1 2

CADM2 is implicated in impulsive personality and numerous other traits by genome- and phenome-wide association studies in humans and mice

Sandra Sanchez-Roige^{*1,2}, Mariela V Jennings¹, Hayley H A Thorpe³, Jazlene E Mallari¹, Lieke 3 C van der Werf¹, Sevim B Bianchi¹, Calvin Lee¹, Travis T Mallard⁴, Samuel A Barnes¹, Jin Yi 4 Wu¹, Amanda M Barkley-Levenson¹, Ely C Boussaty¹, Cedric E Snethlage¹, Danielle Schafer¹, 5 Zeljana Babic¹, Boyer D Winters⁵, Katherine E Watters^{6,7}, Thomas Biederer⁶, on behalf of the 6 23andMe Research Team⁸, James Mackillop⁹, David N Stephens¹⁰, Sarah L Elson⁸, Pierre 7 Fontanillas⁸, Jibran Y Khokhar^{3,11}, Jared W Young¹, Abraham A Palmer^{*1,12} 8

9

10 ¹Department of Psychiatry, University of California San Diego, San Diego, CA, USA, 11 ²Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA. ³Department of Biomedical Sciences, Ontario Veterinary College, University of Guelph, Ontario, 12 Canada, ⁴Psychiatric and Neurodevelopmental Genetics Unit. Center for Genomic Medicine. 13 Massachusetts General Hospital, Boston, MA, 02114, USA. ⁵Department of Psychology, 14 University of Guelph, Guelph, ON, Canada. ⁶Department of Neuroscience, Tufts University 15 School of Medicine, Boston, MA, USA. ⁷Department of Neurology, Yale School of Medicine, 16 New Haven, CT, USA. ⁸23andMe, Inc., Sunnyvale, CA, USA. ⁹Peter Boris Centre for Addictions 17 Research, McMaster University and St. Joseph's Healthcare Hamilton, Hamilton, ON, Canada 18 and Homewood Research Institute, Guelph, ON, Canada, ¹⁰Laboratory of Behavioural and 19 Clinical Neuroscience, School of Psychology, University of Sussex, Brighton, UK.¹¹ Schulich 20 School of Medicine and Dentistry, Western University, London, ON. ¹²Institute for Genomic 21 22 Medicine, University of California San Diego, San Diego, CA, USA.

23

24 *Co-corresponding Authors: Sandra Sanchez-Roige, sanchezroige@ucsd.edu, Abraham A.

Palmer, aap@ucsd.edu 25

Keywords: Cadm2, UPPS-P, BIS, impulsivity, GWAS, PheWAS, animal models, cross-species 26 27 translation

28 ABSTRACT

29 Impulsivity is a multidimensional heritable phenotype that broadly refers to the tendency to act

- 30 prematurely and is associated with multiple forms of psychopathology, including substance use
- 31 disorders. We performed genome-wide association studies (**GWAS**) of eight impulsive
- 32 personality traits from the Barratt Impulsiveness Scale and the short UPPS-P Impulsive
- 33 Personality Scale (N=123,509-133,517 23andMe research participants of European ancestry),
- 34 and a measure of Drug Experimentation (N=130,684). Because these GWAS implicated the
- 35 gene *CADM2*, we next performed single-SNP phenome-wide studies (**PheWAS**) of several of
- 36 the implicated variants in CADM2 in a multi-ancestral 23andMe cohort (N=3,229,317, European;
- N=579,623, Latin American; N=199,663, African American). Finally, we produced *Cadm2*
- 38 mutant mice and used them to perform a Mouse-PheWAS ("MouseWAS") by testing them with a
- 39 battery of relevant behavioral tasks. In humans, impulsive personality traits showed modest
- 40 chip-heritability (~6-11%), and moderate genetic correlations (r_g =.20-.50) with other personality
- 41 traits, and various psychiatric and medical traits. We identified significant associations proximal
- 42 to genes such as *TCF4* and *PTPRF*, and also identified nominal associations proximal to *DRD2*
- 43 and CRHR1. PheWAS for CADM2 variants identified associations with 378 traits in European
- 44 participants, and 47 traits in Latin American participants, replicating associations with risky
- 45 behaviors, cognition and BMI, and revealing novel associations including allergies, anxiety,
- 46 irritable bowel syndrome, and migraine. Our MouseWAS recapitulated some of the associations
- 47 found in humans, including impulsivity, cognition, and BMI. Our results further delineate the role

48 of *CADM*2 in impulsivity and numerous other psychiatric and somatic traits across ancestries

49 and species.

50

51 INTRODUCTION

52 Impulsivity is a multifaceted psychological construct that has been broadly defined as 53 thoughts or actions that are "poorly conceived, prematurely expressed, unduly risky or 54 inappropriate to the situation, and that often result in undesirable consequences" [1]. Impulsivity 55 has been repeatedly associated with numerous psychiatric diseases, including ADHD and 56 substance use disorders [2, 3]. We previously performed genome-wide association studies 57 (GWAS) of impulsive personality traits (n=21,806-22,861) using two of the most widely used 58 impulsivity guestionnaires, the Barratt Impulsiveness Scale (BIS-11; 3 traits) and the Impulsive 59 Personality Scale (UPPS-P; 5 traits), as well as a measure of Drug Experimentation [4]. These 60 traits were partially genetically correlated, suggesting that each impulsivity domain is governed 61 by overlapping but distinct biological mechanisms [4, 5]. Our work also identified significant 62 denetic correlations between impulsivity and numerous psychiatric and substance use traits, in 63 line with the NIMH Research Domain Criteria (**RDoC**), proposing impulsivity as a 64 transdiagnostic endophenotype for psychopathology [6].

65 The cell adhesion molecule 2 (CADM2) gene, which was the most robustly implicated 66 gene in our prior GWAS of impulsivity [4], has also been extensively implicated in other risky 67 and substance use behaviors [7]. CADM2 mediates synaptic plasticity and is enriched in the 68 frontal cortex and striatum, which are regions that regulate reward and inhibitory processes. We 69 and others have implicated this gene in traits that may underlie disinhibition in humans, 70 supporting the observed genetic correlations between impulsivity and personality [8], 71 educational attainment [9], cognition [10], risk-taking [11], substance use [4, 10, 12–15], 72 externalizing psychopathology [16], neurodevelopmental disorders [17, 18], physical activity 73 [19], reproductive health [20, 21], metabolic traits [22], and BMI [23], among others (see GWAS 74 Catalog www.ebi.ac.uk/gwas/). Cadm2 knockout mice have previously been assessed for body 75 weight and energy homeostasis [24] but have never been behaviorally characterized for 76 measures of impulsivity or related behaviors. 77 Here, we took three approaches to elucidate genetic factors related to impulsivity. First, we collaborated with 23andMe, Inc., to extend upon our earlier GWAS of impulsivity [4] by 78 79 increasing our sample size approximately 6-fold (n=123,509-133,517). Second, we performed

single-SNP phenome-wide studies (**PheWAS**) of the 5 single nucleotide polymorphisms (**SNPs**)

in and around *CADM*² that have been most strongly implicated by the current and prior GWAS.

PheWAS were conducted in three ancestral groups (N=3,229,317, European; N=579,623, Latin

83 American; N=199,663, African American) from the 23andMe research cohort, examining close

- 84 to 1,300 traits, most with no published GWAS. Finally, we performed a mouse-PheWAS
- 85 ("MouseWAS") by creating and phenotyping mice harboring a *Cadm*² mutant allele in a broad
- 86 battery of behavioral tasks that included analogous human measures of risk-taking and
- 87 impulsivity, substance use, cognition and BMI.

88 MATERIALS AND METHODS

89 Human Studies

90 GWAS cohort and phenotypes

91 We analyzed data from a cohort of up to 133,517 male and female research participants 92 of European ancestry, a subset of which were analyzed in our prior publications [4, 13, 13, 25, 93 26]. All participants were drawn from the research participant base of 23andMe, Inc., a direct-to-94 consumer genetics company, and were not compensated for their participation. Participants 95 provided informed consent and participated in the research online, under a protocol approved by the external AAHRPP-accredited IRB, Ethical & Independent Review Services 96 97 (www.eandireview.com). During 4 months in 2015 and 14 months from 2018-2020, participants 98 responded to a survey that included up to 139 questions pertaining to aspects of impulsivity and 99 substance use and misuse. To measure impulsive personality, we used five subscales from the 100 UPPS-P ([27, 28]; a 20-item that measures (lack of) Premeditation, (lack of) Perseverance, 101 Positive Urgency, Negative Urgency, and Sensation Seeking; **Table S1**). We also administered 102 the BIS-11 ([29]; a 30-item questionnaire that measures Attentional, Motor, and Nonplanning 103 impulsiveness; **Table S1**). Lastly, we measured Drug Experimentation, defined as the number 104 of substances an individual has used (adapted from the PhenX toolkit [30]; Table S1). We 105 scored UPPS-P. BIS-11 and Drug Experimentation as previously described [4]. We used 106 quantile normalization, since some scores were not normally distributed (Figures S1-3). Only 107 individuals identified as European ancestry based on empirical genotype data [31] were 108 included in this study. Basic demographic information about this sample is presented in Table 109 **S2**. We used Pearson correlation coefficients (*r*) to measure the phenotypic relationships 110 between impulsivity subscales and demographics.

111 Genome-wide association and secondary analyses

DNA extraction and genotyping were performed on saliva samples by CLIA-certified and CAP-accredited clinical laboratories of Laboratory Corporation of America. Quality control, imputation, and genome-wide analyses were performed by 23andMe (**Table S3**; [32, 33]). 23andMe's analysis pipeline performs linear regression assuming an additive model for allelic effects (**Supplementary Material**). Covariates included age (inverse-normal transformed), sex, the top five principal genotype components, and indicator variables for genotyping platforms. *p*values were corrected for genomic control. We examined genotype*sex interactions for

suggestive loci. In addition, for the top loci, we examined African American and Latin American23andMe research participants who had responded to the same survey.

- 121 We used the FUMA web-based platform (version 1.3.6a) and MAGMA v1.08 [34, 35] to 122 explore the functional consequences of the GWAS loci and to conduct gene-based analyses.
- 123 We used LDSC [36] to calculate genetic correlations (r_g) between UPPS-P, BIS and
- 124 Drug Experimentation, and 96 selected traits informed by prior literature.

125 **Phenome-wide association scan (PheWAS) in 23andMe**

126 We performed single-SNP PheWAS for 5 CADM2 SNPs (rs993137, rs62263923, 127 rs11708632, rs818219, rs6803322) using up to 1,291 well-curated self-reported phenotypes 128 from a separate cohort of 23andMe research participants of European (N≤3.229.317), Latin 129 American (N≤579,623) and African American (N≤199,663) ancestries. We excluded traits with 130 <1,000 responses, based on a prior simulation study for PheWAS power analysis [37]. Ancestry 131 was determined by analyzing local ancestry ([31] Supplementary Material). The variants were 132 selected based on our GWAS results and previous literature (Table S4, and Supplementary 133 Material). Genotyped and imputed variant statistics in PheWAS are shown in Table S5.

An overview of the data collection process has been previously described [38]. All regression analyses were performed using R version 3.2.2. We assumed additive allelic effects and included covariates for age (as determined by participant date of birth), sex, and the top five ancestry-specific principal components. We used a 5% FDR correction for multiple testing.

138 MouseWAS

139 Subjects, behavioral characterization, and statistical analyses

140 Our Cadm2 mutant mice were produced at the University of California San Diego, 141 Moores Cancer Center, Transgenic Mouse Core. We used the JM8.N4 cryosperm line 142 (CSD70565 KOMP), which carries a floxed null allele in the *Cadm2* gene (Figure S30), on a 143 C57BL/6N background. We crossed the floxed null allele line with a constitutive CRE driver line 144 (Stock# 014094: The Jackson Laboratory), vielding a global constitutive null allele. We used a 145 heterozygous x heterozygous (**HET**) breeding scheme, which produced homozygous (**HOM**) 146 mutant *Cadm2* mice and their HET and wildtype (**WT**) littermates. Mice were genotyped using 147 allele-specific polymerase chain reaction on ear notch tissue followed by gel electrophoresis 148 [39]. CADM2 protein expression levels were quantified by western blotting (Figure S31).

149 Five separate cohorts of male and female mice were used for these studies. See 150 Supplementary Material for a more detailed description of the tasks and analyses of main 151 variables. Procedures were approved by the University of California San Diego Institutional 152 Animal Care and Use Committee. The UCSD animal facility meets all federal and state 153 requirements for animal care and was approved by the American Association for Accreditation 154 of Laboratory Animal Care. Procedures from cohort 2 were conducted in accordance with the 155 Canadian Council on Animal Care and were approved by the University of Guelph Institutional 156 Animal Care and Use Committee.

157

158 **RESULTS**

159 Genome-wide association analyses and secondary analyses

Self-reported impulsivity and Drug Experimentation scores are shown in Table S6. We
found that ~6-11% of the phenotypic variation of these traits can be explained by common
variants (Table S7). We identified 21 genome-wide significant associations (*p*<5.0E-08) for
UPPS-P (5 traits), BIS (3 traits), and Drug Experimentation (Figure 1; Figures S4-21; Table
S8). Although we tested 9 traits, in keeping with the standards of the field, we did not adjust the
significance threshold. We also detected several nominal associations (*p*<1.0E-06, Table S8);
we discuss some of them in the Supplementary Material.

167 GWAS of UPPS-P

Premeditation. We detected one significant hit (rs2958162, *p*=2.50E-10), located on
chromosome 18 in the *TCF4* gene, which encodes a helix-loop-helix transcription factor and is
widely expressed throughout the body and during development. Polymorphisms in *TCF4* have
been associated with risk-taking and adventurousness [15], alcohol consumption [40],
schizophrenia [41], depression [42, 43], and neuroticism [44, 45] (**Table S9**); *TCF4* is also a
non-GWAS candidate gene for other psychiatric and neurological conditions [46].

Perseverance. We detected one significant association (rs5943997, *p*=1.50E-8) in the
 POLA1 gene on the X chromosome. *POLA1* has been related to blood traits [46] and
 neurodevelopmental disorders [47], but its association with impulsivity is novel.

Positive Urgency. We identified one significant hit (rs143987963, *p*=4.30E-08) on
chromosome 12, near the genes *MDM1* and *RAP1B*; however, inspection of the locus zoom plot
(Figure S9) does not support a robust association.

180 Negative Urgency. We detected three significant hits: rs4840542 (p=1.60E-09), on 181 chromosome 8, in the XKR6 gene; rs5008475 (p=4.90E-09), on chromosome 5, near 182 TMEM161B and MEF2C; and rs7829975, on chromosome 8, near SGK223 and CLDN23 183 (p=5.00E-09). Variants in strong LD with rs4840542 and rs7829975 are highly pleiotropic, and 184 have been previously associated with several traits (**Table S9**), including body mass index 185 (BMI) [48, 49], neuroticism [50, 51], depression [52], blood pressure, and alcohol consumption 186 [53]. XKR6 was also implicated in a recent GWAS of externalizing [16], and a GWAS of anxiety 187 and depression [52].

188 Sensation Seeking. We detected 5 significant associations. First, we again observed a previously reported [4] association with a SNP near CADM2 (rs11288859, p=2.10E-09) on 189 190 chromosome 3. We also detected an association with a SNP in TCF4 (rs2958178, p=3.80E-12). 191 We identified a significant hit in CACNA2D1 (rs38547, p=2.10E-08) on chromosome 18. 192 CACNA2D1 has been previously associated with feeling nervous [50], and levels of sex 193 hormone-binding globulin [54]. Furthermore, we found a significant association (rs1605379, 194 p=3.80E-08) on chromosome 16, near CYLD and SALL1. SNPs in strong LD with rs1605379 195 have been previously identified for risk-taking, adventurousness, and smoking initiation (Table 196 **S9**). Lastly, we found a significant association (rs12600879, p=4.10E-08) on chromosome 17, 197 near TBX21 and OSBPL7. Variants in strong LD with rs12600879 have been associated with 198 BMI [55], but the finding in relation to impulsivity is novel.

199 GWAS of BIS-11

Attentional. We identified one significant association (rs10196237, *p*=1.10E-08) on
 chromosome 2, near the genes *SPHKAP and PID1*. *SPHKAP* has been previously associated
 with educational attainment [9], but the association with impulsivity is novel.

203 *Motor.* We detected one significant association near *CADM2* (rs35614735, p=3.20E-11). 204 We also identified an association (rs111502401, p=2.00E-08), on chromosome 19, near the 205 genes *ZNF229* and *ZNF180*; however, inspection of the regional association is not supportive of 206 a strong association (**Figure S17**).

Nonplanning. We detected 2 variants: rs35614735 (p=4.70E-12) near *CADM*2, which was the same SNP identified for Motor impulsivity; and rs6872863 (p=1.20E-08) in the gene *ELOVL7* on chromosome 5. Variants in strong LD with rs6872863 have been reported for a variety of traits including educational attainment, mathematical ability [9], household income [56], and brain morphology, such as cortical surface area [57] (**Table S9**). However, there is extensive LD in this region, making the association difficult to interpret.

213 GWAS of Drug Experimentation

214 We previously reported [4] a suggestive association (rs2163971, p=3.00E-07) near the 215 *CADM2* gene. In the present study, we identified a nearby SNP that was genome-wide 216 significant (rs35614735, p=2.80E-15). We also report 4 novel hits (rs951740, p=9.70E-10, 217 *PTPRF* on chromosome 1; rs12713405, p=9.70E-09, *BLC11A* on chromosome 2; rs67660520,

218 p=7.60E-09, CADPS2 on chromosome 7; rs7128648, p=2.50E-09, NCAM1 on chromosome 219 11). Intriguingly, PTPRF has been recently associated with problematic prescription opioid use 220 [25] and opioid use disorder [58], as well as smoking initiation/cessation [59], cognition [60], and 221 educational attainment [9] (Table S9). Variants in strong LD with rs67660520 have been 222 associated with ADHD [61], smoking initiation [59], number of sexual partners [15] and BMI [49] 223 (Table S9). NCAM1 variants have been previously associated with alcohol, cannabis and 224 smoking behaviors [59, 63], mathematical ability [9], and anxiety and depression [52], among 225 other traits.

226 Gene-based analyses

227 Similar to the GWAS results, gene-based analyses using MAGMA identified an 228 association (Bonferroni p<2.53E-06; Table S10) between CADM2 and 6 of the 9 traits 229 examined in this paper: Premeditation, Sensation Seeking (UPPS-P); Attentional, Motor and 230 Nonplanning (BIS-11); and Drug Experimentation. TCF4, which was significantly associated 231 with Premeditation and Sensation Seeking in the GWAS, was significantly associated with these 232 traits in the gene-based analysis. MAPT, which has been previously associated with many traits 233 including multiple alcohol-related behaviors [13], was implicated in Negative Urgency. Lastly, 234 KDM4A, which was recently related to problematic opioid use and interacts with selective 235 serotonin reuptake inhibitors and dopaminergic agents [25], was significantly associated with 236 Drug Experimentation.

237 Phenotypic and genetic correlations

A phenotypic and genetic correlation matrix of all 9 traits is shown in **Figure S22** and **Tables S11-12**. Consistent with the literature and our prior work [4, 5, 64, 65], both phenotypic and genetic inter-correlations among the UPPS-P and BIS subscales were high and positive, with the exception of Sensation Seeking and Perseverance, suggesting that these traits may represent relatively different constructs [5, 13, 64]. Drug experimentation was positively and significantly associated with all impulsive personality traits.

All impulsivity traits were phenotypically associated (*r*=-0.34-0.11) with demographic variables (**Table S12**), impulsivity scores being greater in male and younger research participants, compared to female and older participants; and in participants with higher BMI, lower household income, and fewer years of education, as we previously reported [13].

Figure 2 shows a genetic correlation matrix of BIS, UPPS-P, Drug Experimentation and several other phenotypes (full results in **Table S13**). As anticipated, we found positive moderate to high genetic correlations (r_g =0.25-0.79) between virtually all UPPS-P (except Perseverance and Sensation Seeking) and BIS subscales, and Drug Experimentation, and substance use disorders **Table S13**).

We also observed moderate to strong associations between all impulsive subscales (except UPPS-P Perseverance) and other personality traits, such as risk-taking (r_g =0.15-0.65), neuroticism (r_g =-0.23-0.84), and loneliness (r_g =0.17-0.54), particularly for Positive and Negative Urgency. Extraversion was positively associated with Sensation Seeking (r_g =0.34). Externalizing psychopathology, which represents disorders and behaviors characterized by deficits in inhibition, was strongly associated with all impulsivity facets (r_g =0.28-0.92), except Perseverance.

We also identified positive associations with an array of psychiatric phenotypes, including ADHD (r_g =0.20-0.47), depression (r_g =-0.13-0.47) and anxiety (r_g =-0.38-0.61) disorders, and cross-disorder (r_g =0.12-0.44). The associations were again primarily significant for all except Perseverance and Sensation Seeking. Other disorders showed weaker associations (e.g., schizophrenia, r_g =-0.09-0.15) or were only significantly associated with one impulsivity facet [e.g., anorexia nervosa (Perseverance, r_g =-0.16); bipolar disorder (Motor, r_g =0.22)].

267 Most impulsivity subscales were genetically correlated with lower socioeconomic 268 variables [e.g., educational attainment (r_g =-0.49 to -0.16), income (r_g =-0.38 to -0.16), Townsend 269 index (r_g =0.18-0.58)].

270 Metabolic and medical phenotypes, such as BMI (r_g =0.18-0.28), chronic pain (r_g =0.22-271 0.46), insomnia (r_g =0.20-0.42), and coronary artery disease (r_g =0.18-0.30), were genetically 272 correlated with all impulsive subscales (except Perseverance and Sensation Seeking). We also 273 noted negative genetic associations with parental longevity (r_g =-0.17 to -0.32).

274 PheWAS

To explore the impact of specific variants in and around *CADM2*, we performed singleSNP PheWAS using 5 of the most implicated SNPs, independently, against 1,291 traits (Figure
The list of PheWAS association results using the 23andMe cohort after 5% FDR correction is

available in Tables S14 (summary), S15 (Europeans), S16 (Latin American) and S17 (African
Americans).

In European cohorts, *CADM2* variants had been previously identified to be significantly associated with numerous traits (**Table S18**). Most SNPs were highly correlated (R²>0.1) and tagged similar traits (**Figure S23**), but the overlap was incomplete (**Figure S24** and **Table S19**). rs993137, located at 85,449,885 bp on chromosome 3, showed the highest number of associations (378), which we describe below.

We replicated all previously known associations in 23andMe participants of European 285 286 ancestry, identifying signals across all categories tested (Table S15). These included negative 287 associations with risky behavior (e.g., lower risk for adventurousness [β =-0.05, p=1.33E-08], 288 risk-taking tendencies [β =-0.02, p=1.13E-07]) and substance use behaviors (e.g., lower risk for 289 alcohol consumption [β =-0.03, p=2.05E-09] and tobacco initiation (β =-0.02, p=3.66E-12; but see 290 packs per day, β =0.01, p=1.05E-03), as well as negative associations with psychiatric disorders 291 characterized by deficits in impulsivity, such as lower risk for ADHD (β =-0.05, p=2.17E-41). 292 Furthermore, we found positive associations with educational outcomes (e.g., higher 293 educational attainment (β =0.03, p=1.67E-12). Novel findings included positive associations with 294 allergies (β =0.04, p=4.51E-03), anxiety (e.g., panic [β =0.02, p=6.82E-08]), and medical 295 conditions (e.g., IBS [β =0.02, p=8.89E-07]), anemia (β =0.01, p=8.30E-74), hepatitis C (β =-0.06, 296 p=8.36E-10). Intriguingly, we also detected positive associations with pain phenotypes (β =0.02, 297 p=8.37E-12) and a need for a higher dose of pain medication (β =0.01, p=1.02E-06).

For the overlapping phenotypes, UK Biobank PheWAS results [66] largely supported the 23andMe PheWAS findings (except for smoking behaviors). For example, we identified associations with dietary traits (e.g., daily fruit and vegetable intake (β =-0.01, *p*=4.23E-11), pastry frequency (β =0.01, *p*=7.36E-06), sleep quality (β =-0.01, *p*=2.53E-03), and number of pregnancies (β =-0.01, *p*=7.69E-04), among others (**Table S15**, [12]).

303 For the PheWAS of the Latin American cohort, 47 traits were significantly associated 304 with *CADM*2 variants (**Table S16**). The highest number of associations were again observed for 305 rs993137 [67], which are described below. Similarly, although some of the SNPs were 306 correlated (R^2 >0.1; **Figure S24**), the overlap was incomplete (**Figure S26**, **Table S20**). The 307 pattern of associations was consistent with those described in the European cohort. The 308 strongest associations were with risky behaviors, such as adventurousness (β =-0.04, *p*=1.76E-

309 17), risk-taking (β =-0.02, *p*=5.90E-07), alcohol consumption (β =-0.03, *p*=1.41E-12), and

- disorders characterized by high levels of impulsivity, such as ADHD (β =-0.04, *p*=4.74E-10). The
- 311 novel findings were, again, with multiple forms of allergies (e.g., seasonal allergies, β =0.03,
- 312 p=3.0E-04), migraine ($\beta=0.04$, p=1.56E-04), sleep behaviors (e.g., sleep apnea, $\beta=-0.03$,
- 313 *p*=6.76E-04), among others.

314 All findings that were in common between the European and Latin American cohorts

- 315 showed the same direction of effect and similar effect sizes. We did not identify FDR-significant
- 316 associations in the African American cohort (**Table S17**). The effect sizes were generally
- 317 extremely small (Figures S27-28), as is expected for a single gene and complex traits.

318 MouseWAS

Figure 4 summarizes the MouseWAS results across the five cohorts tested. Full
 statistics and additional secondary measures are described in the Supplementary Material and
 Table S20.

322 Cohort 1 - Motivation, inhibition, and risk-taking behavior. No differences in motivation 323 were found between WT and HET mice during the Progressive Breakpoint task $[F_{(1.51)}=0.003,$ 324 p=9.57E-01; Figure 4A]. However, we noted significant genotype differences in behavioral 325 flexibility in the Probabilistic Reversal Learning (PRL) task, as indexed by the number of trials to 326 first reversal $[F_{(1,42)}=4.27, p=4.50E-02; Figure 4B]$, and risky behavior in the mouse lowa 327 Gambling Task (IGT; F_(1.51)=4.70, p=3.50E-02; Figure 4C], HET mice requiring fewer trials to 328 reach criterion and choosing risky options less frequently than WT mice (p < 0.05), respectively. 329 The number of premature responses, on the contrary, were higher in HET mice $[F_{(1.51)}=5.78]$, 330 p=2.00E-02] compared to WT mice (p<0.05; Figure 4D). In the Behavioral pattern monitor 331 (BPM), HET mice exhibited greater exploratory behavior, as shown by an increase in hole-332 pokes $[F_{(1,53)} = 4.88, p=3.20E-02;$ Figure 4E], compared to WT mice (p<0.05), but general levels 333 of activity, such as distance traveled ($F_{(1.53)} = 0.42$, p=5.21E-01; Figure 4F), were similar across 334 the genotypes. Lastly, although the startle response was equal across the groups (Figure 4G), 335 prepulse inhibition (**PPI**) was larger in HET mice compared to WT mice (p < 0.05; Figure 4H), 336 particularly at ISI 25 and 100 in HET mice [F_(1,53)=8.23, p=6.00E-03, F_(1,53)=4.50, p=3.90E-02, 337 respectively].

338 *Cohort 2 - Motoric impulsivity.* The main outcome tested in cohort 2 were premature
 339 responses via the 5-choice serial reaction time task (5CSRTT; Figure 4J-M). Premature

- responses were lower in HOM (p<0.001) and WT (p<0.02) mice compared to HET mice under
- standard conditions ($F_{(2,36)}$ =8.74, *p*=8.06E-04; **Figure 4J**), and compared to both HET (*p*<0.001)
- and WT (p<0.01) mice during a long ITI session (H $_{(2)}$ =16.10, p=3.19E-04; Figure 4L). HOM
- 343 mice were faster at learning the 5CSRTT, requiring fewer days for adequate baseline
- 344 performance ($F_{(2,36)}$ =7.42, *p*=2.00E-03; **Figure 4I**), compared to WT mice (*p*<0.01).
- 345 Cohort 3 General locomotion, anxiety-like behavior, and ethanol consumption. We
- found a significant effect of genotype on the distance traveled in the Open Field [**OF**;
- 347 F_(2,70)=7.525, *p*=1.00E-03; **Figure 4N**], HOM mice showing higher levels of locomotor activity
- 348 than WT mice (p=1.40E-02). No differences in anxiety-like behavior were detected across WT,
- 349 HET or HOM mice in the Elevated Plus Maze (EPM) or Light-Dark Box (LDB) tests (Figure 40-
- 350 **P**, **Table S20**). The total amount of ethanol consumed during the drinking-in-the-dark (**DID**)
- paradigm did not differ between the groups ($[F_{(2,78)}=1.084, p=3.44E-01]$; **Figure 4Q**).
- 352 *Cohort 4 Body weight.* Relative to WT mice, there was a significant reduction in body 353 weight in HOM mice from week 21 onwards (β =-3.74±1.27, *p*=4.00E-03; **Figure 4R**).
- 354 *Cohort 5 Dendrite morphology*. Quantitative analyses of MSN in the NAc revealed no 355 difference in dendritic spine density across the groups (**Figure 5S**).

356 **DISCUSSION**

357 In this study, we performed the largest GWAS of impulsive personality traits to date, we 358 conducted the first multi-ancestral PheWAS exploring the role of CADM2 on a diverse array of 359 traits, and we performed a corresponding MouseWAS using Cadm2 mutant mice to assess its 360 role in impulsivity and other relevant behaviors. We extend on prior findings [4, 5] showing that 361 the genetic architecture of impulsivity facets only partially overlap, providing further support to 362 the idea of impulsivity being a multifaceted construct even at the genetic level. We identified 363 positive genetic correlations across multiple domains, particularly substance use disorders, 364 confirming that NIMH RDoC transdiagnostic domains [6], or endophenotypes, such as impulsive 365 personality traits, can be used to dissect the genetic basis of psychiatric illness and normal 366 functioning. RDoC or transdiagnostic traits are beneficial because they enable translational 367 research and provide a more granular biological understanding of psychiatric disorders. Using 368 mouse and human correlates, we provide further evidence that CADM2 is a robust candidate 369 gene for impulsivity and an important modulator of numerous other psychiatric and somatic 370 traits.

371 We increased the sample size of our prior GWAS of impulsivity by almost 6-fold and 372 identified 21 genome-wide significant loci implicated in impulsive personality and Drug 373 Experimentation. For instance, SNPs located in the gene *TCF4* were implicated in 3 subscales: 374 this gene is also highly pleiotropic for other psychiatric conditions. Furthermore, we identified 375 associations with NCAM1, which, intriguingly, is a critical member of the NTAD (NCAM1-376 TTC12-ANKK1-DRD2) gene cluster [68] and variants correlated with NCAM1 in that cluster 377 have been associated with differences in D2 receptor density [69]. We also detected 378 associations near XKR6 and AFF3, which have been recently implicated in externalizing 379 psychopathology [16], and PTPRF and KDM4A, recently implicated in problematic opioid use 380 [25] and opioid use disorder [58]. Although in this report we focused on CADM2, functional 381 studies of those genes are also warranted. Furthermore, we found nominal evidence for 382 candidate gene studies implicating monoamine neurotransmitters in impulsivity and Drug 383 Experimentation (*DRD2*, *HTR3B*). High impulsivity depends on a neural network that includes 384 the ventral striatum (subsuming the NAc) with top-down control from prefrontal cortical regions, 385 and is modulated by monoamine neurotransmitters including dopamine and serotonin [70]: this 386 is the first GWAS to implicate genes modulating these systems as robust candidate genes for 387 impulsivity.

388 Recent studies have implicated the CADM2 gene in impulsivity and traits associated with 389 reward sensitivity and multiple domains of human health. We confirmed numerous previously 390 reported associations and extended our findings of variants related to CADM2. CADM2 was 391 significantly associated with 4 out of the 9 traits that we measured in GWAS and 6 out of the 9 392 traits that we measured in gene-based analyses. In the PheWAS, CADM2 variants were 393 associated with decreased risk for externalizing psychopathology, but also increased risk for 394 internalizing psychopathology (anxiety, depression, OCD). We also observed novel associations 395 with migraines and various allergies. Others using a similar approach with UK Biobank data 396 have found that this enrichment of associations is higher than expected [12] compared to other 397 genes. These results provide evidence that CADM2 variants are associated with broad health 398 outcomes, but whether this gene affects human health via disruptions in inhibitory control or 399 reward systems, or whether it acts via multiple pathways [71], is still not fully understood.

A relatively unique feature of our study is that, to follow up on the *CADM2* loci implicated in human studies, we generated a *Cadm2* mutant mouse line and used it to perform a PheWASlike study in mice, which we have termed a MouseWAS to emphasize its conceptual similarity to human PheWAS studies. These functional experiments provided information about the causality and directionality of effects in its reported associations. We found evidence that loss of *Cadm2* resulted in *less* risky behavior and *improved* information processing, extending on prior work in humans [4, 10, 16, 62, 69].

407 Cadm2 expression may uniquely contribute to the different domains of impulsivity. The 408 IGT assays preference for high risk, high reward (disadvantageous) choices vs low risk, low 409 reward (advantageous) choices [72]. HET mice exhibited a greater preference for selecting the 410 safe option vs their WT littermates. This finding can be contrasted with the *elevated* premature 411 responses in the 5CSRTT seen in HET vs WT mice, reflective of motoric impulsivity. However, 412 premature responses have also been linked to temporal discrimination, wherein mice and 413 humans overestimating the passage of time exhibit higher premature responses [73, 74]. The 414 preference for less risky options of HET mice in the IGT could reflect their misjudgment of time -415 resulting in higher premature responses – and thus avoidance of higher temporal punishment in 416 the IGT.

We also observed genotype differences in performance that could be indicative of *Cadm2* function in information processing. HET mice exhibiting *better* PPI at the shortest
temporal window (25 ms) supports the premise that these mice have faster processing speeds.

HET mice also showed small increases in hole-poking in the BPM test, which is thought to
 reflect exploration of an environment and information gathering. Finally, we observed that HOM
 mice acquired 5CSRTT faster than WT littermates. Taken together, these results suggest that
 Cadm2 reduction may improve some facets of information processing.

424 Findings from the 5CSRTT provide evidence that Cadm2 deletion improves some 425 information processing and impulsivity outcomes, while being detrimental to others. HET mice 426 were the most likely to commit 5CSRTT premature responses, although HOM mice were 427 surprisingly the least likely to make premature responses. Interestingly, although not significant. 428 there was a consistent elevation in the number of premature responses committed by the HOM 429 mice as the stimulus duration was reduced. This could suggest that HOM mice, like HET mice, 430 may show motoric impulsivity deficits when performing tasks that require greater attentional 431 demand. Compared with WT, HOM mice also showed impaired accuracy performance under 432 RSD conditions, in line with our human findings of CADM2 association with BIS Attentional, and 433 cognitive function by others [10]. The heterogeneity of performance outcomes in the HOM mice 434 further supports a unique but overlapping contribution of genetics across impulsivity domains.

435 In this paper, we have translated measures from human to mice. These studies begin 436 the process of understanding the biological basis of associations identified by GWAS. The 437 methods for measuring impulsivity in humans and mice are fundamentally different. Despite 438 these differences, our MouseWAS identified several measures of impulsivity that were 439 influenced by Cadm2, consistent with our observations in humans. Furthermore, CADM2 has 440 been shown to be implicated in BMI in humans [24, 75] and energy homeostasis in mice [24]; 441 extending on this, we found novel evidence of body weight reductions in adult mutant mice. 442 Interestingly, Cadm2 did not have more general effects on mouse behavior; for instance, we did 443 not observe deficits in anxiety-like behavior or general motivation, as some of the human 444 PheWAS findings revealed. A few other measures were also inconsistent across species, 445 particularly measures of alcohol consumption, where CADM2 showed a role in humans [7, 13, 446 76, 77] but not mice. Lastly, some measures identified by our human PheWAS (e.g., allergies 447 and other medical conditions) were not examined in our MouseWAS. This approach highlights 448 the challenges of using mouse models to further investigate the role of specific genes in 449 behavioral traits.

450 *CADM*2 encodes the immunoglobulin adhesion protein SynCAM 2, which is part of the 451 family of synaptic adhesion molecules known as SynCAMs. Studies have shown the influence

452 of SynCAMs on synaptogenesis [78-82], axon guidance [83], and neuron myelination [84-86], 453 processes that have direct effects on the pathology of neurodevelopmental diseases [56]. 454 CADM2 is strongly expressed in the striatum and frontal cortex, which are core regions that 455 regulate impulsivity [70]. We did not observe changes in spine density in the Nac, which 456 suggests that Cadm2 may not have a role as a postsynaptic organizer of spines in this region, 457 or may have redundant functions that are compensated in the mutant mice by other molecules. 458 Based on *in-silico* analyses in humans, CADM2 expression seems to be greater at earlier 459 stages of development (Figure S29); whether Cadm2 may affect earlier stages of development 460 (prenatal and early postnatal) that are compensated in adulthood has not been investigated in 461 this study.

462 Several limitations of this study are worth noting. The discovery GWAS only includes 463 male and female participants of European ancestry. While we provided exploratory analyses of 464 top variants in other ancestries and by broken down by sex (Supplementary Table 22), larger 465 sample sizes would be needed to perform GWAS separately in males and females. Our results 466 are also biased by potential ascertainment and characteristics of the sample; the 23andMe 467 participant population is more educated and has higher socioeconomic status and lower levels 468 of drug use and impulsivity than the general US population [86]. Replication in additional cohorts 469 with different characteristics is warranted. Moreover, although the traits we studied are extracted 470 via well-established questionnaires, they are self-reported measures, which are different from 471 behavioral phenotypes [87, 88]. Another issue is that, although we tested multiple variants in the 472 CADM2 loci, further conditional analyses are required to determine if this signal and previously 473 reported associations implicating CADM2 loci, including a large non-coding rare deletion in the 474 first intron of CADM2 [71], may tag the same underlying genetic effect. We are also unaware of 475 the sequence of events, and whether there is true pleiotropy or mediation effects has not been 476 examined. The analyses were well-powered for moderate and large effect sizes. Still, for 477 unclear reasons, despite similar minor allele frequencies and imputation quality of the SNPs we 478 tested across all ancestries, we identified no significant associations in the African American 479 cohort. Finally, although our mouse studies detected some discordant cross-species effects of 480 *Cadm2* on behavior, background strain effects [89] or subtle allelic variations (vs whole KO) 481 may explain those differences. While some results are suggestive of additive effects, we were 482 unable to evaluate different genetic models due to lack of sufficient sample sizes for HOM mice. 483 Future multivariate analyses examining paths of commonality and specificity across impulsivity 484 facets may provide further insights not herein examined.

In conclusion, we show that impulsivity facets are extremely polygenic, but of very high transdiagnostic significance. Genetic studies using research participants not ascertained for neuropsychiatric disorders may represent an efficient and cost-effective strategy for elucidating the genetic basis and etiology of genetically complex psychiatric diseases. Using homologous measures of impulsivity in mice and humans across three ancestral backgrounds, we provide evidence of the overarching role of *CADM2* on impulsivity, and a much broader impact on human health.

492 Data availability

We will provide summary statistics for the top 10,000 SNPs upon publication (**Tables S22-30**). Full GWAS summary statistics will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please visit https://research.23andme.com/collaborate/#dataset-access/ for more information and to apply to access the data.

498 Acknowledgements

499 We would like to thank the research participants and employees of 23andMe for making 500 this work possible. The following members of the 23andMe Research Team contributed to this 501 study: Stella Aslibekvan, Adam Auton, Elizabeth Babalola, Robert K. Bell, Jessica Bielenberg, 502 Katarzyna Bryc, Emily Bullis, Daniella Coker, Gabriel Cuellar Partida, Devika Dhamija, 503 Sayantan Das, Teresa Filshtein, Kipper Fletez-Brant, Will Freyman, Karl Heilbron, Pooja M. 504 Gandhi, Karl Heilbron, Barry Hicks, David A. Hinds, Ethan M. Jewett, Yunxuan Jiang, Katelyn 505 Kukar, Keng-Han Lin, Maya Lowe, Jey C. McCreight, Matthew H. McIntyre, Steven J. Micheletti, 506 Meghan E. Moreno, Joanna L. Mountain, Priyanka Nandakumar, Elizabeth S. Noblin, Jared 507 O'Connell, Aaron A. Petrakovitz, G. David Poznik, Morgan Schumacher, Anjali J. Shastri, Janie 508 F. Shelton, Jingchunzi Shi, Suyash Shringarpure, Vinh Tran, Joyce Y. Tung, Xin Wang, Wei 509 Wang, Catherine H. Weldon, Peter Wilton, Alejandro Hernandez, Corinna Wong, Christophe 510 Toukam Tchakouté.

We would also like to thank The Externalizing Consortium for sharing the GWAS
summary statistics of externalizing. The Externalizing Consortium: Principal Investigators:
Danielle M. Dick, Philipp Koellinger, K. Paige Harden, Abraham A. Palmer. Lead Analysts:
Richard Karlsson Linnér, Travis T. Mallard, Peter B. Barr, Sandra Sanchez-Roige. Significant
Contributors: Irwin D. Waldman. The Externalizing Consortium has been supported by the

516 National Institute on Alcohol Abuse and Alcoholism (R01AA015416 -administrative supplement). 517 and the National Institute on Drug Abuse (R01DA050721). Additional funding for investigator 518 effort has been provided by K02AA018755, U10AA008401, P50AA022537, as well as a 519 European Research Council Consolidator Grant (647648 EdGe to Koellinger). The content is 520 solely the responsibility of the authors and does not necessarily represent the official views of 521 the above funding bodies. The Externalizing Consortium would like to thank the following groups 522 for making the research possible: 23andMe, Add Health, Vanderbilt University Medical Center's 523 BioVU, Collaborative Study on the Genetics of Alcoholism (COGA), the Psychiatric Genomics 524 Consortium's Substance Use Disorders working group, UK10K Consortium, UK Biobank, and 525 Philadelphia Neurodevelopmental Cohort.

526 Disclosures

527 MVJ, SBB, YH, SSR and AAP were supported by funds from the California Tobacco-Related Disease Research Program (TRDRP; Grant Number 28IR-0070, T29KT0526 and 528 529 T32IR5226). SBB was also supported by P50DA037844, SSR was also supported by NIH/NIDA 530 DP1DA054394. JM and SSR were supported by the Brain and Behavior Foundation (grant 531 27676) and the Families for Borderline Personality Disorder Research (Beth and Rob Elliott) 532 2018 NARSAD Young Investigator Grant. The content is solely the responsibility of the authors 533 and does not necessarily represent the official views of the National Institutes of Health. SAB 534 was supported by NIH/NIMH grants R01MH108653 and R21MH117518. AMBL was supported 535 by NIH/NIAAA grant K99AA027835. TB acknowledges support by NIH/NIDA R01 DA018928. 536 JM is supported by the Peter Boris Chair in Addictions Research. HHAT is funded through a 537 Natural Science and Engineering Research Council PGS-D scholarship, and studies in Cohort 2 538 were supported by the Canadian Institutes of Health Research Project Grant (PJT-173442 to 539 JYK).

540 PF and SLE are employees of 23andMe, Inc., and hold stock or stock options in 541 23andMe. JY reports having received grant support funding from Sunovion, Heptares, and 542 Gilgamesh, as well as honoraria from Marvel Biotech, none of which were involved in the 543 current project. The other authors report no conflict of interest.

544

545 **REFERENCES**

546

- 547 1. Daruna JH, Barnes PA. A neurodevelopmental view of impulsivity. Impuls. Client Theory 548 Res. Treat., Washington, D.C.: American Psychological Association; 1993. p. 23. 549 2. Coskunpinar A, Dir AL, Cyders MA. Multidimensionality in Impulsivity and Alcohol Use: A 550 Meta-Analysis Using the UPPS Model of Impulsivity. Alcohol Clin Exp Res. 551 2013:37:1441-1450. 552 Jackson JNS, MacKillop J. Attention-Deficit/Hyperactivity Disorder and Monetary Delay 3. 553 Discounting: A Meta-Analysis of Case-Control Studies. Biol Psychiatry Cogn Neurosci 554 Neuroimaging. 2016;1:316-325. 555 Sanchez-Roige S, Fontanillas P, Elson SL, Gray JC, de Wit H, MacKillop J, et al. 4. 556 Genome-wide association studies of impulsive personality traits (BIS-11 and UPPSP) and drug experimentation in up to 22,861 adult research participants identify loci in the 557 558 CACNA11 and CADM2 genes. J Neurosci. 2019:2662-18. 559 Gustavson DE, Friedman NP, Fontanillas P, Elson SL. The Latent Genetic Structure of 5. 560 Impulsivity and Its Relation to Internalizing Psychopathology:11. 561 6. Insel TR. The NIMH Research Domain Criteria (RDoC) Project: Precision Medicine for Psychiatry. Am J Psychiatry. 2014;171:395–397. 562 Arends RM, Pasman JA, Verweij KJH, Derks EM, Gordon SD, Hickie I, et al. Associations 563 7. 564 between the CADM2 gene, substance use, risky sexual behavior, and self-control: A 565 phenome-wide association study:13. Boutwell B, Hinds D, Tielbeek J, Ong KK, Day FR, Perry JRB, et al. Replication and 566 8. 567 characterization of CADM2 and MSRA genes on human behavior. Helivon.
- 568 2017;3:e00349.
- 569 9. 23andMe Research Team, COGENT (Cognitive Genomics Consortium), Social Science
 570 Genetic Association Consortium, Lee JJ, Wedow R, Okbay A, et al. Gene discovery and
 571 polygenic prediction from a genome-wide association study of educational attainment in
 572 1.1 million individuals. Nat Genet. 2018;50:1112–1121.
- 573 10. on behalf of Generation Scotland, Ibrahim-Verbaas CA, Bressler J, Debette S, Schuur M,
 574 Smith AV, et al. GWAS for executive function and processing speed suggests
 575 involvement of the CADM2 gene. Mol Psychiatry. 2016;21:189–197.
- 576 11. Strawbridge RJ, Ward J, Cullen B, Tunbridge EM, Hartz S, Bierut L, et al. Genome-wide
 577 analysis of self-reported risk-taking behaviour and cross-disorder genetic correlations in
 578 the UK Biobank cohort. Transl Psychiatry. 2018;8:39.
- A. Pasman J, Chen Z, Vink JM, Van Den Oever MC, Pattij T, Vries TJD, et al. The *CADM2* gene and behavior: A phenome-wide scan in UK-Biobank. Genetic and Genomic
 Medicine; 2021.
- Sanchez-Roige S, Palmer AA, Fontanillas P, Elson SL, the 23andMe Research Team, the Substance Use Disorder Working Group of the Psychiatric Genomics Consortium, Adams MJ, et al. Genome-Wide Association Study Meta-Analysis of the Alcohol Use Disorders Identification Test (AUDIT) in Two Population-Based Cohorts. Am J Psychiatry. 2019;176:107–118.
- the 23andMe Research Team, The Substance Use Disorders Working Group of the
 Psychiatric Genomics Consortium, International Cannabis Consortium, Pasman JA,
 Verweij KJH, Gerring Z, et al. GWAS of lifetime cannabis use reveals new risk loci,
 genetic overlap with psychiatric traits, and a causal effect of schizophrenia liability. Nat
 Neurosci. 2018;21:1161–1170.
- 592 15. 23and Me Research Team, eQTLgen Consortium, International Cannabis Consortium,
 593 Social Science Genetic Association Consortium, Karlsson Linnér R, Biroli P, et al.
 594 Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million

595 individuals identify hundreds of loci and shared genetic influences. Nat Genet. 596 2019;51:245-257. 597 16. Linnér RK. Multivariate analysis of 1.5 million people identifies genetic associations with 598 traits related to self-regulation and addiction. Nat Neurosci. 2021;24:27. 17. 599 Casey JP, Magalhaes T, Conroy JM, Regan R, Shah N, Anney R, et al. A novel approach 600 of homozygous haplotype sharing identifies candidate genes in autism spectrum disorder. 601 Hum Genet. 2012;131:565-579. 602 ADHD Working Group of the Psychiatric Genomics Consortium (PGC), Early Lifecourse & 18. 603 Genetic Epidemiology (EAGLE) Consortium, 23andMe Research Team, Demontis D, 604 Walters RK, Martin J, et al. Discovery of the first genome-wide significant risk loci for 605 attention deficit/hyperactivity disorder. Nat Genet. 2019;51:63-75. 606 19. Klimentidis YC, Raichlen DA, Bea J, Garcia DO, Wineinger NE, Mandarino LJ, et al. 607 Genome-wide association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple variants including CADM2 and APOE. Int J Obes. 608 609 2018:42:1161-1176. 610 20. Day FR, Helgason H, Chasman DI, Rose LM, Loh P-R, Scott RA, et al. Physical and 611 neurobehavioral determinants of reproductive onset and success. Nat Genet. 2016;48:617-623. 612 21. Mills MC, Tropf FC, Brazel DM, van Zuydam N, Vaez A, eQTLGen Consortium, et al. 613 614 Identification of 371 genetic variants for age at first sex and birth linked to externalising 615 behaviour. Nat Hum Behav. 2021:5:1717-1730. The LifeLines Cohort Study, The ADIPOGen Consortium, The AGEN-BMI Working 616 22. 617 Group, The CARDIOGRAMplusC4D Consortium, The CKDGen Consortium, The GLGC, 618 et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 619 2015;518:197-206. 620 23. Justice AE, Winkler TW, Feitosa MF, Graff M, Fisher VA, Young K, et al. Genome-wide 621 meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for 622 obesity traits. Nat Commun. 2017:8:14977. 623 24. Yan X, Wang Z, Schmidt V, Gauert A, Willnow TE, Heinig M, et al. Cadm2 regulates body 624 weight and energy homeostasis in mice. Mol Metab. 2018;8:180-188. 625 25. Sanchez-Roige S, Fontanillas P, Jennings MV, Bianchi S, Huang Y, Hatoum A, et al. 626 Genome-wide association study of problematic opioid prescription use in 132,113 627 23andMe research participants of European ancestry. Psychiatry and Clinical 628 Psychology; 2021. 629 Sanchez-Roige S, Fontanillas P, Elson SL, the 23andme research team, Pandit A, 26. 630 Schmidt EM, et al. Genome-wide association study of delay discounting in 23,217 adult 631 research participants of European ancestry. Nat Neurosci. 2018;21:16-18. 632 27. Cyders MA, Littlefield AK, Coffey S, Karyadi KA. Examination of a short English version of 633 the UPPS-P Impulsive Behavior Scale. Addict Behav. 2014;39:1372–1376. 634 Whiteside SP, Lynam DR. The Five Factor Model and impulsivity: using a structural 28. 635 model of personality to understand impulsivity. Personal Individ Differ. 2001;30:669-689. 636 29. Patton JH, Stanford MS, Barratt ES. Factor structure of the barratt impulsiveness scale. 637 Wiley Period Inc. 1995;51:768-774. 638 30. VanderBroek L, Acker J, Palmer AA, de Wit H, MacKillop J. Interrelationships among 639 parental family history of substance misuse, delay discounting, and personal substance 640 use. Psychopharmacology (Berl). 2016;233:39-48. 641 31. Durand EY, Do CB, Mountain JL, Macpherson JM. Ancestry Composition: A Novel, 642 Efficient Pipeline for Ancestry Deconvolution:16. 643 Eriksson N, Macpherson JM, Tung JY, Hon LS, Naughton B, Saxonov S, et al. Web-32. 644 Based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits. 645 PLoS Genet. 2010;6:e1000993.

- 646 33. Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR, et al. Identification of 15
 647 genetic loci associated with risk of major depression in individuals of European descent.
 648 Nat Genet. 2016;48:1031–1036.
- 649 34. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and 650 annotation of genetic associations with FUMA. Nat Commun. 2017;8:1826.
- 651 35. Leeuw CA de, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set 652 Analysis of GWAS Data. PLOS Comput Biol. 2015;11:e1004219.
- Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Patterson N, et al. LD Score
 regression distinguishes confounding from polygenicity in genome-wide association
 studies. Nat Genet. 2015;47:291–295.
- 656 37. Verma A, Bradford Y, Dudek S, Lucas AM, Verma SS, Pendergrass SA, et al. A
 657 simulation study investigating power estimates in phenome-wide association studies.
 658 BMC Bioinformatics. 2018;19:120.
- 659 38. Ehm MG, Aponte JL, Chiano MN, Yerges-Armstrong LM, Johnson T, Barker JN, et al.
 660 Phenome-wide association study using research participants' self-reported data provides 661 insight into the Th17 and IL-17 pathway. PLOS ONE. 2017;12:e0186405.
- 662 39. Lloyd KCK. A knockout mouse resource for the biomedical research community: The
 663 KOMP Repository. Ann N Y Acad Sci. 2011;1245:24–26.
- 40. Evangelou E, Gao H, Chu C, Ntritsos G, Blakeley P, Butts AR, et al. New alcohol-related
 genes suggest shared genetic mechanisms with neuropsychiatric disorders. Nat Hum
 Behav. 2019;3:950–961.
- 41. Ikeda M, Takahashi A, Kamatani Y, Momozawa Y, Saito T, Kondo K, et al. Genome-Wide
 Association Study Detected Novel Susceptibility Genes for Schizophrenia and Shared
 Trans-Populations/Diseases Genetic Effect. Schizophr Bull. 2019;45:824–834.
- 42. 23andMe Research Team, Major Depressive Disorder Working Group of the Psychiatric
 671 Genomics Consortium, Howard DM, Adams MJ, Clarke T-K, Hafferty JD, et al. Genome672 wide meta-analysis of depression identifies 102 independent variants and highlights the
 673 importance of the prefrontal brain regions. Nat Neurosci. 2019;22:343–352.
- 43. Yang H, Liu D, Zhao C, Feng B, Lu W, Yang X, et al. Mendelian randomization integrating
 675 GWAS and eQTL data revealed genes pleiotropically associated with major depressive
 676 disorder. Transl Psychiatry. 2021;11:225.
- BIOS consortium, Social Science Genetic Association Consortium, Baselmans BML,
 Jansen R, Ip HF, van Dongen J, et al. Multivariate genome-wide analyses of the wellbeing spectrum. Nat Genet. 2019;51:445–451.
- 45. Hill WD, Weiss A, Liewald DC, Davies G, Porteous DJ, Hayward C, et al. Genetic
 contributions to two special factors of neuroticism are associated with affluence, higher
 intelligence, better health, and longer life. Mol Psychiatry. 2020;25:3034–3052.
- 46. Teixeira JR, Szeto RA, Carvalho VMA, Muotri AR, Papes F. Transcription factor 4 and its association with psychiatric disorders. Transl Psychiatry. 2021;11:19.
- 47. Van Esch H, Colnaghi R, Freson K, Starokadomskyy P, Zankl A, Backx L, et al. Defective
 DNA Polymerase α-Primase Leads to X-Linked Intellectual Disability Associated with
 Severe Growth Retardation, Microcephaly, and Hypogonadism. Am J Hum Genet.
 2019;104:957–967.
- 689 48. Christakoudi S, Evangelou E, Riboli E, Tsilidis KK. GWAS of allometric body-shape
 690 indices in UK Biobank identifies loci suggesting associations with morphogenesis,
 691 organogenesis, adrenal cell renewal and cancer. Sci Rep. 2021;11:10688.
- 49. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, et al. Metaanalysis of genome-wide association studies for body fat distribution in 694 649
 individuals of European ancestry. Hum Mol Genet. 2019;28:166–174.
- 695 50. Nagel M, Watanabe K, Stringer S, Posthuma D, van der Sluis S. Item-level analyses
 696 reveal genetic heterogeneity in neuroticism. Nat Commun. 2018;9:905.

- 697 51. Lo M-T, Hinds DA, Tung JY, Franz C, Fan C-C, Wang Y, et al. Genome-wide analyses for
 698 personality traits identify six genomic loci and show correlations with psychiatric
 699 disorders. Nat Genet. 2017;49:152–156.
- 52. 23andMe Research Team, Thorp JG, Campos AI, Grotzinger AD, Gerring ZF, An J, et al.
 Symptom-level modelling unravels the shared genetic architecture of anxiety and
 depression. Nat Hum Behav. 2021. 15 April 2021. https://doi.org/10.1038/s41562-02101094-9.
- Feitosa MF, Kraja AT, Chasman DI, Sung YJ, Winkler TW, Ntalla I, et al. Novel genetic
 associations for blood pressure identified via gene-alcohol interaction in up to 570K
 individuals across multiple ancestries. PLOS ONE. 2018;13:e0198166.
- The Endometrial Cancer Association Consortium, Ruth KS, Day FR, Tyrrell J, Thompson
 DJ, Wood AR, et al. Using human genetics to understand the disease impacts of
 testosterone in men and women. Nat Med. 2020;26:252–258.
- 710 55. Rask-Andersen M, Karlsson T, Ek WE, Johansson Å. Genome-wide association study of
 711 body fat distribution identifies adiposity loci and sex-specific genetic effects. Nat
 712 Commun. 2019;10:339.
- 56. Hill WD, Davies NM, Ritchie SJ, Skene NG, Bryois J, Bell S, et al. Genome-wide analysis
 identifies molecular systems and 149 genetic loci associated with income. Nat Commun.
 2019;10:5741.
- Alzheimer's Disease Neuroimaging Initiative, Pediatric Imaging, Neurocognition and
 Genetics, Zhao B, Luo T, Li T, Li Y, et al. Genome-wide association analysis of 19,629
 individuals identifies variants influencing regional brain volumes and refines their genetic
 co-architecture with cognitive and mental health traits. Nat Genet. 2019;51:1637–1644.
- 58. Gaddis N, Mathur R, Marks J, Zhou L, Quach B, Waldrop A, et al. Multi-trait genome-wide association study of opioid addiction: *OPRM1* and Beyond. Addiction Medicine; 2021.
- 59. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to 1.2
 million individuals yield new insights into the genetic etiology of tobacco and alcohol use.
 Nat Genet. 2019;51:237–244.
- Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, et al. Study of 300,486
 individuals identifies 148 independent genetic loci influencing general cognitive function.
 Nat Commun. 2018;9:2098.
- ADHD Working Group of the Psychiatric Genomics Consortium, 23andMe Research
 team, Rovira P, Demontis D, Sánchez-Mora C, Zayats T, et al. Shared genetic
 background between children and adults with attention deficit/hyperactivity disorder.
 Neuropsychopharmacology. 2020;45:1617–1626.
- Karlsson Linnér R, Biroli P, Kong E, Meddens SFW, Wedow R, Fontana MA, et al.
 Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nat Genet.
 2019;51:245–257.
- Pritikin JN, Neale MC, Prom-Wormley EC, Clark SL, Verhulst B. GW-SEM 2.0: Efficient,
 Flexible, and Accessible Multivariate GWAS. Behav Genet. 2021;51:343–357.
- MacKillop J, Weafer J, C. Gray J, Oshri A, Palmer A, de Wit H. The latent structure of
 impulsivity: impulsive choice, impulsive action, and impulsive personality traits.
 Psychopharmacology (Berl). 2016;233:3361–3370.
- 65. Gray JC, MacKillop J, Weafer J, Hernandez KM, Gao J, Palmer AA, et al. Genetic
 analysis of impulsive personality traits: Examination of a priori candidates and genomewide variation. Psychiatry Res. 2018;259:398–404.
- 66. Leppert B, Millard LAC, Riglin L, Davey Smith G, Thapar A, Tilling K, et al. A crossdisorder PRS-pheWAS of 5 major psychiatric disorders in UK Biobank. PLOS Genet.
 2020;16:e1008185.

747 67. Metha JA, Brian ML, Oberrauch S, Barnes SA, Featherby TJ, Bossaerts P, et al. 748 Separating Probability and Reversal Learning in a Novel Probabilistic Reversal Learning Task for Mice. Front Behav Neurosci. 2020;13:270. 749 750 68. Mota NR, Araujo-Jnr EV, Paixão-Côrtes VR, Bortolini MC, Bau CHD. Linking dopamine 751 neurotransmission and neurogenesis: the evolutionary history of the NTAD (NCAM1-752 TTC12-ANKK1-DRD2) gene cluster. Genet Mol Biol. 2012;35:912–918. 753 69. Greenbaum L, Ravona-Springer R, Livny A, Shelly S, Ganmore I, Alkelai A, et al. The 754 CADM2 gene is associated with processing speed performance - evidence among 755 elderly with type 2:8. 756 70. Dalley JW, Everitt BJ, Robbins TW. Impulsivity, Compulsivity, and Top-Down Cognitive 757 Control. Neuron. 2011;69:680-694. 758 71. Collins RL, Glessner JT, Porcu E, Niestroj L-M, Ulirsch J, Kellaris G, et al. A cross-759 disorder dosage sensitivity map of the human genome. Genetic and Genomic Medicine; 760 2021. van Enkhuizen J. Henry BL. Minassian A. Perry W. Milienne-Petiot M. Higa KK. et al. 761 72. 762 Reduced Dopamine Transporter Functioning Induces High-Reward Risk-Preference 763 Consistent with Bipolar Disorder. Neuropsychopharmacology. 2014;39:3112–3122. 764 73. Cope ZA, Halberstadt AL, van Enkhuizen J, Flynn AD, Breier M, Swerdlow NR, et al. 765 Premature responses in the five-choice serial reaction time task reflect rodents' temporal 766 strategies: evidence from no-light and pharmacological challenges. Psychopharmacology 767 (Berl), 2016:233:3513-3525. Hsu T-Y, Lee H-C, Lane TJ, Missal M. Temporal Preparation, Impulsivity and Short-Term 768 74. 769 Memory in Depression. Front Behav Neurosci. 2019;13:258. 770 75. Morris J. Genetic variation in CADM2 as a link between psychological traits and 771 obesity:14. 772 76. Clarke T-K, Adams M, Davies G, Howard D, Hall L, Padmanabhan S, et al. Genome-wide 773 association study of alcohol consumption and genetic overlap with other health-related 774 traits in UK Biobank. Mol Psychiatry. 2017:9. Kranzler HR, Zhou H, Kember RL, Smith RV, Justice AC, Damrauer S, et al. Genome-775 77. wide association study of alcohol consumption and use disorder in 274,424 individuals 776 777 from multiple populations. Nat Commun. 2019;10:1499. 778 78. Giza JI, Jung Y, Jeffrey RA, Neugebauer NM, Picciotto MR, Biederer T. The Synaptic 779 Adhesion Molecule SynCAM 1 Contributes to Cocaine Effects on Synapse Structure and 780 Psychostimulant Behavior. Neuropsychopharmacology. 2013;38:628-638. 781 Biederer T, Sara Y, Mozhayeva M, Atasoy D, Liu X, Kavalali ET, et al. SynCAM, a 79. 782 Synaptic Adhesion Molecule That Drives Synapse Assembly. Science. 2002;297:1525-783 1531. 784 80. Fogel AI, Akins MR, Krupp AJ, Stagi M, Stein V, Biederer T. SynCAMs Organize 785 Synapses through Heterophilic Adhesion. J Neurosci. 2007;27:12516–12530. 786 Robbins EM, Krupp AJ, Perez de Arce K, Ghosh AK, Fogel AI, Boucard A, et al. SynCAM 81. 787 1 Adhesion Dynamically Regulates Synapse Number and Impacts Plasticity and 788 Learning. Neuron. 2010;68:894-906. 789 82. Yamada A, Inoue E, Deguchi-Tawarada M, Matsui C, Togawa A, Nakatani T, et al. Necl-790 2/CADM1 interacts with ErbB4 and regulates its activity in GABAergic neurons. Mol Cell 791 Neurosci. 2013;56:234-243. 792 Niederkofler V, Baeriswyl T, Ott R, Stoeckli ET. Nectin-like molecules/SynCAMs are 83. 793 required for post-crossing commissural axon guidance. Development. 2010;137:427-435. 794 Maurel P, Einheber S, Galinska J, Thaker P, Lam I, Rubin MB, et al. Nectin-like proteins 84. 795 mediate axon-Schwann cell interactions along the internode and are essential for 796 myelination. J Cell Biol. 2007;178:861-874.

- Spiegel I, Adamsky K, Eshed Y, Milo R, Sabanay H, Sarig-Nadir O, et al. A central role for Necl4 (SynCAM4) in Schwann cell–axon interaction and myelination. Nat Neurosci. 2007;10:861–869.
- 80. 86. Park J, Liu B, Chen T, Li H, Hu X, Gao J, et al. Disruption of Nectin-Like 1 Cell Adhesion
 801 Molecule Leads to Delayed Axonal Myelination in the CNS. J Neurosci. 2008;28:12815–
 802 12819.
- 803 87. Enticott PG, Ogloff JRP, Bradshaw JL. Associations between laboratory measures of
 804 executive inhibitory control and self-reported impulsivity. Personal Individ Differences.
 805 2006:10.
- 806 88. Lane SD, Cherek DR, Rhoades HM, Pietras CJ, Tcheremissine OV. Relationships
 807 Among Laboratory and Psychometric Measures of Impulsivity: Implications in Substance
 808 Abuse and Dependence:8.
- 809 89. Sittig LJ, Carbonetto P, Engel KA, Krauss KS, Barrios-Camacho CM, Palmer AA. Genetic
 810 Background Limits Generalizability of Genotype-Phenotype Relationships. Neuron.
 811 2016;91:1253–1259.
- 812
- 813

Figure Legends

Figure 1. Porcupine plot displaying 21 genome-wide significant hits for all impulsivity facets and Drug Experimentation. *CADM2* was consistent across 3/8 impulsivity facets [Sensation Seeking (UPPS-P), Motor and Nonplanning impulsivity (BIS-11)] at a genome-wide association level, and with 3 more impulsivity facets [Attentional (BIS-11), Negative Urgency and Premeditation (UPPS-P]), at a gene-based level (**Table S8**). *CADM2* was also associated with risky behavior, such as Drug Experimentation.

Figure 2. Genetic correlations (r_g) between UPPS-P, BIS, and Drug Experimentation, and other substance use, psychiatric, personality, cognitive, metabolic, health, pain, longevity and sleep traits (see **Table S9** for full results). All values survive 5% FDR correction.

Figure 3. FDR-significant associations from *CADM2* single-SNP PheWAS in individuals of European ancestry (**A**) and Latin American ancestry (**B**). Results for the SNP with the highest number of significant PheWAS associations (rs993137) are presented; results for additional SNPs are included in Supplementary Tables 15-17. No FDR-significant findings were detected in individuals of African American ancestry. The size of the dots represents the magnitude of the effect size for each trait. The effect sizes ranged from -0.14 to 0.13 in the European cohort and from -0.08 to 0.16 in the Latin American cohort. Dotted line denotes Bonferroni significance (p<3.79E-05).

Figure 4. MouseWAS examined the complete and partial loss of *Cadm2* function and behavioral consequences in the Probability Breakpoint Ratio Task (PBRT, **A**), Probabilistic Reversal Learning Task (PRLT, **B**), Iowa gambling task (IGT, **C-D**), Behavioral Pattern Monitor task (BPM, **E-F**), Prepulse Inhibition (PPI, **G-H**), 5-Choice Serial Reaction Time Task (5CSRTT) performance (**I-M**), Locomotor (LCM) activity (**N**), Elevated Plus Maze (EPM, **O**), Light-Dark Box (LDB, **P**), and Drinking in the Dark (DID, **Q**); Iongitudinal body weight changes (**R**), and dendrite morphology (DSM) in the nucleus accumbens (**S**). Western blot (**WB**) analysis (**T**) of Cadm2 protein in whole brain, frontal cortex and striatum. WT, wildtype, HET, heterozygote, HOM, homozygote. Sample size by cohort: cohort 1 (WT=25, HET=30, HOM=3), cohort 2 (WT=13; HET=14, HOM=12), cohort 3 (WT=22; HET=44, HOM=12), cohort 4 (WT=29, HT=54, HOM=17), cohort 5 (WT=3, HET=3, HOM=3). In cohort 1, the sample size of the HOM mice deviated from the expected Mendelian frequency for unknown reasons; these animals were excluded from the analyses. Males are represented in circles, females in triangles. * *p*<0.05, ** *p*<0.01, *** *p*<0.001.







