

Investigating the role of the relaxin-3/RXFP3 system in neuropsychiatric disorders and metabolic phenotypes: a candidate gene approach

*Win Lee Edwin Wong^{1,2,3}, Ryan Arathimos⁴, †Cathryn M. Lewis^{4,5}, †Allan H. Young^{3,6},

†Gavin Stewart Dawe^{1,2,7}

1. Department of Pharmacology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

2. Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

3. Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

4. Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

5. Department of Medical and Molecular Genetics, Faculty of Life Sciences and Medicine, King's College London, London, United Kingdom

6. South London & Maudsley NHS Foundation Trust, Bethlem Royal Hospital, Monks Orchard Road, London, United Kingdom

7. Neurobiology Programme, Life Sciences Institute, National University of Singapore, Singapore

†Contributed equally as senior authors

*Correspondence to Win Lee Edwin Wong

Address: Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom, SE5 8AF.

Email: win_lee_edwin.wong@kcl.ac.uk, edwin.wong@u.nus.edu

Abstract

The relaxin-3/RXFP3 system has been implicated in the modulation of depressive- and anxiety-like behaviour in the animal literature; however, there is a lack of human studies investigating this signalling system. We seek to bridge this gap by leveraging the large UK Biobank study to retrospectively assess genetic risk variants linked with this neuropeptidergic system. Specifically, we conducted a candidate gene study in the UK Biobank to test for potential associations between a set of functional, candidate single nucleotide polymorphisms (SNPs) pertinent to relaxin-3 signalling, determined using *in silico* tools, and several outcomes, including depression, atypical depression, anxiety and metabolic syndrome. For each outcome, we used several rigorously defined phenotypes, culminating in subsample sizes ranging from 85,881 to 386,769 participants. Across all outcomes, there were no associations between any candidate SNP and any outcome phenotype, following corrections for multiple testing burden. Regression models comprising several SNPs per relevant candidate gene as exploratory variables further exhibited no prediction of outcome. Our findings corroborate conclusions from previous literature about the limitations of candidate gene approaches, even when based on firm biological hypotheses, in the domain of genetic research for neuropsychiatric disorders.

1 **1. Introduction**

2 Genetics plays a major role in neuropsychiatric conditions, with twin-based heritability
3 estimates of major depressive disorder (MDD) between 40%-50%^{1,2} and between 30%-
4 40%³ for anxiety disorders. Recent single nucleotide polymorphism (SNP)-based
5 heritability measures further corroborate a role for genetics, with MDD estimates ranging
6 between 0.102 – 0.162⁴; however, the full extent of genetic involvement remains poorly
7 defined. Much of the initial focus in this domain was on candidate gene studies, which
8 suggested that variation in several genes was putatively implicated in MDD and anxiety
9 disorders, like the short allele of serotonin transporter gene-linked polymorphic region⁵
10 and Catechol-O-methyltransferase⁶. However, over the last decade, advances in
11 computing power and sequencing technology, together with substantially larger study
12 sample sizes, have enabled a shift towards hypothesis-free genome-wide association
13 studies (GWAS), which have identified novel loci associated with MDD⁷⁻⁹ and anxiety
14 disorders¹⁰⁻¹², further underscoring the contribution of genetics to neuropsychiatric
15 disorders. While GWAS have become the standard approach in this field, there may still
16 be scope for candidate gene approaches, particularly for studies rooted in comprehensive
17 understanding of biological processes.

18 To this end, neuropeptidergic modulation has emerged as a key contributory factor to the
19 aetiology of several mental health disorders; for instance, much evidence implicates
20 corticotropin-releasing factor and Substance P in affective disorders and stress-
21 signaling¹³⁻¹⁵. Relaxin-3 is another such neuropeptide, identified in GABAergic neurons
22 of the rodent nucleus incertus approximately two decades ago^{16,17}. This protein binds to
23 several receptors, including its cognate receptor RXFP3, RXFP4, and to some extent
24 RXFP1, forming a signalling network that anatomical studies suggest is distributed across
25 the hypothalamus, hippocampus, amygdala and several other brain regions implicated in

1 neuropsychiatric behavioural alterations^{18,19}. Moreover, several literature reviews of
2 animal studies have highlighted the role of this relaxin-3/RXFP3 arousal system in
3 regulating behaviours akin to neuropsychiatric conditions and their endophenotypes¹⁹⁻²³;
4 pharmacological and genetic interventions in rodent models have resulted in altered
5 performance on several testing paradigms, as well as induced orexigenic and arousal
6 behaviour, underlining a modulatory role for relaxin-3 signalling.

7 Despite evidence for the role of relaxin-3/RXFP3 in affective disorders and associated
8 behaviours, there is a clear lack of human studies on this neuropeptidergic system. Our
9 recent systematic review identified only five studies²³, the majority of which performed
10 relatively poorly in methodological assessment. Limitations notwithstanding, this review
11 identified a retrospective candidate gene study in a cohort of antipsychotic-treated
12 patients²⁴, which reported several significant associations between candidate SNPs at the
13 RLN3, RXFP3, and RXP4 genes and metabolic phenotypes, including hypertension,
14 dyslipidaemia, and hypercholesterolaemia. Despite these preliminary findings, no follow-
15 up work has been conducted to explore these candidate polymorphisms in a large, more
16 robust cohort, or for any other neuropsychiatric phenotypes linked to the relaxin-
17 3/RXFP3 system.

18 The UK Biobank resource²⁵ offers an opportunity to build on prior work and investigate
19 these variants in a large-scale study. Participants in the UK Biobank have been deeply
20 phenotyped, allowing for the derivation of several outcomes retrospectively. We therefore
21 conducted a candidate gene study in UK Biobank to explore functional SNPs, determined
22 using *in silico* prediction tools, at genes thought to be relevant to relaxin-3 signalling,
23 including RLN3, RXFP3, RXFP4, RXFP1, and RLN2. Our outcomes of interest in this
24 association study include depression, atypical depression, anxiety and metabolic
25 syndrome, each of which has previously been linked to the relaxin-3/RXFP3 across the

1 pertinent literature of animal studies. Several definitions of outcomes were assessed, to
2 ensure comprehensive evaluation of our candidate variants.

3 **2. Methods**

4 ***2.1 Study population***

5 The data used in this study was collected by the UK Biobank, a study of approximately
6 500,000 participants across the United Kingdom²⁵. Genotyping data was available for
7 488,171 individuals, with DNA extracted from whole blood samples then assayed using
8 the Affymetrix UK BiLEVE Axiom Array or the Affymetrix UK Biobank Axiom Array²⁶;
9 genotype imputation was done using IMPUTE4 with Haplotype Reference Consortium
10 (HRC) data as the main reference panel²⁶. The UK Biobank has extensive phenotypic
11 data, including self-report data and various measures from the baseline assessment,
12 electronic health records, and further online data collection, including a mental health
13 questionnaire²⁷. Using these data, we created several definitions of depression subtypes,
14 anxiety, and metabolic parameters, as described below.

15 Prior to analysis, exclusions were applied for each phenotype. Across all depression and
16 anxiety phenotypes, participants who reported taking antipsychotics during baseline
17 assessment interview, or with one of several confounding illnesses (mania, bipolar,
18 schizophrenia, or psychosis) based on ICD-10 diagnostic codes from hospital inpatient
19 data and self-reported professional diagnosis, were excluded. Controls for depression
20 phenotypes were further excluded if they were on antidepressant treatment or fulfilled
21 case definition for the other depression phenotypes, while controls for anxiety phenotypes
22 were further excluded if on anxiolytic treatment or fulfilling case definitions for the other
23 anxiety phenotypes. For metabolic phenotypes, controls were excluded when they had

1 prescriptions of drugs corresponding to any metabolic condition comprising the metabolic
2 syndrome definition (for example statins for hypertension).

3 With respect to the genetic quality control, individuals were excluded from this study
4 when they were outliers in the genetic data for heterozygosity, had a variant call rate of
5 <98%, had a mismatch between their reported sex and genetic sex, or were not recorded
6 as White British on initial assessment centre visit. Moreover, individuals were excluded
7 based on genetic relatedness, whereby only one member in groups of related individuals
8 was used in this study; this was achieved by initial exclusion of individuals using kinship
9 coefficients derived with the KING software, then subsequent addition of one member
10 per related group, selecting individuals with a genetic relatedness of <0.025 with any
11 other participant.

12 This study was conducted under the UK Biobank application number 16577. The UK
13 Biobank study was conducted under generic approval from the NHS National Research
14 Ethics Service (approval letter dated 17th June 2011, Ref 11/NW/0382). All participants
15 gave full informed written consent.

16 ***2.2 Candidate SNP selection***

17 The process used to select candidate SNPs is outlined in **Figure 1**. The National Center
18 for Biotechnology Information Single Nucleotide Polymorphism Database²⁸ was searched
19 for all unmerged SNPs in the genes RLN3, RXFP3, RXFP4, RLN2, and RXFP1 with a
20 minor allele frequency of >0.01, using Entrez²⁹ to interrogate SNP records mapped to
21 these genes. All SNPs were then passed through the Combined Annotation-Dependent
22 Depletion (CADD) and Genome Wide Annotation of Variants (GWAVA) *in silico* tools
23 for identifying variants with potentially functional effects, retaining SNPs above the
24 recommended C-score cut-off of >10 for CADD³⁰ and above the region, TSS, and

1 unmatched scores cut-off of >0.5 for GWAVA³¹. Finally, the remaining SNPs were
2 assessed for linkage disequilibrium using the LDlink web tool³², with an r^2 cut-off
3 of >0.8 used to prune SNPs in strong linkage disequilibrium based on European reference
4 population data from the 1000 Genomes Project³³. This culminated in a final set of 12
5 SNPs across RLN3 (rs74400983, rs1982632, rs78161395, rs6511905), RXFP3
6 (rs9292519, rs171631), RXFP1 (rs62351166, rs7695640, rs11100192), and RLN2
7 (rs72703633, rs11793069, rs72499174). We supplemented this with an additional 3 SNPs
8 from RXFP3 (rs42868, rs7702361) and RXFP4 (rs11264422), which were reported to be
9 associated with metabolic parameters in a previous case-control study²⁴.

10 ***2.3 Phenotype definitions***

11 Several phenotypes were defined for the outcomes of depression, atypical depression,
12 anxiety, metabolic syndrome, and the distinct metabolic conditions that comprise
13 metabolic syndrome.

14 Depression was defined with six different phenotypes: a “broad” phenotype, an “ICD10-
15 coded” phenotype, a “lifetime” phenotype, a CIDI phenotype, a “PHQ-9 definition”
16 phenotype and a “PHQ-9 cut-off” phenotype. The “broad” phenotype, “ICD10-coded”
17 phenotype and “lifetime” phenotype roughly correspond to definitions previously used in
18 a UK Biobank GWAS of depression⁸. In brief, the broad phenotype was defined using
19 self-reported help-seeking behaviour from participant touchscreen responses, while
20 ICD10-coded phenotype was defined using linked hospital admission records from the
21 UK Biobank. The lifetime phenotype was defined using a returned field from Smith et
22 al.³⁴, which was derived using touchscreen responses for both help-seeking behaviour and
23 the presence of low mood or anhedonia. The CIDI phenotype of depression²⁷ was derived
24 using participant responses to the online follow-up mental health questionnaire, based on

1 depression-relevant questions from the World Health Organisation’s (WHO) Composite
2 International Diagnostic Interview Short Form (CIDI-SF)³⁵, modified to consider lifetime
3 history, where participants asked about their worst ever depression episode had to report
4 either anhedonia or depressed mood in their lifetime and meet at least four of eight
5 lifetime depression symptoms for caseness. The two PHQ-9 phenotypes were also
6 derived using responses to the online follow-up mental health questionnaire, from the
7 Patient Health Questionnaire 9-question version for current depression symptoms³⁶ – the
8 “PHQ-9 definition” was based on a major depression diagnosis on the PHQ-9, while the
9 “PHQ-9 cut-off” phenotype was based on a cut-off score of ≥ 10 ^{36,37}.

10 The atypical depression subtype was defined with three different phenotypes, for
11 participants who responded to the online follow-up mental health questionnaire. Caseness
12 for atypical depression was defined as an indication of depression, based on our
13 previously defined CIDI phenotype, PHQ-9 definition phenotype, or PHQ-9 cut-off
14 definition phenotype, as well as indicating both hypersomnia and weight gain during their
15 worst episode of depression on the online follow-up mental health questionnaire³⁸.

16 Anxiety was defined with three different phenotypes: an ICD10-coded phenotype, a
17 “lifetime disorder” phenotype and a “GAD-7 cut-off” phenotype, the latter two of which
18 mirror the definitions used in a previous GWAS of anxiety¹². The ICD10-coded
19 phenotype of anxiety was defined using linked hospital admission records for the main
20 five anxiety disorders. The lifetime disorder phenotype was derived with a combination
21 of self-reported lifetime professional diagnosis for the main five anxiety disorders and
22 responses to the online follow-up mental health questionnaire, based on anxiety-relevant
23 questions from the CIDI. The GAD-7 cut-off phenotype was defined using responses to
24 answers to questions based on the Generalised Anxiety Disorder Assessment (GAD-7),
25 using a cut-off score of ≥ 10 to define caseness³⁹.

1 Metabolic syndrome was assessed using the National Heart, Lung and Blood Institute
2 definition⁴⁰, where caseness was defined as having ≥ 3 of the following risk factors: a
3 waist circumference of ≥ 102 cm for males or ≥ 88 cm for females, hypertension (or
4 antihypertensive drug treatment), hypertriglyceridaemia (or drug treatment for elevated
5 triglycerides), hyperglycaemia (or drug treatment for elevated glucose) and low HDL
6 cholesterol levels (or drug treatment for reduced HDL cholesterol). Each of these risk
7 factors, with the exception of waist circumference, were also assessed as individual
8 phenotypes.

9 The full criteria used to derive these phenotype variables in the UK Biobank, including
10 all relevant field codes used, are detailed in **Supplementary Tables 1, 2 and 3**.

11 **2.4 Statistical analyses**

12 We performed logistic regression to test for association between each candidate gene and
13 outcome. These regression models were adjusted for age, age², sex, genotyping batch,
14 testing centre, and the first six European ancestry principal components (to control for
15 population structure). For each primary outcome of interest (i.e., MDD, anxiety, atypical
16 depression and metabolic syndrome), we corrected for multiple testing using false
17 discovery rate (FDR) across the various phenotype definitions – 6 definitions for MDD, 3
18 definitions for anxiety, 3 definitions for atypical depression and 5 definitions for
19 metabolic syndrome – and compared these q-values against a threshold of 0.01, more
20 stringent than the standard threshold of 0.05 to account for the multiple SNPs tested. To
21 ensure that there were no methodological errors in phenotypic definitions or exclusions,
22 similar regression testing was also conducted for three independent, control SNPs at
23 chromosome 7 from a previous GWAS analysis done for MDD in the UK Biobank:
24 rs3807865, rs1554505, and rs5011432. All three of these polymorphisms were previously

1 significantly associated with a specific MDD outcome phenotype, such that these are
2 solid control points of reference to verify that the correct quality control procedures were
3 implemented in the present study.

4 We also conducted a sensitivity analysis for each candidate gene, in which every
5 candidate SNP at a particular gene (for RLN3, RXFP3, RXFP1, and RLN2) was an
6 explanatory variable in a single multivariable regression model, adjusted for covariates as
7 described above. This was then compared to an analogous regression model with no
8 candidate SNP variables present, using a likelihood ratio test. This was conducted for all
9 outcome phenotypes, with multiple testing accounted for as in the individual SNP
10 analyses, with an FDR correction across the number of phenotype sub-definitions then
11 comparison to an alpha level of 0.01.

12 **3. Results**

13 ***3.1 Study demographics***

14 In total, there were six phenotypes for depression, three phenotypes for anxiety, three
15 phenotypes for atypical depression, and five phenotypes for metabolic indications. Study
16 demographics for case and control groups in each phenotype definition are outlined in
17 **Table 1**. The number of cases and controls within each phenotype varied heavily, as each
18 was derived using a different set of responses in UK Biobank. Across all phenotype
19 definitions, the mean age of participants at baseline ranged from 51.6 years to 59.2 years.

20 ***3.2 SNP associations with depression and anxiety phenotypes***

21 Each candidate SNP was individually investigated for all phenotypes, across all outcomes
22 of interest. When corrected for the number of tests conducted, there were no associations
23 between any individual candidate SNP and any of the six depression phenotypes (**Figure**
24 **2**). Before correction for multiple testing, rs74400983 (unadjusted $p = 0.0232$) and

1 rs78161395 (unadjusted $p = 0.0428$) at RLN3, as well as rs62351166 (unadjusted $p =$
2 0.0388) at RXFP1 were nominally linked with the broad phenotype of depression;
3 however, these did not withstand FDR corrections. No candidate SNPs were associated
4 with any of the three atypical depression phenotypes following corrections for multiple
5 test burden (**Figure 3**), with rs42868 from RXFP3 closest to demonstrating any
6 significant association (unadjusted $p = 0.0297$).

7 In addition to evaluating these candidate SNPs, we also tested the SNPs rs3807865,
8 rs1554505, and rs5011432 at chromosome 7, which have previously been associated with
9 single depression phenotype definitions (that mirror our definitions) in a 2018 GWAS of
10 MDD in the UK Biobank⁸. The associations from this GWAS were largely reproduced in
11 this study, with rs3807865 still significantly associated with the broad phenotype
12 depression ($\beta = 0.033$, $SE = 0.005$, $p = 4.74 \times 10^{-11}$), rs5011432 significantly associated
13 with the lifetime phenotype depression ($\beta = 0.048$, $SE = 0.011$, $p = 1.96 \times 10^{-5}$), and
14 rs1554505 significantly associated with the ICD10-coded depression phenotype ($\beta = -$
15 0.047, $SE = 0.013$, $p = 2.04 \times 10^{-4}$).

16 There were also no associations between any individual candidate SNP and any of the
17 three anxiety phenotypes (**Figure 4**). Before applying our relatively liberal multiple test
18 correction procedure, rs11264422 from RXFP4 was preliminarily associated with both
19 our probable (unadjusted $p = 0.00318$) and ICD10-coded (unadjusted $p = 0.0106$)
20 phenotype for anxiety; however, following FDR corrections, these associations were not
21 significant against our threshold of 0.01.

22 Detailed information for all relevant regression analyses pertinent to depression, atypical
23 depression, and anxiety phenotypes are summarised in **Supplementary Tables 4, 5, and**
24 **6**.

1

2

3 ***3.3 SNP associations with metabolic phenotypes***

4 Candidate SNPs were also tested for associations with metabolic syndrome and four of its
5 constituent phenotypes: hypertension, low HDL cholesterol levels (dyslipidaemia),
6 hypertriglyceridaemia, and hyperglycaemia. After FDR adjustment for the number of
7 sub-phenotypes tested, there were no associations between any of the candidate SNPs and
8 any of the metabolic phenotypes (**Figure 5**). Associations between rs72499174 from
9 RLN2 and two phenotypes – hypertension ($q = 0.0355$) and hypertriglyceridaemia ($q =$
10 0.0704) – were the two closest to the significance cut-offs used. The candidate SNPs
11 rs42868 (unadjusted $p = 0.0289$) and rs11793069 (unadjusted $p = 0.0353$) were also
12 provisionally linked with low HDL cholesterol levels (dyslipidaemia) and hypertension,
13 respectively. However, these associations were not statistically significant after
14 corrections for multiple testing. Detailed information for all pertinent regression analyses
15 involving metabolic phenotypes is summarised in **Supplementary Table 7**.

16 ***3.4 Sensitivity analyses by gene***

17 Association analyses for each gene were also conducted to additionally investigate our
18 candidate SNPs of interest. All SNPs at a particular gene were entered as explanatory
19 variables in a single multivariate logistic regression model, for each outcome phenotype,
20 which was then compared to a null model with no SNPs and only covariates as predictors.
21 Combined SNPs at RLN3, RXFP3, RLN2, and RXFP1 were tested, with RXFP4 not
22 included as there was only one pertinent candidate SNP at this gene. Across all
23 depression, anxiety, atypical depression and metabolic phenotypes, none of the regression
24 models comprising multiple SNP predictors significantly improved fit on the likelihood

1 ratio test, compared to the null models, following correction for multiple testing burden
2 (**Supplementary Table 8, 9, 10**). This further underscores the lack of observed
3 associations between our candidate SNPs and the outcomes of interest in the UK Biobank.

4 **4. Discussion**

5 In this analysis of a large population-based cohort, a set of functional SNPs from several
6 candidate genes relevant to relaxin-3 signalling were investigated for associations with
7 depression, anxiety, atypical depression, and metabolic syndrome, corresponding to a
8 putative link reported in the animal literature. However, across all outcomes, there were
9 no significant associations between any individual candidate SNP and any of the outcome
10 phenotypes, suggesting that common genetic variation in the relaxin-3/RXFP3 system
11 makes little contribution to liability for any of our outcomes of interest. This lack of
12 associations arose in spite of the relatively lenient procedure undertaken to correct for
13 multiple testing, in which an FDR correction was applied across the several phenotype
14 definitions per outcome and compared against a threshold of 0.01, without detailed
15 consideration for the number of SNPs tested. An FDR approach was adopted to enable
16 the detection of any possible signals in this study, yet no significant associations were
17 identified. Furthermore, the present results cannot be attributed to errors in the data
18 preparation and analysis pipeline undertaken in this study, as previously identified
19 GWAS hits in the UK Biobank were reproduced in a sensitivity analysis. A lack of
20 statistical power cannot be ruled out as a potential factor explaining the absence of any
21 significant associations, particularly for certain phenotypes, like the PHQ-9 cut-off
22 phenotype of depression, for which there were only 5,720 cases; however, sample sizes
23 across the different phenotypic subsamples ranged between 85,881 and 386,769,
24 rendering this analysis relatively well powered, especially when compared to the
25 traditional candidate gene literature in this domain. Moreover, we used between three to

1 six phenotypic definitions under each umbrella outcome, a very comprehensive approach
2 to ensure that any findings were not specific to a restricted outcome sub-definition; for
3 instance, the six phenotypes of depression help to circumvent some of the diagnostic
4 heterogeneity^{41,42} of the disorder. The consistent lack of associations across the different
5 phenotypes for each definition, along with the reasonably high power of this UK Biobank
6 analysis, further reinforce the notion that our results are attributable to the absence of any
7 direct link between genetic variations at the relaxin-3/RXFP3 system and the presence of
8 neuropsychiatric and metabolic outcomes.

9 These findings conflict with the only prior candidate gene study in the relaxin-3/RXFP3
10 field, which reported associations between several candidate SNPs and miscellaneous
11 metabolic parameters²⁴. The only variant peripherally pertinent in both studies was
12 rs42868 at RXFP3, which was associated with diabetes and hypercholesterolaemia in the
13 prior study and nominally associated with low HDL cholesterol levels (dyslipidaemia) in
14 this analysis. There are several potential explanations for the discrepancies between these
15 two studies. Firstly, the previous study was conducted using a cohort of 419
16 antipsychotic-treated patients, a niche sample with fewer participants than our present
17 investigation of the UK Biobank. Secondly, the metabolic outcome definitions differed
18 between studies: the prior study used BMI, clinical records of diabetes and
19 hypercholesterolaemia, as opposed to the present study, which used waist circumference,
20 fasting glucose levels, and HDL cholesterol levels; our definitions were derived based on
21 definitions previously used for other analyses of metabolic syndrome in the UK
22 Biobank^{43,44}, and are in line with the standardised definition from the National Heart,
23 Lung, and Blood Institute⁴⁰. Thirdly, the candidate polymorphisms investigated in the two
24 studies did not overlap completely: the prior study made use of the HapMap programme⁴⁵,
25 whereas we leveraged several *in silico* prediction tools that were developed only after the

1 initial 2011 study and explored the additional relaxin family genes RLN2 and RXFP1⁴⁶.
2 Nonetheless, we supplemented our set of SNPs with the prior study's putatively reported
3 significant candidate SNPs, to ensure our investigation encompassed all prior reported
4 associations. Perhaps the biggest contributor to the difference in results was the statistical
5 methodology adopted. While both studies employed multivariable regressions, there were
6 differences in the external variables controlled for; our study controlled for potential
7 genotyping differences, as well as population structure (using principal components), but
8 did not take into account duration of antipsychotic medication, which was particularly
9 pertinent in the prior study given the nature of its cohort. Furthermore, our study
10 corrected for multiple testing burden using FDR, whereas the prior study did not use any
11 multiple test corrections. It is important to note that had the prior study corrected for
12 multiple testing, using for instance a lenient Bonferroni approach across phenotypes or
13 SNPs, no SNP would have been significantly associated. While there is no clear
14 consensus on the appropriate approaches to correcting for numerous tests^{47,48}, a complete
15 lack of correction is not robust considering the number of polymorphisms and outcomes
16 tested. As such, the prior findings are likely to have been false positives generated by
17 chance, and our present results are more reflective of the relationship (or lack thereof)
18 between genetic variation relevant to relaxin-3 signalling and metabolic outcomes.

19 More broadly, these results provide additional support for a shift away from candidate
20 gene studies in psychiatric research. An extensive and well-powered investigation from
21 Border et al.⁴⁹ evaluated 18 commonly studied candidate genes from the last three
22 decades for associations with several depressive phenotypes, finding that there was little
23 evidence of any relationship between the candidate polymorphisms and depression
24 liability. This mirrors conclusions from other rigorous studies and reviews; for example,
25 van de Weijer et al.⁵⁰ highlighted a lack of support for candidate gene associations with

1 well-being, a construct heavily related to neuropsychiatric disorders. Our focused
2 candidate gene study on the relaxin-3 neuropeptidergic system corroborate these
3 conclusions, underlining the importance of proceeding cautiously before undertaking a
4 candidate gene approach. Genetic contributions towards multifactorial disorders, like
5 depression, anxiety, or metabolic syndrome, are extremely complex and polygenic in
6 nature, with our study further demonstrating the minuscule effect sizes that individual
7 genes have on these phenotypes. While candidate gene approaches may still be relevant
8 in particular scenarios, our findings underscore why there has been a shift towards
9 GWAS approaches in understanding the genetic underpinnings of complex diseases.

10 It is important to recognise that a lack of significant associations in this candidate gene
11 study does not rule out a role for the relaxin-3/RXFP3 system in MDD, atypical
12 depression, anxiety or metabolic syndrome. The Border et al. study⁴⁹ explored several
13 key neurotransmitters and peptides that are generally accepted to be involved in the
14 aetiology of MDD, yet also revealed a lack of associations between polymorphisms at the
15 pertinent genes and presence of MDD. For example, SNPs from the HTR2A gene, which
16 codes for the serotonin 5-HT_{2A} receptor, were not significantly associated with MDD.
17 This finding does not, however, invalidate prior evidence outlining 5-HT_{2A} receptor
18 involvement in MDD, with good evidence for the distribution of this protein across brain
19 areas related to MDD and several preclinical studies linking this receptor with depressive-
20 like phenotypes^{51,52}. The apparent inconsistency between null genetic associations and
21 putative pathophysiological involvement is not altogether surprising as these complex,
22 polygenic diseases are the culmination of dynamic interactions between several
23 peptidergic systems, such that variation at only a small subset of genes may not be
24 particularly informative, highlighted by the small effect sizes. Moreover, these candidate
25 gene approaches only explore genetic differences at the DNA level, and do not consider

1 epigenetic factors, which could very feasibly, in this context, alter the expression levels of
2 the various relaxin-3/RXFP3-relevant proteins.

3 The present investigation possesses several strengths. This is the first study to evaluate
4 candidate SNPs pertinent to the relaxin-3/RXFP3 system in the context of depression,
5 anxiety and atypical depression, and the first to assess a relationship with metabolic
6 parameters using a rigorous statistical approach. The relatively large sample size for the
7 analyses is a particular strength relative to the traditional candidate gene literature in this
8 domain. Furthermore, we were cognisant to define several phenotypes for each of our
9 outcomes of interest; the consistency in our null results across the various phenotypic
10 definitions provides some assurance that our findings are reflective of a holistic definition
11 for MDD, atypical depression, anxiety, and metabolic syndrome.

12 There were also several important limitations of note in this study. Firstly, our candidate
13 gene selection methodology is rooted in the use of *in silico* tools, which have certain
14 limitations in predicting variant deleteriousness. We employed two different tools, each
15 based on different statistical methodology and principles, to help overcome the lack of
16 absolute certainty in predicting the effects of variants. The cut-offs used in this study
17 were recommended by the creators of each tool and helped ensure we attained a
18 reasonable number of candidate SNPs to evaluate. There are also several limitations to
19 our outcome phenotype definitions, though the use of proxies to define our disease states
20 of interest was unavoidable. Several of the phenotype definitions for depression and
21 anxiety are based on participant self-reported information, which can be prone to biases
22 that may arise from elements of recall and social desirability⁵³. The broad definition of
23 depression in particular may additionally capture an overly general phenotype, given the
24 frequency of comorbidity between depression and anxiety disorders. Our CIDI phenotype
25 of definition also did not stringently mirror the diagnostic criteria from the WHO CIDI

1 assessment, as one of the nine pertinent depression symptoms (psychomotor agitation or
2 retardation) was not available in the online mental health follow-up questionnaire.
3 Consequently, our CIDI phenotype altered the standard definition of ≥ 5 of 9 symptoms
4 to ≥ 4 of 8 symptoms. While our definitions for atypical depression have been used in
5 previous analyses of the UK Biobank³⁸, they are also of ambiguous reliability, as they
6 were based on answers to questions about hypersomnia and weight gain, which do not
7 comprise the full spectrum of atypical symptoms that characterise this disorder⁵⁴⁻⁵⁶. The
8 limitations inherent in each of these definitions provides further rationale for the wide
9 range of outcome phenotype definitions used in this analysis, including definitions that
10 mirror the validated PHQ-9 and GAD-7 questionnaires^{36,39}. Finally, it is important to
11 acknowledge that the UK Biobank is not a robust representation of the general UK
12 population, with higher participation rates across certain demographics⁵⁷; this may limit
13 the generalisability of these findings.

14 In summary, this candidate gene study revealed that candidate functional polymorphisms
15 at RLN3, RXFP3, RXFP4, RLN2, and RXFP1 had no significant effects on the outcomes
16 of MDD, atypical depression, anxiety, and metabolic syndrome. This lack of associations
17 was consistent across the many phenotypic definitions investigated, ensuring
18 comprehensive analysis, and further confirmed in sensitivity analyses exploring several
19 candidate polymorphisms simultaneously. While the relaxin-3/RXFP3 system may still
20 be involved in the pathophysiology of these diseases, this is unlikely to be reflected in
21 common genetic variation at the DNA level. At a broader level, our findings support prior
22 conclusions for prudent consideration and interpretation of candidate gene studies for
23 neuropsychiatric or metabolic conditions, especially given their complex underpinnings,.

1 **Data availability**

2 The genetic and phenotypic data underlying the results presented in this study are
3 available from the UK Biobank subject to standard procedures, described at the following
4 webpage: <https://www.ukbiobank.ac.uk/enable-your-research>. Detailed information about
5 the UK Biobank data used in this research study can be found at the following webpage:
6 <https://www.ukbiobank.ac.uk/enable-your-research/about-our-data/>.

7 **Conflicts of interest**

8 Gavin S. Dawe is employed by the National University of Singapore and is a co-inventor
9 on patent applications on relaxin-3 B chain stapled peptide agonist and antagonists at
10 RXFP3: “Stapled Peptide Agonists and Their Use in Treatment of Behavioural Disorders”
11 Singapore Patent Application No. 10201709379P filed on 14 November 2017; and
12 “Stapled relaxin-3 B chain peptide antagonists” Singapore Patent Application No.
13 10201904291Y filed on 13 May 2019.

14 Allan H. Young declares the following competing interests. Employed by King’s College
15 London; Honorary Consultant SLaM (NHS UK). Paid lectures and advisory boards for
16 the following companies with drugs used in affective and related disorders: AstraZeneca,
17 Eli Lilly, Lundbeck, Sunovion, Servier, Livanova, Janssen, Allegan, Bionomics,
18 Sumitomo Dainippon Pharma, COMPASS. Consultant to Johnson & Johnson. Consultant
19 to Livanova. Received honoraria for attending advisory boards and presenting talks at
20 meetings organised by LivaNova. Principal Investigator in the Restore-Life VNS registry
21 study funded by LivaNova. Principal Investigator on ESKETINTRD3004: “An Open-
22 label, Long-term, Safety and Efficacy Study of Intranasal Esketamine in Treatment-
23 resistant Depression”. Principal Investigator on “The Effects of Psilocybin on Cognitive
24 Function in Healthy Participants”. Principal Investigator on “The Safety and Efficacy of

1 Psilocybin in Participants with Treatment-Resistant Depression (P-TRD)". UK Chief
2 Investigator for Novartis MDD study MIJ821A12201. Grant funding (past and present):
3 NIMH (USA); CIHR (Canada); NARSAD (USA); Stanley Medical Research Institute
4 (USA); MRC (UK); Wellcome Trust (UK); Royal College of Physicians (Edin); BMA
5 (UK); UBC-VGH Foundation (Canada); WEDC (Canada); CCS Depression Research
6 Fund (Canada); MSFHR (Canada); NIHR (UK). Janssen (UK). No shareholdings in
7 pharmaceutical companies.

8 Cathryn Lewis sits on the Scientific Advisory Board for Myriad Neuroscience, and has
9 received speaker fees from SYNLAB.

10 **Funding statement**

11 Win Lee Edwin Wong's research is supported by the National University of Singapore
12 President's Graduate Fellowship. Gavin S. Dawe's research is supported by the Ministry
13 of Education, Singapore, under its Academic Research Fund Tier 3 Award (MOE2017-
14 T3-1-002), and by the National Medical Research Council, Singapore, under its NMRC
15 NUHS Centre Grant (NMRC/CG/M009/2017_NUH/NUHS). Allan H. Young's
16 independent research is funded by the National Institute for Health and Care Research
17 (NIHR) Maudsley Biomedical Research Centre at South London and Maudsley NHS
18 Foundation Trust and King's College London. Cathryn M. Lewis' independent research
19 is funded by the National Institute for Health and Care Research (NIHR) Maudsley
20 Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and
21 King's College London. Ryan Arathimos' research was part funded by the National
22 Institute for Health Research (NIHR) Maudsley Biomedical Research Centre at South
23 London and Maudsley NHS Foundation Trust and King's College London for the

1 duration of this research project. The views expressed are those of the author(s) and not
2 necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

3 **Acknowledgements**

4 This research has been conducted using the UK Biobank Resource under Application
5 Number 16577. The authors wish to thank Jonathan R. I. Coleman and Gerome Breen for
6 their valuable contributions in support of this study. In addition, we are grateful to all UK
7 Biobank staff and volunteers. Finally, the authors acknowledge use of the research
8 computing facility at King's College London, Rosalind (<https://rosalind.kcl.ac.uk>).

References

- 1 Fernandez-Pujals, A. M. *et al.* Epidemiology and Heritability of Major Depressive Disorder, Stratified by Age of Onset, Sex, and Illness Course in Generation Scotland: Scottish Family Health Study (GS:SFHS). *PLoS One* **10**, e0142197, doi:10.1371/journal.pone.0142197 (2015).
- 2 Kendler, K. S., Ohlsson, H., Lichtenstein, P., Sundquist, J. & Sundquist, K. The Genetic Epidemiology of Treated Major Depression in Sweden. *Am J Psychiatry* **175**, 1137-1144, doi:10.1176/appi.ajp.2018.17111251 (2018).
- 3 Hettema, J. M., Neale, M. C. & Kendler, K. S. A review and meta-analysis of the genetic epidemiology of anxiety disorders. *Am J Psychiatry* **158**, 1568-1578, doi:10.1176/appi.ajp.158.10.1568 (2001).
- 4 Jermy, B. S., Glanville, K. P., Coleman, J. R. I., Lewis, C. M. & Vassos, E. Exploring the genetic heterogeneity in major depression across diagnostic criteria. *Mol Psychiatry* **26**, 7337-7345, doi:10.1038/s41380-021-01231-w (2021).
- 5 Greenberg, B. D., McMahon, F. J. & Murphy, D. L. Serotonin transporter candidate gene studies in affective disorders and personality: promises and potential pitfalls. *Mol Psychiatry* **3**, 186-189, doi:10.1038/sj.mp.4000415 (1998).
- 6 Hettema, J. M. *et al.* Catechol-O-methyltransferase contributes to genetic susceptibility shared among anxiety spectrum phenotypes. *Biol Psychiatry* **64**, 302-310, doi:10.1016/j.biopsych.2008.03.014 (2008).
- 7 Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci* **22**, 343-352, doi:10.1038/s41593-018-0326-7 (2019).
- 8 Howard, D. M. *et al.* Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat Commun* **9**, 1470, doi:10.1038/s41467-018-03819-3 (2018).
- 9 Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet* **50**, 668-681, doi:10.1038/s41588-018-0090-3 (2018).
- 10 Levey, D. F. *et al.* Reproducible Genetic Risk Loci for Anxiety: Results From approximately 200,000 Participants in the Million Veteran Program. *Am J Psychiatry* **177**, 223-232, doi:10.1176/appi.ajp.2019.19030256 (2020).
- 11 Meier, S. M. *et al.* Genetic Variants Associated With Anxiety and Stress-Related Disorders: A Genome-Wide Association Study and Mouse-Model Study. *JAMA Psychiatry* **76**, 924-932, doi:10.1001/jamapsychiatry.2019.1119 (2019).
- 12 Purves, K. L. *et al.* A major role for common genetic variation in anxiety disorders. *Mol Psychiatry* **25**, 3292-3303, doi:10.1038/s41380-019-0559-1 (2020).
- 13 Arborelius, L., Owens, M. J., Plotsky, P. M. & Nemeroff, C. B. The role of corticotropin-releasing factor in depression and anxiety disorders. *J Endocrinol* **160**, 1-12, doi:10.1677/joe.0.1600001 (1999).
- 14 Ebner, K. & Singewald, N. The role of substance P in stress and anxiety responses. *Amino Acids* **31**, 251-272, doi:10.1007/s00726-006-0335-9 (2006).
- 15 Waters, R. P. *et al.* Evidence for the role of corticotropin-releasing factor in major depressive disorder. *Neurosci Biobehav Rev* **58**, 63-78, doi:10.1016/j.neubiorev.2015.07.011 (2015).
- 16 Bathgate, R. A. *et al.* Human relaxin gene 3 (H3) and the equivalent mouse relaxin (M3) gene. Novel members of the relaxin peptide family. *J Biol Chem* **277**, 1148-1157, doi:10.1074/jbc.M107882200 (2002).
- 17 Ma, S. *et al.* Relaxin-3 in GABA projection neurons of nucleus incertus suggests widespread influence on forebrain circuits via G-protein-coupled receptor-135 in the rat. *Neuroscience* **144**, 165-190, doi:10.1016/j.neuroscience.2006.08.072 (2007).

- 18 Ganella, D. E., Ma, S. & Gundlach, A. L. Relaxin-3/RXFP3 Signaling and Neuroendocrine Function - A Perspective on Extrinsic Hypothalamic Control. *Front Endocrinol (Lausanne)* **4**, 128, doi:10.3389/fendo.2013.00128 (2013).
- 19 Ma, S., Smith, C. M., Blasiak, A. & Gundlach, A. L. Distribution, physiology and pharmacology of relaxin-3/RXFP3 systems in brain. *Br J Pharmacol* **174**, 1034-1048, doi:10.1111/bph.13659 (2017).
- 20 Kumar, J. R. *et al.* Relaxin' the brain: a case for targeting the nucleus incertus network and relaxin-3/RXFP3 system in neuropsychiatric disorders. *Br J Pharmacol* **174**, 1061-1076, doi:10.1111/bph.13564 (2017).
- 21 Smith, C. M., Ryan, P. J., Hosken, I. T., Ma, S. & Gundlach, A. L. Relaxin-3 systems in the brain--the first 10 years. *J Chem Neuroanat* **42**, 262-275, doi:10.1016/j.jchemneu.2011.05.013 (2011).
- 22 Smith, C. M. *et al.* Relaxin-3/RXFP3 networks: an emerging target for the treatment of depression and other neuropsychiatric diseases? *Front Pharmacol* **5**, 46, doi:10.3389/fphar.2014.00046 (2014).
- 23 Wong, W. L. E., Dawe, G. S. & Young, A. H. The putative role of the relaxin-3/RXFP3 system in clinical depression and anxiety: A systematic literature review. *Neurosci Biobehav Rev* **131**, 429-450, doi:10.1016/j.neubiorev.2021.09.028 (2021).
- 24 Munro, J. *et al.* Relaxin polymorphisms associated with metabolic disturbance in patients treated with antipsychotics. *J Psychopharmacol* **26**, 374-379, doi:10.1177/0269881111408965 (2012).
- 25 Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* **12**, e1001779, doi:10.1371/journal.pmed.1001779 (2015).
- 26 Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-209, doi:10.1038/s41586-018-0579-z (2018).
- 27 Davis, K. A. S. *et al.* Mental health in UK Biobank - development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. *BJPsych Open* **6**, e18, doi:10.1192/bjo.2019.100 (2020).
- 28 Sayers, E. W. *et al.* Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **49**, D10-D17, doi:10.1093/nar/gkaa892 (2021).
- 29 Maglott, D., Ostell, J., Pruitt, K. D. & Tatusova, T. Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res* **39**, D52-57, doi:10.1093/nar/gkq1237 (2011).
- 30 Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J. & Kircher, M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* **47**, D886-D894, doi:10.1093/nar/gky1016 (2019).
- 31 Ritchie, G. R., Dunham, I., Zeggini, E. & Flicek, P. Functional annotation of noncoding sequence variants. *Nat Methods* **11**, 294-296, doi:10.1038/nmeth.2832 (2014).
- 32 Machiela, M. J. & Chanock, S. J. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31**, 3555-3557, doi:10.1093/bioinformatics/btv402 (2015).
- 33 Genomes Project, C. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65, doi:10.1038/nature11632 (2012).
- 34 Smith, D. J. *et al.* Prevalence and characteristics of probable major depression and bipolar disorder within UK biobank: cross-sectional study of 172,751 participants. *PLoS One* **8**, e75362, doi:10.1371/journal.pone.0075362 (2013).
- 35 Kessler, R. C. *et al.* Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMH-CIDI). *Int J Methods Psychiatr Res* **13**, 122-139, doi:10.1002/mpr.169 (2004).

- 36 Kroenke, K., Spitzer, R. L. & Williams, J. B. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* **16**, 606-613, doi:10.1046/j.1525-1497.2001.016009606.x (2001).
- 37 Manea, L., Gilbody, S. & McMillan, D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. *CMAJ* **184**, E191-196, doi:10.1503/cmaj.110829 (2012).
- 38 Brailean, A., Curtis, J., Davis, K., Dregan, A. & Hotopf, M. Characteristics, comorbidities, and correlates of atypical depression: evidence from the UK Biobank Mental Health Survey. *Psychol Med* **50**, 1129-1138, doi:10.1017/S0033291719001004 (2020).
- 39 Spitzer, R. L., Kroenke, K., Williams, J. B. & Lowe, B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* **166**, 1092-1097, doi:10.1001/archinte.166.10.1092 (2006).
- 40 Institute, N. H. L. a. B. *What Is Metabolic Syndrome?*, <<https://www.nhlbi.nih.gov/health/metabolic-syndrome#:~:text=Metabolic%20syndrome%20is%20a%20group,also%20called%20insulin%20resistance%20syndrome.>> (2022).
- 41 Cai, N., Choi, K. W. & Fried, E. I. Reviewing the genetics of heterogeneity in depression: operationalizations, manifestations and etiologies. *Hum Mol Genet* **29**, R10-R18, doi:10.1093/hmg/ddaa115 (2020).
- 42 Goldberg, D. The heterogeneity of "major depression". *World Psychiatry* **10**, 226-228, doi:10.1002/j.2051-5545.2011.tb00061.x (2011).
- 43 Dregan, A. *et al.* Associations Between Depression, Arterial Stiffness, and Metabolic Syndrome Among Adults in the UK Biobank Population Study: A Mediation Analysis. *JAMA Psychiatry* **77**, 598-606, doi:10.1001/jamapsychiatry.2019.4712 (2020).
- 44 Lind, L. Genome-Wide Association Study of the Metabolic Syndrome in UK Biobank. *Metab Syndr Relat Disord* **17**, 505-511, doi:10.1089/met.2019.0070 (2019).
- 45 Manolio, T. A. & Collins, F. S. The HapMap and genome-wide association studies in diagnosis and therapy. *Annu Rev Med* **60**, 443-456, doi:10.1146/annurev.med.60.061907.093117 (2009).
- 46 Halls, M. L., van der Westhuizen, E. T., Bathgate, R. A. & Summers, R. J. Relaxin family peptide receptors--former orphans reunite with their parent ligands to activate multiple signalling pathways. *Br J Pharmacol* **150**, 677-691, doi:10.1038/sj.bjp.0707140 (2007).
- 47 Sethuraman, A. *et al.* Continued misuse of multiple testing correction methods in population genetics-A wake-up call? *Mol Ecol Resour* **19**, 23-26, doi:10.1111/1755-0998.12969 (2019).
- 48 Streiner, D. L. Best (but oft-forgotten) practices: the multiple problems of multiplicity--whether and how to correct for many statistical tests. *Am J Clin Nutr* **102**, 721-728, doi:10.3945/ajcn.115.113548 (2015).
- 49 Border, R. *et al.* No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression Across Multiple Large Samples. *Am J Psychiatry* **176**, 376-387, doi:10.1176/appi.ajp.2018.18070881 (2019).
- 50 van de Weijer, M. P., Pelt, D. H. M., de Vries, L. P., Baselmans, B. M. L. & Bartels, M. A Re-evaluation of Candidate Gene Studies for Well-Being in Light of Genome-Wide Evidence. *J Happiness Stud* **23**, 3031-3053, doi:10.1007/s10902-022-00538-x (2022).
- 51 Celada, P., Puig, M., Amargos-Bosch, M., Adell, A. & Artigas, F. The therapeutic role of 5-HT_{1A} and 5-HT_{2A} receptors in depression. *J Psychiatry Neurosci* **29**, 252-265 (2004).
- 52 Guiard, B. P. & Di Giovanni, G. Central serotonin-2A (5-HT_{2A}) receptor dysfunction in depression and epilepsy: the missing link? *Front Pharmacol* **6**, 46, doi:10.3389/fphar.2015.00046 (2015).
- 53 Althubaiti, A. Information bias in health research: definition, pitfalls, and adjustment methods. *J Multidiscip Healthc* **9**, 211-217, doi:10.2147/JMDH.S104807 (2016).

- 54 Quitkin, F. M. Depression With Atypical Features: Diagnostic Validity, Prevalence, and Treatment. *Prim Care Companion J Clin Psychiatry* **4**, 94-99, doi:10.4088/pcc.v04n0302 (2002).
- 55 Stewart, J. W., McGrath, P. J., Quitkin, F. M. & Klein, D. F. DSM-IV depression with atypical features: is it valid? *Neuropsychopharmacology* **34**, 2625-2632, doi:10.1038/npp.2009.99 (2009).
- 56 Thase, M. E. Recognition and diagnosis of atypical depression. *J Clin Psychiatry* **68 Suppl 8**, 11-16 (2007).
- 57 Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol* **186**, 1026-1034, doi:10.1093/aje/kwx246 (2017).

Tables

Table 1: Summary of study demographics for case and control groups in each of the phenotypic definitions, across all outcomes.

Phenotype	Total	Number of cases	Number of controls	%Male	Mean age (standard deviation)
Broad depression	367445	131219	236226	53.9%	56.8 (8.01)
ICD10-coded depression	288937	19307	269630	53.0%	57.3 (7.98)
Lifetime depression	85881	23271	62610	52.8%	57.2 (8.01)
CIDI depression	111786	33449	78337	56.1%	56.1 (7.67)
PHQ-9 definition depression	103236	5470	97766	54.5%	56.2 (7.69)
PHQ-9 cutoff depression	105170	5720	99450	54.9%	56.1 (7.7)
CIDI atypical depression	79999	1662	78337	51.3%	56.6 (7.67)
PHQ-9 definition atypical depression	98192	426	97766	54.4%	56.3 (7.66)
PHQ-9 cutoff atypical depression	100011	561	99450	54.4%	56.3 (7.67)
ICD10-coded anxiety	321083	16151	304932	54.3%	57.2 (7.97)
Lifetime disorder anxiety	101836	21045	80791	56.1%	56.2 (7.68)
GAD-7 cutoff anxiety	114667	4472	110195	56.1%	56.1 (7.67)
Metabolic syndrome	305742	104629	201113	53.4%	56.7 (8.02)
Hypertension	335730	235005	100725	54.0%	56.6 (8.05)
Low HDL cholesterol	336469	65720	270749	53.5%	56.7 (8.02)
Hyperglycaemia	331715	25573	306142	53.7%	56.7 (8.03)
Hypertriglyceridaemia	326585	147594	178991	54.7%	56.2 (8.05)

Figures

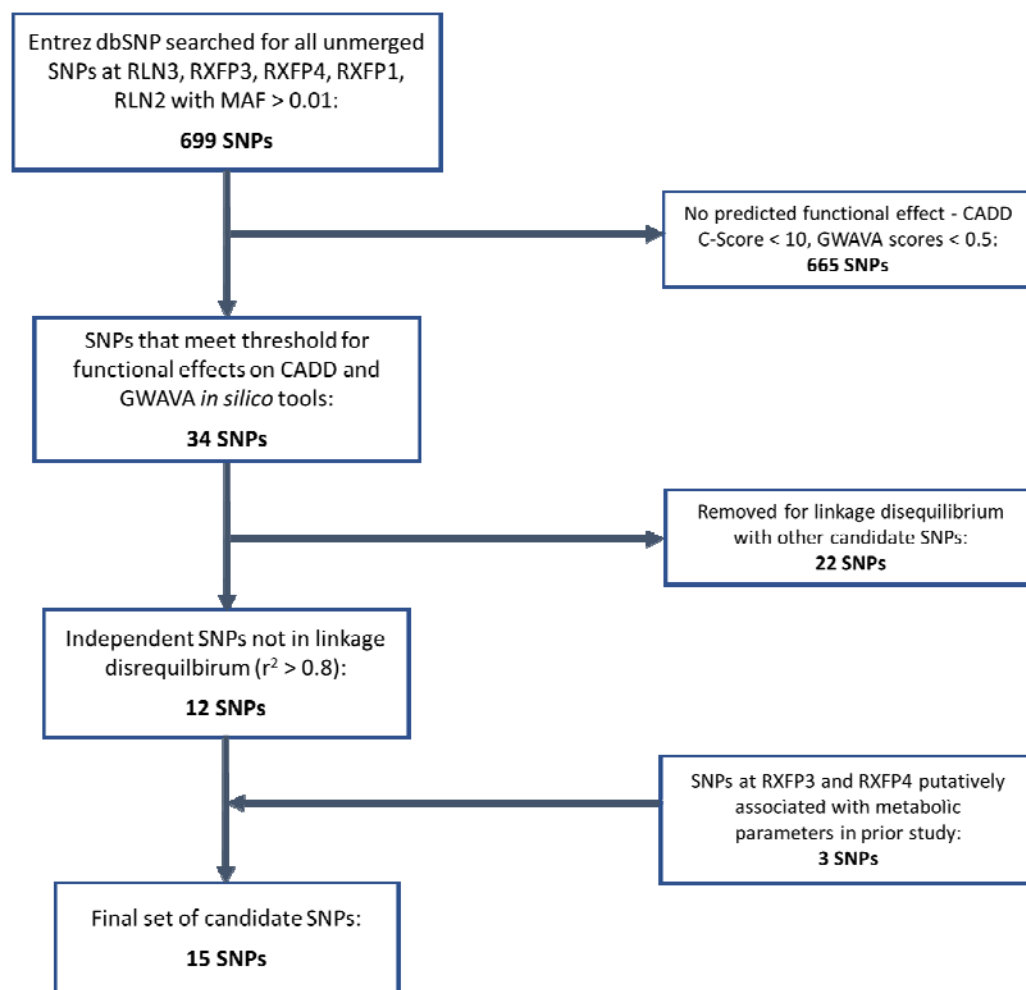


Figure 1: Flowchart outlining the process used to select candidate single nucleotide polymorphisms in this study.

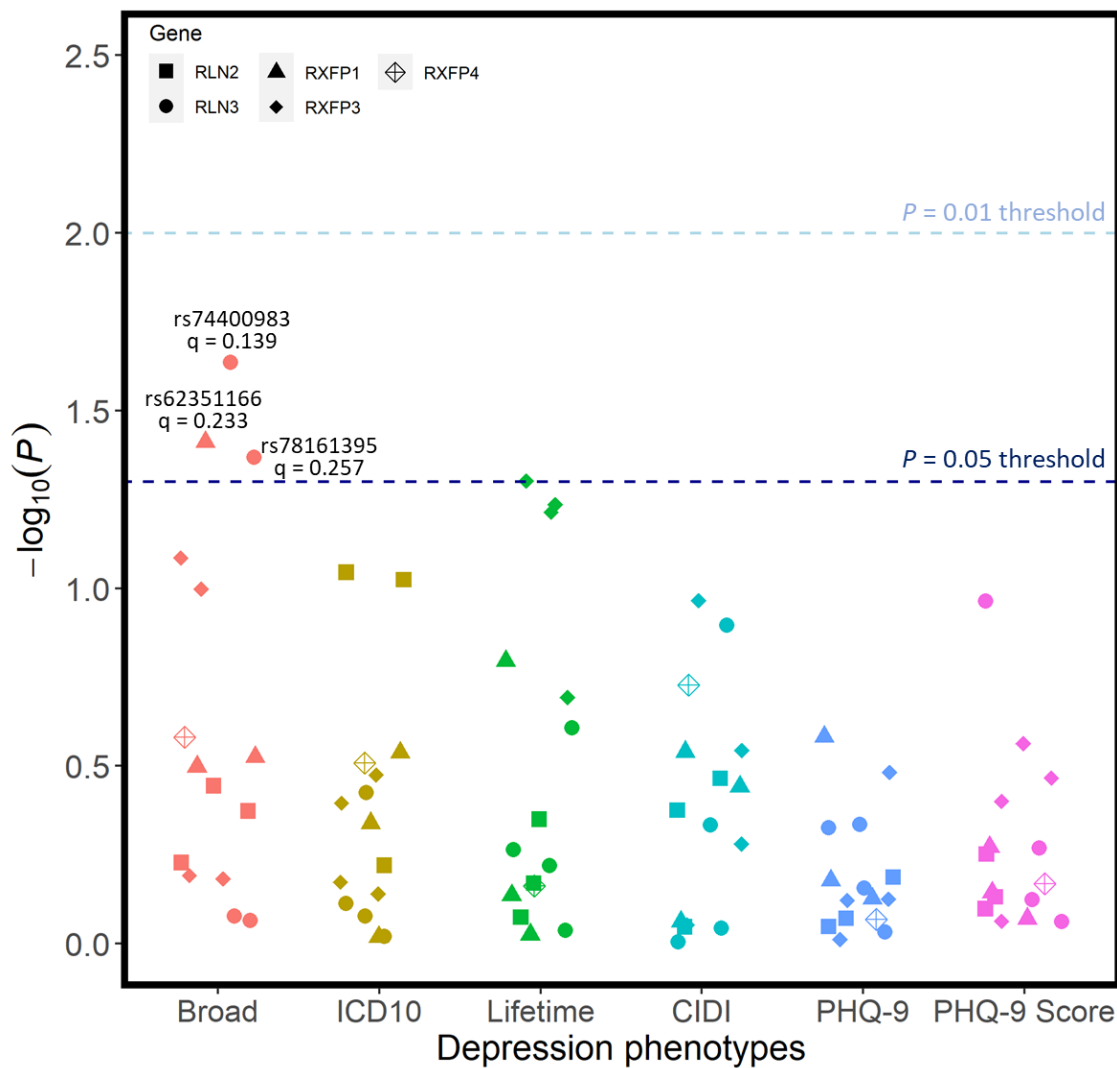


Figure 2: Unadjusted p-values for association tests between all SNPs and the 6 phenotypic definitions of depression. Regression models were adjusted for age, age², sex, genotyping batch, testing centre, and the first six European ancestry principal components. SNPs that were nominally associated with a phenotype are annotated with their corresponding q-value following FDR correction – there were no statistically significant associations.

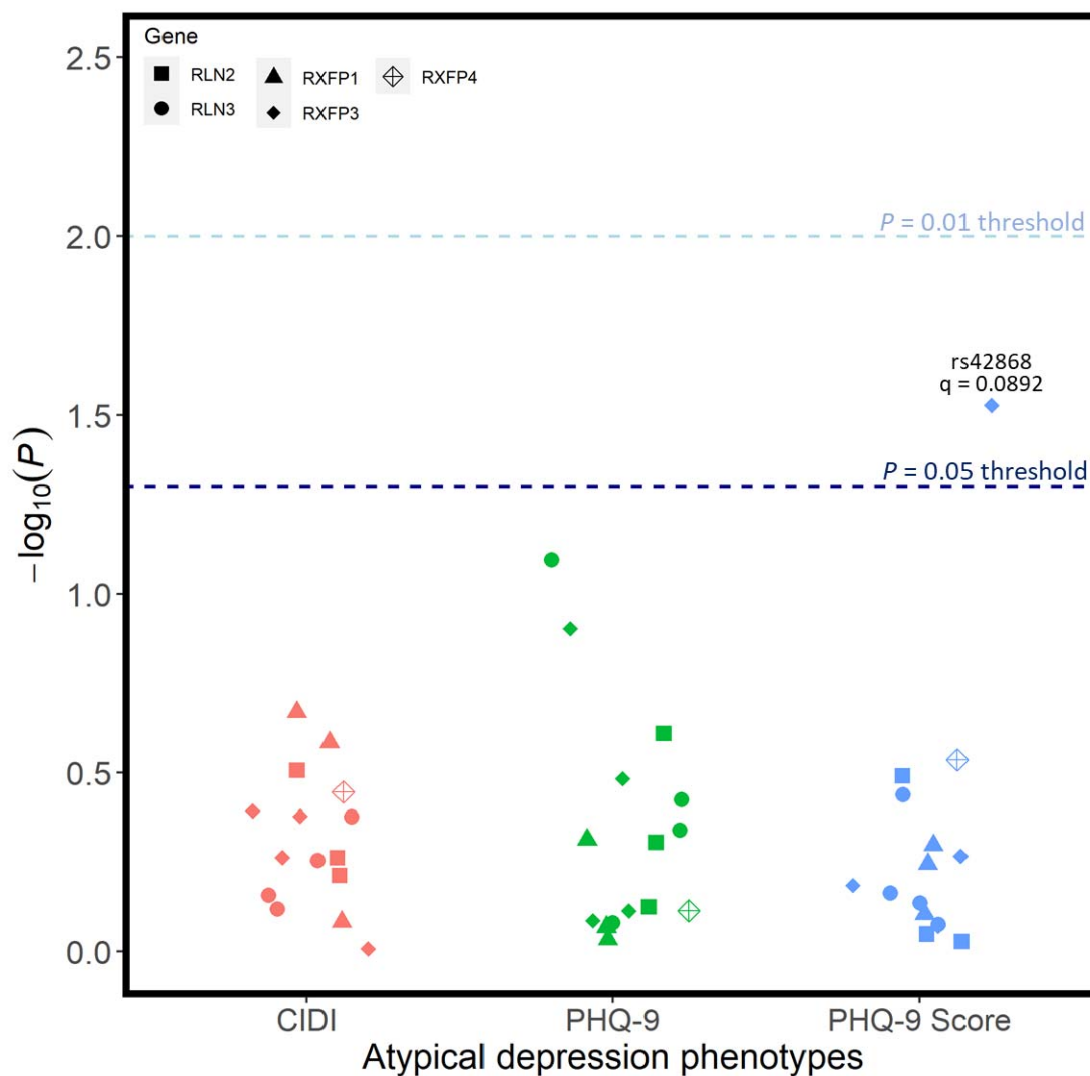


Figure 3: Unadjusted p-values for association tests between all SNPs and the 3 phenotypic definitions of atypical depression. Regression models were adjusted for age, age², sex, genotyping batch, testing centre, and the first six European ancestry principal components. SNPs that were nominally associated with a phenotype are annotated with their corresponding q-value following FDR correction – there were no statistically significant associations.

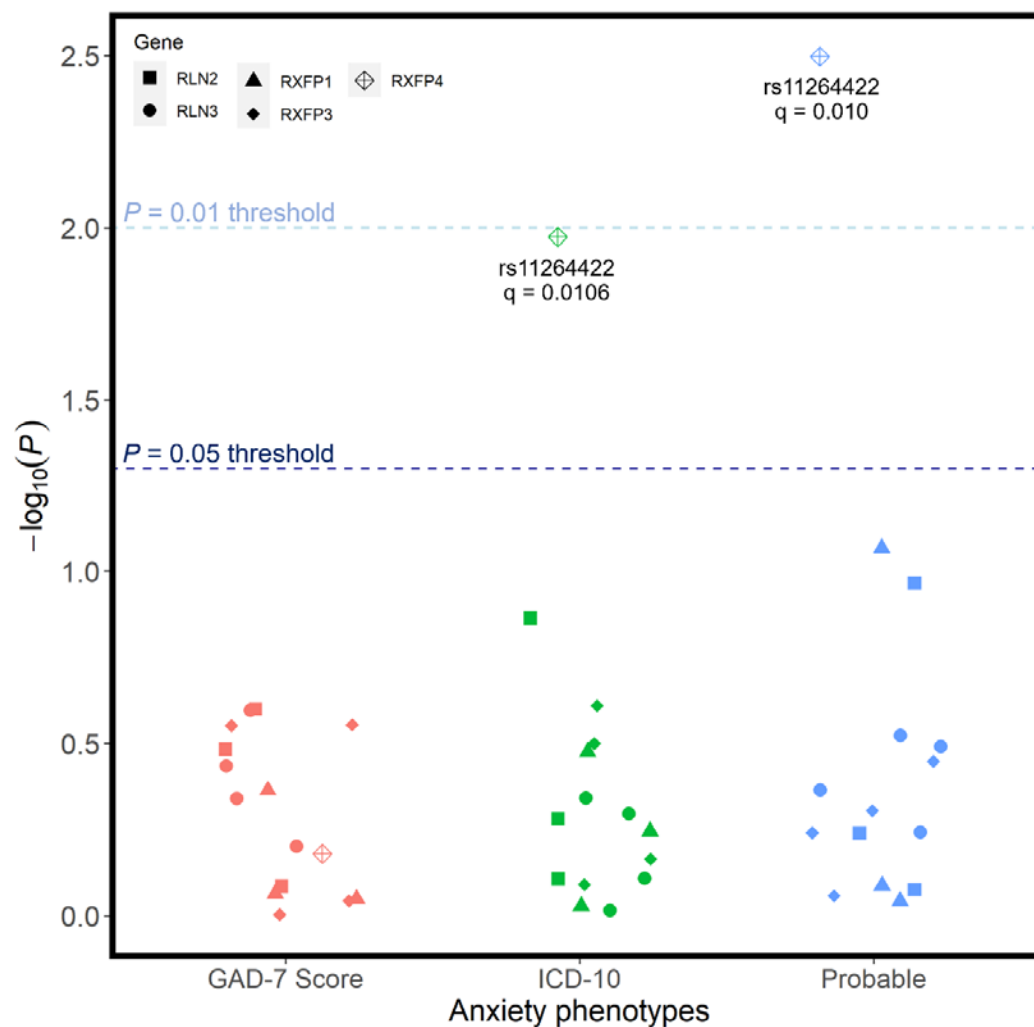


Figure 4: Unadjusted p-values for association tests between all SNPs and the 3 phenotypic definitions of anxiety disorders. Regression models were adjusted for age, age², sex, genotyping batch, testing centre, and the first six European ancestry principal components. SNPs that were nominally associated with a phenotype are annotated with their corresponding q-value following FDR correction – there were no statistically significant associations.

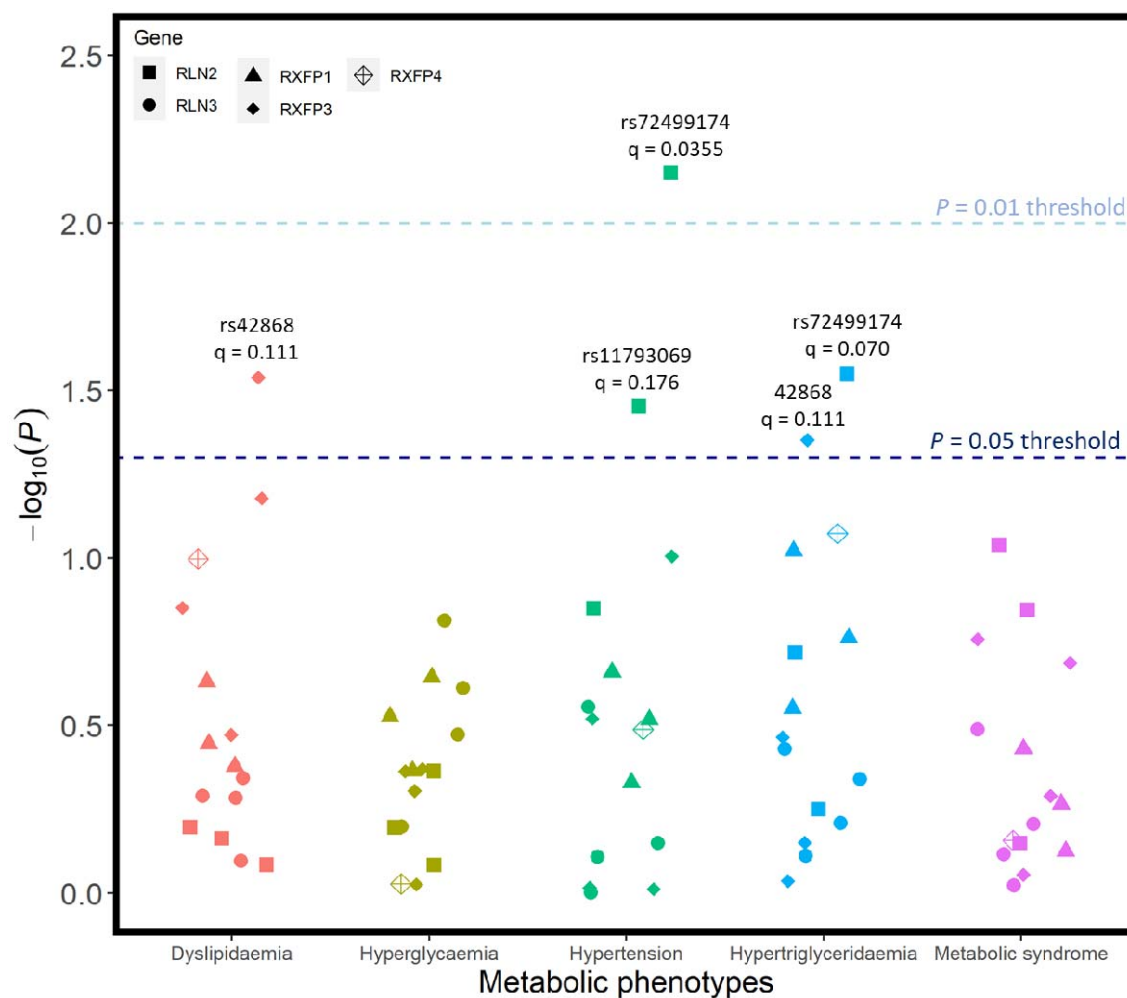


Figure 5: Unadjusted p-values for association tests between all SNPs and metabolic syndrome, as well as 4 sub-outcomes that comprise metabolic syndrome. Regression models were adjusted for age, age², sex, genotyping batch, testing centre, and the first six European ancestry principal components. SNPs that were nominally associated with a phenotype are annotated with their corresponding q-value following FDR correction – there were no statistically significant associations.

Entrez dbSNP searched for all unmerged SNPs at RLN3, RXFP3, RXFP4, RXFP1, RLN2 with MAF > 0.01:

699 SNPs

No predicted functional effect - CADD C-Score < 10, GWAVA scores < 0.5:

665 SNPs

SNPs that meet threshold for functional effects on CADD and GWAVA *in silico* tools:

34 SNPs

Removed for linkage disequilibrium with other candidate SNPs:

22 SNPs

Independent SNPs not in linkage disequilibrium ($r^2 > 0.8$):

12 SNPs

SNPs at RXFP3 and RXFP4 putatively associated with metabolic parameters in prior study:

3 SNPs

Final set of candidate SNPs:

15 SNPs

