eNeuro

Research Article: New Research | Disorders of the Nervous System

Lingual Gyrus surface area is associated with anxiety-depression severity in young adults: a genetic clustering approach

Baptiste Couvy-Duchesne[1,2,3], Lachlan T. Strike[1], Greig I. de Zubicaray[4], Katie L. McMahon[3], Paul M. Thompson[5], Ian B. Hickie[6], Nicholas G. Martin[2] and Margaret J. Wright[1,3]

[1]*Queensland Brain Institute, University of Queensland, Brisbane, 4072, Australia*

[2]*QIMR Berghofer Medical Research Institute, Brisbane, 4006, Australia*

[3]*Centre for Advanced Imaging, University of Queensland, Brisbane, 4072, Australia*

[4]*Institute of Health Biomedical Innovations, Queensland Institute of Technology, Brisbane, 4006, Australia*

[5]*Imaging Genetics Center, Keck School of Medicine, University of Southern California, Marina del Rey, California 90292, USA*

[6]*Brain & Mind Research Institute, University of Sydney, Sydney, 2050, Australia*

DOI: 10.1523/ENEURO.0153-17.2017

Received: 8 May 2017

Revised: 21 December 2017

Accepted: 25 December 2017

Published: 11 January 2018

Author Contributions: BCD designed and performed the analyses. BCD and MJW wrote the manuscript. LTS processed the structural images and derived the genetic clusters. LTS, GIdZ, KLM, PMT, IBH, NGM edited the manuscript. GIdZ, KLM, PMT, NGM, MJW designed the QTIM study.

Funding: http://doi.org/10.13039/100000002HHS | National Institutes of Health (NIH)
R01 HD050735
U54 EB020403

Funding: http://doi.org/10.13039/501100000925Department of Health | National Health and Medical Research Council (NHMRC)
496682
1009064

Conflict of Interest: Authors report no conflict of interest.

The MRI imaging of QTIM was supported by grants from National Institute of Health (NIH) [R01 HD050735] and the National Health and Medical Research Council (NHMRC) [496682, 1009064]. Zygosity typing and collection of SPHERE data were funded by the Australian Research Council (ARC) [A7960034, A79906588, A79801419, DP0212016]. Genotyping was funded by the NHMRC [389891]. Replication data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centres that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Centre for Systems Neuroscience at Washington University.

BCD is supported by a UQ International scholarship and LS by an Australian Postgraduate Award scholarship. PT and MJW are supported in part by NIH grant U54 EB020403 to the ENIGMA Center for Worldwide Medicine, Imaging & Genomics.

Correspondence should be addressed to Baptiste Couvy-Duchesne, Queensland Brain Institute, the University of Queensland, (building 79), 4072 St Lucia, QLD, Australia. E-mail: b.couvyduchesne@uq.edu.au

Cite as: eNeuro 2018; 10.1523/ENEURO.0153-17.2017

Lingual Gyrus surface area is associated with anxiety-depression severity in young adults: a genetic clustering approach

Abbreviated title: Anxiety-depression and lingual gyrus surface area

Baptiste Couvy-Duchesne[1,2,3,*], Lachlan T. Strike[1], Greig I. de Zubicaray[4], Katie L. McMahon[3], Paul M. Thompson[5], Ian B. Hickie[6], Nicholas G. Martin[2], Margaret J. Wright[1,3]

1 Queensland Brain Institute, University of Queensland, 4072 Brisbane, Australia

2 QIMR Berghofer Medical Research Institute, 4006 Brisbane, Australia

3 Centre for Advanced Imaging, University of Queensland, 4072 Brisbane, Australia

4 Institute of Health Biomedical Innovations, Queensland Institute of Technology, 4006 Brisbane, Australia

5 Imaging Genetics Center, Keck School of Medicine, University of Southern California, Marina del Rey, California 90292, United States of America,

6 Brain & Mind Research Institute, University of Sydney, 2050 Sydney, Australia,

* Corresponding author

Author Contributions: BCD designed and performed the analyses. BCD and MJW wrote the manuscript. LTS processed the structural images and derived the genetic clusters. LTS, GIdZ, KLM, PMT, IBH, NGM edited the manuscript. GIdZ, KLM, PMT, NGM, MJW designed the QTIM study.

Correspondence should be addressed to Baptiste Couvy-Duchesne: Queensland Brain Institute, the University of Queensland, (building 79), 4072 St Lucia, QLD, Australia. b.couvyduchesne@uq.edu.au.

LTS: l.strike1@uq.edu.au, NGM: Nick.Martin@qimrberghofer.edu.au, KLM: katie.mcmahon@cai.uq.edu.au, GIdZ: greig.dezubicaray@qut.edu.au, IBH: ian.hickie@sydney.edu.au, PMT: pthomp@usc.edu, MJW: margie.wright@uq.edu.au

This manuscript consists of 38 pages, 4 figures, 3 tables.

The abstract is 224 words long, the significance statement 111, the introduction 597 and the discussion 1489 words.

Acknowledgments: We are very grateful to the twins for their generosity of time and willingness to participate in the Brisbane studies. We thank research nurses Marlene Grace and Ann Eldridge for twin recruitment, research assistants Kori Johnson, Aaron Quiggle, and Natalie Garden, and radiographers Matthew Meredith, Peter Hobden, Kate Borg, Aiman Al Najjar, and Anita Burns for MRI data acquisition, David Butler and Daniel Park for IT support. We also thank Anjali Henders, Lisa Bardsley, Zhen Zhen Zhao, Leanne Wallace the genotype data collection and QC, as well as Scott Gordon for all downstream processing, including further QC and imputation. Informed consent was obtained from all participants, including a parent or guardian for those aged less than 18 years. Participants received an honorarium for each study. The studies were approved by the Human Research Ethics Committees of the QIMR Berghofer Medical Research Institute, the University of Queensland and Uniting Health Care.

Conflict of Interest: The authors declare no competing financial interests.

Funding sources: The MRI imaging of QTIM was supported by grants from National Institute of Health (NIH) [R01 HD050735] and the National Health and Medical Research Council (NHMRC) [496682, 1009064]. Zygosity typing and collection of SPHERE data were funded by the Australian Research Council (ARC) [A7960034, A79906588, A79801419, DP0212016]. Genotyping was funded by the NHMRC [389891]. Replication data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centres that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Centre for Systems Neuroscience at Washington University.

BCD is supported by a UQ International scholarship and LS by an Australian Postgraduate Award scholarship. PT and MJW are supported in part by NIH grant U54 EB020403 to the ENIGMA Center for Worldwide Medicine, Imaging & Genomics.

Abstract

Here we aimed to identify cortical endophenotypes for anxiety/depression. Our data-driven approach used vertex-wise genetic correlations (estimated from a twin sample: 157 monozygotic and 194 dizygotic twin pairs) to parcellate cortical thickness and surface area into genetically homogeneous regions (Chen et al., 2013). In an overlapping twin and sibling sample (N=834; aged 15-29, 66% female), in those with anxiety-depression SPHERE scores (Hickie et al., 2001) above median we found a reduction of surface area in an occipito-temporal cluster, which comprised part of the right lingual, fusiform and parahippocampal gyrii. A similar reduction was observed in the Human Connectome Project sample (N=890, age 22-37, 56.5% female) in those with Adult Self Report DSM-oriented scores (Achenbach et al., 2005) in the 25-95% quantiles. A post-hoc vertex-wise analysis identified the right lingual and, to a lesser extent the fusiform gyrus. Overall, the surface reduction explained by the anxiety-depression scores was modest (r=-0.10, 3rd order spline, and r=-0.040, 1st order spline in the HCP). The discordant results in the top 5% of the anxiety-depression scores may be explained by differences in recruitment between the studies and especially medication screening as anti-depressants may increase the lingual gyrus volume (Jung et al., 2014). However, we could not conclude whether this cortical region was an endophenotype for anxiety-depression as the genetic correlations did not reach significance, which we attribute to the modest effect size (post-hoc statistical power<10%).

92 Significance Statement

93 Endophenotypes may help shed light on the aetiology, cognitive impairment and
94 genetics of psychiatric disorders. Here, we report a non-linear negative
95 association between anxiety-depression and smaller surface area of the occipito-
96 temporal region, which comprises most of the right lingual and fusiform gyri
97 (N=834). This cluster was defined by applying a fuzzy clustering algorithm to a
98 matrix of vertex-wise genetic correlations among cortical surface measures. We
99 replicated this association in an independent sample from the Human
100 Connectome Project (N=890). We could not confirm the presence of a genetic
101 correlation as the effect size of the association was modest (r=-0.10).

102

103 Introduction

104 Endophenotypes may be useful in decomposing or deconstructing a
105 psychiatric disorder, which should result in more successful genetic analyses,
106 redefinition of diagnosis, improved studies of the course of illness and
107 development of pertinent animal models (Gottesman and Gould, 2003).
108 According to a widely accepted definition, an endophenotype needs to satisfy 4
109 core criteria (Gottesman and Gould, 2003): it should be heritable, genetically and
110 phenotypically associated with the illness in the population and state-
111 independent (not a consequence of the disorder). There should also be a genetic
112 correlation between the disease and endophenotype: unaffected family members
113 of affected individuals should also exhibit the endophenotype, to some extent, as
114 they have a higher genetic load for the disease than people who are not closely
115 related to affected individuals. Reliability of the endophenotype is sometimes

included as part of the definition (Goldstein and Klein, 2014) and is implied by the trait heritability (Dohm, 2002).

The endophenotype concept is popular (Glahn et al., 2014), but only a few endophenotypes have been identified for affective disorders (Glahn et al., 2012; Goldstein and Klein, 2014; Bas-Hoogendam et al., 2016). These include: neuroticism (Goldstein and Klein, 2014), personal life events (Boardman et al., 2011) and perceived social support (Kendler and Baker, 2007; Matthews et al., 2016). Magnetic resonance imaging (MRI) brain measures may be promising endophenotypes, as they are objective (not self-reported), reliable (Liem et al., 2015) and heritable (Strike et al., 2015; Whelan et al., 2016). There has been an explosion of proposed brain endophenotypes for depression, and most publications report unreplicated phenotypic associations based on small sample sizes (see (Savitz and Drevets, 2009; Hasler and Northoff, 2011; Zhang et al., 2013) for reviews). To overcome publication bias and propose robust candidate endophenotypes, the Major Depressive Disorder (MDD) group of the ENIGMA consortium (Thompson et al., 2014) has conducted large-scale case-control comparisons, which found several structural brain phenotypes consistently associated with MDD (Schmaal et al., 2015; Schmaal et al., 2016), but somewhat conflicting cortical signatures of depression when stratifying the analysis by age (cut-off 21 yrs.). In addition, another large voxel-wise meta-analysis of cortical grey matter density reported a profile of volume changes associated with MDD (Arnone et al., 2016). However, more work is needed to determine genetic relationships between depression and structural brain measures and to replicate the reported associations.

Here, we aimed to identify cortical thickness (CT) or surface area (SA) endophenotypes for depression and anxiety, which share most of their genetic risk (Kendler et al., 1987; Hettema, 2008). We considered both CT and SA as there is evidence they are influenced by specific environmental and genetic factors, and thus are preferred to combined volumetric measures as they gather more information (Panizzon et al., 2009; Winkler et al., 2010). We performed an exploratory data driven analysis in a large, genetically informative young adult twin sample, followed by a replication in an independent cohort of similar age. The large cohorts allowed us to investigate the presence of non-linear relationships between cortical structures and anxiety-depression level.

We used twin modelling (Martin and Eaves, 1977; Neale and Cardon, 1992) to perform a genetic parcellation of the cortex (Chen et al., 2013), aiming to reduce the data dimension, while ensuring genetic homogeneity of the brain phenotypes studied. We controlled for mean SA/CT to adjust for the confounding effect of overall brain or body size. To measure anxiety-depression, we used a validated continuous measure of anxiety-depression (SPHERE: Somatic and Psychological Health Report) (Hickie et al., 2001) collected before the MRI scan, rather than a DSM-IV lifetime diagnosis (with disease onset not always preceding imaging). We could then get around the question of state-independence (even if a longitudinal design is needed to confirm the direction of causation) and maximised statistical power by providing a larger sample with greater variance across the population. In the replication sample, we used validated Achenbach scales (Achenbach et al., 2005; Achenbach, 2009) collected at the time of imaging. Furthermore, we aimed to characterise sources of covariation between anxiety-depression and cortical measures using twin modelling.

Materials and Methods

Exploratory Analysis

Participants

We analysed data from a total of 834 twins and siblings from the Brisbane Longitudinal Twin Study (BLTS)(Wright and Martin, 2004; Gillespie et al., 2013) who were assessed for symptoms of anxiety-depression and who had undergone brain imaging. Anxiety-Depression was assessed at mean age 17 (SD=2.2, range 10-24) with imaging completed 4.4 years (SD=2.1, range 0.6 to 9.5 years) later, at a mean age of 21 (SD=3.2, range 15-29). The sample (66% female) comprised 23 twin-sibling trios (9 MZ, 14 DZ), 275 complete pairs (101 MZ, 142 DZ and 32 pairs of siblings treated as DZ) and 214 singletons (Table 1). Zygosity of twin pairs was initially determined from DNA using a commercial kit (AmpFISTR Profiler Plus Amplification Kit, ABI) and was later confirmed by genome-wide single nucleotide polymorphism genotyping (Illumina 610K chip).

Brain Imaging

Imaging was conducted using a 4T Bruker Medspec whole-body MRI on 1,161 twins and their siblings aged 15 to 30 (62% female), as part of the QTIM study of brain structures and function (de Zubicaray et al., 2008; Blokland, 2014; Renteria et al., 2014). Participants were all right-handed and were screened for prior mental health diagnoses and anti-depressant use, as well as neurological disorders and loss of consciousness. Structural T1-weighted 3D images were acquired with the following parameters: TR=1500ms, TE=3.35ms, TI=700ms, 240mm FOV, 0.9mm slice thickness, 256 or 240 slices depending on acquisition

orientation: 86% coronal (256 slices), 14% sagittal (240 slices). The raw T1-weighted images were corrected for intensity inhomogeneity with SPM8 (Flandin and Friston, 2008; Wellcome Department of Imaging Neuroscience, 2009). Cortical surfaces were reconstructed using FreeSurfer (v5.3) (Fischl and Dale, 2000; Fischl, 2012), then resampled into a common space, smoothed (N = 2819 iterations, nearest neighbour) and down-sampled, with SA and CT measured at each surface vertex (2562 per hemisphere). Whole brain total SA and mean CT were additionally extracted. : From the total imaging sample (N=1,161), we excluded 11% of participants due to neuroanatomical abnormalities (n=48), excessive head motion during scanning (n=16), or poor quality FreeSurfer cortical surface reconstructions (n=57, using the ENIGMA quality checking procedure that consists of manual QC assisted by automatic detection of outliers (enigma.ini.usc.edu)).

Assessment of anxiety-depression

The anxiety-depression score was calculated using the SPHERE questionnaire (Hickie et al., 2001), which has been administered to 3,312 twins and siblings across 1 to 4 waves of the BLTS. When several SPHERE scores were available, we selected the closest to the time of scan. We previously showed that this score is moderately heritable (Hansell et al., 2012) and has good psychometric properties (stochastic ordering on the sum score, 3 months' test-retest, internal consistency, limited sex bias) (Couvy-Duchesne et al., 2017). In addition, we have shown that the anxiety-depression score from the SPHERE collected in adolescence predicts lifetime DSM-IV MDD diagnoses in young adulthood in the BLTS sample.

Table 1: Detailed family structure of the QTIM and HCP samples

			QTIM	HCP
Total sample size (Individuals) for phenotypic analysis			834	890
Final sample size (individuals) for twin modeling			833*	853** ♣
Incl.	N complete trios		23	184
	Incl.	N MZ pairs + one sibling	9	75
		N DZ pairs + one sibling	14	71
		N sibling trios	0	38
	N complete pairs		275	109
	Incl.	N MZ pairs	101	14
		N DZ pairs	142	19
		N twin-sib pairs	32	38
		N sibling pairs	0	38
	N Singletons (214 individuals)		214	83
	Incl.	N twins	181	33
		N non-twin siblings	33	50

* 1 individual from a family of 4 siblings was excluded for twin modelling (i.e. maximum family size restricted to 3 individuals (twin and/or non-twin siblings)

** 37 individuals from families with 4+ siblings were excluded with no effect on the sample characteristics (mean age 28, sd=3.7, range 22-37, 56.5% females after exclusion)

♣ To maximise the HCP sample 12 half-siblings were included (categorised as siblings in twin analysis); this low number should not affect estimates from twin modelling.

Genetic parcellation of the cortex

Twin modelling contrasts MZ twins (who share the same genetic information and familial environment) with DZ twins (who have the same familial environment but share on average half of their genetic information) and unrelated pairs (independent familial environment and genetic information). It allows the estimation of the proportion of inter-individual differences

attributable to genetic variability in the population (narrow sense heritability or additive genetic effects (A), the individuals' unique environment (E) and either the familial environment (C) or genetic dominance (D) (Martin and Eaves, 1977; Neale and Cardon, 1992; Verweij et al., 2012). Multivariate models can further break down the sources of variances into common and trait specific, providing an estimate of genetic and environmental correlations (Martin and Eaves, 1977; Neale and Cardon, 1992).

Following prior work (Chen et al. 2013), and to reduce the number of brain phenotypes tested (and the burden of multiple testing correction), we estimated the matrices of genetic correlations between the vertex-wise SA and CT measures (2305 left hemisphere, 2308 right hemisphere; medial wall vertices excluded) within each hemisphere. To do this, we used bivariate AE models (Neale and Cardon, 1992) using OpenMx (Boker et al., 2011), which implements Full Information Maximum Likelihood (FIML) and allows some missingness in the outcome variables. Omitting C/D terms in the models may result in a slight overestimation of the genetic correlation, as only moderate shared environmental sources of variance have been detected throughout the cortex (Lenroot et al., 2009; Kremen et al., 2010; Eyler et al., 2011). We included a maximum of 3 individuals (e.g. 1 twin pair and a sibling) per family. This increases power to detect A and C/D compared to the standard twin design (Posthuma and Boomsma, 2000).

Before estimating the genetic correlations, we residualised the CT and SA measures to remove global effects (whole brain total SA, mean CT), as well as effects of sex, age, acquisition orientation. We used a fuzzy clustering algorithm to identify clusters of genetically correlated voxels (Kaufman and Rousseeuw,

1990; Chen et al., 2013) implemented in the R package "cluster" (R Development Core Team, 2012; Maechler et al., 2016) and determined the optimal number of clusters using silhouette coefficients (Kaufman and Rousseeuw, 1990; Chen et al., 2013) (Figure 1). Such coefficients combine cluster cohesion (intra cluster differences) and separation (inter cluster differences). A high coefficient indicates better-separated clusters (Chen et al., 2013). The silhouette coefficients plateaued after 10 to 14 clusters, for each hemisphere and measurement (Figure 1). We restricted our analyses to 12 clusters, which results in a relatively parsimonious parcellation, to facilitate both comparisons across hemisphere and with prior work (Chen et al., 2013). For each cluster, we calculated the mean thickness and surface area, reducing the cortex to a total of 48 phenotypes. See Figure 2 for cluster visualisation and labelling. See extended data 1-4 for correspondence between cortical vertices and SA or CT clusters. In this analysis, we preferred the clusters derived from our QTIM sample to those defined previously in a sample of middle-aged male veterans from the US (Chen et al., 2013). Though, it is interesting to observe that the clusters are largely conserved across these two samples (Couvy-Duchesne, 2017).

Association between cortical measures and anxiety-depression

First, we tested the phenotypic association between the anxiety-depression SPHERE score and both SA and CT for each of the genetic clusters. To capture complex associations, we modelled linear, quadratic and cubic relationships using natural splines, implemented in R (R Development Core Team, 2012). Natural splines are piecewise polynomials (of 3rd order here) that include constraints at the boundaries, forcing the tails of the association to be

linear, thus preventing diverging "unnatural" solutions. All models included the 3 splines orders, required to make the results interpretable. This approach can detect both linear and U-shaped (quadratic) relationships, as well as changes limited to the extreme of the anxiety-depression continuum (cubic), or even a mixture of these. Compared to simple linear modelling, this should provide a better fit to the data and may detect associations otherwise overlooked. In addition, splines are more robust than polynomial regression as the splines orders are constrained to be orthogonal, thus removing issues arising from collinearity between the predictors. To limit the effect of outlier scores on the results, we winsorised four observations with SPHERE sum scores greater than 20 (more than 4 SD away from the mean, in the heavy right tail).

To control for sample relatedness, we used a mixed model that integrates the variance-covariance of observations via a 834x834 kinship matrix (calculated with R package "kinship2" (Therneau and Sinnwell, 2015)). To estimate the parameters of the models, we used the "hglm" package (Ronnegard et al., 2010) that relies on extended quasi-likelihood (Lee and Nelder, 1996; Lee and Lee, 2012; Ronnegard and Lee, 2013). The significance of the anxiety-depression score (and all fixed effects) was tested using Students t-tests (test statistic: $\beta/SD(\beta)$). We included as covariates the linear and non-linear (quadratic and cubic) effects of age (at questionnaire and age at scan), sex, wave, acquisition direction (coronal or sagittal), as well as mean vertex-wise SA/CT in order to adjust for sex and age, which are in turn strongly associated with head/body size/cortical surface area (sex, age) and thickness (age).

To avoid an overly stringent multiple testing correction, we estimated the effective number of independent brain phenotypes from the eigenvalues of the

correlation matrix (Li and Ji, 2005). This method provides an accurate and fast alternative to permutation tests (Li and Ji, 2005). We estimated the effective number of genetic clusters to be 33 (out of a total of 48) (Li and Ji, 2005), after regressing out the covariates (age, sex, acquisition, wave, mean vertex-wise SA/CT) that tend to inflate the coefficients of the correlation matrix. This translates into about 99 independent tests (as we are testing 3 effects: linear, quadratic and cubic), which yields a significance threshold of 5.2E-4 (Sidak, 1967) corresponding to a FWER (Family Wise Error Rate) less than or equal to 5%.

Then, we used bivariate (AE) twin models to estimate the genetic and environmental correlations between cluster measurements and anxiety-depression. Covariates used in the phenotypic association analysis were regressed from the brain measurement and the residuals were used in the twin analysis (see Table 1 for sample composition. In this paper, we only used residuals in multivariate OpenMx models in order to reduce computational time and directly included the covariates in all other models. We tested the significance of the heritability estimates as well as the genetic and environmental correlations using a likelihood ratio test on nested models (1 degree of freedom tests). We corrected for the number of tests performed. Twin analyses were performed in OpenMx (Boker et al., 2011). We also conducted a post-hoc analysis to highlight which voxels from significantly associated genetic clusters were driving the phenotypic association with anxiety-depression.

Post-hoc association analyses in the QTIM sample

In addition to the main analysis, we report the association of the SPHERE scores with the surface and thickness measurements derived from the Desikan anatomical atlas (Desikan et al., 2006), as well as a vertex-wise analysis. This facilitates comparison with previous studies and shows that the reported associations are not an artefact of our genetic parcellation. The same models and covariates were used as for the main analysis. We estimated the significance threshold to be 1.5E-4 (108 effective phenotypes) for the associations with the anatomical cortical regions. For the vertex-wise analysis, we used the brain-wide significance threshold (1.6E-5, based on 10,000 effective phenotypes) calculated by Medland et al., (Medland et al., 2014). Both significance thresholds took into account that we tested 3 splines orders.

Replication Analysis

Participants

We used MRI and Achenbach Adult Self Report (ASR) questionnaire (Achenbach, 2009) data for 890 participants (mean age 28 SD 3.7 range 22-37, 56.5% female) from the Human Connectome Project (HCP) (Van Essen et al., 2012b; Van Essen et al., 2013) to replicate our results. In twin modelling, a maximum of 3 participants per family was included, resulting in 184 complete trios, 109 pairs and 83 singletons (see Table 1 for detailed breakdown).

We utilised pre-processed T1-weighted structural scans from the HCP sample (Marcus et al., 2011). Minimal processing of the structural images by the HCP team consists of removing spatial artifacts and T1 alignment (Van Essen et al., 2012a; Glasser et al., 2013), using FSL (Jenkinson et al., 2002; Jenkinson et al., 2012) and FreeSurfer (Fischl, 2012). We applied the same techniques used for

QTIM data to produce the vertex-wise and genetic cluster measures from the pre-processed HCP surfaces.

We used the two ASR scores that were available: the anxiety-depression syndrome based scale (Achenbach, 2009) and a DSM-oriented scale that we constructed by summing the DSM-oriented scales for both anxiety and depression (Achenbach et al., 2003). The correlation between the DSM-oriented anxiety and depression scales was 0.66 (95%CI 0.62 0.70) and mostly driven by common genetic sources of variance: rG=0.87 (95%CI 0.70 1.00), rE=0.54 (95%CI 0.42 0.64). These high phenotypic and genetic correlations suggest that combining the 2 DSM-oriented scales is a valid approach. Participants reported an average syndrome score of 5.7 (SD=5.2, range 0-33) and an average DSM-oriented score of 8.0 (SD=5.6, range 0-35). The phenotypic correlation between the DSM-oriented and syndrome anxiety-depression score was 0.90 (95%CI 0.89 0.91), with comparable genetic and environmental correlations (rG=0.91, 95%CI 0.84 0.95, rE=0.89, 95%CI 0.85 0.92). Notably, the genetic correlation between the two scores was significantly different from 1 (p-value=1.2E-4), suggesting that they may have some unique genetic sources of variance. All ASR scales have previously shown good test-retest reliability and internal consistency (Achenbach et al., 2005). In addition, the ASR syndrome based scale is heritable through adolescence (Nivard et al., 2015) and captures a stable construct across age and sex (Fonseca-Pedrero et al., 2012). DSM-oriented and syndrome based scales appear to comparably predict affective disorder diagnoses (Najman et al., 2008; Dingle et al., 2010; Dingle et al., 2011).

Replication Analysis

We considered for replication all significant associations between the anxiety-depression (SPHERE) score and genetic clusters identified in the exploratory analysis. As before, we tested for linear, quadratic and cubic association with the brain phenotypes. We also corrected for sex, acquisition variables, total SA or CT, linear and non-linear effects of age as well as familial relatedness using a kinship matrix (Therneau and Sinnwell, 2015). We controlled for multiple testing in the replication analysis by estimating the effective number of independent phenotypes carried in the replication step (Li and Ji, 2005). We further corrected for testing linear, quadratic and cubic relationships and for considering the two ASR anxiety-depression scores. ASR scores more than 4 standard deviations from the mean were winsorised to limit the influence of outliers (heavy right tail) on the results.

We then explored the voxel-wise associations within replicating clusters and further decomposed the associations into their genetic and environmental components using bivariate AE twin models in OpenMx (Boker et al., 2011). As before, we regressed out the covariates before twin modelling.

Code accessibility

The code/software described in the paper is freely available online at [retracted for double-blind review]. The code is also accessible as extended data.

Results

Exploratory Analysis

Across the 48 brain clusters only one (cluster 6R for SA) comprising the lower part of the occipital cortex was associated with the anxiety-depression

score from the SPHERE, after correcting for multiple tests (third order spline association: β=-0.037, SD=8.6E-3, rP=-0.10, p-value=2.4E-05, Table 2). Surface area in this cluster appeared to peak for participants with a median SPHERE score (score = 2.8) and to be substantially decreased (up to 1 SD lower) in those with a high-anxiety depression score (Figure 3a). Though, we should be cautious interpreting the increasing surface area for participants with low SPHERE scores, as it may be an artefact of plotting only the significant spline order. Indeed, this increase is not observed when plotting the effect of all splines order (dotted line, Figure 3a). We are confident the model is not over fitted as the 3 splines order only accounted for a 1.6% of the score variance and we did not observe large standard errors for the estimates, which often indicate colinearity.

Fitting AE models, heritability was estimated to be 0.55 (95%CI: 0.32 0.73) for the 6R cluster for SA and 0.28 (95%CI: 0.14 0.41) for anxiety-depression (SPHERE), but neither the genetic nor environmental correlations between these traits reached significance (cubic correlations: rG=-0.068, p-value=0.35 and rE=-0.084, p-value=0.11; linear correlations: rG=-0.20, pvalue=0.35 and rE=0.0091, p-value=0.91). In comparison, heritability of the other SA clusters ranged from 0.39 to 0.71 (0.32 to 0.63 for the CT clusters).

The 6R cluster for SA consists of a total of 94 vertices and is located in the occipital cortex. It comprises most of the fusiform gyrus (56.6% or 43 vertices) and the parahippocampal gyrus (56.5% or 13 vertices) as well as the lower part of the lingual gyrus (33.8% of the gyrus or 22 vertices) and the medial part of the lateral occipital cortex (12.6% or 12 vertices). To identify if one or more regions of the 6R cluster were driving the cubic association between SA and the anxiety-depression SPHERE score we restricted our analysis to the 94 voxels in the 6R

cluster. Using a significance threshold that accounts for the extra number of tests performed (p-value<4.2E.4), 61 voxels were found to be driving the non-linear association between SA of the cluster 6R and the anxiety-depression SPHERE score (p-value range: 5.9E-06 3.8E-04, correlation range: -0.13 -0.091, Figure 4a). When decomposing the phenotypic association using an AE model, none of the genetic or environmental correlations was significantly different from zero.

435

436 Table 2: Summary of discovery (QTIM sample) and replication (HCP sample) analysis

		First order spline (linear) *Beta (se) r*	*p-value*	Second order spline (quadratic) *Beta (se) r*	*p-value*	Third order spline (cubic) *Beta (se) r*	*p-value*
QTIM	SPHERE scale	0.010 (6.0E-3) r=0.030	0.094	-0.021 (6.5E-3) r=-0.069	1.6E-3	-0.037 (8.6E-3) r=-0.10	2.4E-5
HCP	DSM-oriented scale	-0.18 (0.060) r=-0.040	3.2E-3	0.015 (0.12) r=0.0022	0.89	0.15 (0.120) r=0.022	0.13
	ASR syndrome based scale	-0.032 (0.076) r=-0.0064	0.68	-0.044 (0.10) r=-0.0088	0.67	0.079 (0.15) r=0.012	0.61

437 Significant associations (after multiple testing correction) are reported in bold.

438

439 Replication Analysis

440 We found a significant (first order spline) association between the DSM-
441 oriented score for anxiety-depression and SA of the 6R cluster (Table 2), using a
442 significance threshold of 8.5E-3, which corresponds to a FWER<5% considering
443 6 independent tests (3 tests of association for 2 anxiety-depression scores). This
444 association supports our finding in the QTIM data set, of a reduction of SA (r=-
445 0.040, p-value=3.2E-3) in participants with moderate anxiety-depression score
446 (25-95 percentile) (Figure 3b), which we reported above.

447 In the HCP dataset, SA for the 6R cluster was moderately heritable (0.63;
448 95%CI 0.52 - 0.72), after regressing mean SA and the other covariates, as well as
449 the anxiety-depression scores: 0.36 for the DSM oriented ASR (95%CI: 0.20 -
450 0.51) and 0.41 for the syndrome based ASR (95%CI 0.24 - 0.56). When breaking
451 down the phenotypic association between DSM-oriented scale and 6R SA, neither
452 genetic nor environmental components reached significance (linear correlations:
453 rG=-0.090, p-value=0.37 and rE=-0.032, p-value=0.59). Likewise, when using the
454 ASR syndrome based scale (rG=-0.078, p-value=0.56 and rE=-0.059, p-
455 value=0.49). Heritability of other SA clusters ranged from 0.46 to 0.82 in the HCP
456 (0.42 to 0.72 for CT clusters).

457 We then tested the association between each voxel of the cluster with the
458 DSM-oriented scale. We estimated the number of effective independent voxels to
459 be 7 (Li and Ji, 2005) after regressing out the covariates' effect on the vertex-
460 wise measurements. Thus, we used a significance threshold of 3.6E-3, to account
461 for 13 independent tests (6 previously and 7 for the voxel-wise testing).
462 Nineteen (out of 94) voxels located in the medial posterior part of the cluster
463 survived multiple testing correction (-log(p-value)>2.4, r in -0.040 -0.068, Figure

464 4b). Thirteen of these were located in the lingual gyrus and were also associated
465 with the anxiety-depression scores in QTIM. A further 3 neighbouring vertices
466 from the fusiform gyrus were also consistently associated to anxiety-depression
467 across the two samples (Table 3). As previously, vertex-wise genetic and
468 environmental correlations did not reach significance.
469
470 Table 3: Summary of post-hoc vertex-wise association testing for SA of the 6R cluster

	N total vertices in each gyrus	N vertices for cluster 6R		N vertices associated in vertex-wise analysis- QTIM	N vertices associated in vertex-wise replication analysis - HCP
		N	MNI coordinates		
Parahippocampal	23	13	x in 14 33 y in -64 -50 z in -19 -6	13	0
Lingual	65	22	x in 7 32 y in -107 -68 z in -4 6	22	13 #
Fusiform	76	43	x in 24 43 y in -102 -58 z in -19 3	22	3•
Lateral Occipital	95	12	x in 21 45 y in -109 -91 z in -8 0	0	2

471 # All vertices associated in QTIM. Vertices correspond to numbers 144, 570, 573, 2220, 2244,
472 2252, 2253, 2258, 2262, 2264, 2265, 2266, 2269 of the FreeSurfer fsaverage4 parcellation (MNI
473 coordinates: x in 7 20, y in -107 -82 and z in -3 6).
474 • All vertices associated in QTIM. Vertices correspond to number 2263, 2267 and 2270 of the
475 FreeSurfer fsaverage4 parcellation (MNI coordinates: x in 24 27, y in -98 -92, z in 0 3).
476
477 Post-hoc association analyses

The association analysis between the SPHERE scores and anatomical regions from the Desikan atlas returned one significant association with surface area of the right lingual gyrus (p=9.5E-5). As per the main QTIM analysis, the association was negative and with the 3rd order spline (β=-330.7, SD=83.9, r=-0.13). No significant association was found at a vertex-level for surface area or cortical thickness.

Discussion

Here, we used twin modelling to derive a genetic parcellation of two characteristics of the cortical ribbon (i.e. surface area and thickness), reducing each hemisphere to 12 measurements. The surface area of one region (cluster 6R) comprising part of the lingual, fusiform, parahippocampal and lateral occipital gyri showed a significant association with the anxiety-depression score from the SPHERE in the QTIM sample. We partially replicated this association in an independent sample of similar age, using data from the HCP, in which we found an association between the DSM-oriented score and surface area in the 6R cluster. Both associations suggested lower surface area in participants with anxiety-depression scores above average (50-95% quantiles). A post-hoc vertex-wise analysis suggested that the lingual vertices (and to a lesser extent fusiform) might drive the association. We cannot conclude whether the observed association is specific to the right hemisphere due to differences in the cortical parcellation between hemispheres (Figure 2, in the left hemisphere the occipital cortex is composed of a unique cluster).

Interestingly, a meta-analysis (472 cases, 680 controls, from 12 published studies, mean age 43, partially medicated) identified significant reductions of GM

density (Ashburner and Friston, 2000) (often considered as a proxy for volume) in right lingual, fusiform and parahipocampal gyri (Arnone et al., 2016). The authors found little evidence for publication bias and the voxel-based morphometry approach cannot be confounded by head/body size. This echoes results from the meta-analysis by the ENIGMA-MDD group, which reported a significant reduction of right lingual gyrus surface area in adolescent depression (237 cases, 294 controls) (Schmaal et al., 2016) (non-significant association for fusiform or parahippocampal). Though the meta-analysis by ENIGMA did not correct for total SA and included medicated cases, it is also not totally independent from the present study as it included QTIM participants (26 cases, 140 controls). Even so, in other work, right lingual GM density was identified as a predictor of anti-depressant response, with higher density predicting better response to first treatment (Jung et al., 2014). The non-respondent group showed lower performance in 3 aspects of the Stroop test (neutral word, colour word and error control), and a post-hoc analysis in cases showed right lingual density to be associated with better error control (Stroop), and non-verbal memory (Rey-Kim memory test). This is coherent with neurological case reports that highlight the crucial role of the lingual gyrus in visual memory (Bogousslavsky et al., 1987). Furthermore, impaired visual memory has been associated with 1st episode MDD (Lee et al., 2012) and with adolescent MDD (Baune et al., 2014); lower Stroop accuracy has also been reported in adult MDD (Snyder, 2013), while more research is needed in pediatric depression (Vilgis et al., 2015). Finally, reduced right lingual volume was also reported in adult MDD cases (partially medicated) (Lee et al., 2011), following traumatic brain injury (Maller et al., 2014).

The reported right lingual gyrus volume (or density) reduction in depression may be driven by reduced surface area, at least in early adulthood, which may be associated with visual memory and attention deficits in depression. In addition, the right lingual surface area may be associated with anti-depressant response and cognitive functions in MDD patients (Jung et al., 2014), which might explain why the findings are not consistent across samples that may differ in term of medication status, response to treatment or cognitive abilities. For example, screening in QTIM may exclude individuals with extreme levels of anxiety-depression, and we may be only observing part of the association.

Reduction of right fusiform gyrus volume has also been reported in off-medication adult patients (Lener et al., 2016), but several large studies of older (medicated) participants attributed this reduction to reduced cortical thickness (Canu et al., 2015; Schmaal et al., 2016). In addition, cortical thickness of the right fusiform in MDD cases with comorbid generalised anxiety was even more reduced (Canu et al., 2015). The right fusiform gyrus has a central role in face perception (Haxby et al., 2000), in neurological case reports of prosopagnosis (face blindness) (Whiteley and Warrington, 1977; Damasio et al., 1990; Derenzi et al., 1994), behavioural studies (Rhodes, 1993) and imaging studies (McCarthy et al., 1997; Haxby et al., 2000). Lower right fusiform grey matter density was observed in developmental prosopagnosis (Garrido et al., 2009) and in congenital prosopagnosia (Behrmann et al., 2007). In addition, both studies reported an association between right fusiform volume and performance in face identification (Behrmann et al., 2007; Garrido et al., 2009). Reviews on adult depression suggest a generally reduced accuracy of face expression evaluation,

coupled to increased attention and response bias in expression evaluation toward sadness (Bourke et al., 2010).

In adolescent depression, more research is needed to confirm the mixed evidence of impairment in face (and face expression) processing (Vilgis et al., 2015). Lower fusiform gyrus surface area in young adults might relate to the poorer face recognition and processing of face expressions reported in adult depression but we did not have the cognition data to test such hypothesis. However, our results only weakly point to the fusiform gyrus and do not align with prior studies that only reported a reduced fusiform thickness associated with MDD (Canu et al., 2015; Schmaal et al., 2016). Overall, we cannot rule out that medication or sample age might explain some of these differences, as well as the hypothesis of delayed cortical development in depressed participants, implying a normalisation of cortical surface in mid-adulthood that leaves a more permanent decrease in thickness (Schmaal et al., 2016).

In line with prior publications (Panizzon et al., 2009; Chen et al., 2013) that reported significant heritability of cortical thickness and surface, we confirmed that the SA of the 6R cluster is significantly heritable ($h2_{QTIM}$=0.55, $h2_{HCP}$=0.63), even after correcting for total SA, suggesting that this region is influenced by specific genetic sources of variance, that do not contribute to global SA. In addition, all anxiety-depression scores were comparably heritable ($h2_{SPHERE}$=0.28, $h2_{DSM\text{-}ASR}$=0.36, $h2_{ASR}$=0.41), in line with the known heritability of depression from twin and family studies (Sullivan et al., 2000; Kendler et al., 2006; Polderman et al., 2015). In the bivariate analysis, because of the modest phenotypic association, we had limited statistical power to determine whether common genetic or environmental factors drove the observed phenotypic

association. Indeed, a post hoc power analysis indicates that we had, at best, 10% power to detect the observed genetic correlation in QTIM and HCP (taking into account observed heritability and effect sizes, assuming an AE model and a risk α=5%). A combined analysis of the two samples would not confer a statistical power greater than 22%. The limited power may explain the instability of the correlations estimated in both samples, even if differences in anxiety-depression questionnaire, sample composition or recruitment could also be at play.

The reported heritabilities of the scores are consistent with the diathesis-stress model of anxiety-depression, in that genetic liability and environmental factors both contribute to the risk of depression (Rosenthal, 1963), sometimes in a multiplicative manner (Peyrot et al., 2014; Musliner et al., 2015; Mullins et al., 2016; Colodro-Conde et al., 2017). More generally, the identification of brain endophenotypes should lead to refine this model by providing insight into the brain networks and behavioural mechanisms that are influenced by the genetic and environmental risk factors. We still lack evidence to provide a credible interpretation for the association reported here, which may require reconciling evidence produced from complementary types of imaging. For example, reduced glucose metabolism in the right lingual gyrus has been linked to sleep anomalies in depression (Germain et al., 2004; Fitzgerald et al., 2008) but more work is need to investigate whether structural cortical markers may tag the same brain pathways and contribute to the same symptoms.

Our study has several limitations. Firstly, we did not perform a strict replication as QTIM and HCP differed in the anxiety-depression scores used, time difference between scoring and scanning, and shape of the relationship (3rd vs. 1st order splines). Thus, differences in studies design (different timeframes,

603 medication, population genetics and environment, ascertainment bias…) might
604 be partially responsible for the different effect size observed in QTIM and HCP. In
605 addition, the SPHERE and ASR both aim to measure anxiety-depression but they
606 are composed of different questions and may be sensitive to different ranges of
607 symptoms or different stages of the disorder (e.g. subclinical for the SPHERE and
608 clinical for the DSM based ASR). Furthermore, the association between 6R SA
609 and anxiety-depression scores in the HCP was only observed using the DSM-
610 oriented score, which we attribute to differences in scores distribution making it
611 harder to detect non-linear associations. Indeed, despite high (phenotypic and
612 genetic) correlations between the scores (r=0.90, rG=0.91), the correlation
613 between spline orders was lower (r=0.79 between first order splines of the 2
614 anxiety-depression scores, 0.56 between second order and 0.73 between third
615 order). This may reflect differences of scale and distribution. Differences in
616 genetic sources of variances (genetic correlation different from 1) are unlikely to
617 completely explain the differences in results, as the correlation remained high.
618 Nonetheless the different results between QTIM and HCP as well as the
619 inconsistent associations in the HCP sample using different scores remain
620 important limitations of this study. These limitations call for further
621 investigation of the relationship between lingual gyrus and anxiety-depression.
622 The small effect size of the association in QTIM and HCP can also be seen
623 as a limitation, as lingual gyrus SA does not explain much of the anxiety-
624 depression variability and would have a very limited predictive power. However,
625 our result is in line with the effect sizes observed in larger case-control studies
626 that suggest that no single brain marker of depression explains much of the
627 disorder (Schmaal et al., 2016). A parallel to this is the small SNP effects

628 observed/found in GWAS of complex traits (Visscher et al., 2017). Thus, despite
629 the small association between cortical regions and the trait/disease, their impact
630 on brain networks and functions may be much larger. That is, a small association
631 does not necessarily preclude a biological or cognitive relevance.

632 Further limitations include our clustering approach, which may make it
633 harder to detect localised structural changes associated with anxiety-depression,
634 as the thickness and surface are averaged over large cortical regions. In addition,
635 we did not investigate sex specific changes that may reflect depression subtypes
636 with different aetiology (Kendler et al., 2001; Kendler and Gardner, 2014). As per
637 the hypothesis of age specific markers reported by the ENIGMA-MDD (Schmaal
638 et al., 2016), age was not significant in our analysis, which prevented us from
639 testing for an age interaction with anxiety-depression scores. However larger
640 samples would be required to overcome the multiple testing correction burden
641 from a vertex-wise analysis or the loss of power resulting from study
642 stratification. Finally, the interpretation of our results is limited by the relative
643 lack of robust research on cognitive and imaging aspects of pediatric depression
644 and of normal neuro-development. The Research Domain Criteria (RDoC) may
645 help in making connections between cognition, imaging, genetics and psychiatric
646 illnesses.

647 In summary, we proposed a candidate endophenotype for depression,
648 which is heritable, and shows a replicable phenotypic association with anxiety-
649 depression scores. The vertex-wise post-hoc analysis suggested that a reduction
650 of SA in the ventral part of the lingual gyrus could drive the observed association.
651 Longitudinal studies beginning in adolescence are emerging (Luby et al., 2016;
652 Schmaal et al., 2017) and will help clarify the temporal and causal relationships

between brain development, cognition and mental health.

References

Achenbach TM (2009) Achenbach system of empirically based assessment (ASEBA): Development, findings, theory, and applications. In: University of Vermont, Research Center of Children, Youth & Families.

Achenbach TM, Dumenci L, Rescorla LA (2003) Ratings of Relations Between DSM-IV Diagnostic Categories and Items of the Adult Self-Report (ASR) and Adult Behavior Checklist (ABCL). In.

Achenbach TM, Bernstein A, Dumenci L (2005) DSM-oriented scales and statistically based syndromes for ages 18 to 59: Linking taxonomic paradigms to facilitate multitaxonomic approaches. J Pers Assess 84:49-63.

Arnone D et al. (2016) Computational meta-analysis of statistical parametric maps in major depression. Hum Brain Mapp 37:1393-1404.

Ashburner J, Friston KJ (2000) Voxel-based morphometry - The methods. NeuroImage 11:805-821.

Bas-Hoogendam JM, Blackford JU, Bruhl AB, Blair KS, van der Wee NJ, Westenberg PM (2016) Neurobiological Candidate Endophenotypes of Social Anxiety Disorder. Neurosci Biobehav Rev.

Baune BT, Fuhr M, Air T, Hering C (2014) Neuropsychological functioning in adolescents and young adults with major depressive disorder - A review. Psychiat Res 218:261-271.

Behrmann M, Avidan G, Gao F, Black S (2007) Structural Imaging reveals anatomical alterations in inferotemporal cortex in congenital prosopagnosia. Cerebral Cortex 17:2354-2363.

Blokland GAM (2014) Genetic effects on the cerebellar role in working memory: Same brain, different genes? NeuroImage 86:392-403.

Boardman JD, Alexander KB, Stallings MC (2011) Stressful Life Events and Depression among Adolescent Twin Pairs. Biodemogr Soc Biol 57:53-66.

Bogousslavsky J, Miklossy J, Deruaz JP, Assal G, Regli F (1987) Lingual and fusiform gyri in visual processing: a clinico-pathologic study of superior altitudinal hemianopia. J Neurol Neurosurg Psychiatry 50:607-614.

Boker S, Neale M, Maes H, Wilde M, Spiegel M, Brick T, Spies J, Estabrook R, Kenny S, Bates T, Mehta P, Fox J (2011) OpenMx: An Open Source Extended Structural Equation Modeling Framework. Psychometrika 76:306-317.

Bourke C, Douglas K, Porter R (2010) Processing of facial emotion expression in major depression: a review. Aust N Z J Psychiatry 44:681-696.

Canu E, Kostić M, Agosta F, Munjiza A, Ferraro P, Pesic D, Copetti M, Peljto A, Tosevski D, Filippi M (2015) Brain structural abnormalities in patients with major depression with or without generalized anxiety disorder comorbidity. Official Journal of the European Neurological Society 262:1255-1265.

Chen CH, Fiecas M, Gutierrez ED, Panizzon MS, Eyler LT, Vuoksimaa E, Thompson WK, Fennema-Notestine C, Hagler DJ, Jr., Jernigan TL, Neale MC, Franz CE,

699 Lyons MJ, Fischl B, Tsuang MT, Dale AM, Kremen WS (2013) Genetic
700 topography of brain morphology. Proc Natl Acad Sci U S A 110:17089-
701 17094.
702 Colodro-Conde L, Couvy-Duchesne B, Zhu G, Coventry WL, Byrne EM, Gordon S,
703 Wright MJ, Montgomery GW, Madden PAF, Major Depressive Disorder
704 Working Group of the Psychiatric Genomics C, Ripke S, Eaves LJ, Heath AC,
705 Wray NR, Medland SE, Martin NG (2017) A direct test of the diathesis-
706 stress model for depression. Mol Psychiatry.
707 Couvy-Duchesne B (2017) Breaking down the genetics of depression using brain
708 endophenotypes. In: Queensland Brain Institute, p 370. PhD Thesis: The
709 University of Queensland.
710 Couvy-Duchesne **B**, Davenport TA, Martin NG, Wright MJ, Hickie IB (2017)
711 Validation and psychometric properties of the Somatic and Psychological
712 HEalth REport (SPHERE) in a young Australian-based population sample
713 using nonparametric Item Response Theory. BMC Psychiatry In Press.
714 Damasio AR, Tranel D, Damasio H (1990) Face agnosia and the neural substrates
715 of memory. Annu Rev Neurosci 13:89-109.
716 de Zubicaray GI, Chiang MC, McMahon KL, Shattuck DW, Toga AW, Martin NG,
717 Wright MJ, Thompson PM (2008) Meeting the challenges of neuroimaging
718 genetics. Brain imaging and behavior 2:258-263.
719 Derenzi E, Perani D, Carlesimo GA, Silveri MC, Fazio F (1994) Prosopagnosia Can
720 Be Associated with Damage Confined to the Right-Hemisphere - an Mri
721 and Pet Study and a Review of the Literature. Neuropsychologia 32:893-
722 902.
723 Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL,
724 Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ (2006) An
725 automated labeling system for subdividing the human cerebral cortex on
726 MRI scans into gyral based regions of interest. NeuroImage 31:968-980.
727 Dingle K, Alati R, Williams GM, Najman JM, Bor W, Clavarino A (2010) The ability
728 of YSR DSM-oriented depression scales to predict DSM-IV depression in
729 young adults: A longitudinal study. J Affect Disorders 121:45-51.
730 Dingle K, Clavarino A, Williams GM, Bor W, Najman JM, Alati R (2011) Predicting
731 depressive and anxiety disorders with the YASR internalising scales
732 (empirical and DSM-oriented). Social Psychiatry and Psychiatric
733 Epidemiology 46:1313-1324.
734 Dohm MR (2002) Repeatability estimates do not always set an upper limit to
735 heritability. Functional Ecology 16:273-280.
736 Eyler LT et al. (2011) Genetic and Environmental Contributions to Regional
737 Cortical Surface Area in Humans: A Magnetic Resonance Imaging Twin
738 Study. Cerebral Cortex 21:2313-2321.
739 Fischl B (2012) FreeSurfer. NeuroImage 62:774-781.
740 Fischl B, Dale AM (2000) Measuring the thickness of the human cerebral cortex
741 from magnetic resonance images. Proceedings of the National Academy of
742 Sciences 97:11050-11055.
743 Fitzgerald PB, Laird AR, Maller J, Daskalakis ZJ (2008) A meta-analytic study of
744 changes in brain activation in depression. Human Brain Mapping 29:683-
745 695.
746 Flandin G, Friston KJ (2008) Statistical parametric mapping (SPM). Scholarpedia
747 3(4):6232.

748 Fonseca-Pedrero E, Sierra-Baigrie S, Lemos-Giraldez S, Paino M, Muniz J (2012)
749 Dimensional Structure and Measurement Invariance of the Youth Self-
750 Report Across Gender and Age. J Adolescent Health 50:148-153.
751 Garrido L, Furl N, Draganski B, Weiskopf N, Stevens J, Tan GCY, Driver J, Dolan RJ,
752 Duchaine B (2009) Voxel-based morphometry reveals reduced grey
753 matter volume in the temporal cortex of developmental prosopagnosics.
754 Brain 132:3443-3455.
755 Germain A, Nofzinger EA, Kupfer DJ, Buysse DJ (2004) Neurobiology of non-REM
756 sleep in depression: Further evidence for hypofrontality and thalamic
757 dysregulation. Am J Psychiat 161:1856-1863.
758 Gillespie NA, Henders AK, Davenport TA, Hermens DF, Wright MJ, Martin NG,
759 Hickie IB (2013) The Brisbane Longitudinal Twin Study: Pathways to
760 Cannabis Use, Abuse, and Dependence project-current status, preliminary
761 results, and future directions. Twin Res Hum Genet 16:21-33.
762 Glahn DC, Knowles EE, McKay DR, Sprooten E, Raventos H, Blangero J,
763 Gottesman, II, Almasy L (2014) Arguments for the sake of
764 endophenotypes: examining common misconceptions about the use of
765 endophenotypes in psychiatric genetics. Am J Med Genet B
766 Neuropsychiatr Genet 165B:122-130.
767 Glahn DC, Curran JE, Winkler AM, Carless MA, Kent JW, Jr., Charlesworth JC,
768 Johnson MP, Goring HH, Cole SA, Dyer TD, Moses EK, Olvera RL, Kochunov
769 P, Duggirala R, Fox PT, Almasy L, Blangero J (2012) High dimensional
770 endophenotype ranking in the search for major depression risk genes.
771 Biol Psychiatry 71:6-14.
772 Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J,
773 Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinson M, Consortium
774 WU-MH (2013) The minimal preprocessing pipelines for the Human
775 Connectome Project. NeuroImage 80:105-124.
776 Goldstein BL, Klein DN (2014) A review of selected candidate endophenotypes
777 for depression. Clin Psychol Rev 34:417-427.
778 Gottesman, II, Gould TD (2003) The endophenotype concept in psychiatry:
779 etymology and strategic intentions. Am J Psychiatry 160:636-645.
780 Hansell NK, Wright MJ, Medland SE, Davenport TA, Wray NR, Martin NG, Hickie
781 IB (2012) Genetic co-morbidity between neuroticism, anxiety/depression
782 and somatic distress in a population sample of adolescent and young
783 adult twins. Psychol Med 42:1249-1260.
784 Hasler G, Northoff G (2011) Discovering imaging endophenotypes for major
785 depression. Mol Psychiatr 16:604-619.
786 Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system
787 for face perception. Trends Cogn Sci 4:223-233.
788 Hettema JM (2008) What is the genetic relationship between anxiety and
789 depression? Am J Med Genet C 148c:140-146.
790 Hickie IB, Davenport TA, Hadzi-Pavlovic D, Koschera A, Naismith SL, Scott EM,
791 Wilhelm KA (2001) Development of a simple screening tool for common
792 mental disorders in general practice. Med J Aust 175 Suppl:S10-17.
793 Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for
794 the robust and accurate linear registration and motion correction of brain
795 images. NeuroImage 17:825-841.

796 Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM (2012) FSL.
797 NeuroImage 62:782-790.
798 Jung J, Kang J, Won E, Nam K, Lee MS, Tae WS, Ham BJ (2014) Impact of lingual
799 gyrus volume on antidepressant response and neurocognitive functions
800 in Major Depressive Disorder: A voxel-based morphometry study. J Affect
801 Disorders 169:179-187.
802 Kaufman L, Rousseeuw PJ (1990) Finding groups in data: an introduction to
803 cluster analysis. New York ; Chichester, England ; Brisbane, Qld: Wiley.
804 Kendler KS, Baker JH (2007) Genetic influences on measures of the environment:
805 a systematic review. Psychological Medicine 37:615-626.
806 Kendler KS, Gardner CO (2014) Sex differences in the pathways to major
807 depression: a study of opposite-sex twin pairs. Am J Psychiatry 171:426-
808 435.
809 Kendler KS, Heath AC, Martin NG, Eaves LJ (1987) Symptoms of anxiety and
810 symptoms of depression. Same genes, different environments? Arch Gen
811 Psychiatry 44:451-457.
812 Kendler KS, Gardner CO, Neale MC, Prescott CA (2001) Genetic risk factors for
813 major depression in men and women: similar or different heritabilities
814 and same or partly distinct genes? Psychol Med 31:605-616.
815 Kendler KS, Gatz M, Gardner CO, Pedersen NL (2006) A Swedish national twin
816 study of lifetime major depression.(Author Abstract). Am J Psychiat
817 163:109.
818 Kremen WS, Prom-Wormley E, Panizzon MS, Eyler LT, Fischl B, Neale MC, Franz
819 CE, Lyons MJ, Pacheco J, Perry ME, Stevens A, Schmitt JE, Grant MD,
820 Seidman LJ, Thermenos HW, Tsuang MT, Eisen SA, Dale AM, Fennema-
821 Notestine C (2010) Genetic and environmental influences on the size of
822 specific brain regions in midlife: The VETSA MRI study. NeuroImage
823 49:1213-1223.
824 Lee HY, Tae WS, Yoon HK, Lee BT, Paik JW, Son KR, Oh YW, Lee MS, Ham BJ
825 (2011) Demonstration of decreased gray matter concentration in the
826 midbrain encompassing the dorsal raphe nucleus and the limbic
827 subcortical regions in major depressive disorder: An optimized voxel-
828 based morphometry study. J Affect Disorders 133:128-136.
829 Lee RSC, Hermens DF, Porter MA, Redoblado-Hodge MA (2012) A meta-analysis
830 of cognitive deficits in first-episode Major Depressive Disorder. J Affect
831 Disorders 140:113-124.
832 Lee W, Lee Y (2012) Modifications of REML algorithm for HGLMs. Statistics and
833 Computing 22:959-966.
834 Lee Y, Nelder JA (1996) Hierarchical Generalized Linear Models. Journal of the
835 Royal Statistical Society Series B (Methodological) 58:619-678.
836 Lener MS, Kundu P, Wong E, Dewilde KE, Tang CY, Balchandani P, Murrough JW
837 (2016) Cortical abnormalities and association with symptom dimensions
838 across the depressive spectrum. J Affect Disord 190:529-536.
839 Lenroot RK, Schmitt JE, Ordaz SJ, Wallace GL, Neale MC, Lerch JP, Kendler KS,
840 Evans AC, Giedd JN (2009) Differences in Genetic and Environmental
841 Influences on the Human Cerebral Cortex Associated With Development
842 During Childhood and Adolescence. Human Brain Mapping 30:163-174.
843 Li J, Ji L (2005) Adjusting multiple testing in multilocus analyses using the
844 eigenvalues of a correlation matrix. Heredity 95:221-227.

Liem F, Merillat S, Bezzola L, Hirsiger S, Philipp M, Madhyastha T, Jancke L (2015) Reliability and statistical power analysis of cortical and subcortical FreeSurfer metrics in a large sample of healthy elderly. NeuroImage 108:95-109.

Luby JL, Belden AC, Jackson JJ, Lessov-Schlaggar CN, Harms MP, Tillman R, Botteron K, Whalen D, Barch DM (2016) Early Childhood Depression and Alterations in the Trajectory of Gray Matter Maturation in Middle Childhood and Early Adolescence. Jama Psychiat 73:31-38.

Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K (2016) cluster: Cluster Analysis Basics and Extensions. In.

Maller JJ, Thomson RHS, Pannek K, Bailey N, Lewis PM, Fitzgerald PB (2014) Volumetrics relate to the development of depression after traumatic brain injury. Behavioural Brain Research 271:147-153.

Marcus DS, Harwell J, Olsen T, Hodge M, Glasser MF, Prior F, Jenkinson M, Laumann T, Curtiss SW, Van Essen DC (2011) Informatics and data mining tools and strategies for the human connectome project. Front Neuroinform 5:4.

Martin NG, Eaves LJ (1977) The genetical analysis of covariance structure. Heredity (Edinb) 38:79-95.

Matthews T, Danese A, Wertz J, Odgers CL, Ambler A, Moffitt TE, Arseneault L (2016) Social isolation, loneliness and depression in young adulthood: a behavioural genetic analysis. Social Psychiatry and Psychiatric Epidemiology 51:339-348.

McCarthy G, Puce A, Gore JC, Allison T (1997) Face-specific processing in the human fusiform gyrus. J Cogn Neurosci 9:605-610.

Medland SE, Jahanshad N, Neale BM, Thompson PM (2014) Whole-genome analyses of whole-brain data: working within an expanded search space. Nat Neurosci 17:791-800.

Mullins N et al. (2016) Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. Psychol Med 46:759-770.

Musliner KL, Seifuddin F, Judy JA, Pirooznia M, Goes FS, Zandi PP (2015) Polygenic risk, stressful life events and depressive symptoms in older adults: a polygenic score analysis. Psychol Med 45:1709-1720.

Najman JM, Heron MA, Hayatbakhsh MR, Dingle K, Jamrozik K, Bor W, O'Callaghan MJ, Williams GM (2008) Screening in early childhood for risk of later mental health problems: A longitudinal study. J Psychiatr Res 42:694-700.

Neale MC, Cardon LR (1992) Methodology for genetic studies of twins and families: Kluwer Academic Pub.

Nivard MG, Dolan CV, Kendler KS, Kan KJ, Willemsen G, van Beijsterveldt CE, Lindauer RJ, van Beek JH, Geels LM, Bartels M, Middeldorp CM, Boomsma DI (2015) Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. Psychol Med 45:1039-1049.

Panizzon MS, Fennema-Notestine C, Eyler LT, Jernigan TL, Prom-Wormley E, Neale M, Jacobson K, Lyons MJ, Grant MD, Franz CE, Xian H, Tsuang M, Fischl B, Seidman L, Dale A, Kremen WS (2009) Distinct genetic influences on cortical surface area and cortical thickness. Cereb Cortex 19:2728-2735.

Peyrot WJ, Milaneschi Y, Abdellaoui A, Sullivan PF, Hottenga JJ, Boomsma DI, Penninx BW (2014) Effect of polygenic risk scores on depression in childhood trauma. The British journal of psychiatry : the journal of mental science 205:113-119.

Polderman TJ, Benyamin B, de Leeuw CA, Sullivan PF, van Bochoven A, Visscher PM, Posthuma D (2015) Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nat Genet 47:702-709.

Posthuma D, Boomsma DI (2000) A note on the statistical power in extended twin designs. Behav Genet 30:147-158.

R Development Core Team (2012) R: A Language and Environment for Statistical Computing. In. Vienna, Austria: R Foundation for Statistical Computing.

Renteria ME, Hansell NK, Strike LT, McMahon KL, de Zubicaray GI, Hickie IB, Thompson PM, Martin NG, Medland SE, Wright MJ (2014) Genetic architecture of subcortical brain regions: common and region-specific genetic contributions. Genes Brain Behav 13:821-830.

Rhodes G (1993) Configural coding, expertise, and the right hemisphere advantage for face recognition. Brain Cogn 22:19-41.

Ronnegard L, Lee Y (2013) Exploring the potential of hierarchical generalized linear models in animal breeding and genetics. J Anim Breed Genet 130:415-416.

Ronnegard L, Shen X, Alam M (2010) hglm: A Package for Fitting Hierarchical Generalized Linear Models. The R Journal 2(2):20-28.

Rosenthal D (1963) A suggested conceptual framework. In: The Genain quadruplets (Rosenthal D, ed), p 505—516. New York: Basic Books.

Savitz JB, Drevets WC (2009) Imaging phenotypes of major depressive disorder: genetic correlates. Neuroscience 164:300-330.

Schmaal L, Yucel M, Ellis R, Vijayakumar N, Simmons JG, Allen NB, Whittle S (2017) Brain Structural Signatures of Adolescent Depressive Symptom Trajectories: A Longitudinal Magnetic Resonance Imaging Study. J Am Acad Child Adolesc Psychiatry 56:593-601 e599.

Schmaal L et al. (2015) Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. Mol Psychiatry.

Schmaal L et al. (2016) Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. Mol Psychiatry.

Sidak Z (1967) Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. J Am Stat Assoc 62:626-633.

Snyder HR (2013) Major Depressive Disorder Is Associated With Broad Impairments on Neuropsychological Measures of Executive Function: A Meta-Analysis and Review. Psychological Bulletin 139:81-132.

Strike LT, Couvy-Duchesne B, Hansell NK, Cuellar-Partida G, Medland SE, Wright MJ (2015) Genetics and Brain Morphology. Neuropsychol Rev 25:63-96.

Sullivan PF, Neale MC, Kendler KS (2000) Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry 157:1552-1562.

Therneau TM, Sinnwell J (2015) kinship2: Pedigree Functions. In.

Thompson PM et al. (2014) The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. Brain Imaging Behav 8:153-182.

943 Van Essen DC, Glasser MF, Dierker DL, Harwell J, Coalson T (2012a) Parcellations
944 and hemispheric asymmetries of human cerebral cortex analyzed on
945 surface-based atlases. Cereb Cortex 22:2241-2262.
946 Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, Consortium
947 WU-MH (2013) The WU-Minn Human Connectome Project: an overview.
948 NeuroImage 80:62-79.
949 Van Essen DC et al. (2012b) The Human Connectome Project: a data acquisition
950 perspective. NeuroImage 62:2222-2231.
951 Verweij KJ, Mosing MA, Zietsch BP, Medland SE (2012) Estimating heritability
952 from twin studies. Methods in molecular biology 850:151-170.
953 Vilgis V, Silk TJ, Vance A (2015) Executive function and attention in children and
954 adolescents with depressive disorders: a systematic review. Eur Child
955 Adoles Psy 24:365-384.
956 Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, Yang J (2017)
957 10 Years of GWAS Discovery: Biology, Function, and Translation. Am J
958 Hum Genet 101:5-22.
959 Wellcome Department of Imaging Neuroscience (2009) SPM8. In. London, UK:
960 Queen Square.
961 Whelan CD, Hibar DP, van Velzen LS, Zannas AS, Carrillo-Roa T, McMahon K,
962 Prasad G, Kelly S, Faskowitz J, deZubiracay G, Iglesias JE, van Erp TG, Frodl
963 T, Martin NG, Wright MJ, Jahanshad N, Schmaal L, Samann PG, Thompson
964 PM, Alzheimer's Disease Neuroimaging I (2016) Heritability and
965 reliability of automatically segmented human hippocampal formation
966 subregions. NeuroImage 128:125-137.
967 Whiteley AM, Warrington EK (1977) Prosopagnosia: a clinical, psychological, and
968 anatomical study of three patients. J Neurol Neurosurg Psychiatry 40:395-
969 403.
970 Winkler AM, Kochunov P, Blangero J, Almasy L, Zilles K, Fox PT, Duggirala R,
971 Glahn DC (2010) Cortical thickness or grey matter volume? The
972 importance of selecting the phenotype for imaging genetics studies.
973 NeuroImage 53:1135-1146.
974 Wright MJ, Martin NG (2004) Brisbane adolescent twin study: outline of study
975 methods and research projects. Australian Journal of Psychology 56:65-
976 78.
977 Zhang WN, Chang SH, Guo LY, Zhang KL, Wang J (2013) The neural correlates of
978 reward-related processing in major depressive disorder: a meta-analysis
979 of functional magnetic resonance imaging studies. J Affect Disord
980 151:531-539.
981

982

Figure 1: Silhouette coefficients of the clustering scenarios

Vertical dashed line corresponds to 12 clusters per hemisphere and measurement, which we used to parcellate the cortex.

Figure 2: Summary of genetic parcellation of the cortical thickness and surface area

Clusters are labelled 1L to 12L for the left hemisphere and 1R to 12R for the right one. Extended Data Figure 2-1, Figure 2-2, Figure 2-3, and Figure 2-4 describe the genetic clusters used in this analysis.

Figure 3: Non-linear associations between anxiety-depression and surface area of the genetic cluster 6R

The solid line represents the regression effect of the significant spline order. The dashed line combines the association for the linear, quadratic and cubic splines and the anxiety-depression SPHERE score and the SA of the 6R cluster. The y-axis corresponds to SA after removing the effect of the intercept and all other covariates. The vertical dashed bars indicate the 5, 25, 50, 75 and 95% quantiles of the scores distributions

a) Association between SA and cluster 6R SA and the SPHERE anxiety-depression score (QTIM sample). The reduction of surface area is observed in participants with SPHERE score greater than 2.9 (median score, see solid line: 3rd order natural spline). We interpret the increase of the (solid) regression line for SPHERE score below median as an artefact of plotting the significant spline order only. This effect is not observed when the 3 splines order are combined

b) Association between SA for cluster 6R and the DSM-oriented (left panel) and ASR syndrome based (right panel) anxiety-depression scores. A consistent reduction of surface area was observed in participants with ASR DSM-oriented scores between 4 and 19 (25% - 95% centiles; see solid line: 1st order natural spline). No significant association was found with the ASR syndrome based scale.

Figure 4: Vertex-wise phenotypic association of surface area with anxiety-depression scores within cluster 6R

Bottom view (top) and medial view (bottom) of the right hemisphere. The left panels show the cubic effect sizes (correlations) for each vertex of the 6R cluster for SA (94 vertex). The right panels show the significance (p-value) of the association in –log10 scale. Vertices represent the intersections of the triangular mesh reconstructed by FreeSurfer to model cortical surfaces.

a) Vertex-wise association between SA and anxiety-depression (SPHERE score) in the QTIM sample (N=833). The significance threshold of 4.2E-4 is reached for p-values greater than 3.4 in the log scale.

b) Vertex-wise association between SA and anxiety-depression (DSM-oriented score) in the HCP sample (N= 890).

Extended data

Figure 2-1: Description of vertices in each of left CT clusters.

Figure 2-2: Description of vertices in each of left SA clusters.

Figure 2-3: Description of vertices in each of right CT clusters.

Figure 2-4: Description of vertices in each of right SA clusters.

1032 Extended Data 1: R code used for all the analyses