Uncovering the Genetic Architecture of Broad Antisocial Behavior through a Genome-Wide Association Study Meta-analysis.

Authors: Jorim J. Tielbeek, Emil Uffelmann, Benjamin S. Williams, Lucía Colodro-Conde, Éloi Gagnon, Travis T. Mallard, Brandt Levitt, Philip R. Jansen, Ada Johansson, Hannah Sallis, Giorgio Pistis, Gretchen R.B. Saunders, Andrea G Allegrini, Kaili Rimfeld, Bettina Konte, Marieke Klein, Annette M. Hartmann, Jessica E Salvatore, Ilja M. Nolte, Ditte Demontis, Anni Malmberg, S. Alexandra Burt, Jeanne Savage, Karen Sugden, Richie Poulton, Kathleen Mullan Harris, Scott Vrieze, Matt McGue, William G. Iacono, Nina Roth Mota, Jonathan Mill, Joana F. Viana, Brittany L Mitchell, Jose J Morosoli, Till Andlauer, Isabelle Ouellet-Morin, Richard E. Tremblay, Sylvana Côté ,Jean-Philippe Gouin, Mara Brendgen, Ginette Dionne, Frank Vitaro, Michelle K Lupton, Nicholas G Martin, COGA Consortium, Spit for Science Working Group, Enrique Castelao, Katri Räikkönen, Johan Eriksson, Jari Lahti, Catharina A Hartman, Albertine J. Oldehinkel, Harold Snieder, Hexuan Liu, Martin Preisig, Alyce Whipp, Eero Vuoksimaa ,Yi Lu, Patrick Jern, Dan Rujescu, Ina Giegling, Teemu Palviainen, Jaakko Kaprio, Kathryn Paige Harden, Marcus R. Munafò, Geneviève Morneau-Vaillancourt, Robert Plomin, Essi Viding, Brian B. Boutwell, Fazil Aliev, Danielle Dick, Arne Popma, Stephen V Faraone, Anders D. Børglum, Sarah E Medland, Barbara Franke, Michel Boivin, Jean-Baptiste Pingault, Jeffrey C Glennon, James C. Barnes, Simon E. Fisher, Terrie E. Moffitt, Avshalom Caspi, Tinca JC Polderman, Danielle Posthuma, *Broad Antisocial Behavior Consortium collaborators*

Correspondence to: Jorim J. Tielbeek, Department of Complex Trait Genetics, Vrije Universiteit Amsterdam, De Boelelaan 1085, 1081 HV, Amsterdam, The Netherlands.

E-mail: j.j.tielbeek@vu.nl

Conflict of Interest Statement

BF has received educational speaking fees from Medice.

Abstract

Despite the substantial heritability of antisocial behavior (ASB), specific genetic variants robustly associated with the trait have not been identified. The present study by the Broad Antisocial Behavior Consortium (BroadABC) meta-analyzed data from 28 discovery samples (N = 85,359) and five independent replication samples (N = 8,058) with genotypic data and broad measures of ASB. We identified the first significant genetic associations with broad ASB, involving common intronic variants in the forkhead box protein P2 (FOXP2) gene (lead SNP rs12536335, P = 6.32 x 10-10). Furthermore, we observed intronic variation in Foxp2 and one of its targets (Cntnap2) distinguishing a mouse model of pathological aggression (BALB/cJ strain) from controls (BALB/cByJ strain). The SNP-based heritability of ASB was 8.4% (s.e.= 1.2%). Polygenic-risk-score (PRS) analyses in independent samples revealed that the genetic risk for ASB was associated with several antisocial outcomes across the lifespan, including diagnosis of conduct disorder, official criminal convictions, and trajectories of antisocial development. We found substantial genetic correlations of ASB with mental health (depression rg = 0.63, insomnia rg = 0.47), physical health (overweight rg = 0.19, waist-to-hip ratio rg = 0.32), smoking (rg = 0.54), cognitive ability (intelligence rg= -0.40), educational attainment (years of schooling rg = -0.46) and reproductive traits (age at first birth rg= -0.58, father's age at death rg= -0.54). Our findings provide a starting point towards identifying critical biosocial risk mechanisms for the development of ASB.

Main

Antisocial behaviors (ASB) are disruptive acts characterized by covert and overt hostility and violation of the rights and safety of others[1]. The emotional, social, and economic costs incurred by victims of antisocial behavior are far-reaching, ranging from victims' psychological trauma to reduced productivity when victims miss work to costs incurred by taxpayers in order to staff and run a justice system[2,3]. ASB has been recognized not merely as a social problem, but also as a mental health economic priority[4]. In addition of causing harm to others, those with ASB are themselves at elevated risk of criminal convictions as well as mental health and substance abuse problems[5]. Moreover, given the relative stability of ASB[6], it is important to also examine personality traits potentially tied to overt behaviors. Previous meta-analyses demonstrated that the Five-Factor Model of personality (FFM), particularly the domains of Agreeableness, Conscientiousness, and Neuroticism, is potentially critical for better illuminating the correlates and causes of ASB[7,8]. Given all this, it is a research imperative to illuminate the mechanisms underlying the pathogenesis, emergence, and persistence of ASB.

Toward this end, statistical genetic studies have consistently revealed the relevance of environmental and genetic risk factors in the genesis of inter-individual differences in ASB. Family studies - mostly conducted in samples of European ancestry - have demonstrated a considerable heritable component for ASB, with estimates of approximately 50%[9] across studies. The increasing availability of genome-wide data along with data on dimensional ASB measures facilitates in building more advanced explanatory models aimed at identifying trait-relevant genetic variants, that could serve as moderators of socio-environmental factors and vice versa. Moreover, while heritability estimates can differ across subtypes of ASB (e.g., significantly higher twin-based heritability estimates for aggressive forms (65%) versus non-aggressive, rule-breaking forms (48%) of antisocial behavior[10]), these subtypes are genetically correlated (r_g = .38)[11].

Measuring antisocial behavior, a broad view

Considering multiple forms of ASB together increases power of genetic analysis and may improve our ability to detect new genetic variants. Here, we thus examine a broadly defined construct of antisocial behaviors, an approach that has successful precedents. Large-scale genomic studies have indicated

substantial genetic overlap among psychiatric disorders[12]. A recent genome-wide meta-analysis across eight neuropsychiatric disorders revealed extensive pleiotropic genetic effects (N = 232,964 cases and 494,162 controls)[13,14]. The study found that 109 out of the total 146 contributing loci were associated with at least two psychiatric disorders, suggesting broad liability to these conditions. Moreover, the Externalizing Consortium recently conducted a multivariate analysis of large-scale genome-wide association studies (GWAS) of seven externalizing-related phenotypes (N= ~1.5 million) and found 579 genetic associations with a general liability to externalizing behavior[15]. Although these very large multivariate approaches are crucial in enhancing genetic discovery across phenotypes, they do not detect all the genetic variation relevant to individual disorders. Since ASB is a critical issue for psychiatry and for society, the present study uniquely focuses on (severe) forms of ASB and persistence over the lifespan. To do so, we initiated the Broad Antisocial Behavior Consortium (BroadABC), to perform large-scale meta-analytical genetic analyses, utilizing a broad range of phenotypic ASB measures (e.g., conduct disorder symptoms, aggressive behavior, and delinquency). In our first meta-analysis[16], we demonstrated that effect sizes for SNPs with suggestive evidence of association with ASB were small, as anticipated for most polygenic traits. Still, we found that the collective effect across all of the included variants (typically referred to as 'SNP heritability') explained roughly 5% of the total variation in ASB[16], which is in line with meta-analyses of the ACTION[17] and EAGLE[18] consortium.

To date, however, no previous GWAS meta-analysis targeting broad ASB detected SNPs or genes that are well-replicated. The polygenic architecture of ASB underscores the importance of employing very large samples to yield sufficient power to detect genetic loci of small effect size. Therefore, we substantially boost statistical power by quadrupling the sample size and adding new cohorts to the BroadABC consortium. Since ASB is a critical issue for psychiatry and for society, the present study uniquely focuses on (severe) forms of ASB and persistence over the lifespan.

In our meta-analysis, we also include the results of a GWAS study of Disruptive Behavior Disorders (DBDs) in the context of Attention-Deficit/Hyperactivity Disorder (ADHD), which identified three genome-wide significant loci for DBDs[19]. The present study considers multiple measures of antisocial behaviors in people with and without psychiatric diagnoses across 28 samples to reveal the genetic

underpinnings of ASB phenotypes typically studied in psychology, psychiatry, and criminology. These larger samples allow well-powered genetic correlation analyses and improved polygenic risk scores (PRS). Five independent cohorts (total N = 8,058) were employed to validate the ASB PRS in different populations, at different developmental stages, and for different ASB phenotypes. Moreover, we conducted a follow-up analysis by using a mouse model of pathological aggression. Since ASB is known to correlate phenotypically with an array of cognitive and health problems[20–23], we tested for genetic overlap between ASB and a range of other traits and disorders, including anthropometric, cognitive, reproductive, neuropsychiatric, and smoking.

Results

Meta-analysis on broad ASB identifies association with common variants in FOXP2

After quality control and imputation to the Haplotype Reference Consortium or 1000 Genomes Project reference panel (see **Online Methods**), 85,359 individuals from 28 cohorts and a maximum of 7,392,849 variants were available for analysis. We carried out a pooled-sex GWAS meta-analysis for the broad ASB phenotype with METAL[24] and found one genome-wide significant locus, on chromosome 7 (chromosome band 7q31.1, **Fig. 1A**, **Supplementary Table 3**). The top lead SNP was rs12536335 ($P = 6.32 \times 10^{-10}$; **Fig. 1B and 1C**), located in an intronic region upstream of one of the transcriptional start-sites for the forkhead box protein P2 (*FOXP2*) gene[25,26]. Consistent with this finding, a gene-based association test carried out with MAGMA[27], identified a significant association for *FOXP2* ($P = 7.43 \times 10^{-7}$, **Supplementary Note 3, Supplementary Figure 1, Supplementary Table 6**). The *FOXP2* gene has been related to the development of speech and language[28], yet is also implicated in a wide range of other traits and diagnoses[29] (see **Fig. 1D**). MAGMA generalized gene-set and tissue-specific gene-set analyses (sex-combined) yielded no significant gene-sets after Bonferroni-correction for multiple testing. The top gene-set for generalized gene-set analysis was activated NTRK2 signals through RAS signaling pathway (**Supplementary Table 7)**, while the top tissue-specific gene expression was the hypothalamus (**Supplementary Table 8**). We next ran sex-specific GWAS meta-

analyses. These analyses did not identify SNPs that reached genome-wide significance (**Supplementary Tables 4 and 5**).

Mouse model of pathological aggression

Whole genome sequencing analysis of SNVs in aggressive antisocial BALB/cJ mice compared to BALB/cByJ mice controls revealed differences between these lines located in introns of *Foxp2* (rs241912422) and *Cntnap2* (rs212805467; rs50446478; rs260305923; rs242237534), a well-studied neural target of this transcription factor.

Heritability and Polygenic Scoring

SNP heritability

To assess the proportion of variance in liability for broad ASB explained by all measured SNPs, we computed the SNP-based heritability (h^2_{SNP}), which was estimated to be 8.4% (s.e. = 1.2%) by LD score regression (LDSC)[30].

Figure 1: SNP-based results from the GWAS meta-analysis of broad ASB.

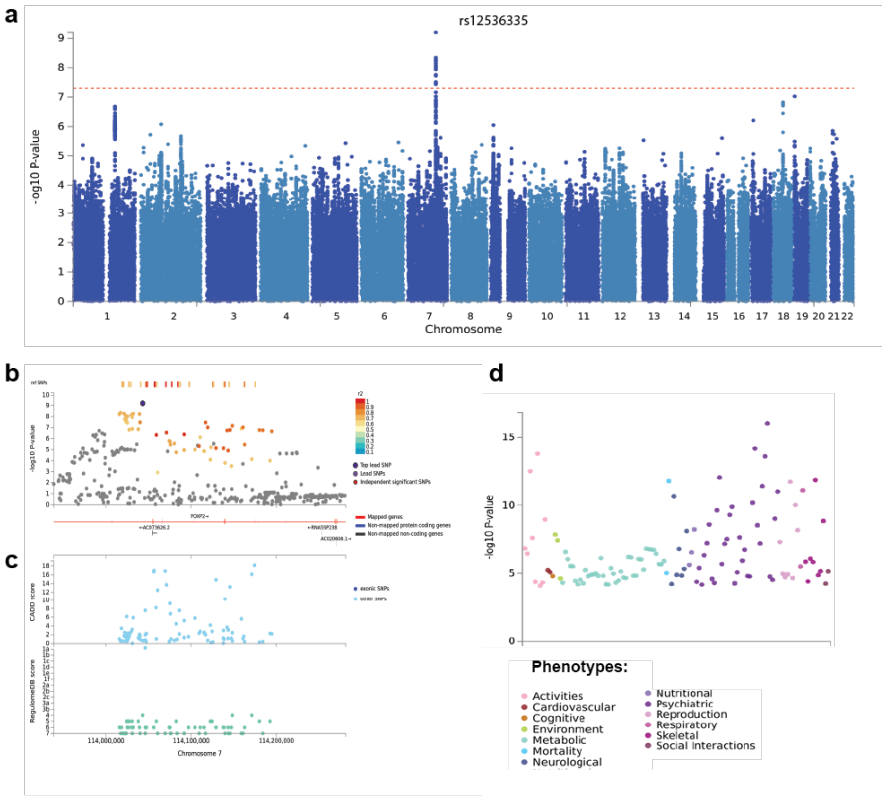

Figure 1. A. Manhattan plot of the GWAS meta-analysis (N = 85,359) of a broad antisocial behavior phenotype, showing the negative log10-transformed P value for each SNP. SNP two-sided P values from a linear model were calculated using METAL[24], weighting SNP associations by sample size. **B**. Regional association plot around chromosome 7:114043159 with functional annotations of SNPs in LD of lead SNP rs12536335 (shown in purple). The plot displays GWAS P-value plotted against its chromosomal position, where colors represent linkage disequilibrium and r^2 values with the most significantly associated SNP. **C.** The plot displays CADD scores (Combined Annotation Dependent Depletion) and RegulomeDB scores of these SNPs. **D**. PheWAS plot showing the significance of associations of common variation in the *FOXP2* gene with a wide range of traits and diagnoses based on MAGMA gene-based tests (with Bonferroni corrected P-value: 1.05e-5), as obtained from GWASAtlas (https://atlas.ctglab.nl).

Figure 2: Polygenic risk score (PRS) associations of broad ASB with six antisocial outcomes in five cohorts.

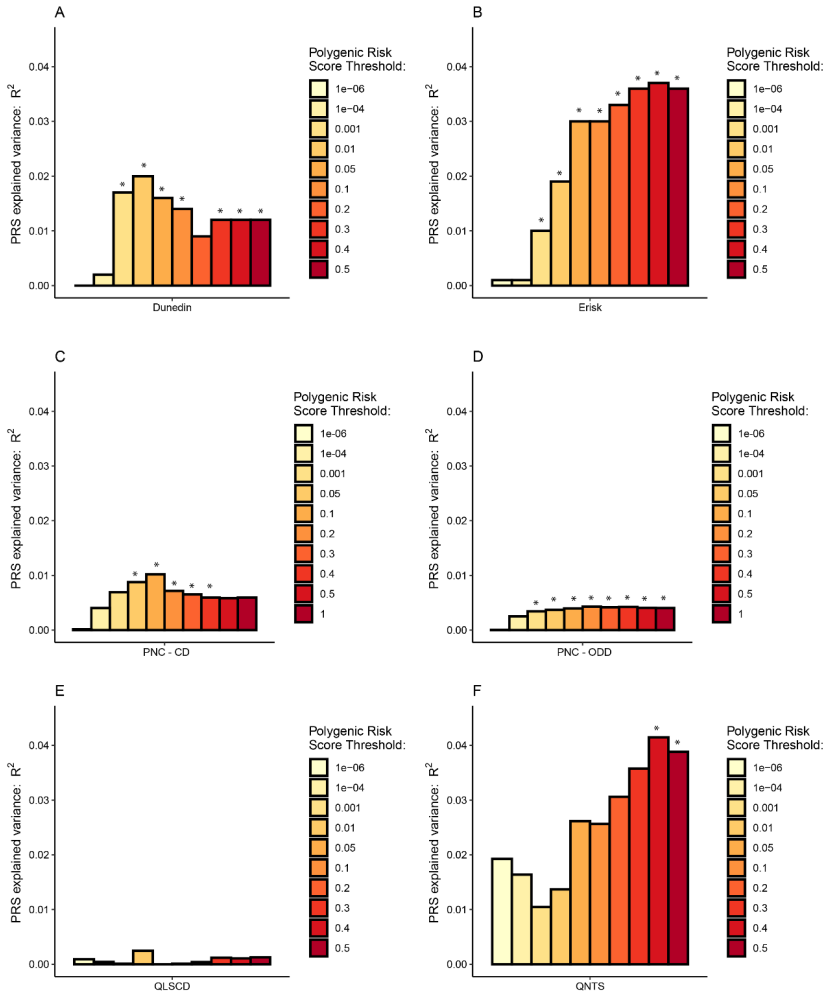

Figure 2. Bar charts illustrating the proportion of variance (incremental R^2, or ΔR^2) explained by the PRSs. PRSs are shown for broad ASB associated with childhood ASB in the Dunedin Longitudinal Study [A], with externalizing behavior in the E-Risk Study [B], with Conduct Disorder [C] and Oppositional Defiant Disorder [D] in the Philadelphia Neurodevelopmental Cohort Study, with ASB in the Quebec Longitudinal Study of Children's

Development Study [E], and with time-aggregated ASB in the Quebec Newborn Twin Study [F]. Asterisks (*) show statistical significance after applying a Bonferroni correction on the 22 tested phenotypes at $P < 0.0023$.

Polygenic Risk Scoring in five independent cohorts

To assess how well the PRS derived from our ASB GWAS meta-analysis predicts other measures of antisocial behavior, we carried out PRS analyses in five independent cohorts (Supplementary Note 7).

Dunedin Longitudinal Study

In New Zealand, participants were derived from the Dunedin Longitudinal Study[31] (N=1,037, assessed 14 times from birth to age 45 years). We tested nine phenotypes and found significant associations with the BroadABC-based PRS for two: childhood ASB and official-records of juvenile convictions. Although not surviving Bonferroni adjustment, we found nominal significant ($P < 0.05$) association with the BroadABC-based PRS for eight phenotypes. We did not find evidence for a PRS association with partner violence. Lastly, we compared individuals grouped into the following four distinct developmental trajectories of antisocial behavior using general growth mixture modeling: low antisocial behavior across childhood through adulthood, childhood-limited antisocial behavior, adolescent-onset antisocial behavior, and life-course persistent antisocial behavior[32]. Individuals following the life-course persistent (LCP) antisocial trajectory were characterized by the highest levels of genetic risk (see Supplementary Figure 2); the nominally significant higher PRS of the LCP trajectory group compared to the low ASB group ($P = 0.032$ and $P = 0.049$, for P-value thresholds 0.05 and 0.1 respectively) did not survive Bonferroni adjustment. For a full report of the findings in the Dunedin cohort, see Supplementary Table 9 and Supplementary Note 8.

Environmental Risk Longitudinal Twin Study (E-Risk)

In England and Wales, participants were included from the E-Risk Study (N=2,232, assessed five times from birth to age 18 years). We tested eight phenotypes and found significant associations for seven. PRS analyses revealed significant associations with parent- and teacher-reported antisocial behavior up to age 12 years, conduct disorder diagnosis up to age 12 years, with the externalizing spectrum at age

18 years, and with official records of criminal convictions up to age 22 years. For a full report of the findings in the E-risk Study, see Supplementary Table 10 and Supplementary Note 8.

Philadelphia Neurodevelopmental Cohort (PNC)

In the United States, participants were included from the PNC Study (N=4,201). We tested two phenotypes and found significant associations for both. We found that higher PRS for ASB were associated with symptom counts of both conduct disorder ($P < 0.0001$, delta R^2=1.0%, Supplementary Table 11) and oppositional defiant disorder ($P < 0.0001$, delta R^2=0.4%, Supplementary Table 12).

Quebec Longitudinal Study of Children's Development (QLSCD)

In Canada, participants were included from the QLSCD study (N=599). We tested one phenotype and did not find a significant association ($P > 0.05$, Supplementary Table 13) between PRS and the score on a self-report questionnaire related to conduct disorder, delinquency, and broad antisocial behavior in young adults (age range= 18-19 years).

Quebec Newborn Twin Study (QNTS)

In Canada, participants were derived from the QNTS study (N=341). We tested two phenotypes and found a significant association for one. We computed a factor score based upon five teacher-rated assessments of ASB in youngsters during primary school (age range= 6-12 years). We found that higher PRS were associated with a higher factor score of ASB ($P = 0.001$, for P-value thresholds .4, adjusted delta R^2=3.9%, Supplementary Table 14). We failed to find evidence for an association between PRS and self-reported antisocial behavior in young adults ($P > 0.05$).

Genetic correlations through LD score regression

ASB is known to correlate with an array of phenotypes[20–22]. At the same time there has been a growing availability of publicly accessible genetic data across these phenotypes. To test whether these phenotypic associations are also reflected in genetic correlations we performed analyses with LDSC in a selection of 73 traits and diagnoses (Supplementary Table 15). We found strong correlations between ASB and reproductive traits (e.g. younger age of first birth ($rg = -0.58$, s.e. = 0.06, $P = 2.93 \times 10^{-15}$)),

cognitive traits (e.g. fewer years of schooling (rg = -0.49, s.e. = 0.06, P = 1.94×10^{-10})), anthropometric traits (e.g. increased waist-to-hip ratio (rg = 0.32, s.e. = 0.05, P = 5.59×10^{-6})), neuropsychiatric traits (e.g. more depressive symptoms (rg = 0.63, s.e. = 0.07, P = 2.45×10^{-16})) and smoking related traits (e.g. ever smoked (rg = 0.54, s.e. = 0.08, P = 1.48×10^{-7})). It is important to emphasize here that correlation, even when genetic, does not necessarily imply causation.

Figure 3: Genetic correlations of traits and diseases that were significantly associated with ASB

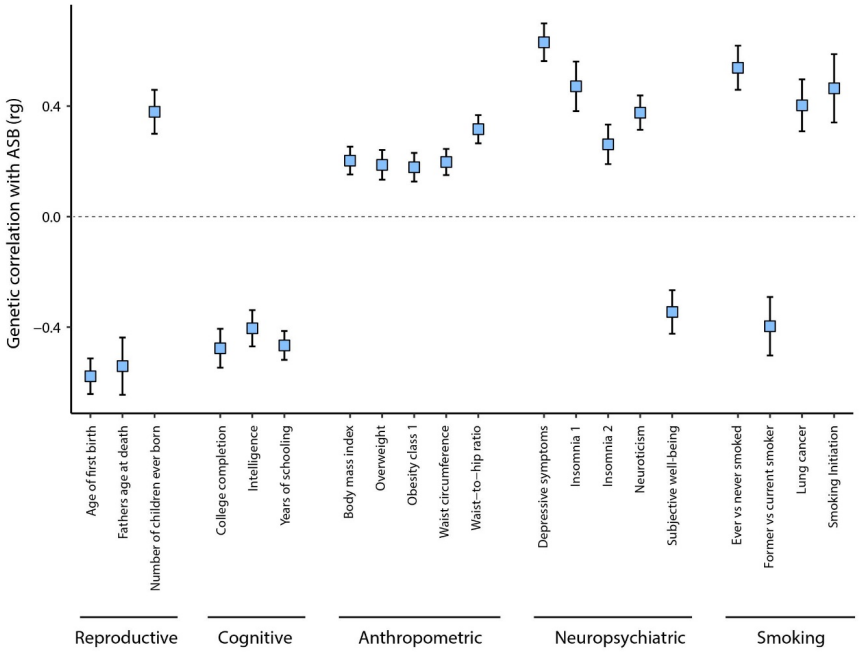

Figure 3. Significant genetic correlations of ASB with previously published results of other traits and diseases, computed using cross-trait LD Score Regression in LDHub, Bonferroni-corrected P-value: 0.00068 (bars represent 95% confidence intervals).

Commented [TJ1]: This Figure still needs to be updated (the results for ADHD will be added)

Discussion

Our GWAS meta-analysis of broad ASB in 85,359 individuals from population cohorts and those with a clinical diagnosis related to ASB, revealed one novel associated locus on chromosome 7 (7:114043159, rs12536335), residing in the forkhead box P2 (*FOXP2)* gene. The lead SNP is relatively proximal (~14kb upstream) to an important enhancer region located 330 kb downstream of the first transcriptional start site (TSS1) of the gene[26]. This SNP is also in the vicinity (~8kb upstream) of a second transcriptional start site (TSS2) of *FOXP2* that can drive expression of alternative transcripts. The *FOXP2* gene is expressed in sensory, limbic, and motor circuits of the brain, as well as the lungs, heart, and gut[26]. It encodes a transcription factor that acts as a regulator of numerous target genes and has been implicated in multiple aspects of brain development (e.g. neuronal growth, synaptic plasticity)[33]. *FOXP2* was first identified two decades ago when rare heterozygous mutations of the gene were linked to a monogenic disorder involving speech motor deficits, accompanied by impairments in expressive and receptive language[34,35(p2)]. Nevertheless, there is scant evidence that common *FOXP2* variants play a large role in the creation of interindividual differences in language function[36,37]. Though prior behavioral research[38,39,40] reported a link between language problems and ASB, it is premature to over-interpret the *FOXP2* findings here. SNPs at this locus have been associated, through GWAS, with a range of externalizing traits, including ADHD[41], cannabis use disorder[42], and generalized risk tolerance[43]. Given the involvement of SNPS at this locus in different behavioral traits and diagnoses, and considering the small effect sizes, it is clear that the association of FOXP2 variation with ASB has limited explanatory value on its own. That said, nothing yet precludes the possibility that this SNP may help to yield deeper insights once placed in broader context by future research.

In the present study we also compared the BALB/cJ strain, a mouse model of pathological aggression, to BALB/cByJ controls, and found intronic variants in *Foxp2* and one of its downstream targets, *Cntnap2.* Previous studies in human cellular models have shown that the protein encoded by *FOXP2* can directly bind to regulatory regions in the *CNTNAP2* locus to repress its expression[44]. Interestingly, mice with cortical-specific knockout of *Foxp2* have been reported to show abnormalities in social behaviors[45].Although these findings may indicate that the intronic SNVs are relevant to the behavioral

differences between the strains, further evidence is needed to show that the variants actually have functional relevance for the mouse phenotype. Future studies may utilize complementary data comparing gene expression in the two mouse lines or could investigate functional impact (e.g. do they map to credible enhancer regions, are they likely to alter binding for transcription factors?) of the SNVs identified.

Contrary to previous BroadABC GWAS analyses, we did not find evidence for sex-specific genetic effects in the present study. Although we did have access to sex-specific data in considerable subsets (N = 22,322 males, N = 26,895 females), the power to detect new variants employing such sample sizes is still limited. Compared to our previous study, we found that the variance explained in independent samples by PRS based on the resulting summary statistics has substantially increased from 0.21% to 3.9%. Essentially, we found consistent links of our ASB PRS with multiple antisocial phenotypes at different developmental stages, from different reporting sources, and reflecting measurements from different disciplines (psychology, psychiatry, criminology). These links were found in individuals from New Zealand, Britain, the United States, and Canada, born 30 years apart. We also show that our ASB PRS were more strongly associated with more severe and persistent types of ASB.

Notwithstanding the increase of effect size of the PRS, and calculations yielding a more precise estimate, the variance explained by the PRS was still relatively small, which was expected in light of the low SNP heritability of 8.4%. Given the highly polygenic architecture of ASB, contributing SNPs have low average effect sizes, thus leading to limited predictive power in independent samples. New PRS methods along with further increasing sample sizes will likely further increase the amount of variance accounted for by the PRS. Moreover, the association may be enhanced by improving the quality of phenotype measurements, which is reflected by our PRS results demonstrating the most robust association with high quality measurement of ASB (using a factor score based upon multiple assessments). Aggregating data from measurements across ages, as opposed to the measures assessed at a single time point, can lead to more reliable trait measures and to better prediction[46]. Phenotypically, adding more extreme ASB phenotypes to the GWAS meta-analysis might also lead to more explained variance. In addition, the inclusion of clinical samples displaying extreme ASB phenotypes (e.g.,

[multiple] homicide, sexual assaults, etc.) in GWAS studies could help increase the ecological validity of genetic findings to forensic populations. Thus, future efforts of the BroadABC will continue to focus on more severe forms of ASB and its persistence across the lifespan. Moreover, by considering genetically correlated traits through multi-trait GWAS methods[47] and multi-trait PRS methods[48] it might be possible to boost power for discovery through GWAS meta-analysis and PRS prediction. Lastly, a major limitation of the present study is that our GWAS results are limited to individuals of European ancestry. This Eurocentric bias may lead to more accurate predictions in individuals with European ancestry, compared to non-Europeans, thus potentially increasing disparities in outcomes related to ASB[49,50]. To realize the full and equitable potential of polygenic risk, future genetic studies on ASB should also include non-European samples.

Developmental criminological research findings, such as the influential developmental taxonomy theory by Moffitt[51,52], have suggested the existence of distinctive offending patterns across the life-course[53]. These developmental trajectories of ASB are thought to have different underlying etiological processes, with relatively more variance explained by genetic factors for life-course-persistent offending as compared to the more socially influenced adolescence-limited offending. Barnes, Beaver and Boutwell have previously produced evidence that heritability estimates were not uniform across different offending groups, suggesting that the causal processes may vary across offending patterns[54,55]. In the present study we found a trend of higher PRS for ASB showing a stronger association with the life-course-persistent trajectory of ASB as compared to the low ASB group. The life-course-persistent trajectory is also known to be associated with profound brain alterations and diminished neurological health[56]. These findings are important since they have the capacity to help improve the current understanding of downstream neurobiological mechanisms relevant to the etiology of antisocial development[56]. Sufficiently powered future studies should thus aim to further elucidate the genetic risk and protective factors that underlie different offending trajectories[57].

Our genetic correlation analyses confirmed previously reported[16,23,58] correlations between ASB and a wide range of traits and diagnoses. The relatively small GWAS sample sizes of some traits, however, coupled with wide confidence intervals (such as agreeableness, rg = -0,81, s.e. = 0,47) call for larger

samples in order to achieve more precise estimates concerning the genetic overlap between personality and ASB. It seems worthwhile to mention again here that partially overlapping genetic architectures does not necessarily provide causal insights of any kind. In this case they merely signify the presence of some potentially shared biological mechanisms linking the conditions[59]. One can reasonably conclude, though, there are likely common underlying genetic factors which operate to increase a general vulnerability to a range of psychopathologies. These comorbid effects are in line with findings in the Dunedin Study demonstrating that life-course-persistent offenders are characterized by several pathological risk factors, related to domains of parenting, neurocognitive development, and temperament[52]. This signifies the importance of investigating pleiotropy and considering the complex etiology of the broader ASB phenotype. Large-scale collaborations, such as the BroadABC, will facilitate the expansion of epidemiological studies capable of further exploring the interaction of genetic risk and socio-environmental risks, and how these contribute to the multifaceted origin of ASB.

Methods

Samples

The meta-analysis included 21 new discovery samples of the BroadABC with GWAS data on a continuous measure of ASB, totaling 50,252 participants: The National Longitudinal Study of Adolescent to Adult Health[60] (ADH), Avon Longitudinal Study of Parents and Children[61–63] (ALSPAC), Brain Imaging Genetics[64] (BIG), CoLaus|PsyCoLaus[65], Collaborative Study on the Genetics of Alcoholism[66] (COGA), Finnish Twin Cohort[67] (FinnTwin), The Genetics of Sexuality and Aggression[68] (GSA), Minnesota Center for Twin and Family Research[69] (MCTFR), Phenomics and Genomics Sample[70] (PAGES), eight samples of the QIMR Berghofer Medical Research Institute (QIMR; 16Up project [16UP[71]], Twenty-Five and Up Study [25UP[72]], Genetics of Human Agency [GHA[73]], Prospective Imaging Study of Ageing [PISA[74]], Semi-Structured Assessment for the Genetics of Alcoholism SSAGA Phase 2 [SS2[75]], Genetic Epidemiology of Pathological Gambling [GA[76]], Twin 89 Study [T89[77]], and Nicotine Study [NC[78]]), Spit for Science[79] (S4S), two samples (from different genotype platforms) of the Twin Early Development Study[80] (TEDS), and the TRacking Adolescents' Individual Lives Survey[81] (TRAILS).

We complemented the above data with GWAS summary statistics on case-control data on disruptive behavior disorders from the recently published Psychiatric Genetics Consortium/iPSYCH consortium meta-analysis, which included data from seven cohorts (Cardiff sample, CHOP cohort, IMAGE-I & IMAGE-II samples, Barcelona sample, Yale-Penn cohort, and the Danish iPSYCH cohort), totaling 3,802 cases and 31,305 controls[19].

We observed a high genetic correlation between the 21 meta-analyzed BroadABC samples and the 7 Psychiatric Genetics Consortium/iPSYCH samples, with the 'Effective N' as weight ($r_g = 0.93$, $P = 9.04 \times 10^{-8}$), indicating strong overlap of genetic effects. Hence, we continued with the combined 28 samples (N= 85,359) for all analyses.

All included studies were approved by local ethics committees, and informed consent was obtained from all of the participants. All study participants were of European ancestry. Full details on demographics, measurements, sample analysis, and quality control are provided in Supplementary Table 1.

Genome-wide association analysis and quality control of individual cohorts

In all 28 discovery samples, genetic variants were imputed using the reference panel of the Haplotype Reference Consortium (HRC) or the 1000G Phase 1 version 3 reference panel. The regression analyses were adjusted for age at measurement, sex, and the first ten principal components. To harmonize the imputation, data preparation, and genome-wide association (GWA) analyses, a specific analysis protocol (Supplementary Note 1) was followed in the 18 BroadABC discovery samples. Further details on the genotyping (platform and quality control criteria), imputation, and GWA analyses for each cohort are provided in Supplementary Table 2.

Two semi-independent analysts (JJT & EU) performed stringent within-cohort quality control, filtering out poor performing SNPs. SNPs were excluded if they met any of the following criteria: study-specific minor allele frequency (MAF) corresponding to a minor allele count (MAC) <100, poor imputation quality ((INFO/R2) score <0.6), and/or Hardy–Weinberg equilibrium $P < 5 \times 10^{-6}$. Moreover, we excluded SNPs and indels that were ambiguous (A/T or C/G with MAF >0.4), duplicated,

monomorphic, multiallelic, or reference-mismatched (Supplementary Note 2, Supplementary Table 17). Then, we visually inspected the distribution of the summary statistics by creating quantile–quantile plots and Manhattan plots for the cleaned summary statistics from each cohort (Supplementary Notes 4, 5 and 6). Discrepancies between the results files of the two semi-independent analysts were examined and errors corrected.

Meta-analyses on combined and sex-specific samples

A meta-analysis of the GWAS results of the 28 discovery samples (N = 85,359) was performed through fixed-effects meta-analysis in METAL, using SNP P-values weighted by sample size. After combining all cleaned GWAS data files, meta-analysis results were filtered to exclude any variants with N < 30,000. Consequently, we removed 2,134,049 SNPs, resulting in 7,392,849 SNPs available for analysis. To investigate sex-specific genetic effects, we also ran the meta-analysis in the datasets for which we had sex-specific data (N = 50,252). However, sex-specific SNP heritabilities, as estimated with LD Score Regression, were small and non-significant (3.7% (s.e. = 2.2%) for males and 1.0% (s.e. = 1.8%) for females). Due to the non-significant sex-specific heritability estimates, the genetic correlation of male and female ASB could not be estimated reliably and no sex-specific follow-up analyses were conducted.

Whole-genome sequencing based on genetic differences between the BALB/c strains

Through whole-genome sequencing, we identified single nucleotide variants that distinguish aggressive BALB/cJ mice from control BALB/cByJ strains[82]. Sequencing libraries were prepared from high-quality genomic DNA using the TruSeq DNA PCR-Free kit (Illumina) and ultra-deep whole genome sequencing (average 30X read-depth across the genome) was performed on a HiSeq X Ten System (Illumina). We developed an efficient data processing and quality control pipeline. Briefly, raw sequencing data underwent stringent quality control and was aligned to either the mm10 (BALB/cJ versus BALB/cByJ strain comparison). Isaac[83] was used to align reads and call single nucleotide variations (SNVs). We excluded SNVs that were covered by less than 20 reads, and that were not present in both animals from the same strain. SnpEff[84] was used to annotate SNVs and explore

functional effects on gene function. SNVs differing between the two strains were annotated to a total of 1573 genes, which were subdivided into three different categories (intronic/exonic non-coding and synonymous variants (1422 genes), untranslated regions (90 genes), missense mutations and splicing variants (61 genes)).

Polygenic Risk Score Analyses

Polygenic risk scores (PRS) were created for ASB using all available SNPs of the discovery dataset[85,86]. PRS were computed as the weighted sum of the effect-coded alleles per individual. We calculated the PRS for subjects of five independent datasets, selected for their detailed phenotypes related to antisocial outcomes: (1) the Dunedin Study[46], (2) the E-risk study[87], (3) the Philadelphia Neurodevelopmental Cohort[88], (4) the Quebec Longitudinal Study of Child Development[89], and (5) the Quebec Newborn Twin Study[90]. All individuals were of European ancestry. To maintain uniformity across target cohorts, we adhered to the following parameters: Clumping was performed by removing markers in linkage disequilibrium, utilizing the following thresholds: maximum r2 = 0.2, window size = 500 kb. We excluded variants within regions of long-range LD[91] (including the Major Histocompatibility Complex, see Supplementary Table 16 for exact regions). Second generation PLINK[92] was employed to construct PRS for each phenotype, at the following 10 thresholds: $P < 1 \times 10^{-6}$, $P < 1 \times 10^{-4}$, $P < 1 \times 10^{-3}$, $P < 1 \times 10^{-2}$, $P < 0.05$, $P < 0.1$, $P < 0.2$, $P < 0.3$, $P < 0.4$, $P < 0.5$. To correct for multiple testing, we applied a Bonferroni correction on the 22 tested phenotypes ($\alpha = 0.00227$).

Genetic correlation analysis

To estimate the genetic correlation between ASB and a range of other phenotypes, we employed Linkage Disequilibrium Score Regression (LDSC)[30] through the LD Hub web portal (http://ldsc.broadinstitute.org/ldhub/)[93]. LD Hub, which is a centralized database of summary-level GWAS results, offers the screening of hundreds of traits. To guard against the possibility of chance findings, we focused on domains of traits that have been previously reported to be comorbid with ASB. Genetic correlations of ASB were thus calculated by selecting 68 phenotypes of health, physiological,

personality, disorder and disease relevant outcomes in LD Hub. In addition, we manually ran genetic correlations analyses for a selection of traits that were not included in the LD Hub framework. Given their relevance for this study, we employed LDSC to examine the genetic correlation of ASB with FFM domains agreeableness, openness to experience and conscientiousness, by using data of the Genetics of Personality Consortium[94]. Similarly, we used data of the most recent GWAS meta-analysis on neuroticism[95] to compute the genetic correlation with ASB. Lastly, while excluding the iPSYCH/PGC samples (given the extensive sample overlap), we computed the genetic correlation of ASB with ADHD[41]. We corrected for multiple testing by applying a Bonferroni correction on the 73 tested genetic correlations ($\alpha = 0.0007$).

Acknowledgements

G.M-V. was supported by a Doctoral Research Scholarship from the FRQSC. M.B., I.O-M., and J-P Gouin are supported by the Canada Research Chair Program. K.R. is supported by a Sir Henry Wellcome Postdoctoral Fellowship. ADB and DD were supported by grants from the Lundbeck Foundation (R102-A9118, R155-2014-1724 and R248-2017-2003), the EU FP7 Program (Grant No. 602805, "Aggressotype") and H2020 Program (Grant No. 667302, "CoCA"), NIMH (1U01MH109514-01. JL, AM, and KR acknowledge the following funders: Academy of Finland, the Signe and Ane Gyllenberg foundation, Juho Vainio foundation, Yrjö Jahnsson foundation, Jalmari and Rauha Ahokas foundation, Sigrid Juselius Foundation and The Finnish Society of Sciences and Letters. DD is supported by NIH R01 AA015416 (Finnish Twin Study), P50 AA022537 (Alcohol Research Center), R25 AA027402 (VCU GREAT), R34 AA027347 (Personalized Risk Assessment), R01 AA028064

(Parental Marital Discord, PI: JS), and U10 AA008401 (COGA) from the National Institute on Alcohol Abuse and Alcoholism (NIAAA), and by R01 DA050721 (Externalizing Consortium) from the National Institute on Drug Abuse (NIDA). BF, MK, and NRM were also supported by funding from the European Community's Horizon 2020 Programme (H2020/2014 – 2020) under grant agreements n° 728018 (Eat2beNICE) and n° 847879 (PRIME). They also received relevant funding from the Netherlands Organization for Scientific Research (NWO) for the Dutch National Science Agenda NeurolabNL project (grant 400-17-602). HMS and MRM are members of the Medical Research Council (MRC) Integrative Epidemiology Unit at the University of Bristol (MC_UU_00011/7). This work is also supported by the NIHR Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care. AW has been supported by funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 602768 for the ACTION consortium. JK has been supported by the Academy of Finland Academy professorship program (grants 265240 and 263278). LC-C was supported by a QIMR Berghofer Research Fellowship. JJM was supported by a QIMR Berghofer International PhD Scholarship. ML was funded by an NHMRC Boosting Dementia Leadership Fellowship (APP1140441). SEM and NGM were funded by NHMRC investigator grants (APP1172917 and APP1172990).

Additional acknowledgements for each study cohort are described in **Supplementary Table 1**.

References

1. Allen JJ, Anderson CA. Aggression and Violence: Definitions and Distinctions. In: *The Wiley Handbook of Violence and Aggression*. American Cancer Society; 2017:1-14. doi:10.1002/9781119057574.whbva001

2. Rivenbark JG, Odgers CL, Caspi A, et al. The high societal costs of childhood conduct problems: evidence from administrative records up to age 38 in a longitudinal birth cohort. *Journal of Child Psychology and Psychiatry*. 2018;59(6):703-710. doi:10.1111/jcpp.12850

3. Brewin CR, Andrews B, Rose S, Kirk M. Acute stress disorder and posttraumatic stress disorder in victims of violent crime. *Am J Psychiatry*. 1999;156(3):360-366. doi:10.1176/ajp.156.3.360

4. Eme R. Male Life-Course-Persistent Antisocial Behavior: The Most Important Pediatric Mental Health Problem. *Archives of Pediatrics & Adolescent Medicine*. 2010;164(5):486-487. doi:10.1001/archpediatrics.2010.44

5. Abram KM, Zwecker NA, Welty LJ, Hershfield JA, Dulcan MK, Teplin LA. Comorbidity and continuity of psychiatric disorders in youth after detention: a prospective longitudinal study. *JAMA Psychiatry*. 2015;72(1):84-93. doi:10.1001/jamapsychiatry.2014.1375

6. Loeber R. The stability of antisocial and delinquent child behavior: a review. *Child development*. 1982;53(6):1431-1446. doi:10.1111/j.1467-8624.1982.tb03465.x

7. Jones SE, Miller JD, Lynam DR. Personality, antisocial behavior, and aggression: A meta-analytic review. *Journal of Criminal Justice*. 2011;39(4):329-337. doi:10.1016/j.jcrimjus.2011.03.004

8. Miller JD, Lynam D. Structural Models of Personality and Their Relation to Antisocial Behavior: A Meta-Analytic Review*. *Criminology*. 2001;39(4):765-798. doi:10.1111/j.1745-9125.2001.tb00940.x

9. Polderman TJC, Benyamin B, de Leeuw CA, et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*. 2015;47(7):702-709. doi:10.1038/ng.3285

10. Burt SA. Are there meaningful etiological differences within antisocial behavior? Results of a meta-analysis. *Clin Psychol Rev*. 2009;29(2):163-178. doi:10.1016/j.cpr.2008.12.004

11. Burt SA, Klump KL. Etiological distinctions between aggressive and non-aggressive antisocial behavior: results from a nuclear twin family model. *J Abnorm Child Psychol*. 2012;40(7):1059-1071. doi:10.1007/s10802-012-9632-9

12. Smoller JW, Andreassen OA, Edenberg HJ, Faraone SV, Glatt SJ, Kendler KS. Psychiatric genetics and the structure of psychopathology. *Molecular Psychiatry*. 2019;24(3):409-420. doi:10.1038/s41380-017-0010-4

13. Lee PH, Anttila V, Won H, et al. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell*. 2019;179(7):1469-1482.e11. doi:10.1016/j.cell.2019.11.020

14. Abdellaoui A, Smit DJA, van den Brink W, Denys D, Verweij KJH. Genomic relationships across psychiatric disorders including substance use disorders. *Drug Alcohol Depend*. 2021;220:108535. doi:10.1016/j.drugalcdep.2021.108535

15. Linnér RK, Mallard TT, Barr PB, et al. Multivariate genomic analysis of 1.5 million people identifies genes related to addiction, antisocial behavior, and health. *bioRxiv*. Published online October 16, 2020:2020.10.16.342501. doi:10.1101/2020.10.16.342501

16. Tielbeek JJ, Johansson A, Polderman TJC, et al. Genome-Wide Association Studies of a Broad Spectrum of Antisocial Behavior. *JAMA Psychiatry*. 2017;74(12):1242-1250. doi:10.1001/jamapsychiatry.2017.3069

17. Ip H, Laan C, Brikell I, et al. *Genetic Association Study of Childhood Aggression across Raters, Instruments and Age*.; 2019. doi:10.1101/854927

18. Pappa I, Pourcain BS, Benke K, et al. A genome-wide approach to children's aggressive behavior: The EAGLE consortium. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 2016;171(5):562-572. doi:10.1002/ajmg.b.32333

19. Demontis D, Walters RK, Rajagopal VM, et al. Risk variants and polygenic architecture of disruptive behavior disorders in the context of attention-deficit/hyperactivity disorder. *Nature Communications*. 2021;12(1):576. doi:10.1038/s41467-020-20443-2

20. Jackson DB, Vaughn MG. Promoting health equity to prevent crime. *Prev Med*. 2018;113:91-94. doi:10.1016/j.ypmed.2018.05.009

21. Burt SA, Donnellan MB. Personality correlates of aggressive and non-aggressive antisocial behavior. *Personality and Individual Differences*. 2008;44(1):53-63. doi:10.1016/j.paid.2007.07.022

22. Piquero AR, Daigle LE, Gibson C, Piquero NL, Tibbetts SG. Research Note: Are Life-Course-Persistent Offenders At Risk for Adverse Health Outcomes? *Journal of Research in Crime and Delinquency*. 2007;44(2):185-207. doi:10.1177/0022427806297739

23. Tielbeek JJ, Boutwell BB. Exploring the Genomic Architectures of Health, Physical Traits and Antisocial Behavioral Outcomes: A Brief Report. *Front Psychiatry*. 2020;11:539. doi:10.3389/fpsyt.2020.00539

24. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26(17):2190-2191. doi:10.1093/bioinformatics/btq340

25. MacDermot KD, Bonora E, Sykes N, et al. Identification of FOXP2 Truncation as a Novel Cause of Developmental Speech and Language Deficits. *Am J Hum Genet*. 2005;76(6):1074-1080.

26. Becker M, Devanna P, Fisher SE, Vernes SC. Mapping of Human FOXP2 Enhancers Reveals Complex Regulation. *Front Mol Neurosci*. 2018;11:47. doi:10.3389/fnmol.2018.00047

27. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput Biol*. 2015;11(4). doi:10.1371/journal.pcbi.1004219

28. den Hoed J, Fisher SE. Genetic pathways involved in human speech disorders. *Current Opinion in Genetics & Development*. 2020;65:103-111. doi:10.1016/j.gde.2020.05.012

29. Watanabe K, Stringer S, Frei O, et al. A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet*. 2019;51(9):1339-1348. doi:10.1038/s41588-019-0481-0

30. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*. 2015;47(3):291-295. doi:10.1038/ng.3211

31. Poulton R, Moffitt TE, Silva PA. The Dunedin Multidisciplinary Health and Development Study: overview of the first 40 years, with an eye to the future. *Soc Psychiatry Psychiatr Epidemiol*. 2015;50(5):679-693. doi:10.1007/s00127-015-1048-8

32. Odgers CL, Moffitt TE, Broadbent JM, et al. Female and male antisocial trajectories: from childhood origins to adult outcomes. *Dev Psychopathol*. 2008;20(2):673-716. doi:10.1017/S0954579408000333

33. den Hoed J, Devaraju K, Fisher SE. Molecular networks of the FOXP2 transcription factor in the brain. *EMBO Rep*. 2021;22(8):e52803. doi:10.15252/embr.202152803

34. Lai CSL, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*. 2001;413(6855):519-523. doi:10.1038/35097076

35. Lai CSL, Gerrelli D, Monaco AP, Fisher SE, Copp AJ. FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain*. 2003;126(Pt 11):2455-2462. doi:10.1093/brain/awg247

36. Uddén J, Hultén A, Bendtz K, et al. Toward Robust Functional Neuroimaging Genetics of Cognition. *J Neurosci*. 2019;39(44):8778-8787. doi:10.1523/JNEUROSCI.0888-19.2019

37. Deriziotis P, Fisher SE. Speech and Language: Translating the Genome. *Trends Genet*. 2017;33(9):642-656. doi:10.1016/j.tig.2017.07.002

38. Warr-Leeper G, Wright NA, Mack A. Language disabilities of antisocial boys in residential treatment. *Behavioral Disorders*. 1994;19(3):159-169.

39. Cohen NJ, Menna R, Vallance DD, Barwick MA, Im N, Horodezky NB. Language, social cognitive processing, and behavioral characteristics of psychiatrically disturbed children with previously identified and unsuspected language impairments. *J Child Psychol Psychiatry*. 1998;39(6):853-864.

40. Brownlie E, Beitchman J, Escobar M, et al. Early Language Impairment and Young Adult Delinquent and Aggressive Behavior. *Journal of abnormal child psychology*. 2004;32:453-467. doi:10.1023/B:JACP.0000030297.91759.74

41. Demontis D, Walters RK, Martin J, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet*. 2019;51(1):63-75. doi:10.1038/s41588-018-0269-7

42. A large-scale genome-wide association study meta-analysis of cannabis use disorder - The Lancet Psychiatry. Accessed December 14, 2020. https://www.thelancet.com/journals/lanpsy/article/PIIS2215-0366(20)30339-4/fulltext

43. Karlsson Linnér R, Biroli P, Kong E, 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(2):245-257. doi:10.1038/s41588-018-0309-3

44. Vernes SC, Newbury DF, Abrahams BS, et al. A functional genetic link between distinct developmental language disorders. *N Engl J Med*. 2008;359(22):2337-2345. doi:10.1056/NEJMoa0802828

45. Medvedeva VP, Rieger MA, Vieth B, et al. Altered social behavior in mice carrying a cortical Foxp2 deletion. *Human Molecular Genetics*. 2019;28(5):701-717. doi:10.1093/hmg/ddy372

46. Caspi A, Houts RM, Ambler A, et al. Longitudinal Assessment of Mental Health Disorders and Comorbidities Across 4 Decades Among Participants in the Dunedin Birth Cohort Study. *JAMA Netw Open*. 2020;3(4). doi:10.1001/jamanetworkopen.2020.3221

47. Turley P, Walters RK, Maghzian O, et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature Genetics*. 2018;50(2):229-237. doi:10.1038/s41588-017-0009-4

48. Maier RM, Zhu Z, Lee SH, et al. Improving genetic prediction by leveraging genetic correlations among human diseases and traits. *Nature Communications*. 2018;9(1):989. doi:10.1038/s41467-017-02769-6

49. Whose genomics? *Nat Hum Behav*. 2019;3(5):409-410. doi:10.1038/s41562-019-0619-1

50. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 2019;51(4):584-591. doi:10.1038/s41588-019-0379-x

51. Moffitt TE. *Adolescence-Limited and Life-Course-Persistent Antisocial Behavior: A Developmental Taxonomy*.; 1993.

52. Moffitt TE. Male antisocial behaviour in adolescence and beyond. *Nature Human Behaviour*. 2018;2(3):177-186. doi:10.1038/s41562-018-0309-4

53. Jennings WG, Reingle JM. On the number and shape of developmental/life-course violence, aggression, and delinquency trajectories: A state-of-the-art review. *Journal of Criminal Justice*. 2012;40(6):472-489. doi:10.1016/j.jcrimjus.2012.07.001

54. Barnes JC, BEAVER K, Boutwell B. Examining the genetic underpinnings of Moffitt's developmental taxonomy: A behavioral genetic analysis. *Criminology*. 2011;49:923-954. doi:10.1111/j.1745-9125.2011.00243.x

55. Motz RT, Barnes JC, Caspi A, et al. Does contact with the justice system deter or promote future delinquency? Results from a longitudinal study of British adolescent twins. *Criminology*. 2020;58(2):307-335. doi:https://doi.org/10.1111/1745-9125.12236

56. Carlisi CO, Moffitt TE, Knodt AR, et al. Associations between life-course-persistent antisocial behaviour and brain structure in a population-representative longitudinal birth cohort. *The Lancet Psychiatry*. 2020;7(3):245-253. doi:10.1016/S2215-0366(20)30002-X

57. Wertz J, Caspi A, Belsky DW, et al. Genetics and Crime: Integrating New Genomic Discoveries Into Psychological Research About Antisocial Behavior. *Psychol Sci*. 2018;29(5):791-803. doi:10.1177/0956797617744542

58. Tielbeek JJ, Barnes JC, Popma A, et al. Exploring the genetic correlations of antisocial behaviour and life history traits. *BJPsych Open*. 2018;4(6):467-470. doi:10.1192/bjo.2018.63

59. Martin J, Taylor MJ, Lichtenstein P. Assessing the evidence for shared genetic risks across psychiatric disorders and traits. *Psychol Med*. 2018;48(11):1759-1774. doi:10.1017/S0033291717003440

60. Harris KM, Halpern CT, Whitsel EA, et al. Cohort Profile: The National Longitudinal Study of Adolescent to Adult Health (Add Health). *International Journal of Epidemiology*. 2019;48(5):1415-1415k. doi:10.1093/ije/dyz115

61. Boyd A, Golding J, Macleod J, et al. Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*. 2013;42(1):111-127. doi:10.1093/ije/dys064

62. Northstone K, Lewcock M, Groom A, et al. The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019. *Wellcome Open Res*. 2019;4. doi:10.12688/wellcomeopenres.15132.1

63. Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol*. 2013;42(1):97-110. doi:10.1093/ije/dys066

64. Franke B, Vasquez AA, Veltman JA, Brunner HG, Rijpkema M, Fernández G. Genetic variation in CACNA1C, a gene associated with bipolar disorder, influences brainstem rather than gray matter volume in healthy individuals. *Biol Psychiatry*. 2010;68(6):586-588. doi:10.1016/j.biopsych.2010.05.037

65. Preisig M, Waeber G, Vollenweider P, et al. The PsyCoLaus study: methodology and characteristics of the sample of a population-based survey on psychiatric disorders and their association with genetic and cardiovascular risk factors. *BMC Psychiatry*. 2009;9:9. doi:10.1186/1471-244X-9-9

66. The Collaborative Study on the Genetics of Alcoholism. *Alcohol Health Res World*. 1995;19(3):228-236.

67. Kaprio J. The Finnish Twin Cohort Study: an update. *Twin Res Hum Genet*. 2013;16(1):157-162. doi:10.1017/thg.2012.142

68. Johansson A, Jern P, Santtila P, et al. The Genetics of Sexuality and Aggression (GSA) twin samples in Finland. *Twin Res Hum Genet*. 2013;16(1):150-156. doi:10.1017/thg.2012.108

69. Miller MB, Basu S, Cunningham J, et al. The Minnesota Center for Twin and Family Research Genome-Wide Association Study. *Twin Res Hum Genet*. 2012;15(6):767-774. doi:10.1017/thg.2012.62

70. Rujescu D, Hartmann AM, Giegling I, et al. Genome-Wide Association Study in Vestibular Neuritis: Involvement of the Host Factor for HSV-1 Replication. *Front Neurol*. 2018;9:591. doi:10.3389/fneur.2018.00591

71. Mitchell BL, Kirk KM, McAloney K, et al. 16Up: Outline of a Study Investigating Wellbeing and Information and Communication Technology Use in Adolescent Twins. *Twin Res Hum Genet*. 2020;23(6):345-357. doi:10.1017/thg.2020.83

72. Mitchell BL, Campos AI, Rentería ME, et al. Twenty-Five and Up (25Up) Study: A New Wave of the Brisbane Longitudinal Twin Study. *Twin Res Hum Genet*. 2019;22(3):154-163. doi:10.1017/thg.2019.27

73. Morosoli JJ, Colodro-Conde L, Barlow FK, Medland SE. Investigating perceived heritability of mental health disorders and attitudes toward genetic testing in the United States, United

677 Kingdom, and Australia. *Am J Med Genet B Neuropsychiatr Genet*. Published online September
678 25, 2021. doi:10.1002/ajmg.b.32875

679 74. Lupton MK, Robinson GA, Adam RJ, et al. A prospective cohort study of prodromal Alzheimer's
680 disease: Prospective Imaging Study of Ageing: Genes, Brain and Behaviour (PISA). *Neuroimage*
681 *Clin*. 2021;29:102527. doi:10.1016/j.nicl.2020.102527

682 75. Madden PA, Bucholz KK, Dinwiddie SH, et al. Nicotine withdrawal in women. *Addiction*.
683 1997;92(7):889-902.

684 76. Slutske WS, Meier MH, Zhu G, Statham DJ, Blaszczynski A, Martin NG. The Australian Twin
685 Study of Gambling (OZ-GAM): Rationale, Sample Description, Predictors of Participation, and a
686 First Look at Sources of Individual Differences in Gambling Involvement. *Twin Research and*
687 *Human Genetics*. 2009;12(1):63-78. doi:10.1375/twin.12.1.63

688 77. Knopik VS, Heath AC, Madden PAF, et al. Genetic effects on alcohol dependence risk: re-
689 evaluating the importance of psychiatric and other heritable risk factors. *Psychol Med*.
690 2004;34(8):1519-1530. doi:10.1017/s0033291704002922

691 78. Saccone SF, Pergadia ML, Loukola A, et al. Genetic Linkage to Chromosome 22q12 for a Heavy-
692 Smoking Quantitative Trait in Two Independent Samples. *Am J Hum Genet*. 2007;80(5):856-
693 866.

694 79. Dick DM, Nasim A, Edwards AC, et al. Spit for Science: launching a longitudinal study of genetic
695 and environmental influences on substance use and emotional health at a large US university.
696 *Front Genet*. 2014;5:47. doi:10.3389/fgene.2014.00047

697 80. McAra L, McVie S. Youth crime and justice: Key messages from the Edinburgh Study of Youth
698 Transitions and Crime. *Criminology & Criminal Justice*. 2010;10(2):179-209.
699 doi:10.1177/1748895809360971

700 81. Oldehinkel AJ, Rosmalen JG, Buitelaar JK, et al. Cohort Profile Update: the TRacking
701 Adolescents' Individual Lives Survey (TRAILS). *Int J Epidemiol*. 2015;44(1):76-76n.
702 doi:10.1093/ije/dyu225

703 82. Jager A, Amiri H, Bielczyk N, et al. Cortical control of aggression: GABA signalling in the anterior
704 cingulate cortex. *Eur Neuropsychopharmacol*. 2020;30:5-16.
705 doi:10.1016/j.euroneuro.2017.12.007

706 83. Raczy C, Petrovski R, Saunders CT, et al. Isaac: ultra-fast whole-genome secondary analysis on
707 Illumina sequencing platforms. *Bioinformatics*. 2013;29(16):2041-2043.
708 doi:10.1093/bioinformatics/btt314

709 84. Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of
710 single nucleotide polymorphisms, SnpEff. *Fly (Austin)*. 2012;6(2):80-92. doi:10.4161/fly.19695

711 85. Choi SW, Mak TSH, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses.
712 *Nature Protocols*. 2020;15(9):2759-2772. doi:10.1038/s41596-020-0353-1

713 86. Wray NR, Lin T, Austin J, et al. From Basic Science to Clinical Application of Polygenic Risk
714 Scores: A Primer. *JAMA Psychiatry*. Published online September 30, 2020.
715 doi:10.1001/jamapsychiatry.2020.3049

87. Wertz J, Agnew-Blais J, Caspi A, et al. From Childhood Conduct Problems to Poor Functioning at Age 18 Years: Examining Explanations in a Longitudinal Cohort Study. *J Am Acad Child Adolesc Psychiatry*. 2018;57(1):54-60.e4. doi:10.1016/j.jaac.2017.09.437

88. Satterthwaite TD, Connolly JJ, Ruparel K, et al. The Philadelphia Neurodevelopmental Cohort: A publicly available resource for the study of normal and abnormal brain development in youth. *Neuroimage*. 2016;124(Pt B):1115-1119. doi:10.1016/j.neuroimage.2015.03.056

89. Orri M, Boivin M, Chen C, et al. Cohort Profile: Quebec Longitudinal Study of Child Development (QLSCD). *Soc Psychiatry Psychiatr Epidemiol*. 2021;56(5):883-894. doi:10.1007/s00127-020-01972-z

90. Boivin M, Brendgen M, Dionne G, et al. The Quebec Newborn Twin Study at 21. *Twin Res Hum Genet*. 2019;22(6):475-481. doi:10.1017/thg.2019.74

91. Price AL, Weale ME, Patterson N, et al. Long-Range LD Can Confound Genome Scans in Admixed Populations. *Am J Hum Genet*. 2008;83(1):132-135. doi:10.1016/j.ajhg.2008.06.005

92. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7. doi:10.1186/s13742-015-0047-8

93. Zheng J, Erzurumluoglu AM, Elsworth BL, et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics*. 2017;33(2):272-279. doi:10.1093/bioinformatics/btw613

94. de Moor MHM, Costa PT, Terracciano A, et al. Meta-analysis of genome-wide association studies for personality. *Mol Psychiatry*. 2012;17(3):337-349. doi:10.1038/mp.2010.128

95. Nagel M, Jansen PR, Stringer S, et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet*. 2018;50(7):920-927. doi:10.1038/s41588-018-0151-7