

## ORIGINAL ARTICLE

## Genetic contributions to self-reported tiredness

V Deary<sup>1,7</sup>, SP Hagenaars<sup>2,3,4,7</sup>, SE Harris<sup>2,5</sup>, WD Hill<sup>2,3</sup>, G Davies<sup>2,3</sup>, DCM Liewald<sup>1,2</sup>, International Consortium for Blood Pressure GWAS, CHARGE Consortium Aging and Longevity Group, CHARGE Consortium Inflammation Group, AM McIntosh<sup>4</sup>, CR Gale<sup>2,3,6,7</sup> and IJ Deary<sup>2,3,7</sup>

Self-reported tiredness and low energy, often called fatigue, are associated with poorer physical and mental health. Twin studies have indicated that this has a heritability between 6 and 50%. In the UK Biobank sample ( $N = 108\,976$ ), we carried out a genome-wide association study (GWAS) of responses to the question, 'Over the last two weeks, how often have you felt tired or had little energy?' Univariate GCTA-GREML found that the proportion of variance explained by all common single-nucleotide polymorphisms for this tiredness question was 8.4% (s.e. = 0.6%). GWAS identified one genome-wide significant hit (Affymetrix id 1:64178756\_C\_T;  $P = 1.36 \times 10^{-11}$ ). Linkage disequilibrium score regression and polygenic profile score analyses were used to test for shared genetic aetiology between tiredness and up to 29 physical and mental health traits from GWAS consortia. Significant genetic correlations were identified between tiredness and body mass index (BMI), C-reactive protein, high-density lipoprotein (HDL) cholesterol, forced expiratory volume, grip strength, HbA1c, longevity, obesity, self-rated health, smoking status, triglycerides, type 2 diabetes, waist-hip ratio, attention deficit hyperactivity disorder, bipolar disorder, major depressive disorder, neuroticism, schizophrenia and verbal-numerical reasoning (absolute  $r_g$  effect sizes between 0.02 and 0.78). Significant associations were identified between tiredness phenotypic scores and polygenic profile scores for BMI, HDL cholesterol, low-density lipoprotein cholesterol, coronary artery disease, C-reactive protein, HbA1c, height, obesity, smoking status, triglycerides, type 2 diabetes, waist-hip ratio, childhood cognitive ability, neuroticism, bipolar disorder, major depressive disorder and schizophrenia (standardised  $\beta$ 's had absolute values  $< 0.03$ ). These results suggest that tiredness is a partly heritable, heterogeneous and complex phenomenon that is phenotypically and genetically associated with affective, cognitive, personality and physiological processes.

*Molecular Psychiatry* (2018) **23**, 609–620; doi:10.1038/mp.2017.5; published online 14 February 2017

## INTRODUCTION

'Hech, sirs! But I'm wabbit, I'm back frae the toon; I ha'ena dune pechin'—jist let me sit doon.'

From *Glesca*' By William Dixon Cocker (1882–1970)

The present study examines genetic contributions to how the UK Biobank's participants answered the question, 'Over the last two weeks, how often have you felt tired or had little energy?' Ideal questionnaire items do not have conjunctions, but the 'or' is understandable here, and it may even allow capture of both peripheral and central fatigue. The first and last authors of the present study grew up in South Lanarkshire in Scotland, where fatigue was often self-reported in terms of feeling 'wabbit'. The Scots word wabbit encompasses both peripheral fatigue, the muscle weakness after a long walk, and central fatigue, the reduced ability to initiate and/or sustain mental and physical activity, such as we might experience while having flu. Throughout the paper, we refer mainly to the single English words 'fatigue' and/or 'tiredness' as the construct captured by the question, but the Scottish vernacular word is a good reminder of the subjective 'feel' of fatigue.

Fatigue is a common complaint. In a Dutch adult, general population survey with 9375 respondents, 4.9% reported short-term fatigue ( $< 6$  months duration), 30.5% chronic fatigue ( $> 6$  months duration) and 1% fulfilled diagnostic criteria for chronic fatigue syndrome (CFS).<sup>1</sup> These findings are similar to a London-based survey of general practice patients in England, aged 18–45 years, with 15 283 respondents, where 36.7% reported substantial fatigue, 18.3% substantial fatigue of  $> 6$  months duration and 1% fulfilled the diagnostic criteria for CFS.<sup>2</sup> Two other large surveys of US workers and community-dwelling adults aged 51 years and over report fatigue rates of 37.9% (2-week period prevalence) and 31.2% (1-week period prevalence) respectively.<sup>3,4</sup> In an early review of fatigue epidemiology, Lewis and Wessely<sup>5</sup> argue that fatigue 'is best viewed on a continuum', and the continuous distribution of fatigue in the general population is supported by the Pawlikowska *et al.*<sup>2</sup> study. Fatigue is also a common presentation in primary care. In a survey of 1428 consultations to 89 general practitioners in Ireland, fatigue prevalence was 25% and the main reason for attendance in 6.5%.<sup>6</sup>

Demographically, higher levels of self-reported fatigue are associated with female sex, lower socioeconomic status<sup>1</sup> and poorer self-rated health status.<sup>7</sup> There are less clear associations

<sup>1</sup>Department of Psychology, Northumbria University, Newcastle, UK; <sup>2</sup>Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK; <sup>3</sup>Department of Psychology, University of Edinburgh, Edinburgh, UK; <sup>4</sup>Division of Psychiatry, University of Edinburgh, Edinburgh, UK; <sup>5</sup>Medical Genetics Section, University of Edinburgh, Centre for Genomic and Experimental Medicine and MRC Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, UK and <sup>6</sup>MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK. Correspondence: Dr V Deary, Department of Psychology, Northumbria University, Newcastle NE1 8ST, UK.

E-mail: vincent.deary@northumbria.ac.uk

or Professor IJ Deary, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK.

E-mail: i.deary@ed.ac.uk

<sup>7</sup>These authors contributed equally to the work.

Received 6 April 2016; revised 1 December 2016; accepted 13 December 2016; published online 14 February 2017

between age and fatigue, with some studies reporting a small but significant positive correlation between age and fatigue,<sup>2</sup> whereas others report no association<sup>6</sup> or a negative association.<sup>1,7</sup> There is a clearer link with the Fried phenotype of frailty in older adults,<sup>8</sup> which has significant associations with mortality. The frailty phenotype comprises weakness (as measured by grip strength), weight loss, reduced mobility, reduced walking speed and fatigue.<sup>9</sup>

Fatigue is associated with a number of lifestyle-related factors and conditions. Smoking is a risk factor for fatigue<sup>10</sup> and fatigue has strong cross-sectional associations with type 2 diabetes<sup>11</sup> and increased body mass index (BMI).<sup>12</sup> Fatigue is consistently associated with poorer physical and mental health status. It is one of the most common symptom complaints of cancer patients. For those undergoing treatment, prevalence estimates vary between 25 and 99%, and 25 and 30% of survivors report long-term fatigue.<sup>13</sup> Fatigue is also a significant symptom of, to name just a few conditions, primary biliary cirrhosis,<sup>14</sup> multiple sclerosis,<sup>15</sup> rheumatoid arthritis,<sup>16</sup> primary Sjogren's syndrome<sup>17</sup> and Parkinson's disease.<sup>18</sup> It is associated with chronic disease in general, and there is a linear relationship between number of chronic diseases and self-reported fatigue.<sup>19</sup> Fatigue is also associated with depression,<sup>20</sup> with self-reported psychological distress,<sup>2,21</sup> and with the personality trait of neuroticism.<sup>22,23</sup>

Research into the biological mechanisms of fatigue has focussed on a few key areas. Fatigue is associated with the cytokine-mediated inflammatory response, particularly interleukin-1beta and interleukin-6. These latter have been shown, for instance, to be elevated in cancer patients,<sup>24</sup> and administration of interferon-alpha produces depression and/or fatigue in the majority of patients receiving it as a treatment.<sup>25</sup> Hypothalamic-pituitary-adrenal axis dysregulation in the form of hypocortisolaemia, blunted diurnal variation and blunted stress reactivity have been found in the cross-sectional studies of CFS patients (see Tomas *et al.*<sup>26</sup> for a recent review). Other popular candidate aetiopathological mechanisms for fatigue include serotonin pathways, circadian dysregulation, autonomic dysfunction,<sup>17</sup> 5HT neurotransmitter dysregulation, alterations in ATP metabolism and vagal afferent activation.<sup>24</sup> Some authors have suggested that, rather than being located with one biological system, fatigue represents a systemic dysregulation of the interaction between these systems.<sup>27</sup>

At the other end of the biopsychosocial spectrum, psychosocial models of fatigue focus on the role that the individual's response to their symptom may serve in perpetuating it. For instance, in a cross-sectional study of 149 patients with multiple sclerosis, illness-related cognitions and behaviours were associated with a higher level of fatigue independent of neurological impairment.<sup>28</sup> More integrative models are predicated on the notion of allostatic load, the psychophysiological work done to adjust to stress and its impact upon the body's self-regulatory systems. As such, these models complement biological accounts of fatigue and provide potential pathways for integrating psychosocial and biological findings.<sup>29</sup> Multifactorial accounts and models of fatigue exist in multiple sclerosis,<sup>30</sup> primary biliary cirrhosis,<sup>14</sup> obesity,<sup>31</sup> diabetes,<sup>32</sup> frailty<sup>33</sup> and cancer.<sup>13</sup> These multifactorial models postulate that fatigue is likely to be the product of physiological factors (generic, such as inflammation and/or disease specific such as hyperglycaemia in diabetes), psychosocial factors (for example, emotional distress), lifestyle and behavioural factors (for example, reduced activity), illness consequences (for example, sleep disturbance and weakness) and the interaction of these contributors.

Research into the genetics and epigenetics of fatigue has tended to focus on genes associated with the biological mechanisms described above, and done so usually within fatiguing illnesses such as those listed above. Candidate gene studies have suggested several genes to be involved in CFS, particularly genes involved in the immune system and

Hypothalamic-pituitary-adrenal axis. Reviewing this literature, Landmark-Høyvik *et al.*<sup>34</sup> suggest that findings are inconclusive, and are hampered by phenotypic heterogeneity, lack of power and poor study design. For example, the candidate gene studies included in the Landmark-Høyvik paper had sample sizes between 2 and 248 individuals, and the results have not been replicated. Twin studies have shown the heritability of fatigue to be between 6 and 50% with a higher concordance in monozygotic twins than dizygotic twins.<sup>35,36</sup> One of these studies<sup>36</sup> did not show any sex-specific patterns of genetic influences in a Swedish sample, whereas the other one<sup>35</sup> showed differences in the amount of variance explained by the genetic effect for males and females. Around half the variance in males was explained by the genetic effects compared to only a fifth of the variance in females. In a study of fatigue, insomnia and depression in 3758 twins (893 monozygotic pairs and 884 dizygotic pairs), the best model was a common pathway model, suggesting that the high association between the symptoms (correlations of 0.35–0.44) was mediated by an underlying common factor whose variation was 49% genetic and 51% unique environmental.<sup>37</sup> This study showed that unique specific variance in fatigue was 38% genetic and 62% unique environmental, which supports a previous study, suggesting that fatigue is largely attributable to additive genetic factors.<sup>38</sup> Genome-wide association studies (GWAS) have shown an association between single-nucleotide polymorphisms (SNPs) in genes associated with impaired cognitive abilities (*GRIK2*,  $P=1.26 \times 10^{-11}$ )<sup>39,40</sup> and the circadian clock (*NPAS2*, not genome-wide significant)<sup>40</sup> and CFS, but this was in a sample of just 42 cases of CFS and 38 controls, lacking statistical power to detect genome-wide findings.

To sum up in the words of Landmark-Høyvik *et al.*:<sup>41</sup> 'fatigue can be conceptualised as a final common end point for psychological and biological processes. Fatigue is therefore both heterogeneous (occurring across different conditions) and multifactorial'. Given that, it could be argued that it is futile to search for a shared genetic contribution to tiredness, as it may not exist. However, in line with other fatigue research programmes<sup>42</sup> and the research cited above, we judge that the best way to approach this complexity is to conduct large, well-designed studies focussing on specific areas of the biopsychosocial spectrum, and that to date no large study has done this at the genetic level.

Tiredness can be the result of external factors—such as poor sleep—or inherent factors—such as personality traits or poor health. It is therefore important that we clarify the scope of the present study in terms of what questions we can ask, and how definitively we can answer them. By averaging tiredness across a large sample and performing a GWAS, the present study will primarily pick upon the genetic links between tiredness and inherent factors. These are likely to be various, so we will seek to bring some clarity to what we consider to be the likely genetic heterogeneity of tiredness by posing the following questions:

1. Is there a direct genetic contribution to self-reported tiredness *per se*, not accounted for the factors in questions 2–4 below?
2. Is tiredness genetically linked to proneness to health-related traits?
3. Is tiredness genetically linked to a systemic proneness to poor health?
4. Is there a genetic relationship between the personality trait of neuroticism and tiredness?

With regard to questions 2 and 3, it is important to remember that a positive answer may constitute more than the obvious conclusion that the presence of an illness phenotype is inevitably accompanied by tiredness. Our analyses of UK Biobank data capture genetic predisposition to illness rather than its actual presence; we will address this in our polygenic prediction sensitivity analysis.

The aim of the present study, then, is to understand further the genetic contribution to self-reported tiredness and/or low energy. We conducted a genome-wide association analysis, in the UK Biobank sample, of a response to a single item question: ‘Over the past two weeks, how often have you felt tired or had little energy?’ On the basis of the foregoing literature overview, we also investigated pleiotropy with physical- and mental health-related traits, and we specifically investigated pleiotropy with factors associated with allostatic load as a first step towards answering question three above. The current study design, using > 100 000 UK Biobank participants, directly addresses the main limitation from the previous studies by substantially increasing the sample size. In addition, we complement this with SNP-based heritability estimates of tiredness, sex- and age-stratified analysis, and an examination of the genetic overlap of tiredness with many health-related traits.

## MATERIALS AND METHODS

### Study design and participants

UK Biobank is a large resource for identifying determinants of human diseases in middle-aged and older individuals (<http://www.ukbiobank.ac.uk>).<sup>43</sup> A total of 502 655 community-dwelling individuals aged between 37 and 73 years were recruited in the United Kingdom between 2006 and 2010. Baseline assessment included cognitive testing, personality self-report, and physical and mental health measures. For the present study, genome-wide genotyping data were available for 112 151 participants (58 914 females and 53 237 males) after quality control (see below). They were aged from 40 to 73 years (mean = 56.9 years, s.d. = 7.9). UK Biobank received ethical approval from the Research Ethics Committee (REC reference for UK Biobank is 11/NW/0382). This study has been completed under UK Biobank application 10279. Figure 1 shows the study flow for the present report.

### Procedures

**Tiredness.** Participants were asked the question, ‘Over the past two weeks, how often have you felt tired or had little energy?’ Possible answers were: ‘Not at all/Several days/More than half the days/Nearly every day/Do not know/Prefer not to answer’. This question was asked as part of the Mental Health Questionnaire, which consists of items from the Patient Health Questionnaire.<sup>44</sup> Participants answering with ‘Do not know’ or ‘Prefer not to answer’ were excluded, resulting in a four-category variable for tiredness ranging from ‘Not at all’ to ‘Nearly every day’. We will refer to

this question in the rest of the paper as ‘tiredness’, but we ask the reader to bear in mind the question as it was asked, that is, its referring to tiredness and/or low energy.

**Genotyping and quality control.** The interim release of UK Biobank included genotype data for 152 729 individuals, of whom 49 979 were genotyped using the UK BiLEVE array and 102 750 using the UK Biobank axiom array. These arrays have over 95% content in common. Details of the array design, genotyping procedures and quality control details have been published elsewhere.<sup>45,46</sup>

**Imputation.** An imputed data set was made available as part of the UK Biobank interim data release. The 1000 Genomes phase 3 and UK10K haplotype reference panels were merged and the genotype data imputed to this merged reference panel. Further details can be found at the following URL: <http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=157020>. Autosomal variants with a minor allele frequency  $\leq 0.1\%$  and an imputation quality score of  $< 0.1$  were excluded from further analysis ( $N \sim 17.3M$  SNPs).

**Curation of summary results from GWAS consortia on health-related variables.** Published summary results from international GWAS consortia were gathered to derive genetic correlations using the linkage disequilibrium (LD) score regression method and perform polygenic profile score analysis between the UK Biobank tiredness variable and the genetic predisposition to multiple health-related traits. Details of the health-related variables, the consortia’s websites, key references for each consortium and number of subjects included in each consortium’s GWAS are given in Supplementary Table 1.

### Statistical analysis

**Phenotypic correlations.** Spearman’s rank correlation coefficients were calculated between responses to the tiredness question, grip strength, forced expiratory volume in 1 s, height, BMI, self-rated health, verbal-numerical reasoning and neuroticism, all of which were measured phenotypes in UK Biobank. Details on measurements of these phenotypes can be found in the Supplementary Material.

**Genetic association analysis.** A total of 111 749 individuals answered the tiredness question and had genotypic information. After visual inspection of the distribution of the UK Biobank tiredness variable no exclusions were made. Prior to analysis, tiredness was adjusted for age, sex, assessment centre, genotyping batch and array, and 10 principal components for population stratification. Genotype–phenotype association analyses were conducted using SNPTEST v2.5.1 (ref. 47) and can be found at the following URL: [https://mathgen.stats.ox.ac.uk/genetics\\_software/snpstest/snpstest.html#introduction](https://mathgen.stats.ox.ac.uk/genetics_software/snpstest/snpstest.html#introduction). An additive model was specified using the ‘frequentist 1’ option. Genotype uncertainty was accounted for by analysing genotype dosages.

Genetic association analyses were also performed on the following UK Biobank phenotypes to perform further analyses: self-rated health,<sup>48</sup> grip strength, forced expiratory volume in 1 s, neuroticism,<sup>49</sup> verbal-numerical reasoning.<sup>50</sup> We specifically examined whether any variants associated with tiredness were also associated with grip strength, self-rated health and neuroticism because we judged these to provide some coverage of physical and mental resilience in UK Biobank.

**Estimation of SNP-based heritability.** To estimate the proportion of variance explained by all common SNPs in tiredness, univariate GCTA-GREML analysis was performed.<sup>51</sup> This analysis included only unrelated individuals, using a relatedness cutoff of 0.025 in the generation of the genetic relationship matrix.

**Gene-based association analysis.** Gene-based associations were derived using MAGMA,<sup>52</sup> using the summary GWAS statistics for tiredness. SNPs were assigned to 18 062 genes using the National Center for Biotechnology Information build 37.3. The gene boundary was defined as the start and stop site of each gene. To account for LD between the SNPs used, the European panel of the 1000 Genomes data (phase 1, release 3) was used. A Bonferroni correction was used to control for 18 062 tests ( $\alpha = 0.05/18\ 062$ ;  $P < 2.768 \times 10^{-6}$ ).

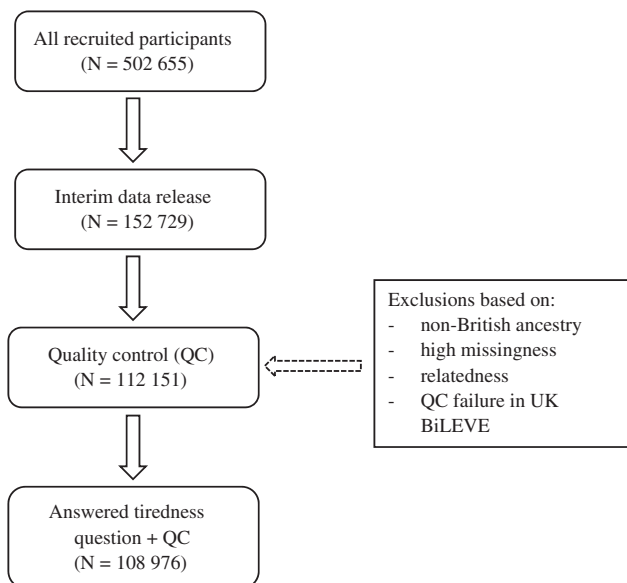


Figure 1. Flow diagram of participant selection.

**Partitioned heritability.** The summary statistics from the GWAS on tiredness was partitioned into functional categories using the same data processing pipeline as Finucane *et al.*<sup>53</sup> more details on this method can be found in the Supplementary Materials.

**Shared genetic aetiology: LD score regression and polygenic profiling.** Genetic associations between tiredness and health-related variables from GWAS consortia were computed using two methods, LD score regression and polygenic profile score analysis. Each provides a different metric to infer the existence of loci contributing to pairs of traits. LD score regression was used to derive genetic correlations between two traits; this tests the degree to which the polygenic architecture of one trait overlaps with that of other traits. Polygenic profile score analysis was used to test the extent to which individual differences in the tiredness phenotype in UK Biobank could be predicted by polygenic profile scores predictive of the health-related traits from other GWAS consortia. Both of these methods are dependent each on trait being polygenic in nature, that is, where a large number of variants of small effect contribute towards phenotypic variation.<sup>54,55</sup> Bivariate LD score regression was performed between tiredness and 29 health-related traits. Polygenic profile score analysis was performed on 26 of the 29 health-related traits, as this method requires independent samples to provide the summary GWAS information from which the polygenic profile score is computed.

**Bivariate LD score regression:** This was used to quantify the extent of genetic overlap between tiredness in UK Biobank and 29 health-related traits.<sup>55,56</sup> This technique examines the correlational structure of the SNPs found across the genome. In the present study, LD score regression was used to derive genetic correlations between tiredness and health-related traits using the GWAS results of 25 large GWAS consortia and four UK Biobank phenotypes. The data processing pipeline devised by Bulik-Sullivan *et al.*<sup>55</sup> was followed. To ensure that the genetic correlation for the Alzheimer's disease phenotype was not driven by a single locus or biased the fit of the regression model, a 500-kb region centred on the *APOE* locus was removed and this phenotype was re-run. This additional model is referred to in the tables and figures as 'Alzheimer's disease (500 kb)'.

**Polygenic profile scores:** The UK Biobank genotyping data were recoded from numeric (1,2) allele coding to standard ACGT coding using a bespoke programme developed by one of the present authors (DCML).<sup>46</sup> Polygenic profile scores were created for 25 health-related traits in all genotyped participants using PRSice.<sup>57</sup> Prior to creating the scores, SNPs with a minor allele frequency <0.01 were removed and clumping was used to obtain SNPs in linkage equilibrium with an  $r^2 < 0.25$  within a 200 bp window. Five polygenic profile scores were created for each trait including SNPs according to their significance of association with the relevant trait at  $P$ -value thresholds of  $P < 0.01$ ,  $P < 0.05$ ,  $P < 0.1$ ,  $P < 0.5$ , and all SNPs.

Regression models were used to examine the association between the polygenic profile scores and tiredness in UK Biobank, adjusting for age at measurement, sex, genotyping batch and array, assessment centre, and the first 10 principal components for population stratification. All associations were corrected for multiple testing using the false discovery rate (FDR) method.<sup>58</sup> Sensitivity analyses were performed to test whether the results were confounded by individual's neuroticism levels, their self-rated health scores or a diagnosis of major depressive disorder. This was done by adjusting the models for the neuroticism and self-rated health scores. Individuals with a probable diagnosis of major depressive disorder were excluded from the sensitivity analysis, based on the diagnostic method formulated by Smith *et al.*<sup>59</sup> Further details can be found in the Supplementary Material. To examine whether any association between polygenic profile score for type 2 diabetes and tiredness was confounded by having had a diagnosis of type 2 diabetes, all individuals with a self-reported doctor's diagnosis of type 2 diabetes were excluded from that specific sensitivity analysis (Supplementary Material). Multivariate regression was performed using all FDR significant polygenic profile scores and earlier described covariates.

**Comparison of gene-based analysis results within UK Biobank.** Gene-based associations for tiredness were compared with gene-based results for other UK Biobank health-related traits that, in the present report's results, showed a statistically significant genetic correlation with tiredness, using previously described methods.<sup>52</sup>

**Age- and sex-stratified analysis.** On the basis of the age and sex distribution for tiredness (Supplementary Figure 1), further analyses examining

potential age and sex effects were performed. The sample was split by sex, as well as the following three age groups for each sex: 40 to < 50 years, 50 to < 60 years and 60 to < 70 years, one male aged > 70 years was excluded from these analyses. The analysis included heritability estimates for the eight different groups, genome-wide association analysis, and genetic correlations with BMI and waist-hip ratio, as these summary data were available separately for males and females. All models were adjusted for age (sex-stratified analyses), sex (age-stratified analysis) and the previously mentioned covariates (assessment centre, genotyping batch and array, and 10 principal components for population stratification).

**Code availability.** The code used to run the analysis is available from the authors upon request.

## RESULTS

### Phenotypic correlations

A total of 108 976 individuals from UK Biobank with genotypic data answered the question 'Over the past two weeks, how often have you felt tired or had little energy?', referred to hereinafter as 'tiredness'. There were 51 416 individuals who answered 'not at all', 44 208 individuals responded 'several days', 6404 individuals answered 'more than half the days' and 6948 individuals responded 'nearly every day'. Correlations indicated that individuals who reported feeling more tired tended to have lower grip strength, lower lung function, poorer self-rated health, lower scores for verbal-numerical reasoning and shorter stature (Table 1). Correlations indicated that individuals who reported feeling more tired tended to have a higher BMI and higher neuroticism scores (Table 1). Absolute effect sizes ranged from very small to moderate (Supplementary Figures 2a–f). The mean scores and distribution for each of these variables at each level of tiredness are shown in Supplementary Figure 2 and Supplementary Table 2.

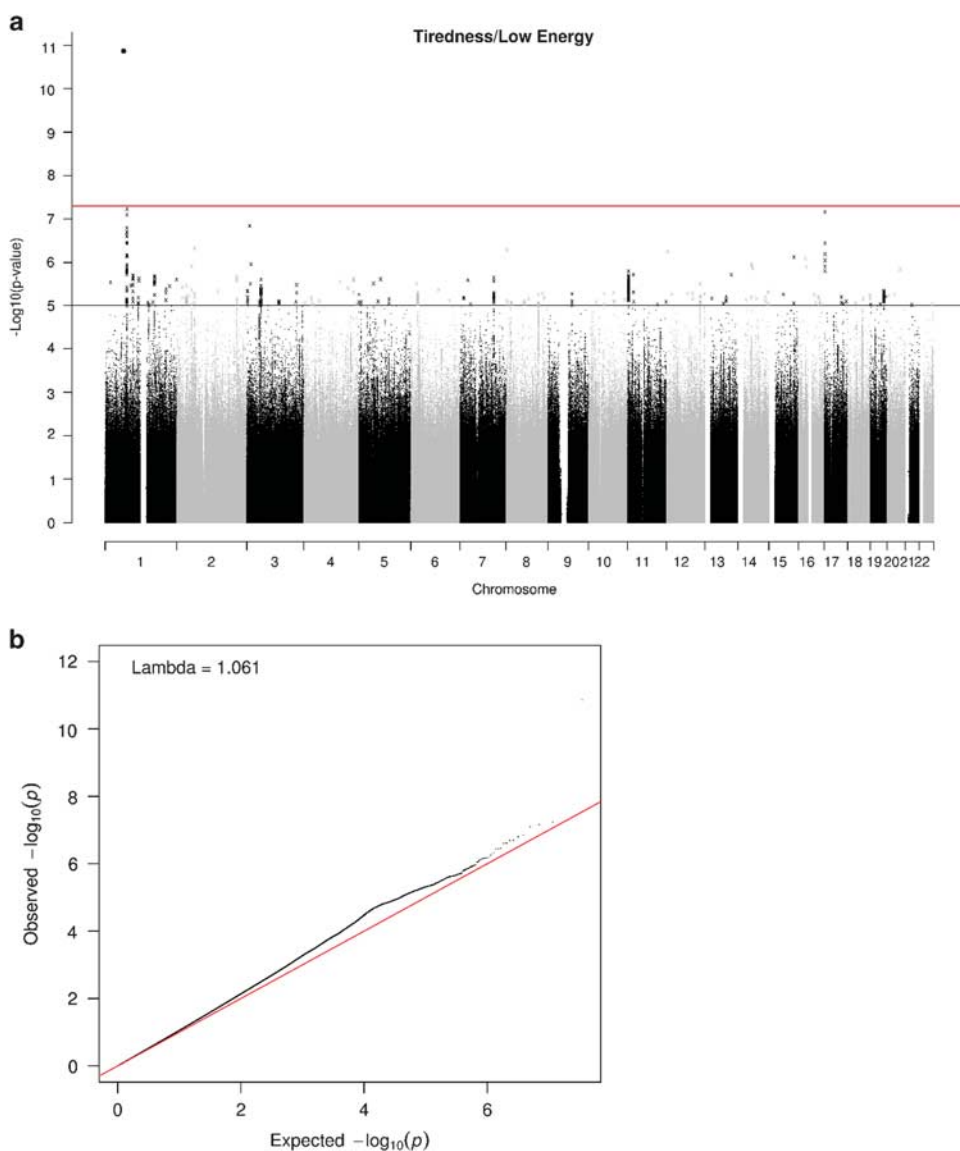
### Genome-wide association study

There was one genome-wide significant SNP (Affymetrix id 1:64178756\_C\_T;  $P = 1.36 \times 10^{-11}$ ) on chromosome 1 (Figure 2). This SNP is not in a gene and does not have an rs id. It has both a low minor allele frequency (0.001) and a low imputation quality score (0.43). It is not in a peak with other SNPs. Therefore, this result should be treated with caution. Two suggestive peaks were identified on chromosomes 1 and 17, with the lowest  $P$ -values of  $5.88 \times 10^{-8}$  (rs142592148; an intronic SNP in *SLC44A5*) and  $6.86 \times 10^{-8}$  (rs7219015; an intronic SNP in *PAFAH1B1*) for each peak, respectively. The peak on chromosome 1 contains three genes (*CRYZ*, *TYW3* and *SLC44A5*). The peak on chromosome 17 contains one gene (*PAFAH1B1*). The *CRYZ*/*TYW3* locus has previously been associated with circulating resistin levels, a hormone associated with insulin resistance, inflammation, and risk of type 2 diabetes and cardiovascular disease.<sup>60</sup> *SLC44A5* encodes a solute carrier protein and is important for metabolism of lipids and

**Table 1.** Spearman phenotypic correlations between tiredness (responses to the question, 'Over the past two weeks, how often have you felt tired or had little energy?') and physical and mental health

	Tiredness
Self-rated health (N = 108 648)	−0.35
Grip strength (N = 108 573)	−0.12
Forced expiratory function in 1 s (N = 101 823)	−0.08
Height (N = 108 796)	−0.09
BMI (N = 108 681)	0.10
Verbal-numerical reasoning (N = 35 101)	−0.04
Neuroticism (N = 105 456)	0.39

Abbreviation: BMI, body mass index. All correlations had  $P < 0.001$ .



**Figure 2.** (a) Manhattan and (b) Q-Q plot of *P*-values of the SNP-based association analysis of tiredness (responses to the question, ‘Over the past two weeks, how often have you felt tired or had little energy?’). The red line on the Manhattan plot indicates the threshold for genome-wide significance ( $P < 5 \times 10^{-8}$ ); the grey line on the Manhattan plot indicates the threshold for suggestive significance ( $P < 1 \times 10^{-5}$ ). SNP, single-nucleotide polymorphism.

lipoproteins, and has been associated with birth weight in cattle.<sup>61</sup> *PAFAH1B1* encodes a subunit of an enzyme that has important roles in brain development and spermatogenesis. Mutations in this gene cause the neurological disorder lissencephaly.<sup>62</sup>

#### SNP-based heritability estimate

Using GCTA-GREML common SNPs were found to explain 8.4% (s.e. 0.6%) of the phenotypic variation of tiredness as measured in UK Biobank.

#### Gene-based association analysis

Gene-based association analysis identified five genes, *DRD2*, *PRRC2C*, *C3orf84*, *ANO10* and *ASXL3*, that attained genome-wide

**Table 2.** The genome-wide significant genes from the UK Biobank tiredness phenotype and the significance values for the same genes using the neuroticism, SRH and grip phenotypes, also in the UK Biobank sample

CHR	Gene	Tiredness <i>P</i>	SRH <i>P</i>	Grip <i>P</i>	Neuroticism <i>P</i>
11	<i>DRD2</i>	$2.94 \times 10^{-7}$	0.012	0.156	$9.69 \times 10^{-9}$
1	<i>PRRC2C</i>	$1.43 \times 10^{-6}$	0.002	0.314	0.020
3	<i>C3orf84</i>	$1.45 \times 10^{-6}$	$7.38 \times 10^{-5}$	0.001	0.016
3	<i>ANO10</i>	$1.52 \times 10^{-6}$	0.058	0.728	0.001
18	<i>ASXL3</i>	$2.67 \times 10^{-6}$	$1.03 \times 10^{-6}$	0.052	$1.36 \times 10^{-5}$

Abbreviations: grip, grip strength; SRH, self-rated health. Unmodified *P*-values are shown for all phenotypes.

significance for tiredness (Table 2 and Supplementary Table 3) following correction for multiple comparisons. *DRD2* encodes a dopamine receptor and has previously been associated with psychiatric illnesses.<sup>63</sup> Alternative splicing of *PRRC2C* has been associated with lung cancer.<sup>64</sup> Mutations in *ANO10* cause cerebellar ataxias.<sup>65</sup> *ASXL3* encodes a polycomb protein and mutations in this gene are associated with intellectual disability, feeding problems and distinctive facial features.<sup>66</sup>

In addition, each of these genes was also nominally significant in a GWAS of neuroticism, with *DRD2* being genome-wide significant in both phenotypes.<sup>49</sup> A comparison with a UK Biobank GWAS of self-rated health<sup>22</sup> showed that four of the five genes (*DRD2*, *PRRC2C*, *C3orf84* and *ASXL3*) were significant across phenotypes. Grip strength showed less overlap, with only *C3orf84* being nominally associated with grip strength. Of the five genes examined here, *C3orf84* was associated with each of these four phenotypes. Whereas the genetic correlations between these traits (see below) are likely to encompass multiple genic and non-genic regions, as well as unique points of overlap between pairs of phenotypes, the variants found in the *C3orf84* represent a point of the genome where the genetic architecture of these four traits converges.

#### Partitioned heritability

From the full baseline model using 52 annotations, only evolutionarily conserved regions were found to be enriched for tiredness (Supplementary Figure 3). This annotation contained only 2.6% of the SNPs from the summary statistics, but they collectively explained 40% of the heritability of tiredness (s.e.= 11%, enrichment metric = 15.34, s.e.= 4.05,  $P=0.0004$ ). By clustering the histone marks into tissue-specific categories, we found significant enrichment for variants found in the central nervous system (Supplementary Figure 4). This category contained 15% of the SNPs and explained 45% of the heritability (s.e.= 8%, enrichment metric = 3.02, s.e.= 0.54,  $P=0.0002$ ).

#### Genetic correlations between tiredness and physical and mental health traits

LD score regression was used to test whether genetic variants associated with health-related traits also contribute towards tiredness in UK Biobank. Table 3 and Figure 3 show these genetic correlations. Positive significant (FDR corrected) genetic correlations were found between tiredness and BMI ( $r_g=0.20$ ), C-reactive protein ( $r_g=0.17=0.02$ ), HbA1c ( $r_g=0.25$ ), obesity ( $r_g=0.21$ ), smoking status ( $r_g=0.20$ ), triglycerides ( $r_g=0.13$ ), type 2 diabetes ( $r_g=0.18$ ), waist-hip ratio ( $r_g=0.28$ ), attention deficit hyperactivity disorder ( $r_g=0.27$ ), bipolar disorder ( $r_g=0.14$ ), major depressive disorder ( $r_g=0.59$ ), neuroticism ( $r_g=0.62$ ) and schizophrenia ( $r_g=0.25$ ). Negative significant (FDR corrected) genetic correlations were found between high-density lipoprotein (HDL) cholesterol ( $r_g=-0.11$ ), forced expiratory volume in 1 s ( $r_g=-0.12$ ), grip strength ( $r_g=-0.16$ ), longevity ( $r_g=-0.39$ ), self-rated health ( $r_g=-0.78$ ) and verbal-numerical reasoning ( $r_g=-0.14$ ). These genetic correlations suggest that there are common genetic associations between tiredness and multiple physical- and mental health-related traits. Supplementary Table 8 shows the genetic correlations between traits associated with the concept of allostatic load (blood pressure, BMI, cholesterol, C-reactive protein, HbA1c, obesity, triglycerides and waist-hip ratio), indicating genetic overlap between these traits.

#### Polygenic prediction

The full results including all five thresholds can be found in Supplementary Table 4, as well as the number of SNPs included for the five thresholds in each trait. Table 4 shows the results for the polygenic profile scores analyses, using the most predictive

**Table 3.** Genetic correlations between tiredness documented in the UK Biobank data set and the health-related variables collected from GWAS consortia

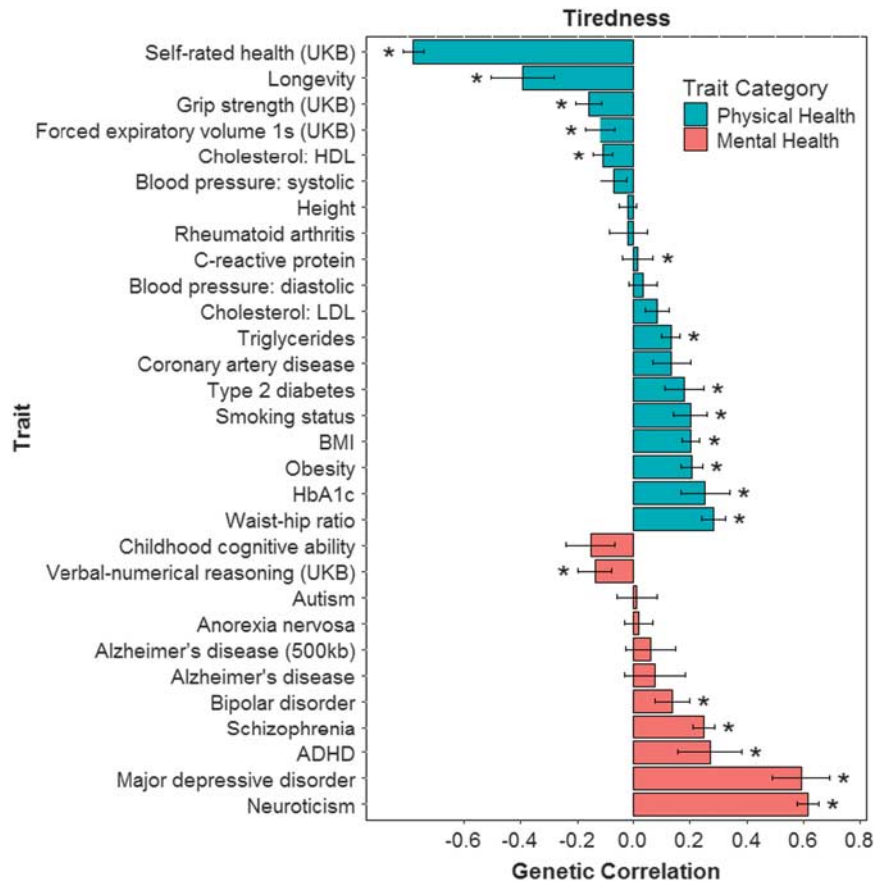
Trait category	Traits from GWAS consortia	$r_g$	s.e.	P
Physical health	Blood pressure: diastolic	0.0332	0.0502	0.5083
	Blood pressure: systolic	-0.0698	0.0478	0.1444
	BMI	0.2024	0.0322	<b><math>3.18 \times 10^{-10}</math></b>
	Cholesterol: HDL	-0.1087	0.0373	<b>0.0036</b>
	Cholesterol: LDL	0.0829	0.0413	0.0449
	Coronary artery disease	0.1338	0.067	0.0459
	C-reactive protein	0.0165	0.054	<b>0.0021</b>
	Grip strength <sup>a</sup>	-0.1596	0.0482	<b>0.0009</b>
	HbA1c	0.2536	0.0857	<b>0.0031</b>
	Height	-0.0201	0.0297	0.4980
	Longevity	-0.3943	0.1096	<b>0.0003</b>
	Forced expiratory volume 1s <sup>a</sup>	-0.1181	0.0538	<b>0.0281</b>
	Obesity	0.2063	0.0381	<b><math>6.31 \times 10^{-8}</math></b>
	Rheumatoid arthritis	-0.0181	0.0674	0.7885
	Self-rated health <sup>a</sup>	-0.7780	0.0349	<b><math>7.30 \times 10^{-110}</math></b>
	Smoking status	0.2009	0.0603	<b>0.0009</b>
	Triglycerides	0.1324	0.0332	<b><math>6.62 \times 10^{-5}</math></b>
	Type 2 diabetes	0.1784	0.0689	<b>0.0097</b>
	Waist-hip ratio	0.2834	0.0417	<b><math>1.09 \times 10^{-11}</math></b>
	Mental health	ADHD	0.2694	0.1116
Alzheimer's disease		0.0762	0.1079	0.4801
Alzheimer's disease (500 kb)		0.0613	0.0872	0.4816
Anorexia nervosa		0.0192	0.0492	0.6967
Autism		0.0129	0.0695	0.8522
Bipolar disorder		0.1382	0.0605	<b>0.0223</b>
Childhood cognitive ability		-0.1528	0.0891	0.0864
Major depressive disorder		0.5902	0.1015	<b><math>6.03 \times 10^{-9}</math></b>
Neuroticism		0.6150	0.038	<b><math>7.34 \times 10^{-59}</math></b>
Schizophrenia		0.2490	0.0386	<b><math>1.14 \times 10^{-10}</math></b>
Verbal-numerical reasoning <sup>a</sup>		-0.1379	0.0596	<b>0.0206</b>

Abbreviations: ADHD, attention deficit hyperactivity disorder; BMI, body mass index; FDR, false discovery rate; GWAS, genome-wide association study; HDL, high-density lipoprotein; LDL, low-density lipoprotein. Statistically significant  $P$ -values (after false discovery rate correction; threshold:  $P=0.0281$ ) are shown in bold. <sup>a</sup>GWAS based on UK Biobank data.

threshold for each trait. Higher polygenic profile scores for 10 physical health traits predicted increased tiredness (significant standardised  $\beta$ 's between 0.008 and 0.026) in UK Biobank: BMI, low-density lipoprotein (LDL) cholesterol, coronary artery disease, C-reactive protein, HbA1c, obesity, smoking status, triglycerides, type 2 diabetes and waist-hip ratio. Higher polygenic profile scores for HDL cholesterol and height predicted lower tiredness (significant standardised  $\beta$ 's between  $\beta=-0.016$  and  $\beta=-0.008$ , respectively).

Of the mental health traits, higher polygenic profile scores for bipolar disorder, neuroticism, major depressive disorder and schizophrenia were associated with increased tiredness (standardised  $\beta$ 's between 0.008 and 0.028). Polygenic profile scores for childhood cognitive ability showed a negative association with tiredness ( $\beta=-0.011$ ).

Sensitivity analysis showed that, when controlling for neuroticism, the associations between tiredness and polygenic profile scores for BMI, obesity, type 2 diabetes, cholesterol (HDL and LDL), C-reactive protein, HbA1c, triglycerides, waist-hip ratio, childhood



**Figure 3.** Barplot of genetic correlations (s.e.) calculated using LD regression between tiredness in UK Biobank and mental and physical health measures from GWAS consortia. \* $P < 0.0281$ . ADHD, attention deficit hyperactivity disorder; BMI, body mass index; GWAS, genome-wide association study; HDL, high-density lipoprotein; LD, linkage disequilibrium; LDL, low-density lipoprotein.

cognitive ability and schizophrenia remained significant (after FDR correction), indicating that these associations are not wholly confounded by scores for neuroticism. Similar analyses controlling for self-rated health indicated that the following associations with tiredness are not wholly confounded by self-rated health: bipolar disorder, neuroticism, major depressive disorder and schizophrenia. Supplementary Table 5 shows the adjusted results and the percentage of attenuation in standardised  $\beta$ 's for the models.

When excluding individuals with a probable diagnosis of major depressive disorder ( $N = 7364$ ) from all individuals with sufficient information about their mental health to make a probable depression diagnosis (full model,  $N = 31\,523$ ), the associations between tiredness and eight polygenic profile scores remained significant (after FDR correction), compared to nine in the full model, the association between type 2 diabetes and tiredness became non-significant after excluding individuals with probable diagnosis of major depressive disorder (Supplementary Table 6). When excluding individuals with a type 2 diabetes diagnosis ( $N = 725$ ), the association between tiredness and the polygenic risk score for diabetes remained significant, indicating that this association is independent of self-reported morbidity for that disorder.

A multivariate regression model including 17 significant polygenic profile scores (BMI, HDL cholesterol, LDL cholesterol, coronary artery disease, C-reactive protein, HbA1c, height, obesity, smoking status, triglycerides, type 2 diabetes, waist-hip ratio, bipolar disorder, childhood cognitive ability, major depressive disorder, neuroticism and schizophrenia) showed that polygenic profile scores for the following traits contributed independently to

the association with tiredness: BMI, HDL cholesterol, triglycerides, waist-hip ratio, childhood cognitive ability, major depressive disorder, neuroticism and schizophrenia. The scores together accounted for 0.25% of the variance in tiredness (Supplementary Table 7).

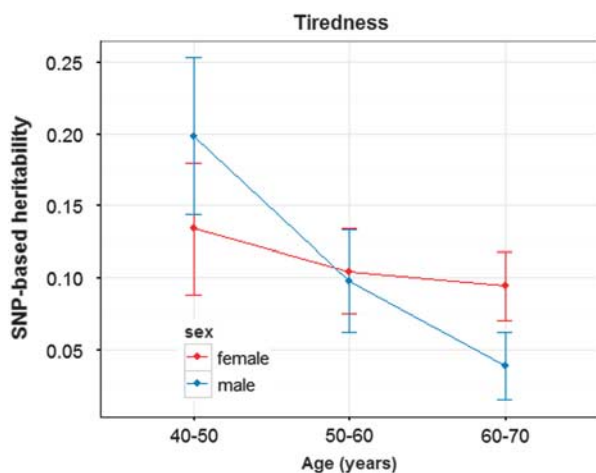
#### Age- and sex-stratified analysis

GCTA-GREML analysis was used to test for possible differences in the heritability estimates for tiredness in different age/sex groups. The proportion of variance in tiredness explained by all common genetic variants using GCTA-GREML was 9.4% (s.e. = 1%,  $N = 57\,165$ ) in females and 8.2% (s.e. = 1%,  $N = 51\,811$ ) in males. Figure 4 shows heritability estimates for the three age groups in men and women (40 to < 50, 50 to < 60 years, and 60 to < 70 years). The greatest differences can be seen between males aged 40 and 50 years ( $h^2 = 19.8\%$ , s.e. = 6%,  $N = 10\,798$ ) and males aged 60–70 years ( $h^2 = 3.8\%$ , s.e. = 2%,  $N = 24\,467$ ). Genome-wide association analysis indicated no significant sex differences between males and females, but did show some significant age differences in males (Supplementary Figures 5). Genetic correlations between BMI and tiredness were not significantly different for males ( $r_g = 0.15$ , s.e. = 0.06, 95% confidence interval (CI) = 0.04–0.26,  $P = 0.0074$ ) and females ( $r_g = 0.26$ , s.e. = 0.05, 95% CI = 0.16–0.36,  $P = 4.25 \times 10^{-7}$ ), as the confidence intervals were overlapping. Also, no significant differences in the genetic correlations were found between waist-hip ratio and tiredness for males ( $r_g = 0.30$ , s.e. = 0.08, 95% CI = 0.14–0.45,  $P = 0.0003$ ) and females ( $r_g = 0.271$ , s.e. = 0.06, 95% CI = 0.15–0.40,  $P = 2.2 \times 10^{-5}$ ), with overlapping confidence intervals.

**Table 4.** Associations between polygenic profile scores of health-related traits created from GWAS consortia summary data, and the UK Biobank tiredness phenotype controlling for age, sex, assessment centre, genotyping batch, and array and 10 principal components for population structure

Trait category	Trait	Threshold	$\beta$	P	
Physical health	Blood pressure: diastolic	0.1	-0.0028	0.3619	
	Blood pressure: systolic	0.1	-0.0025	0.4077	
	BMI	1	0.0280	<b><math>4.90 \times 10^{-20a}</math></b>	
	Cholesterol: HDL	0.5	-0.0163	<b><math>8.49 \times 10^{-8a}</math></b>	
	Cholesterol: LDL	0.5	0.0081	<b>0.0077<sup>a</sup></b>	
	Coronary artery disease	0.5	0.0084	<b>0.0061</b>	
	C-reactive protein	1	0.0130	<b><math>2.10 \times 10^{-5a}</math></b>	
	Forced expiratory volume 1 s	0.01	-0.0059	0.0529	
	Longevity	0.05	-0.0067	0.0297	
	HbA1c	1	0.0090	<b>0.0033<sup>a</sup></b>	
	Height	1	-0.0077	<b>0.0154</b>	
	Obesity	1	0.0236	<b><math>1.20 \times 10^{-14a}</math></b>	
	Rheumatoid arthritis	0.1	-0.0016	0.5926	
	Smoking status	0.5	0.0086	<b>0.0071</b>	
	Triglycerides	0.5	0.0209	<b><math>1.06 \times 10^{-11a}</math></b>	
	Type 2 diabetes	1	0.0120	<b>0.0002<sup>a,b</sup></b>	
	Waist-hip ratio	1	0.0258	<b><math>7.85 \times 10^{-17a}</math></b>	
	Mental health	ADHD	1	0.0042	0.1647
		Alzheimer's disease	0.05	-0.0052	0.0889
Anorexia nervosa		0.5	0.0048	0.1169	
Autism		1	-0.0018	0.5593	
Bipolar disorder		0.01	0.0081	<b>0.0076<sup>c</sup></b>	
Childhood cognitive ability		0.1	-0.0112	<b>0.0002<sup>a</sup></b>	
Major depressive disorder		1	0.0185	<b><math>2.25 \times 10^{-9c}</math></b>	
Neuroticism		0.1	0.0183	<b><math>2.00 \times 10^{-9c}</math></b>	
Schizophrenia		1	0.0283	<b><math>2.31 \times 10^{-19a,c}</math></b>	

Abbreviations: ADHD, attention deficit hyperactivity disorder; BMI, body mass index; FDR, false discovery rate; GWAS, genome-wide association study; HDL, high-density lipoprotein; LDL, low-density lipoprotein. FDR-corrected statistically significant values ( $P=0.0255$ ) are shown in bold. The associations between the polygenic profile scores with the largest effect size (threshold) and tiredness are presented. Threshold is the  $P$ -value threshold with the largest effect size. <sup>a</sup>Results remain significant after controlling for neuroticism scores. <sup>b</sup>Results remain significant after excluding individuals with type 2 diabetes ( $\beta=0.0105$ ,  $P=0.00076$ ). <sup>c</sup>Results remain significant after controlling for self-rated health.



**Figure 4.** Age- and sex-stratified heritability estimates with s.e.'s for tiredness. SNP, single-nucleotide polymorphism.

**DISCUSSION**

In the present study, shared genetic aetiology was identified between tiredness and longevity, grip strength, multiple metabolic indicators, smoking status, neuroticism, childhood cognitive ability, depression and schizophrenia. These analyses, combining data from the UK Biobank and many GWAS consortia, provide the first estimate of the overlap in the genetic variants contributing to the heritability of tiredness and these physical and mental health-related traits and disorders. Tiredness demonstrated a significant SNP-based heritability of 8.4%.

In answer to our first research question—is there a direct genetic contribution to self-reported tiredness *per se*?—we found that, whereas there was no large influence on tiredness from common genetic variants, five genes attained genome-wide significance for tiredness: *DRD2*, *PRRC2C*, *ANO10*, *ASXL3* and *C3orf84*. The latter is an uncharacterised protein representing a point of genetic convergence between tiredness, neuroticism, grip strength and self-rated health. *DRD2*, *PRRC2C*, *ANO10* and *ASXL3* have previously been associated with psychiatric illnesses,<sup>18</sup> lung cancer,<sup>19</sup> cerebellar ataxias<sup>20</sup> and intellectual disability,<sup>21</sup> respectively. The three genes within the suggestive peak on chromosome 1 (*CRYZ*, *TYW3* and *SLC44A5*) have previously been associated with insulin resistance, inflammation, risk of type 2 diabetes, metabolism of lipids and lipoproteins, and cardiovascular disease.<sup>15,16</sup> These genes are consistent with the identification of regions associated with both tiredness and metabolic irregularities, and perhaps more broadly with 'metabolic syndrome' and 'allostatic load'. *PAFAH1B*, within a suggestive peak on chromosome 17 has important roles in brain development.<sup>17</sup> This is consistent with the identification of regions associated with both tiredness and cognitive traits, and the finding of significant enrichment for variants found in the central nervous system.

Evolutionarily conserved regions were found to be enriched for association with tiredness, consistent with findings for other quantitative traits including disease status,<sup>53</sup> suggesting that these are important loci where common additive SNPs cluster to produce phenotypic variation in many traits, as explored in more detail in the paper by Hill *et al.*<sup>67</sup> The range of factors— affective, cognitive, behavioural and physical—that are genetically associated with tiredness is in itself remarkable, and confirms the observation of Landmark-Høyvik *et al.*,<sup>34</sup> quoted in the introduction, that the related construct of fatigue is aetiologically



heterogeneous and multifactorial. No overlap has been found with genes (*GRIK2*, *NPAS2*) identified in previous GWAS of fatigue;<sup>39,40</sup> however, the sample size of these studies was too small to have enough power to detect statistically significant differences. The present study did not show significant associations for candidate genes previously identified.<sup>34</sup>

The results of the present study add to the body of evidence that tiredness has a genetic underpinning.<sup>35,36,38</sup> This study estimated the SNP-based heritability of self-reported tiredness at 8.4%, and also examined age- and sex-specific heritability. No differences were found between males and females, but the results suggested a higher heritability in males aged between 40 and 50 years, compared to males between 60 and 70 years. Previous twin studies have shown inconsistent results regarding the sex-specific heritability. One study reported a higher heritability for prolonged fatigue (fatigue for more than one month) in males,<sup>36</sup> whereas another reported a higher heritability for 'interfering' fatigue (fatigue for >5 days) in females.<sup>35</sup> One study reported no sex differences in the heritability of chronic fatigue.<sup>36</sup> In summary, the answer to our first question is that, whereas tiredness is, as expected, largely causally heterogeneous, there may be a small but significant, direct genetic contribution to tiredness proneness.

In answer to question two—is tiredness genetically linked to proneness to health-related traits?—we can answer in the affirmative. The range of factors— affective, cognitive, behavioural and physical—that are genetically associated with tiredness is in itself remarkable, and confirms the observation of Landmark-Høyvik *et al*,<sup>34</sup> quoted in the introduction, that the related construct of fatigue is aetiologically heterogeneous and multifactorial. This may seem a relatively trivial finding whether we assume that it simply reflects the sum of genetic factors that are primarily associated with other more specific phenotypes, which cause tiredness in one way or another. However, it is important to recall that the biobank data capture illness propensity rather than actual morbidity. In our sensitivity analysis, we controlled for the presence of type 2 diabetes and found that the genetic link between type 2 diabetes and tiredness remained significant. This would indicate that, for this health marker at least, tiredness and illness proneness are genetically related irrespective of the presence of morbidity. Similarly, the genetic association between tiredness and longevity would argue for a non-trivial link between self-reported tiredness and a more general tendency to poor health.

This takes us into the territory of question three—is tiredness genetically linked to a systemic proneness to poor health? To begin to answer this, we examined the genetic associations between tiredness and putative markers of allostatic load. Tiredness showed significant shared heritability with a range of factors associated with the metabolic syndrome<sup>68</sup> including cholesterol, triglycerides, HbA1c, waist-hip ratio, BMI, obesity and type 2 diabetes. Several of these factors are biomarkers of allostatic load.<sup>61,62</sup> The concept of allostatic load has been used in the context of both physical and mental ill health, including the symptom of fatigue. Conceptually, allostatic load represents the cumulative, physiological 'wear and tear' of a prolonged response to a stressor. Allostatic load has been shown to be a reliably-measurable multi-variate construct,<sup>69,70</sup> with first-order factors comprising cardiovascular, immune, metabolic, anthropometric and neuroendocrine markers.<sup>29</sup> It is hypothesised that, in response to threats to homeostasis, the body's self-regulatory mechanisms, such as the sympathetic-adrenal medullary axis and the hypothalamic-pituitary-adrenal axis, have the potential to 'over-compensate and eventually collapse upon themselves',<sup>29</sup> with consequences for morbidity and mortality.

In our current analysis, we included metabolic (cholesterol, HbA1c and triglycerides), anthropometric (waist-hip ratio, BMI and obesity) and cardiovascular/respiratory (diastolic and systolic

blood pressure, and forced expiratory volume) markers of allostatic load. The results showed significant shared genetic aetiology, as measured by both LD score regression and polygenic profile score analysis, between tiredness and most of the metabolic and anthropometric markers, though not the cardiovascular/respiratory markers. This raises the possibility that the genetic overlap between tiredness and these physiological factors may be due to a biological propensity to an over-compensatory physiological stress response. This suggestion will require further investigation, because a more parsimonious explanation of these links would be that there is a genetic link between tiredness and multiple, separate genetic determinants of poor physical health. However, the substantial genetic correlations between these traits, and between these traits and tiredness provide some evidence that the allostatic load concept does have coherence at the genetic level.

Fourth, to answer the question on the genetic associations between tiredness and the personality trait of neuroticism, which is the tendency to experience negative affective states, these were indeed strongly correlated, both phenotypically and genetically. This may represent a separate route to fatigue, a predominately affective one, and/or it may overlap with the physiological factors described above. A recent paper by Gale *et al*,<sup>49</sup> also using this UK Biobank sample, supports that the physiological and affective dimensions of poor health overlap in neuroticism. That paper showed that polygenic profile scores for several physical and mental health traits—BMI, coronary artery disease, smoking status, bipolar disorder, borderline personality, major depressive disorder, negative affect and schizophrenia—significantly predicted neuroticism. In the present study, when tiredness polygenic profile score analyses were adjusted for neuroticism, the associations between tiredness and mental health disorders (bar schizophrenia) were largely attenuated, whereas most of the metabolic and anthropometric associations remained significant. This suggests that it is the propensity to neuroticism, rather than the specific propensity to these disorders, that accounts or mediates the tiredness associated with mood disorders. Watson and Pennebaker,<sup>71</sup> discussing competing models of how negative affectivity is related to self-reported physical and emotional well-being, found that it is associated as much with the former as with the latter, and that negative affectivity might better be conceptualised as a general tendency to experience both somatic and emotional distress. This concept of a general tendency to what they termed somatopsychic distress could explain the pleiotropy observed in the present study between neuroticism and tiredness.

That neuroticism may also be a distinct route to fatigue is supported by the fact that when the polygenic profile score analysis is adjusted for self-rated health, all associations between polygenic profile scores for physical health and tiredness are attenuated to the point of non-significance, whereas the relationship between tiredness, and polygenic profiles for neuroticism, major depressive disorder and bipolar disorder remain significant. This is consistent with the study of Gale *et al*,<sup>49</sup> investigating shared genetic aetiology between neuroticism and physical and mental health, where there were more and stronger genetic associations between neuroticism and mental health than between neuroticism and physical health. If we take self-rated health to be a marker, to some extent, of actual physical health (and the study by Harris *et al*.<sup>48</sup> would indicate that it is), then this would suggest that when physical health is adjusted for, polygenic profile scores for neuroticism and its associated negative affective states, continue to make a unique contribution to tiredness.

These proposed affective and physiological routes to fatigue may not be mutually exclusive. The allostatic load model, and the multifactorial models of fatigue described in the introduction, postulate that individual differences in personality, cognition and behavioural responses to stress, and socio-cultural factors, affect the physiological stress response. An increased propensity to

experience distress, as captured in the concept of neuroticism, would imply that there is increasing propensity to over-respond to stressors and thus to physiological dysregulation. In the study by Gale *et al.*,<sup>49</sup> the only significant result between neuroticism and physical disease state was a significant association with the polygenic risk score for coronary artery disease, which is suggestive of an overlap of affective and cardiovascular stress responses. This multifactorial understanding of fatigue would also allow us to incorporate childhood cognitive ability (reduced ability to problem solve), smoking status (at least one study has found smoking and allostatic load interaction effects<sup>29</sup>) and grip strength (reduced overall system integrity/vigour) into a more general model that is suggestive of shared genetic variance between stress proneness (neuroticism, reduced cognitive ability and reduced vigour), the physiological response to stress (biomarkers), behavioural responses (smoking), self-reported tiredness, disease and mortality.<sup>72</sup> However, as with our discussion of allostatic load, these links are suggestive not conclusive, and will require further empirical and theoretical investigation.

The large sample size of the present study is a strength of this study, providing powerful and robust tests of shared genetic aetiology between tiredness, and physical and mental health. A second strength is that all genetic samples were processed on the same platform at the same location. The use of summary data from many international GWAS consortia provided foundations for a comprehensive examination of shared genetic aetiology between tiredness and a wide range of health-related phenotypes.

The study has some limitations. The amount of variance explained by the polygenic profile score analysis was small, which would be expected as not all SNPs are genotyped. The SNPs that were genotyped do not necessarily accurately tag the causal genetic variants. All analyses were restricted to individuals of white British ancestry, because the sample does not have enough power to generalise results for individuals with different backgrounds. Also, the sample consisted of middle- and older-aged adults, thus limiting its generalisability to the adult population as a whole. However, as mentioned in the introduction, there are no clear age-related differences in levels of self-reported fatigue. This could be taken as an indication that the phenomenon is fairly stable across the adult life course, at least at the level of phenotype. Whether the genetic determinants are different in younger adults is a topic for future research.

A further limitation of the present study is the fact that tiredness was measured by self-report; that is, that we were looking for objective correlates of a subjective construct. However, as Wessely<sup>73</sup> observed, an objective measure of fatigue is 'an unattainable holy grail'. Almost all the studies cited in the introduction have used subjective self-reports. The self-report measures used vary widely, with there being several validated fatigue measures, and many of the reported studies use either double- or single-item questionnaires and/or single item visual analogue scales. This in itself may account for some of the inconsistency in fatigue research, though the demographic studies cited at the beginning of this article, using a wide variety of measures from single questions<sup>7</sup> to a well-validated fatigue questionnaire,<sup>2</sup> produced similar findings. However, our findings of genetic associations of fatigue will need replication with better validated multi-item measures.

Perhaps a more serious concern is the one signalled in the introduction: that fatigue is too causally heterogeneous a trait to meaningfully study at the genetic level. To address this, it is worth situating this research in the context of other recent and ongoing fatigue research, such as the recently announced National Institute of Health Mechanisms of Fatigue programme. The latter, while acknowledging that fatigue 'is a common co-morbid condition in a multitude of disease conditions', is attempting to define whether 'molecular, cellular or imaging signatures of fatigue can be defined'.<sup>42</sup> Like the psychosocial and biological fatigue research

cited in our introduction, this work is predicated on the notion that, whereas fatigue shows up as a response to many physical and psychosocial stressors, its determinants may be shared across conditions. As such, we should distinguish between the effective cause of fatigue (the illness/stressors that set it going), and the material and formal causes (the bodily and psychosocial processes that produce and maintain the phenomenon).<sup>74</sup> Whereas the effective causes might be various, the material and formal causes are likely to be more limited and shared across individuals and precipitating conditions. Even if this is not the case, we judge that the present study has gone some way to specifying the nature of the heterogeneity of tiredness at the genetic level, and that there are several non-trivial insights that will require further investigation, specifically: the links between tiredness and illness proneness as distinct from actual morbidity; the genetic coherence of the allostatic load concept and its contribution to tiredness; and the nature of the shared genetic and phenotypic links between tiredness and the personality trait of neuroticism. In terms of the genetic contributions to this complex phenomenon, the current study is probably best seen as the first attempt to use a large and relatively well-powered GWAS to identify these areas for future research.

### Summary

Being genetically predisposed to a range of mental and physical health complaints also predisposes individuals to report that they are more tired or lacking in energy. This study confirms that self-reported tiredness is a partly heritable, heterogeneous and complex phenomenon that is phenotypically and genetically associated with affective, cognitive, personality, health and physiological processes. This study also served as a first step in testing some genetic hypotheses from the allostatic load model, finding suggestive links between tiredness and three genes on chromosome one associated with allostatic processes and considerable genetic overlap between tiredness and allostatic markers. We can foresee more tests of these links as more genome-wide genotyping data become available.

### CONFLICT OF INTEREST

UJD is a participant in UK Biobank. The remaining authors declare no conflict of interest.

### ACKNOWLEDGMENTS

This research has been conducted using the UK Biobank Resource. The work was undertaken in The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC, Age UK (Disconnected Mind Project) and Medical Research Council (MRC) is gratefully acknowledged.

### REFERENCES

- van't Leven M, Zielhuis GA, van der Meer JW, Verbeek AL, Bleijenberg G. Fatigue and chronic fatigue syndrome-like complaints in the general population. *Eur J Public Health* 2010; **20**: 251–257.
- Pawlikowska T, Chalder T, Hirsch S, Wallace P, Wright D, Wessely S. Population based study of fatigue and psychological distress. *BMJ* 1994; **308**: 763–766.
- Ricci JA, Chee E, Lorandeanu AL, Berger J. Fatigue in the US workforce: prevalence and implications for lost productive work time. *J Occup Environ Med* 2007; **49**: 1–10.
- Meng H, Hale L, Friedberg F. Prevalence and predictors of fatigue among middle-aged and older adults: evidence from the Health and Retirement study. *J Am Geriatr Soc* 2010; **58**: 2033.
- Lewis G, Wessely S. The epidemiology of fatigue: more questions than answers. *J Epidemiol Community Health* 1992; **46**: 92.
- Cullen W, Kearney Y, Bury G. Prevalence of fatigue in general practice. *Irish J Med Sci* 2002; **171**: 10–12.

- 7 Dolan P, Kudrna L. More yawns, less yawns: fresh evidence on tiredness by age and other factors. *J Gerontol B Psychol Sci Soc Sci* 2013; **70**: 576–580.
- 8 Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J *et al*. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* 2001; **56**: M146–M156.
- 9 Bandeen-Roche K, Xue Q-L, Ferrucci L, Walston J, Guralnik JM, Chaves P *et al*. Phenotype of frailty: characterization in the women's health and aging studies. *J Gerontol A Biol Sci Med Sci* 2006; **61**: 262–266.
- 10 Theorell-Haglöw J, Lindberg E, Janson C. What are the important risk factors for daytime sleepiness and fatigue in women. *Sleep* 2006; **29**: 751–757.
- 11 Fritschi C, Quinn L. Fatigue in patients with diabetes: a review. *J Psychosom Res* 2010; **69**: 33–41.
- 12 Jarosz PA, Davis JE, Yarandi HN, Farkas R, Feingold E, Shippings SH *et al*. Obesity in urban women: associations with sleep and sleepiness, fatigue and activity. *Womens Health Issues* 2014; **24**: e447–e454.
- 13 Bower JE. Cancer-related fatigue [mdash] mechanisms, risk factors, and treatments. *Nat Rev Clin Oncol* 2014; **11**: 597–609.
- 14 Griffiths L, Jones DE. Pathogenesis of primary biliary cirrhosis and its fatigue. *Digestive Dis* 2014; **32**: 615–625.
- 15 Krupp LB, Alvarez LA, LaRocca NG, Scheinberg LC. Fatigue in multiple sclerosis. *Arch Neurol* 1988; **45**: 435–437.
- 16 Hewlett S, Cockshott Z, Byron M, Kitchen K, Tipler S, Pope D *et al*. Patients' perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. *Arthritis Rheum* 2005; **53**: 697–702.
- 17 Ng W-F, Bowman SJ. Primary Sjögren's syndrome: too dry and too tired. *Rheumatology* 2010; **49**: 844–853.
- 18 Alves G, Wentzel-Larsen T, Larsen J. Is fatigue an independent and persistent symptom in patients with Parkinson disease? *Neurology* 2004; **63**: 1908–1911.
- 19 Franssen PM, Bültmann U, Kant I, van Amelsvoort LG. The association between chronic diseases and fatigue in the working population. *J Psychosom Res* 2003; **54**: 339–344.
- 20 Demyttenaere K, De Fruyt J, Stahl SM. The many faces of fatigue in major depressive disorder. *Int J Neuropsychopharmacol* 2005; **8**: 93–105.
- 21 Bültmann U, Kant I, Schröder C, Kasl S. The relationship between psychosocial work characteristics and fatigue and psychological distress. *Int Arch Occup Environ Health* 2002; **75**: 259–266.
- 22 Fernández-Muñoz JJ, Morón-Verdasco A, Cigarán-Méndez M, Muñoz-Hellín E, Pérez-de-Heredia-Torres M, Fernández-de-las-Peñas C. Disability, quality of life, personality, cognitive and psychological variables associated with fatigue in patients with multiple sclerosis. *Acta Neurol Scand* 2015; **132**: 118–124.
- 23 Kitamura H, Shindo M, Tachibana A, Honma H, Someya T. Personality and resilience associated with perceived fatigue of local government employees responding to disasters. *J Occup Health* 2013; **55**: 1–5.
- 24 Barsevick A, Frost M, Zwinderman A, Hall P, Halyard M, Consortium G. I'm so tired: biological and genetic mechanisms of cancer-related fatigue. *Qual Life Res* 2010; **19**: 1419–1427.
- 25 Felger JC, Cole SW, Pace TW, Hu F, Woolwine BJ, Doho GH *et al*. Molecular signatures of peripheral blood mononuclear cells during chronic interferon- $\alpha$  treatment: relationship with depression and fatigue. *Psychol Med* 2012; **42**: 1591–1603.
- 26 Tomas C, Newton J, Watson S. A review of hypothalamic-pituitary-adrenal axis function in chronic fatigue syndrome. *ISRN Neurosci* 2013; **2013**: 784520.
- 27 Klimas NG, Broderick G, Fletcher MA. Biomarkers for chronic fatigue. *Brain Behav Immun* 2012; **26**: 1202–1210.
- 28 Skerrett TN, Moss-Morris R. Fatigue and social impairment in multiple sclerosis: the role of patients' cognitive and behavioral responses to their symptoms. *J Psychosom Res* 2006; **61**: 587–593.
- 29 Juster R-P, McEwen BS, Lupien SJ. Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neurosci Biobehav Rev* 2010; **35**: 2–16.
- 30 Van Kessel K, Moss-Morris R. Understanding multiple sclerosis fatigue: a synthesis of biological and psychological factors. *J Psychosom Res* 2006; **61**: 583–585.
- 31 Dixon JB. The effect of obesity on health outcomes. *Mol Cell Endocrinol* 2010; **316**: 104–108.
- 32 Lasselín J, Layé S, Dexpert S, Aubert A, Gonzalez C, Gin H *et al*. Fatigue symptoms relate to systemic inflammation in patients with type 2 diabetes. *Brain Behav Immun* 2012; **26**: 1211–1219.
- 33 Brown PJ, Roose SP, Fieo R, Liu X, Rantanen T, Sneed JR *et al*. Frailty and depression in older adults: a high-risk clinical population. *Am J Geriatr Psychiatry* 2014; **22**: 1083–1095.
- 34 Landmark-Høyvik H, Reinertsen KV, Loge JH, Kristensen VN, Dumeaux V, Fosså SD *et al*. The genetics and epigenetics of fatigue. *PM R* 2010; **2**: 456–465.
- 35 Schur E, Afari N, Goldberg J, Buchwald D, Sullivan PF. Twin analyses of fatigue. *Twin Res Hum Genet* 2007; **10**: 729–733.
- 36 Sullivan PF, Evengard B, Jacks A, Pedersen NL. Twin analyses of chronic fatigue in a Swedish national sample. *Psychol Med* 2005; **35**: 1327–1336.
- 37 Hur Y-M, Burri A, Spector TD. The genetic and environmental structure of the covariation among the symptoms of insomnia, fatigue, and depression in adult females. *Twin Res Hum Genet* 2012; **15**: 720–726.
- 38 Hickie I, Kirk K, Martin N. Unique genetic and environmental determinants of prolonged fatigue: a twin study. *Psychol Med* 2009; **39**: 259–268.
- 39 Schlauch K, Khaiboullina S, De Meirleir K, Rawat S, Petereit J, Rizvanov A *et al*. Genome-wide association analysis identifies genetic variations in subjects with myalgic encephalomyelitis/chronic fatigue syndrome. *Transl Psychiatry* 2016; **6**: e730.
- 40 Smith AK, Fang H, Whistler T, Unger ER, Rajeevan MS. Convergent genomic studies identify association of GRIK2 and NPAS2 with chronic fatigue syndrome. *Neuropsychobiology* 2011; **64**: 183–194.
- 41 Landmark-Høyvik H, Reinertsen K, Loge J, Fosså S, Børresen-Dale A, Dumeaux V. Alterations of gene expression in blood cells associated with chronic fatigue in breast cancer survivors. *Pharmacogenomics J* 2009; **9**: 333–340.
- 42 The National Institute of Mental Health. Request for Information (RFI): Input on a Planned NIH Common Fund Mechanisms of Fatigue Program 2016.
- 43 Sudlow C, Gallacher J, Allen N, Burton P, Danesh J *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* 2015; **12**: e1001779.
- 44 Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. *JAMA* 1999; **282**: 1737–1744.
- 45 Wain LV, Shrine N, Miller S, Jackson VE, Ntalla I, Artigas MS *et al*. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir Med* 2015; **3**: 769–781.
- 46 Hagenaaers SP, Harris SE, Davies G, Hill WD, Liewald DC, Ritchie SJ *et al*. Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N = 112 151) and 24 GWAS consortia. *Mol Psychiatry* 2016; **21**: 1624–1632.
- 47 Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007; **39**: 906–913.
- 48 Harris SE, Hagenaaers SP, Davies G, Hill WD, Liewald DC, Ritchie SJ *et al*. Molecular genetic contributions to self-rated health. *Int J Epidemiol* 2016.
- 49 Gale C, Hagenaaers SP, Davies G, Hill WD, Liewald DC, Cullen B *et al*. Pleiotropy between neuroticism and physical and mental health: findings from 108 038 men and women in UK Biobank. *Transl Psychiatry* 2016; **6**: e791.
- 50 Davies G, Marioni RE, Liewald DC, Hill WD, Hagenaaers SP, Harris SE *et al*. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N = 112 151). *Mol Psychiatry* 2016; **21**: 758–767.
- 51 Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR *et al*. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010; **42**: 565–569.
- 52 de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* 2015; **11**: e1004219.
- 53 Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh PR *et al*. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* 2015; **47**: 1228–1235.
- 54 Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF *et al*. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
- 55 Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N *et al*. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; **47**: 291–295.
- 56 Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR *et al*. An atlas of genetic correlations across human diseases and traits. *Nat Genet* 2015; **47**: 1236–1241.
- 57 Euesden J, Lewis CM, O'Reilly PF. PRSice: polygenic risk score software. *Bioinformatics* 2015; **31**: 1466–1468.
- 58 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R. Stat Soc B* 1995; **57**: 289–300.
- 59 Smith DJ, Nicholl BI, Cullen B, Martin D, Ul-Haq Z, Evans 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* 2013; **8**: e75362.
- 60 Qi Q, Menzaghi C, Smith S, Liang L, de Rekeneire N, Garcia ME *et al*. Genome-wide association analysis identifies TYW3/CRYZ and NDST4 loci associated with circulating resistin levels. *Hum Mol Genet* 2012; **21**: 4774–4780.
- 61 Sugimoto M, Watanabe T, Sugimoto Y. The molecular effects of a polymorphism in the 5'UTR of solute carrier family 44, member 5 that is associated with birth weight in holsteins. *PLoS ONE* 2012; **7**: e41267.

- 62 Clark GD. Platelet-activating factor acetylhydrolase and brain development. *Enzymes* 2015; **38**: 37–42.
- 63 Hoenicka J, aragüés M, Ponce G, Rodríguez-Jiménez R, Jiménez-Arriero MA, Palomo T. From dopaminergic genes to psychiatric disorders. *Neurotox Res* 2007; **11**: 61–71.
- 64 de Miguel FJ, Sharma RD, Pajares MJ, Montuenga LM, Rubio A, Pio R. Identification of alternative splicing events regulated by the oncogenic factor SRSF1 in lung cancer. *Cancer Res* 2014; **74**: 1105–1115.
- 65 Sailer A, Houlden H. Recent advances in the genetics of cerebellar ataxias. *Curr Neurol Neurosci Rep* 2012; **12**: 227–236.
- 66 Russell B, Graham JM. Expanding our knowledge of conditions associated with the ASXL gene family. *Genome Med* 2013; **5**: 1–3.
- 67 Hill WD, Davies G, Harris SE, Hagenaars SP, Liewald D, Penke L et al. Molecular genetic aetiology of general cognitive function is enriched in evolutionarily conserved regions. *Transl Psychiatry* 2016; **6**: e980.
- 68 Alberti KG, Zimmet P, Shaw J, IDF Epidemiology Task Force Consensus Group. The metabolic syndrome—a new worldwide definition. *Lancet* 2005; **366**: 1059–1062.
- 69 Booth T, Starr JM, Deary I. Modeling multisystem biological risk in later life: allostatic load in the Lothian birth cohort study 1936. *Am J Hum Biol* 2013; **25**: 538–543.
- 70 Seeman T, Gruenewald T, Karlamangla A, Sidney S, Liu K, McEwen B et al. Modeling multisystem biological risk in young adults: the Coronary Artery Risk Development in Young Adults Study. *Am J Hum Biol* 2010; **22**: 463–472.
- 71 Watson D, Pennebaker JW. Health complaints, stress, and distress: exploring the central role of negative affectivity. *Psychol Rev* 1989; **96**: 234–254.
- 72 Deary IJ. Looking for 'system tegrity' in cognitive epidemiology. *Gerontology* 2012; **58**: 545–553.
- 73 Wessely S. Chronic fatigue: symptom and syndrome. *Ann Int Med* 2001; **134**: 838–843.
- 74 Aristotle. *Physics*. Oxford University Press: USA, 1999.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2018

Supplementary Information accompanies the paper on the Molecular Psychiatry website (<http://www.nature.com/mp>)