nature medicine

ARTICLES

https://doi.org/10.1038/s41591-018-0021-y

Convergence of placenta biology and genetic risk for schizophrenia

Gianluca Ursini[1,2,3], Giovanna Punzi[1,2], Qiang Chen[1], Stefano Marenco[4,5], Joshua F. Robinson[6], Annamaria Porcelli[2], Emily G. Hamilton[6], Marina Mitjans[7], Giancarlo Maddalena[2], Martin Begemann[7], Jan Seidel[7], Hidenaga Yanamori[8], Andrew E. Jaffe[1,9], Karen F. Berman[4], Michael F. Egan[10], Richard E. Straub[1], Carlo Colantuoni[11,12,13], Giuseppe Blasi[2], Ryota Hashimoto[8,14], Dan Rujescu[15], Hannelore Ehrenreich[7], Alessandro Bertolino[2] and Daniel R. Weinberger[1,3,11,12,13,16]*

Defining the environmental context in which genes enhance disease susceptibility can provide insight into the pathogenesis of complex disorders. We report that the intra-uterine environment modulates the association of schizophrenia with genomic risk (in this study, genome-wide association study–derived polygenic risk scores (PRSs)). In independent samples from the United States, Italy, and Germany, the liability of schizophrenia explained by PRS is more than five times greater in the presence of early-life complications (ELCs) compared with their absence. Patients with ELC histories have significantly higher PRS than patients without ELC histories, which is confirmed in additional samples from Germany and Japan. The gene set composed of schizophrenia loci that interact with ELCs is highly expressed in placenta, is differentially expressed in placentae from complicated in comparison with normal pregnancies, and is differentially upregulated in placentae from male compared with female offspring. Pathway analyses reveal that genes driving the PRS-ELC interaction are involved in cellular stress response; genes that do not drive such interaction implicate orthogonal biological processes (for example, synaptic function). We conclude that a subset of the most significant genetic variants associated with schizophrenia converge on a developmental trajectory sensitive to events that affect the placental response to stress, which may offer insights into sex biases and primary prevention.

Schizophrenia is a complex disabling disorder that occurs in all populations, with a lifetime morbidity risk of around 0.7–0.8%[1] and a higher incidence in males compared with females[2]. The high heritability of the disorder indicates a major role for genetic variants in its etiology[3,4]; however, non-genetic influences involving the intra-uterine environment have been repeatedly implicated in explaining at least part of the non-shared environmental contribution to the disorder[4–6].

Animal studies have shown that exposure to environmental insults in utero leads to altered response to stress postnatally, with effects on brain development and behavior that are partly mediated by gene expression changes in placenta[7–9], a key environmentally sensitive organ during development[9,10]. Studies in animals also reveal that males are more vulnerable than females to prenatal adversities[8,9].

An important role for the intra-uterine environment in the etiology of schizophrenia is consistent with the disorder's putative neurodevelopmental origins[11] and is also supported by many epidemiological studies. For example, the prevalence of schizophrenia increases in offspring of mothers who were in the second trimester during influenza epidemics; in a prospective study, maternal respiratory infection during pregnancy increased the risk for schizophrenia in the offspring threefold to sevenfold[5,12]. More generally, schizophrenia has been associated with a number of early-life complications (ELCs), that is, potentially adverse events that occur during pregnancy and labor, at delivery, and early in neonatal life[5,12,13]. Meta-analyses of this body of literature have found that ELCs increase risk by 1.5- to 2-fold[13], a greater effect than any common genetic variant. Studies of ELCs in high-risk individuals (that is, offspring of parents affected with schizophrenia) suggest an interactive role for genetic background[13], which is consistent with preliminary evidence of a relationship between ELCs, hypoxia-related genes, and risk for schizophrenia[13–15].

[1]Lieber Institute for Brain Development, Johns Hopkins Medical Campus, Baltimore, MD, USA. [2]Group of Psychiatric Neuroscience, Department of Basic Medical Science, Neuroscience and Sense Organs, Aldo Moro University, Bari, Italy. [3]Departments of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [4]Clinical and Translational Neuroscience Branch, National Institute of Mental Health, Intramural Research Program, National Institutes of Health, Bethesda, MD, USA. [5]Human Brain Collection Core, National Institute of Mental Health, Intramural Research Program, National Institutes of Health, Bethesda, MD, USA. [6]Center for Reproductive Sciences, Department of Obstetrics, Gynecology and Reproductive Sciences, University of California, San Francisco, San Francisco, CA, USA. [7]Clinical Neuroscience, Max Planck Institute of Experimental Medicine, DFG Research Center for Nanoscale Microscopy and Molecular Physiology of the Brain, Göttingen, Germany. [8]Department of Psychiatry, Osaka University Graduate School of Medicine, Osaka, Japan. [9]Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. [10]Merck Research Laboratories, Merck and Co., Inc., Whitehouse Station, NJ, USA. [11]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [12]Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [13]Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD, USA. [14]Molecular Research Center for Children's Mental Development, United Graduate School of Child Development, Osaka University, Osaka, Japan. [15]Department of Psychiatry, Psychotherapy, and Psychosomatics, Martin Luther University of Halle-Wittenberg, Halle, Germany. [16]McKusick Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA. *e-mail: drweinberger@libd.org

Table 1 | Sample composition and summary statistics

	CONT		SCZ		ELCs and diagnosis		PRS1 and diagnosis							PRS1 and ELCs						PRS1*ELCs interaction	
							Whole sample			ELCs absence		ELCs presence		Whole sample		CONT		SCZ			
	N	ELCs	*N*	ELCs	*z*	*P*	*N*	R^2	*P*	*N*	R^2	*N*	R^2	*P*	*t*	*P*	*t*	*P*	*t*	*P*	*t*	*P*
scz_lie_eur	267	0.67	234	0.66	−0.38	0.70	501	**0.06**	**2e−07**	167	0.01	334	**0.11**	**5e−09**	0.68	0.49	−0.68	0.50	**2.21**	**0.028**	**2.87**	**0.004**
scz_bari_eur	182	0.51	91	0.47	−0.51	0.61	273	**0.02**	**0.04**	138	<.01	135	**0.09**	**0.001**	1.14	0.26	−0.57	0.57	**2.88**	**0.003**	**2.17**	***0.015***
scz_munc_eur	398	0.15	521	0.24	**3.54**	**4e−04**	919	**0.04**	**5e−08**	733	**0.02**	**1e−04**	186	**0.11**	**2e−05**	0.91	0.36	−1.64	0.10	1.60	*0.055*	**2.12**	**0.017**
scz_osak_asi			172	0.60			172			69			103							**1.79**	**0.047**		
scz_gras_eur			1,020	0.32			1,020			690			330							**1.70**	**0.044**		
merged1	847	0.39	846	0.38	−0.33 **(1.97)**	0.74 **(0.05)**	1693	**0.04**	**1e−14**	1038	**0.02**	**2e−04**	655	**0.10**	**2e−15**	1.49	0.14	−1.58	0.10	**3.25**	**0.001**	**4.02**	**6e−05**
merged2			2038	0.37			2038			1281			757							**3.86**	**1e−04**		

Columns 2 and 4 report number (*N*) of healthy control subjects (CONT) and patients with schizophrenia (SCZ) in the discovery sample (*scz_lie_eur*), in the four replication samples, and in the merged samples (*merged1*: cases and controls; *merged2*: cases only); columns 3 and 5 ("ELCs") report frequency of ELCs in controls and patients with schizophrenia (notably, ELC assessment was different among samples, whereas the scoring system was the same; see Methods and Supplementary information, "Considerations about the assessment and the frequency of early life complications (ELCs)" note, for details). Columns 6 (*z*) and 7 (*P*) report the statistics of the association of ELCs with case–control status (notably, *merged1* sample values within brackets refer to the statistics after adjustment for sample effect). Columns 8–16 report the statistics (R^2: Nagelkerke R^2 variance explained and *P*) for the relationship between PRS1 and case–control status in the whole sample (columns 9 and 10), and in individuals without (columns 12 and 13) and with (columns 15 and 16) a history of ELCs. Columns 17–22 report the statistics for the relationship between PRS1 and ELC history in the whole sample (columns 17 and 18), in controls (columns 19 and 20), and in patients with schizophrenia (columns 21 and 22). Columns 23 and 24 report the statistics for the interaction between PRS1 and ELCs on case–control status. All of the statistics were generated using multiple logistic regression, adjusting for population stratification (ten PCs); Nagelkerke R^2 were calculated by comparison of a full model (covariates + PRS) with a reduced model (covariates only); significant results ($P < 0.05$) are highlighted in bold; one-sided *P* values in replication samples are indicated in italics (all other *P* values are two-sided).

Genome-wide association studies (GWASs) indicate that genetic risk for schizophrenia across heterogeneous samples is conferred by many small-effect alleles throughout the genome[16]. Studies of rare chromosomal defects showing greater penetrance also imply a myriad of susceptibility genes[17–19], indicating that the genetic architecture of the disorder is heterogeneous, consistent with polygenic mechanisms[16,20]. Although current GWASs are not designed to detect complex genetic and environmental heterogeneity[16], we hypothesized that the most significant GWAS associations might achieve their statistical status by converging on early developmental mechanisms sensitive to environmental factors that are also relatively common among patients. In this article, we analyze the role played by the intra-uterine and perinatal environment in modulating the association of schizophrenia with genomic risk, with emphasis on the placental transcriptome.

Results

Interaction of polygenic risk score and early-life complications history on case–control status. We first investigated whether the intra-uterine and perinatal environment modulates the association of schizophrenia with genomic risk. Specifically, we explored the interaction between genomic risk for schizophrenia and history of ELCs on the likelihood that a subject is a patient or a control, that is, case–control status. Genomic risk for schizophrenia was measured as the polygenic risk score (PRS)[21] based on GWAS significant alleles ($P < 5 \times 10^{-8}$, PRS1; single nucleotide polymorphisms (SNPs) in Supplementary Table 1)[16], whereas ELC history was assessed with the McNeil–Sjöström scale[22,23], defining a positive or negative ELC history based on the presence or absence of at least one potentially serious complication (that is, presence or absence of ELCs with McNeil–Sjöström scale severity level ≥ 4) as in previous reports[22,24,25]. In the discovery sample of 501 individuals from the United States (*scz_lie_eur*: 267 healthy subjects and 234 patients with schizophrenia, all Caucasian; see Table 1 and Supplementary Table 2 for sample information), the polygenic risk profile score constructed from alleles showing significant ($P < 5 \times 10^{-8}$) association with schizophrenia (PRS1) in a meta-analysis of Psychiatric Genetic Consortium (PGC) GWAS datasets, after excluding the *scz_lie_eur* sample, was positively associated with case–control status ($N = 501$, $t = 5.347$, $P = 2 \times 10^{-7}$), so that patients had greater genetic risk compared with controls; PRS1 accounted for approximately 6% of risk prediction (Nagelkerke's pseudo $R^2 = 0.060$; Fig. 1a and Supplementary Table 3). In this sample, ELCs were not significantly different among healthy subjects and patients with schizophrenia ($z = -0.38$, $P = 0.704$). However, multiple logistic regression revealed a significant interaction between PRS1 and severe ELCs on case–control status ($t = 2.87$, $P = 0.004$; Fig. 1b and Table 1); results of the multiple regression also indicated that the ELC history was associated with schizophrenia when covarying for genetic risk score ($t = 2.11$, $P = 0.03$), whereas PRS1 was not associated with schizophrenia when covarying for ELCs ($t = 1.18$, $P = 0.24$). This result suggests that the association between genetic risk and schizophrenia was affected by ELC history. Indeed, when analyzing the relationship between PRS1 and case–control status in the absence and in the presence of ELCs, we found that the liability of schizophrenia explained by the genetic risk score was highly significant in the context of ELCs ($N = 334$, Nagelkerke's pseudo $R^2 = 0.112$, $t = 5.97$, $P = 5.02 \times 10^{-9}$), but not in the absence of them ($N = 167$, Nagelkerke's pseudo $R^2 = 0.008$, $t = 1.07$, $P = 0.28$; Fig. 1b). We evaluated the same relationship in the context of each severity level of ELCs; strikingly, whereas in the absence of potentially serious ELCs (weights 0–3) cases and controls were not different in PRS1, the two groups became clearly differentiated as the severity of ELCs increased (Fig. 1c). These results were not affected by the inclusion or exclusion of the top GWAS

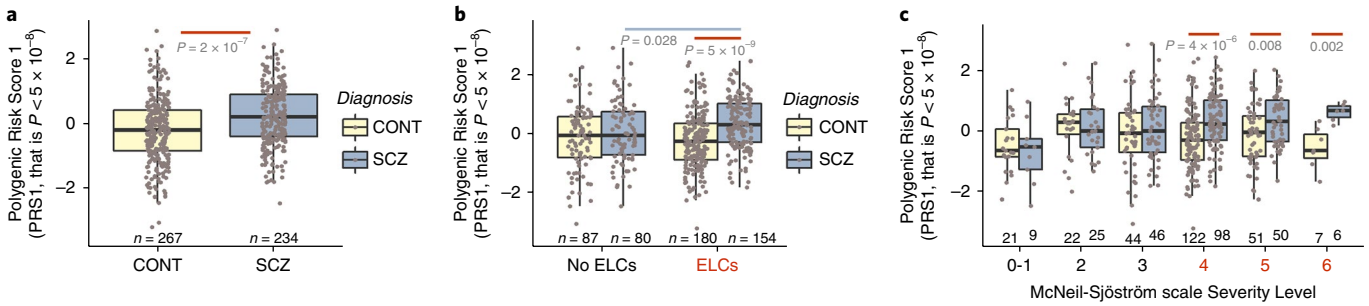

Fig. 1 | PRS1, ELC history, and schizophrenia in the *scz_lie_eur* sample ($N = 501$). a, Association between genomic risk score (PRS1), constructed from alleles showing significant association with schizophrenia (with GWAS $P < 5 \times 10^{-8}$), and case–control status. **b**, Interaction between PRS1 and ELCs on case–control status: shown are genomic risk scores (PRS1) of controls (CONT) and patients with schizophrenia (SCZ), in the absence of ELCs (left) and in the presence of ELC history (right). **c**, Relation between PRS1 and case–control status in the context of ELCs with different severity levels; ELCs with severity scores ≥ 4 (in red) are considered harmful or relevant factors in fetal stress, whereas ELCs with severity scores of 0, 1, 2, and 3 (in black) are not. All of the statistics were generated using multiple logistic regression, adjusting for population stratification (ten PCs) (see also main text and Table 1 for detailed statistics). Boxplot centers depict median; lower and upper hinges correspond to 25th and 75th percentiles; whiskers extend from hinges to the smallest and larger values no further than 1.5*IQR from the 25th and 75th percentiles. Results in the other clinical samples are shown in Supplementary Figs. 3 and 4.

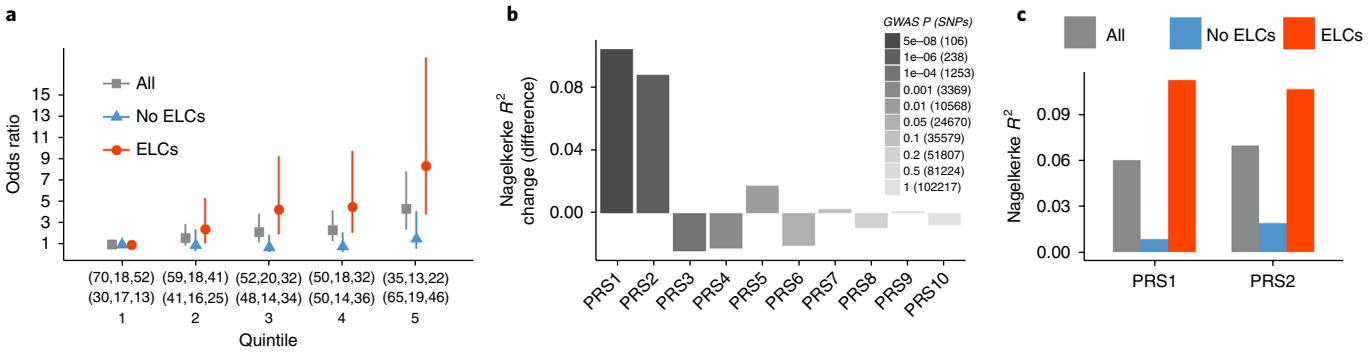

Fig. 2 | Liability of schizophrenia explained by genomic risk in the context of ELC history, in the *scz_lie_eur* sample ($N = 501$). a, OR for schizophrenia by PRS1 quintile and ELC history: the PRSs constructed from alleles showing significant association with schizophrenia (with GWAS $P < 5 \times 10^{-8}$) were converted to quintiles (1 = lowest PRS1, 5 = highest PRS1), and the ORs were estimated for affected status for each quintile with reference to the lowest risk quintile, in the whole sample (gray square) and in the absence (blue triangles) and presence (orange dots) of ELCs. Bars correspond to 95% confidence intervals. The numbers within brackets on the *x* axis refer to the individuals within each quintile, in the whole sample, in the absence of ELCs and in the presence of ELCs, in controls (first row) and in patients with schizophrenia (second row). **b**, Change of variance explained by genomic risk in the presence of ELCs compared with the absence of ELCs: proportion of variance (Nagelkerke R^2) of schizophrenia risk was calculated, by comparison of a full model (covariates + PRS) with a reduced model (covariates only), in the presence and absence of ELC history, for each of the ten different PRSs constructed from alleles showing association with schizophrenia at different threshold of significance (GWAS *P* value threshold and number of SNPs contributing to each PRS are reported in the grayscale legend). Shown is the difference (*y* axis) between Nagelkerke R^2 in the presence of ELCs and Nagelkerke R^2 in the absence of ELC history, for each PRS (*x* axis). **c**, Proportion of variance of schizophrenia by PRS1 and PRS2, and ELC history: shown is variance of case–control status (Nagelkerke R^2, *y* axis) explained by PRS1 and PRS2, in the whole sample (gray bars), in the absence of ELCs (blue bars), and in the presence of ELCs (orange bars). All of the statistics were generated using multiple logistic regression, adjusting for population stratification (ten PCs). Results in the other clinical samples are shown in Supplementary Figs. 3 and 4.

significant variant in the extended major histocompatibility complex (MHC) region (chr6: 25–34 Mb; Supplementary Table 4).

We then grouped individuals in quintiles based on their PRS1 levels, and we determined odds ratios (ORs) of being affected with schizophrenia associated with being in each PRS1 quintile, compared with the lowest PRS1 quintile. We also stratified our sample by ELC history to further represent the capacity of PRS1 to predict risk for schizophrenia in the context of ELCs. The OR increased with higher PRS1 quintiles only in the sample with ELCs, so that having the highest PRS1 quintile corresponded to an OR of 8.36 (95% confidence interval [CI]: 3.79–18.47, $P = 3.22 \times 10^{-8}$) in the presence of ELCs, and only 1.55 (95% CI: 0.59–4.07, $P = 0.37$) in the absence of ELCs (Fig. 2a), compared with having the lowest PRS1 quintile.

We further analyzed whether the interaction between genomic risk and ELCs was specific for the PRS constructed with GWAS significant alleles (PRS1) or was also found with other PRS levels (that is, PRS2–10) constructed from alleles showing association with schizophrenia at lesser statistical thresholds (that is, not GWAS significant). Interestingly, the ELC-dependent change in schizophrenia risk variance, explained by PRS, gradually decreased when considering different PRSs constructed from variants showing association with schizophrenia at the lower thresholds of significance (Fig. 2b). Specifically, only the first two scores, constructed from the alleles showing the strongest clinical association with schizophrenia (PRS1: $P < 5 \times 10^{-8}$; PRS2: $P < 1 \times 10^{-6}$), interact with ELCs on case–control status, and the variance in risk explained by them is much

higher in individuals with a history of ELCs, compared with those without (Fig. 2b,c and Supplementary Figs. 1 and 2). The other scores, PRS3–10, do not show any interaction with ELCs, and the variance explained by them is not influenced by a history of ELCs (Fig. 2b and Supplementary Figs. 1 and 2). This is consistent with the possibility that the latter scores, involving a greater number of putative susceptibility genes, capture a greater number of genetic risk variants acting in a simply cumulative way, whereas the aggregate effect of the GWAS significant SNPs (PRS1) and the almost GWAS significant SNPs (PRS2) is more conditioned on the history of ELCs. These results raise the possibility that the reason PRS1 and PRS2 loci achieve their privileged statistical significance status in this heterogeneous clinical sample is because of this interaction. From another perspective, the data show that patients with a history of ELCs have greater PRS1 than patients without ELCs ($N=234$, $t=2.21$, $P=0.028$), whereas this relationship is not seen in healthy subjects ($N=267$, $t=-0.68$, $P=0.50$). Maternal PRSs were available on a subsample of healthy mothers ($N=160$) of schizophrenic offspring and were not associated with ELCs in their offspring ($t=0.08$, $P=0.94$; Supplementary Table 5). Similarly, paternal PRSs were available only for a subsample of fathers ($N=136$) and were also not significantly associated with ELCs in the offspring ($t=1.40$, $P=0.16$; Supplementary Table 5). These results suggest that the interaction between genomic risk and ELCs is mainly driven by the fetal genome and is relatively independent of gene–environment interactions related to parental genomes per se.

We sought to replicate our findings in several additional, independent samples. We first analyzed two case–control samples: an Italian sample of 273 subjects (*scz_bari_eur*) and a German sample of 919 subjects (*scz_munc_eur*) (see Table 1 and Supplementary Table 2 for sample information). As in the discovery sample, PRS1 was positively associated with case–control status in both samples (*scz_bari_eur*: $N=273$, $R^2=0.02$, $t=2.11$, $P=0.036$; *scz_munc_eur*: $N=919$, $R^2=0.04$, $t=5.51$, $P=5\times10^{-8}$; Supplementary Table 3). ELCs were not differentially distributed between healthy subjects and patients with schizophrenia in the *scz_bari_eur* sample ($z=-0.51$, $P=0.61$), but there was a significant association of ELC history with schizophrenia in the larger *scz_munc_eur* sample ($z=3.54$, $P=0.0004$, OR = 1.85, 95% CI: 1.32–2.61). In both case–control samples, PRS again showed a significant interaction with ELCs in predicting case–control status (*scz_bari_eur*: $t=2.17$, $P_{\text{one-sided}}=0.015$; *scz_munc_eur*: $t=2.12$, $P_{\text{one-sided}}=0.017$; Table 1 and Supplementary Fig. 3a,b). When analyzing the relationship between PRS1 and case–control status in the context of ELCs, this PRS was once again associated with schizophrenia only in the presence of ELCs in the *scz_bari_eur* sample ($N=135$, $t=3.38$, $P=0.001$), and not in their absence ($N=138$, $t=-0.11$, $P=0.91$), whereas in the *scz_munc_eur* sample, the association between PRS1 and case–control status was significant both in the absence ($N=733$, $t=3.88$, $P=0.0001$) and in the presence of ELCs ($N=186$, $t=4.45$, $P=2\times10^{-5}$; Table 1 and Supplementary Fig. 3a,b). However, in both samples the variance of case–control status explained by PRS1 was much higher in individuals with a history of ELCs (*scz_bari_eur*: $R^2=0.09$; sc*z_munc_eur*: $R^2=0.11$) than in those without such history (*scz_bari_eur*: $R^2=0.0001$; *scz_munc_eur*: $R^2=0.02$; Supplementary Fig. 3c,d), and again subjects who experienced ELCs who were in the upper quintile for PRS1 had the highest OR (*scz_bari_eur*: OR = 6.67, 95% CI: 1.6–27.6, $P=0.005$; *scz_munc_eur*: OR = 14.17, 95% CI: 4.0–49.9, $P=5.03\times10^{-6}$; Supplementary Fig. 3e,f). These replication analyses also confirmed that PRS1 was positively associated with ELCs only in cases (*scz_bari_eur*: $N=91$, $t=2.88$, $P_{\text{one-sided}}=0.003$; *scz_munc_eur*: $N=521$, $t=1.60$, $P_{\text{one-sided}}=0.0547$), but not in controls (*scz_bari_eur*: $N=182$, $t=-0.57$, $P=0.57$; *scz_munc_eur*: $N=398$, $t=-1.64$, $P=0.10$; Table 1 and Supplementary Fig. 3a,b).

We further tested the relationship between genomic risk for schizophrenia and ELC history in two more samples of only patients (total $N=1,192$): another independent German sample of 1,020 patients with schizophrenia, namely, the Göttingen Research Association for Schizophrenia (GRAS) data collection (*scz_gras_eur*), and a Japanese sample of 172 patients with schizophrenia (*scz_osak_asi*) (Table 1 and Supplementary Table 2). In regard to the latter, it should be noted that the PRS derived from the European Caucasian sample of the recent GWAS study of schizophrenia has much less association with schizophrenia in Asian samples[16], as would be expected because the correlation between genotypes at nearby markers (that is, the linkage disequilibrium structure) is different across populations[26]. However, because many of the alleles comprising the score likely monitor ancient haplotypes, an association with ELCs might still be expected. As in the three other samples, we again found that PRS1 was associated with ELCs in both samples of patients with schizophrenia (*scz_gras_eur*: $N=1,020$, $t=1.70$, $P_{\text{one-sided}}=0.044$; *scz_osak_asi*: $N=172$, $t=1.79$, $P_{\text{one-sided}}=0.047$; Supplementary Fig. 4), so that patients with a history of complications had higher PRSs than patients who did not experience ELCs.

We also performed analyses in merged samples of cases and controls (*scz_lie_eur*, *scz_bari_eur*, *scz_munc_eur*) and of only cases (*scz_lie_eur*, *scz_bari_eur*, *scz_munc_eur*, *scz_osak_asi*, *scz_gras_eur*), after normalization of PRSs in each population. In these merged samples, we confirmed the interaction of PRS1 and ELCs on case–control status ($N=1,693$, $t=4.02$, $P=6.18\times10^{-5}$; Supplementary Fig. 5) and the relationship between PRS1 and ELCs in patients with schizophrenia ($N=2,038$, $t=3.86$, $P=0.0001$; Supplementary Figs. 5 and 6). Also, in the merged samples, only PRS1 and PRS2 interact with ELCs on case–control status, and only PRS1 and PRS2 are positively associated with ELCs in patients with schizophrenia (Supplementary Figs. 5 and 6). Again, the positive association between genomic risk and ELCs was not present in controls, where the trend was actually negative ($P=0.10$; Table 1), which is compatible with a pattern of a gene–environment interaction. Sensitivity analyses with sex, age, paternal and maternal ages, maternal stress, socioeconomic status, and substance use as covariates and related interaction terms (in addition to genetic principal components) confirmed the same results (Supplementary Tables 6–8). These consistent results in five independent samples support the hypothesis that these top PRSs are relevant for the etiopathogenesis of schizophrenia, particularly in the context of ELCs, whereas other PRSs (that is, PRS3–10) may capture polygenic mechanisms of schizophrenia not directly related to ELCs.

Expression of schizophrenia risk-associated genes in placenta. Even though several recent studies show preferential regulation of many schizophrenia risk genes in fetal brain[27–29], the relationship between the PRSs and ELCs that we found in five independent clinical samples from diverse ancestries points to the intra-uterine context as a likely place where some risk genes for schizophrenia and environmental adversity intersect, with implications not limited to the brain. Because PRS1 and PRS2 risk SNPs are associated with expression of nearby genes across many different tissues (Supplementary information, see "Screening of PRS1 and PRS2 SNPs for eQTLs across different tissues" note), we tested whether the genes mapping to the loci showing the strongest association with schizophrenia and interacting with ELCs (that is, PRS1 and PRS2 genes; Fig. 2b and Supplementary Table 9) were more highly expressed in the placenta, compared with randomly selected genes contributing to PRSs constructed from alleles showing association with schizophrenia at lesser thresholds of significance ($P>1\times10^{-6}$), which do not show an interaction with ELCs (that is, PRS3–10). We analyzed RNA sequencing data from placental tissue, generated in the Epigenome Roadmap Project (GSE16368), and found relatively greater expression of the PRS1 and PRS2 genes ($N=1,643$ genes), compared with same size set of genes randomly selected from PRS3–10 genes ($N=18,029$ genes), in multiple placental tissue

Table 2 | Differential expression of schizophrenia risk genes in placentae from complicated pregnancies

Dataset	Condition	Tissue	*N*	*P* value of gene set test	χ^2 test	
					P value	χ^2
GSE24129	Pre-eclampsia	Chorionic villi	16	**3.5e−04**	**0.002**	**7.93**
GSE24129	IUGR	Chorionic villi	16	**0.019**	**0.0007**	**10.21**
GSE35574	Pre-eclampsia	Chorionic tissue	59	**0.04**	0.062	2.36
GSE35574	IUGR	Chorionic tissue	75	**0.007**	**0.03**	**3.56**
GSE10588	Pre-eclampsia	Chorionic tissue	43	**0.003**	**0.002**	**8.46**
GSE25906	Pre-eclampsia	Chorionic tissue	60	**0.02**	**0.03**	**3.40**
GSE12216	Pre-eclampsia	Chorionic tissue	16	**0.01**	**0.01**	**4.82**
GSE40182	Pre-eclampsia	CTB	20	**0.009**	**0.04**	**3.04**
GSE12767	Pre-eclampsia	First trimester chorionic villi	12	**0.003**	**0.005**	**6.78**
GSE25861	Pre-eclampsia/IUGR	Microvascular endothelium	8	**0.006**	**0.04**	**3.002**
GSE65271	CTB invasiveness	CTB	7	**0.005**	**0.002**	**8.30**
GSE28619	Hepatitis (liver)	Liver	22	0.136	0.10	1.70
GSE41804	Hepatitis (liver)	Liver	40	0.285	0.20 (opposite direction)	0.69
GSE27411	HP gastritis (corpus)	Stomach corpus	9	0.377	0.45	0.01
GSE27411	HP gastritis (antrum)	Stomach antrum	9	0.453	0.34	0.17
GSE42955	Dilatative cardiomyopathy	Heart	17	0.172	0.40	0.07
GSE3586	Dilatative cardiomyopathy	Heart	28	0.283	0.10 (opposite direction)	1.63
GSE4172	Dilatative cardiomyopathy	Heart	12	0.246	0.42	0.04
GSE4483	Hypoxia	Second trimester astrocytes	6	0.470	0.18	0.85
GSE26420	MIBP1 overexpression	HEK293 cells	6	0.263	0.002 (opposite direction)	8.01
GSE64699	IUGR	Adipocytes from UC-MSC lines	28	0.109	0.37	0.12

Genes mapping to loci showing the strongest association with schizophrenia (GWAS $P < 5 \times 10^{-8}$e−08: PRS1; $P < 1 \times 10^{-6}$: PRS2) and interacting with ELCs were tested for enrichment among the differentially expressed genes in pre-eclamptic and IUGR placental samples compared with controls, and in non-invasive cytotrophoblasts (CTBs), in nine independent datasets (11 differential expression analyses: rows 1–11), and in datasets from liver, stomach (HP: *Helicobacter pylori*), heart, and cells of embryonic origins (last 10 rows). Sample sizes are reported in the fourth column. The table shows the results of the gene set test (Wilcoxon test) analysis using the moderated *t*-statistics from each differential expression analysis, and the χ^2 test results from the gene set enrichment analyses (see also Supplementary information, "Differential expression of schizophrenia risk genes in placenta" and "Sensitivity analyses for placental enrichment" notes). Significant results ($P < 0.05$) with consistent directionality are highlighted in bold.

compartments: amnion ($N=4$ samples, $P=1 \times 10^{-4}$), basal plate ($N=4$, $P=1\times10^{-4}$), chorion ($N=4$, $P=3\times10^{-4}$), villi ($N=4$, $P=1\times10^{-5}$), trophoblast ($N=4$, $P=1 \times 10^{-5}$; second trimester: $N=2$, $P=3 \times 10^{-5}$; third trimester: $N=2$, $P=1.6 \times 10^{-6}$; Supplementary Table 10). These results indicate that, as predicted, genes mapping to GWAS significant loci that interact with ELCs are more abundantly expressed in placenta than are genes in the other GWAS loci, which do not interact with ELCs.

Differential expression of schizophrenia risk-associated genes in placentae from complicated pregnancies. To elaborate on a specific role for the placenta in mediating the interaction between schizophrenia risk genes and ELCs, we next analyzed whether the PRS1 and PRS2 genes were differentially expressed in placentae from pregnancies complicated with pre-eclampsia and/or intrauterine growth restriction (IUGR) compared with normal placental controls, and in contrast with the PRS3–10 genes. The ELCs interacting with genomic risk for schizophrenia represent heterogeneous conditions, spanning pregnancy, labor, delivery, and early neonatal life; however, ischemic disease processes, with impaired trophoblast invasion and deficient remodeling of the maternal spiral arteries, as well as an altered inflammatory response, may represent common denominators in the mechanisms underlying many ELCs[30–32] (also including perinatal complications that are often the result of pathological processes starting during pregnancy[30,33]). We therefore analyzed gene expression data from placentae with pre-eclampsia and IUGR, because they represent two paradigmatic placental diseases, characterized by ischemic processes, with impaired migration of extravillous trophoblasts and impaired spiral artery remodeling[32,34,35], and often associated with an altered inflammatory response of the placenta[31,36,37]. Pre-eclampsia and IUGR are multifactorial syndromes and indeed are frequently linked with many other ELCs detected in our samples, including diabetes, obesity, alcohol use, vaginal bleeding, maternal smoking, preterm birth and other adverse birth outcomes, and perinatal morbidity[38–41]. They are themselves classic severe ELCs (severity level ≥ 4 in the McNeil–Sjöström scale) that have been linked with increased risk for schizophrenia, and also where the primary affected cells have been isolated and studied ex vivo[13,42,43]. In analyzing eight publicly available datasets, we consistently detected enrichment of the PRS1 and PRS2 genes (Table 2 and Supplementary Table 9) among the genes differentially expressed in the fetal portion of placentae from pregnancies complicated with pre-eclampsia and IUGR, specifically in pre-eclamptic (GSE24129: $P=3.5 \times 10^{-4}$; GSE35574: 0.04; GSE10588: 0.003; GSE25906: 0.02) and IUGR (GSE24129: $P=0.019$; GSE35574: 0.007; GSE12216: 0.01) chorionic tissue from term placentae, in pre-eclamptic cytotrophoblast (GSE40182: $P=0.009$) and first trimester chorionic villi (GSE12767: $P=0.003$), and in microvascular endothelium from IUGR/pre-eclamptic pregnancies (GSE25861: $P=0.006$). We observed that PRS1 and PRS2 genes tend to be upregulated (positive *t*-statistics) in multiple placental samples from pre-eclampsia and IUGR, compared with placental controls (Supplementary information, see "Differential expression of schizophrenia risk genes in placenta" note). Because the PRS1 and PRS2 genes were among the highly expressed placental genes, we then performed a

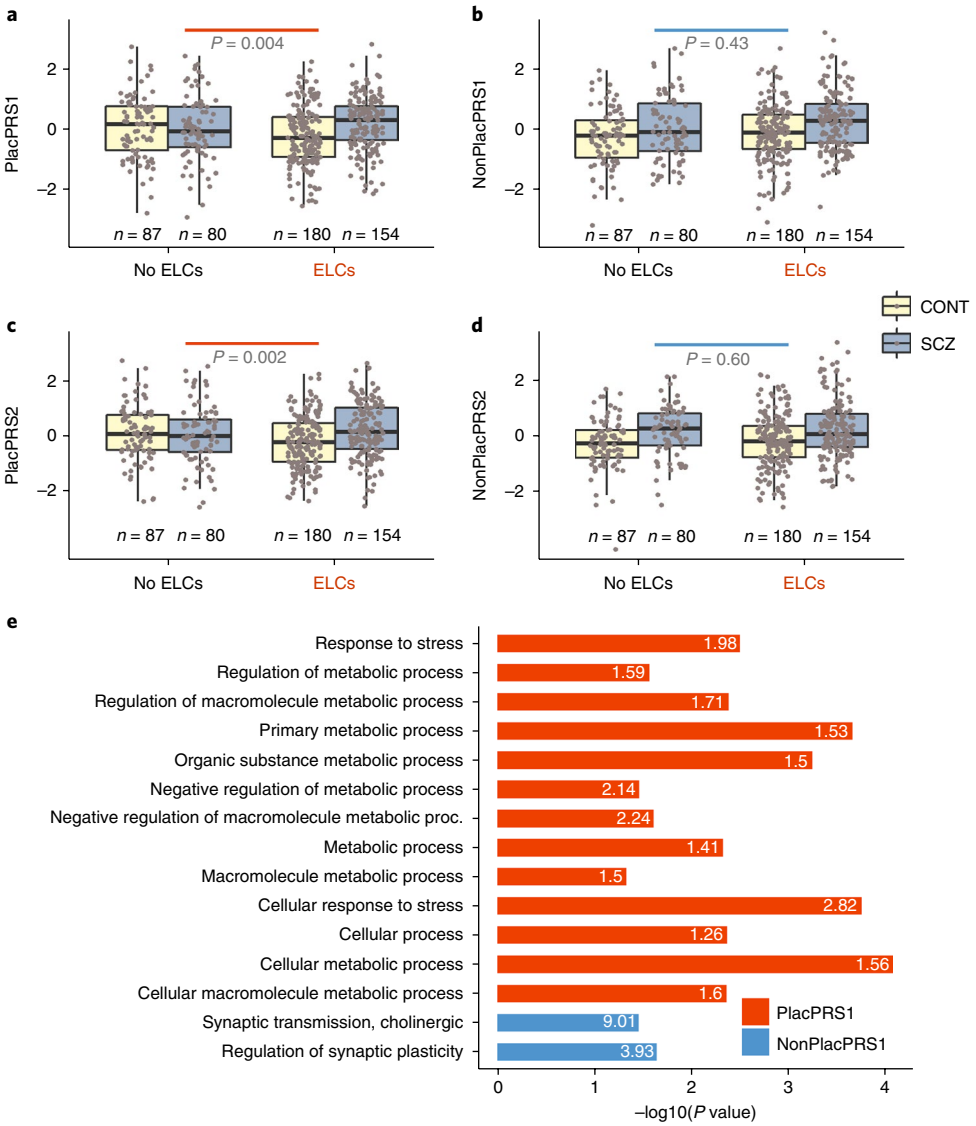

Fig. 3 | Placental and non-placental genomic risk for schizophrenia. a–d, Using GWAS SNPs marking loci containing genes highly and differentially expressed in pre-eclamptic/IUGR placental samples, we created new PRSs (PlacPRSs) and compared their interaction with ELCs with PRSs derived from the SNPs marking the remaining GWAS significant loci (NonPlacPRSs). The figure shows the interaction between PRSs and ELCs on case–control status in the US discovery sample (*scz_lie_eur*: $N = 501$): shown are PlacPRS1 (**a**), NonPlacPRS1 (**b**), PlacPRS2 (**c**), and NonPlacPRS2 (**d**) of controls (CONT) and patients with schizophrenia (SCZ), in the absence of ELCs (left side of each boxplot) and in the presence of ELC history (right). Results in the other clinical samples are shown in Supplementary Figs. 7–9. All the statistics were generated using multiple logistic regression, adjusting for population stratification (ten PCs); the *P* values refer to the significance of the interaction of ELCs with PlacPRSs (orange lines, **a**,**c**) and with NonPlacPRSs (blue lines, **b**,**d**); boxplot centers depict median; lower and upper hinges correspond to 25th and 75th percentiles; whiskers extend from hinges to smallest and larger values no further than 1.5*IQR from the 25th and 75th percentiles. **e**, Biological processes Gene Ontology terms enriched for PlacPRS1 genes ($N = 130$, orange bars) and NonPlacPRS1 genes ($N = 707$, blue bars) (see also Supplementary Tables 9, 15–18 and Supplementary Figs. 10–20): bars depict negative logarithm of the *P* values, and white numbers within bars correspond to fold enrichment.

sensitivity analysis controlling for average gene expression level, and the results were consistent (Supplementary Table 11).

We considered the possibility that differential expression of PRS1 and PRS2 genes in complicated placentae might be a nonspecific response to pathology or stress in adult or fetal tissue. We therefore performed the same differential expression analyses on datasets from tissues with diseases likely unrelated to schizophrenia, such as hepatitis (GSE28619, GSE41804), *Helicobacter pylori* (HP) gastritis infection (GSE27411), and dilatative cardiomyopathy (GSE42955, GSE4172), as well as in datasets from embryonal cells (GSE4483, GSE26420, GSE64699). The PRS1 and PRS2 genes were not enriched among the genes differentially expressed in the pathological compared with normal condition in any of these datasets, from adult tissues and embryonic cells (Table 2; Supplementary Table 9; and Supplementary information, see "Sensitivity analyses for placental enrichment" note).

Taken together, these results converge on the conclusion that schizophrenia GWAS significant risk-associated genes that interact with ELCs are highly expressed in the placenta during early life and dynamically modulated in the placenta during biological stress, as reflected in their differential expression in placentae from complicated pregnancies, and that these associations are relatively placental specific.

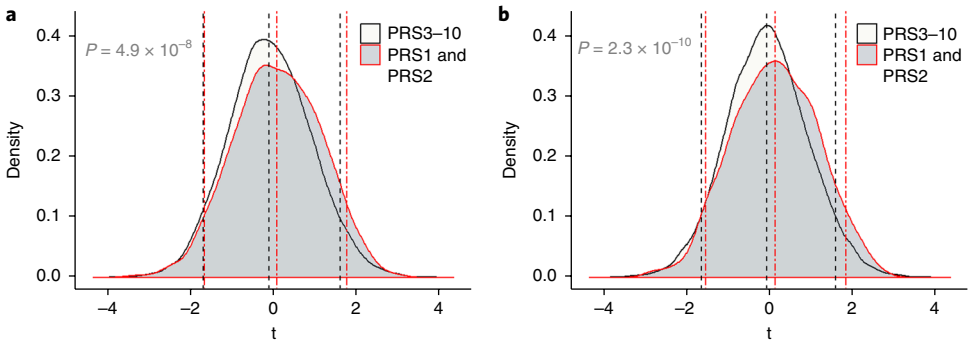

Fig. 4 | Upregulation of schizophrenia risk genes in male compared with female placentae. a,b, PRS1 and PRS2 genes were tested for enrichment among the differentially expressed genes in placentae from male compared with female offspring in two placental datasets (**a**: GSE35574: $N = 40$, 17 females and 23 males; **b**: GSE25906: $N = 37$, 16 females and 21 males). Shown are the density plots of the *t*-statistics, from the differential expression analysis, of PRS1 and PRS2 genes (dark gray area), and of all of the other genes in PRSs (PRS3–10, ivory area) constructed from lesser statistical thresholds (negative *t*-statistics = more expressed in females; positive *t*-statistics = more expressed in males; $P_{\text{one-sided}}$ from the Wilcoxon 'geneSetTest' statistics at the top of the graphic). Dotted lines depict 95% confidence intervals and median of the moderated *t*-statistics from the differential expression analysis (multiple regression) of PRS1 and PRS2 genes (red double-dotted lines) and of PRS3–10 genes (black dotted lines).

Genes highly and differentially expressed in placenta drive the interaction between polygenic risk scores and early-life complications on schizophrenia risk. The enrichment of expression in the placenta of genes in schizophrenia GWAS significant loci provides circumstantial evidence that the interaction of these loci with ELCs on risk for schizophrenia arises at least in part because of primary effects in the placenta. To achieve a more direct test of this possibility, we created new PRSs based on the GWAS SNPs marking loci-containing genes highly expressed in normal placentae and dynamically modulated in placentae from complicated pregnancies (Table 2 and Supplementary Table 9), and compared their interaction with ELCs with PRSs derived from the SNPs marking the remaining GWAS significant loci, first in our discovery sample (*scz_lie_eur*; Fig. 3a–d). The PRSs from the former group significantly interact with ELCs in increasing risk for schizophrenia (PlacPRS1 (PRS1 "placental" subset based on 56 SNPs): $t = 2.86$, $P = 0.004$; PlacPRS2 (PRS2 "placental" subset based on 112 SNPs): $t = 3.10$, $P = 0.002$; Fig. 3a,c), whereas those from the latter do not (NonPlacPRS1 (PRS1 "non-placental" subset based on 49 SNPs): $t = 0.78$, $P = 0.43$; NonPlacPRS2 (PRS2 "non-placental" subset based on 125 SNPs): $t = -0.53$, $P = 0.60$; Fig. 3b,d). Similar results were found in both other case–control samples (Supplementary Figs. 7 and 8). Analyses on the merged samples of cases and controls (*scz_lie_eur*, *scz_bari_eur*, *scz_munc_eur*: $N = 1{,}693$) confirm these results; PlacPRS1 ($t = 3.19$, $P = 0.0014$) and PlacPRS2 ($t = 3.28$, $P = 0.0011$) significantly interact with ELCs on case–control status, whereas NonPlacPRS1 and NonPlacPRS2 do not (Supplementary Fig. 9).

To verify the specificity of these interactions to placenta gene expression, we calculated PRSs based on the genes highly expressed in various adult and fetal tissues/embryonic cells, and differentially expressed in these organs during pathological/stress compared with the normal condition, employing the same procedure that we used for the computation of PlacPRSs and NonPlacPRSs (Supplementary information, see "Sensitivity analyses for placental enrichment" note). We also calculated brain PRSs, based on SNPs in PRS1 and PRS2 loci associated with methylation quantitative trait loci in adult brain[29] and with chromatin interaction in fetal brain[27]. In all of these sensitivity analyses, the PRSs comprising SNPs marking loci having genes highly expressed in these diverse adult and fetal tissues, and dynamically regulated in adult and fetal brain, as in the pathology of heart, liver, and stomach, and in pathological cells of embryonic origins, do not significantly interact with ELCs on risk for schizophrenia (all $P > 0.16$ after false discovery rate correction; results are in Supplementary Tables 12–14), whereas only the SNPs mapping to the loci highly expressed and differentially expressed in placenta do.

Biological insights about placental-enriched genes associated with ELCs. Both PlacPRS1 and PlacPRS2 genes are significantly enriched for many pathways related to metabolic and cellular stress and hypoxia, particularly to "unfolded protein response", "mitochondrial dysfunction", and "HIF1α signaling" (Supplementary Figs. 10 and 11; Supplementary Table 15), whereas not a single significant pathway enrichment could be obtained from the remaining PRS1 and PRS2 genes (NonPlacPRS1 and NonPlacPRS2), as well as from the whole PRS1 and PRS2 gene sets, consistent with the analogously negative results of the original analysis of the GWAS significant loci[16] (Supplementary information, see "Pathway, functional and co-expression analyses" note). Notably, the pathways (Supplementary Figs. 10 and 11; Supplementary Table 15), biological functions and processes (Fig. 3e; Supplementary Figs. 12–14; Supplementary Tables 16, 17), and cellular compartments (Supplementary Figs. 15 and 16; Supplementary Table 17) implicated in PlacPRS genes are virtually orthogonal to those highlighted in other analyses of schizophrenia loci, such as synaptic function, calcium signaling, fragile X–associated proteins, and chromatin remodeling[16]. Interestingly, and in contrast, genes in the NonPlacPRSs do tend to implicate some of these brain-relevant functions. These results suggest that the loci containing the schizophrenia-associated genes dynamically modulated and most enriched in the placenta contribute to schizophrenia risk at least in part by influencing the fetal/placental response to stress (Supplementary Figs. 17–19), as exemplified by the cellular stress response factor HSF1[44,45] being the main transcriptional regulator of genes in PlacPRS2 (Supplementary Fig. 18 and Supplementary Table 18). Moreover, co-expression analyses reveal that the PlacPRS1 and PlacPRS2 genes are significantly co-expressed with immune response genes, in contrast with NonPlacPRS1 and NonPlacPRS2 genes, as well as similarly sized gene sets of non-schizophrenia-associated genes in the same datasets (Supplementary Tables 19 and 20; Supplementary Fig. 20; Supplementary information, see "Pathway, functional and co-expression analyses" note).

The suggestion of a distinct and orthogonal biology for the placental component of genomic risk raises the question whether genetic prediction might be enhanced by deconstructing genomic risk into discrete subcompartments that represent alternate risk biologies. An exploratory analysis revealed that the aggregate effect on prediction accuracy of the SNPs contributing to PRS3–10

(which include PRS1 loci) is higher when separating the contribution of PRS1 (Supplementary Fig. 21; Supplementary information, see "Variance of schizophrenia liability explained by 'decomposed' PRS's" note for details). This is particularly true in the context of a history of ELCs, for each PRS. These results suggest that decomposing PRSs based on early environmental exposure and placental genetic risk may increase the prediction accuracy of genetic variation for schizophrenia.

Sex-specific analyses. The interaction between ELCs and genetic risk for schizophrenia is consistent with a body of literature pointing to the placenta as a mediator of stress effects on the developing brain[7–9]. Animal studies also have shown that the outcomes of altered placental functioning on neurodevelopment are substantially sex-specific, with males more vulnerable than females to prenatal adversity[8,9]. Epidemiological studies of schizophrenia suggest that incidence is higher in males than in females[2,46,47], despite the prevalence being similar across sexes[1,2], likely because of higher mortality in males[48]. Consistently, most evidence suggests that males have an earlier age of onset of schizophrenia[49,50], which is also a predictor of a worse prognosis[49–51], and is plausibly linked with a higher sensitivity to early developmental risk factors. These observations raise the possibility that expression of schizophrenia risk-associated genes may be different in placentae of male compared with female offspring, and this might relate to the greater incidence of developmental disorders like schizophrenia among males[2,52]. We therefore tested whether PRS1 and PRS2 genes, which interact with ELCs on case–control status, are differentially expressed in placentae from male compared with female offspring. Analyses on placental samples from the two datasets with sex information revealed that PRS1 and PRS2 genes are highly significantly enriched among the genes differentially expressed, and specifically upregulated, in placentae from male compared with female offspring (GSE35574: $N=40$, 17 females and 23 males, $P=4.9\times10^{-8}$, Fig. 4a; GSE25906: $N=37$, 16 females and 21 males, $P=2.3\times10^{-10}$; Fig. 4b). In the same datasets, the relative upregulation was also present in male preeclamptic placentae (GSE35574: $P=0.01$; GSE25906: $P=0.001$). Analogous analyses in a heart dataset (GSE4172) and a fetal lung dataset (GSE14334) with sex information did not reveal upregulation of the PRS1 and PRS2 genes in males compared with females ($P>0.40$; Supplementary information, see "Sensitivity analyses for placental enrichment" note). These data suggest a sex-biased role for the placenta in expressing genetic risk for schizophrenia.

Discussion

In this study, we show that exposure to ELCs represents an early environmental context that influences cumulative genomic risk for schizophrenia derived from GWAS significant loci. More to the point, the set of genes within these genomic loci that show interaction with intra-uterine and perinatal complications is highly expressed in placenta, and the same set of genes displays differential enrichment in this organ in abnormal invasive placental states. These results suggest that the most significant genetic variants detected by current GWASs[16] contribute to risk for schizophrenia at least partly by converging on a developmental trajectory sensitive to intra-uterine and perinatal adversity, and linked with abnormal placentation. Moreover, the strikingly relative enrichment of expression of schizophrenia risk genes in placentae from male compared with female offspring suggests that gene–environment interactions influencing placental biology may contribute to the higher incidence of schizophrenia in males compared with females[2].

Our results indicate a link between placental biology, ELCs, and schizophrenia, even as the syndrome is diagnosed during early adult life, which resonates with a broader evolutionary perspective and the developmental trajectory of schizophrenia. Schizophrenia is thought to be a condition on which the human species has a monopoly, and the delayed emergence of the clinical disorder has been posited to reflect the relatively late maturation of highly evolved neural functions, such as prefrontal cortical circuitry[11]. Interestingly, the evolutionary complexity of the primate placenta shows parallels with the phylogenetically remarkable expansion of the human brain, particularly prefrontal cortical regions that are among the most affected in schizophrenia[53,54]; both placental complexity and brain expansion come with higher rates of ELCs in humans than in other species[53,55]. Our results are consistent with the possibility that some of the common genes implicated in schizophrenia risk—through diverse biological mechanisms—regulate the physiology of the placenta, the risk of ELCs, and thereby secondarily the development of the brain, potentially interacting with other mechanisms of gene regulation acting primarily within fetal brain[27,56].

Despite many studies that have stressed the role of prenatal development and early-life events in affecting risk for brain disorders like schizophrenia[5,12,13,57], as well as autism[58,59], the mechanisms by which this may happen have been elusive. Genetic research has been successfully focused on detecting GWAS significant variants, but the difficulty of collecting environmental data has hindered defining the developmental context in which these common variants may have their critical effects. Our results underscore the importance of assessing early environmental factors such as obstetrical complications, in addition to genetic risk, to fully investigate their joint effect on susceptibility to neurodevelopmental disorders. Our results also point to the placenta as a crucial mediator of this interaction in relation to schizophrenia in particular, but likely to other neurodevelopmental disorders in general, underscoring the need for further research on placenta physiology in the context of brain development and genomic risk. Pursuing this path should advance the role of prenatal care for reducing the burden of psychiatric illness and may identify new strategies[60] for placental health as a form of primary prevention of schizophrenia, perhaps particularly in males with high genetic risk.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41591-018-0021-y.

Received: 30 August 2017; Accepted: 16 March 2018;
Published online: 28 May 2018

References

1. Saha, S., Chant, D., Welham, J. & McGrath, J. A systematic review of the prevalence of schizophrenia. *PLoS Med.* **2**, e141 (2005).
2. McGrath, J., Saha, S., Chant, D. & Welham, J. Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol. Rev.* **30**, 67–76 (2008).
3. Sullivan, P. F. The genetics of schizophrenia. *PLoS Med.* **2**, e212 (2005).
4. Lichtenstein, P. et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* **373**, 234–239 (2009).
5. Patterson, P. H. Neuroscience. *Maternal effects on schizophrenia risk. Science* **318**, 576–577 (2007).
6. Gottesman, I. I. & Wolfgram, D. L. *Schizophrenia Genesis: The Origins of Madness* (Freeman, 1991).
7. Mikaelsson, M. A., Constancia, M., Dent, C. L., Wilkinson, L. S. & Humby, T. Placental programming of anxiety in adulthood revealed by Igf2-null models. *Nat. Commun.* **4**, 2311 (2013).
8. Bronson, S. L. & Bale, T. L. Prenatal stress-induced increases in placental inflammation and offspring hyperactivity are male-specific and ameliorated by maternal antiinflammatory treatment. *Endocrinology* **155**, 2635–2646 (2014).
9. Bronson, S. L. & Bale, T. L. The placenta as a mediator of stress effects on neurodevelopmental reprogramming. *Neuropsychopharmacology* **41**, 207–218 (2016).
10. Bonnin, A. & Levitt, P. Placental source for 5-HT that tunes fetal brain development. *Neuropsychopharmacology* **37**, 299–300 (2012).
11. Weinberger, D. R. From neuropathology to neurodevelopment. *Lancet* **346**, 552–557 (1995).

12. Brown, A. S. & Derkits, E. J. Prenatal infection and schizophrenia: a review of epidemiologic and translational studies. *Am. J. Psychiatry* **167**, 261–280 (2010).
13. Cannon, M., Jones, P. B. & Murray, R. M. Obstetric complications and schizophrenia: historical and meta-analytic review. *Am. J. Psychiatry* **159**, 1080–1092 (2002).
14. Nicodemus, K. K. et al. Serious obstetric complications interact with hypoxia-regulated/vascular-expression genes to influence schizophrenia risk. *Mol. Psychiatry* **13**, 873–877 (2008).
15. Schmidt-Kastner, R., van Os, J., Esquivel, G., Steinbusch, H. W. & Rutten, B. P. An environmental analysis of genes associated with schizophrenia: hypoxia and vascular factors as interacting elements in the neurodevelopmental model. *Mol. Psychiatry* **17**, 1194–1205 (2012).
16. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
17. Stefansson, H. et al. Common variants conferring risk of schizophrenia. *Nature* **460**, 744–747 (2009).
18. Ingason, A. et al. Copy number variations of chromosome 16p13.1 region associated with schizophrenia. *Mol. Psychiatry* **16**, 17–25 (2011).
19. Purcell, S. M. et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* **506**, 185–190 (2014).
20. International Schizophrenia Consortium et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).
21. Wray, N. R., Goddard, M. E. & Visscher, P. M. Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Res.* **17**, 1520–1528 (2007).
22. McNeil, T. F. et al. Obstetric complications in histories of monozygotic twins discordant and concordant for schizophrenia. *Acta Psychiatr. Scand.* **89**, 196–204 (1994).
23. McNeil, T. F., Cantor-Graae, E. & Sjostrom, K. Obstetric complications as antecedents of schizophrenia: empirical effects of using different obstetric complication scales. *J. Psychiatr. Res.* **28**, 519–530 (1994).
24. Ursini, G. et al. Stress-related methylation of the catechol-O-methyltransferase Val 158 allele predicts human prefrontal cognition and activity. *J. Neurosci.* **31**, 6692–6698 (2011).
25. Verdoux, H., Sutter, A. L., Glatigny-Dallay, E. & Minisini, A. Obstetrical complications and the development of postpartum depressive symptoms: a prospective survey of the MATQUID cohort. *Acta Psychiatr. Scand.* **106**, 212–219 (2002).
26. Wall, J. D. & Pritchard, J. K. Haplotype blocks and linkage disequilibrium in the human genome. *Nat. Rev. Genet.* **4**, 587–597 (2003).
27. Won, H. et al. Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **538**, 523–527 (2016).
28. Jaffe, A. E. et al. Developmental regulation of human cortex transcription and its clinical relevance at single base resolution. *Nat. Neurosci.* **18**, 154–161 (2015).
29. Jaffe, A. E. et al. Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. *Nat. Neurosci.* **19**, 40–47 (2016).
30. Burton, G. J. & Jauniaux, E. The cytotrophoblastic shell and complications of pregnancy. *Placenta* **60**, 134–139 (2017).
31. Cotechini, T. & Graham, C. H. Aberrant maternal inflammation as a cause of pregnancy complications: a potential therapeutic target?. *Placenta* **36**, 960–966 (2015).
32. Maltepe, E. & Fisher, S. J. Placenta: the forgotten organ. *Annu. Rev. Cell Dev. Biol.* **31**, 523–552 (2015).
33. Maltepe, E., Bakardjiev, A. I. & Fisher, S. J. The placenta: transcriptional, epigenetic, and physiological integration during development. *J. Clin. Invest.* **120**, 1016–1025 (2010).
34. Leavey, K. et al. Unsupervised placental gene expression profiling identifies clinically relevant subclasses of human preeclampsia. *Hypertension* **68**, 137–147 (2016).
35. Redman, C. Pre-eclampsia: a complex and variable disease. *Pregnancy Hypertens.* **4**, 241–242 (2014).
36. Ferguson, K. K., Meeker, J. D., McElrath, T. F., Mukherjee, B. & Cantonwine, D. E. Repeated measures of inflammation and oxidative stress biomarkers in preeclamptic and normotensive pregnancies. *Am. J. Obstet. Gynecol.* **216**, 527. e521–527.e529 (2017).
37. Redman, C. W., Sacks, G. P. & Sargent, I. L. Preeclampsia: an excessive maternal inflammatory response to pregnancy. *Am. J. Obstet. Gynecol.* **180**, 499–506 (1999).
38. Hwang, S. S. et al. Maternal substance use disorders and infant outcomes in the first year of life among Massachusetts singletons, 2003-2010. *J. Pediatr.* **191**, 69–75 (2017).
39. Mol, B. W. J. et al. Pre-eclampsia. *Lancet* **387**, 999–1011 (2016).
40. Lisonkova, S. & Joseph, K. S. Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease. *Am. J. Obstet. Gynecol.* **209**, 544.e541–544.e512 (2013).
41. Wollmann, H. A. Intrauterine growth restriction: definition and etiology. *Horm. Res.* **49** (Suppl. 2), 1–6 (1998).
42. Dalman, C., Allebeck, P., Cullberg, J., Grunewald, C. & Koster, M. Obstetric complications and the risk of schizophrenia: a longitudinal study of a national birth cohort. *Arch. Gen. Psychiatry* **56**, 234–240 (1999).
43. Byrne, M., Agerbo, E., Bennedsen, B., Eaton, W. W. & Mortensen, P. B. Obstetric conditions and risk of first admission with schizophrenia: a Danish national register based study. *Schizophr. Res.* **97**, 51–59 (2007).
44. Xiao, X. et al. HSF1 is required for extra-embryonic development, postnatal growth and protection during inflammatory responses in mice. *EMBO J.* **18**, 5943–5952 (1999).
45. Hashimoto-Torii, K. et al. Roles of heat shock factor 1 in neuronal response to fetal environmental risks and its relevance to brain disorders. *Neuron* **82**, 560–572 (2014).
46. McGrath, J. et al. A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. *BMC Med.* **2**, 13 (2004).
47. Aleman, A., Kahn, R. S. & Selten, J. P. Sex differences in the risk of schizophrenia: evidence from meta-analysis. *Arch. Gen. Psychiatry* **60**, 565–571 (2003).
48. Hjorthoj, C., Sturup, A. E., McGrath, J. J. & Nordentoft, M. Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. *Lancet Psychiatry* **4**, 295–301 (2017).
49. Ochoa, S., Usall, J., Cobo, J., Labad, X. & Kulkarni, J. Gender differences in schizophrenia and first-episode psychosis: a comprehensive literature review. *Schizophr. Res. Treatment* **2012**, 916198 (2012).
50. Abel, K. M., Drake, R. & Goldstein, J. M. Sex differences in schizophrenia. *Int. Rev. Psychiatry* **22**, 417–428 (2010).
51. Angermeyer, M. C., Kuhn, L. & Goldstein, J. M. Gender and the course of schizophrenia: differences in treated outcomes. *Schizophr. Bull.* **16**, 293–307 (1990).
52. Bloom, B., Cohen, R. A. & Freeman, G. Summary health statistics for U.S. children: National Health Interview Survey, 2010. *Vital Health Stat.* **10**, 1–80 (2011).
53. Burton, G. J., Moffett, A. & Keverne, B. Human evolution: brain, birthweight and the immune system. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **370**, 20140061 (2015).
54. Teffer, K. & Semendeferi, K. Human prefrontal cortex: evolution, development, and pathology. *Prog. Brain Res.* **195**, 191–218 (2012).
55. Rosenberg, K. R. & Trevathan, W. R. An anthropological perspective on the evolutionary context of preeclampsia in humans. *J. Reprod. Immunol.* **76**, 91–97 (2007).
56. Sekar, A. et al. Schizophrenia risk from complex variation of complement component 4. *Nature* **530**, 177–183 (2016).
57. Niemela, S., et al. Prenatal nicotine exposure and risk of schizophrenia among offspring in a national birth cohort. *Am. J. Psychiatry* **73**, 799–806 (2016).
58. Glasson, E. J. et al. Perinatal factors and the development of autism: a population study. *Arch. Gen. Psychiatry* **61**, 618–627 (2004).
59. Stoner, R. et al. Patches of disorganization in the neocortex of children with autism. *N. Engl. J. Med.* **370**, 1209–1219 (2014).
60. Ross, R. G. et al. Perinatal phosphatidylcholine supplementation and early childhood behavior problems: evidence for CHRNA7 moderation. *Am. J. Psychiatry* **173**, 509–516 (2016).

Acknowledgements

We are grateful to the Lieber and Maltz families for their visionary support that funded the analytic work of this project. We thank all of the participants in the study and their families. We thank Joo Heon Shin for help with the gene expression analyses, and Barbara Gelao, Marina Mancini, Raffaella Romano, Rita Masellis, and Grazia Caforio for help with data acquisition. We also thank Sally Cheung and John Meyer for data management, and Susan Fisher for guidance with placental datasets and review of the manuscript. We thank the Psychiatry Genomics Consortium for providing the statistics for PRS calculation, and Thomas F. McNeil for providing the scale for ELC scoring. We also thank all of the authors of the publicly available placental datasets that have been used in this study. The collection of the ELC and genetic data for the American samples was supported by direct funding from the Intramural Research Program of the NIMH to the Clinical Brain Disorders Branch (D.R.W., PI, protocol 95-M-0150, NCT00001486, annual report number: ZIA MH002942-05), with supplemental analytic support from the Clinical and Translational Neuroscience Branch (K.F.B., PI). G.U. received partial support from P50MH094268. The collection of ELCs and genetic data for the Italian sample has been supported by the NARSAD grant no. 20337 and the "Ricerca Finalizzata" grant no. PE-2011-02347951 (both awarded to A.B.). The GRAS data collection has been supported by funding from the Max Planck Society, the Max Planck Förderstiftung, the DFG (CNMPB), EXTRABRAIN EU-FP7, and the Niedersachsen-Research Network on Neuroinfectiology (N-RENNT) (to H.E.).

Author contributions

G.U., G.P., and D.R.W. designed the study and interpreted the results. G.U., G.P., J.F.R., E.G.H., A.E.J., and C.C. carried out statistical analyses. G.U., Q.C., M.M., R.E.S., H.E., and D.R.W. organized and performed genotyping, imputation, and risk profile scoring. G.U., S.M., M.B., J.S., K.F.B., M.F.E., R.E.S., G.B., R.H., D.R., H.E., A.B., and D.R.W. organized and carried out subject recruitment and biological material collection in the discovery sample and in the replication samples, whereas G.U., G.P., S.M., A.P., G.M., M.B., H.Y., R.H., D.R., and H.E. carried out ELC assessment. J.F.R. and E.G.H. contributed to the collection of the placental tissue used in the RNA-sequencing analysis and, together with G.U., G.P., C.C., and D.R.W., interpreted the results of the gene set enrichment analyses in placental samples from complicated pregnancies compared with controls. G.U., G.P., and D.R.W. drafted the manuscript, and all authors contributed to the final version of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41591-018-0021-y.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to D.R.W.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Compliance with ethical requirements All procedures performed in the clinical samples were in accordance with the ethical standards of the institutional research committees and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in this study.

Samples and genotyping.

- *Discovery sample (USA): scz_lie_eur*. A total of 501 Caucasian unrelated adult subjects, with good-quality genetics data and ELC history information (as described later), were selected from participants in the Clinical Brain Disorders Branch Sibling Study of Schizophrenia at the National Institute of Mental Health (NIMH; Clinical Brain Disorders Branch, protocol 95-M-0150, NCT00001486, Annual Report number: ZIA MH002942-05). The sample included 234 patients who met *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (DSM-IV) criteria for schizophrenia and 267 healthy subjects (see Supplementary Table 2 for details). The Institutional Review Board of the NIMH Intramural Program approved the study, and written informed consent was obtained from the participants after complete description of the study. Exclusion and inclusion criteria have been previously reported[61]. Genotyping was done using Illumina BeadChips (510/610/660/2.5).
- *First replication sample (Italy): scz_bari_eur*. A total number of 273 Italian Caucasian unrelated adult subjects from the Region of Puglia, Italy (91 schizophrenia cases and 182 controls; see Supplementary Table 2 for details) with availability of genetics data and ELC history information, similar to the discovery sample, entered the study. The Institutional Review Board of University of Bari "Aldo Moro", Bari (Italy), approved protocols and procedures, and written informed consent was obtained from the participants after complete description of the study. Exclusion and inclusion criteria were similar to the discovery sample, as reported elsewhere[62]. Genotyping was done using Illumina BeadChips (510/610/660/2.5).
- *Second replication sample (Germany): scz_munc_eur*. A total of 919 Caucasian unrelated adult subjects entered the study (521 schizophrenia cases and 298 controls; see Supplementary Table 2 for details). Cases were ascertained from the Munich area of Germany, as described previously[17]. The controls were unrelated volunteers randomly selected from the general population of Munich. All were screened to exclude a history of psychosis/central neurological disease either personally or in a first-degree relative. All participants gave written informed consent, and the ethic committee of the Ludwig Maximilians University, Munich (Germany), approved the human subjects protocols. Genotyping was done with the Illumina 317 K array.
- *Third replication sample (Germany): scz_gras_eur*. The GRAS collection included 1,020 unrelated adult patients with schizophrenia (see Supplementary Table 2 for details), recruited across 23 German hospitals. Cases completed a structured clinical interview and were diagnosed with DSM-IV schizophrenia or schizoaffective disorder[63,64]. The study was approved by the Georg-August-University ethics committee and Ethics Committee of the University of Göttingen, Göttingen (Germany). All participants gave written informed consent. Genotyping was done with a semicustom Axiom myDesign genotyping array (Affymetrix, Santa Clara, CA, USA), based on a CEU (Caucasian residents of European ancestry from Utah, USA) marker backbone and a custom marker set.
- *Fourth replication sample (Eastern Asia): scz_osak_asi*. A total of 172 Japanese unrelated adult patients who met DSM-IV criteria for schizophrenia (see Supplementary Table 2 for details) with availability of genetics data and ELC exposure information, similar to the discovery sample, entered the study. The Institutional Review Board of University of Osaka, Osaka (Japan), approved protocols and procedures, and written informed consent was obtained from the participants after complete description of the study. Exclusion and inclusion criteria were similar to the other samples[65]. Genotyping of this sample was done using the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA).

Quality control for genotyping. Quality control was performed using PLINK (version 1.07; http://pngu.mgh.harvard.edu/purcell/plink/)[66], consistent with previous reference, for the *scz_lie_eur*, *scz_bari_eur*, *scz_munc_eur*, and *scz_osak_asi* samples[16]. Participants with missing rate higher than 2% and extreme heterozygosity values (±3 s.d.) were removed. SNPs with missing rate higher than 2% and difference in missingness between cases and controls > 0.02 were also removed. In addition, SNPs were excluded if they failed Hardy–Weinberg equilibrium ($P < 10^{-6}$ in controls or $P < 10^{-10}$ in cases) and if they had minor allele frequency less than 1%. Prephasing was done before imputation using SHAPEIT, and imputation was done using IMPUTE2 with Phase I 1000 genome as the reference panel[67,68]. The quality-control procedure in the *scz_gras_eur* sample was consistent with the other samples, as described previously[69].

Derivation of polygenic risk profile scores. Cumulative genetic risk profile scores (PRSs)[21] were calculated for each individual, as described elsewhere[20]. In brief, PRSs are a measure of genomic risk[21] calculated as the weighted sum of risk alleles for schizophrenia from the recent GWAS study[16,20]. We thus multiplied the natural log of the OR of each index SNP, from this recent schizophrenia GWAS[16], by the imputation probability for the corresponding reference allele at each variant, and summed the products over all variants, so that each subject had whole genome PRSs as originally described for this measure[20,21]. The PGC provided ORs and index SNPs for each sample. For the *scz_lie_eur* sample, ORs and 102K index SNPs were derived from a meta-analysis of PGC GWAS datasets excluding our discovery sample (PGC 2014, non-*scz_lie_eur* PGC GWAS)[16]. Consistently, for the *scz_munc_eur* and the *scz_gras_eur* samples (also included in the PGC GWAS[16]), ORs and index SNPs were derived from a meta-analysis of PGC GWAS datasets, excluding, respectively, the *scz_munc_eur* and the *scz_gras_eur* samples. For the *scz_bari_eur* and *scz_osak_asi* samples, ORs and index SNPs were derived from the PGC GWAS datasets, because these samples are not included in the PGC GWAS dataset[16]. Consistent with the standard procedure for PRS calculation[16,20], only autosomal SNPs were included in the analysis, to prevent any bias related to sex in the PRS calculation. We performed a linkage disequilibrium (LD) pruning and clumping of the SNPs, discarding variants within 500 kb of, and in $r^2 \geq 0.1$ with, another (more significant) marker, as reported elsewhere[16,20]. Ten PRSs (PRS1–10) were calculated using subsets of SNPs selected according to the GWAS *P* value thresholds of association with schizophrenia: 5e−08 (PRS1), 1e−06 (PRS2), 1e−04 (PRS3), 0.001 (PRS4), 0.01 (PRS5), 0.05 (PRS6), 0.1 (PRS7), 0.2 (PRS8), 0.5 (PRS9), and 1 (PRS10). SNPs in sets with lower *P* values are also in sets with higher *P* values (for example, SNPs in PRS1 are included in PRS2, SNPs in PRS2 are included in PRS3, and so on). A detailed list of SNPs included in PRS1 and PRS2 is provided in Supplementary Table 1. We performed all of the analyses both including and excluding the top GWAS significant SNP in the extended MHC locus (hg19 coordinates: chr6: 25–34 Mb), with similar results (Supplementary Table 4). For additional analyses (Supplementary Fig. 21; Supplementary information, see "Variance of schizophrenia liability explained by 'decomposed' PRS's" note), we also calculated PRSs from sets of SNPs with higher *P* values (PRS2–10) without including SNPs in sets with the lowest *P* values (PRS1).

Assessment of early-life complications. ELCs are here referred to as "somatic complications and conditions occurring during pregnancy, labor-delivery and the neonatal period" potentially harmful for the offspring, with special focus on the central nervous system[22,23]. These events are also referred to elsewhere as "obstetric complications"[14,22,23] and, despite their potential frequent occurrence[70], do not lead to negative outcomes in most cases. We assessed ELCs through medical records, when available, and personal interviews, described as follows:

- *scz_lie_eur*, and *scz_bari_eur*: We used mainly maternal recall based on an extensive personal interview, which has been repeatedly shown to represent a reliable method for obtaining ELC history, when used in a careful and detailed manner[71,72]. Specifically, we used a well-standardized and validated questionnaire[14], based on all the items scored with the McNeil–Sjöström scale for obstetric complications[22,23]. It covers the entire duration of pregnancy and early neonatal life, and also contains indicators of reliability and an assessment of the seriousness of each complication.
- *scz_osak_asi*: We used mainly medical records. When medical records were not exhaustive, we interviewed the mothers of the patients; the histories were again scored directly based on the McNeil–Sjöström metrics[23].
- *scz_munc_eur* and *scz_gras_eur*: We used medical records, including all the discharge letters of patients, and personal interviews. Differently from the questionnaires used in the other samples, these interviews did not contain all the items included in the McNeil–Sjöström scale[23], thus increasing the risk for incomplete information. History of ELCs from the available information was again scored using the McNeil–Sjöström scale[22,23].

In the McNeil–Sjöström scale[22,23] each ELC is assigned a severity level of 1–6, based on the degree of inferred potential harm to the offspring central nervous system. ELCs with severity weight ≥ 4 are considered potentially clearly harmful or relevant factors in fetal stress. The McNeil–Sjöström scale in the context of maternal recollection has been well validated in comparison with hospital records[22]. As in other reports[22,24,25], we defined a positive history of ELCs based on the presence of at least one serious ELC (severity level ≥ 4), and we identified the severity level of each individual as the highest severity level of all of the ELCs occurring in that individual. GWAS-derived PRSs were unknown to both the individuals who provided the information about ELCs and to the researchers who collected and evaluated them. Individuals were excluded if the information provided was incomplete or inconsistent (for example, contradictory answers to related questions), or if the presence of a complication was certain but a severity weight could not be confidently determined. The frequency of ELCs in our samples may be not representative of the general populations (Supplementary information, see "Considerations about the assessment and the frequency of early life complications (ELCs)" note). Supplementary Table 21 contains a list of the ELCs detected in each sample.

Statistical analysis of the interaction between polygenic risk scores and early-life complications on case–control status. All statistical analyses were performed in the 'R' environment[73]. To test the central hypothesis of this study (that is, the interaction between PRS1 and ELCs on case–control status),

we used multiple logistic regression, with the following model: Diagnosis ~ PRS + ELCs + PRS*ELCs + covariates. We also used multiple logistic regressions to confirm the association of each PRS with case–control status in our samples (Diagnosis ~ PRS + covariates) and to verify whether ELCs were associated with schizophrenia (Diagnosis ~ ELCs + covariates). In the presence of an interaction between PRS and ELCs, we performed post hoc analyses to evaluate the relationship between PRS and case–control status, in the presence and absence of ELC history (that is, stratifying the sample, based on ELC history), using the same model described earlier (Diagnosis ~ PRS + covariates). For all of these analyses, we report in the main text the *P* values and the *t*-statistics associated with our variable of interest (that is, PRS*ELCs, or PRS). To evaluate goodness of fit of these logistic models (Diagnosis ~ PRS + covariates) in the whole sample, in the absence and presence of ELC history, we calculated the Nagelkerke R^2, by comparison of a full model (covariates + PRS) with a reduced model (covariates only). Similarly, in the presence of an interaction between PRS and ELCs on case–control status, we performed post hoc analyses to test the relation between PRS and ELCs, separately in controls and in patients with schizophrenia (that is, after stratifying the sample for diagnosis), using multiple logistic regression (ELC history ~ PRS + covariates). Consistently with the interaction between PRS and ELCs, we found a positive relation between PRS and ELCs only in patients with schizophrenia; we further explored this relationship in the two replication samples of only patients (*scz_gras_eur*, *scz_osak_asi*). In each analysis, we used 10 ancestry-based principal components as covariates, to avoid potential confounding effects of population stratification, consistent with previous work[16]. We performed sensitivity analyses adding sex, age, maternal and paternal ages, maternal stress, history of substance use, and socioeconomic status, as covariates, and also their interaction with PRS and ELCs, as recommended to properly exclude the role of confounders[74]. We also performed sensitivity analyses, in each sample, by excluding the individuals with mothers with a history of substance use. Results were consistent and are reported in Supplementary Tables 6–8. For the analyses in the merged sample, PRS scores were normalized by subtracting the mean and dividing for the s.d. in each sample; for this analysis, we added the sample as a covariate.

Selection of PRS1 and PRS2 genes. To define genes mapping to the PRS1 and PRS2 loci for gene set analyses, we used two alternative criteria:

- PGC LD regions: We considered, as PRS1 and PRS2 genes, all of the University of California, Santa Cruz (UCSC) genes overlapping the LD regions associated with each SNP ($R^2 > 0.6$), as reported in a previous reference[16] and on the PGC website (http://www.med.unc.edu/pgc/downloads);
- Distance: We considered, as PRS1 and PRS2 genes, all of the UCSC genes mapping 500 kb ± the index SNPs of each PRS in the discovery sample (*scz_lie_eur*). We use this criterion, in addition to the "traditional" LD criterion, on the grounds that LD differs among populations, as we analyzed multiple samples. Moreover, the LD regions associated with each SNP have a huge variability: for example, 2 out of the 108 GWAS significant schizophrenia-risk SNPs have an LD region that spans only 1 bp (rs4766428, rs117074560)[16]. Further, it has been shown that GWAS SNPs are often associated with expression of genes that are not their nearest genes and are outside the associated LD regions[16,75]. Finally, the distance of 500 kb ± the index SNPs is within the range commonly used for detection of *cis*-expression quantitative trait loci (eQTL) [75] and is the same dimension used to calculate PGC loci eQTL in the original PGC report [16]. This criterion allowed us to distribute 21,028 out of 23,056 UCSC genes among the 10 PRSs.

Differences between the two list of genes (reported in Supplementary Table 9a,b) are related not only to the criterion adopted for SNP selection (distance or LD), but also to the fact that the PGC loci associated with schizophrenia at $P < 5 \times 10^{-8}$ are defined based on combining the primary GWAS and the supplementary deCODE data, whereas SNPs for PRS calculation are derived from the primary GWAS only[16]. Because only SNPs mapping to autosomal chromosomes are used for schizophrenia PRS construction[16,20], we excluded from our analysis genes that were irrelevant to our question, that is, genes mapping to mitochondrial DNA, and X- and Y-chromosome genes or other genes mapping to loci not used for PRS calculation. After exclusion of the genes on sex chromosomes and on mitochondrial DNA, and genes undetected in the expression datasets analyzed, the final number of PRS1 and PRS2 genes was 1,643 in the list based on distance (matching 325 out of the 348 genes assigned to the 108 schizophrenia GWAS significant loci[16]), and 589 in the gene list based on LD (matching 334 out of the 348 genes assigned to the 108 schizophrenia GWAS significant loci[16]). In both gene lists, PRS1 genes are a subset of PRS2 genes (therefore referred in the text as PRS1 and PRS2 genes). We performed all of the gene set analyses, with PRS1 and PRS2 genes defined with both criteria (LD and distance), and found consistent results (Supplementary Table 9). In the main text, we report results with the PRS1 and PRS2 genes defined based on the distance criterion (Table 2).

Expression of PRS1 and PRS2 genes in placenta. The proprietary placental tissue used for this analysis was collected at the University of California, San Francisco. Methods of collection were approved by the University of California, San Francisco Institutional Review Board, and informed consent was obtained from all donors. The RNA-sequencing datasets related to second trimester and term placental tissues (amnion, basal plate, chorion, villi) and isolated cells (trophoblasts) are publicly available (GSE16368) on the National Institutes of Health Roadmap Epigenomics Project website (http://www.roadmapepigenomics.org/). Expression levels of each gene were quantified by determining reads per kilobase of transcript per million values[76]. Based on our primary hypothesis and on the fact that PRS1 and PRS2 risk SNPs are associated with expression of nearby genes across many different tissues (Supplementary information, see "Screening of PRS1 and PRS2 SNPs for eQTLs across different tissues"), we tested whether the genes mapping to the loci showing the strongest association with schizophrenia and interacting with ELCs are more expressed in placenta compared with randomly selected genes contributing to the other PRSs constructed from alleles showing association with schizophrenia at lesser thresholds of significance, which do not show an interaction with ELCs. To perform this analysis, we used the function 'geneSetTest' in the R package 'limma'[77], using the gene expression data from the RNA-sequencing analysis in placenta. With this function, we calculated a *P* value from a Wilcoxon test to verify the hypothesis that the selected set of genes (PRS1 and PRS2) tends to be more highly ranked in expression compared with randomly selected sets of genes of the same size (from the PRS3–10 genes). Results are reported in Supplementary Table 10. We also performed a further analysis testing the enrichment of the genes overlapping the PRS1 and PRS2 loci using the function 'findOverlaps' in the R package 'GenomicRanges'[78] (either including or excluding genes with reads per kilobase of transcript per million expression > 0.01) and, as predictable, we obtained similar results.

Differential expression of PRS1 and PRS2 genes in placentae from complicated pregnancies. We searched for enrichment of the PRS1 and PRS2 genes among genes differentially expressed in placental samples from complicated pregnancies compared with controls. We interrogated placental datasets from the Gene Expression Omnibus public repository. Datasets were chosen for analysis if they included all of the following: a comparison between placental samples from complicated pregnancies and controls, more than one sample per group (that is, comparisons between pooled RNA samples were discarded), and expression data for at least half of the PRS genes. We found eight datasets that met these criteria, comparing control versus diseased (pre-eclampsia and IUGR) placenta cells. A dataset on normal cultured cytotrophoblasts was also chosen as cells were induced into different states of invasiveness. Because two of these datasets contain a comparison of controls versus pre-eclamptic and controls versus IUGR placentae, we were able to perform 11 differential expression analyses. In each dataset, we dropped probes that map to multiple genes and, when more than one probe per gene was present, we selected the one with the highest mean expression. We used the function 'eBayes' in the R package 'limma'[77] to attribute a moderated *t*-statistic to each gene related to differential expression (using the covariates provided by each reporting group); then we applied the 'geneSetTest' function on the moderated *t*-statistics (results are reported in Table 2) testing whether the selected set of genes (that is, those related to PRS1 and PRS2) tends to be more highly ranked in differential expression compared with randomly selected genes of the same size from the other GWAS loci (PRS3–10). We also used a χ^2 test to confirm whether PRS1 and PRS2 genes were enriched for differentially expressed genes compared with the remaining genes (Table 2).

We chose different thresholds for gene expression to exclude that low-expressed genes could affect the significance of the results. Also, because PRS1 and PRS2 genes are relatively highly expressed in placenta, we performed sensitivity analysis, adjusting for average gene expression: for this purpose, we assigned each gene a moderated *t*-statistic from the differential expression analyses, an 'in-set' value "1" to PRS1 and PRS2 genes, and an 'inset' value "0" to the remaining genes (PRS3–PRS10); we then analyzed the relationship between the *t*-statistics and the 'in-set' variable, after covarying for average gene expression (see results in Supplementary Table 11). Importantly, in addition to an enrichment analysis of PRS1 and PRS2 genes based on genes mapping 500 kb ± the index SNPs of PRS1 and PRS2, we also calculated the enrichment results for the PRS1 and PRS2 genes, defined as the UCSC genes overlapping the LD regions associated with each SNP ($R^2 > 0.6$), as reported on the PGC website (http://www.med.unc.edu/pgc/downloads) (Supplementary Table 9). We finally tested whether the PRS1 and PRS2 genes are enriched among the genes that are differentially expressed in placentae from male compared with female offspring. Among the placental datasets selected in our study, only three ("GSE25861", "GSE35574", "GSE25906") contained sex information; one of them ("GSE25861") included one female sample. Therefore, we limited this analysis to the remaining two datasets, "GSE35574" and "GSE25906" (Fig. 4). Also, in this case, we performed sensitivity analyses adjusting for average gene expression (Supplementary Table 11).

To confirm the specificity of our findings in the placentae from complicated pregnancies, we performed similar analyses in available datasets from normal/affected organs and in embryonic cells under distress (results for all of these sensitivity analyses are reported in Table 2, Supplementary Tables 9 and 11, and Supplementary information, see "Sensitivity analyses for placental enrichment" note).

Placental-enriched risk profile scoring. We calculated PRSs based on the GWAS SNPs marking loci-containing genes highly expressed in placenta and differentially expressed in placentae from complicated pregnancies, and compared their interaction with ELCs to PRSs derived from the SNPs marking the remaining

GWAS significant loci. For this purpose, we selected the loci-containing genes differentially expressed ($P < 0.05$) in at least four of the eight datasets analyzed and the loci containing genes with expression in the upper decile both in trophoblast and in villi. This gave us a list of 56 SNPs for PRS1 and 112 SNPs for PRS2, as marking loci with genes dynamically modulated and enriched in placenta; we then calculated PRSs based on these SNPs ("Placental" risk profile scores, PlacPRS1 and PlacPRS2) and on the remaining SNPs in these PRS1 and PRS2 loci that did not show high or differential expression in placental tissues (NonPlacPRS1 and NonPlacPRS2), as described earlier (see earlier "Derivation of polygenic risk profile scores" section). In an analogous way and to address the organ specificity of the PlacPRS interaction with ELCs, we calculated "TissuePRSs" and "NonTissuePRSs" based on SNPs marking PRS1 and PRS2 loci-containing genes highly and differentially expressed in adult and fetal tissues, or associated with methylation quantitative trait loci in postmortem human dorsolateral prefrontal cortex[29], or with chromatin interactions in fetal brain[27]. We then analyzed the interaction of these TissuePRSs and NonTissuePRSs with ELCs on case–control status. Results of these sensitivity analyses are reported in Supplementary Tables 12–14 and in the Supplementary information (see "Sensitivity analyses for placental enrichment" note).

Pathway and functional analyses. We investigated whether the placenta-enriched genes mapping to the loci of PlacPRS1 and PlacPRS2 are enriched for particular biological features, compared with the remaining genes mapping to the PRS1 and PRS2 loci (NonPlacPRS1 and NonPlacPRS2). Data were analyzed through QIAGEN's Ingenuity Pathway Analysis (QIAGEN, Redwood City, CA, USA; http://www.qiagen.com/ingenuity). The software determines the pathways and biological processes enriched for a given set of genes by considering the number of focus genes that participate in each process and the total number of genes that are known to be associated with that process in the selected reference set. We performed the Ingenuity Pathway Analysis "core" analysis, using default parameters (reference set: Ingenuity Knowledge Base; relationships: direct and indirect; node types: all; data sources: all; confidence: experimentally observed and high; species: human, mouse, and rat; tissues and cell lines: all; mutations: all). We chose a *P* value calculation based on the Benjamini–Hochberg method of accounting for multiple testing in the canonical pathway and functional analyses. In addition, we used the Panther tool[79] on the Gene Ontology database (http://geneontology.org) for statistical overrepresentation testing, to further explore whether PlacPRS1 and PlacPRS2 genes and NonPlacPRS1 and NonPlacPRS2 genes show differences in enrichment among Gene Ontology terms associated with molecular functions, biological processes, and cellular components. In this analysis, the *P* value calculation is based by default on the Bonferroni method of accounting for multiple testing.

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data and code availability. To protect the privacy of the study participants, the genetic and ELC data generated and analyzed during this study are available from the corresponding author on reasonable request, together with the codes used for the analyses. The placental datasets and the other gene expression datasets analyzed in this study are available on the Gene Expression Omnibus repository (https://www.ncbi.nlm.nih.gov/geo/) under the accession codes provided in this article.

References

61. Rasetti, R. et al. Altered cortical network dynamics: a potential intermediate phenotype for schizophrenia and association with ZNF804A. *Arch. Gen. Psychiatry* **68**, 1207–1217 (2011).
62. Ursini, G. et al. BDNF rs6265 methylation and genotype interact on risk for schizophrenia. *Epigenetics* **11**, 1–13 (2016).
63. Begemann, M. et al. Modification of cognitive performance in schizophrenia by complexin 2 gene polymorphisms. *Arch. Gen. Psychiatry* **67**, 879–888 (2010).
64. Ribbe, K. et al. The cross-sectional GRAS sample: a comprehensive phenotypical data collection of schizophrenic patients. *BMC Psychiatry* **10**, 91 (2010).
65. Ohi, K. et al. Glutamate networks implicate cognitive impairments in schizophrenia: genome-wide association studies of 52 cognitive phenotypes. *Schizophr. Bull.* **41**, 909–918 (2015).
66. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Human Genet.* **81**, 559–575 (2007).
67. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* **1**, 457–470 (2011).
68. Delaneau, O., Marchini, J. & Zagury, J. F. A linear complexity phasing method for thousands of genomes. *Nat. Methods* **9**, 179–181 (2012).
69. Stepniak, B. et al. Accumulated environmental risk determining age at schizophrenia onset: a deep phenotyping-based study. *Lancet Psychiatry* **1**, 444–453 (2014).
70. Danilack, V. A., Nunes, A. P. & Phipps, M. G. Unexpected complications of low-risk pregnancies in the United States. *Am. J. Obstet. Gynecol.* **212**, 809. e1–809.e6 (2015).
71. Hewson, D. & Bennett, A. Childbirth research data: medical records or women's reports? *Am. J. Epidemiol.* **125**, 484–491 (1987).
72. O'Callaghan, E., Larkin, C. & Waddington, J. L. Obstetric complications in schizophrenia and the validity of maternal recall. *Psychol. Med.* **20**, 89–94 (1990).
73. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2008).
74. Keller, M. C. Gene x environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biol. Psychiatry* **75**, 18–24 (2014).
75. Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
76. Roadmap Epigenomics Consortium et al. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
77. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
78. Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).
79. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* **8**, 1551–1566 (2013).

nature research

Corresponding author(s): Daniel R. Weinberger

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Quality control for genotyping was performed using PLINK (version 1.07; http://pngu.mgh.harvard.edu/purcell/plink/, PMID: 17701901). Pre-phasing was done before imputation using SHAPEIT, and imputation was done using IMPUTE2 with Phase I 1000 genome as reference panel (PMID: 22384356, 22138821).
Data analysis	All statistical analyses were performed in the 'R' environment. We used the following 'R' packages: 'psych' Version = 1.5.8 (to calculate biserial correlation coefficients), 'limma' (PMID: 25605792), 'GenomicRanges' (PMID: 23950696), 'stats', 'BioMart' (PMID: 19617889, to identify genes mapping to specific loci). We performed pathway and functional analyses using the QIAGEN's Ingenuity® Pathway software (www.qiagen.com/ingenuity) and the Panther tool (release 20160321) on the Gene Ontology (GO) database (http://geneontology.org).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

In order to protect the privacy of the study participants, the genetic and ELCs data generated and analyzed during the current study are available from the corresponding author on reasonable request, together with the codes used for the analyses. The placental datasets and the other gene expression datasets analyzed in this study are available on the Gene Expression Omnibus repository (https://www.ncbi.nlm.nih.gov/geo/) under the accession codes provided in the manuscript.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences

Study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size in the discovery sample was determined based on all the individuals with available good quality obstetrical histories data and genetic data, that were participants in the Clinical Brain Disorders Branch (CBDB) Sibling Study of Schizophrenia at the National Institute of Mental Health (NIMH; Clinical Brain Disorders Branch, protocol 95-M-0150, NCT00001486, Annual Report number: ZIA MH002942-03 CTNB). We performed replication analyses in 4 more independent samples, to support the reliability of our findings. Sample sizes in the replication samples were determined based on all the individuals with available good quality obstetrical histories data (i.e. assessable with the McNeil-Sjostrom scale) and genetic data in the Psychiatric Genetic Consortium, and in our collaboration network. To further support the reliability of our findings, we also performed the analyses in the merged sample.
Data exclusions	We excluded from our analysis individuals for which it was not possible to establish the presence or absence of at least one serious early life complication (ELC), based on the available data. For genotyping quality control, we excluded individuals with missing rate higher than 2% and extreme heterozygosity values (±3 SD). For calculation of polygenic risk scores, we excluded SNPs if they failed Hardy-Weinberg equilibrium test (P < 10−6 in controls or P < 10−10 in cases) and if they had minor allele frequency less than 1%. Quality control for genotyping was performed consistent with previous reference (PMID: 25056061). Since only SNPs mapping to autosomal chromosomes are used for schizophrenia PRS construction, we excluded - from the selection of genes mapping to PRS loci - the genes that were irrelevant to our question, i.e. genes mapping to mitochondrial DNA, and X and Y-chromosome genes or other genes mapping to loci not used for PRS's calculation, thus avoiding the risk of overinflating p-values (https://bioconductor.org/packages/release/bioc/vignettes/GOstats/inst/doc/GOstatsHyperG.pdf). All the study participants were unrelated.
Replication	The interaction between polygenic score for schizophrenia (PRS) and ELCs found in the discovery sample was reliably replicated in two more independent samples.The relationship between PRS and ELCs in patients with schizophrenia, detected in the discovery sample, was reliably replicated in four independent samples.
Randomization	In each analysis, we used 10 ancestry-based principal components as covariates, to avoid potential confounding effects of population stratification, consistent with previous work (PMID: 25056061). We performed sensitivity analyses adding sex and age, maternal and paternal ages, maternal stress, substance use history and socioeconomic status, as covariates, and also their interaction with PRS and ELCs, as recommended to properly exclude the role of confounders (PMID: 24135711).
Blinding	The investigators who performed the recruitment and the clinical evaluation of controls and patients with schizophrenia were blinded to their PRS's and ELCs. PRS's were unknown to both the individuals who provided the information about ELCs and to the researchers who collected and evaluated them. ELCs histories were unknown to the investigators that calculate PRS's.

Materials & experimental systems

Policy information about availability of materials

n/a	Involved in the study
☒	☐ Unique materials
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Research animals
☐	☒ Human research participants

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	No cell lines were used. We just analyzed cell line data publicly available.
Authentication	No cell lines were used. We just analyzed cell line data publicly available.
Mycoplasma contamination	No cell lines were used. We just analyzed cell line data publicly available.
Commonly misidentified lines (See ICLAC register)	No cell lines were used. We just analyzed cell line data publicly available.

Human research participants

Policy information about studies involving human research participants

Population characteristics	All relevant information on human research participants are provided in the On-line methods section and in Table S2.

Method-specific reporting

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ Magnetic resonance imaging