This protocol is based on the methodology published by de Lima et al 20201. The creation of a Biologically informed Polygenic Risk Score start from defining the critical determinant of the biological functionality that better represents the research question. In this case we were interested in genes that were co-expressed with the serotonin transporter gene (5-HTT, GRCh37.p13/ENSG00000108576) in amygdala, during early life (fetal life/infancy). The expression-based polygenic risk score was created considering genes co-expressed with the serotonin transporter gene (5-HTT-ePRS) in amygdala, according to the protocol previously described by Silveira et al(2017).2 and Miguel et al(2019).3.

\Online resources/databases used for the expression-based polygenic risk score:

(A) GeneNetwork (<http://genenetwork.org>);

(B) BrainSpan (<http://www.brainspan.org>);

(C) NCBI Variation Viewer (<https://www.ncbi.nlm.nih.gov/variation/view>);

(D) The Genotype-Tissue Expression (GTEx) (<https://gtexportal.org/home/>).

**GeneNetwork**(A) was used to generate a list of genes co-expressed with the serotonin 5-HTT receptor gene in the amygdala in mice (Mulligan et al., 2016).4, retaining only the genes with absolute value of co-expression correlation higher or equal to 0.5.

The gene list generated by **GeneNetwork**(A) was then filtered using **BrainSpan**(B) to identify consensus transcripts enriched in the fetal and childhood human brain (Miller et al., 2014)5. Since we were interested in genes that were active during early developmental periods, we selected autosomal transcripts expressed in the amygdala with at least 1.5-fold more expressed during fetal and child development (all fetal samples and up to the first 5 years of age) as compared to adult samples. The final list of genes is made available here as table 1, and the gene network is represented in figure 1.

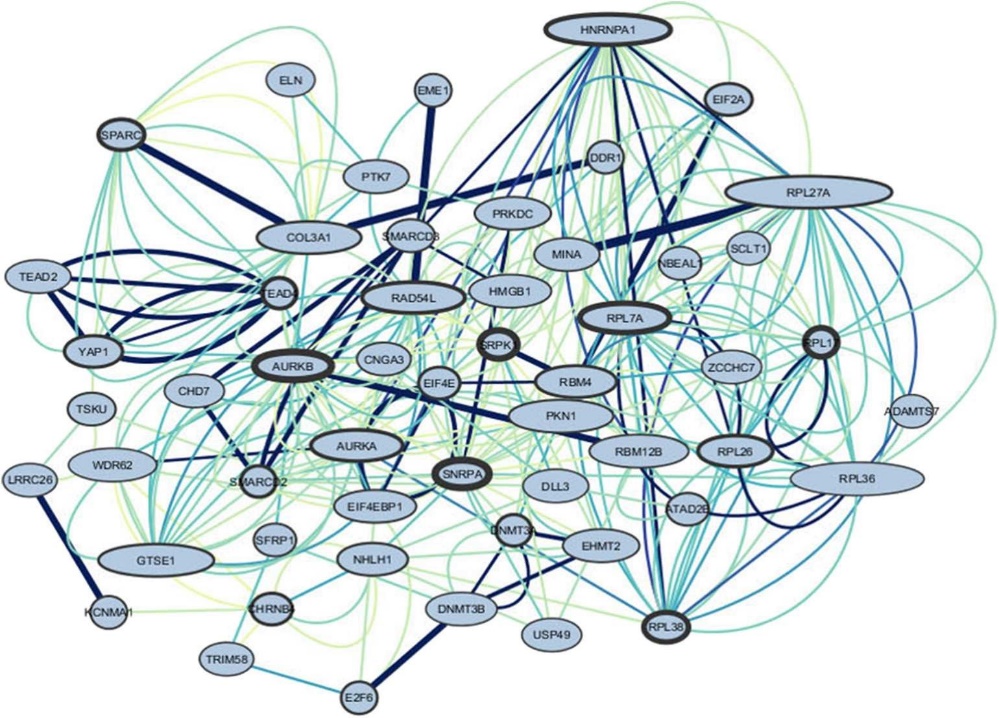


FIGURE 5 | Genes interactions of the amygdala 5-HTT co-expression network. The border of the genes represents the out-degree of these nodes, meaning the number of outgoing relationships. The larger the border, the stronger the relationship of the gene to other genes. The size of the nodes represents the in-degree, i.e., the number of incoming relationships with neighbors. The bigger the node, the higher the relationships of other genes to the target gene. The edges represent

co-expression. Color and thickness were used to identify the most co-expressed genes, darker, and thicker represent higher co-expression.

Figure 1: Representation of the gene’s interactions of the amygdala 5-HTT co-expression network. The larger the border, the stronger the relationship of the gene to other genes. The size of the nodes represents the number of incoming relationships with neighbors. The bigger the node, the higher the relationships of other genes to the target gene. The edges represent co-expression. Color and thickness were used to identify the most co-expressed genes, darker, and thicker represent higher co-expression.

Based on their functional annotation in the National Center for Biotechnology Information, U.S. National Library of Medicine, **NCBI Variation Viewer** (C), using GRCh37.p13, we gathered all the SNPs from these genes and merged this list with the SNPs from the **GTEx data**(D) in human amygdala to form a list of common SNPs. The list of common SNPs was then filtered by subjecting it to linkage disequilibrium clumping (r2 < 0.25), which resulted in a total of 463 SNPs.

In ePRS calculation, alleles at a given cis-SNP were weighed by the estimated effect of the genotype on gene expression (expression quantitative trait loci from GTEx, in which the effect allele is the alternative allele). Final ePRS was obtained by summation over all SNPs accounting for the sign of correlation coefficient between the genes and 5HTT gene expression. For more information on ePRS calculation, see Hari Dass et al. (2019)6. The summation of these values from the total number of SNPs provides the amygdala 5-HTT-ePRS score (Table 1 shows the final gene list). Figure 2 illustrates the steps involved in the creation of the ePRS.

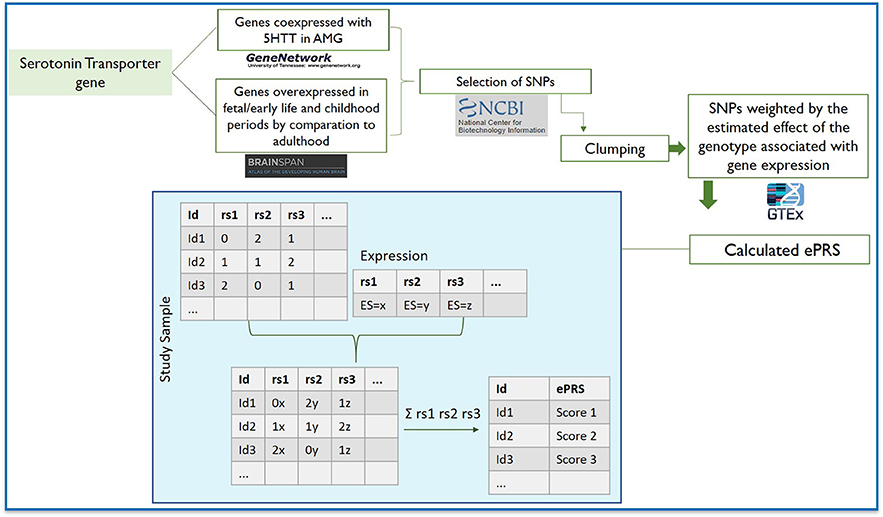


Figure 2. Flowchart depicting the steps involved in creating the expression-based polygenic risk score based on genes co-expressed with serotonin transporter gene in amygdala using gene co-expression databases: GeneNetwork was used to generate a co-expression matrix with 5-HTT gene in the amygdala in mice (absolute value of the co-expression correlation r ≥ 0.5); BrainSpan was then used to identify consensus human transcripts from this list; BrainSpan was also used for selecting genes differentially expressed at ≥1.5-fold during child and fetal development as compared to adulthood within the same brain areas. Based on their functional annotation in the National Center for Biotechnology Information, U.S. National Library of Medicine, we gathered all the existing SNPs from these genes and subjected this list of SNPs to linkage disequilibrium clumping. We used a count function of the number of alleles at a given SNP (rs1, rs2…) weighted by the estimated effect of the genotype associated with gene expression. The sum of these values from the total number of SNPs provides the amygdala 5-HTT ePRS score.

References

1. de Lima RMS, Barth B, Arcego DM, et al. Amygdala 5-HTT Gene Network Moderates the Effects of Postnatal Adversity on Attention Problems: Anatomo-Functional Correlation and Epigenetic Changes. *Front Neurosci-Switz.* 2020;14(198).

2. Silveira PP, Pokhvisneva I, Parent C, et al. Cumulative prenatal exposure to adversity reveals associations with a broad range of neurodevelopmental outcomes that are moderated by a novel, biologically informed polygenetic score based on the serotonin transporter solute carrier family C6, member 4 (SLC6A4) gene expression. *Dev Psychopathol.* 2017;29(5):1601-1617.

3. Miguel PM, Deniz BF, Deckmann I, et al. Prefrontal cortex dysfunction in hypoxic-ischaemic encephalopathy contributes to executive function impairments in rats: Potential contribution for attention-deficit/hyperactivity disorder. *The world journal of biological psychiatry : the official journal of the World Federation of Societies of Biological Psychiatry.* 2017:1-14.

4. Mulligan MK, Mozhui K, Prins P, Williams RW. GeneNetwork: a toolbox for systems genetics. In: *Systems Genetics.* Springer; 2017:75-120.

5. Miller JA, Ding S-L, Sunkin SM, et al. Transcriptional landscape of the prenatal human brain. *Nature.* 2014;508(7495):199-206.

6. Hari Dass SA, McCracken K, Pokhvisneva I, et al. A biologically-informed polygenic score identifies endophenotypes and clinical conditions associated with the insulin receptor function on specific brain regions. *EBioMedicine.* 2019;42:188-202.