

## ORIGINAL ARTICLE

## ASD and schizophrenia show distinct developmental profiles in common genetic overlap with population-based social communication difficulties

B St Pourcain<sup>1,2,3,22</sup>, EB Robinson<sup>4,5,22</sup>, V Anttila<sup>4,5</sup>, BB Sullivan<sup>4,5</sup>, J Maller<sup>4</sup>, J Golding<sup>6</sup>, D Skuse<sup>7</sup>, S Ring<sup>1,2</sup>, DM Evans<sup>1,8</sup>, S Zammit<sup>1,9</sup>, SE Fisher<sup>3,10</sup>, BM Neale<sup>4,5,11</sup>, RJA Anney<sup>11</sup>, S Ripke<sup>4,5,12</sup>, MV Hollegaard<sup>13</sup>, T Werge<sup>14,15,16</sup>, iPSYCH-SSI-Broad Autism Group, A Ronald<sup>17</sup>, J Grove<sup>14,18,19,20</sup>, DM Hougaard<sup>13</sup>, AD Børglum<sup>14,18,19</sup>, PB Mortensen<sup>14,19,21</sup>, MJ Daly<sup>4,5,23</sup> and G Davey Smith<sup>1,2,23</sup>

Difficulties in social communication are part of the phenotypic overlap between autism spectrum disorders (ASD) and schizophrenia. Both conditions follow, however, distinct developmental patterns. Symptoms of ASD typically occur during early childhood, whereas most symptoms characteristic of schizophrenia do not appear before early adulthood. We investigated whether overlap in common genetic influences between these clinical conditions and impairments in social communication depends on the developmental stage of the assessed trait. Social communication difficulties were measured in typically-developing youth (Avon Longitudinal Study of Parents and Children,  $N \leq 5553$ , longitudinal assessments at 8, 11, 14 and 17 years) using the Social Communication Disorder Checklist. Data on clinical ASD (PGC-ASD: 5305 cases, 5305 pseudo-controls; iPSYCH-ASD: 7783 cases, 11 359 controls) and schizophrenia (PGC-SCZ2: 34 241 cases, 45 604 controls, 1235 trios) were either obtained through the Psychiatric Genomics Consortium (PGC) or the Danish iPSYCH project. Overlap in genetic influences between ASD and social communication difficulties during development decreased with age, both in the PGC-ASD and the iPSYCH-ASD sample. Genetic overlap between schizophrenia and social communication difficulties, by contrast, persisted across age, as observed within two independent PGC-SCZ2 subsamples, and showed an increase in magnitude for traits assessed during later adolescence. ASD- and schizophrenia-related polygenic effects were unrelated to each other and changes in trait-disorder links reflect the heterogeneity of genetic factors influencing social communication difficulties during childhood versus later adolescence. Thus, both clinical ASD and schizophrenia share some genetic influences with impairments in social communication, but reveal distinct developmental profiles in their genetic links, consistent with the onset of clinical symptoms.

*Molecular Psychiatry* advance online publication, 3 January 2017; doi:10.1038/mp.2016.198

## INTRODUCTION

The phenotypic overlap between autism spectrum disorder (ASD) and schizophrenia is complex and dates back to Kanner in 1943.<sup>1</sup> Individuals affected by either condition display deficits in the ability to initiate and maintain reciprocal interaction.<sup>2</sup> This includes impairments in social cognition<sup>3,4</sup> but also poor social competence<sup>5</sup> affecting verbal and nonverbal communication skills. Recent cross-disorder genetic analyses highlighted the continuity of psychiatric phenotypes beyond current diagnostic boundaries.<sup>6</sup> The nature of shared genetic influences between childhood

neurodevelopmental disorders, such as ASD, and adult-onset psychiatric illnesses, like schizophrenia, however, remains less well understood.

ASD represent a group of neurodevelopmental conditions with a typical age of onset before the age of 3 years affecting ~ 1 to 2% of children.<sup>7,8</sup> Core features include deficits in social interaction and communication, as well as highly restricted interests and/or stereotyped repetitive behaviours.<sup>2</sup> By contrast, schizophrenia is an adult-onset psychiatric illness with a typical first-time diagnosis between 16 and 30 years. The disorder has a lifetime prevalence of

<sup>1</sup>MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK; <sup>2</sup>School of Social and Community Medicine, University of Bristol, Bristol, UK; <sup>3</sup>Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; <sup>4</sup>Analytic and Translational Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA; <sup>5</sup>Stanley Center for Psychiatric Research and Medical and the Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA; <sup>6</sup>Centre for Child and Adolescent Health, University of Bristol, Bristol, UK; <sup>7</sup>Behavioural and Brain Sciences, Institute of Child Health, University College London, London, UK; <sup>8</sup>University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, QLD, Australia; <sup>9</sup>MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK; <sup>10</sup>Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands; <sup>11</sup>Department of Psychiatry, Harvard Medical School, Boston, MA, USA; <sup>12</sup>Department of Psychiatry and Psychotherapy, Charité - Universitätsmedizin Berlin, Campus Mitte, Berlin, Germany; <sup>13</sup>Statens Serum Institut, Department of Congenital Disorders, Copenhagen, Denmark; <sup>14</sup>The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Aarhus, Denmark; <sup>15</sup>Institute of Biological Psychiatry, MHC Sct. Hans, Mental Health Services Copenhagen, Copenhagen, Denmark; <sup>16</sup>Institute of Psychological Sciences, Faculty of Medicine and Health Sciences, University of Copenhagen, Copenhagen, Denmark; <sup>17</sup>Department of Psychological Sciences, Birkbeck, University of London, London, UK; <sup>18</sup>Department of Biomedicine, Aarhus University, Aarhus, Denmark; <sup>19</sup>Centre for Integrative Sequencing, iSEQ, Aarhus University, Aarhus, Denmark; <sup>20</sup>Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark and <sup>21</sup>National Centre for Register-based Research, Aarhus University, Aarhus, Denmark. Correspondence: Dr B St Pourcain, Language and Genetics Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, Nijmegen 6525 XD, The Netherlands. E-mail: Beate.StPourcain@mpi.nl

<sup>22</sup>These authors contributed equally to this work.

<sup>23</sup>These authors codirected the study.

Received 15 March 2016; revised 10 June 2016; accepted 1 August 2016

~1%<sup>9</sup> and is characterised by hallucinations, delusions, disorganised speech or behaviour, apathy and lack of emotional reactivity.<sup>2</sup> Both ASD and schizophrenia are highly heritable<sup>10,11</sup> and recent studies have linked different types of genetic variation including common variants,<sup>10,12,13</sup> as well as rare inherited<sup>14,15</sup> and *de novo* variation<sup>16,17</sup> to risk of illness in both conditions. Contemporary research strongly supports a genetic overlap between ASD and schizophrenia for rare copy number variants<sup>18</sup> and rare *de novo* mutation events<sup>16</sup> with converging evidence for gene sets involved in synaptic function.<sup>16</sup> The role of shared common genetic risk between ASD and schizophrenia, however, is less clear. Common genetic influences account for 25 to 33%<sup>19</sup> of total liability to schizophrenia and up to 49% of total liability to ASD.<sup>10,12</sup> Despite this, the common genetic overlap between ASD and schizophrenia is small compared with the overlap between psychiatric adult-onset only disorders.<sup>12,20</sup>

The framework of Research Domain Criteria (RDoC), including social communication difficulties, now actively facilitates the study of functional dimensions spanning the full range of human behaviour from normal to abnormal and across development.<sup>21</sup> Common disorders, due to their polygenic architecture, can be understood as quantitative traits.<sup>22</sup> For ASD, following the findings of earlier twin studies,<sup>23,24</sup> there is now molecular evidence for shared common genetic influences with social communication difficulties during childhood.<sup>25</sup> The genetic continuity of social interaction and communication deficits in schizophrenia has not yet been observed though it can be hypothesised that such common genetic links exist given the impairments in social cognition within first-degree relatives of schizophrenia patients.<sup>3</sup>

Impaired abilities in social communication in affected children are heritable (twin- $h^2 = 0.74$ )<sup>26</sup> and a large part of these genetic influences can be captured through common single-nucleotide polymorphisms (SNPs; SNP- $h^2 \leq 0.45$ ).<sup>27</sup> Beside some stable genetic influences,<sup>28</sup> genetic factors underlying social interaction impairments and social communication difficulties vary during development,<sup>27,28</sup> especially for common variation.<sup>27</sup> Thus, we hypothesise that also the genetic overlap between social communication difficulties and clinically recognised disorder may change during childhood and adolescence.

The primary aim of this study is to examine the nature of common polygenic influences in ASD and schizophrenia through their genetic overlap with phenotypic symptoms in the general population that are shared between both conditions, but differ according to developmental stage. We predict that if social communication difficulties are part of a common shared aetiology between ASD and schizophrenia, trait-disorder relationships for both conditions should follow similar patterns. Dissimilar patterns due to independent genetic influences would be expected for a non-shared genetic aetiology. Here, we report developmental profiles in common genetic overlap for both ASD and schizophrenia with respect to longitudinal measures of social communication difficulties within the general population. Analyses are based on the largest publicly available genome-wide data for ASD<sup>29</sup> and schizophrenia,<sup>13</sup> in addition to a large Danish ASD sample from the iPSYCH project and a deeply-phenotyped UK birth cohort, the Avon Longitudinal Study of Parents and Children (ALSPAC).<sup>30,31</sup>

## MATERIALS AND METHODS

### Genome-wide summary statistics

*Population-based social communication difficulties.* Genome-wide association studies (GWASs) were carried out in ALSPAC participants, a UK population-based longitudinal pregnancy-ascertained birth cohort (estimated birth date: 1991–1992).<sup>30,31</sup> Ethical approval was obtained from the ALSPAC Law-and-Ethics Committee (IRB00003312) and the Local Research-Ethics Committees, written informed consent was obtained from a parent

or individual with parental responsibility and assent was obtained from child participants.

ALSPAC children were genotyped using the Illumina HumanHap550 quad-chip and imputation was performed on 8237 children and 477,482 SNP genotypes using a 1000 Genomes reference (Phase1\_v3, <http://www.1000genomes.org/>)<sup>32</sup> (Supplementary Methods).

Quantitative social communication problems in ALSPAC children were assessed with the 12-item Social Communication Disorder Checklist (SCDC; score-range: 0 to 24).<sup>26</sup> The SCDC is a brief screening instrument of social reciprocity and verbal/nonverbal communication (for example, 'Not aware of other people's feelings'), with high reliability and good validity,<sup>26</sup> which has been extensively investigated.<sup>26,27,33</sup> Higher SCDC scores reflect more social communication deficits and are positively skewed (Supplementary Figure 1). Mother-reported scores for children and adolescents were repeatedly measured at 8, 11, 14 and 17 years (Supplementary Table 1) and are inter-correlated (Spearman's  $\rho$ : 0.39 to 0.57, Supplementary Table 2). Information on phenotypic and genotypic data was available for 4175 to 5553 children (Table 1).

SCDC scores were residualised for sex, age and the two most significant ancestry-informative principal components<sup>34</sup> and then rank-transformed (Supplementary Figure 2). Transformed scores showed similar correlation patterns as untransformed scores (Pearson's  $r$ : 0.38 to 0.61, Supplementary Table 2).

Genome-wide single marker summary statistics were generated by regressing rank-transformed residuals on allele dosages using SNPTTEST<sup>35</sup> (without genomic control-based correction<sup>36</sup>).

*Clinical ASD.* The Psychiatric Genomics Consortium (PGC) has completed a genome-wide scan of 5305 ASD cases and their parents (PGC-ASD), all of European ancestry (2015 freeze; summary results at <http://www.med.unc.edu/pgc/>). An ASD diagnosis was confirmed using research standard diagnoses and expert clinical consensus diagnoses. 94.1% of all ASD cases had also a diagnosis of autism from the Autism Diagnostic Interview-Revised<sup>37</sup> and/or the Autism Diagnostic Observation Schedule.<sup>38</sup> Genome-wide data were imputed to a 1000 Genomes reference (Phase1\_v3) and genetic association studied using a case and pseudo-control design.<sup>29</sup> This design is robust to population stratification as pseudo-controls are based on un-transmitted parental alleles, and thus cases and pseudo-controls are ancestrally matched. To replicate findings, we analysed ASD GWAS summary results in the Danish iPSYCH project (iPSYCH-ASD: 7783 ASD cases, 11 359 controls) using samples from the Danish Neonatal Screening Biobank hosted by Statens Serum Institute (Supplementary Methods). The iPSYCH-ASD project aims to genotype all Danish individuals with available DNA from bloodspots and an ASD diagnosis (International Classification of Diseases<sup>39</sup>) in their medical record. iPSYCH-ASD has been genotyped using the Illumina Infinium PsychArray BeadChip and genotypes were imputed to a 1000 Genomes template (Phase1\_v3). This study has been approved by the Danish research ethical committee system.

Note that also a small number of ALSPAC children with clinical ASD ( $N \leq 36$ ) has been included in this study (Supplementary Methods).

*Clinical schizophrenia.* A large PGC mega-analysis on schizophrenia has been carried out studying individuals of predominantly European descent<sup>13</sup> (Summary results at <http://www.med.unc.edu/pgc/>). Cases met diagnostic criteria for either schizophrenia or schizoaffective disorder.<sup>13</sup> Here, we investigated two non-overlapping schizophrenia subsets: (1) PGC-SCZ1 (11 958 cases, 12 710 controls), constructed as part of the first PGC mega-analysis of schizophrenia,<sup>13,40</sup> and (2) PGC-SCZ2i, containing novel PGC-SCZ2 cases and controls not included in PGC-SCZ1 (22 283 cases, 32 894 controls, 1235 trios).<sup>13</sup> In addition, we studied the combined PGC-SCZ2 sample (PGC-SCZ1+PGC-SCZ2i: 34 241 cases, 45 604 controls, 1235 trios) of the second PGC mega-analysis of schizophrenia.<sup>13</sup> As PGC-SCZ2 contains 1836 cases and 3383 controls from East Asia, we also studied a PGC-SCZ2 sample of European ancestry only (PGC-SCZ2-Eur: 32 405 cases, 42 221 controls, 1235 trios). Genome-wide data were imputed to a 1000 Genomes template (Phase1\_v3).

The studied population-based and clinical samples (Table 1) contain no sample overlap.

*Other adult-onset disorders.* To analyse the specificity of genetic overlap between SCDC scores and schizophrenia, we studied further adult-onset psychiatric disorders, such as major depressive disorder (MDD) and bipolar disorder (BIP; Supplementary Methods).

**Table 1.** Genome-wide summary statistics

Sample	Source	Phenotype/diagnosis	Ethnicity	N	LD score correlation	Study design
ALSPAC	General population	Mother-reported SCDC scores	White European	5553 (8 years) 5462 (11 years) 5060 (14 years) 4175 (17 years)	$r_g$ with respect to SCDC scores	Polygenic effect of risk-increasing alleles on SCDC scores <sup>a</sup>
PGC-ASD iPSYCH-ASD	Clinical sample Clinical sample	ASD ASD	White European White European	5305 cases; 5305 pseudo-controls 7783 cases; 11 359 matched controls	Discovery sample Replication sample	Discovery sample <sup>b,c</sup>
PGC-SCZ1	Clinical sample, subset of PGC-SCZ2	Schizophrenia or schizoaffective disorder	White European	11 958 cases; 12 710 controls	Independent sample	Independent sample <sup>c</sup>
PGC-SCZ2i	Clinical sample, subset of PGC-SCZ2	Schizophrenia or schizoaffective disorder	Predominantly white European	22 283 cases; 32 894 controls; 1235 trios (including 1836 cases and 3383 controls of East Asian ancestry)	Independent sample	Independent sample <sup>c</sup>
PGC-SCZ2	Clinical sample	Schizophrenia or schizoaffective disorder	Predominantly white European	34 241 cases; 45 604 controls; 1235 trios (including 1836 cases and 3383 controls of East Asian ancestry)	Combined sample	Combined sample <sup>b,c</sup>
PGC-SCZ2-Eur	Clinical sample	Schizophrenia or schizoaffective disorder	White European	32 405 cases; 42 221 controls; 1235 trios	Combined sample <sup>d</sup>	—

Abbreviations: ALSPAC, Avon Longitudinal study of Parents and Children; ASD, autism spectrum disorder; iPSYCH-iPADS, iPSYCH-SSI-BROAD Autism project; LD, linkage disequilibrium; PGC, Psychiatric Genomics Consortium; PGC-SCZ1, Samples of the first PGC mega-analysis of SCZ; PGC-SCZ2i, PGC-SCZ2 samples not analysed within PGC-SCZ1; PGC-SCZ2, Samples of the second PGC mega-analysis of SCZ (PGC-SCZ1 +PGC-SCZ2i); PGC-SCZ2-Eur, PGC-SCZ2 participants of European ancestry only (exclusion of 1836 cases and 3383 controls from East Asia);<sup>13</sup> PGS, polygenic scores;  $r_g$ , genetic correlation; SCDC, Social Communication Disorder Checklist; SCZ, schizophrenia. All samples were imputed to a 1000 genomes reference (Phase1\_v3); Note that there is no overlap between population-based and clinical samples. <sup>a</sup>ALSPAC is target sample. <sup>b</sup>Largest publicly available set of genome-wide summary statistics. <sup>c</sup>Sensitivity analysis. <sup>d</sup>Clinical samples are training sets.

**Statistical methods**

Linkage disequilibrium (LD) score regression<sup>41</sup> was applied to estimate the cumulative effect of common SNPs on either variation in developmental SCDC scores or risk to disorder (SNP- $h^2$ ), using GWAS statistics and exploiting LD patterns in the genome. LD score correlation<sup>20</sup> analysis was carried out to estimate genetic correlations ( $r_g$ ) between SCDC scores and clinical conditions, or among clinical conditions, that is, the extent to which two phenotypes share common genetic factors, based on GWAS statistics. All analyses were performed with LDSC software<sup>20,41</sup> using HapMap3 markers<sup>42</sup> (Supplementary Methods).

Polygenic risk scores (PGS)<sup>43,44</sup> were analysed to estimate the explained phenotypic variance in social communication difficulties due to risk-increasing alleles for clinical disorder. Using a range of  $P$ -value thresholds ( $0.001 < P_T \leq 1$ ), PGS for ASD (based on PGC-ASD), schizophrenia (based on PGC-SCZ2) and schizophrenia subsamples (based on PGC-SCZ1 and PGC-SCZ2i) were generated in ALSPAC (Table 1) using imputed genotypes (1000 Genomes, Phase1\_v3, INFO > 0.8). For this, common autosomal signals observed in clinical samples (with MAF > 0.01 in ALSPAC) were clumped ( $LD-r^2 > 0.25$ ,  $\pm 500$  kb) consistent with current guidelines<sup>45</sup> using PLINK,<sup>46</sup> excluding duplicate SNPs (Supplementary Methods). Rank-transformed SCDC scores were regressed on Z-standardised PGS (Ordinary least square regression, R software Rv3.2.2, <https://cran.r-project.org/>), and the proportion of phenotypic variance explained by each PGS predictor reported as adjusted regression  $R^2$ . Note that assuming an infinitely large clinical 'discovery' sample, the regression  $R^2$  is equivalent to the product of  $r_g$  squared and the heritability of the explained Z-standardised trait.<sup>44</sup>

Mixed Poisson regression (Rlme4 library) was utilised to test the trend in common genetic overlap longitudinally using untransformed SCDC scores. Repeatedly assessed SCDC score counts were regressed on ASD- and schizophrenia-PGS with overdispersion being accounted for through the random error part.<sup>47</sup> Models included fixed effects for ASD-PGS and schizophrenia-PGS, sex, age at assessment, as well as random intercepts. Beta-coefficients for PGS quantify here the increase in natural-log SCDC scores for each increase in one standard deviation of PGS. Differences in sample-dropout across time were accounted for through bootstrapping, generating parametric 95%-bootstrap confidence intervals ( $N_{\text{bootstrap}} = 500$ ). We have not estimated adjusted  $R^2$ -related measures due to the difficulty of defining the residual variance for non-Gaussian responses, especially within a mixed model context.<sup>48</sup>

Genome-wide Complex Trait Analysis (GCTA)<sup>49,50</sup> was utilised to estimate SNP- $h^2$  and genetic correlations among SCDC scores, as published previously,<sup>27</sup> for comparison only (Supplementary Methods).

Attrition analysis in ALSPAC studied the relationship between SCDC-missingness at each assessed age and PGS for clinical ASD and schizophrenia (Supplementary Methods).

**RESULTS**

**SNP-heritabilities for social communication difficulties and psychiatric disorder**

Genome-wide analyses of population-based SCDC scores at 8, 11, 14 and 17 years provided little evidence for bias in GWAS statistics due to population stratification. The estimated LDSC- $h^2$  intercepts were consistent with one, ranging from 0.988 (s.e. = 0.0067) to 1.009 (s.e. = 0.0070; Table 2). In subsequent analyses LDSC- $h^2$  intercepts were thus constrained to one, including LDSC correlation analyses.

Cumulative influences of SNPs on variation in SCDC scores were strongest at the age of 8, 11 and 17 years with LDSC- $h^2$  estimates of 0.19 (s.e. = 0.06), 0.17 (s.e. = 0.07) and 0.30 (s.e. = 0.11), respectively (Table 2). The estimates were lower, however, at 14 years (LDSC- $h^2$  = 0.08 (s.e. = 0.06)). These LDSC-based findings mirrored closely GCTA- $h^2$  estimates using GREML (Table 2), although latter might potentially be biased.<sup>51</sup> SCDC scores shared furthermore genetic factors across development (GREML  $r_g$  = 0.38 (s.e. = 0.16) to 0.95 (s.e. = 0.34),  $P_{\text{min}} = 2 \times 10^{-7}$ ), as previously reported,<sup>27</sup> with lower correlations across wider age gaps (Supplementary Table 3).

A common genetic basis for ASD has been described earlier<sup>12,25</sup> including PGC-ASD (liability-scale LDSC- $h^2$  = 0.23 (s.e. = 0.03))<sup>25</sup> and iPSYCH-ASD (liability-scale LDSC- $h^2$  = 0.14 (s.e. = 0.03)),<sup>25</sup> with strong evidence for similar polygenic architectures among samples ( $r_g$  = 0.74 (s.e. = 0.07),  $P < 10^{-20}$ ).<sup>25</sup> Also, it is known<sup>12,13,40</sup> that



**Table 2.** LD-score regression and GCTA results for SCDC scores in ALSPAC

SCDC score	LD score regression				GCTA <sup>a</sup>		
	Unconstrained <sup>b</sup>	Constrained <sup>c</sup>	$\lambda_{GC}$	Mean $\chi^2$	N	$h^2(SE)^a$	N <sup>d</sup>
	Intercept(SE)	$h^2(SE)$					
8 y	0.992 (0.0067)	0.19 (0.06)	1.023	1.022	5553	0.24 (0.07)	5137
11 y	1.000 (0.0065)	0.17 (0.07)	1.014	1.019	5462	0.17 (0.07)	5058
14 y	0.988 (0.0067)	0.08 (0.06)	1.005	1.009	5060	0.08 (0.07)	4735
17 y	1.009 (0.0070)	0.30 (0.11)	1.029	1.025	4175	0.45 (0.08)	3978

Abbreviations: ALSPAC, Avon Longitudinal study of Parents and Children; GCTA, genome-wide complex trait analysis;  $h^2$ , SNP heritability; LD, linkage disequilibrium; SCDC, Social Communication Disorder Checklist; y, age at assessment in years;  $\lambda_{GC}$ , Genomic inflation factor. <sup>a</sup>Findings correspond closely to previously published estimates.<sup>27</sup> <sup>b</sup>LD score regression using an unconstrained intercept. <sup>c</sup>LD score regression constraining the intercept for the SNP- $h^2$  estimation to one. <sup>d</sup>Differences compared with the total sample *N* are due to the exclusion of individuals with a relatedness of  $\geq 2.5\%$ .

common genetic factors influence schizophrenia liability. Liability-scale LDSC-SNP- $h^2$  estimates for PGC-SCZ1, PGC-SCZ2i, PGC-SCZ2Eur and PGC-SCZ2 were 0.31 (s.e.=0.02), 0.24 (s.e.=0.01), 0.25 (s.e.=0.01) and 0.25 (s.e.=0.01), respectively (assumed population-prevalence of 0.01), with strong evidence for shared genetic factors among independent samples (PGC-SCZ1 and PGC-SCZ2i:  $r_g=0.96$  (s.e.=0.024),  $P < 10^{-20}$ ).

#### Genetic correlations between social communication difficulties and psychiatric disorder

As part of a two-stage analysis design (Table 1), we used constrained LD score correlation to study the genetic overlap between psychiatric disorder and social communication problems during development. Genetic correlations between rank-transformed social communication difficulties and clinical ASD decreased in point estimates with progressing age of the trait (Figure 1a, Supplementary Table 4). For PGC-ASD, the genetic link with SCDC scores was strongest at 8 years ( $r_g=0.34$  (s.e.=0.15),  $P=0.027$ )<sup>25</sup> and attenuated by 17 years ( $r_g=0.01$  (s.e.=0.12),  $P=0.94$ ). This pattern was replicated in iPSYCH-ASD ( $r_g=0.35$  (s.e.=0.13),  $P=0.008$  and  $r_g=0.02$  (s.e.=0.10),  $P=0.81$ , respectively, Supplementary Table 4). In contrast, common genetic links between schizophrenia and social communication difficulties during childhood and adolescence persisted and increased in point estimates (Figure 1b). Within PGC-SCZ1, genetic overlap with SCDC scores started to emerge at 8 years ( $r_g=0.20$  (s.e.=0.08),  $P=0.01$ ) and was strongest at 17 years ( $r_g=0.24$  (s.e.=0.08),  $P=0.004$ ; Figure 1b, Supplementary Table 4). The genetic link during later adolescence was replicated in PGC-SCZ2i (age 17:  $r_g=0.15$  (s.e.=0.06),  $P=0.011$ , Figure 1b) and also observed in the combined PGC-SCZ2 sample (PGC-SCZ1+PGC-SCZ2i:  $r_g=0.18$  (s.e.=0.06),  $P=0.003$ , Supplementary Table 4). These findings were not affected by the presence of a small proportion of individuals of Asian origin (PGC-SCZ2-Eur:  $r_g=0.18$  (s.e.=0.06),  $P=0.004$ , Supplementary Table 4). Importantly, other PGC adult-onset disorders, such as MDD and BIP, showed no correlations with SCDC scores (Age 17: MDD- $r_g=-0.05$  (s.e.=0.11),  $P=0.65$  and BIP- $r_g=0.04$  (s.e.=0.08),  $P=0.62$ , Supplementary Table 4) suggesting that findings are specific to schizophrenia. Note that LD-score correlations between schizophrenia and ASD ( $r_g=0.20$  (s.e.=0.05),  $P=0.00011$ ) were modest, compared with considerably stronger links between schizophrenia and other adult-onset disorders (for example, BIP- $r_g=0.76$  (s.e.=0.04),  $P=6.5 \times 10^{-70}$ , Supplementary Table 5), as previously reported.<sup>20</sup>

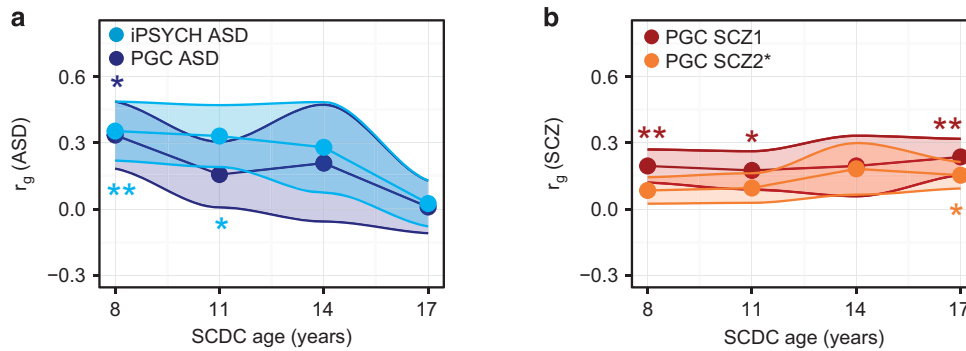
For comparison, we also analysed trait-disorder overlap using LD score correlation without constraining intercepts (Supplementary Table 4). In the presence of genetic links,

unconstrained  $r_g$ -point estimates were, overall, in close correspondence with constrained estimates, but had wider standard errors.

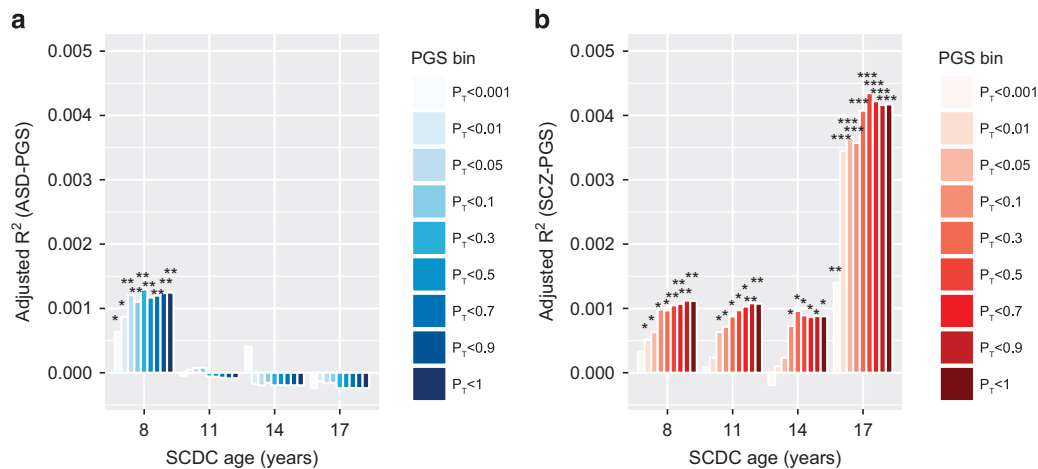
#### Polygenic scores for risk-increasing alleles predicting social communication difficulties

To provide an absolute measure of shared genetic influences between traits and clinically recognised conditions, we assessed the phenotypic variance in rank-transformed social communication difficulties due to risk-increasing alleles using polygenic scoring<sup>43,44</sup> (Table 1). Alleles more common in ASD cases than in pseudo-controls were only associated with variation in SCDC scores at 8 years (PGC-ASD: adjusted  $R^2_{max}=0.13\%$ ,  $P_{min}=0.0042$ , Figure 2a, Supplementary Table 6). In contrast, alleles more often present in schizophrenia cases than controls explained predominantly variation in social communication difficulties at 17 years, based on risk alleles in both PGC-SCZ subsamples (PGC-SCZ1: adjusted  $R^2_{max}=0.26\%$ ,  $P_{min}=0.00058$ ; PGC-SCZ2i: adjusted  $R^2_{max}=0.19\%$ ,  $P_{min}=0.0028$ ; Supplementary Table 7) and the combined PGC-SCZ2 sample (adjusted  $R^2_{max}=0.43\%$ ,  $P_{min}=0.000012$ , Figure 2b, Supplementary Table 6). Excluding ALSPAC children with a clinical ASD diagnosis had little influence on the reported changes in genetic effect (Supplementary Table 8). Importantly, adjustment of ASD-PGS and schizophrenia-PGS for each other did not affect the nature of these findings, suggesting the independence of ASD- and schizophrenia-related polygenic influences (Supplementary Table 6).

To assess developmental trends in common genetic trait-disorder overlap, we modelled the effect of ASD-PGS and schizophrenia-PGS on untransformed SCDC scores longitudinally. Applying a mixed Poisson model, we found evidence for age-specific changes in genetic effects for both ASD-PGS and schizophrenia-PGS (Supplementary Table 9). For example, at  $P_T < 0.05$  (Figure 3), a threshold shown to predict schizophrenia case-ness in independent samples,<sup>13</sup> the effect of ASD-PGS decreased with progressing age of the trait (ASD-PGS  $\times$  SCDC-age: Beta = -0.0031 (s.e.=0.0014),  $P=0.019$ , 95%-bootstrapped confidence interval: -0.0057 to -0.00035), while the effect of schizophrenia-PGS increased (schizophrenia-PGS  $\times$  SCDC-age: Beta = 0.0029 (s.e.=0.0014),  $P=0.030$ , 95%-bootstrapped confidence-interval: 0.00047 to 0.0054). Consistent with the findings for rank-transformed scores, ASD-related polygenic influences on SCDC score counts were strongest during childhood (age 8: Beta = 0.047 (s.e.=0.017),  $P=0.0056$ ; age 17: Beta = 0.019 (s.e.=0.018),  $P=0.29$ ), while schizophrenia-related polygenic effects were more pronounced during later adolescence (age 8: Beta = 0.046 (s.e.=0.017),  $P=0.0080$ ; age 17: Beta = 0.072 (s.e.=0.018),



**Figure 1.** Genetic correlations between (a) clinical ASD and (b) clinical schizophrenia and SCDC scores during development. Genetic correlations between clinical disorder and rank-transformed SCDC scores in ALSPAC (at 8, 11, 14 and 17 years) were estimated cross-sectionally using LD-score correlation analysis<sup>20</sup> and are shown with their standard errors (shaded). Standard error distributions for SCDC scores across age were approximated using loess.  $P$ -values indicate the probability that the true genetic correlation is different from zero ( $*P \leq 0.05$ ,  $**P \leq 0.01$ ). ALSPAC, Avon Longitudinal study of Parents and Children; ASD, autism spectrum disorder; iPSYCH-ASD, iPSYCH-SSI-BROAD Autism project; LD, linkage disequilibrium; PGC, Psychiatric Genomics Consortium; PGC-ASD, ASD collection of the PGC; PGC-SCZ1, Samples of the first PGC mega-analysis of SCZ; PGC-SCZ2i, PGC-SCZ2 samples not analysed within PGC-SCZ1; SCDC, Social Communication Disorder Checklist; SCZ, schizophrenia.



**Figure 2.** Proportion of variance in SCDC scores explained by polygenic scores for (a) clinical ASD and (b) clinical schizophrenia. Polygenic scores were constructed in ALSPAC based on the largest publicly available samples for ASD (PGC-ASD) and schizophrenia (PGC-SCZ2) as a training set, and then Z-standardised. The proportion of explained phenotypic variance in rank-transformed SCDC scores (adjusted regression  $R^2$ ) is displayed cross-sectionally at 8, 11, 14 and 17 years and given with respect to ASD-PGS (a) and schizophrenia-PGS (b). Nine different  $P$ -value thresholds  $P_T$  for selecting risk alleles (PGS bins) in clinical samples are displayed. Starred  $P$ -values indicate the strength of the association ( $*P \leq 0.05$ ,  $**P \leq 0.01$ ). ALSPAC, Avon Longitudinal study of Parents and Children; ASD, autism spectrum disorder; PGC-ASD, ASD collection of the PGC; PGC, Psychiatric Genomics Consortium; PGC-SCZ2, Samples of the second PGC mega-analysis of SCZ; PGS, polygenic scores; PGS bin, Z-standardised polygenic scores according to threshold  $P_T$ ; SCDC, Social Communication Disorder Checklist; SCZ, schizophrenia.

$P = 0.000056$ ). Similar developmental changes in genetic overlap were also found for other PGS thresholds (Supplementary Table 9).

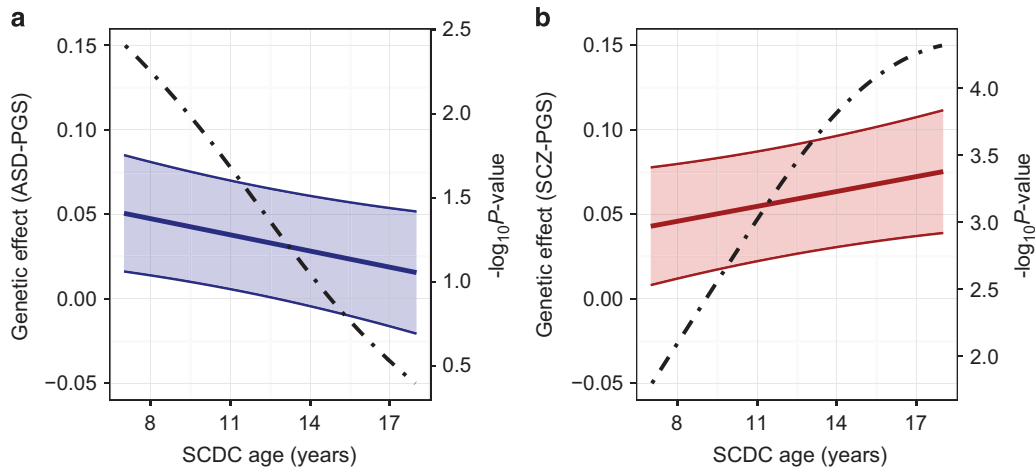
#### Attrition in ALSPAC

Analyses of SCDC-missingness in ALSPAC were carried out to investigate potential sources of bias (Supplementary Table 10). Using for simplicity a PGS threshold of  $P_T < 0.05$ , there was little evidence for a relationship between sample-dropout and ASD-PGS, especially after adjustment for maternal educational level (age 8: odds ratio = 0.99 (s.e. = 0.03),  $P = 0.82$ ), although there was support for an association with schizophrenia-PGS (age 17: odds ratio = 1.10 (s.e. = 0.03),  $P = 0.000050$ ), consistent with previous studies.<sup>52</sup>

#### DISCUSSION

This study provided evidence for shared common genetic overlap between social communication difficulties and both ASD and schizophrenia, but does not imply a shared genetic susceptibility between these clinical conditions. Instead, we identified distinct patterns in genetic trait-disorder relationships, largely consistent with the onset of clinical symptoms. Genetic links were driven by independent polygenic influences and showed opposite trends in magnitude with progressing age of the population-based trait, as supported by longitudinal analyses.

Genetic overlap with ASD was strongest for social communication difficulties during middle childhood ( $r_g \sim 33\%$ ), in line with recent cross-sectional studies,<sup>25</sup> while those with schizophrenia was strongest for social communication difficulties during later



**Figure 3.** Developmental changes in genetic effects of polygenic scores for (a) clinical ASD and (b) clinical schizophrenia on SCDC scores. Polygenic scores (PGS) were constructed in ALSPAC based on the largest publicly available samples for ASD (PGC-ASD) and schizophrenia (PGC-SCZ2) as a training set, and then Z-standardised. A  $P$ -value threshold of  $P_T < 0.05$  for selecting risk alleles in clinical samples is displayed. Using a mixed Poisson regression framework, longitudinal measures of untransformed SCDC score counts were regressed on ASD-PGS and schizophrenia-PGS simultaneously allowing for changes in genetic effects over time. Repeatedly assessed SCDC score counts in ALSPAC were available at 8, 11, 14 and 17 years of age with individual ages ranging between 7 to 18 years. Genetic effects for ASD-PGS (a) and their 95% confidence intervals (shaded) as well as schizophrenia-PGS (b) and their 95% confidence intervals (shaded) were estimated across development, and show the increase in SCDC log counts per standard deviation in PGS score. A dotted line indicates the  $P$ -value of the genetic effect. ALSPAC, Avon Longitudinal study of Parents and Children; ASD, autism spectrum disorder; PGC-ASD, ASD collection of the PGC; PGC, Psychiatric Genomics Consortium; PGC-SCZ2, Samples of the second PGC mega-analysis of SCZ; SCDC, Social Communication Disorder Checklist; SCZ, schizophrenia.

adolescence ( $r_g \sim 18\%$ ). Complementary estimates were provided by polygenic scoring analyses. Up to 0.13% phenotypic variation in social communication difficulties could be explained by ASD risk-increasing alleles during childhood and up to 0.43% phenotypic variation by schizophrenia risk-increasing alleles during later adolescence, independently of each other. The genetic overlap with social communication difficulties during later adolescence was not observed for other adult-onset disorders, such as BIP, despite their strong genetic links with psychosis,<sup>12</sup> making unspecific age-related influences unlikely. Thus, our findings suggest that social communication impairments are part of the genetically influenced phenotypic spectrum of schizophrenia.

Changes in genetic overlap over time need to be viewed within the context of cohort-specific sampling properties and clinical sample power. For instance, it is possible that the genetic overlap between schizophrenia and social communication difficulties has been underestimated, as SCDC-missingness, and more generally study non-participation,<sup>52</sup> has been related to common genetic risk for schizophrenia. In contrast, there was little evidence for a link between SCDC missingness and common ASD risk. In addition, mother-report of social communication difficulties may have contributed to enhanced variance sharing among population-based traits, and thus underestimated the true variation in child genetic effects. Finally, the studied clinical discovery sets differed in their inherent power. For example, the power<sup>44</sup> of ASD-PGS (PGC-ASD  $\sim 5000$  trios) was only 0.58 compared with 0.99 for schizophrenia-PGS (PGC-SCZ2,  $N \sim 80\,000$ ), assuming a liability-scale  $SNP-h^2$  of 0.23 for ASD and 0.25 for schizophrenia, a disease prevalence of 0.01, a type-I-error rate of 0.05 and a population-based target sample of 5000 individuals. Under attrition, such as a decrease of  $\sim 1000$  ALSPAC participants during later adolescence, the power of ASD-PGS would further drop (to 0.49), while the power of schizophrenia-PGS remains largely unaffected (0.99). Longitudinal analyses, adjusting for differences in population-based sample numbers across time through bootstrapping, suggested, however, that the

observed developmental changes in polygenic risk effects are robust, even in the presence of sample dropout.

Our results have direct relevance for the definition of RDoC<sup>21</sup> within a developmental context. The lack of support for shared polygenic effects between ASD and schizophrenia, with respect to social communication impairments, is in agreement with recent studies. Molecular analyses of PGC samples reported modest correlations between ASD and schizophrenia<sup>20,29</sup> ( $r_g \sim 0.20$ ), confirmed within this study, and twin research suggested little genetic overlap between autistic traits and psychotic experiences.<sup>53</sup> The absence of shared aetiological factors strengthens furthermore positions suggesting that the exact nature of social deficits implicated within ASD and schizophrenia differs from each other.<sup>54</sup> Here we show that common genetic variation underlying complex disorders can be dissected through temporal changes in the genetic architecture of behavioural symptoms that are shared between disorders. Thus, a developmental analysis of genetic relationships between population-based and clinical samples can be informative with regard to the dimensional nature of psychiatric illness without discarding the aetiology of different disorders, a concern often raised with respect to RDoC.<sup>21</sup>

The identification of distinct patterns in common genetic overlap between social communication difficulties and psychiatric illness is consistent with the presence of multiple distinct genetic influences contributing to variation in social communication behaviour during development. While genetic factors underlying SCDC scores across  $\sim 3$ -year intervals are stable and shared by at least 80%, only  $\sim 50\%$  of common genetic influences are shared across 10 years intervals.<sup>27</sup> This may suggest developmental (but not rapid) changes in the phenotype capture by the SCDC with progressing age. For example, it is possible that the behavioural phenotypes influencing SCDC scores at age 8 or 11 years are, in terms of average composition, different from those influencing the scale at age 17. Social communication abilities comprise many components, such as social interaction, social cognition, pragmatic and language processing skills (<http://www.asha.org/>), some of which will vary during child and adolescent development,

including changes in the social-cognitive understanding of friendship and peer interaction.<sup>55</sup> One might envisage that these phenotypic changes reflect distinct genetic factors driving different stages of postnatal brain development.<sup>56</sup> In addition, social communication difficulties have been linked to behavioural problems.<sup>57</sup> Note that the SCDC has a high sensitivity but a lower specificity in discriminating ASD from the non-ASD patients in the presence of other clinical disorders.<sup>26</sup> Thus, the SCDC is likely to capture multiple behavioural and cognitive dimensions related to social communication problems during the course of child and adolescent development, spanning around 10 years, which give rise to distinct patterns in trait-disorder overlap. This poses questions on the nature of genetic influences affecting variation in social communication impairments across development that will require exploration with longitudinal genome-wide approaches and biological network analyses.

## CONCLUSIONS

Social communication difficulties are phenotypically shared with both ASD and schizophrenia and show common genetic overlap with both disorders. These polygenic links manifest, however, as distinct developmental profiles and do not imply a shared genetic susceptibility between these clinical conditions.

## CONFLICT OF INTEREST

TW has acted as lecturer and advisor to the H. Lundbeck A/S. The remaining authors declare no conflict of interest.

## ACKNOWLEDGMENTS

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. We also thank the Psychiatric Genomics Consortium for providing access to genome-wide summary statistics for clinical ASD and schizophrenia samples. This publication is the work of the authors and they will serve as guarantors for the contents of this paper. The UK Medical Research Council and the Wellcome Trust (102215/2/13/2) and the University of Bristol provide core support for ALSPAC. The ALSPAC GWAS data was generated by Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. Autism Speaks (7132) provided support for the analysis of autistic-trait related data in ALSPAC to BStP. EBR was funded by the NIMH grant 1K01MH099286-01A1 and NARSAD Young Investigator grant 22379. BStP and SEF are supported by the Max Planck Society. The iPSYCH project is funded by the Lundbeck Foundation and the universities and university hospitals of Aarhus and Copenhagen. Genotyping of iPSYCH and PGC samples was supported by grants from the Lundbeck Foundation, the Stanley Foundation, the Simons Foundation (SFARI 311789 to MJD), and NIMH (5U01MH094432-02 to MJD). High-performance computer capacity for handling and statistical analysis of iPSYCH data was provided by the Centre for Integrative Sequencing, iSEQ, Aarhus University, Denmark.

## REFERENCES

- 1 Kanner L. Autistic disturbances of affective contact. *Nerv Child* 1943; **2**: 217–250.
- 2 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Association: Washington, DC, USA, 1994.
- 3 Penn DL, Sanna LJ, Roberts DL. Social Cognition in Schizophrenia: An Overview. *Schizophr Bull* 2008; **34**: 408–411.
- 4 Baron-Cohen S. Theory of mind and autism: a review. In: Glidden LM (ed). *International Review of Research in Mental Retardation: Autism* vol. 23. Academic Press: San Diego, CA, USA, 2001, pp 169–184.
- 5 Dickinson D, Bellack AS, Gold JM. Social/Communication Skills, Cognition, and Vocational Functioning in Schizophrenia. *Schizophr Bull* 2007; **33**: 1213–1220.
- 6 Owen MJ. New approaches to psychiatric diagnostic classification. *Neuron* 2014; **84**: 564–571.
- 7 Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008. Principal investigators, centers for disease control and prevention. prevalence of autism spectrum disorders. *MMWR* 2002; **2012**: 1–19.

- 8 Baron-Cohen S, Scott FJ, Allison C, Williams J, Bolton P, Matthews FE et al. Prevalence of autism-spectrum conditions: UK school-based population study. *Br J Psychiatry* 2009; **194**: 500–509.
- 9 Saha S, Chant D, Welham J, McGrath J. A systematic review of the prevalence of schizophrenia. *PLoS Med* 2005; **2**: e141.
- 10 Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB et al. Most genetic risk for autism resides with common variation. *Nat Genet* 2014; **46**: 881–885.
- 11 Shih RA, Belmonte PL, Zandi PP. A review of the evidence from family, twin and adoption studies for a genetic contribution to adult psychiatric disorders. *Int Rev Psychiatry Abingdon Engl* 2004; **16**: 260–283.
- 12 Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* 2013; **45**: 984–994.
- 13 Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–427.
- 14 Krumm N, Turner TN, Baker C, Vives L, Mohajeri K, Witherspoon K et al. Excess of rare, inherited truncating mutations in autism. *Nat Genet* 2015; **47**: 582–588.
- 15 Loohuis LMO, Ori AP, Staats KA, Wang T, Richards AL et al. Genome-wide burden of deleterious coding variants increased in schizophrenia. *Nat Commun* 2015; **6**: 7501.
- 16 Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P et al. De novo mutations in schizophrenia implicate synaptic networks. *Nature* 2014; **506**: 179–184.
- 17 Iossifov I, O’Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* 2014; **515**: 216–221.
- 18 Doherty JL, Owen MJ. Genomic insights into the overlap between psychiatric disorders: implications for research and clinical practice. *Genome Med* 2014; **6**: 29.
- 19 Sullivan PF, Daly MJ, O’Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet* 2012; **13**: 537–551.
- 20 Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet* 2015; **47**: 1236–1241.
- 21 Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med* 2013; **11**: 126.
- 22 Plomin R, Haworth CMA, Davis OSP. Common disorders are quantitative traits. *Nat Rev Genet* 2010; **10**: 872.
- 23 Robinson EB. Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Arch Gen Psychiatry* 2011; **68**: 1113.
- 24 Lundström S, Chang Z, Rastam M, Gillberg C, Larsson H, Anckarsäter H et al. Autism spectrum disorders and autisticlike traits: similar etiology in the extreme end and the normal variation. *Arch Gen Psychiatry* 2012; **69**: 46–52.
- 25 Robinson EB, St Pourcain B, Anttila V, Kosmicki JA, Bulik-Sullivan B, Grove J et al. Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nat Genet* 2016; **48**: 552–555.
- 26 Skuse DH, Mandy WPL, Scourfield J. Measuring autistic traits: heritability, reliability and validity of the Social and Communication Disorders Checklist. *Br J Psychiatry* 2005; **187**: 568–572.
- 27 St Pourcain B, Skuse DH, Mandy WP, Wang K, Hakonarson H, Timpson NJ et al. Variability in the common genetic architecture of social-communication spectrum phenotypes during childhood and adolescence. *Mol Autism* 2014; **5**: 18.
- 28 Holmboe K, Rijdsdijk FV, Hallett V, Happé F, Plomin R, Ronald A. Strong Genetic Influences on the Stability of Autistic Traits in Childhood. *J Am Acad Child Adolesc Psychiatry* 2014; **53**: 221–230.
- 29 Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013; **381**: 1371–1379.
- 30 Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J et al. Cohort Profile: The ‘Children of the 90s’—the Index Offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2013; **42**: 111–127.
- 31 Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Smith GD et al. Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* 2013; **42**: 97–110.
- 32 Consortium T 1000 GP. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491**: 56–65.
- 33 St Pourcain B, Wang K, Glessner JT, Golding J, Steer C, Ring SM et al. Association between a high-risk autism locus on 5p14 and social communication spectrum phenotypes in the general population. *Am J Psychiatry* 2010; **167**: 1364–1372.
- 34 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**: 904–909.



- 35 Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007; **39**: 906–913.
- 36 Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999; **55**: 997–1004.
- 37 Lord C, Rutter M, Le Couteur A. Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord* 1994; **24**: 659–685.
- 38 Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC *et al*. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord* 2000; **30**: 205–223.
- 39 The ICD-10 Classification of Mental and Behavioural Disorders. *Clinical Descriptions and Diagnostic Guidelines*, 1 edn. World Health Organization: Geneva, Switzerland, 1992.
- 40 Consortium TSPG-WAS (GWAS). Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**: 969–976.
- 41 Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium *et al*. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; **47**: 291–295.
- 42 International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations. *Nature* 2010; **467**: 52–58.
- 43 Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF *et al*. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
- 44 Dudbridge F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet* 2013; **9**: e1003348.
- 45 Wray NR, Lee SH, Mehta D, Vinkhuyzen AAE, Dudbridge F, Middeldorp CM. Research Review: polygenic methods and their application to psychiatric traits. *J Child Psychol Psychiatry* 2014; **55**: 1068–1087.
- 46 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D *et al*. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* 2007; **81**: 559–575.
- 47 Gelman A, Hill J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press: New York, NY, USA, 2006.
- 48 Nakagawa S, Schielzeth H. A general and simple method for obtaining R<sup>2</sup> from generalized linear mixed-effects models. *Methods Ecol Evol* 2013; **4**: 133–142.
- 49 Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR *et al*. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010; **42**: 565–569.
- 50 Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* 2012; **28**: 2540–2542.
- 51 Kumar SK, Feldman MW, Rehkopf DH, Tuljapurkar S. Limitations of GCTA as a solution to the missing heritability problem. *Proc Natl Acad Sci* 2016; **113**: E61–E70.
- 52 Martin J, Tilling K, Hubbard L, Stergiakouli E, Thapar A, Smith GD *et al*. Association of genetic risk for schizophrenia with nonparticipation over time in a Population-Based Cohort Study. *Am J Epidemiol* 2016; **183**: 1149–1158.
- 53 Taylor MJ, Robinson EB, Happé F, Bolton P, Freeman D, Ronald A. A longitudinal twin study of the association between childhood autistic traits and psychotic experiences in adolescence. *Mol Autism* 2015; **6**.
- 54 Abu-Akel AM, Wood SJ, Hansen PC, Apperly IA. Perspective-taking abilities in the balance between autism tendencies and psychosis proneness. *Proc R Soc B* 2015; **282**: 20150563.
- 55 Berndt TJ. Friendship quality and social development. *Curr Dir Psychol Sci* 2002; **11**: 7–10.
- 56 Johnson MH. Development of human brain functions. *Biol Psychiatry* 2003; **54**: 1312–1316.
- 57 Robinson EB, Munir K, Munafò MR, Hughes M, McCormick MC, Koenen KC. Stability of autistic traits in the general population: further evidence for a continuum of impairment. *J Am Acad Child Adolesc Psychiatry* 2011; **50**: 376–384.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

Supplementary Information accompanies the paper on the Molecular Psychiatry website (<http://www.nature.com/mp>)