Journal of Affective Disorders 191 (2016) 123–131

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Genome-wide linkage on chromosome 10q26 for a dimensional scale of major depression

Emma E.M. Knowles [a,b,*], Jack W. Kent Jr. [c], D. Reese McKay [a,b], Emma Sprooten [a,b], Samuel R. Mathias [a,b], Joanne E. Curran [g], Melanie A. Carless [c], Marcio A.A. de Almeida [g], H.H. Goring Harald [g], Tom D. Dyer [g], Rene L. Olvera [d], Peter T. Fox [e,f], Ravi Duggirala [g], Laura Almasy [g], John Blangero [g], David C. Glahn [a,b]

[a] Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA
[b] Olin Neuropsychiatric Research Center, Institute of Living, Hartford Hospital, Hartford, CT, USA
[c] Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX, USA
[d] Department of Psychiatry, University of Texas Health Science Center San Antonio, Texas Center San Antonio, San Antonio, TX, United States
[e] Research Imaging Institute, University of Texas Health Science Center San Antonio, San Antonio, TX, United States
[f] South Texas Veterans' Healthcare System, 7400 Merton Minter, San Antonio, TX 78229, USA
[g] South Texas Diabetes and Obesity Institute, University of Texas Health Science Center at San Antonio & University of Texas of the Rio Grande Valley, Brownsville, TX, United States

ARTICLE INFO

Article history:
Received 25 August 2015
Received in revised form
27 October 2015
Accepted 9 November 2015
Available online 17 November 2015

ABSTRACT

Major depressive disorder (MDD) is a common and potentially life-threatening mood disorder. Identifying genetic markers for depression might provide reliable indicators of depression risk, which would, in turn, substantially improve detection, enabling earlier and more effective treatment. The aim of this study was to identify rare variants for depression, modeled as a continuous trait, using linkage and post-hoc association analysis. The sample comprised 1221 Mexican–American individuals from extended pedigrees. A single dimensional scale of MDD was derived using confirmatory factor analysis applied to all items from the Past Major Depressive Episode section of the Mini-International Neuropsychiatric Interview. Scores on this scale of depression were subjected to linkage analysis followed by QTL region-specific association analysis. Linkage analysis revealed a single genome-wide significant QTL (LOD = 3.43) on 10q26.13, QTL-specific association analysis conducted in the entire sample revealed a suggestive variant within an intron of the gene *LHPP* (rs11245316, $p = 7.8 \times 10^{-04}$; LD-adjusted Bonferroni-corrected $p = 8.6 \times 10^{-05}$). This region of the genome has previously been implicated in the etiology of MDD; the present study extends our understanding of the involvement of this region by highlighting a putative gene of interest (*LHPP*).

1. Introduction

Major depressive disorder (MDD) is a common and potentially life-threatening mood disorder (Sullivan et al., 2000). It affects 16.2% of individuals in the US during their lifetime (Kessler et al., 2003), and incurs great economic cost ($83.1 billion per annum in the US) (Greenberg et al., 2003). The illness also places an immense burden on the sufferer, such that the impact of MDD on wellbeing and functioning is in line with that seen in other major chronic conditions (e.g., arthritis and diabetes mellitus) (Wells et al., 1989). Moreover, functional impairments remain even after the remission of a depressive episode (Hays et al., 1995). Unsurprisingly, the World Health Organization (WHO) cites MDD as a leading cause of disability worldwide (World Health Organization, 2012). Current methods of diagnosing and treating MDD are symptom based, that is, diagnosis is made based on the presence of symptoms outlined in the DSM (American Psychiatric Association, 1994) and successful treatment is defined by the reduction and eventual remission of those symptoms (Binder and Holsboer, 2005). Relying on symptoms alone, without regard for the etiological roots of a disorder, makes for mediocre diagnostic reliability (Bromet et al., 1986; Kendler et al., 1993; Keller et al., 1996) and inadequate treatment (U.S. Department of Health and Human Services, 1993). The effectiveness of anti-depressant pharmacotherapy is hampered by our limited understanding of the

* Corresponding author at: Olin Neuropsychiatry Research Center, Whitehall Research Building, Institute of Living, 200 Retreat Ave, Hartford, CT 06106, United States.
E-mail address: emma.knowles@yale.edu (E.E.M. Knowles).

http://dx.doi.org/10.1016/j.jad.2015.11.012
0165-0327/

biological basis of MDD (Binder and Holsboer, 2005; Lopez et al., 2007); indeed the administration of anti-depressant medications results in remission in only one third of patients (Rush et al., 2006). Identifying risk variants for depression would enhance our understanding of the etiology of MDD which in turn would enable earlier and more reliable detection as well as, potentially, the development of new and more effective therapies (Miller and O'Callaghan, 2013; Flint and Kendler, 2014).

Heritability estimates of MDD vary around 0.37 (Sullivan et al., 2000; Kendler et al., 2006), indicating a substantial influence of genes on MDD risk. However, attempts to isolate specific genes which mediate MDD risk have been met with difficulty (Flint and Kendler, 2014; Cohen-Woods et al., 2013): meta-analysis suggests that many of the early candidate gene studies were false positives (Bosker et al., 2011; Wray et al., 2012), and numerous genome-wide association (GWA) studies, including the latest mega-analysis of over nine thousand depressed individuals from the Psychiatric Genetics Consortium, have struggled to attain genome-wide significant results (Wray et al., 2012; Sullivan et al., 2009; Rietschel et al., 2010; Lewis et al., 2010; Muglia et al., 2010; Kohli et al., 2011; Shyn et al., 2011; Shi et al., 2011; Hek et al., 2013; Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, 2012; Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium et al., 2013). As a consequence it has been suggested that even larger sample sizes are necessary for the identification of risk variants for MDD (Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium et al., 2013). However, linkage, a method that ostensibly measures rare in addition to common variation, has isolated numerous genome-wide significant loci in relatively small samples (Neff et al., 2009; Breen et al., 2011; Pergadia et al., 2011; Glahn et al., 2012). Thus, while intuitively the assertion that greater sample sizes are needed to isolate genes for MDD makes good sense (particularly given that increasing sample size has worked for other disorders e.g. schizophrenia (Ripke et al., 2013)), it is also possible that the degree of genetic (and also phenotypic) heterogeneity is greater for MDD than for other disorders (Collins and Sullivan, 2013) and as a consequence increasing sample sizes might only serve to compound the problem. Therefore a complementary approach would be to focus on reducing genetic heterogeneity using, for example, a family-based approach when searching for MDD risk genes.

In order to effectively account for the phenotypic heterogeneity associated with MDD it is critical to develop optimal MDD phenotypes (Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium et al., 2013). MDD is typically treated as a categorical trait, it is assumed that MDD reflects the tail end of an underlying normal distribution of mood, and that diagnosis occurs when a threshold for liability is crossed. It seems plausible that the genes which moderate behavior at the tail end of the distribution are the same as those that underlie the regulation of normal mood (Luft, 2002) and by dichotomizing the MDD distribution, we ignore a substantial proportion of variance that would contribute to gene-finding efforts. Conceptualizing MDD as a continuous dimension would capture this important information which would confer greater sensitivity and power to detect genes (Duggirala et al., 1997).

Thus, the present study, we report on univariate linkage and association analysis of a dimensional scale of depression derived from the Past Major Depressive Episode section of the Mini-International Neuropsychiatric Interview (MINI) within extended-pedigree data.

2. Methods

2.1. Participants

The sample comprised 1221 Mexican–American individuals from extended pedigrees (132 families, average size 9.32 people, range = 1–129). The sample was 63% female and had a mean age of 46.01 (SD = 15.10; range = 18–97). Individuals in this San Antonio Family Study cohort have actively participated in research for over 18 years and were randomly selected from the community with the constraints that they are of Mexican–American ancestry, part of a large family, and live within the San Antonio region (see (Olvera et al., 2011) for recruitment details). All participants provided written informed consent on forms approved by the institutional review board at the University of Texas Health Science Center of San Antonio.

2.2. Diagnostic assessment

All participants received the Mini-International Neuropsychiatric Interview (MINI) (Sheehan et al., 1998), which is a semi-structured interview augmented to include items on lifetime diagnostic history. Masters- and doctorate-level research staff, with established reliability ($\kappa \geq 0.85$) for affective disorders, conducted all interviews. All subjects with possible psychopathology were discussed in case conferences that included licensed psychologists or psychiatrists. Lifetime consensus diagnoses were determined based on available medical records, the MINI interview, and the interviewer's narrative.

2.3. Data analysis

2.3.1. Depression modeling: confirmatory factor analysis

All items from the Past Major Depressive Episode (A3a–g) section of the MINI were modeled using a single factor score; Table 1 outlines each of these items. This enabled the categorical outcomes associated with the A3a–g MINI items to be modeled as a unitary quantitative trait. It is important to note that because the factor model included all items from the *past* major depressive episode section, the resultant score should be thought of as a lifetime rating of depression not a reflection of current symptom severity. the Specifically, a single-factor model was built using confirmatory factor analysis in Mplus (Fig. 1). Family structure was taken into account using the *cluster* command, under the cluster command in Mplus standard errors in the model are adjusted in

Table 1
Descriptive statistics for each item for the past major depressive episode section of the MINI.

When you felt depressed or uninterested:	% Responded no	h^2	SE
Was your appetite decreased or increased nearly every day? Did your weight decrease or increase without trying intentionally?	0.663	0.270	0.074
Did you have trouble sleeping nearly every night?	0.659	0.203	0.087
Did you talk or move more slowly than normal or were you fidgety, restless or having trouble sitting still almost every day?	0.725	0.177	0.243
Did you feel tired or without energy almost every day?	0.691	0.290	0.074
Did you feel worthless or guilty almost every day?	0.730	0.193	0.224
Did you have daily difficulty concentrating or making decisions?	0.744	0.245	0.100
Did you repeatedly consider hurting yourself, feel suicidal, or wish that you were dead?	0.840	0.339	0.153

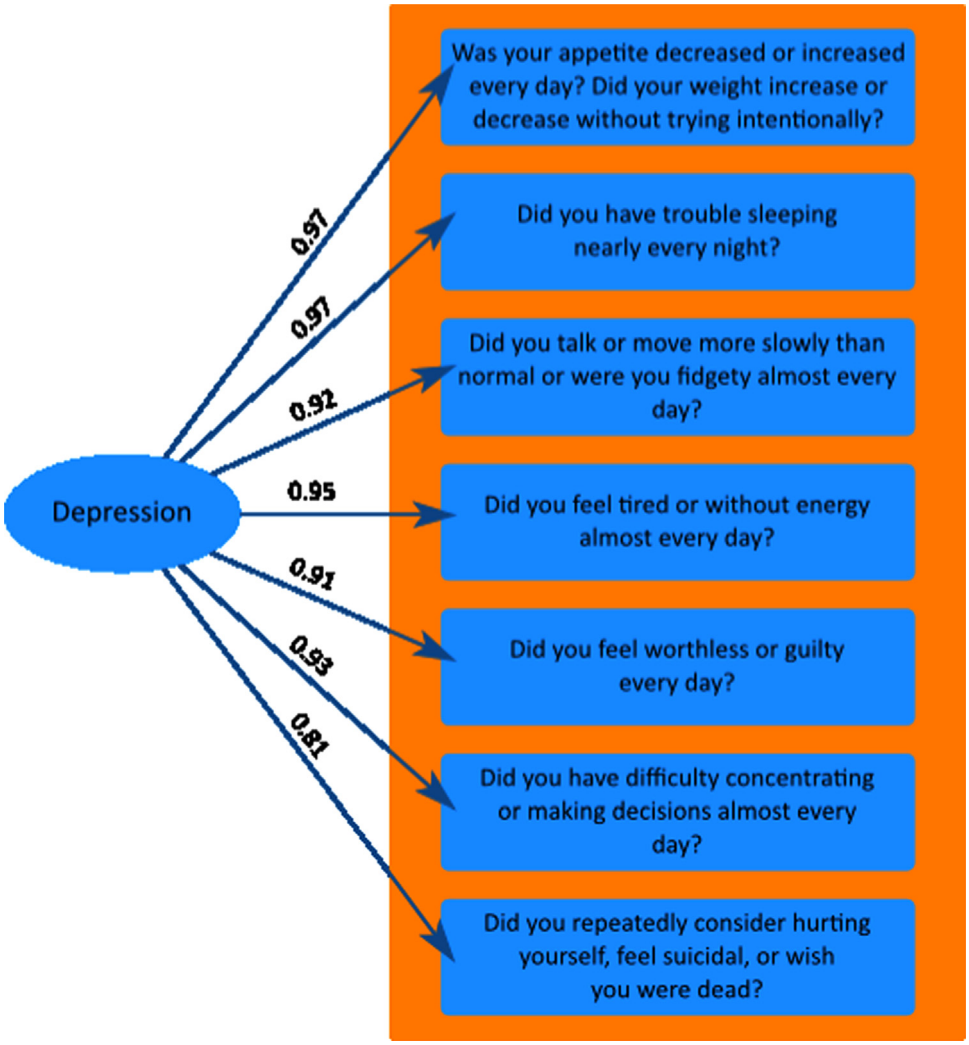

Fig. 1. One-factor confirmatory factor model of all items from the past major depressive episode section of the MINI.

accordance with non-independence in the data, in this way family ID is treated as a nuisance covariate. Because the questionnaire items have categorical rather than continuous outcomes, factor analysis was applied to tetrachoric correlations derived from the raw phenotypic data. The resultant factor score (mean=0.15, standard deviation=0.59) was subjected to an inverse normalization to ensure normality.

2.3.2. Genotyping

Subjects were genotyped for approximately one million SNPs using Illumina HumanHap550v3, HumanExon510Sv1, Human1Mv1 and Human1M-Duov3 BeadChips, according to the Illumina Infinium protocol (Illumina, San Diego, CA). SNP loci were checked for Mendelian consistency utilizing SimWalk2 (Sobel and Lange, 1996). SNPs or samples exhibiting high calling rate failures or requiring excessive blanking (i.e., if $<95\%$ of the genotypes are retained) were eliminated from analyses. Missing genotypes were imputed according to Mendelian laws based on available pedigree data using MERLIN (Abecasis et al., 2002). Maximum likelihood techniques, accounting for pedigree structure, were used to estimate allelic frequencies (Boehnke, 1991). For linkage analyses, multipoint identity-by-descent (IBD) matrices were calculated based on 28,387 SNPs selected from the 1 M GWAS panel as follows. Using genotypes for 345 founders, SNPs on each chromosome were selected to be at least 1 kb apart, MAF≥5%, and LD within a 100 kb sliding window not exceeding |rho|=0.15. The resulting selection averaged 7–8 SNPs/centimorgan. For each centimorgan location in the genome, multipoint IBD probability matrices were calculated using a stochastic Markov Chain Monte Carlo procedure implemented in the computer package, LOKI (Heath, 1997).

2.3.3. Quantitative genetic analyses

All genetic analyses were performed in SOLAR (Almasy and Blangero, 1998). SOLAR implements a maximum likelihood variance decomposition to determine the contribution of genes and environmental influence to a trait by modeling the covariance among family members as a function of expected allele sharing given the pedigree. In the simplest such decomposition, the additive genetic contribution to a trait is represented by the heritability, or h^2, index. Univariate variance decomposition analysis was applied to the continuous measure of depression. The trait was normalized using an inverse Gaussian transformation. Age, age^2, sex and their interactions were included as covariates.

2.3.4. Linkage and association analyses

Quantitative trait linkage analysis was performed to localize specific chromosomal locations influencing MDD (Almasy and Blangero, 1998). Model parameters were estimated using maximum likelihood. The hypothesis of significant linkage was assessed by comparing the likelihood of a classical additive polygenic model with that of a model allowing for both a polygenic component and a variance component due to linkage at a specific chromosomal location (as evidenced by the location-specific

identity-by-descent probability matrix). The LOD score, given by the log10 of the ratio of the likelihoods of the linkage and the polygenic null models, served as the test statistic for linkage. Genome-wide thresholds for linkage evidence were computed for this exact pedigree structure and density of markers, using a method derived from Feingold et al. (1993): a LOD of 1.69 is required for suggestive significance (likely to happen by chance less than once in a genome-wide scan) and a LOD of 2.9 is required for genome-wide significance.

Genomic regions meeting genome-wide significance for linkage were investigated in greater detail using association analysis of the MDD confirmatory factor score and the genetic variants encapsulated by the linkage peak. Statistical significance levels were established according to the effective number of tested variants given the linkage disequilibrium (LD) structure in the region. To this end, the pairwise genotypic correlations were calculated in an effort to establish the effective number of independent tests carried out during association analysis. This method, by Moskvina and Schmidt (Moskvina and Schmidt, 2008), is considered to be conservative and entails computing the eigenvalues of the genotypic correlation matrix. A corrected *P*-value is obtained from a Bonferroni correction based on the nominal alpha (=0.05) and the total number of independent tests.

3. Results

3.1. Confirmatory factor analysis

All MINI items were shown to be significantly heritabile (Table 1). The bivariate tetrachoric correlations (see Table S1) were uniformly moderate to high with little discriminability between items, suggesting a single underlying dimension. A one-factor model fit the data excellently ($X^2_{12}=13.82$, $p=0.129$, RMSEA=0.020 (0.000–0.040) $p=0.995$, CFI=1.000, TLI=1.000, WRMR=0.617).

Table 2
Estimates for the Top Five SNPs from the QTL-specific Association Analysis in (a) the Entire Sample and (b) the Multiplex MDD Pedigree for the Continuous Depression Score.

Entire sample						
SNP	χ^2	*p-Value*	β	Variance explained	MAF	HWE *p-Value*
rs11245316	11.28	0.000783	0.15	0.009	0.213	0.88
rs3884528	9.09	0.002565	−0.11	0.008	0.466	0.18
rs4578341	7.51	0.006127	0.12	0.008	0.219	0.98
rs1123988	7.38	0.00658	0.11	0.007	0.006	0.86
rs859556	7.30	0.00688	−0.58	0.006	0.006	0.96
Multiplex pedigree						
rs7913161	17.84	0.000024	0.69	0.150	0.163	0.38
rs4995180	16.19	0.000057	1.35	0.139	0.025	0.45
rs7906808	15.35	0.000089	1.95	0.133	0.006	0.66
rs7906939	14.28	0.000158	1.52	0.131	0.010	0.69
rs7095366	14.17	0.000167	1.18	0.116	0.021	0.51

3.2. Heritability and linkage analysis

The score derived from the factor model was deemed to be significantly heritable ($h^2=0.21$, $p=2.3\times10^{-05}$). Significant univariate linkage was detected for the depression trait on chromosome 10 at 153 cM (LOD=3.43; Fig. 2). The majority of this linkage signal originated from a single multiplex MDD pedigree within the data ($h^2=0.33$, $p=1.7\times10^{-02}$, LOD=1.54) and the top LOD for the multiplex pedigree within the region encapsulated by the linkage peak met the criteria for suggestive significance (LOD=1.84, 152 cM).

3.3. Association analysis

Association analysis was conducted using all variants within the linkage peak (defined as 150–154 cM) and the continuous factor score of (Table 2 and Fig. 3), the peak-wide (LD-adjusted Bonferroni-corrected) significance level=8.7×10^{-05} (975 SNPs, 590.69 effective SNPs). For association analysis run in the entire sample, the top-ranked variant was suggestively significant

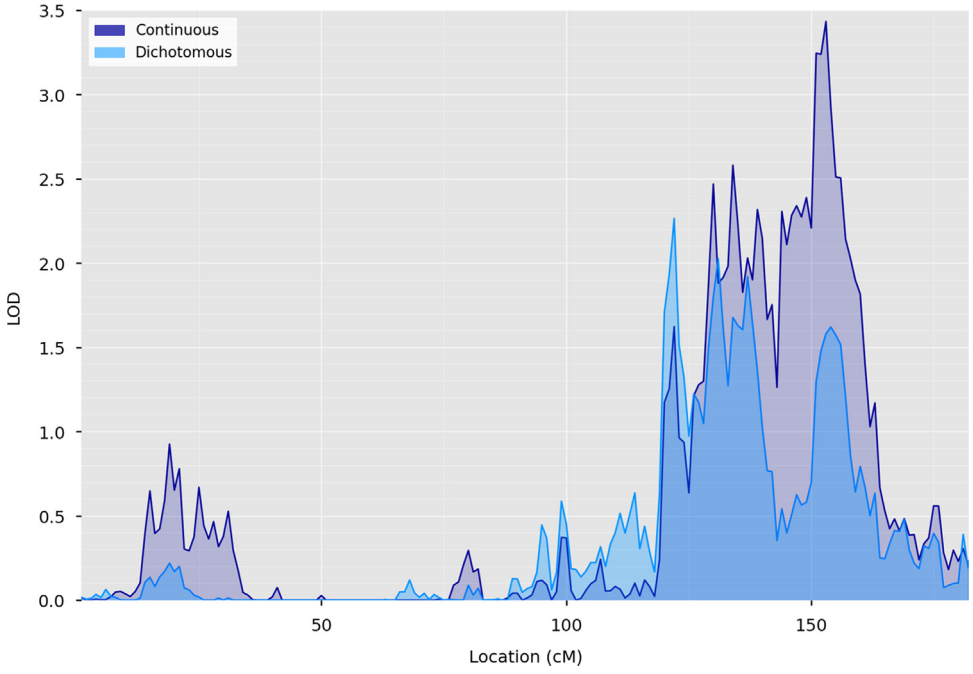

Fig. 2. Chromosome 10 multipoint plot for the univariate linkages of the continuous depression factor score (dark blue) and also the dichotomous measure of depression derived from precisely the same items (light blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

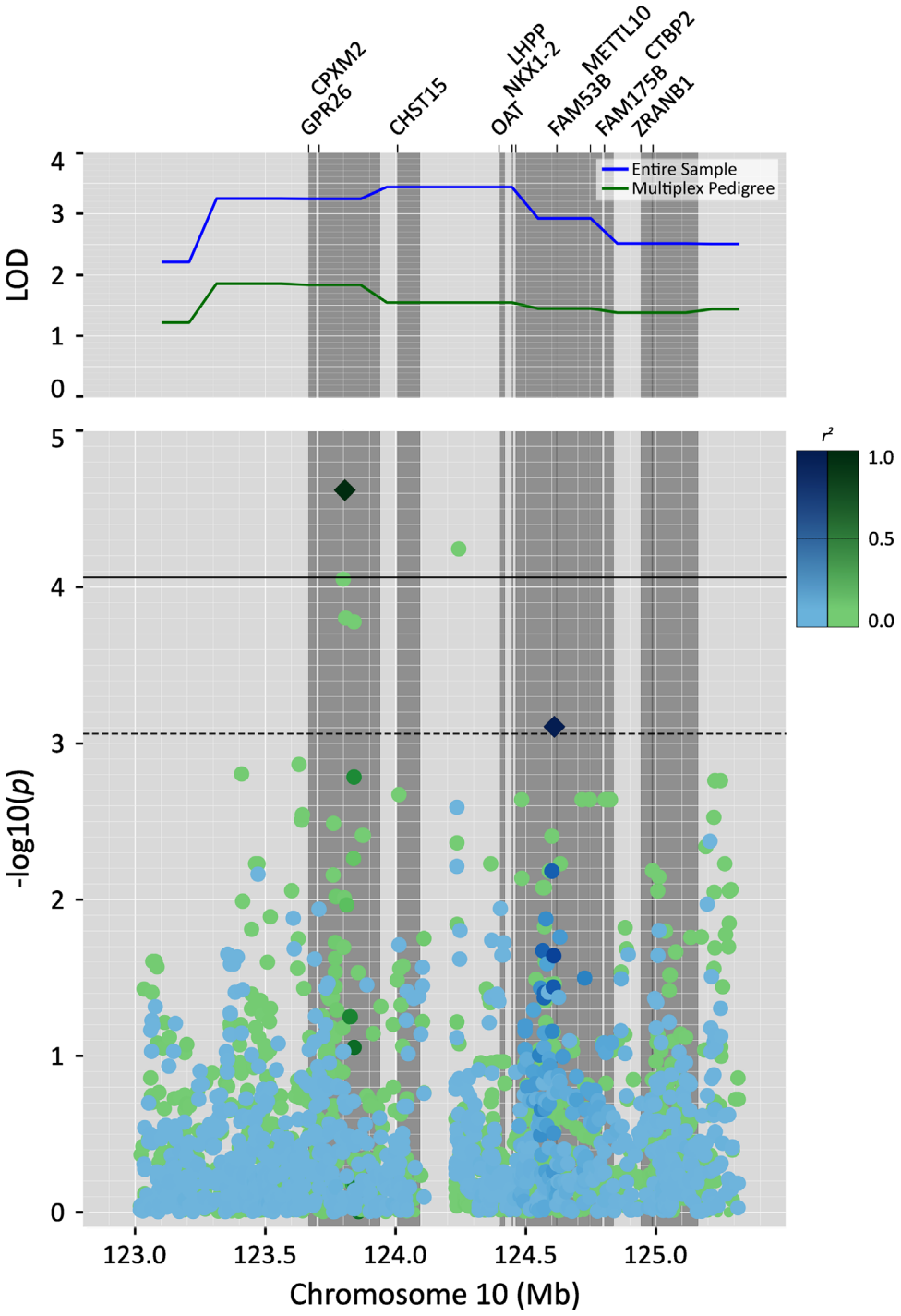

Fig. 3. QTL-specific association analysis within the QTL on chromosome 10 for the continuous depression factor score in the entire sample (blue) and the multiplex MDD pedigree (green). The top plot shows the linkage signal in the entire sample and the multiplex pedigree. The plot below shows the results of association analysis in the same region. Intergenic regions are pale gray and genes are represented by the dark gray bars. The top ranked variant in each subject group is represented by a diamond and the degree of linkage disequilibrium is represented by the color scale. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(rs11245316, $\chi^2=11.28$, $p=7.8\times10^{-04}$) and located within an intron of the gene *LHPP* (phospholysine phosphohistidine inorganic pyrophosphate phosphatase). Association analysis run only in the multiplex MDD pedigree from which the majority of the linkage signal originated revealed a variant that met peak-wide significance (rs7913161, $\chi^2=17.84$, $p=2.4\times10^{-05}$) within an intron of the gene *CPXM2* (carboxypeptidase X (M14 family), member 2). When the SNP rs7913161 was included as a covariate in the linkage analysis of the continuous depression factor score in the multiplex pedigree, the LOD score observed without the covariate (LOD=1.84) was reduced substantially (LOD=0.46). This linkage conditional on association test gives additional support for the involvement of rs7913161 in depression risk within the multiplex MDD as it implicates the variant in the linkage model for MDD within the pedigree. Association for this variant was not significant in either the entire sample ($\chi^2=1.68$, $p=0.19$) or in any other individual pedigree (the next best association for rs7913161 in any other pedigree did not reach suggestive significance, $\chi^2=3.97$,

$p=0.04$), suggesting that the variant is likely marking a functional and rare variant present only in the multiplex pedigree.

3.4. Validation of the continuous factor score measure of depression

The factor score derived from the one-factor model of depression correlates highly with the analogous dichotomous diagnosis of depression (indicating the presence or absence of a depressive episode over the lifetime of an individual), derived from the same section of the MINI ($r_{phenotypic}=0.87$ ($p=1.22\times10^{-217}$), $r_{genetic}=1.00$ ($p=2.46\times10^{-06}$)). This high correlation is unsurprising given that the two traits are derived from precisely the same items. However, to provide further validation of the factor model, we ran linkage analysis for the dichotomous diagnostic trait within the region of chromosome 10 where we observed a genome-wide significant peak for the continuous factor score. For this analysis, the dichotomous diagnostic trait was transformed into a normally-distributed liability trait based on disease prevalence following standard quantitative genetic practice (Falconer, 1996) (pp. 299–309). The analogous liability measure from the same items in the entire sample did not exhibit genome-wide significant linkage but neared suggestive significance ($h^2=0.38$, $p=4.7\times10^{-06}$, LOD=1.58 at 153 cM; Fig. 2). Moreover, association analysis with the top ranked variant (for the continuous factor score) and the dichotomous measure exhibited some signal without reaching suggestive significance ($\chi^2=3.93$, $p=4.7\times10^{-02}$), while the top-ranked variant (in the multiplex pedigree) also reached significance for the dichotomous measure of depression ($\chi^2=16.23$, $p=5.6\times10^{-05}$). Thus, while the continuous measure of depression derived from the single-factor model overlaps almost completely with the dichotomous measure derived from the same items; the use of a continuous measure was shown to be more successful for gene-finding efforts.

It is of note that the distribution of our continuous factor score is bimodal which is in line with a number of unaffected individuals within the data. Indeed, the kurtosis score (-0.5887) indicates a platykurtic distribution. However, previous work indicates that positive kurtosis, not negative, may inflate the false-positive rate for linkage (Blangero et al., 2001; Blangero et al., 2007). Nonetheless, we ran an emprical LOD adjustment routine in SOLAR which calculates an adjustment factor by which to multiply the peak LOD from linkage analysis; 10,000 simulations run on the inverse normalized factor score calculated an adjustment factor of 1.08 where an adjustment factor >1 means that our LOD of 3.43 is in fact somewhat conservative. Thus, we are satisfied that our trait, despite being negatively skewed, has not resulted in an inflated LOD score in the present paper.

4. Discussion

Repeated attempts to identify genetic influences on MDD using genome-wide association have been met with difficulty. Conversely, several genome-wide significant loci have been identified using linkage analysis (Breen et al., 2011; Pergadia et al., 2011), including in the same region of chromosome 10, and more specifically in the same gene *LHPP*, as identified here (Neff et al., 2009). Moreover, a recent whole genome sequence study has also highlighted the role of the gene *LHPP* in MDD risk (CONVERGE consortium, 2015). The present study extends the previous literature by supporting the role of 10q26.13 (and possibly *LHPP*) in risk for MDD. Moreover, the present study identifies a novel and interesting gene, *CPXM2*, in a newly identified large multiplex MDD pedigree from whom the majority of the linkage signal originates–though this finding needs replication before it can be considered a risk gene for MDD. The linkage conditional on association test for the variant within CPXM2 suggests that it is partially responsible (either directly or via LD with another variant) for the linkage signal within the multiplex pedigree (Biernacka and Cordell, 2007). Indeed, it is likely that the variant identified in the pedigree is marking a functional and rare variant that exists only in this pedigree. Thus, the present study, through a combination of linkage in extended pedigrees and a dimensional index of depression, highlights a two interesting genes for MDD risk and, potentially, the role of rare variation in risk for the illness.

LHPP encodes the protein phospholysine phosphohistidine inorganic pyrophosphate phosphatase (Lhpp) (Seal and Binkley, 1957) and is highly expressed in brain (Neff et al., 2009; Yokoi et al., 2003). Neff and colleagues have previously implicated the gene LHPP in MDD risk using a combination of linkage and association analysis (Neff et al., 2009). However, the LHPP associations were dependent on *HTR1A* genotype, which is not a finding that we were able to replicate in the present paper. First, many of the *LHPP* variants identified by Neff and colleagues as being associated with MDD risk are not present in our sample, although those that are present are in partial LD with our top-ranked variant (rs12265012, $r^2=0.21$; rs10794134, $r^2=0.17$) (Neff et al., 2009). Second, while Neff and colleagues showed an interaction between HTR1A (and specifically the 1019 C > G genotype, rs2495, which is not present in out sample) and *LHPP*, we did not. Three *HTR1A* variants are present in our sample, and an interaction term between our top-ranked *LHPP* variant and any of the *HTR1A* variants was not significant when included as covariates in a polygenic model of the depression factor score (rs10052087, $\chi^2=0.75$, $p=0.39$; rs6449693, $\chi^2=0.01$, $p=0.93$; rs6294, $\chi^2=0.75$, $p=0.39$); and none of the *HTR1A* variants were significantly associated with the continuous depression score in isolation (rs10052087, $\chi^2=0.46$, $p=0.50$; rs6449693, $\chi^2=2.16$, $p=0.14$; rs6294, $\chi^2=0.46$, $p=0.50$). Relatively little is known about the function of *LHPP*. A single study implicates the role of Lhpp in thyroid function (Koike et al., 2006), which could be interesting given that thyroid function is thought to mediate the function of certain anti-depressants (Altshuler et al., 2001). However, an explicit relationship between thyroid function, Lhpp and MDD is not apparent based on current research.

The gene *CPXM2* is a member of the metallocarboxypeptidase A family of digestive enzymes and is highly expressed in brain, particularly in the hippocampus, hypothalamus, choroid plexus and throughout the cerebral cortex (Xin et al., 1998). *CPXM2* is distinct from other gene-family members as it lacks the active site residues necessary for enzyme function and as a consequence it may fulfill an alternative role as a phospholipid binding protein (Xin et al., 1997). In rats *CPE*, a paralog of *CPXM1* (which is highly similar to *CPXM2* in that it also lacks the catalytic activity found in other carboxypetidases (Lei et al., 1999)), mediates dopamine transporter (DAT) expression such that co-expression of *CPE* and *DAT* results in increased dopamine reuptake in brain (Zhang et al., 2009). Also, a variant of *CPXM2* is suggestively associated with cognitive decline in schizophrenia (Hashimoto et al., 2013), where cognitive ability, and more specifically cognitive impairment in schizophrenia, is thought to be modulated by dopaminergic signaling (Nieoullon, 2002; Goldman-Rakic et al., 2004; Knowles et al., 2014). Insofar as the role of dopamine is well established in MDD (Nestler and Carlezon, 2006; Dunlop and Nemeroff, 2007; Tye et al., 2013) and that the gene *CPXM2*, or at least very similar genes in the same family, appears to influence dopamine functioning in the brain, the present paper highlights a new candidate gene for MDD in a newly established large multiplex MDD pedigree that warrants further investigation.

The association for the variant rs7913161 in *CPXM2* in the larger sample was low and not significant, which could suggests that the variant is marking a rare and functional variant that exists only within the multiplex MDD pedigree, one which likely makes up a haplotype of many variants. The authors of the most recent mega-

analysis GWA study (GWAS) from the PGC, which comprised approximately seventy thousand subjects, cite the need for even greater sample sizes and increased power in order to detect genetic variants for depression (Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium et al., 2013). However, it is also possible that genetic heterogeneity for MDD is greater than for other disorders, meaning that it will be necessary to reduce genetic heterogeneity (for example, by studying a group of genetically-homogenous kindred) in order to find risk genes. The present study makes a case for the latter approach; namely, the use of whole genome sequence data in extended pedigrees. The common disease-rare variant hypothesis states that the genetic causes of common, polygenic disorders such as depression are likely to be rare in the population. Clearly the use of rare variation alone will not solve the power issues highlighted by the PGC. However, identifying a rare functional variant (with a large effect size) in only a handful of affected individuals can be sufficient to verify that a given gene is involved in an illness. Data from the 1000 Genomes Project confirm that rare ($<1\%$) variants constitute the vast majority (73%) of polymorphic sites in humans (Marth et al., 2011). A key factor for identification of specific rare functional variants is detecting sufficient copies of that variant for statistical inference (Kent et al., 2007; Blangero et al., 2009). Pedigree-based studies represent an implicit enrichment strategy for identifying rare variants as Mendelian transmissions from parents to offspring maximize the chance that multiple copies of rare variants exist in the pedigree. Family-based cohorts have substantially greater power than unrelated cases to detect rare genetic effects given an equivalent number of sampled individuals (Li et al., 2006; Saad and Wijsman, 2014). For example, genes for hypertension have been identified for blood pressure in the general population by focusing research efforts on an extended pedigree with a rare form of hypertension (Tobin et al., 2008). Indeed, rare deleterious mutations are known to occur in genes that also harbor common variants with modest effects on disease risk (Bodmer and Bonilla, 2008). For example, 11 of 30 genes with common variants associated with lipid levels also carry known rare alleles of large effect in Mendelian dyslipidemias (Cohen et al., 2006; Romeo et al., 2007). Furthermore, rare variants may contribute to loci identified through common variation (Ji et al., 2008).

Another issue the PGC highlights in the hunt for depression genes is the possibility that the depression phenotypes used in genetic studies are 'suboptimal' (Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium et al., 2013) (p. 9). This observation dovetails with the Research Domain Criteria (RDoC) strategy that was recently proposed by the NIMH. This strategy encourages researchers to focus their efforts on developing new ways of classifying psychopathology by developing a dimension-based taxonomy of functioning that encompasses behavior, neuroscience and genetics (Cuthbert and Insel, 2010; Insel et al., 2010). MDD is typically treated as a categorical trait it is assumed that MDD reflects the tail end of an underlying normal distribution of mood, and that diagnosis occurs when a threshold for liability is crossed. It seems plausible that the genes which moderate behavior at the tail end of the distribution are the same as those that underlie the regulation of normal mood (Luft, 2002) and by dichotomizing the MDD distribution, we ignore a substantial proportion of variance that would contribute to gene-finding efforts. Thus the present paper is in line with the RDoC strategy whereby depression is represented as a continuum or dimension and moreover the use of a continuous measure of depression derived from a single-factor model of interview items versus the analogous dichotomous measure from the same items was shown to be more successful in the present study. The notion that continuous models of complex disorders derived from commonly used questionnaires can be used in association studies represents a significant advancement over studies that rely entirely on a diagnostic endpoint. Other studies have examined this notion in detail, and developed multidimensional models of depressive symptomatology (Uher et al., 2008; Korszun et al., 2004; Kendler et al., 2013). The present work focussed specifically on the MINI and as such fewer dimensions were derived (indeed, inspection of the correlations between items strongly supports the existence of a single dimension in the data used in the present study; Table S1), however the utility of the present study over those published previously is the inclusion of genetic data which allowed the identification of possible candidate genes (*LHPP* and *CPXM2*).

In summary, the present study represents advancement in our understanding in the genetics of depression in two ways. First, it confirms the probable involvement of a gene previously implicated in illness risk (*LHPP*) and, through the use of a multiplex MDD pedigree, it highlights a novel risk gene (*CPXM2*), which warrants further investigation. Second, it draws attention to an alternative methodology for the hunt for depression genes, which is focusing on rare variation in a multiplex MDD pedigree combined with the use of dimensional indices of MDD symptomatology.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jad.2015.11.012.

References

Abecasis, G.R., Cherny, S.S., Cookson, W.O., Cardon, L.R., 2002. Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. Nat. Genet. 30, 97–101.

American Psychiatric Association, 1994. Diagnostic and statistical manual of mental disorders: DSM-IV. American Psychiatric Association, Washington, D.C.

Almasy, L., Blangero, J., 1998. Multipoint quantitative-trait linkage analysis in general pedigrees. Am. J. Hum. Genet. 62, 1198–1211.

Altshuler, L.L., Bauer, M., Frye, M.A., Gitlin, M.J., Mintz, J., Szuba, M.P., et al., 2001. Does thyroid supplementation accelerate tricyclic antidepressant response? A review and meta-analysis of the literature. Am. J. Psychiatry 158, 1617–1622.

Binder, E.B., Holsboer, F., 2005. Pharmacogenomics. Handb. Exp. Pharmacol. 169, 527–546.

Bromet, E.J., Dunn, L.O., Connell, M.M., Dew, M.A., Schulberg, H.C., 1986. Long-term reliability of diagnosing lifetime major depression in a community sample. Arch. Gen. Psychiatry 43, 435–440.

Bosker, F.J., Hartman, C.A., Nolte, I.M., Prins, B.P., Terpstra, P., Posthuma, D., et al., 2011. Poor replication of candidate genes for major depressive disorder using genome-wide association data. Mol. Psychiatry 16, 516–532.

Breen, G., Webb, B.T., Butler, A.W., van den Oord, E.J., Tozzi, F., Craddock, N., et al., 2011. A genome-wide significant linkage for severe depression on chromosome 3: the depression network study. Am. J. Psychiatry 168, 840–847.

Blangero, J., Williams, J.T., Almasy, L., 2001. Variance component methods for detecting complex trait loci. In: Rao, D.C., Province, M.A. (Eds.), Genetic Dissection of Complex Traits. Academic Press, San Diego, pp. 151–181.

Blangero, J., Williams, J.T., Almasy, L., Williams-Blangero, S., 2007. Mapping genes influencing human quantitative trait variation. In: Crawford, Michael H. (Ed.), Anthropological Genetics: Theory, Methods and Applications. Cambridge University Press, New York, NY, pp. 306–333, Chapter 11.

Biernacka, J.M., Cordell, H.J., 2007. Exploring causality via identification of SNPs or haplotypes responsible for a linkage signal. Genet. Epidemiol. 31, 727–740.

Blangero, J., Goring, H.H., Kent, J.W., Williams, J.T., Peterson, C.P., Almasy, L., et al., 2009. Quantitative trait nucleotide analysis using bayesian model selection. 2005. Hum. Biol. 81, 829–847.

Bodmer, W., Bonilla, C., 2008. Common and rare variants in multifactorial susceptibility to common diseases. Nat. Genet. 40, 695–701.

Cohen-Woods, S., Craig, I.W., McGuffin, P., 2013. The current state of play on the molecular genetics of depression. Psychol. Med. 43, 673–687.

Collins, A.L., Sullivan, P.F., 2013. Genome-wide association studies in psychiatry: What have we learned? Br. J. Psychiatry 202, 1–4.

CONVERGE consortium, 2015. Sparse whole-genome sequencing identifies two loci for major depressive disorder. Nature 523, 588–591.

Cohen, J.C., Pertsemlidis, A., Fahmi, S., Esmail, S., Vega, G.L., Grundy, S.M., et al., 2006. Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels. Proc. Natl. Acad. Sci. USA 103, 1810–1815.

Cuthbert, B.N., Insel, T.R., 2010. Toward new approaches to psychotic disorders: the NIMH research domain criteria project. Schizophr. Bull. 36, 1061–1062.
Duggirala, R., Williams, J.T., Williams-Blangero, S., Blangero, J., 1997. A variance component approach to dichotomous trait linkage analysis using a threshold model. Genet. Epidemiol. 14, 987–992.
Dunlop, B.W., Nemeroff, C.B., 2007. The role of dopamine in the pathophysiology of depression. Arch. Gen. Psychiatry 64, 327–337.
Flint, J., Kendler, K.S., 2014. The genetics of major depression. Neuron 81, 484–503.
Feingold, E., Brown, P.O., Siegmund, D., 1993. Gaussian models for genetic linkage analysis using complete high-resolution maps of identity by descent. Am. J. Hum. Genet. 53, 234–251.
Falconer, D.S., 1996. Introduction to quantitative genetics. Longman, Harlow.
Greenberg, P.E., Kessler, R.C., Birnbaum, H.G., Leong, S.A., Lowe, S.W., Berglund, P.A., et al., 2003. The economic burden of depression in the united states: How did it change between 1990 and 2000? J. Clin. Psychiatry 64, 1465–1475.
Glahn, D.C., Curran, J.E., Winkler, A.M., Carless, M.A., Kent Jr, J.W., Charlesworth, J.C., et al., 2012. High dimensional endophenotype ranking in the search for major depression risk genes. Biol. Psychiatry 71, 6–14.
Goldman-Rakic, P.S., Castner, S.A., Svensson, T.H., Siever, L.J., Williams, G.V., 2004. Targeting the dopamine D1 receptor in schizophrenia: insights for cognitive dysfunction. Psychopharmacology 174, 3–16.
Hays, R.D., Wells, K.B., Sherbourne, C.D., Rogers, W., Spritzer, K., 1995. Functioning and well-being outcomes of patients with depression compared with chronic general medical illnesses. Arch. Gen. Psychiatry 52, 11–19.
Hek, K., Demirkan, A., Lahti, J., Terracciano, A., Teumer, A., Cornelis, M.C., et al., 2013. A genome-wide association study of depressive symptoms. Biol. Psychiatry 73, 667–678.
Hashimoto, R., Ikeda, M., Ohi, K., Yasuda, Y., Yamamori, H., Fukumoto, M., et al., 2013. Genome-wide association study of cognitive decline in schizophrenia. Am. J. Psychiatry 170, 683–684.
Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D.S., Quinn, K., et al., 2010. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am. J. Psychiatry 167, 748–751.
Ji, W., Foo, J.N., O'Roak, B.J., Zhao, H., Larson, M.G., Simon, D.B., et al., 2008. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nat. Genet. 40, 592–599.
Kessler, R.C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K.R., et al., 2003. The epidemiology of major depressive disorder: results from the national comorbidity survey replication (NCS-R). JAMA 289, 3095–3105.
Kendler, K.S., Neale, M.C., Kessler, R.C., Heath, A.C., Eaves, L.J., 1993. The lifetime history of major depression in women. reliability of diagnosis and heritability. Arch. Gen. Psychiatry 50, 863–870.
Keller, M.B., Hanks, D.L., Klein, D.N., 1996. Summary of the DSM-IV mood disorders field trial and issue overview. Psychiatr. Clin. N. Am. 19, 1–28.
Kendler, K.S., Gatz, M., Gardner, C.O., Pedersen, N.L., 2006. A swedish national twin study of lifetime major depression. Am. J. Psychiatry 163, 109–114.
Kohli, M.A., Lucae, S., Saemann, P.G., Schmidt, M.V., Demirkan, A., Hek, K., et al., 2011. The neuronal transporter gene SLC6A15 confers risk to major depression. Neuron 70, 252–265.
Koike, E., Toda, S., Yokoi, F., Izuhara, K., Koike, N., Itoh, K., et al., 2006. Expression of new human inorganic pyrophosphatase in thyroid diseases: its intimate association with hyperthyroidism. Biochem. Biophys. Res. Commun. 341, 691–696.
Knowles, E.E.M., Mathias, S.R., McKay, D.R., Sprooten, E., Blangero, J., Almasy, L., et al., 2014. Genome-wide analyses of working memory: A review. Current Behavioral Neuroscience Reports. 2014[Journal, Electronic].
Kent Jr, J.W., Dyer, T.D., Goring, H.H., Blangero, J., 2007. Type I error rates in association versus joint linkage/association tests in related individuals. Genet. Epidemiol. 31, 173–177.
Korszun, A., Moskvina, V., Brewster, S., Craddock, N., Ferrero, F., Gill, M., et al., 2004. Familiality of symptom dimensions in depression. Arch. Gen. Psychiatry 61, 468–474.
Kendler, K.S., Aggen, S.H., Neale, M.C., 2013. Evidence for multiple genetic factors underlying DSM-IV criteria for major depression. JAMA Psychiatry 70, 599–607.
Lopez, M.F., Compton, W.M., Grant, B.F., Breiling, J.P., 2007. Dimensional approaches in diagnostic classification: a critical appraisal. Int. J. Methods Psychiatr. Res. 16 (Suppl. 1), S6–S7.
Lewis, C.M., Ng, M.Y., Butler, A.W., Cohen-Woods, S., Uher, R., Pirlo, K., et al., 2010. Genome-wide association study of major recurrent depression in the U.K. population. Am. J. Psychiatry 167, 949–957.
Luft, F.C., 2002. Hypertension as a complex genetic trait. Semin. Nephrol. 22, 115–126.
Lei, Y., Xin, X., Morgan, D., Pintar, J.E., Fricker, L.D., 1999. Identification of mouse CPX-1, a novel member of the metallocarboxypeptidase gene family with highest similarity to CPX-2. DNA Cell Biol. 18, 175–185.
Li, M., Boehnke, M., Abecasis, G.R., 2006. Efficient study designs for test of genetic association using sibship data and unrelated cases and controls. Am. J. Hum. Genet. 78, 778–792.
Miller, D.B., O'Callaghan, J.P., 2013. Personalized medicine in major depressive disorder – opportunities and pitfalls. Metabolism 62 (Suppl. 1), S34–S39.
Muglia, P., Tozzi, F., Galwey, N.W., Francks, C., Upmanyu, R., Kong, X.Q., et al., 2010. Genome-wide association study of recurrent major depressive disorder in two european case-control cohorts. Mol. Psychiatry 15, 589–601.
Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, 2012. A mega-analysis of genome-wide association studies for major depressive disorder. Mol. Psychiatry 18, 497–511.
Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke, S., Wray, N.R., Lewis, C.M., Hamilton, S.P., Weissman, M.M., et al., 2013. A mega-analysis of genome-wide association studies for major depressive disorder. Mol. Psychiatry 18, 497–511.
Moskvina, V., Schmidt, K.M., 2008. On multiple-testing correction in genome-wide association studies. Genet. Epidemiol. 32, 567–573.
Marth, G.T., Yu, F., Indap, A.R., Garimella, K., Gravel, S., Leong, W.F., et al., 2011. The functional spectrum of low-frequency coding variation. Genome Biol. 12, R84 2011-12-9-r84.
Neff, C.D., Abkevich, V., Packer, J.C., Chen, Y., Potter, J., Riley, R., et al., 2009. Evidence for HTR1A and LHPP as interacting genetic risk factors in major depression. Mol. Psychiatry 14, 621–630.
Nieoullon, A., 2002. Dopamine and the regulation of cognition and attention. Prog. Neurobiol. 67, 53–83.
Nestler, E.J., Carlezon Jr., W.A., 2006. The mesolimbic dopamine reward circuit in depression. Biol. Psychiatry 59, 1151–1159.
Olvera, R.L., Bearden, C.E., Velligan, D.I., Almasy, L., Carless, M.A., Curran, J.E., Williamson, D.E., Duggirala, R., Blangero, J., Glahn, D.C., 2011. Common genetic influences on depression, alcohol, and substance use disorders in mexican-american families. Am. J. Med. Genet. Part B Neuropsychiatr. Genet. 156B, 561–568.
Pergadia, M.L., Glowinski, A.L., Wray, N.R., Agrawal, A., Saccone, S.F., Loukola, A., et al., 2011. A 3p26-3p25 genetic linkage finding for DSM-IV major depression in heavy smoking families. Am. J. Psychiatry 168, 848–852.
Rush, A.J., Trivedi, M.H., Wisniewski, S.R., Nierenberg, A.A., Stewart, J.W., Warden, D., et al., 2006. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am. J. Psychiatry 163, 1905–1917.
Rietschel, M., Mattheisen, M., Frank, J., Treutlein, J., Degenhardt, F., Breuer, R., et al., 2010. Genome-wide association-, replication-, and neuroimaging study implicates HOMER1 in the etiology of major depression. Biol. Psychiatry 68, 578–585.
Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J.L., Kahler, A.K., Akterin, S., et al., 2013. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat. Genet. 45, 1150–1159.
Romeo, S., Pennacchio, L.A., Fu, Y., Boerwinkle, E., Tybjaerg-Hansen, A., Hobbs, H.H., et al., 2007. Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. Nat. Genet. 39, 513–516.
Sobel, E., Lange, K., 1996. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am. J. Hum. Genet. 58, 1323–1337.
Sullivan, P.F., Neale, M.C., Kendler, K.S., 2000. Genetic epidemiology of major depression: Review and meta-analysis. Am. J. Psychiatry 157, 1552–1562.
Sullivan, P.F., de Geus, E.J., Willemsen, G., James, M.R., Smit, J.H., Zandbelt, T., et al., 2009. Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. Mol. Psychiatry 14, 359–375.
Shyn, S.I., Shi, J., Kraft, J.B., Potash, J.B., Knowles, J.A., Weissman, M.M., et al., 2011. Novel loci for major depression identified by genome-wide association study of sequenced treatment alternatives to relieve depression and meta-analysis of three studies. Mol. Psychiatry 16, 202–215.
Shi, J., Potash, J.B., Knowles, J.A., Weissman, M.M., Coryell, W., Scheftner, W.A., et al., 2011. Genome-wide association study of recurrent early-onset major depressive disorder. Mol. Psychiatry 16, 193–201.
Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., et al., 1998. The mini-international neuropsychiatric interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatry 59 (Suppl. 20), S22–S33, quiz 34-57.
SEAL, U.S., BINKLEY, F., 1957. An inorganic pyrophosphatase of swine brain. J. Biol. Chem. 228, 193–199.
Saad, M., Wijsman, E.M., 2014. Power of family-based association designs to detect rare variants in large pedigrees using imputed genotypes. Genet. Epidemiol. 38, 1–9.
Tye, K.M., Mirzabekov, J.J., Warden, M.R., Ferenczi, E.A., Tsai, H.C., Finkelstein, J., et al., 2013. Dopamine neurons modulate neural encoding and expression of depression-related behaviour. Nature 493, 537–541.
Tobin, M.D., Tomaszewski, M., Braund, P.S., Hajat, C., Raleigh, S.M., Palmer, T.M., et al., 2008. Common variants in genes underlying monogenic hypertension and hypotension and blood pressure in the general population. Hypertension 51, 1658–1664.
U.S. Department of Health and Human Services, 1993. Public Health Service Agency for Healthcare Policy and Research. Depression guideline panel. depression in primary care: Volume 2. Treatment of major depression. clinical practice guideline, number 5. rockville, MD. AHCPR publication no. 93-0551. April 1993. 1993[Laws/Statutes].
Uher, R., Farmer, A., Maier, W., Rietschel, M., Hauser, J., Marusic, A., et al., 2008. Measuring depression: Comparison and integration of three scales in the GENDEP study. Psychol. Med. 38, 289–300.
Wells, K.B., Stewart, A., Hays, R.D., Burnam, M.A., Rogers, W., Daniels, M., et al., 1989. The functioning and well-being of depressed patients. results from the medical outcomes study. JAMA 262, 914–919.
World Health Organization, 2012. Depression fact sheet number 369. 2012 (11/06)[Web Page].
Wray, N.R., Pergadia, M.L., Blackwood, D.H., Penninx, B.W., Gordon, S.D., Nyholt, D. R., et al., 2012. Genome-wide association study of major depressive disorder: New results, meta-analysis, and lessons learned. Mol. Psychiatry 17, 36–48.
Xin, X., Day, R., Dong, W., Lei, Y., Fricker, L.D., 1998. Identification of mouse CPX-2, a novel member of the metallocarboxypeptidase gene family: CDNA cloning,

mRNA distribution, and protein expression and characterization. DNA Cell Biol. 17, 897–909.
Xin, X., Varlamov, O., Day, R., Dong, W., Bridgett, M.M., Leiter, E.H., et al., 1997. Cloning and sequence analysis of cDNA encoding rat carboxypeptidase D. DNA Cell Biol. 16, 897–909.
Yokoi, F., Hiraishi, H., Izuhara, K., 2003. Molecular cloning of a cDNA for the human phospholysine phosphohistidine inorganic pyrophosphate phosphatase. J. Biochem. 133, 607–614.
Zhang, H., Li, S., Wang, M., Vukusic, B., Pristupa, Z.B., Liu, F., 2009. Regulation of dopamine transporter activity by carboxypeptidase E. Mol. Brain 2 (10), 6606-2–6606-10.