Twin Research and Human Genetics
Volume 20 ▌ Number 4 ▌ pp. 267–270 ▌ © The Author(s) 2017 ▌ doi:10.1017/thg.2017.30

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. Penninx[2,b]

[1]*Department of Biological Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands*
[2]*Department of Psychiatry, Amsterdam Public Health Research Institute, VU University Medical Center/GGz inGeest, Amsterdam, the Netherlands*
[3]*Avera 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).

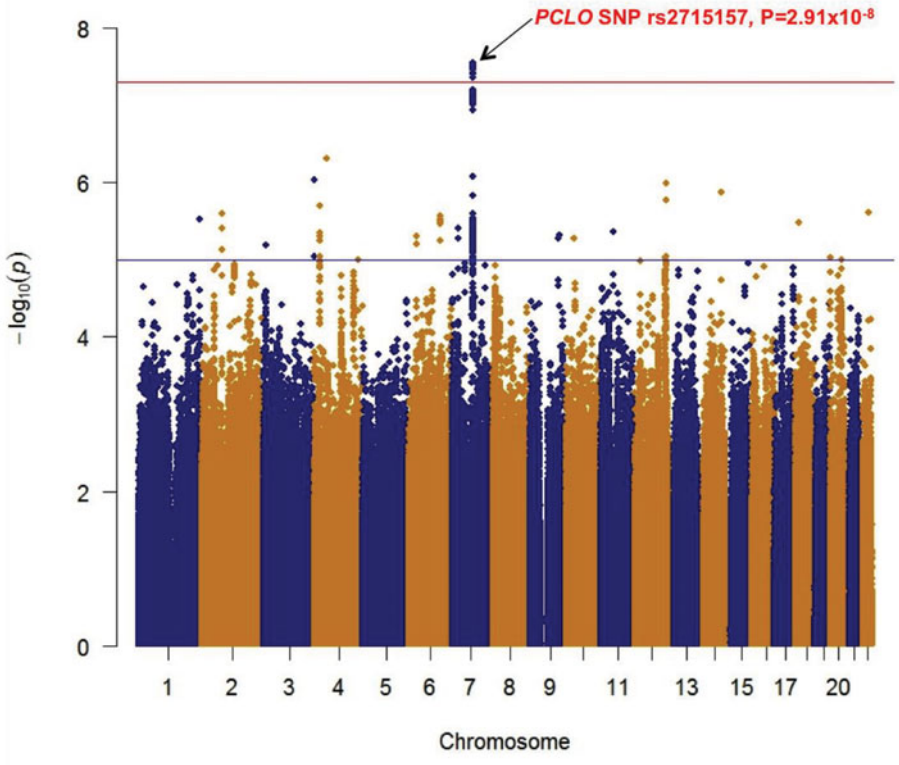
The present sample consisted of 6,507 participants (63.1% females) of European ancestry from NESDA and NTR. Lifetime MDD diagnoses according to *DSM-IV* were ascertained using the Composite Interview Diagnostic Instrument. Healthy controls were screened based on absence of any lifetime psychiatric disorder (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.milanechi@ggzingeest.nl

[a] These authors contributed equally to this work.
[b] These authors contributed equally to this work.

Downloaded from https://www.cambridge.org/core. Vrije Universiteit, on 09 Aug 2017 at 12:58:35, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/thg.2017.30



B

SNP	Position[a]	Annotation	Effect Allele	EAF	OR (95% CI)	P
rs2715157	82468374	intron	A	0.48	0.79 (0.73-0.86)	2.91×10^{-8}
rs2371214	82454672	Intron	C	0.50	1.27 (1.17-1.38)	3.08×10^{-8}
rs2715151	82456663	intron	C	0.48	0.79 (0.73-0.86)	3.21×10^{-8}
rs2371215	82457494	intron	T	0.48	0.79 (0.73-0.86)	3.21×10^{-8}
rs2715153	82457858	intron	C	0.48	0.79 (0.73-0.86)	3.21×10^{-8}
rs2715154	82460866	intron	G	0.48	0.79 (0.73-0.86)	3.21×10^{-8}
7:82457498	82457498	intron	A	0.48	0.79 (0.73-0.86)	3.32×10^{-8}
rs2522836	82473053	intron	A	0.50	1.27 (1.16-1.38)	3.33×10^{-8}
rs2522831	82448100	intron	C	0.50	1.27 (1.16-1.38)	3.44×10^{-8}
rs2522832	82450364	3'-UTR	G	0.50	1.27 (1.16-1.38)	3.44×10^{-8}
rs2715161	82482281	intron	C	0.48	0.79 (0.73-0.86)	3.56×10^{-8}
rs2715147	82448405	intron	C	0.48	0.79 (0.73-0.86)	3.89×10^{-8}
rs2715148	82450035	3'-UTR	A	0.48	0.79 (0.73-0.86)	3.89×10^{-8}
rs2247523	82454404	intron	G	0.50	1.26 (1.16-1.37)	3.96×10^{-8}
rs2522835	82471496	intron	T	0.48	0.79 (0.73-0.86)	4.02×10^{-8}
rs2888019	82503409	intron	T	0.48	0.79 (0.73-0.86)	4.50×10^{-8}
rs1986742	82503442	intron	C	0.48	0.79 (0.73-0.86)	4.50×10^{-8}

FIGURE 1

(Colour online) A. Genome-wide association results for MDD in NTR-NESDA cohorts. The *y* axis denotes the -log10 (*p* value) for association. The *x* axis gives the physical position of SNPs across the genome. The red line indicates the threshold of genome-wide significance $p < 5 \times 10^{-8}$, and the blue line indicates the threshold of $p < 5 \times 10^{-5}$. B. Genome-wide significant loci in the GWAS for MDD in NTR-NESDA cohorts.

Note: EAF = effect allele frequency; *OR* = odds ratio; 95% CI = 95% confidence interval. [a]SNP positions are according to NCBI Human Genome Build 37.

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

Downloaded from https://www.cambridge.org/core. Vrije Universiteit, on 09 Aug 2017 at 12:58:35, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/thg.2017.30

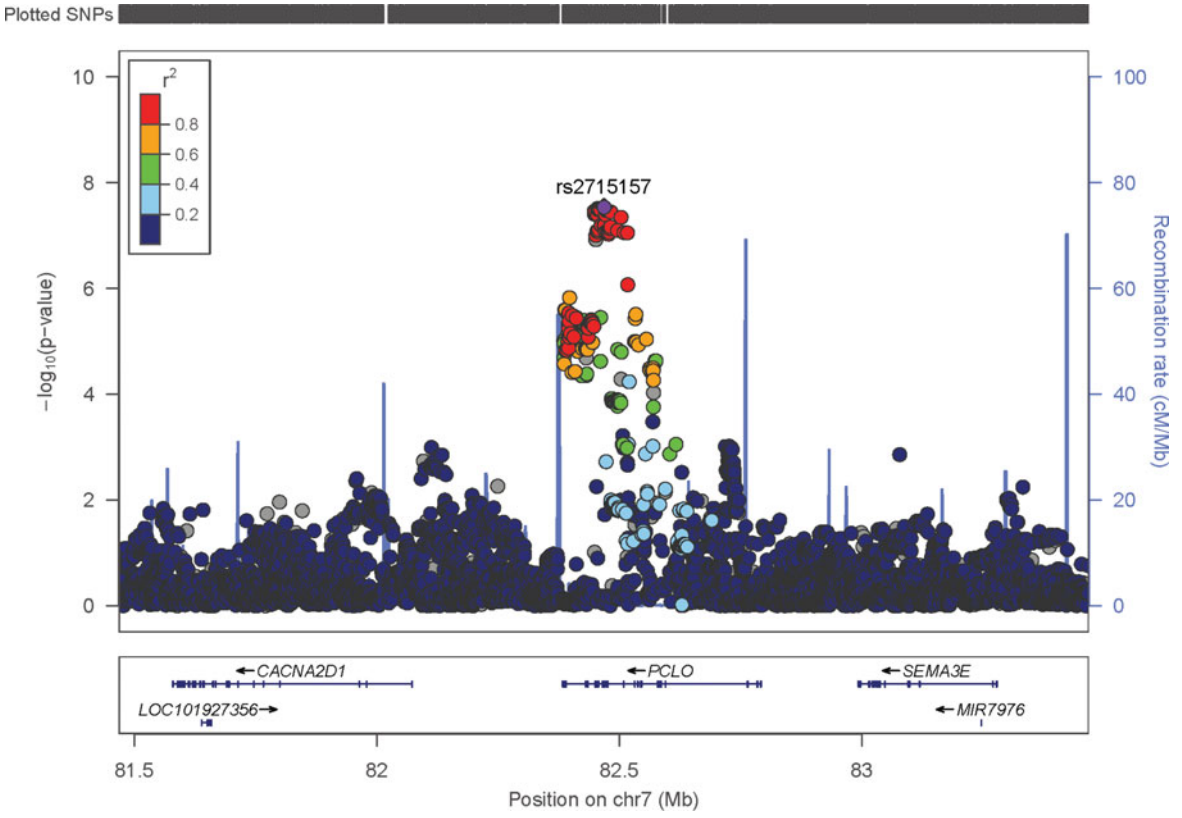


FIGURE 2

(Colour online) Regional association plot of *PCLO* region. The -log10 *p* values (*y* axis) of the SNPs are shown according to their chromosomal positions (*x* axis). The estimated recombination rates from the 1,000 Genomes Project March 2012 release are shown as blue lines, and the genomic locations of genes within the regions of interest in the NCBI Build 37 human assembly are shown as arrows. SNP color represents LD with the most highly associated SNP. The figure was created with LocusZoom (http://csg.sph.umich.edu/locuszoom/).

Note: Mb = megabases.

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 ($r^2 > 0.8$; 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,

Downloaded from https://www.cambridge.org/core. Vrije Universiteit, on 09 Aug 2017 at 12:58:35, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/thg.2017.30

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 *PCLO* gene. Our findings, together with previous biological evidence (Ahmed et al., 2015; Waites et al., 2011), suggest the importance of the *PCLO* gene to MDD, which is worth further replication and functional studies.

Conflict of Interest

None.

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).

Supplementary material

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

References

Ahmed, M. Y., Chioza, B. A., Rajab, A., Schmitz-Abe, K., Al-Khayat, A., Al-Turki, S., … Mochida, G. H. (2015). Loss of PCLO function underlies pontocerebellar hypoplasia type III. *Neurology, 84*, 1745–1750.

Boomsma, D. I., Beem, A. L., van den Berg, M., Dolan, C. V., Koopmans, J. R., Vink, J. M., … Slagboom, P. E. (2000). Netherlands twin family study of anxious depression (NETSAD). *Twin Research, 3*, 323–334.

Boomsma, D. I., Willemsen, G., Sullivan, P. F., Heutink, P., Meijer, P., Sondervan, D., … Penninx, B. W. (2008). Genome-wide association of major depression: Description of samples for the GAIN major depressive disorder study: NTR and NESDA biobank projects. *European Journal of Human Genetics, 16*, 335–342.

Hek, K., Mulder, C. L., Luijendijk, H. J., van Duijn, C. M., Hofman, A., Uitterlinden, A. G., & Tiemeier, H. (2010). The PCLO gene and depressive disorders: Replication in a population-based study. *Human Molecular Genetics, 19*, 731–734.

Hyde, C. L., Nagle, M. W., Tian, C., Chen, X., Paciga, S. A., Wendland, J. R., … Winslow, A. R. (2016). Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nature Genetics, 48*, 1031–1036.

Lee, S. H., Ripke, S., Neale, B. M., Faraone, S. V., Purcell, S. M., Perlis, R. H., … Wray, N. R. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics, 45*, 984–994.

Okbay, A., Baselmans, B. M., De Neve, J. E., Turley, P., Nivard, M. G., Fontana, M. A., … Cesarini, D. (2016). Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics, 48*, 624–633.

Sullivan, P. F., de Geus, E. J., Willemsen, G., James, M. R., Smit, J. H., Zandbelt, T., … Penninx, B. W. (2009). Genome-wide association for major depressive disorder: A possible role for the presynaptic protein piccolo. *Molecular Psychiatry, 14*, 359–375.

Verbeek, E. C., Bevova, M. R., Bochdanovits, Z., Rizzu, P., Bakker, I. M., Uithuisje, T., … Heutink, P. (2013). Resequencing three candidate genes for major depressive disorder in a Dutch cohort. *PLoS One, 8*, e79921.

Waites, C. L., Leal-Ortiz, S. A., Andlauer, T. F., Sigrist, S. J., & Garner, C. C. (2011). Piccolo regulates the dynamic assembly of presynaptic F-actin. *Journal of Neuroscience, 31*, 14250–14263.

Downloaded from https://www.cambridge.org/core. Vrije Universiteit, on 09 Aug 2017 at 12:58:35, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/thg.2017.30