1   **Genome-wide association study of circadian rhythmicity in 71 500 UK Biobank participants and**

2   **polygenic association with mood instability.**

3   Amy Ferguson[1], Laura M. Lyall[1], Joey Ward[1], Rona J. Strawbridge[1,2], Breda Cullen[1], Nicholas Graham[1],

4   Claire L. Niedzwiedz[1], Keira J.A. Johnston[1], Daniel MacKay[1], Stephany M. Biello[3], Jill P. Pell[1], Jonathan

5   Cavanagh[1], Andrew M. McIntosh[4], Aiden Doherty[5,6], Mark E.S. Bailey[7], Donald M. Lyall[1], Cathy A.

6   Wyse[8], Daniel J. Smith[1]

7   [1]Institute of Health & Wellbeing, University of Glasgow, Scotland, UK.

8   [2]Department of Medicine Solna, Karolinska Institute, Stockholm, Sweden.

9   [3]Institute of Neuroscience and Psychology, University of Glasgow, Scotland, UK.

10   [4]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Scotland, UK.

11   [5]Big Data Institute, Nuffield Department of Population Health, BHF Centre of Research Excellence,

12   University of Oxford, Oxford, UK.

13   [6]Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford,

14   Oxford, UK.

15   [7]School of Life Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow,

16   Scotland, UK.

17   [8]Department of Molecular and Cellular Therapeutics, Irish Centre for Vascular Biology, Royal College

18   of Surgeons in Ireland (RCSI), Dublin, Ireland.

19

20 **Abstract**

21 **Background:**

22 Circadian rhythms are fundamental to health and are particularly important for mental wellbeing.

23 Disrupted rhythms of rest and activity are recognised as risk factors for major depressive disorder

24 and bipolar disorder.

25 **Methods:**

26 We conducted a genome-wide association study (GWAS) of low relative amplitude (RA), an objective

27 measure of circadian rhythmicity derived from the accelerometer data of 71 500 UK Biobank

28 participants. Polygenic risk scores (PRS) for low RA were used to investigate potential associations

29 with psychiatric phenotypes.

30 **Outcomes:**

31 Two independent genetic loci were associated with low RA, within genomic regions for Neurofascin

32 (*NFASC*) and Solute Carrier Family 25 Member 17 (*SLC25A17*). A secondary GWAS of RA as a

33 continuous measure identified a locus within Meis Homeobox 1 (*MEIS1*). There were no significant

34 genetic correlations between low RA and any of the psychiatric phenotypes assessed. However, PRS

35 for low RA was significantly associated with mood instability across multiple PRS thresholds (at PRS

36 threshold 0·05: OR=1·02, 95% CI=1·01-1·02, p=9·6x10^{-5}), and with major depressive disorder (at PRS

37 threshold 0·1: OR=1·03, 95% CI=1·01-1·05, p=0·025) and neuroticism (at PRS threshold 0·5:

38 Beta=0·02, 95% CI=0·007-0·04, p=0·021).

39 **Interpretation:**

40 Overall, our findings contribute new knowledge on the complex genetic architecture of circadian

41 rhythmicity and suggest a putative biological link between disrupted circadian function and mood

42 disorder phenotypes, particularly mood instability, but also major depressive disorder and

43 neuroticism.

46 **Keywords:**

47 circadian rhythmicity; mood instability; relative amplitude; gwas; polygenic risk score

48 **Introduction**

49 Circadian rhythms are variations in physiology and behaviour that recur approximately every 24

50 hours.[1] They include rhythms of body temperature, hormone release, activity, concentration, mood,

51 eating and sleeping. Circadian rhythmicity is co-ordinated centrally by the suprachiasmatic nucleus in

52 the anterior hypothalamus[2] and plays a fundamental role in homeostasis and the maintenance of

53 both physical and mental wellbeing.[2,3] Disruption to circadian rhythmicity is associated with a range

54 of adverse health outcomes, including cardiovascular disease, obesity, diabetes and some cancers,[4–6]

55 as well as increased risk for major depressive disorder (MDD) and bipolar disorder (BD).[7–9]

56

57 Circadian rhythmicity is regulated by both exogenous environmental stimuli ("zeitgebers") and by

58 genetic factors.[10] Several core circadian clock genes are involved in autoregulatory

59 transcription/translational feedback loops that maintain cell-cycle function.[11] However, the control

60 of circadian rhythms is likely to be polygenic, with regulatory genes and pathways still to be

61 identified.[3,12]

62

63 To date, the most commonly used measure of circadian phenotypes has been subjectively-reported

64 chronotype, defined as an individual's preference for morning or evening wakefulness and activity.[13]

65 Evening chronotype is more likely to be associated with adverse health outcomes.[3,14–16] Recently,

66 genome wide association studies (GWAS) of chronotype, self-reported sleep duration, and

67 accelerometer-derived sleep traits have identified several independent genetic loci previously

68 implicated in the regulation of circadian function (including *PER2*, *PER3*, *RSG16*, *AK5*, *FBXL13*), in

69 addition to novel associated genetic loci.[17–22]

70 However, as a subjective measure, chronotype can be vulnerable to response bias. It may also have

71 inconsistent associations with more objective measures of circadian rhythmicity.[23] In a recent study

72 of over 91 488 individuals in the UK Biobank cohort, we derived objective measures of rest-activity

73 rhythmicity from accelerometer data.[24] We found that low relative amplitude (RA), a measure of an

74 individual's rest-activity rhythm, was associated with several mood disorder phenotypes.[24] We now

75 extend this work by conducting a GWAS of low RA in the largest sample known to date from UK

76 Biobank. We also assess the degree of genetic correlation between low RA and several psychiatric

77 phenotypes, including attention deficit hyperactivity disorder (ADHD), BD, MDD, mood instability,

78 post-traumatic stress disorder (PTSD), schizophrenia, anxiety and insomnia. Further, within a group

79 of UK Biobank participants who were not part of the primary GWAS study (up to 141 000 individuals),

80 we test for association between increased polygenic risk score (PRS) for low RA and mood disorder

81   phenotypes (specifically BD, MDD, generalised anxiety disorder (GAD), mood instability and

82   neuroticism).

83

84   **Methods**

85   ***Participants and ethical approval***

86   Over 502 000 United Kingdom (UK) residents aged 37-73 years were recruited to the UK Biobank

87   cohort from 2006-2010. At one of 22 assessment centres across the UK, participants completed a

88   range of lifestyle, demographic, health, mood, cognitive and physical assessments and

89   questionnaires.[25] Here, we used data from 91 448 participants who also provided accelerometer

90   data that passed quality control (details below) and who had available data on all covariates

91   included within fully and/or partially adjusted models for use in the GWAS. UK Biobank obtained

92   informed consent from all participants and this study was conducted under generic approval from

93   the NHS National Research Ethics Service (approval letter dated 13 May 2016, Ref 16/NW/0274) and

94   under UK Biobank approvals for applications 12761 (PI Cathy Wyse; accelerometer data for use in

95   GWAS) and 6553 (PI Daniel Smith; correlations with psychiatric traits).

96   ***Accelerometry data collection and pre-processing***

97   In 2013, 240 000 UK Biobank participants were invited to wear an accelerometer for seven days as

98   part of a physical activity monitoring investigation.[26] Of these, 103 720 (43%) accepted and returned

99   the accelerometer to UK Biobank after use. Participants received a wrist-worn Axivity AX3 triaxial

100   accelerometer in the post and were asked to wear the device on their dominant wrist continuously

101   for seven days, while continuing with their normal activities. At the end of the seven-day period,

102   participants were instructed to return the accelerometer to UK Biobank using a prepaid envelope.

103   Accelerometers were calibrated to local gravity. Devices recorded data at a sampling rate of 94-

104   104Hz, and data were resampled to 100Hz offline. Periods where no data were recorded for >1s

105   were coded as missing, and machine noise was removed using a Butterworth low-pass filter (cut-off

106   20Hz). Raw activity intensity data were combined into five second epochs. Further details on data

107   pre-processing[26] are available from UK Biobank at

108   http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=131600.

109   ***Circadian rest-activity rhythmicity (relative amplitude)***

110   From the summary five second epoch data, a measure of relative amplitude (RA) was calculated

111   using Clocklab Version 6 (Actimetrics). This accelerometer-derived activity measure has

112   demonstrated reliability and validity.[27] RA is used commonly as a non-parametric measure of rest-

113   activity rhythm amplitude. It is defined as the relative difference between the most active

4

114 continuous 10-hour period (M10) and the least active continuous 5-hour period (L5) in an average

115 24-hour period, using the formula below:[28]

$$RA = \frac{(M10 - L5)}{(M10 + L5)}$$

116 M10 is the mean activity during the continuous 10 hour period containing maximum activity in each

117 24 hour recording period (midnight to midnight). L5 is the mean activity for the corresponding 5

118 hour period containing the minimum activity within the same recording period. For each individual,

119 the RA data point was the mean RA value across all included 24-hour periods (seven days). RA ranges

120 from 0 to 1, with higher values indicating greater distinction between activity levels during the most

121 and least active periods of the day.

122 Participants who provided accelerometer data for less than 72 hours or who did not provide data for

123 all one-hour periods within the 24-hour cycle were excluded from analyses. Participants were also

124 excluded if their data was identified by UK Biobank as having poor calibration. We also excluded

125 over 10 000 participants with data flagged by UK Biobank as unreliable (unexpectedly small or large

126 size) and participants whose wear-time overlapped with a daylight savings clock change.[24] In the

127 current sample (N=91,870), mean RA was 0·87 (SD=0·06; range 0·121-0·997), similar to previously

128 reported values in healthy populations,[8] however the distribution of RA values observed in our

129 sample was negatively skewed (Supplemental Figure 1).

130 ***Genotyping and imputation***

131 UK Biobank released genotypic data for over 500 000 participants using two genotyping arrays

132 specifically designed for UK Biobank with 95% shared marker content.[29] Approximately 10% of these

133 participants were genotyped using Applied Biosystems UK BiLEVE Axiom array by Affymetrix, with

134 the remaining participants being genotyped using Applied Biosystems UK Biobank Axiom Array.

135 Phasing on the autosomes was done using SHAPEIT3 using the 1000 Genomes Phase 3 dataset as a

136 reference panel. Imputation of single nucleotide polymorphism (SNP) genotypes was carried out

137 using IMPUTE4; the merged UK10K and 1000 Genomes Phase 3 reference panel, as used for the UK

138 Biobank interim genotype data release. Stringent quality control was applied to the data, described

139 in an open access document.[29]

140 ***Primary GWAS of low relative amplitude***

141 Our primary GWAS was a study of cases of low RA defined as cases with a mean RA greater than two

142 standard deviations below the overall mean RA, with the remaining participants as controls.[24] Before

143 proceeding with genetic analyses, exclusions were applied to the data. Individuals were removed

144   according to UK Biobank genomic analysis exclusions, failure of quality control, gender mismatch,

145   sex chromosome aneuploidy, ethnicity (not Caucasian), lack of accelerometry data, plus other

146   accelerometry exclusions, as noted above. For related individuals (first cousins or closer), a single

147   individual was randomly selected from each pair of related individuals for inclusion in the analysis.

148   After these (and the acclelerometry-based) exclusions, 71 500 individuals were available for GWAS.

149   Data was further refined by removing SNPs with an imputation score of <0·8, minor allele frequency

150   of <0·01 and Hardy-Weinberg equilibrium $p<1x10^{-6}$ resulting in 7 969 123 variants remaining.

151   The primary association analysis was conducted using logistic regression in PLINK;[30] an additive

152   allelic effects model was used with sex, age, genotyping array, and the first eight genetic principal

153   components as covariates. For the GWAS, genome-wide significance was less than $p<5x10^{-8}$.

Secondary GWAS of RA as a continuous trait

155   The BOLT-LMM method allows the inclusion of related individuals within GWAS. This method relaxes

156   the assumptions of the standard GWAS, as used above, by using a mixture of two Gaussian priors

157   and is a generalisation of a standard mixed model. This mixed model accounts for both relatedness

158   and population stratification; this results in greater power when compared to principal component

159   analysis.[31] As above, individuals were removed according to UK Biobank genomic analysis exclusions,

160   failure of quality control, gender mismatch, sex chromosome aneuploidy, ethnicity (not Caucasian),

161   and lack of accelerometry data. After these exclusions, 77 440 individuals were available for this

162   GWAS. As above, genome-wide significance was less than $p<5x10^{-8}$. Note that due to the imbalance

163   between cases and controls available for low RA, we were unable to use the BOLT-LMM approach

164   for the primary GWAS.[31]

Expression quantitative trait locus (eQTL) analysis

166   The lead SNP from each locus, identified by GWAS, was assessed for the possibility of expression

167   quantitative trait loci (eQTLs). This genotype-specific gene expression was assessed using the GTEx

168   portal. The portal was also used to investigate tissue-specific expression for the implicated genes.[32]

Gene-based analysis

170   The summary statistics from both the primary and secondary GWAS analyses were uploaded to

171   FUMA web application for gene-based analyses.[33] Gene-based analyses were carried out based on

172   the MAGMA method using all genetic associations within the summary statistics.[33,34] For these

173   analyses genome-wide significance was set at $p<5x10^{-5}$.

Genetic correlations between low RA and psychiatric phenotypes

175   Linkage Disequilibrium Score Regression (LDSR) was applied to the GWAS summary statistics to

176   provide an estimate of SNP heritability (h^2SNP).[35,36] LDSR was also used to investigate genetic

177   correlations between low RA and anxiety, ADHD, BD, MDD, mood instability, PTSD, schizophrenia

178   and insomnia. The LD scores for these disorders were obtained using the summary statistics from

179   the Psychiatric Genomics Consortium, CNCR-Complex Trait Genomics group, and UK Biobank. ADHD,

180   BD, MDD, schizophrenia and insomnia were analysed using LD Hub.[36]

181   ***Psychiatric diagnoses, neuroticism and mood instability***

182   a) Bipolar Disorder, Major Depressive Disorder and Generalised Anxiety Disorder

183   A mental health questionnaire (MHQ) was developed by a mental health research reference group

184   to collect additional mental health phenotype data in UK Biobank and was administered during

185   2016-2017.[37] The MHQ was used to obtain information about individuals' lifetime experiences of

186   psychiatric disorders. The questions were based on a modified Composite International Diagnostic

187   Interview Short Form (CIDI-SF). Lifetime depression (referred to here as 'lifetime MDD'), 'lifetime BD'

188   and lifetime generalised anxiety disorder (referred to as 'lifetime GAD') were evaluated based on

189   answers provided by participants to the online MHQ. Therefore, these measurements represent a

190   likelihood of diagnosis with the disorder of interest, rather than confirmed diagnoses.[37] Individuals

191   who self-reported BD or MDD were excluded from the control groups. The 'lifetime' variables were

192   also mutually exclusive, designated cases for one variable were excluded from both the case and

193   control groups for the other variables of interest. Due to these exclusions, the number of

194   observations for each association tested was different.

195   b) Neuroticism

196   To define neuroticism a score was taken from the 12-item neuroticism scale of the Eysenck

197   Personality Questionnaire-Revised Short Form (EPQ-R-S).[38,39] Individuals were given a score of 0 or 1

198   for a "no or yes" answer to each item, with total score from 0 to 12.

199   c) Mood instability

200   A "mood instability" outcome measure was also obtained from the EPQ-R-S questionnaire: for one

201   question participants were asked *"Does your mood often go up and down?"* and given the option to

202   answer "yes", "no", "don't know" or "prefer not to answer"[38]. Individuals who selected "don't know"

203   or "prefer not to answer" were coded as missing; this allowed the generation of a categorical mood

204   instability variable where those who answered "yes" were designated as cases and participants who

205   answered "no" were controls, those answering "don't know" and "prefer not to answer" were

206   excluded.[40]

207

208   ***Association between PRS for low RA and affective disorder phenotypes***

209   Associations between higher PRS for low RA and psychiatric diagnoses were examined in up to 76

210   018 individuals who had completed the MHQ and who were not included in the primary GWAS.

211   Similarly, associations between low RA PRS and mood instability/neuroticism were examined in

212   between 91 248 and 140 504 individuals (depending on the dependent variable) not included in the

213   GWAS. PRS including SNPs at 6 different significance thresholds ($p < 5 \times 10^{-8}$, $p < 5 \times 10^{-5}$, $p < 0·01$, $p < 0·05$,

214   $p < 0·1$, $p < 0·5$) were divided into quartiles, with the exception of $p < 5 \times 10^{-8}$ which was divided into

215   tertiles as only three different PRS were generated for this threshold. The top and bottom quantiles

216   were compared in logistic regression models that were adjusted for age, sex, Townsend deprivation

217   index,[41] genotype array and the first eight genetic principal components. False discovery rate (FDR)

218   correction was applied.[42,43]

219

220   **Results**

221   ***GWAS of low RA***

222   Our primary analysis was a case-control GWAS of low RA. The GWAS data showed only a slight

223   deviation in test statistics compared to the null ($\lambda_{GC}=1·016$, Figure 1); this deviation may be due to

224   the polygenic architecture of low relative amplitude. SNP heritability (h^2_{SNP}) accounted for less than 1%

225   of the population variance ($h^2_{SNP}=0·0067$, se=0·0054). The Manhattan plot for low RA GWAS is

226   presented in Figure 1. Two independent genomic loci on chromosomes 1 and 22 were associated

227   with low RA at genome-wide significance, including three SNPs (described in Supplemental Table 1).

228   These SNPs highlighted two candidate gene loci: Neurofascin (*NFASC*) on chromosome 1 and Solute

229   Carrier Family 25 Member 17 (*SLC25A17*) on chromosome 22 (Supplemental Figure 2). As each of

230   these SNPs is an intronic variant, the exact effect of each polymorphism is unclear.

231   **[insert Figure 1 here]**

232   ***GWAS of continuous RA***

233   As a secondary analysis, we performed a GWAS of a continuous measure of RA using a BOLT-LMM

234   model. The BOLT-LMM GWAS showed a slight deviation in the test statistics compared to the null

235   ($\lambda_{GC}=1·054$, Figure 2), again consistent with a polygenic architecture for RA. The h^2_{SNP} for RA as a

236   continuous measure accounted for greater than 8% of the population variance ($h^2_{SNP}=0·085$,

237   se=0·00035). Five SNPs, all localised to one locus on chromosome 2, were associated with

238   continuous RA at genome-wide significance (described in Supplemental Table 2). These SNPs

239   highlight the Meis Homeobox 1 (*MEIS1*) gene. Again, as noted above, these are intronic SNPs and

240   their exact effects are not known.

8

241   *Expression quantitative trait loci (eQTL) analysis*

242   The lead SNPs from both GWAS were assessed for potential eQTLs. Only the lead SNP found on

243   chromosome 22 (rs9611417) was associated with the expression of a nearby gene. Being

244   heterozygous at rs9611417 was associated with lower expression of *RANGAP1* gene in oesophageal

245   mucosa in comparison to rs9611417 C allele homozygotes (Beta=-0·43, p=7·2x10^{-5}, Supplemental

246   Figure 5). Information on the influence of G homozygotes was not available.

247   *Gene-based analysis of RA*

248   Gene-based analyses of both low RA and continuous RA were undertaken. The gene-based analysis

249   of low RA identified two genes significantly associated with low RA: Forkhead Box J1 (*FOXJ1*) on

250   chromosome 17, and Zinc Finger FYVE-type Containing 21 (*ZFYVE21*) on chromosome 14

251   (Supplemental Figure 3). The gene set analysis of continuous RA identified three genes: Copine 4

252   (*CPNE4*) and Chromosome 3 open reading frame 62 (*C3orf62*) on chromosome 3, and Renalase

253   (*RNLS*) on chromosome 10 (Supplemental Figure 4).

254   *Genetic correlation between low RA and psychiatric phenotypes*

255   There was preliminary evidence of genetic correlation between low RA and insomnia (r_g=0·90,

256   se=0·42, p=0·033), suggesting that the biology underlying low RA may be associated with the

257   regulation of sleep (Table 1). However, on FDR correction this was not statistically significant. There

258   were no significant genetic correlations identified between low RA and ADHD, anxiety, BD, MDD,

259   mood instability, PTSD and schizophrenia.

260   **[insert Table 1 here]**

261   *Association between PRS for low RA and affective disorder phenotypes*

262   The findings of analyses assessing the association between low RA PRS and several mood disorder-

263   related phenotypes are presented in Table 2. Positive associations were identified between

264   increased PRS and mood instability at all PRS thresholds, with the exception of genome-wide

265   significance (p<5x10^{-8}). For MDD, small positive associations were found for the low RA PRS at the

266   top three significance thresholds (OR=1·02-1·03), which remained significant after FDR correction

267   (p=0·025-0·05). A positive association with neuroticism was found for the highest threshold (p=0·004,

268   FDR adjusted p=0·021). However, other associations between the remaining PRS thresholds and

269   neuroticism score were not significant.

270   **[insert Table 2 here]**

271

Discussion

We present the first large-scale genetic study of an objective measure of circadian rest-activity rhythmicity in humans, as well as the first study to examine how common risk SNPs for circadian disruption might be associated with mood disorder phenotypes. The primary GWAS of low RA identified three genome-wide significant SNPs within two independent loci; two of the SNPs highlighted were found to be in relatively high LD (r^2=0·66-1·00) in many populations[44] and are potentially tagging a single underlying functional variant influencing low RA. The secondary GWAS of RA as a continuous measure also identified five genome-wide significant SNPs at a single locus, within the *MEIS1* gene on chromosome 2. Again, the SNPs identified are in medium to high LD with each other (r^2=0·39-1·00) in many populations.[45]

Genes of interest

Of considerable interest is that one of the genes highlighted by GWAS was *NFASC*, a gene encoding the neurofascin protein.[46] Neurofascin is an L1 family immunoglobulin cell adhesion molecule that interacts with several proteins to anchor voltage-gated Na^+ channels to the intracellular skeleton in neurons.[46] It is involved in neurite outgrowth and organization of axon initial segments (AIS) during early development.[46] These AIS complexes (comprising neurofascin, ankyrin G (encoded by *ANK3*), gliomedin and betaIV spectrin) are important for the generation of action potentials and for the maintenance of neuronal integrity.[47] Notably, polymorphisms in *ANK3* are robustly associated with BD.[48] The direct binding of neurofascin to ankyrin G at the AIS therefore represents a potentially important biological link between circadian rhythmicity and BD. The *NFASC* SNPs identified by the GWAS were intronic variants and the *NFASC* transcript undergoes extensive alternative splicing with not all variants being functionally categorised.[46] The precise influence these variants have on the *NFASC* gene is currently unclear.

One of the genome-wide significant SNPs from the primary GWAS is located within *SLC25A17*. This gene encodes a peroxisomal solute carrier membrane protein that transports several cofactors from the cytosol to the peroxisomal matrix.[49] Variants of this gene have previously been associated with autism spectrum disorder and schizophrenia.[50] The *SLC25A17* gene is also involved in adrenomyeloneuropathy,[51] an inherited condition in which long chain fatty acids accumulate in the central nervous system (CNS) disrupting several brain functions.[52] The SNP identified in the GWAS of low RA (rs9611417) associated with lower expression of *RANGAP1*, a GTPase activator protein involved in nuclear transport.[53] *RANGAP1* is 439kb downstream of rs9611417 and shows relatively high expression in the brain (Supplemental Figure 6).

305   The GWAS of RA as a continuous measure highlighted SNPs within the *MEIS1* gene. This gene

306   encodes a homeobox transcription factor (TF) protein crucial for the normal development of several

307   tissues, including the CNS.[54,55] The *MEIS1* gene is also associated with myeloid leukaemia[54] and

308   restless leg syndrome 7 (a sleep-wake disorder).[54,55]

309

310   ***Gene-based analyses***

311   The gene-based analysis of low RA identified two genes: *FOXJ1* on chromosome 17, and *ZFYVE21* on

312   chromosome 14. *FOXJ1* encodes a forkhead TF protein which has a role in cell differentiation in

313   respiratory, reproductive, immune, and CNS tissues. *FOXJ1* is required for the formation of cilia.[56] In

314   mouse models, *FOXJ1* was reported to be important for neurogenesis within the forebrain and

315   olfactory bulb.[56] The *ZFYVE21* gene encodes the zinc-finger FYVE-type containing 21 protein,

316   involved in cell migration and adhesion.[57] The potential involvement of this gene in the brain and

317   circadian function is currently unclear.

318   For the continuous measure of RA, three genes were identified in the gene-based analysis: *CPNE4*

319   and *C3orf62* (chromosome 3); and *RNLS* on (chromosome 10). The *CPNE4* gene encodes a calcium-

320   dependent phospholipid binding protein involved in membrane trafficking and may be involved in

321   intracellular calcium-mediated processes.[58] Deletion of this gene has been associated with earlier

322   age-of-onset of Alzheimer's disease.[59] Currently, the C3orf62 gene is an uncharacterised protein

323   coding gene that has not been functionally annotated.[60]

324   RNLS encodes a flavin adenine dinucleotide-dependent amine oxidase, known as Renalase, an

325   enzyme hormone secreted from the kidney into the bloodstream.[61] Renalase is involved in mediating

326   cardiac function and blood pressure by influencing heart rate and has been associated with

327   hypertension, chronic kidney failure and type 1 diabetes prediction.[61–64] It is worth noting that

328   disrupted circadian rhythmicity has been associated with both diabetes and hypertension in several

329   studies,[4,65,66] and that RNLS has been associated with both treatment outcome and episode

330   recurrence in BD.[67]

331

332   ***Polygenic risk for low RA and psychiatric phenotypes***

333   We found little evidence of genetic correlation between low RA and psychiatric phenotypes, even

334   though disruption of circadian rhythmicity is a core feature of mood disorders and polymorphisms in

335   circadian clock genes have been associated with BD within case-control studies.[68,69] Further, as noted

336   above, within the UK Biobank cohort we recently found that lower RA was associated with several

337   adverse mental health outcomes.[24] In the current study, we found some evidence for association

11

338   between greater polygenic risk for low RA and both MDD and neuroticism. Across several PRS

339   thresholds, we identified a more robust association between increasing PRS for low RA and the

340   phenotype of mood instability. Mood instability is a common symptom that cuts across several

341   psychiatric disorders.[70] The possibility of a direct link between genetic loading for circadian

342   disruption and the experience of dysregulated or unstable mood is therefore of considerable

343   interest and merits further investigation.

344   *Limitations*

345   Several limitations are acknowledged. UK Biobank may not be fully representative of the general UK

346   population,[71] with a possible under-representation of individuals with psychiatric disorders. The

347   scores created using the genome-wide SNPs did not show any associations with lifetime mood

348   disorder diagnoses, mood instability or neuroticism outcomes tested. As only 3 genome-wide

349   significant SNPs were included in these analyses, they may be underpowered. The PRSs (at other

350   thresholds) showed relatively small effects overall on the traits investigated and appeared to explain

351   only a small proportion of the variance within the traits. In terms of phenotyping, the mood

352   instability phenotype was a self-reported subjective measure that may be influenced by response

353   bias. Further, we were unable to identify a direction of causality between RA and the psychiatric

354   phenotypes: approaches to causality such as Mendelian randomisation (MR) were not possible due

355   to the small number of genome-wide significant SNPs identified by both GWAS analyses.[72]

356   **Conclusions**

357   Overall, our findings contribute substantial new knowledge on the genetic architecture of circadian

358   rhythmicity and how this overlaps with mood disorder phenotypes, particularly mood instability.

359   Several of the genetic variants identified are located within or close to genes which may have a role

360   in the pathophysiology of mood disorders. It is hoped that these findings will act as stimulus for

361   future work assessing the complex biological relationships between circadian function and

362   psychiatric disorders.

363

381 **Contributors**

382 All authors contributed substantively to this work. DJS, MESB, JW, LML and AF were involved in

383 study conception and design. CAW, JW, LML and AF were involved in data organisation and

384 statistical analyses. DJS, MESB DML and CAW were involved in application to UK Biobank and data

385 coordination. AF, LML and DJS drafted the report. All authors were involved in reviewing and editing

386 of the manuscript and approved it. DJS, JW, LML and AF had full access to all the data in the study

387 and take responsibility for the integrity and accuracy of analyses.

388

389 **Declaration of interests**

390 No declaration of interests.

391

392

393   **References**

394   1.   McClung, C. A. Circadian genes, rhythms and the biology of mood disorders. *Pharmacol. Ther.*
395   **114,** 222–32 (2007).

396   2.   Reppert, S. M. & Weaver, D. R. Molecular analysis of mammalian circadian rhythms. *Annu.*
397   *Rev. Physiol.* **63,** 647–676 (2001).

398   3.   Merikanto, I. *et al.* Circadian preference towards morningness is associated with lower slow
399   sleep spindle amplitude and intensity in adolescents. *Sci. Rep.* **7,** 1–12 (2017).

400   4.   Reutrakul, S. & Knutson, K. L. Consequences of circadian disruption on cardiometabolic health.
401   *Sleep Med. Clin.* **10,** 455–468 (2015).

402   5.   Wulff, K., Gatti, S., Wettstein, J. G. & Foster, R. G. Sleep and circadian rhythm disruption in
403   psychiatric and neurodegenerative disease. *Nat. Rev. Neurosci.* **11,** 589–599 (2010).

404   6.   Sigurdardottir, L., E. A. Circadian Disription, Sleep Loss and Prostate Cancer risk: A Systematic
405   Review of Epidemiological Studies. *Cancer Epidemiol Biomarkers Prev* **21,** 1002–1011 (2012).

406   7.   Burton, C. *et al.* Activity monitoring in patients with depression: A systematic review. *J. Affect.*
407   *Disord.* **145,** 21–28 (2013).

408   8.   Bullock, B. & Murray, G. Reduced amplitude of the 24 hour activity rhythm: A biomarker of
409   vulnerability to bipolar disorder? *Clin. Psychol. Sci.* **2,** 86–96 (2014).

410   9.   Ng, T. H. *et al.* Sleep-wake disturbance in interepisode bipolar disorder and high-risk
411   individuals: A systematic review and meta-analysis. *Sleep Med. Rev.* **20,** 46–58 (2015).

412   10. Charrier, A., Olliac, B., Roubertoux, P. & Tordjman, S. Clock Genes and Altered Sleep – Wake
413   Rhythms : Their Role in the Development of Psychiatric Disorders. *Int. J. Mol. Sci.* **18,** 1–22
414   (2017).

415   11. Koike, N. *et al.* Transcriptional Architecture and Chromatin Landscape of the Core Circadian
416   Clock in Mammals Nobuya. *Science (80-.).* **338,** 349–354 (2012).

417   12. Zhang, E. E. *et al.* A Genome-Wide RNAi Screen for Modifiers of the Circadian Clock in Human
418   Cells. *Cell* **139,** 199–210 (2009).

419   13. Alloy, L. B., Ng, T. H., Titone, M. K. & Boland, E. M. Circadian Rhythm Dysregulation in Bipolar
420   Spectrum Disorders. *Curr. Psychiatry Rep.* (2017). doi:10.1007/s11920-017-0772-z

421 14. Corruble, E. *et al.* Morningness-eveningness and treatment response in major depressive

422   disorder. *Chronobiol. Int.* **31,** 283–289 (2014).

423 15. Dmitrzak-Węglarz, M. *et al.* Chronotype and sleep quality as a subphenotype in association

424   studies of clock genes in mood disorders. *Acta Neurobiol. Exp. (Wars).* **76,** 32–42 (2016).

425 16. Goel, N., Basner, M., Rao, H. & Dinges, D. F. Circadian Rhythms, Sleep Deprivation, and

426   Human Performance. *Prog Mol Biol Transl Sci.* **119,** 155–190 (2014).

427 17. Hu, Y. *et al.* GWAS of 89,283 individuals identifies genetic variants associated with self-

428   reporting of being a morning person. *Nat. Commun.* **7,** 1–9 (2016).

429 18. Jones, S. E. *et al.* Genome-Wide Association Analyses in 128 , 266 Individuals Identifies New

430   Morningness and Sleep Duration Loci. *PLoS Genet.* **12,** 1–19 (2016).

431 19. Lane, J. M. *et al.* Genome-wide association analysis identifies novel loci for chronotype in

432   100,420 individuals from the UK Biobank. *Nat. Commun.* **7,** 1–10 (2016).

433 20. Jones, S. E. *et al.* Genetic studies of accelerometer-based sleep measures in 85,670

434   individuals yield new insights into human sleep behaviour. *bioRxiv* 303925 (2018).

435   doi:10.1101/303925

436 21. Dashti, H. *et al.* GWAS in 446,118 European adults identifies 78 genetic loci for self-reported

437   habitual sleep duration supported by accelerometer-derived estimates. *bioRxiv* 274977

438   (2018). doi:10.1101/274977

439 22. Jones, S. E. *et al.* Genome-wide association analyses of chronotype in 697,828 individuals

440   provides new insights into circadian rhythms in humans and links to disease. *bioRxiv* 303941

441   (2018). doi:10.1101/303941

442 23. Taillard, J., Philip, P., Coste, O., Sagaspe, P. & Bioulac, B. The circadian and homeostatic

443   modulation of sleep pressure during wakefulness differs between morning and evening

444   chronotypes. *J. Sleep Res.* **12,** 275–282 (2003).

445 24. Lyall, L. M. *et al.* Association of disrupted circadian rhythmicity with mood disorders,

446   subjective wellbeing, and cognitive function: a cross-sectional study of 91 105 participants

447   from the UK Biobank. *The Lancet Psychiatry* **0,** (2018).

448 25. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide

449   range of complex diseases of middle and old age. *PLoS Med.* **12,** e1001779 (2015).

450   26.   Doherty, A. *et al.* Large scale population assessment of physical activity using wrist worn

451   accelerometers: The UK biobank study. *PLoS One* **12,** 1–14 (2017).

452   27.   Sadeh, A. The role and validity of actigraphy in sleep medicine: An update. *Sleep Med. Rev.* **15,**

453   259–267 (2011).

454   28.   Van Someren, E. J. W. *et al.* Circadian rest-activity rhythm disturbances in Alzheimer's disease.

455   *Biol. Psychiatry* **40,** 259–270 (1996).

456   29.   Bycroft, C. *et al.* Genome-wide genetic data on ~ 500 , 000 UK Biobank participants. (2017).

457   30.   Purcell, S. *et al.* REPORT PLINK : A Tool Set for Whole-Genome Association and Population-

458   Based Linkage Analyses. **81,** 559–575 (2007).

459   31.   Loh, P.-R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large

460   cohorts. *Nat. Genet.* **47,** 284–90 (2015).

461   32.   GTEx Consortium, T. Gte. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45,**

462   580–5 (2013).

463   33.   Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and

464   annotation of genetic associations with FUMA. *Nat. Commun.* **8,** 1826 (2017).

465   34.   de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set

466   Analysis of GWAS Data. *PLOS Comput. Biol.* **11,** e1004219 (2015).

467   35.   Bulik-sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat.*

468   *Publ. Gr.* **47,** 1236–1241 (2015).

469   36.   Zheng, J. *et al.* LD Hub: A centralized database and web interface to perform LD score

470   regression that maximizes the potential of summary level GWAS data for SNP heritability and

471   genetic correlation analysis. *Bioinformatics* **33,** 272–279 (2017).

472   37.   Davis, K. A. S. *et al.* Mental health in UK Biobank : development , implementation and results

473   from an online questionnaire completed by 157 366 participants *. *BJPsych Open* **4,** 83–90

474   (2018).

475   38.   Eysenck, S. B. G., Eysenck, H. J. & Barrett, P. A revised version of the psychoticism scale. *Pers.*

476   *Individ. Dif.* **6,** 21–29 (1985).

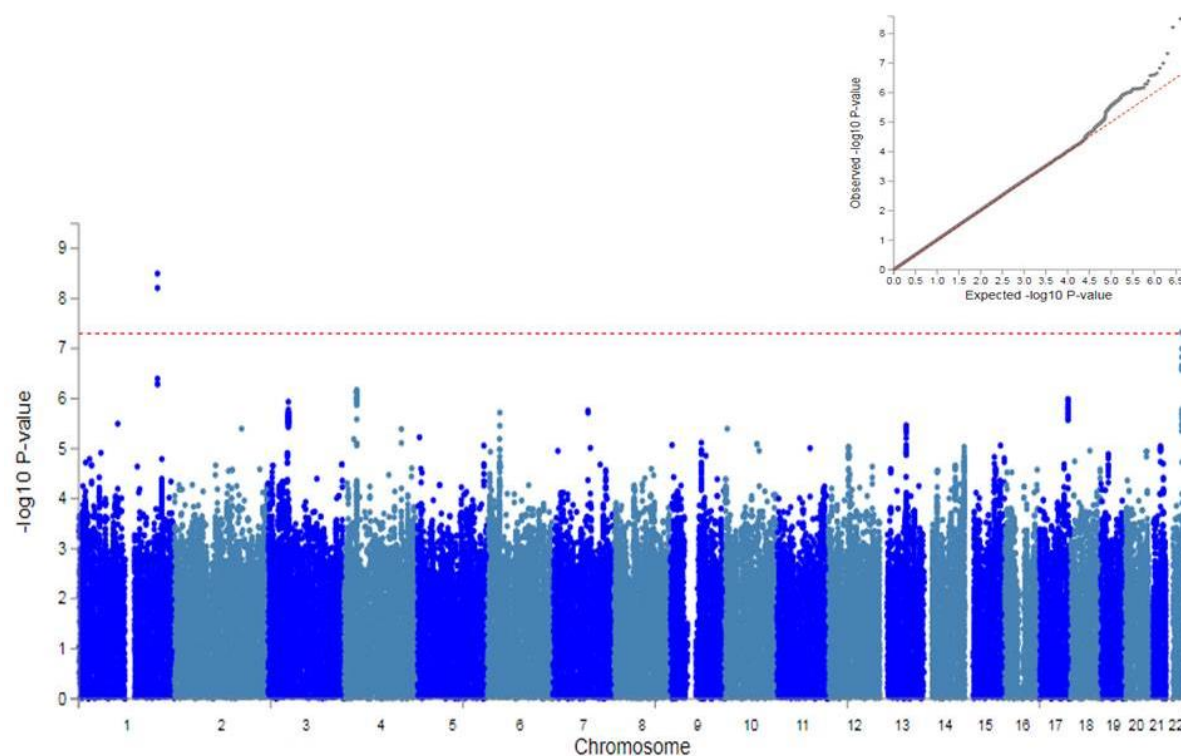477   39.   Smith, D. J. *et al.* Genome-wide association study of neuroticism in UK Biobank. (2015).

478   40.   Ward, J. *et al.* Genome-wide analysis in UK Biobank identifies four loci associated with mood

479   instability and genetic correlation with major depressive disorder, anxiety disorder and

480   schizophrenia. *Transl. Psychiatry* **7,** 1264 (2017).

481   41.   Townsend, P. Deprivation. *J. Soc. Policy* **16,** 125–146 (1987).

482   42.   Pike, N. Using false discovery rates for multiple comparisons in ecology and evolution. 278–

483   282 (2011). doi:10.1111/j.2041-210X.2010.00061.x

484   43.   Benjamini, Y. & Hochberg, Y. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a

485   practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57,** 289–300 (1995).

486   44.   rs147964682 (SNP) - Linkage disequilibrium - Homo sapiens - Ensembl genome browser 92.

487   Available at:

488   http://www.ensembl.org/Homo_sapiens/Variation/HighLD?db=core;r=1:204927176-

489   204928176;v=rs147964682;vdb=variation;vf=32409503. (Accessed: 9th May 2018)

490   45.   rs113851554 (SNP) - Linkage disequilibrium - Homo sapiens - Ensembl genome browser 92.

491   Available at:

492   http://www.ensembl.org/Homo_sapiens/Variation/HighLD?db=core;r=2:66522932-

493   66523932;v=rs113851554;vdb=variation;vf=21204591;second_variant_name=rs142412330.

494   (Accessed: 4th June 2018)

495   46.   Taylor, A. M., Saifetiarova, J. & Bhat, M. A. Postnatal Loss of Neuronal and Glial Neurofascins

496   Differentially Affects Node of Ranvier Maintenance and Myelinated Axon Function. *Front. Cell.*

497   *Neurosci.* **11,** (2017).

498   47.   Thaxton, C. *et al.* In Vivo Deletion of Immunoglobulin Domains 5 and 6 in Neurofascin (Nfasc)

499   Reveals Domain-Specific Requirements in Myelinated Axons. *J. Neurosci.* **30,** 4868–4876

500   (2010).

501   48.   Ferreira, M. A. R. *et al.* Collaborative genome-wide association analysis supports a role for

502   ANK3 and CACNA1C in bipolar disorder. *Nat. Genet.* **40,** 1056–8 (2008).

503   49.   Agrimi, G., Russo, A., Scarcia, P. & Palmieri, F. The human gene *SLC25A17* encodes a

504   peroxisomal transporter of coenzyme A, FAD and NAD $^{+}$. *Biochem. J.* **443,** 241–247 (2012).

505   50.   The Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium.

506   Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a

507   novel locus at 10q24.32 and a significant overlap with schizophrenia. *Mol. Autism* **8,** 21 (2017).

508   51.   Van Veldhoven, P. P. Biochemistry and genetics of inherited disorders of peroxisomal fatty
509   acid metabolism. *J. Lipid Res.* **51,** 2863–95 (2010).

510   52.   Moser, A. B. *et al.* Plasma very long chain fatty acids in 3,000 peroxisome disease patients
511   and 29,000 controls. *Ann. Neurol.* **45,** 100–110 (1999).

512   53.   Bischoff, F. R., Krebber, H., Kempf, T., Hermes, I. & Ponstingl, H. Human RanGTPase-activating
513   protein RanGAP1 is a homologue of yeast Rna1p involved in mRNA processing and transport.
514   *Proc. Natl. Acad. Sci. U. S. A.* **92,** 1749–53 (1995).

515   54.   Wang, Q. *et al.* Regulation of MEIS1 by distal enhancer elements in acute leukemia. *Leukemia*
516   **28,** 138–146 (2014).

517   55.   Xiong, L. *et al.* MEIS1 intronic risk haplotype associated with restless legs syndrome affects its
518   mRNA and protein expression levels. *Hum. Mol. Genet.* **18,** 1065–74 (2009).

519   56.   Jacquet, B. V *et al.* Specification of a Foxj1-dependent lineage in the forebrain is required for
520   embryonic-to-postnatal transition of neurogenesis in the olfactory bulb. *J. Neurosci.* **31,**
521   9368–82 (2011).

522   57.   Nagano, M. *et al.* ZF21 protein, a regulator of the disassembly of focal adhesions and cancer
523   metastasis, contains a novel noncanonical pleckstrin homology domain. *J. Biol. Chem.* **286,**
524   31598–609 (2011).

525   58.   Tomsig, J. L. & Creutz, C. E. Copines: a ubiquitous family of Ca(2+)-dependent phospholipid-
526   binding proteins. *Cell. Mol. Life Sci.* **59,** 1467–77 (2002).

527   59.   Szigeti, K. *et al.* Genome-Wide Scan for Copy Number Variation Association with Age at Onset
528   of Alzheimer's Disease. *J. Alzheimer's Dis.* **33,** 517–523 (2013).

529   60.   Gene: C3orf62 (ENSG00000188315) - Summary - Homo sapiens - Ensembl genome browser
530   92. Available at:
531   http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000188315;r=3:4926860
532   2-49277909. (Accessed: 30th May 2018)

533   61.   Xu, J. *et al.* Renalase is a novel, soluble monoamine oxidase that regulates cardiac function
534   and blood pressure. *J. Clin. Invest.* **115,** 1275–80 (2005).

535   62.   Lv, Y.-B. *et al.* Association of Renalase SNPs rs2296545 and rs2576178 with the Risk of
536   Hypertension: A Meta-Analysis. *PLoS One* **11,** e0158880 (2016).

18

537   63.   Frohnert, B. I. *et al.* Prediction of type 1 diabetes using a genetic risk model in the Diabetes

538   Autoimmunity Study in the Young. *Pediatr. Diabetes* **19,** 277–283 (2018).

539   64.   Desir, G. V. Renalase deficiency in chronic kidney disease, and its contribution to

540   hypertension and cardiovascular disease. *Curr. Opin. Nephrol. Hypertens.* **17,** 181–185 (2008).

541   65.   Merikanto, I. *et al.* Evening types are prone to depression. *Chronobiol. Int.* **30,** 719–725

542   (2013).

543   66.   Merikanto, I., Suvisaari, J., Lahti, T. & Partonen, T. Eveningness relates to burnout and

544   seasonal sleep and mood problems among young adults. *Nord. J. Psychiatry* **70,** 72–80 (2016).

545   67.   Fabbri, C. & Serretti, A. Genetics of long-term treatment outcome in bipolar disorder. *Prog.*

546   *Neuro-Psychopharmacology Biol. Psychiatry* **65,** 17–24 (2016).

547   68.   Geoffroy, P. A. *et al.* Sleep in remitted bipolar disorder: A naturalistic case-control study using

548   actigraphy. *J. Affect. Disord.* **158,** 1–7 (2014).

549   69.   McCarthy, M. J. *et al.* Genetic and clinical factors predict lithium's effects on PER2 gene

550   expression rhythms in cells from bipolar disorder patients. *Transl. Psychiatry* **3,** e318-8 (2013).

551   70.   Broome, M. R., Saunders, K. E. A., Harrison, P. J. & Marwaha, S. Mood instability: Significance,

552   definition and measurement. *Br. J. Psychiatry* **207,** 283–285 (2015).

553   71.   Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK

554   Biobank Participants With Those of the General Population. *Am. J. Epidemiol.* **186,** 1026–1034

555   (2017).

556   72.   Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal inference

557   in epidemiological studies. *Hum. Mol. Genet.* **23,** R89-98 (2014).

558

559   **Figures**

560   Figure 1. SNP Manhattan plot and QQ plot (inset) of low RA GWAS (N=2 700 cases verses N=68 300

561   controls)

562
563

564   Red line of Manhattan plot represents genome-wide significance ($p<5\times10^{-8}$).

565

566   Figure 2. SNP Manhattan plot and QQ plot (inset) of continuous RA GWAS (N=77 440)

567

568   Red line of Manhattan plot represents genome-wide significance ($p<5\times10^{-8}$).

21
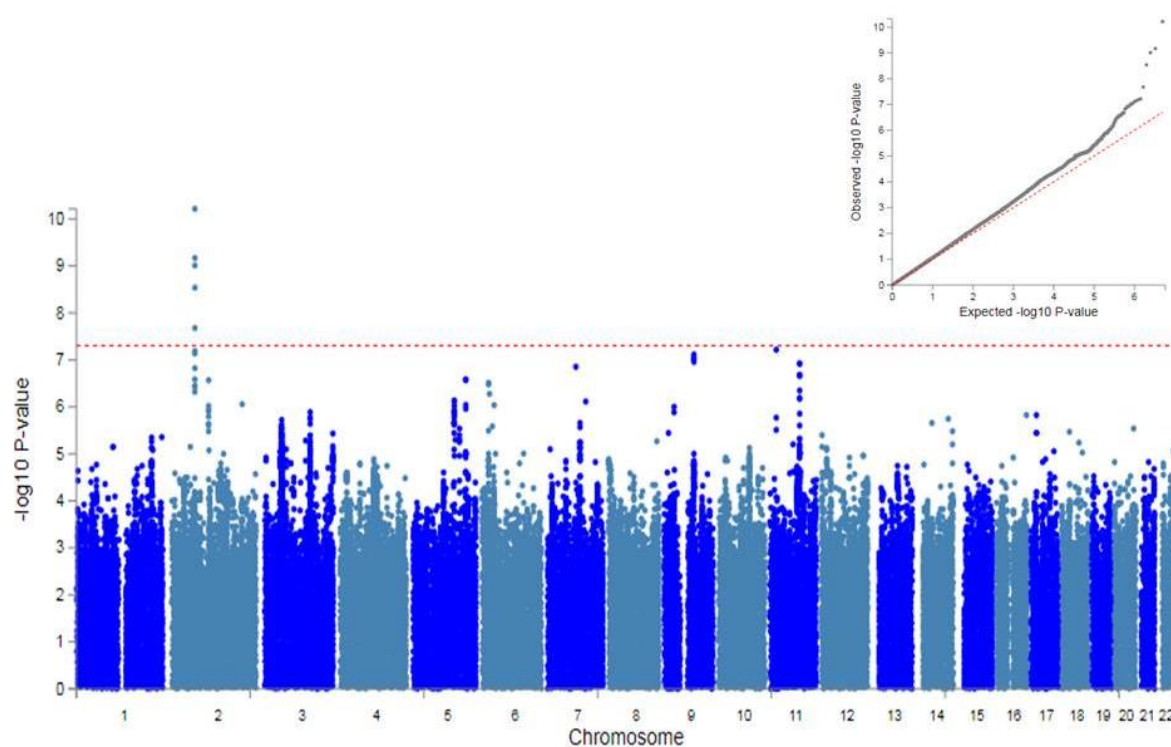
569   Table 1. Genetic correlations between low relative amplitude and ADHD, anxiety, BD, MDD, mood

570   instability, PTSD, schizophrenia and insomnia

Phenotype	r_g	se	z	p	p FDR corrected
ADHD	0·3476	0·4541	0·7656	0·444	0·932
Anxiety	-0·0037	0·4224	-0·0088	0·993	1·000
BD	-0·0626	0·1842	-0·3396	0·7341	1·000
MDD	0·0045	0·2527	0·0178	0·9858	1·000
Mood Instability	-0·1587	0·2713	-0·5847	0·5587	0·939
PTSD	0·7396	0·6201	1·1926	0·233	0·713
Schizophrenia	0·1541	0·1353	1·139	0·2547	0·713
Insomnia	0·8996	0·4222	2·1308	**0·0331**	0·278

571
572   Rg is the genetic correlation with low RA, se is the standard error of the correlation, z represents test
573   statistic. p is the uncorrected p value and p FDR corrected has been adjusted for multiple testing.

574

575

576   Table 2. Associations between low RA PRS and psychiatric phenotypes

PRS p threshold	Outcome (Cases/Controls)	OR (95% CI)	p uncorrected	p FDR corrected
$p<5\times10^{-8}$	BD	0·99 (0·92,1·06)	0·748	0·785
$p<5\times10^{-5}$	(406/37 699)	1·05 (0·98, 1·12)	0·206	0·746
$p<0·01$		1·03 (0·96, 1·11)	0·355	0·746
$p<0·05$		1·04 (0·96, 1·12)	0·309	0·746
$p<0·1$		1·02 (0·94, 1·10)	0·617	0·785
$p<0·5$		1·02 (0·93, 1·11)	0·754	0·785
$p<5\times10^{-8}$	MDD	1·00 (0·99, 1·02)	0·812	0·805
$p<5\times10^{-5}$	(9 543/24 317)	1·00 (0·98, 1·02)	0·966	0·805
$p<0·01$		1·01 (0·99, 1·03)	0·395	0·494
$p<0·05$		1·02 (1·00, 1·04)	**0·03**	**0·05**
$p<0·1$		1·03 (1·01, 1·05)	**0·005**	**0·025**
$p<0·5$		1·03 (1·00, 1·05)	**0·021**	**0·05**
$p<5\times10^{-8}$	GAD	0·97 (0·95, 1·00)	0·092	0·3
$p<5\times10^{-5}$	(2 587/23 564)	0·98 (0·96, 1·01)	0·274	0·548
$p<0·01$		0·99 (0·97, 1·02)	0·729	0·729
$p<0·05$		1·01 (0·98, 1·04)	0·475	0·713
$p<0·1$		1·03 (0·99, 1·06)	0·1	0·3
$p<0·5$		1·01 (0·97, 1·04)	0·699	0·729
$p<5\times10^{-8}$	Mood Instability	1·00 (0·99, 1·01)	0·913	0·94
$p<5\times10^{-5}$	(78 710/91 248)	1·01 (1·00, 1·02)	**0·019**	**0·0096**

23

PRS p threshold	Outcome	Beta (95% CI)	p uncorrected	p FDR corrected
p<0·01		1·01 (1·01, 1·02)	**9·5x10^{-5}**	**2·2x10^{-4}**
p<0·05		1·02 (1·01, 1·02)	**3·6x10^{-6}**	**9·6x10^{-5}**
p<0·1		1·01 (1·01, 1·02)	**8·3x10^{-5}**	**5·9x10^{-4}**
p<0·5		1·02 (1·01, 1·03)	**1·2x10^{-6}**	**1·5x10^{-5}**

PRS p threshold	Outcome	Beta (95% CI)	p uncorrected	p FDR corrected
p<5x10^{-8}	Neuroticism score (140 504)	-0·004 (-0·02, 0·007)	0·456	0·399
p<5x10^{-5}		0·01 (-0·003, 0·02)	0·124	0·134
p<0·01		0·01 (-0·002, 0·02)	0·098	0·134
p<0·05		0·01 (-0·005, 0·03)	0·059	0·134
p<0·1		0·01 (-0·003, 0·02)	0·128	0·134
p<0·5		0·02 (0·007, 0·04)	**0·004**	**0·021**

577

578 **Supplemental**

579 Supplemental Figure 1. Relative amplitude histogram indicating (N=2 700 Cases and N=68 300
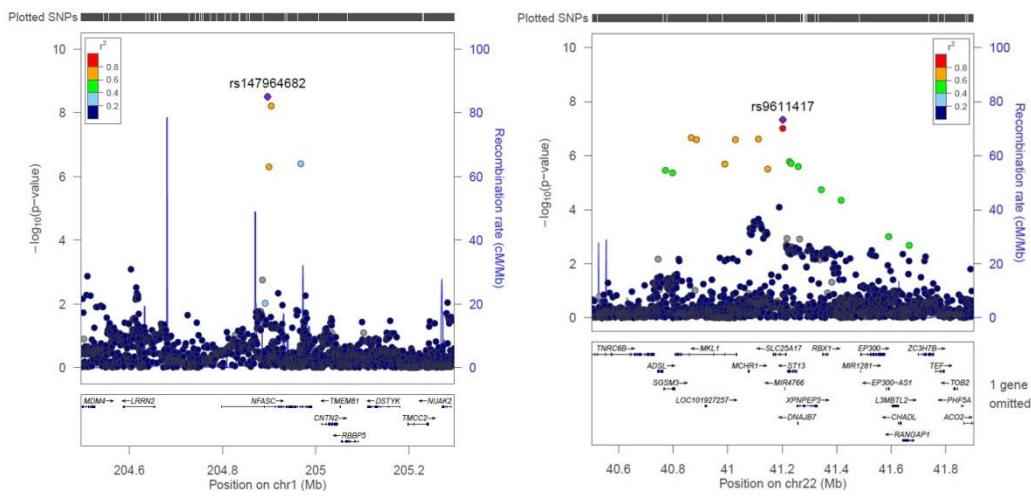
580 Controls)
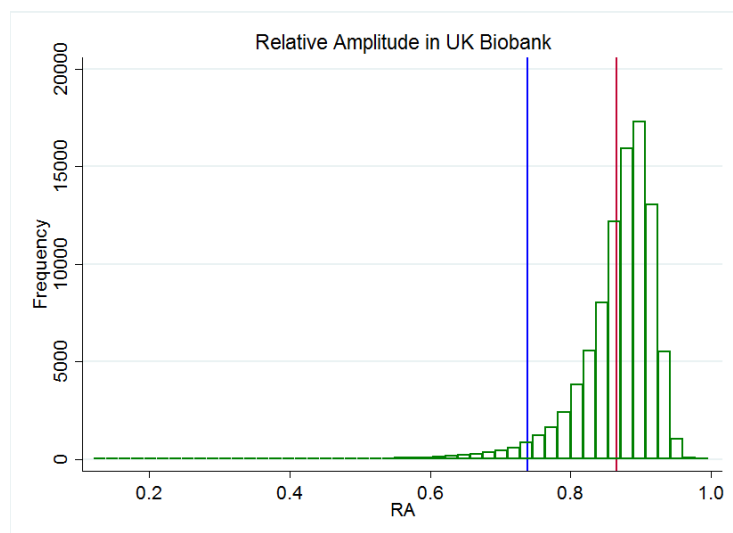
581

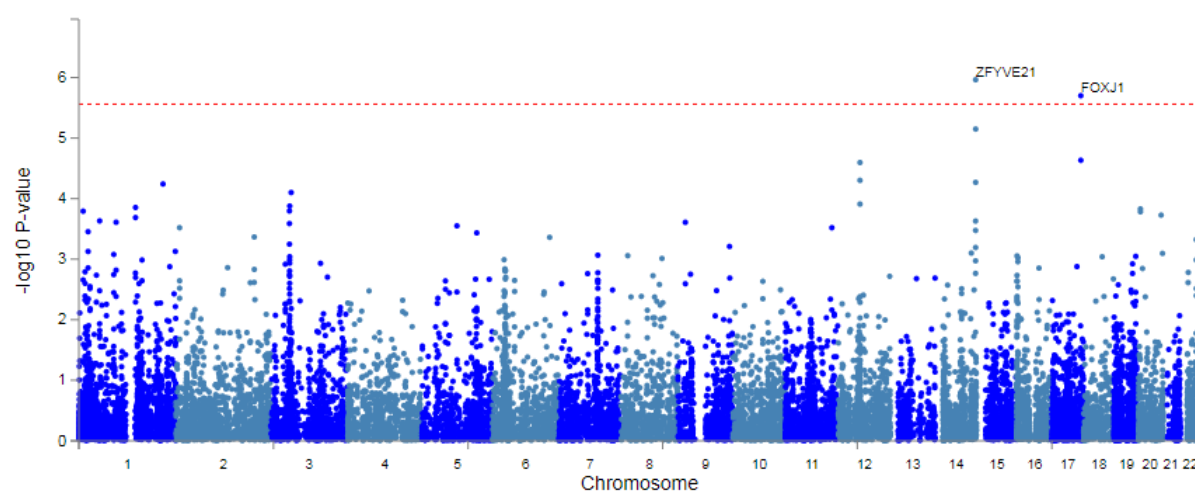582 Red line represents mean value of RA. Blue line is two standard deviations from the mean,

583 designating cases for use in the primary GWAS.

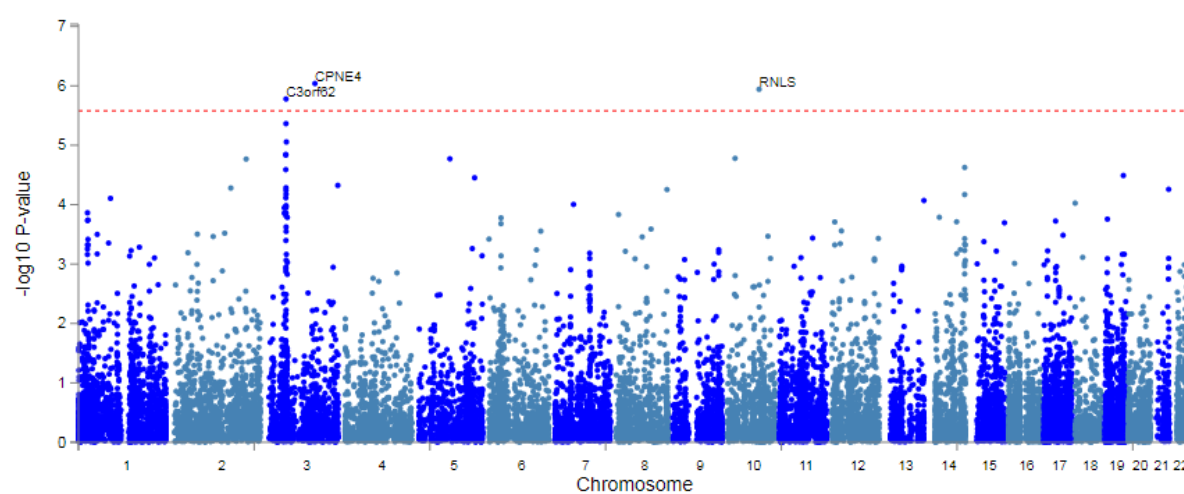584   Supplemental Figure 2. Regional plots of *NFASC* and *SLC25A17*

585

586   Regional plots of SNPs produced by FUMA.[33]

587   Supplemental Figure 3. Low RA gene-based analysis
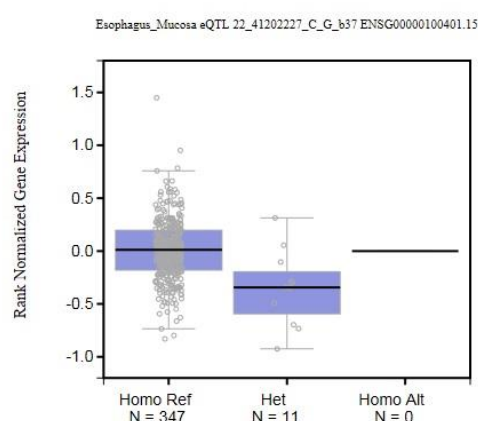

588

589   Red line represents genome-wide significance ($p < 5 \times 10^{-8}$).

590   Supplemental Figure 4. Continuous RA gene-based analysis

591

592   Red line represents genome-wide significance ($p < 5 \times 10^{-8}$).

593

594   Supplemental Figure 5. eQTLs of rs9611417 box plot

Esophagus_Mucosa eQTL 22_41202227_C_G_b37 ENSG00000100401.15

595

596   Homo Ref: rs9611417 CC, Het: rs9611417 CG, Homo Alt: rs9611417 GG. Obtained from GTex

597   portal.[32]

598

599   Supplemental Figure 6. Tissue specific expression of *RANGAP1*

ENSG00000100401.15 Gene Expression
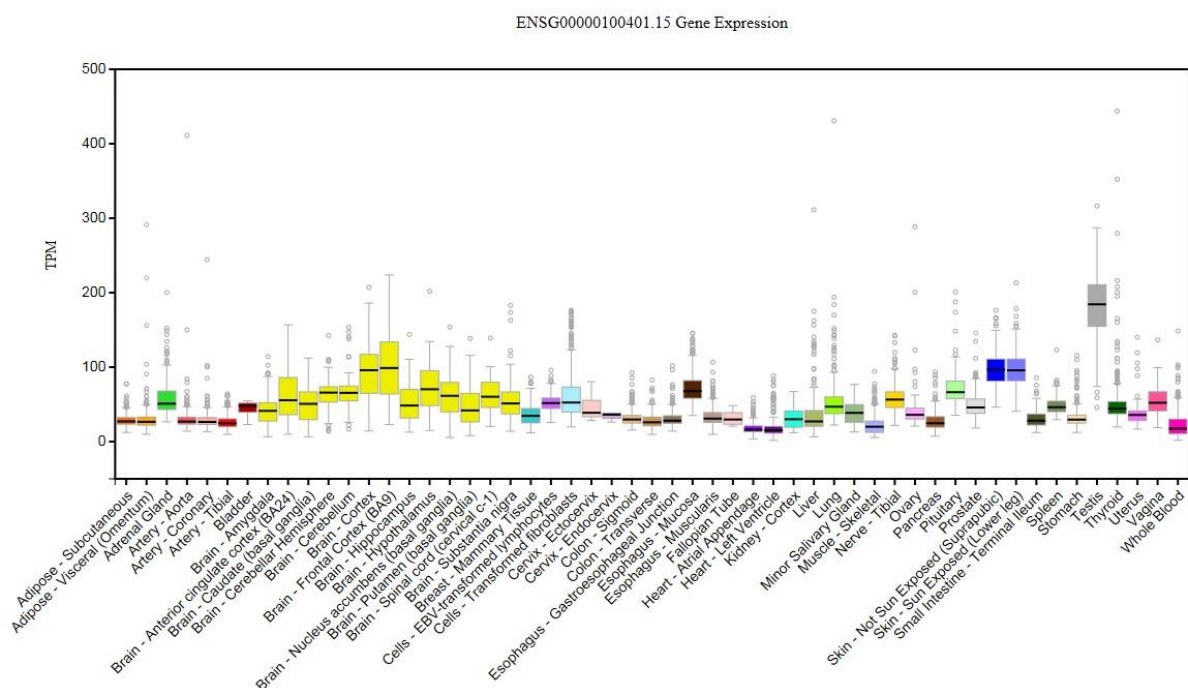
600

601   Obtained from GTex portal.[32]

602   Supplemental Table 1. Genome-wide significant loci associated with low RA

SNP	Chr	Position	A1/A2	MAF	Beta	p	Nearby gene
rs147964682	1	204 896 804	G/C	0·012	0·584	**3·179x10^{-9}**	*NFASC*
rs146042826	1	204 904 528	A/G	0·013	0·568	**6·17x10^{-9}**	*NFASC*
rs9611417	22	41 202 227	G/C	0·012	0·5616	**4·753x10^{-8}**	*SLC25A17*

603

604   Supplemental Table 2. Genome-wide significant loci associated with continuous RA using BOLT-LMM

SNP	Chr	Position	A1/A2	Info	Beta	p	Nearby gene
rs113851554	2	66 750 564	G/T	0·93	0·0048	6·10x10^{-11}	*MEIS1*
2:66782432_AC_A	2	66 782 432	AC/A	0·92	0·0048	9·80x10^{-10}	*MEIS1*
rs11679120	2	66 785 180	G/A	0·93	0·0045	2·10x10^{-8}	*MEIS1*
rs115087496	2	66 793 725	G/C	0·92	0·0047	2·90x10^{-9}	*MEIS1*
rs142412330	2	66 802 493	T/TCTC	0·93	-0·0036	6·80x10^{-10}	*MEIS1*

605