Genome-Wide Significance for PCLO as a Gene for Major Depressive Disorder

Hamdi Mbarek,1,a Yuri Milaneschi,2,a Jouke-Jan Hottenga,1 Lannie Ligthart,1 Eco J. C. de Geus,1 Erik A. Ehli,3 Gonneke Willemsen,1 Gareth E. Davies,3 Jan H. Smit,2 Dorret I. Boomsma,1,b and Brenda W. J. H. Penninx2,b

1Department of Biological Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
2Department of Psychiatry, Amsterdam Public Health Research Institute, VU University Medical Center/GGz in Geest, Amsterdam, the Netherlands
3Avera Institute for Human Genetics, Sioux Falls, South Dakota, USA

In 2009, the first genome-wide association study (GWAS) for major depressive disorder (MDD) highlighted an association with PCLO locus on chromosome 7, although not reaching genome-wide significance level. In the present study, we revisited the original GWAS after increasing the overall sample size and the number of interrogated SNPs. In an analysis comparing 1,942 cases with lifetime diagnosis of MDD and 4,565 controls, PCLO showed a genome-wide significant association with MDD at SNP (rs2715157, \( p = 2.91 \times 10^{-8} \)) and gene-based (\( p = 1.48 \times 10^{-7} \)) level. Our results confirm the potential role of the PCLO gene in MDD, which is worth further replication and functional studies.

**Keywords:** major depressive disorder, genome-wide association study, genetics, PCLO

In 2009, as part of the Genetic Association Information Network (GAIN) initiative, we published (Sullivan et al., 2009) the first genome-wide association study (GWAS) for major depressive disorder (MDD). The participants were assembled from two large Dutch cohorts: the Netherlands Study of Depression and Anxiety (NESDA) and the Netherlands Twin Registry (NTR). A comparison between 1,738 cases with lifetime diagnosis of DSM-IV MDD and 1,802 screened controls highlighted a signal overlapping the region of PCLO gene, although variants identified did not reach genome-wide significance. Importantly, our result received strong support (though not reaching formal significance) from Australian data similarly ascertained from a population twin registry (Sullivan et al., 2009). Based on the promising role of PCLO encoding a cytoskeletal protein located in the presynaptic active zone, the original finding was followed by replication efforts, with positive (Hek et al., 2010) and negative (Verbeek et al., 2013) results. In the present study, we revisited the original GWAS after introducing important improvements. The overall sample size was almost doubled, adding newly genotyped participants from the original cohorts, consisting mainly of a large number of controls carefully screened using several sources of information including longitudinal assessments and family history (Boomsma et al., 2008). The number of interrogated SNPs was increased by re-genotyping the majority of participants on a newer array and by imputing variants with a denser reference panel (1,000 Genomes Phase 3).

The present sample consisted of 6,507 participants (63.1% females) of European ancestry from NESDA and NTR. Lifetime MDD diagnoses according to DSM-IV were ascertained using the Composite Interview Diagnostic Instrument. Healthy controls were screened based on absence of any lifetime psychiatric disorder (NESDA), no report of MDD, no known first-degree relatives with MDD, and a low factor score based on a multivariate analysis of depressive complaints, anxiety, neuroticism, and somatic anxiety (Boomsma et al., 2000). After applying stringent quality control criteria (Supplementary Methods), we performed a GWAS comparing 1,942 cases and 4,565 controls,

**RECEIVED 6 April 2017; accepted 20 April 2017. First published online 25 May 2017.**

**ADDRESS FOR CORRESPONDENCE:** Hamdi Mbarek, Department of Biological Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, the Netherlands.
E-mail: h.mbarek@vu.nl/y.milaneschi@ggzingeest.nl

[a] These authors contributed equally to this work.
[b] These authors contributed equally to this work.
adjusting for gender and 10 ancestry-informative principal components. Family structure was taken into account using a sandwich estimator implemented in Plink.

The estimated LD score regression intercept was 1.03, suggesting no appreciable inflation of the test statistics attributable to unaccounted stratification (Figure S1). The analysis identified 17 genome-wide significant SNPs ($p < 5 \times 10^{-8}$; Figure 1) spanning the PCLO gene on chromosome 7 with rs2715157 emerging as the SNP most significantly associated with MDD ($p = 2.91 \times 10^{-8}$); its minor allele (A) was negatively associated with MDD ($OR = 0.79$). The PCLO lead SNP identified in the original GWAS, rs2715148 showed the same effect size as in the original analysis but was now genome-wide significantly associated with MDD ($p = 3.89 \times 10^{-8}$; $OR = 0.79$). Similarly, the non-synonymous coding SNP rs2522833 reported the
same effect size of the original GWAS but a smaller \( p \) value \((OR = 1.26, p = 8.29 \times 10^{-8})\). These SNPs were highly correlated with rs2715157 \((r^2 > 0.8; \text{Figure 2})\). Furthermore, in silico functional annotation of the top SNPs showed possible functional effects. The PCLO SNP rs2715157 alters the sequence of six protein-binding motifs including Pou3f2, indicating a possible effect on the activation of corticotrophin releasing hormone. It also bound to the P300 protein known to regulate gene transcription. SNPs rs2751161 and rs2522835 were contained in a DNase I hypersensitive site suggesting open chromatin, although they did not alter any of the transcription factor binding sites present in the promoter flanking region. Together, these data indicate these SNPs might have direct functional roles (Table S1). Finally, \( p \) values of analyzed SNPs were used to perform a gene-based test (Supplementary Methods), and PCLO showed a significant association with MDD \((p = 1.48 \times 10^{-7})\), even considering a stringent multiple comparison adjustment accounting for 20,000 genes \((p = 2.5 \times 10^{-6})\).

These findings provide evidence for an association between the PCLO gene and MDD that now clearly reaches genome-wide significance. An important factor was the increase in sample size. Indeed, the effect size detected for SNPs present in both the original and current analyses were exactly the same, but the improved precision in their estimation impacted the test significance. Allele frequencies were also the same for SNPs present in both analyses, suggesting no probable effect on results due to technical differences between the genotyping platforms used in the two GWASs. Nevertheless, imputation allowed us to examine additional SNPs in the same LD block previously not included.

It is important to note that the effectiveness of sample size increase may have been amplified by the clinical homogeneity of the sample and the detailed phenotyping procedure. Recently, the assembling of massive samples has enabled the success of GWAS for depression (Hyde et al., 2016; Okbay et al., 2016), overcoming the problematic phenotypic heterogeneity of this trait. Nevertheless, large collaborative studies clearly identified substantial heterogeneity among different MDD datasets as indexed by cross-dataset SNP correlations smaller than one (Lee et al., 2013), possibly indicating the presence of dataset specific variants. Therefore, collecting a larger sample size, although on a smaller scale,
in more homogeneous populations still constitutes a valuable complementary effort aimed at clarifying specific genetic signals that might be overshadowed by the heterogeneity arising from the addition of each new dataset in meta-analytic studies.

To conclude, in our current genome-wide study of MDD, we found confirming evidence of a genome-wide significant association with the \textit{PCLO} gene. Our findings, together with previous biological evidence (Ahmed et al., 2015; Waites et al., 2011), suggest the importance of the \textit{PCLO} gene to MDD, which is worth further replication and functional studies.

\textbf{Conflict of Interest}

None.

\textbf{Acknowledgments}

Netherland Twin Register and Netherlands Study of Depression and Anxiety (NESDA): Funding was obtained from the Netherlands Organization for Scientific Research (NWO) and MagW/ZonMW grants Middelgroot-911-09-032, Spinozapremie 56-464-14192, Geestkracht program of the Netherlands Organization for Health Research and Development (ZonMW 10-000-1002), Center for Medical Systems Biology (CSMB, NOW Genomics), Genetic influences on stability and change in psychopathology from childhood to young adulthood (ZonMW 912-10-020), NBIC/BioAssist/RK(2008.024), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL, 184.021.007), VU University’s Institute for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam (NCA); the European Science Council (ERC Advanced, 230374). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health, Rutgers University Cell and DNA Repository (NIMH U24 MH068457-06), the Avera Institute for Human Genetics, Sioux Falls, South Dakota (USA) and the National Institutes of Health (NIH R01 HD042157-01A1, MH081802, Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995).

\textbf{Supplementary material}

To view supplementary material for this article, please visit https://doi.org/10.1017/thg.2017.30

\textbf{References}


