright, G. K. Hovingh, A. Metspalu, M. Uusitupa, P. K. Magnusson, K. O. Kyvik, J. Kaprio, J. F. Price, G. V. Dedoussis, P. Deloukas, P. Meneton, L. Lind, M. Boehnke, A. R. Shuldiner, C. M. van Duijn, A. D. Morris, A. Toenjes, P. A. Peyser, J. P. Beilby, A. Korner, J. Kuusisto, M. Laakso, S. R. Bornstein, P. E. Schwarz, T. A. Lakka, R. Rauramaa, L. S. Adair, G. D. Smith, T. D. Spector, T. Illig, U. de Faire, A. Hamsten, V. Gudnason, M. Kivimaki, A. Hingorani, S. M. Keinanen-Kiukaanniemi, T. E. Saaristo, D. I. Boomsma, K. Stefansson, P. van der Harst, J. Dupuis, N. L. Pedersen, N. Sattar, T. B. Harris, F. Cucca, S. Ripatti, V. Salomaa, K. L. Mohlke, B. Balkau, P. Froguel, A. Pouta, M. R. Jarvelin, N. J. Wareham, N. Bouatia-Naji, M. I. McCarthy, P. W. Franks, J. B. Meigs, T. M. Teslovich, J. C. Florez, C. Langenberg, E. Ingelsson, I. Prokopenko and I. Barroso (2012). "Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways." Nat Genet **44**(9): 991-1005.

Sihvola, E., R. J. Rose, D. M. Dick, T. Korhonen, L. Pulkkinen, A. Raevuori, M. Marttunen and J. Kaprio (2011). "Prospective relationships of ADHD symptoms with developing substance use in a population-derived sample." Psychol Med **41**(12): 2615-2623.

Silveira, P. P., M. Agranonik, H. Faras, A. K. Portella, M. J. Meaney, R. D. Levitan, V. Maternal Adversity and T. Neurodevelopment Study (2012). "Preliminary evidence for an impulsivity-based thrifty eating phenotype." Pediatr Res **71**(3): 293-298.

Silveira, P. P., I. Pokhvisneva, H. Gaudreau, A. Rifkin-Graboi, B. F. P. Broekman, M. Steiner, R. Levitan, C. Parent, J. Diorio and M. J. Meaney (2018). "Birth weight and catch up growth are associated with childhood impulsivity in two independent cohorts." Sci Rep **8**(1): 13705.

Snowdon, D. A., L. H. Greiner and W. R. Markesbery (2000). "Linguistic ability in early life and the neuropathology of Alzheimer's disease and cerebrovascular disease. Findings from the Nun Study." Ann N Y Acad Sci **903**: 34-38.

bioRxiv preprint doi: https://doi.org/10.1101/289983; this version posted January 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Talbot, K., H. Y. Wang, H. Kazi, L. Y. Han, K. P. Bakshi, A. Stucky, R. L. Fuino, K. R. Kawaguchi, A. J. Samoyedny, R. S. Wilson, Z. Arvanitakis, J. A. Schneider, B. A. Wolf, D. A. Bennett, J. Q. Trojanowski and S. E. Arnold (2012). "Demonstrated brain insulin resistance in Alzheimer's disease patients is associated with IGF-1 resistance, IRS-1 dysregulation, and cognitive decline." J Clin Invest **122**(4): 1316-1338.

Tobacco and C. Genetics (2010). "Genome-wide meta-analyses identify multiple loci associated with smoking behavior." Nat Genet **42**(5): 441-447.

Vasiliadis, H. M., F. B. Diallo, L. Rochette, M. Smith, D. Langille, E. Lin, S. Kisely, E. Fombonne, A. H. Thompson, J. Renaud and A. Lesage (2017). "Temporal Trends in the Prevalence and Incidence of Diagnosed ADHD in Children and Young Adults between 1999 and 2012 in Canada: A Data Linkage Study." Can J Psychiatry: 706743717714468.

Warde-Farley, D., S. L. Donaldson, O. Comes, K. Zuberi, R. Badrawi, P. Chao, M. Franz, C. Grouios, F. Kazi, C. T. Lopes, A. Maitland, S. Mostafavi, J. Montojo, Q. Shao, G. Wright, G. D. Bader and Q. Morris (2010). "The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function." Nucleic Acids Res **38**(Web Server issue): W214-220.

Wray, N. R. and M. E. Goddard (2010). "Multi-locus models of genetic risk of disease." Genome Med **2**(2): 10.

Wray, N. R., S. H. Lee, D. Mehta, A. A. Vinkhuyzen, F. Dudbridge and C. M. Middeldorp (2014). "Research review: Polygenic methods and their application to psychiatric traits." J Child Psychol Psychiatry **55**(10): 1068-1087.

Wray, N. R. and R. Maier (2014). "Genetic Basis of Complex Genetic Disease: The Contribution of Disease Heterogeneity to Missing Heritability." Current Epidemiology Reports **1**(4): 220-227.

Zheng, J., A. M. Erzurumluoglu, B. L. Elsworth, J. P. Kemp, L. Howe, P. C. Haycock, G. Hemani, K. Tansey, C. Laurin, G. Early, C. Lifecourse Epidemiology Eczema, B. S. Pourcain, N. M. Warrington, H. K. Finucane, A. L. Price, B. K. Bulik-Sullivan, V. Anttila, L. Paternoster, T. R. Gaunt, D. M. Evans and B. M. Neale (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**(2): 272-279.

Figure 1

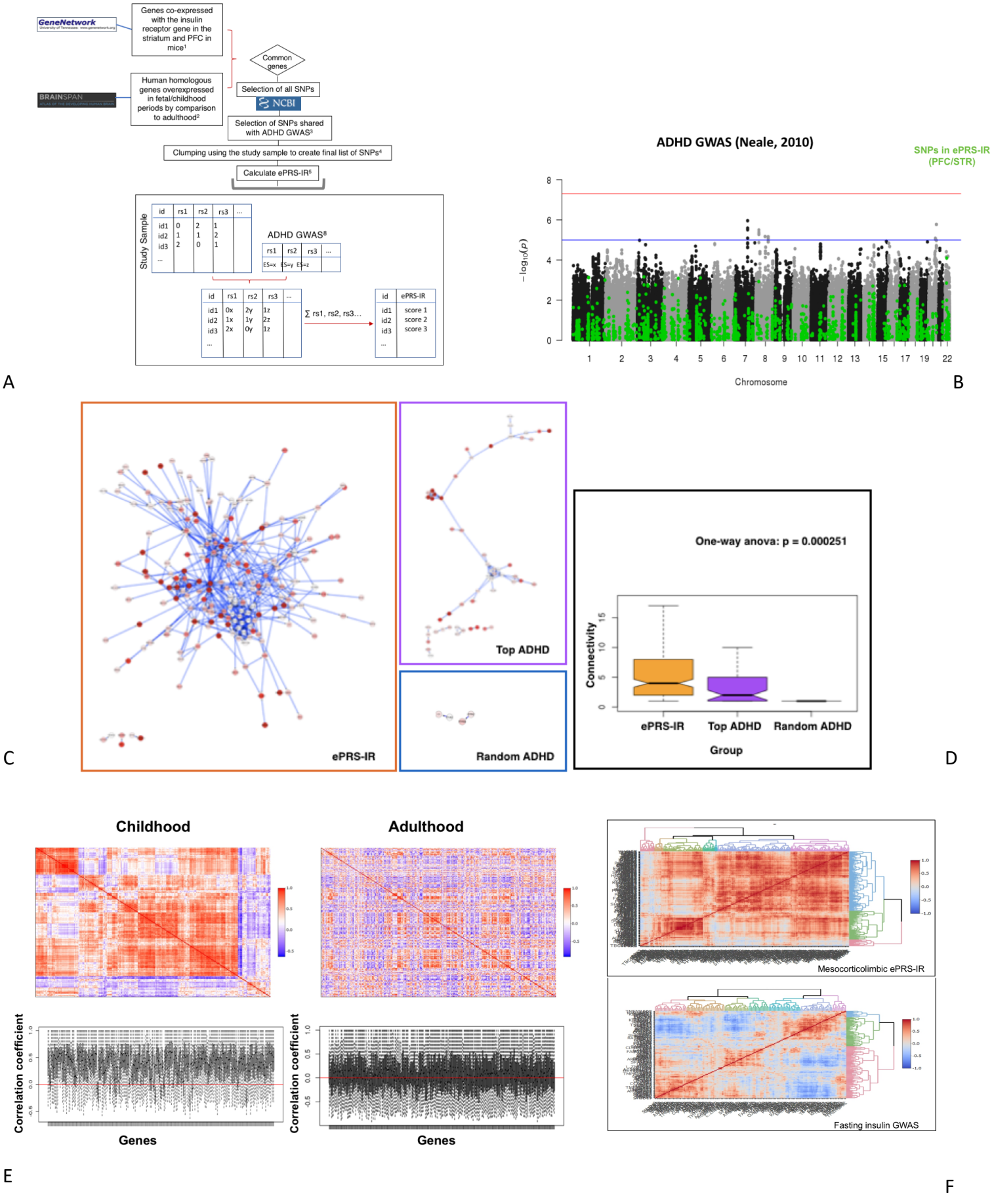
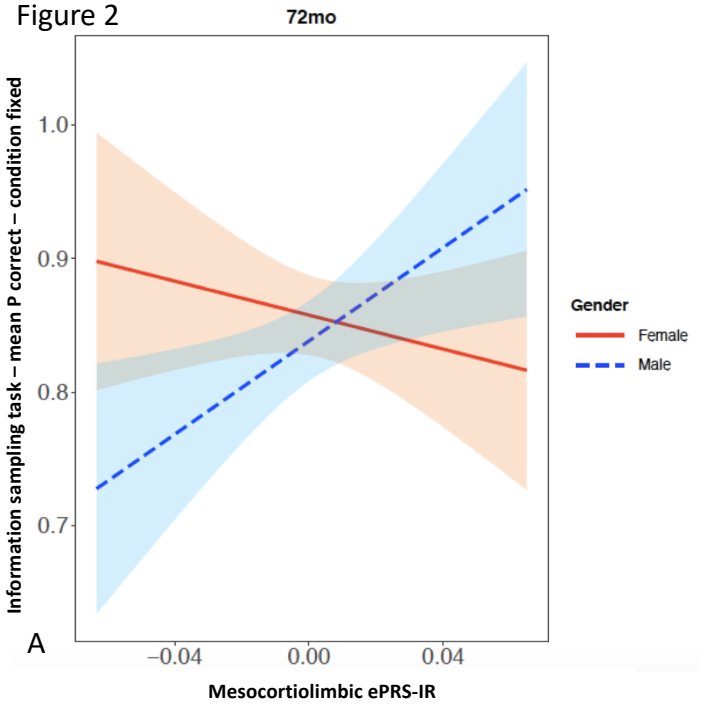
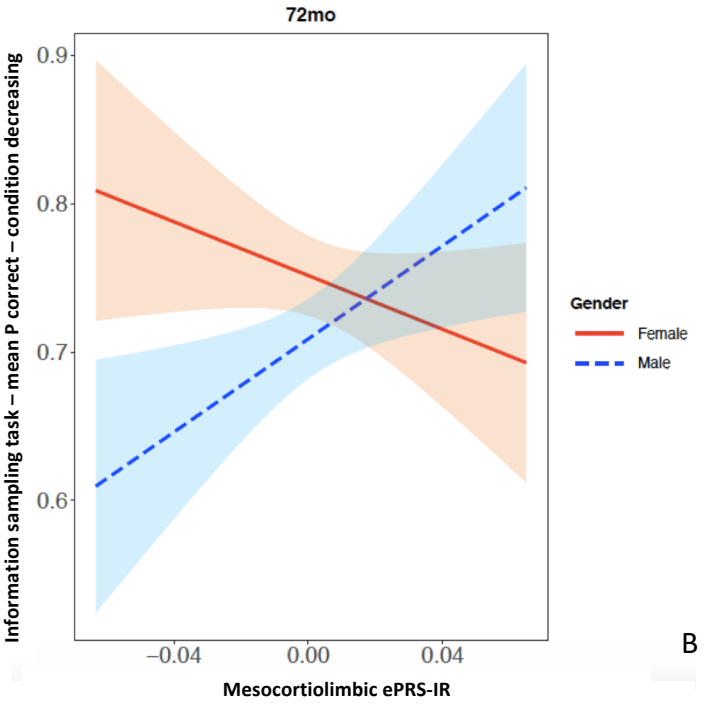

GeneNetwork
Genes co-expressed with the insulin receptor gene in the striatum and PFC in mice[1]
BRAINSPAN
Human homologous genes overexpressed in fetal/childhood periods by comparison to adulthood[2]
Common genes
Selection of all SNPs
NCBI
Selection of SNPs shared with ADHD GWAS[3]
Clumping using the study sample to create final list of SNPs[4]
Calculate ePRS-IR[5]
Study Sample
ADHD GWAS[8]
∑ rs1, rs2, rs3…
ADHD GWAS (Neale, 2010)
SNPs in ePRS-IR (PFC/STR)
−log10(p)
Chromosome
A
B
ePRS-IR
Top ADHD
Random ADHD
One-way anova: p = 0.000251
Connectivity
ePRS-IR
Top ADHD
Random ADHD
Group
C
D
Childhood
Adulthood
Correlation coefficient
Genes
Mesocorticolimbic ePRS-IR
Fasting insulin GWAS
E
F

Figure 2



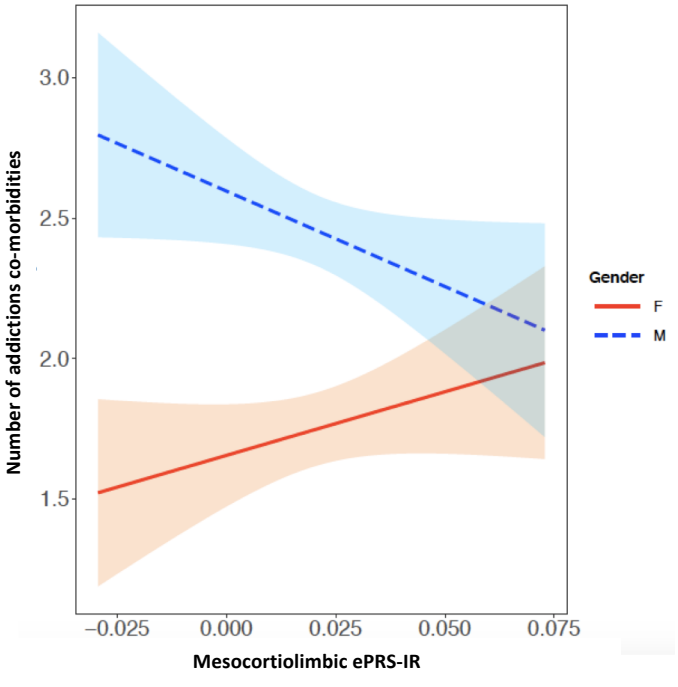
A



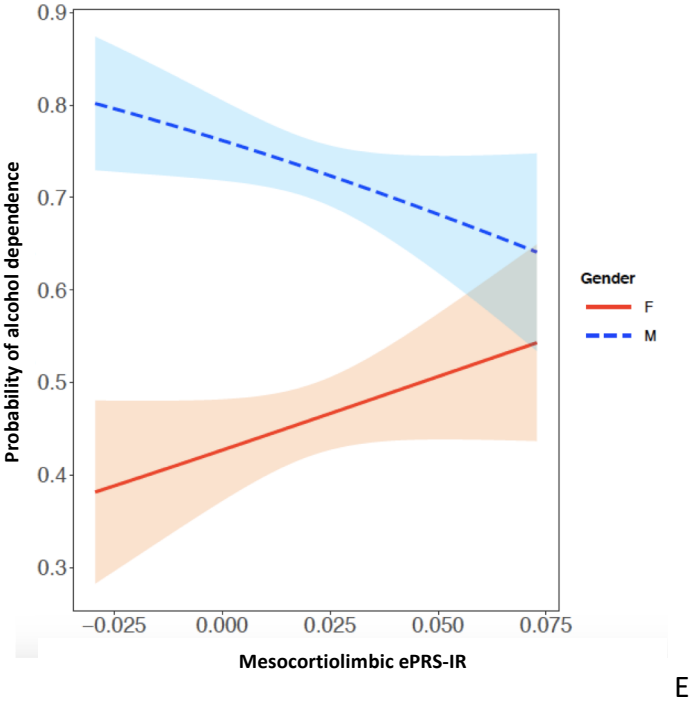
B

PRS	Number of SNPs	P-value	Beta estimate
ePRS_IR (STR+PFC)	1897	0.006	2.479
Random ePRS	1854	0.6333	0.52802
$PRS_{ADHD2010}$	1897	0.6281	-0.12687
$PRS_{ADHD2019}$	1897	0.21644	-33.73129
$PRS_{Smoking}$	1897	0.3315	-9.87592

C



D



E

Figure 3

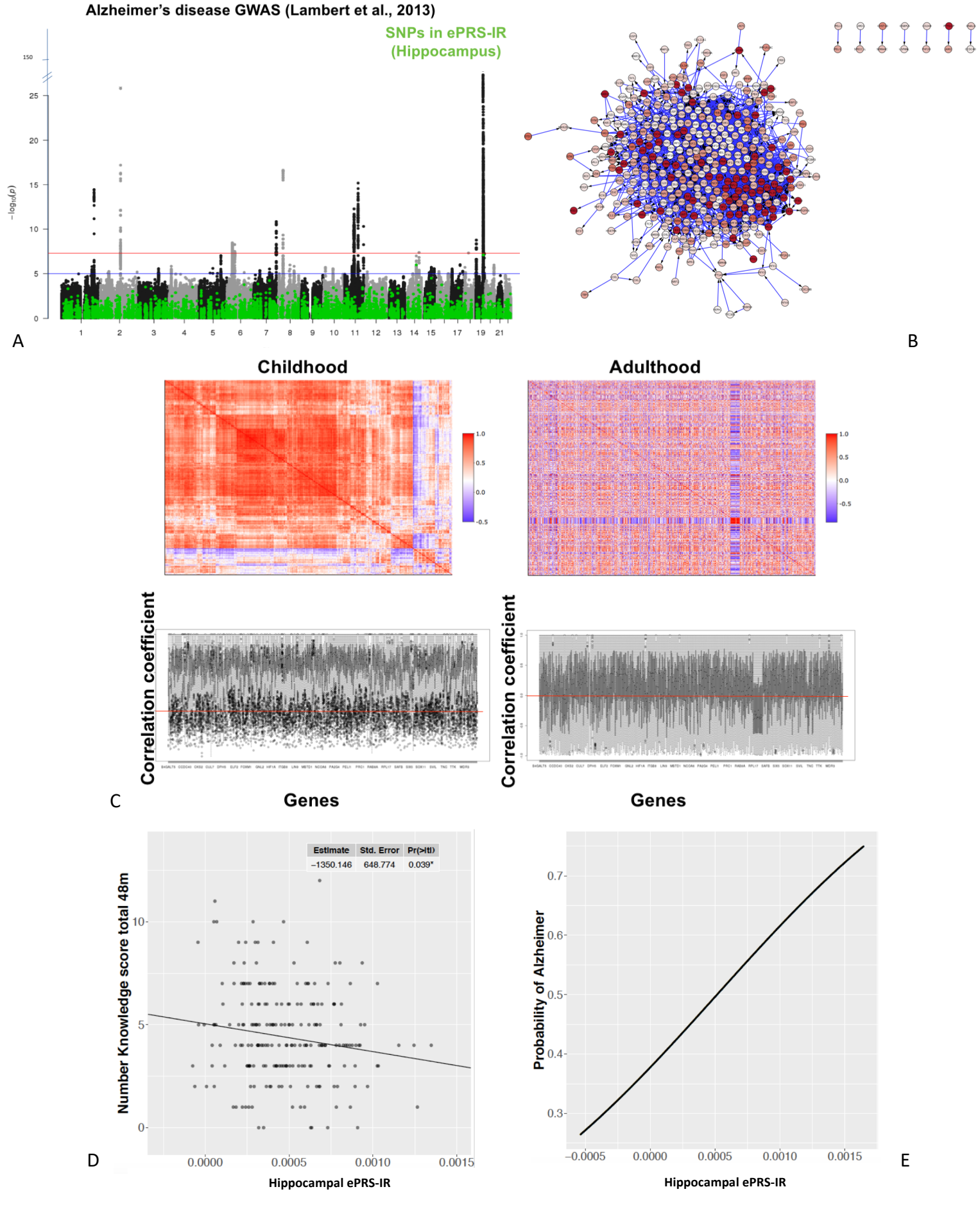
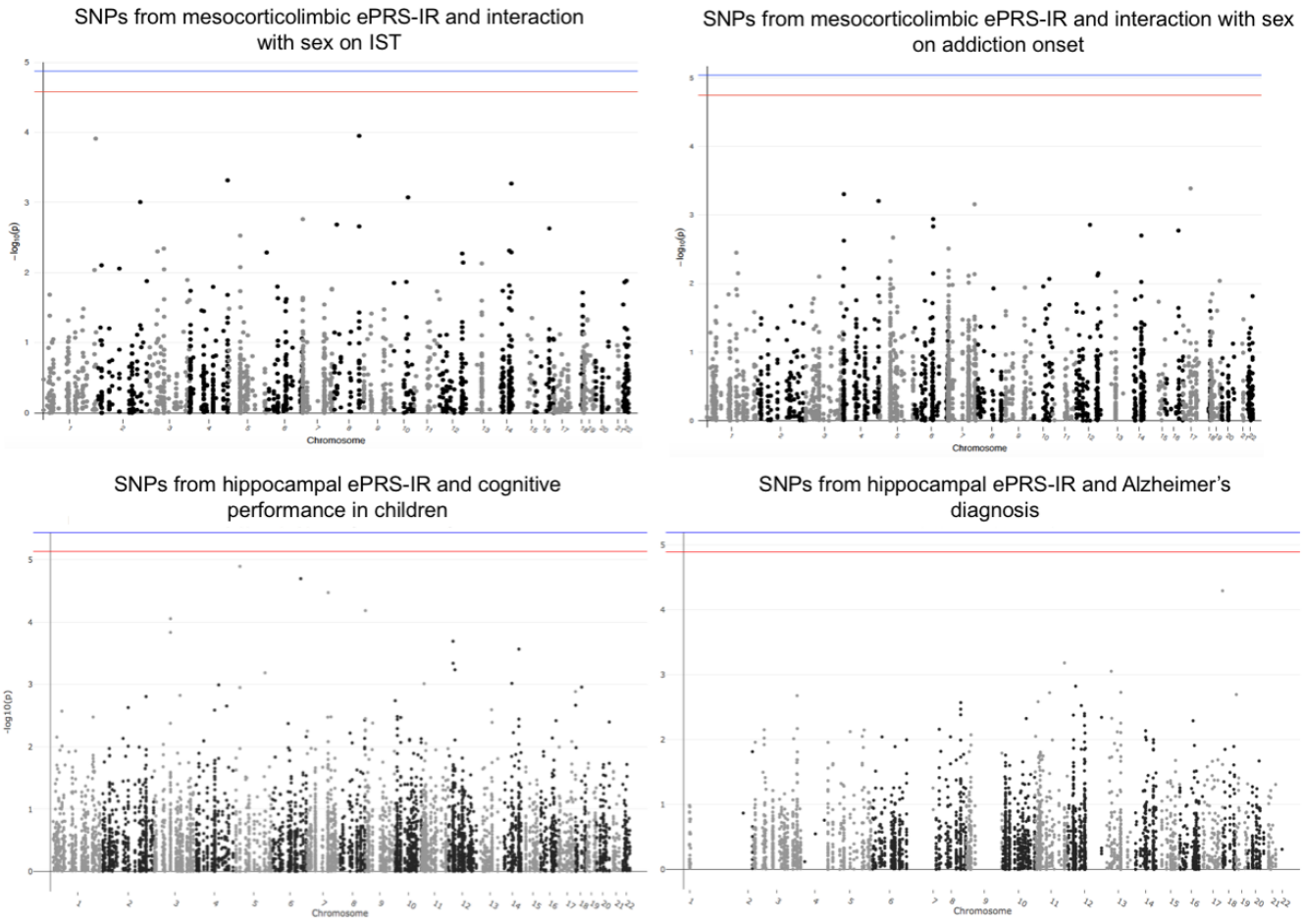

Alzheimer's disease GWAS (Lambert et al., 2013)
SNPs in ePRS-IR
(Hippocampus)
−log10(p)
A
B
Childhood
Adulthood
Correlation coefficient
Genes
C
Estimate
Std. Error
Pr(>|t|)
−1350.146
648.774
0.039*
Number Knowledge score total 48m
Hippocampal ePRS-IR
D
Probability of Alzheimer
Hippocampal ePRS-IR
E

Figure 4



A

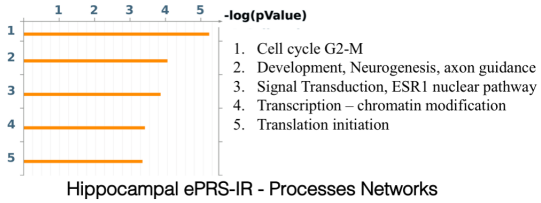
Hippocampal ePRS-IR

Mesocorticolimbic ePRS-IR

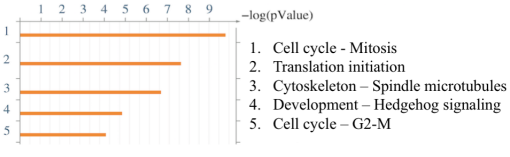
Quartile	1	2	3	4
1	22	14	12	17
2	17	13	18	17
3	16	18	13	18
4	10	20	22	13

B

Mesocorticolimbic ePRS-IR - Processes Networks



Hippocampal ePRS-IR - Processes Networks



C

Mesocorticolimbic ePRS-IR

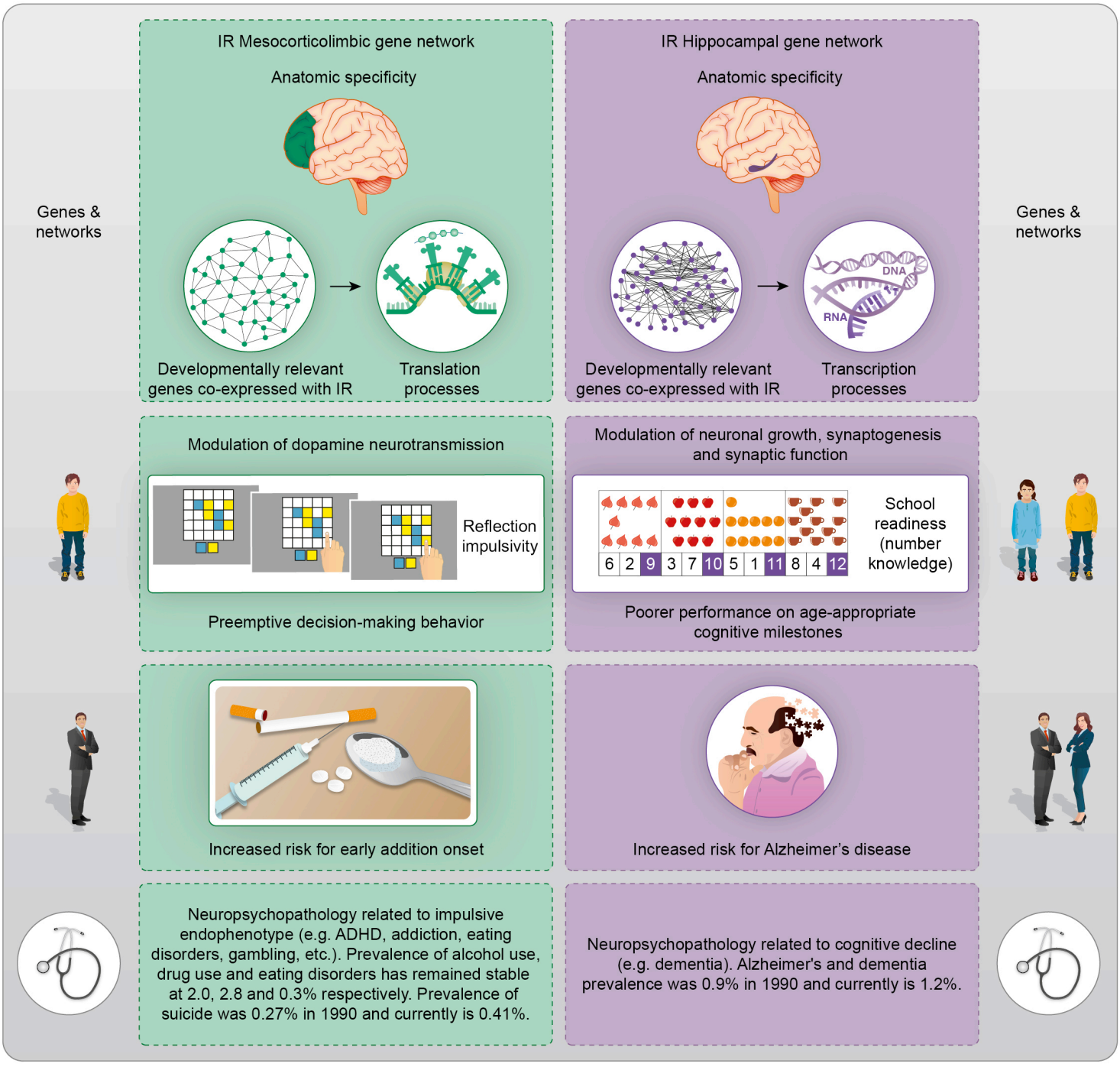
Transcription factor	Ratio	P-Value
E2F1	3.33	8.42e-40
ETS1	2.634	1.72e-37
Oct-3/4	3.087	1.19e-32
FOXP3	2.34	5.27e-30
C-Myc	3.516	3.71e-28
SOX2	3.107	1.10e-27
E2F4	4.473	3.47e-24
GATA-2	3.62	3.57e-23
CREB1	2.571	3.71e-21
NANOG	3.016	1.33e-20
N-Myc	3.551	3.34e-20
GABP alpha	2.478	1.60e-19
SP1	3.735	2.36e-16
p53	3.687	3.19e-15
ZFX	2.956	3.33e-14

Hippocampal ePRS-IR FDR=0.05

Transcription factor	Ratio	P-Value
C-Myc	4.4	7.9e-63
CREB1	2.6	2.6e-39
E2F1	5.6	1.2e-33
SOX2*	6.5	3.5e-29
Oct-3/4	4.4	5.0e-28
NANOG	5.6	4.5e-26
SP1	3.6	3.7e-25
E2F4	8.0	7.0e-25
YY1	4.2	3.3e-22
p53	3.7	1.0e-19
CUX1*	3.5	3.0e-17
E2F3	8.9	8.9e-16
FOXMI*	9.9	1.9e-13
ESR1	3.0	3.6e-13
STAT3	3.8	1.0e-12

D

Figure 5


IR Mesocorticolimbic gene network
Anatomic specificity
Developmentally relevant genes co-expressed with IR
Translation processes
IR Hippocampal gene network
Anatomic specificity
DNA
RNA
Developmentally relevant genes co-expressed with IR
Transcription processes
Genes & networks
Genes & networks
Modulation of dopamine neurotransmission
Reflection impulsivity
Preemptive decision-making behavior
Modulation of neuronal growth, synaptogenesis and synaptic function
6 2 9 3 7 10 5 1 11 8 4 12
School readiness (number knowledge)
Poorer performance on age-appropriate cognitive milestones
Increased risk for early addition onset
Increased risk for Alzheimer's disease
Neuropsychopathology related to impulsive endophenotype (e.g. ADHD, addiction, eating disorders, gambling, etc.). Prevalence of alcohol use, drug use and eating disorders has remained stable at 2.0, 2.8 and 0.3% respectively. Prevalence of suicide was 0.27% in 1990 and currently is 0.41%.
Neuropsychopathology related to cognitive decline (e.g. dementia). Alzheimer's and dementia prevalence was 0.9% in 1990 and currently is 1.2%.