EBioMedicine xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Paper

Shared Genetics and Couple-Associated Environment Are Major Contributors to the Risk of Both Clinical and Self-Declared Depression

Yanni Zeng, MSc [a,*], Pau Navarro, PhD [b], Charley Xia, MSc [b], Carmen Amador, PhD [b], Ana M. Fernandez-Pujals, MSc [a], Pippa A. Thomson, PhD [d,e], Archie Campbell, MA [c], Reka Nagy, MSc [b], Toni-Kim Clarke, PhD [a], Jonathan D. Hafferty, MD [a], Blair H. Smith, MD [c,f], Lynne J. Hocking, PhD [c,g], Sandosh Padmanabhan, PhD [c,h], Caroline Hayward, PhD [b], Donald J. MacIntyre, MD [a], David J Porteous, PhD [c,d,e], Chris S. Haley, PhD [b,i], Andrew M. McIntosh, MD [a,c,d]

[a] Division of Psychiatry, University of Edinburgh, Edinburgh, UK
[b] MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
[c] Generation Scotland, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK
[d] Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK
[e] Medical Genetics Section, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
[f] Division of Population Health Sciences, University of Dundee, Dundee, UK
[g] Division of Applied Health Sciences, University of Aberdeen, Aberdeen, UK
[h] Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK
[i] The Roslin Institute and Royal (Dick) School of Veterinary Sciences, University of Edinburgh, UK

ARTICLE INFO

Article history:
Received 12 September 2016
Received in revised form 1 November 2016
Accepted 2 November 2016
Available online xxxx

Keywords:
Major depressive disorder
Self-declared depression
SNP heritability
Couple effect
Family environment
Linear mixed modeling

ABSTRACT

Background: Both genetic and environmental factors contribute to risk of depression, but estimates of their relative contributions are limited. Commonalities between clinically-assessed major depressive disorder (MDD) and self-declared depression (SDD) are also unclear.
Methods: Using data from a large Scottish family-based cohort (GS:SFHS, N = 19,994), we estimated the genetic and environmental variance components for MDD and SDD. The components representing the genetic effect associated with genome-wide common genetic variants (SNP heritability), the additional pedigree-associated genetic effect and non-genetic effects associated with common environments were estimated in a linear mixed model (LMM).
Findings: Both MDD and SDD had significant contributions from components representing the effect from common genetic variants, the additional genetic effect associated with the pedigree and the common environmental effect shared by couples. The estimate of correlation between SDD and MDD was high (r = 1.00, se = 0.20) for common-variant-associated genetic effect and lower for the additional genetic effect from the pedigree (r = 0.57, se = 0.08) and the couple-shared environmental effect (r = 0.53, se = 0.22).
Interpretation: Both genetics and couple-shared environmental effects were major factors influencing liability to depression. SDD may provide a scalable alternative to MDD in studies seeking to identify common risk variants. Rarer variants and environmental effects may however differ substantially according to different definitions of depression.

1. Introduction

Depression has a pattern of familial aggregation, which implies the influence of genetic effects, common environmental effects shared by relatives, or both. The genetic component (heritability) has been estimated by a twin study of major depressive disorder (MDD) to be 37% (Sullivan et al., 2000). The SNP heritability (heritability attributed to common genetic variants) of MDD varies across populations and samples (21%–32%) (Lubke et al., 2012; Lee et al., 2013). Subsequently, a 'children of twins' study found a significantly greater risk of depression in children of depressed monozygotic (MZ) twins than in the offspring of their non-depressed twin. This implies a potential environmental effect of parental depression on offspring (Singh et al., 2011). Studies have also shown that having a partner with psychiatric disorder may increase an individual's risk of MDD (Joutsenniemi et al., 2011; Desai et al., 2012), but meta-analytic studies suggest no effect of the shared sibling

* Corresponding author at: Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK, EH10 5HF.
E-mail address: Y.Zeng-6@sms.ed.ac.uk (Y. Zeng).

http://dx.doi.org/10.1016/j.ebiom.2016.11.003

environment and other studies have postulated more complex relationships (Olino et al., 2006). Whilst each of these studies separately provided evidence for the genetic and familial environmental components in depression, a precise separation of these potential effects should involve estimating them simultaneously in the same model and has yet to be achieved.

The accurate separation and estimation of the genetic and environment components on liability to depression provide crucial information, as it reveals the upper limit of the genetic effects, the probability of true positive results from genetic studies and the potential for accurate risk predictions for depression (Makowsky et al., 2011; Tenesa and Haley, 2013). Genetic studies attempting to map causal variants have been performed for various definitions of depression. These include clinically-assessed depression, self-report of clinical diagnosis of depression and self-reported depressive symptoms (consortium, 2015; Major Depressive Disorder Working Group of the Psychiatric et al., 2013; Hyde et al., 2016; Okbay et al., 2016), but the findings are generally inconsistent. Although some of the inconsistent findings were probably due to the limited power of the original studies (Flint and Kendler, 2014), there may also be intrinsic heterogeneity across depression definitions. This is further supported by the fact that studies show very different estimates of the narrow-sense heritability (h_n^2) for several depression definitions or related traits to MDD ($h_n^2 = 37\%$ (Sullivan et al., 2000)): perceived stress: 44% (Bogdan and Pizzagalli, 2009); nine depression definitions in women: 21%–45% (Kendler et al., 1992); depressive symptom scores in childhood: 79% (Thapar and Mcguffin, 1994) and depressive symptoms in an elderly population: 69% in women and 64% in men (McGue and Christensen, 2003). In fact, even for MDD, the genetic correlation of MDD phenotypes between independent datasets was relatively low compared with other psychiatric disorders (Gratten et al., 2014).

Because of the heterogeneity across depression definitions, there has been a long debate about the correct phenotype for depression genetic studies. Studies using clinically-assessed depression could provide findings that are directly informative for clinical application. However, the resources required for such data collection are generally very high (consortium, 2015). As an alternative, measuring self-reported depression requires fewer resources and this phenotype is rapidly becoming available for many population-based datasets (Okbay et al., 2016; Hyde et al., 2016). To date, the largest published GWAS of major depression has yielded 15 significant loci (7 loci before meta-analysis) for a self-reported clinical diagnosis (Hyde et al., 2016).

Given that important progress is being made from GWASs on different depression definitions, it becomes increasingly important to understand the similarities and dissimilarities across different definitions. Particularly, the difference in genetic and environmental loadings might underpin the inconsistent results from genetic studies across different depression definitions. Therefore, dissecting the phenotypic variance of each depression definition and understanding the similarities and dissimilarities between those phenotypes in the context of both genetic and common environmental components is particularly important for interpreting the results from published depression studies and for informing about genetically relevant depression phenotypes for future studies.

In this study we sought to partition the phenotypic variation of the diagnosed depression (MDD) and the self-declared depression (SDD) into its genetic and familial environment components using Linear Mixed Modeling (LMM). Specifically, we utilized data from Generation Scotland: Scottish Family Health Study (GS:SFHS), a large Scottish cohort with extensive family relationship information and genome-wide genotype data to answer two questions. First, when simultaneously considering multiple genetic effects and familial environmental effects in the model, what are major contributions to variation in MDD and SDD, respectively? Second, what is the contribution of each of the identified major contributing components to the overall correlation between MDD and SDD?

2. Materials and Methods

The Tayside Research Ethics Committee (reference 05/S1401/89) provided ethical approval for the study. In GS: SFHS, participants gave written consent, after having an opportunity to discuss the project, and before any data or samples were collected.

The details of their consent status are recorded in the study database. All consent forms and study protocols were approved by the Research Ethics Committee.

2.1. Datasets

Generation Scotland: Scottish Family Health Study (GS: SFHS) contains 21,387 subjects ($N_{male} = 8772$, $N_{female} = 12{,}615$; $Age_{mean} = 47.2$), who were recruited from the registers of collaborating general practices. At least one first-degree relative aged 18 or over was required to be identified for each participant (Smith et al., 2006). Genotyping data were generated using the Illumina Human OmniExpressExome -8- v1.0 array (Gunderson, 2009). Details of genotyping are described elsewhere (Smith et al., 2006). Population outlier individuals were removed from the sample (Amador et al., 2015). Quality control (QC) of genotyped SNPs used inclusion thresholds: missing SNPs per individual ≤ 2%, SNP genotype call rate ≥ 98%, minor allele frequency (MAF) > 1% and Hardy-Weinberg equilibrium P value $> 1 \times 10^{-6}$. In total, 561,125 genotyped autosomal SNPs passed QC criteria and were available for 19,994 participants ($N_{male} = 8221$, $N_{female} = 11{,}773$, $Age_{mean} = 47.4$).

2.1.1. Phenotypes

Lifetime Diagnosis of MDD: The Structured Clinical Interview for DSM-IV(SCID) was used (First et al., 2012): participants who screened positive (21.7%) for the questions “Have you ever seen anybody for emotional or psychiatric problems? IF YES: What was that for? (What treatment(s) did you get? Any medications?) IF NO: Was there ever a time when you, or someone else, thought you should see someone because of the way you were feeling or acting” were invited to continue to an interview using the SCID modules for mood disorders (First et al., 2002). Participants who screened positive but refused to undergo the structured clinical interview (N = 507, 2.4%) and those with a diagnosis of bipolar disorder (N = 76) were excluded from the study. More details of phenotyping procedures are described elsewhere (Fernandez-Pujals et al., 2015).

Self-declared depression (SDD): the participants were invited to answer the following question “please mark an X in the box if you have been affected by depression”.

2.2. Partitioning the Phenotypic Variation

Based on the framework of the Genomic-relationship-matrix restricted maximum likelihood (GREML) method, Xia et al. (2016) developed a method to estimate h_g^2 (proportion of additive genetic variance contributed by common genetic variants over the total phenotypic variance, namely SNP heritability), h_p^2 (representing the additional additive genetic effect contributed by pedigree associated variation), h_n^2 (proportion of the total additive genetic variance over the total phenotypic variance, namely narrow-sense heritability) and a number of familial environmental components simultaneously (Xia et al., 2016). This was performed by fitting variance-covariance matrices representing common genetic effects, pedigree-related-genetic effects, and current and past family environmental effects simultaneously in the mixed linear model(Xia et al., 2016), building on previous work by Zaitlen et al. (Zaitlen et al., 2013). This approach enables estimation of the contribution of each genetic and family environmental component and here we applied it to MDD and SDD.

In detail, for each trait, two genomic relationship matrices, $\boldsymbol{G}$ (genomic relationship matrix) and $\boldsymbol{K}$ (kinship matrix created by modifying $\boldsymbol{G}$

Please cite this article as: Zeng, Y., et al., Shared Genetics and Couple-Associated Environment Are Major Contributors to the Risk of Both Clinical and Self-Declared Depression, EBioMedicine (2016), http://dx.doi.org/10.1016/j.ebiom.2016.11.003