Supplementary material

Appendix S1. Genotyping Data

*Genotyping data*

*DNA collection and extraction*

The purity of DNA and its concentration were tested using PicoGreen DNA assay (Invitrogen Quant-iT™ PicoGreen™ dsDNA Assay Kit Cat#P7589). Concentration normalizations were carried out in accordance with the specifications for genotyping on microarrays. The sample included 992 participants aged 10 years old and who consented to providing blood samples for genotyping.

*Genotyping & Quality Control*

Participants were excluded on the basis of <95% of data (12 participants), genetic duplicates (7 participants), sex mismatches (4 participants). Génome Québec genotyped a total of 978 participants DNA samples using a customized chip based on the Illumina Infinium PsychArray-24v1.1 Beadchip. A SNP was excluded if it had a minor allele frequency (MAF) less than 1%, or in deviation from Hardy-Weinberg equilibrium (HWE) (*p* < 1×10-6), or with ambiguous strand information.

A population stratification procedure aimed to identify genetic outliers. Thus, ten multidimensional scaling (MDS) were generated using the pairwise genetic identity-by-state matrix; SNPs were filtered according to the removal of variants with HWE test p<.001 or a MAF <5%. Pruning of the remaining SNPs occurred using windows of 200 variants, a step size of 100, and a linkage disequilibrium threshold of *r*2<.2. A total of 134 genetic outliers were identified through population stratification and these participants were excluded. Moreover, 5 individuals were eliminated as outliers because of high autosomal heterozygosity.

*Imputation*

Imputation against a reference panel of haplotypes (1000 Genomes Phase 3) was conducted to increase the number of SNPs included within the smaller genotyped sample. SHAPEIT2 was used to estimate haplotypes (Delaneau, Marchini, & Zagury, 2012). Using reference data, SNPs were imputed using IMPUTE2 in 5 mega-basepair chunks with 500 kilobase buffers. SNP variants with HWE (*p* < 1×10-6), MAF <1% and INFO metric <0.8 were excluded. In total, 8 407 807 SNPs and 816 participants in the QLSCD were included, after quality control and imputation.

Table S1. Bivariate correlations between study’s main variables and sex and socioeconomic factors (confounders) based on weighted valuesa

|  |  |  |
| --- | --- | --- |
|  | sex | SES |
| PGS resid | -.040 (*p*=.283) | -.085 (*p*=.022) |
| Cum. Maltreatment (prospective) | -.094 (*p*=.025) | -.297 (*p*<.001) |
| Cum. Maltreatment (retrospective) | .068 (*p*=.098) | -.190 (*p*<.001) |
| any lifetime deprivation (prospective) | -.073 (*p*=.066) | -.227 (*p*<.001) |
| any lifetime deprivation (retrospective) | .025 (*p*=.540) | -.169 (*p*<.001) |
| any lifetime threat (prospective) | .003 (*p*=.929) | -.261 (*p*<.001) |
| any lifetime threat (retrospective) | .133 (*p*<.001) | -.127 (*p*<.001) |
| depressive symptoms (20-23 years) | .183 (*p*<.001) | -.036 (*p*=.357) |

*Note*. SES refers to the mean socio-economic status of participants from 5 months to 5 years old based on parental reports.

abased on maximum available samples ranging from *n*=541-721.

Data were compiled from the final master file of the Quebec Longitudinal Study of Child Development (1998–2021), © Gouvernement du Quebec, Institut de la Statistique du Quebec.

Table S2. Included versus excluded participants on key variables based on unweighted valuesab

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
|  | | | | |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  | Prospective | | |  | Retrospective | |  |
|  |  | Participantsc  (n=639) | Non-participants (n=82) | *p-value* |  | Participants (n=591) | Non-participants (n=130) | *p-value* |
| *Child characteristics* | |  |  |  |  |  |  |  |
| Male, N(%) | | 261(40.8) | 62(75.6) | <.001 |  | 228(38.6) | 95(73.1) | <.001 |
| Birth weight, (kg), M(SD) | | 3.42(.47) | 3.43(.61) | .847 |  | 3.41(.47) | 3.47(.54) | 0.224 |
| Non-Canadian, N(%) | | 126(19.9) | 18(22.0) | .664 |  | 116(19.8) | 28(21.7) | .624 |
| Externalizing behaviorsd, M(SD) | | 5.18(2.98) | 5.51(3.07) | .356 |  | 5.14(2.95) | 5.55(3.17) | .182 |
| Internalizing behaviorse, M(SD) | | .88(1.16) | 1.20(1.29) | .025 |  | 0.89(1.16) | 1.09(1.27) | .090 |
| *Family Characteristics* | |  |  |  |  |  |  |  |
| Maternal age at birth, M(SD) | | 29.16(5.06) | 27.98(5.52) | .069 |  | 29.26(5.08) | 27.97(5.22) | .011 |
| Maternal depressionf, M(SD) | | 1.21(1.21) | 1.45(1.22) | .097 |  | 1.21(1.19) | 1.39(1.28) | .148 |
| Non-intact family (single or blended), N(%) | | 107(16.8) | 21(25.6) | .050 |  | 99(16.8) | 29(22.5) | .127 |
| Family socioeconomic statusg, M(SD) | | .17(.93) | -.34(1.00) | <.001 |  | .17(.94) | -.15(1.00) | <.001 |
| *Note.* Data were compiled from the final master file of the Québec Longitudinal Study of Child Development (1998-2021), Québec Government, Québec Statistics Institute.  aVariables were measured when the child was 5 months of age, unless otherwise indicated.  bBased on the target sample of 721 participants (all participants with a valid polygenic risk score). cbased on a sample of 639 participants that had valid data on at least one prospective maltreatment indicator. dAssessed at 29 months, missing values were replaced with 17 months; 10 items from the Behavior Questionnaire (e.g., cannot sit still, is agitated) (Collet et al., 2022), scores range from 0-18.  eAssessed at 29 months, missing values were replaced with 17 months; 6 items from the Behavior Questionnaire (e.g., is too fearful or anxious) (Collet et al., 2022), scores range from 0-8. fAssessed using a shortened version (12 items) of the Center for Epidemiologic Studies-Depression (Poulin, C. *et al.*, 2005). Scores were standardized to range from 0-10. gStandardized index (z-score) based on annual gross income, parental education level and occupational prestige (Willms D.J., & Shields, M., 1996). | | | | | | | | |
|  |
|  |
|  |
|  |
|  |