nature genetics

ARTICLES

https://doi.org/10.1038/s41588-018-0320-8

Multivariate genome-wide analyses of the well-being spectrum

Bart M. L. Baselmans[1,2], Rick Jansen[3,4], Hill F. Ip[1], Jenny van Dongen[1,2], Abdel Abdellaoui[2,5], Margot P. van de Weijer[1], Yanchun Bao[6], Melissa Smart[6], Meena Kumari[6], Gonneke Willemsen[1,2,4], Jouke-Jan Hottenga[1,2,4], BIOS consortium[7], Social Science Genetic Association Consortium[7], Dorret I. Boomsma[1,2,4], Eco J. C. de Geus[1,2,4], Michel G. Nivard[1,2,8*] and Meike Bartels[1,2,4,8*]

We introduce two novel methods for multivariate genome-wide-association meta-analysis (GWAMA) of related traits that correct for sample overlap. A broad range of simulation scenarios supports the added value of our multivariate methods relative to univariate GWAMA. We applied the novel methods to life satisfaction, positive affect, neuroticism, and depressive symptoms, collectively referred to as the well-being spectrum (N_{obs} = 2,370,390), and found 304 significant independent signals. Our multivariate approaches resulted in a 26% increase in the number of independent signals relative to the four univariate GWAMAs and in an ~57% increase in the predictive power of polygenic risk scores. Supporting transcriptome- and methylome-wide analyses (TWAS and MWAS, respectively) uncovered an additional 17 and 75 independent loci, respectively. Bioinformatic analyses, based on gene expression in brain tissues and cells, showed that genes differentially expressed in the subiculum and GABAergic interneurons are enriched in their effect on the well-being spectrum.

In the past decade, genome-wide association studies (GWAS) have provided insight into the genetic basis of quantitative variation in complex traits[1]. With summary statistics of these GWAS becoming public and the development of linkage disequilibrium score regression (LDSC)[2,3], genetic correlations between traits can be systematically estimated (for example, the Brainstorm consortium[4]). Leveraging this widely observed genetic overlap between traits, we introduce two novel methods for multivariate genome-wide-association meta-analysis, in which we define a multivariate model as a model in which the effect of a single SNP is considered for multiple traits: (i) *N*-weighted multivariate GWAMA (N-GWAMA), with a unitary effect of the SNP on all traits, and (ii) model-averaging GWAMA (MA-GWAMA), in which we relaxed the assumption of a unitary effect of the SNP on all traits. Both methods are well equipped to deal with (unknown) sample overlap. The dependence between effect sizes (error correlation) induced by possible sample overlap is estimated from the univariate GWAMA by using LDSC[2,3]. Furthermore, the univariate LDSC intercept is used to correct for population stratification and cryptic relatedness. Both methods have advantages over existing methods. In contrast with MultiPhen[5], CCA (mv-PLINK)[6], Combined-PC[7], and mv-BIMBAM[8], both methods can be applied without a need for individual-level genotypic data, because only GWAS or GWAMA summary statistics are required. Additionally, in contrast with S_{Hom}[9], N-GWAMA and MA-GWAMA take a more precise estimate of the error correlation into account. In contrast with MTAG[10], MA-GWAMA, similarly to S_{Het}[9], generates trait-specific estimates for each SNP, allowing a certain degree of heterogeneity (Methods). Finally, in contrast with TATES[11], both N-GWAMA and MA-GWAMA generate effect sizes for the multivariate effect, whereas TATES only generates a *P* value. The absence of a signed statistic in TATES complicates or even prohibits polygenic prediction.

Results

Simulations. We performed simulations to elucidate in which scenarios N-GWAMA and MA-GWAMA outperform univariate GWAMA and when N-GWAMA outperforms MA-GWAMA, or the reverse. GWAMA summary statistics were simulated for a range of different scenarios. For each scenario, we simulated four heritable traits (h^2_{SNP} = 30%) and varied the genetic correlation between the four traits from 0.1 to 0.9 (Methods). We sampled 80,000 causal SNPs for 80,000 individuals from the UK Biobank[12]. For each trait, 40,000 individuals were simulated, and the sample overlap between the traits ranged from 0 to 25,000 individuals. Four univariate GWAMAs were performed on the generated data. We chose parameters that far exceeded the reported h^2_{SNP} for many complex traits, which allowed us to simulate at smaller sample sizes (N = 80,000), reducing the computational burden.

We found that in the presence of genetic correlations equal to or higher than 0.5, both N-GWAMA and MA-GWAMA outperformed univariate GWAMA (Supplementary Fig. 1 and Supplementary Table 1). The added value of multivariate analysis disappeared when traits showed lower genetic correlations (≤0.4). We performed another four simulation scenarios in which we varied the SNP heritability together with the sample size. We kept the product of the h^2_{SNP} and the sample size constant to consider a realistic SNP heritability, which implies that the expectation of the *z* statistic remains constant (Methods). We simulated data with h^2_{SNP} of 40% (N = 30,000), 20% (N = 60,000), and 10% (N = 120,000), and we kept the genetic correlations between traits constant (at r_g = 0.7). Finally,

[1]Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands. [2]Amsterdam Public Health Research Institute, Amsterdam UMC, Amsterdam, the Netherlands. [3]Department of Psychiatry, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands. [4]Amsterdam Neuroscience, Amsterdam UMC, Amsterdam, the Netherlands. [5]Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands. [6]Institute for Social and Economic Research, University of Essex, Colchester, UK. [7]A list of members and affiliations appears in the Supplementary Note. [8]These authors jointly supervised this work: Michel G. Nivard, Meike Bartels. *e-mail: m.g.nivard@vu.nl; m.bartels@vu.nl

we simulated four traits with a genetic correlation of 0.7 ($N = 40{,}000$), without sample overlap. In all scenarios, both multivariate methods outperformed univariate GWAMA (Supplementary Table 1). To validate MA-GWAMA, we simulated data for which the assumption of a unitary effect of the SNP on all traits was relaxed (Methods). We found that in the scenario in which a SNP had an effect on at least three out of four traits, N-GWAMA and MA-GWAMA performed equally well. However, when a SNP had an effect on two out of four traits or one out of four traits, MA-GWAMA outperformed N-GWAMA (Supplementary Table 2). Of note, in scenarios in which SNPs influence fewer than half of the traits under consideration, univariate methods such as GWAMA can outperform multivariate N-GWAMA and MA-GWAMA.

Application to the well-being spectrum. Following our longstanding research interest[13–16], we applied our methods to the following traits: life satisfaction, positive affect, neuroticism, and depressive symptoms. Although the high phenotypic and genetic correlations between these traits are strongly suggestive of a common underlying biology, most research is still characterized by separate analyses. Acknowledging this, we performed both N-GWAMA and MA-GWAMA ($N_{obs} = 2{,}370{,}390$, Supplementary Table 3) of these four traits to increase the power to identify associated genetic variants.

Our analyses leveraged published univariate GWAMAs of life satisfaction[15,17] ($N_{obs} = 80{,}852$; two studies), positive affect[12,15,17] ($N_{obs} = 410{,}603$; three studies), neuroticism[12,15,17,18] ($N_{obs} = 582{,}989$; six studies), and depressive symptoms[12,15,17,19,20] ($N_{obs} = 1{,}295{,}946$; ten studies). Overall, the mean genetic correlations between different measures of the same trait were higher (life satisfaction, $r_g = 0.68$; positive affect, $r_g = 0.9$; neuroticism, $r_g = 0.84$; depressive symptoms, $r_g = 0.89$) than the mean genetic correlation between measures of different traits ($r_g = 0.7$; Fig. 1). This finding justified our two-stage approach of first performing meta-analyzing the datasets measuring the same traits (life satisfaction, positive affect, neuroticism, and depressive symptoms) and then meta-analyzing the four resulting datasets into what we refer to as the well-being spectrum ($N_{obs} = 2{,}370{,}390$; Supplementary Fig. 2). For the purpose of the multivariate GWAMA, we reversed the estimated SNP effects on neuroticism and depressive symptoms to ensure a positive correlation with life satisfaction and positive affect. The dependence between effect sizes (error correlation) induced by sample overlap was estimated from the genome-wide summary statistics obtained from the univariate GWAMA analyses using LDSC[2,3] (Methods and Fig. 1). Knowledge of the error correlation between univariate meta-analyses allowed for dependent samples to be meta-analyzed, providing a gain in power while guarding against inflated type-1-error rates (Methods).

Multivariate GWAMA results. In our N-GWAMA, we identified 231 independent (250-kb-window linkage disequilibrium (LD) >0.1) loci associated with the well-being spectrum (Fig. 2a and Supplementary Table 4), whereas MA-GWAMA identified 148 (life satisfaction), 191 (positive affect), 263 (neuroticism), and 239 (depressive symptoms) loci (Fig. 2b–e), some of which overlapped, resulting in 289 independent signals (Supplementary Tables 4–8). The overlap in genome-wide-significant loci divided by the geometric mean of the number of loci discovered for the traits was highly consistent with the genetic correlation between the traits (Supplementary Table 9). Of these 289 independent MA-GWAMA signals, 181 were within a 50-kb window of the independent signals present in the N-GWAMA analysis (78.3%). Considering both multivariate methods, we found 304 independent genome-wide signals associated with the well-being spectrum. This is a 26% increase compared with the independent signals found in the univariate GWAMAs (life satisfaction, positive affect, neuroticism, and depressive symptoms; Supplementary Table 10 and Supplementary Fig. 3a–d). The low LD-score intercepts for all analyses confirmed that the inflation in test statistics was due to an increase in signal rather than population stratification or inaccurate accounting for sample overlap (Methods and Supplementary Table 11).

We performed a lookup for the genome-wide-significant loci reported in published studies of related traits. We identified 26 loci in proximity (<250 kb) to the 44 genome-wide-significant loci (59.1%) reported for major depressive disorder (MDD)[21]. Additionally, we identified 58 loci in proximity to the 79 loci identified by using an alternative multivariate method considering well-being-spectrum traits (73.4%)[10]. Using height as a negative control ($r_g \sim 0.05$ with our included traits), we identified 37 loci in proximity to the 697 loci associated with height (5.3%)[22].

Polygenic prediction. We compared the predictive power of polygenic scores constructed from univariate GWAMA against those from N-GWAMA and MA-GWAMA. Prediction of measures of life satisfaction, positive affect, neuroticism, and depressive symptoms was performed in samples of The Netherlands Twin Register ($_{mean}N > 8{,}100$) and Understanding Society ($_{mean}N > 8{,}846$)[17,23]. We evaluated the predictive power of each polygenic score by its incremental R^2 value, defined as the increase in R^2 of the regression, including the polygenic score as an independent variable together with a set of controls (age, age^2, sex, and ten principal components) over a regression omitting the polygenic score. Univariate GWAMA polygenic scores had an incremental R^2 value of 0.13% for life satisfaction, 0.49% for positive affect, 1.53% for neuroticism, and 1.22% for depressive symptoms. The corresponding N-GWAMA and MA-GWAMA had larger incremental R^2 (for life satisfaction, 0.94% and 0.92%, respectively; for positive affect, 1.10% and 1.06%, respectively; for neuroticism, 1.68% and 1.61%, respectively; and for depressive symptoms, 1.64% and 1.63%, respectively). On average, N-GWAMA improved prediction by 59%, and MA-GWAMA improved prediction by 55% (Supplementary Fig 4 and Supplementary Table 12).

Bioinformatics. Given the equal performance of both multivariate GWAMAs, and the assumption of TWAS, MWAS, and LDSC that the test statistics follow a standard normal distribution under the null hypothesis of no effect, we chose to perform the bioinformatics analyses with the N-GWAMA results.

Transcriptome- and methylome-wide analyses. Multivariate GWAMA aggregates the effect of a single SNP across multiple traits, informed by prior knowledge of the genetic correlation between these traits. Next, we aggregated the effect across multiple SNPs on the basis of prior knowledge that some of these SNPs influence the expression level of a gene transcript or the methylation level at a CpG site (methylation quantitative trait loci (mQTL)) measured in whole blood. These methods (known as TWAS and MWAS) enable identification of genes involved in complex traits[24–26]. In TWAS, we uncovered 97 transcript–trait associations (45 loci) significant at a Bonferroni-corrected significance level ($P < 5.2 \times 10^{-6}$). For 17 TWAS hits (14 loci), the corresponding locus (1,000 kb around the transcript) did not contain a significant N-GWAMA SNP. For 49 out of the 97 transcripts (30 loci), the maximum LD between the TWAS model SNPs and N-GWAMA top SNP in the corresponding locus was larger than 0.8 (Supplementary Table 13). Furthermore, we found 913 CpG methylation–trait associations mapping to 141 loci at a Bonferroni-corrected significance level. For 75 out of 913 CpG methylation–trait associations (36 loci), the corresponding locus did not contain an N-GWAMA significant signal. For 396 CpG methylation–trait associations (83 loci), the maximum LD between the MWAS model SNPs and an N-GWAMA top SNP was larger than 0.8 (Supplementary Table 14).

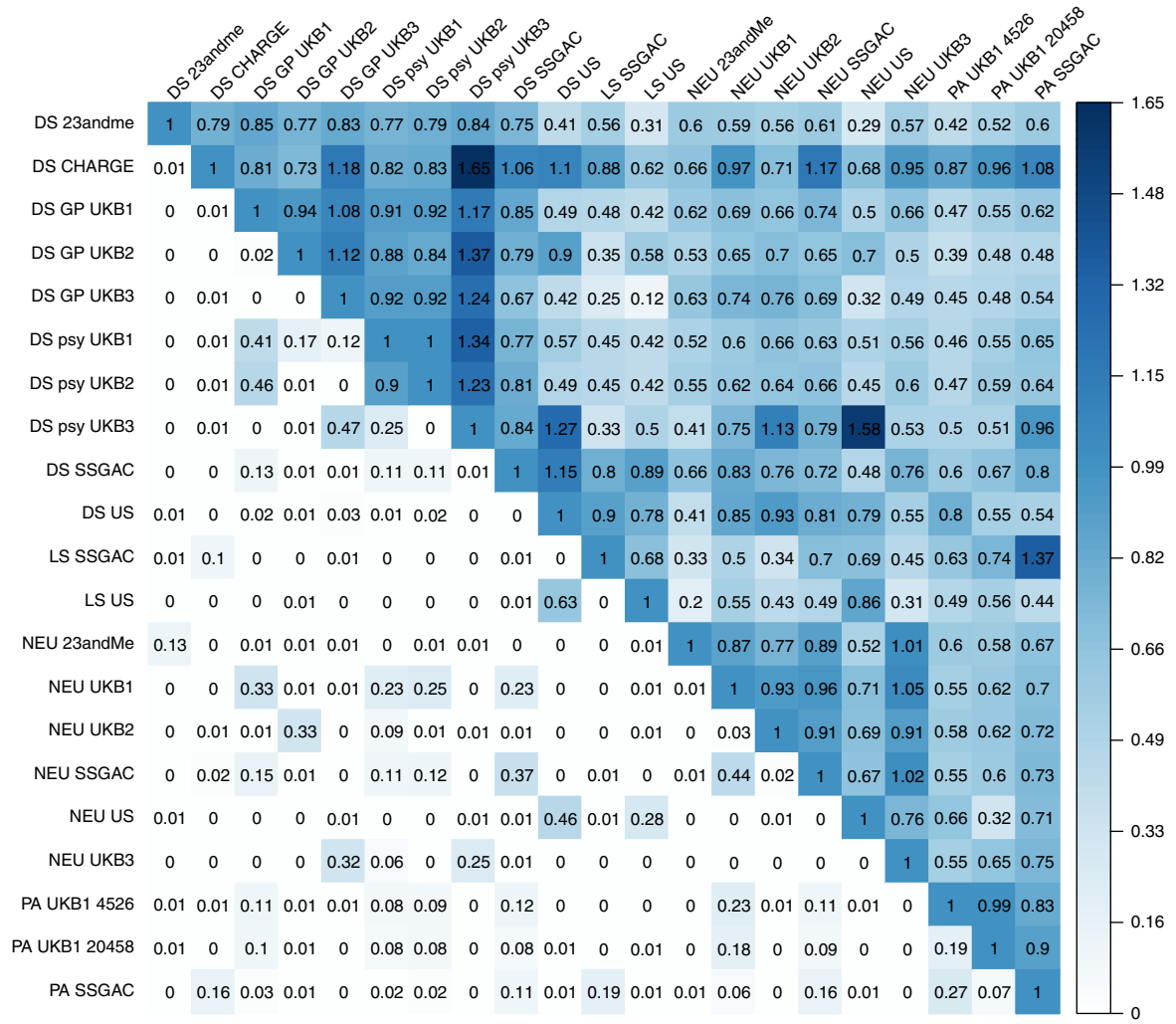

Fig. 1 | Genetic correlations and error correlations (cross-trait intercepts) between the included GWAMA data sets. Upper triangle, genetic correlations. Lower triangle, error correlation representing the magnitude of inflation due to population stratification. UKB1 represents Caucasian UK Biobank participants living in the UK; UKB2 represents Caucasian UK Biobank participants living in the UK that are relatives from UKB1; and UKB3 represents Caucasian UK Biobank participants not living in the UK. Sample sizes of the included traits and cohort descriptions can be found in Supplementary Table 3. Estimates of genetic correlations as estimated in LDSC are not bounded at 1. LS, life satisfaction; PA, positive affect; NEU, neuroticism; DS, depressive symptoms.

A locus of particular interest was found within the major histocompatibility complex (MHC). Recent work has identified three individual signals related to schizophrenia in the MHC region, one of which is linked to complement 4 (*C4A*) gene expression and synapse elimination during puberty[27]. The genome-wide-significant signal for the well-being spectrum in the MHC region was not in strong LD with lead expression quantitative trait loci (eQTLs) for *C4A* gene expression. Instead, a second independent signal tagged by rs13194504 was associated with both schizophrenia and well-being. TWAS results for the MHC region implicated the expression of *ZKSCAN4* in the etiology of well-being (Supplementary Fig. 5).

Stratified LD-score regression. We performed further biological annotation, using stratified LDSC[2,3]. Our first analysis aimed to confirm the involvement of the central nervous system in the etiology of the well-being spectrum. Our second analysis aimed to pinpoint specific locations in the brain. Our final analysis used single-cell sequencing data to identify specific cell-type involvement.

We considered the enrichment in the N-GWAMA-derived SNP set of 220 genomic annotations (33 brain and 187 nonbrain annotations), which reflected the locations of four specific histone marks (H3K4me1, H3K4me3, H3K27ac, or H3K9ac) in 54 tissues in their effect on the well-being spectrum[28]. This allowed for detection of, for example, enrichment of regions of the genome that are histone modified in the prefrontal cortex. Such enrichment would suggest involvement of processes in the prefrontal cortex in the etiology of the well-being spectrum. Our analyses identified significant enrichment of 69 annotations characterized by 32 histone marks in ten brain tissues (Supplementary Table 15 and Supplementary Figure 6). Of note, the top 15 significant annotations involved brain tissues. Among these brain tissues are the midfrontal and inferior-temporal lobes, fetal brain, cingulate and angular gyri, germinal matrix, hippocampus anterior caudate, substantia nigra, and the neurosphere.

To more accurately pinpoint brain regions where genes relevant to the well-being spectrum are differentially expressed, we computed stratified LD scores based on differential gene expression in an anatomically comprehensive set of 210 brain regions, based on 3,707 measurements in six human brains[29]. For each brain region, genes were selected that showed higher expression than that in all other regions (global differential gene expression). The LD scores were significantly enriched at false discovery rate (FDR) <0.05 at multiple gyri in the cortex (Supplementary Table 16). Differential

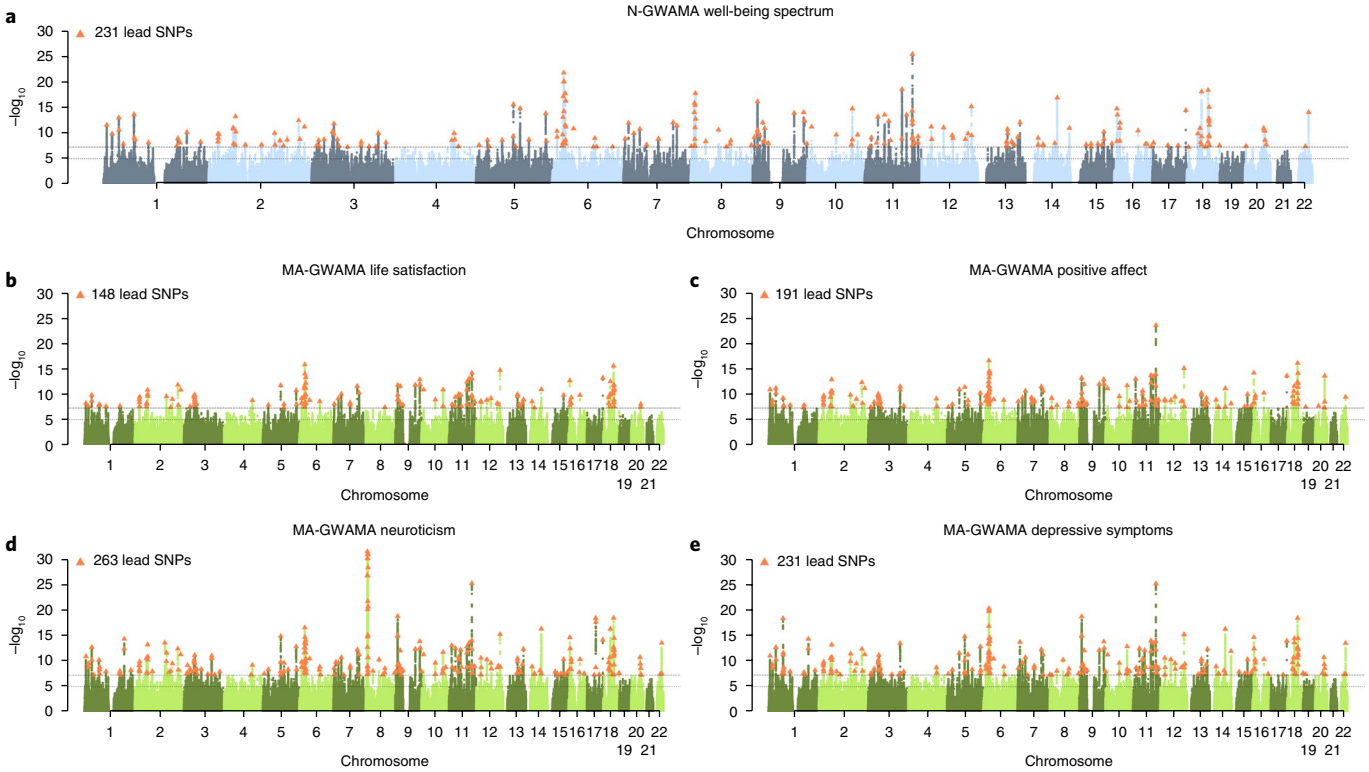

Fig. 2 | Manhattan plots of *N*-weighted and model-averaging GWAMA. a, N-GWAMA of the well-being spectrum. **b**–**e**, MA-GWAMA of life satisfaction (**b**), positive affect (**c**), neuroticism (**d**), and depressive symptoms (**e**). All plots in all panels are based on the same set of SNPs. The *x* axis represents the chromosomal position, and the *y* axis represents the significance on a $-\log_{10}$ scale. Sample sizes of the included traits are displayed in Supplementary Table 3. Each approximately independent genome-wide-significant association ('lead SNP') is marked by a triangle ($P < 5 \times 10^{-8}$).

gene expression appeared to be driven mainly by transcriptional differences between gross anatomical structures in the brain (cortex, subcortical structures, brainstem, and cerebellum). To identify regions related to the well-being spectrum within these structures, we divided the 210 regions into four sets (brainstem, cortex, subcortex, and cerebellum) according to their locations and computed differential gene expression across the regions within each structure (local differential gene expression).

Our results showed a significant (Bonferroni corrected) enrichment of the N-GWAMA signal for genes specifically expressed in the subiculum ($Z = 3.47$, $P < 0.001$; Fig. 3a–c and Supplementary Tables 17–20). The subiculum is considered part of the hippocampal formation and plays a key role in hippocampal–cortical interaction[30], inhibition of the hypothalamic–pituitary–adrenal axis, and the human response to stress[31]. We repeated the analyses using GWAMA summary statistics of educational attainment (EA)[32] and schizophrenia[33], two traits that are relatively weakly genetically correlated with the well-being spectrum ($r_g = -0.15$, $P = 1.81 \times 10^{-10}$, and $r_g = 0.34$, $P = 2.54 \times 10^{-64}$) but for which the brain has been implicated in their etiologies. In particular, we wanted to determine whether the signal observed in the subiculum was specific to the well-being spectrum. As a negative control, we considered the enrichment of genes differentially expressed in all brain regions on height[22]. We found no enrichment of genes differentially expressed in the subiculum on educational attainment ($Z = 1.251$; $P = 0.105$) but found an effect on schizophrenia ($Z = 2.938$; $P = 0.002$). No region was significantly enriched in the effect on height ($P > 0.05$). All results of the differential gene expression analysis were mapped to the MNI coordinates at which the tissue samples were obtained, thus allowing for future integration of our findings with other neuroimaging modalities (Fig. 3a–c and Supplementary Table 21).

Finally, we obtained the publicly available matrix of gene counts generated on the basis of single nuclei ($N = 14{,}963$) from the prefrontal cortex and hippocampus of multiple human donors by Habib et al.[34]. We divided these nuclei into seven types of neurons, two subtypes of astrocytes, oligodendrocytes, oligodendrocyte precursors cells, microglia, endothelial cells, and unclassified cells (hippocampus and prefrontal cortex), and computed cell-type-specific genes for the different types of neurons (Methods). Using LDSC, we tested the enrichment of all cell types in the N-GWAMA. Significant enrichment was observed for GABAergic interneurons sampled from the hippocampus and prefrontal cortex (GABA1; $Z = 3.42$; $P = 3.64 \times 10^{-6}$ and GABA2 $Z = 3.7$; $P = 6.54 \times 10^{-7}$; Fig. 3d and Supplementary Table 22).

Discussion

We introduced N-GWAMA and MA-GWAMA, two novel methods for conducting multivariate analysis of GWAMA summary statistics of related traits. Whereas previous univariate analyses of traits in the well-being spectrum have been moderately successful, we gained power by using multivariate analyses. N-GWAMA and MA-GWAMA identified 304 loci associated with some but not all traits in the well-being spectrum, and provided flexibility in terms of model specification. Of note, model averaging can be extended to incorporate other multivariate models such as MTAG[10] or models specified in Genomic SEM[35]. Model averaging can in fact incorporate any multivariate GWAMA or GWAS model for which the per-SNP model fit can be expressed in terms of an AICc fit statistic. N-GWAMA and MA-GWAMA are complementary to each other and can thus be used together to identify genetic variants associated with clusters of genetically correlated traits. We illustrate the power gain in multivariate GWAMA over univariate GWAMA for traits

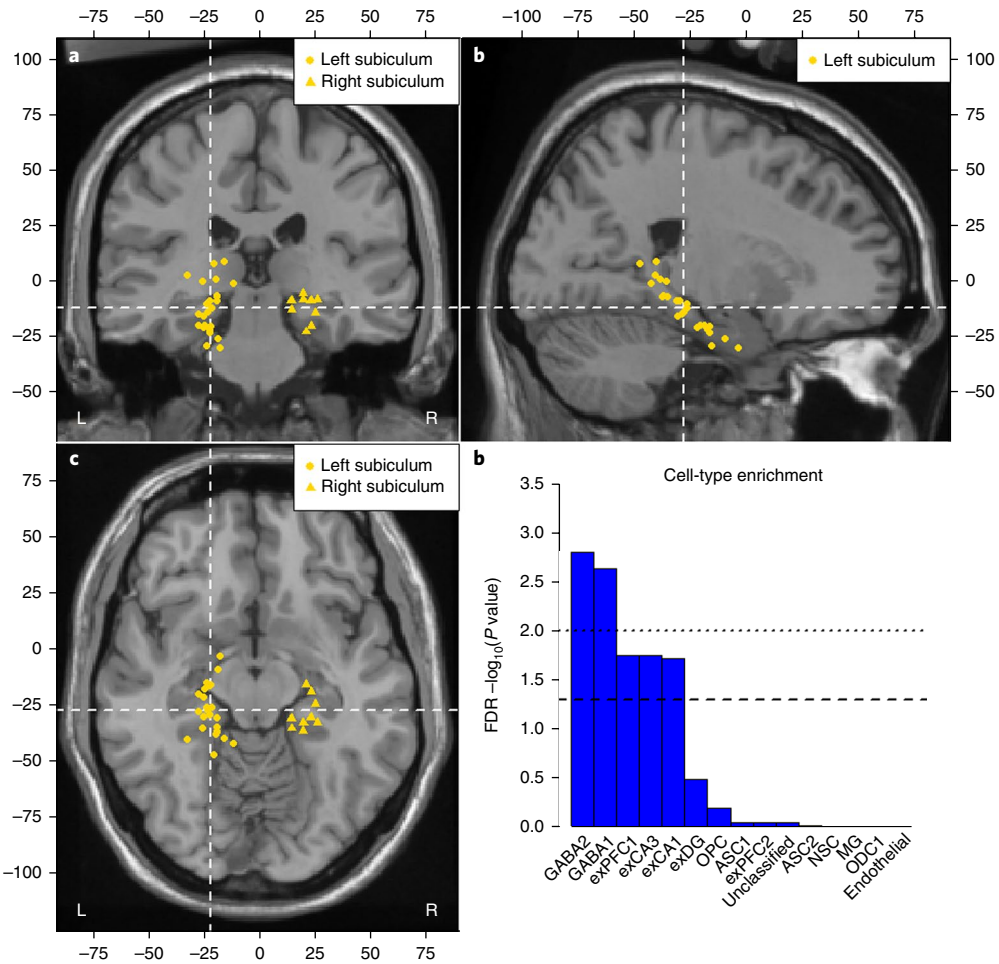

Fig. 3 | Local differential gene expression between subcortical structures and enrichment of individual cell-type enrichment. Local differential gene expression between subcortical structures identifies enrichment of genes specifically expressed in the subiculum ($Z = 3.47$, $P < 0.001$) in their effect on the well-being spectrum. **a**–**c**, Brain tissues: coronial view (**a**), sagittal view (**b**), and axial view (**c**). The locations of the samples of brain tissues that were used to measure gene expression by Hawrylycz et al.[29] are projected to a standard MNI template brain ('Colin27'). The figure is centered on the averaged MNI coordinates of brain samples, which are part of the annotation 'left subiculum' ($x = 77$, $y = 90$, and $z = 60$). **d**, Bar graph representing the cell-type-specific enrichment of GABAergic neurons (GABA1; $Z = 3.42$; $P = 3.64 \times 10^{-6}$ and GABA2 $Z = 3.7$; $P = 6.54 \times 10^{-7}$; one-sided test). Dotted line indicates significant at the FDR < 0.01 level; dashed line indicates significance at the FDR < 0.05 level. GABA, GABAergic interneurons; exPFC, glutamatergic neurons from the prefrontal cortex; exCA1/3, pyramidal neurons from the hippocampus CA region; exDG, granule neurons from the hip denate gyrus; OPC, oligodendrocyte precursors cells; ASC, astrocytes; NSC, neural stem cells; MG, microglia; ODC1, oligodendrocytes. N-GWAMA results were used for all analyses ($N_{obs} = 2{,}370{,}390$).

genetically correlated above 0.4, by using simulations. Empirical application showed that polygenic scores based on multivariate GWAMA outperformed polygenic scores based on univariate GWAMA. Aside from the advantages of the novel methods, it should be noted that both techniques estimate the error covariance between test statistics from LDSC and therefore inherit the assumptions from LDSC. This implies that all limitations that apply to LDSC also apply to our methods. For instance, a previous study has found, through simulations, that the univariate LDSC intercept can be biased at large sample sizes coupled with a high (SNP) heritability[36].

We used TWAS and MWAS to identify additional loci related to variation in complex traits, such as well-being, by aggregating the effects across multiple SNPs on the basis of prior knowledge that some of these SNPs influence the expression level of a gene transcript or the methylation level at a CpG site. By leveraging the N-GWAMA results, LD-score regression, and an atlas of brain gene expression, we were able to pinpoint brain regions where region-specific gene expression exists for genes enriched in their effect on well-being. We report evidence for enrichment of genes differentially expressed in the subiculum. Furthermore, we found enrichment for GABAergic interneurons. In the regions for which cell types were available (hippocampus and prefrontal cortex), we found cell-type-specific enrichment for the well-being spectrum. However, it stands to reason that the same cell-type-specific enrichment in other regions might have existed but may have been missed. Gene expression is known to vary systematically among cell types within the brain[37] (for example, neurons, microglia, astrocytes) and developmental phases[38] (prenatally, childhood, adulthood, and old age), and probably even among subtypes of a single cell type. Differences in gene expression across or within cell types may induce differences between regions, because cell-type composition might differ between regions. This limitation must be addressed in future well-being research, capitalizing on ongoing efforts to categorize gene expression across the human brain at increased (single-cell) resolution. Single-cell sequencing (for example, drop-seq-based anatomically comprehensive survey of the brain), based on donors deceased at different ages, could disentangle cell-type-specific from region-specific differential gene expression as well as age-specific

gene expression[39]. The results of our new multivariate GWAMA methods can be meaningfully mapped to brain regions on the basis of a coordinate system used within multiple other neuroscientific disciplines, thus facilitating future integration of genetic and neuroscientific research on the well-being spectrum.

URLs. N-GWAMA and MA-GWAMA software, https://github.com/baselmans/multivariate_GWAMA/; TWAS and MWAS software, http://bbmri.researchlumc.nl/atlas/#data/; BIOS consortium, http://wiki.bbmri.nl/wiki/BIOS_start-/; TWAS and MWAS prediction models, http://bbmri.researchlumc.nl/atlas/#data/; Brain Map gene expression brain regions, http://www.brain-map.org/.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41588-018-0320-8.

Received: 12 July 2018; Accepted: 27 November 2018;
Published online: 14 January 2019

References

1. Visscher, P. M. et al. 10 Years of GWAS discovery: biology, function, and translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
2. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
3. Bulik-Sullivan, B. K. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
4. Anttila, V. et al. Analysis of shared heritability in common disorders of the brain. *Science* **360**, 6395 (2018).
5. O'Reilly, P. F. et al. MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. *PLoS One* **7**, e34861 (2012).
6. Ferreira, M. A. R. & Purcell, S. M. A. multivariate test of association. *Bioinformatics* **25**, 132–133 (2009).
7. Aschard, H. et al. Maximizing the power of principal-component analysis of correlated phenotypes in genome-wide association studies. *Am. J. Hum. Genet.* **94**, 662–676 (2014).
8. Stephens, M. Unified framework for association analysis with multiple related phenotypes. *PLoS One* **8**, e65245 (2013).
9. Zhu, X. et al. Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. *Am. J. Hum. Genet.* **96**, 21–36 (2015).
10. Turley, P. et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat. Genet.* **50**, 229–237 (2018).
11. van der Sluis, S., Posthuma, D. & Dolan, C. V. TATES: efficient multivariate genotype-phenotype analysis for genome-wide association studies. *PLoS Genet.* **9**, e1003235 (2013).
12. Bycroft, C. et al. Genome-wide genetic data on ~500,000 UK Biobank participants. Preprint at https://doi.org/10.1101/166298 (2017).
13. Bartels, M. & Boomsma, D. I. Born to be happy? The etiology of subjective well-being. *Behav. Genet.* **39**, 605–615 (2009).
14. Rietveld, C. A. et al. Molecular genetics and subjective well-being. *Proc. Natl Acad. Sci. USA* **110**, 9692–9697 (2013).
15. Okbay, A. et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* **48**, 624–633 (2016).
16. Baselmans, B. M. L. & Bartels, M. A genetic perspective on the relationship between eudaimonic -and hedonic well-being. *Sci. Rep.* **8**, 1–10 (2018).
17. Brice, J., Buck, N. & Prentice-Lane, E. (eds.). *British Household Panel Survey User Manual* (University of Essex, Colchester, 1993).
18. Lo, M.-T. et al. Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. *Nat. Genet.* **49**, 152–156 (2017).
19. Hyde, C. L. et al. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat. Genet.* **48**, 1031–1036 (2016).
20. Hek, K. et al. A genome-wide association study of depressive symptoms. *Biol. Psychiatry* **73**, 667–678 (2013).
21. Wray, N. R. et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668 (2018).
22. Wood, A. R. et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–1186 (2014).
23. Willemsen, G. et al. The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. *Twin. Res. Hum. Genet.* **16**, 271–281 (2013).
24. Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
25. Qi, T. et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nat. Commun.* **9**, 2282 (2018).
26. Gusev, A. et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
27. Sekar, A. et al. Schizophrenia risk from complex variation of complement component 4. *Nature* **530**, 177–183 (2016).
28. Finucane, H. K. et al. Partitioning heritability by functional annotation using GWAS summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
29. Hawrylycz, M. J. et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
30. O'Mara, S. M., Commins, S., Anderson, M. & Gigg, J. The subiculum: a review of form, physiology and function. *Prog. Neurobiol.* **64**, 129–155 (2001).
31. Herman, J. P. & Mueller, N. K. Role of the ventral subiculum in stress integration. *Behav. Brain Res.* **174**, 215–224 (2006).
32. Okbay, A. et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–542 (2016).
33. Ripke, S. et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
34. Habib, N. et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).
35. Grotzinger, A. D. et al. Genomic SEM provides insights into the multivariate genetic architecture of complex traits. Preprint at https://doi.org/10.1101/305029 (2018).
36. Vlaming, R. de, Johannesson, M., Magnusson, P. K. E., Ikram, M. A. & Visscher, P. M. Equivalence of LD-score regression and individual-level-data methods. Preprint at https://doi.org/10.1101/211821 (2017).
37. Darmanis, S. et al. A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl Acad. Sci. USA* **112**, 7285–7290 (2015).
38. Kang, H. J. et al. Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).
39. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).

Acknowledgements

We thank all participants in the cohort studies. This work was supported by the Netherlands Organization for Scientific Research (NWO: MagW/ZonMW grants 904-61-090, 985-10-002,904-61-193,480-04-004, 400-05-717, NWO-bilateral agreement 463-06-001, NWO-VENI 451-04-034, Addiction-31160008, Middelgroot-911-09-032, Spinozapremie 56-464-14192), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007), the VU University's Institute for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam (NCA), the European Science Council (ERC Advanced, 230374), the Avera Institute for Human Genetics, Sioux Falls, South Dakota (USA) and the National Institutes of Health (NIH, R01D0042157-01A). Part of the genotyping was funded by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health (NIMH, MH081802) and by the Grand Opportunity grants 1RC2MH089951-01 and 1RC2 MH089995-01 from the NIMH. Part of the analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org/), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003), the Dutch Brain Foundation, and the department of Behavioural and Movement Sciences of the VU University Amsterdam. M.B. is/was financially supported by a senior fellowship of the (EMGO+) Institute for Health and Care and a VU University Research Chair position. This work is supported by an ERC consolidator grant (WELL-BEING 771057 PI Bartels). M.G.N. is supported by a ZonMw grant: 'Genetics as a research tool: A natural experiment to elucidate the causal effects of social mobility on health' (pnr: 531003014), ZonMw project: 'Can sex- and gender-specific gene expression and epigenetics explain sex-differences in disease prevalence and etiology?' (pnr:849200011) and grant R01AG054628 02S. Understanding Society is an initiative funded by the Economic and Social Research Council (ES/H029745/1) and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The BIOS and SSGAC consortia are acknowledged as banner-coauthors for the key role their previous work played. A detailed description of their role and membership appears in the Supplementary Note.

Author Contributions

M.B., M.G.N., and B.M.L.B. oversaw the study. The theory underlying N-GWAMA and MA-GWAMA was developed by M.G.N., with contributions from B.M.L.B and M.B.

Simulations were performed by B.M.LB. and M.G.N. The N-GWAMA and MA-GWAMA software was developed by B.M.L.B., H.F.I., and M.G.N. Data analyses were conducted by B.M.L.B., R.J., H.F.I., J.v.D., A.A., M.P.v.d.W., Y.B., and M.G.N. Data curation was done by R.J., Y.B., MS., M.K., G.W., J.-J.H., E.J.C.d.G., D.I.B., and M.B. The manuscript was written by B.M.L.B., M.G.N., and M.B., with helpful contributions from E.J.C.d.G. All authors provided input and revisions for the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41588-018-0320-8.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to M.G.N. or M.B.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

***N*-weighted GWAMA.** We obtained summary statistics from univariate GWAMAs of life satisfaction, positive affect, neuroticism, and depressive symptoms[12,15,17–20]. We used the tool DIST[40] to impute HapMap2 summary statistics to the 1000Genomes Phase1 reference. To quantify the dependence between the univariate GWAMAs, we estimated the cross-trait LD-score intercept (CTI)[2,3]:

$$CTI = \frac{N_s * r_p}{\sqrt{N_1 N_2}}$$

where N_s is the sample overlap, N_1 is the sample size for trait 1, N_2 is the sample size for trait 2, and r_p is the phenotypic correlation between trait one and two. The CTI is approximately equal to the covariance between the test statistics obtained in univariate GWAMAs of trait 1 and 2. We assume that the estimated CTI is equal to the true CTI, though the uncertainty in the estimated CTI is generally low. Given the estimated covariance between effect sizes, we can meta-analyze the four dependent GWAMAs and obtain a multivariate test statistic per SNP *k*:

$$Z_k = \frac{\sum_{i=1}^{4} (w_{ik} * Z_{ik})}{\sqrt{\sum_{i=1}^{4} (w_{ik} * V_i) + \sum_{i=1}^{4} \sum_{j=1}^{4} (\sqrt{w_{ik} * w_{jk}} * C_{i,j})\ (j \neq i)}}$$

where w_{ik} is the square root of the sample size times the heritability for trait i, Z_{ik} is the test statistic of SNP *k* in the GWAMA of trait *i*, V_{ik} is the variance of the test statistic for SNP *k* in the GWAMA of trait *i* (that is 1 given that *Z* is a standardized test statistic), and $C_{i,j,k}$ is the covariance between (standardized) test statistics for SNP *k* between GWAMA of trait *i* and trait *j* (where *C* equals CTI obtained from cross-trait LD-score regression between trait *i* and trait *j*, and *V* is the univariate LD-score intercept). Under the null hypothesis (no heritability), the test statistics have unit variance, and the covariance $C_{i,j,}$ is equal to a correlation. The multivariate test statistic Z_k, is a weighted sum of test statistics, all of which follow a normal distribution under their respective null distributions. The statistic Z_k follows a standard normal distribution under the null hypothesis of no effect.

Model averaging GWAMA. Consider the following model:

$$\beta = MVN(\gamma X + e, V)$$

where β ($1 \times n$) is a multivariate normal vector of effect sizes obtained from the regression of *n* standardized traits on a standardized genotype (SNP). The matrix V (nxn) is the variance–covariance matrix of effect sizes, matrix X a design matrix (pxn), and γ the corresponding vector of parameters ($1 \times p$). The indexed *p* denotes the number of variables included in the means model of the response vector β, and *e* is the error term.

In this context, a regular GWAMA restricts the design matrix X to a unit vector (that is, we model a single genetic effect that is assumed identical across cohorts, and any observed variation is attributed to sample fluctuation). Generally, matrix V is diagonal and contains the squared standard errors of elements in β. A regular GWAMA is the most restricted model one can consider. However, when considering multivariate GWAMA (in which the elements in β reflect SNP effects on separate yet correlated traits), this model might be too restrictive. Even when traits have a substantial genetic correlation, not all genetic effects need to be shared between traits or be identical in magnitude. The least restrictive model is to consider the SNP effects in β independently (that is, run univariate GWAMA of the correlated traits). In between the most restrictive and least restrictive model, a manifold of models can be specified, equating the effects in γ across combinations of traits while allowing it to differ between other combinations of traits. These models can be specified by ways of the design matrix X.

One could consider a manifold (z) of models (m), each with a different design matrix X.

$$\beta_1 = MVN(\gamma_1 X_1 + e, V)$$
$$\beta_2 = MVN(\gamma_2 X_2 + e, V)$$
$$\beta_z = MVN(\gamma_z X_z + e, V)$$

When considering *i* correlated traits, a simple expansion of X is to allow for two vectors ($p = 2$), a unit vector and a second vector that is coded dichotomously (0,1), where the coding varies over each of the *z* models. Other coding, based on analysis of the genetic correlation between traits (that is, PCA or Cholesky decomposition), can be applied to summary statistics and included in the average. Practically, this approach allows for the existence of two distinct genetic effects. This procedure results in $0.5 \times p^2$ models. The 1–degree of freedom model with a unit vector for X and $(0.5 \times p^2 - 1)$ 2–degrees of freedom models with a unit vector and a second vector that codes for all possible combinations of pairs of *k* traits. However, simply considering *m* models for all SNPs across the genome results in a prohibitive increase of the already substantial multiple testing burden. Given *z* possible models, each of which predicts a different vector γ, and uncertainty for the predicted elements in γ, a possible way forward is to average the model predictions. The models are weighted by the relative proportion of evidence for each model. Specifically, the weights can be based on the AICc[41] information criteria. The AICc for model *m* equals:

$$AICc_m = -\ln(logLik_m) + 2k_m + \frac{2p_m(p_m + 1)}{z - p_m - 1}$$

For each AICc, we compute the delta (Δ_m) to the best (that is, lowest) AICc value, and from these values, we compute the model weights (g) for the *p* models as:

$$g_m = \frac{\exp\left(-\frac{1}{2}\Delta_m\right)}{\sum_{m=1}^{z} \exp\left(-\frac{1}{2}\Delta_m\right)}$$

We predict the vector β using each of the models

$$\hat{\beta}_m = \gamma_m X_m$$

One can aggregate the prediction over all models as:

$$\beta_a = \sum_{m=1}^{z} \frac{\hat{\beta}_m * g_m}{\sum_{m=1}^{z} g_m}$$

And we aggregate the uncertainty within and between models to obtain $var(\beta_a)$:

$$var(\beta_a) = \left[\sum_{m=1}^{z} g_m \sqrt{\mathrm{var}(\hat{\beta}_m) + (\hat{\beta}_m - \hat{\bar{\beta}})^2}\right]^2$$

The resulting vector β_a contains the model averaged effect sizes for the effect of a particular SNP on the traits subjected to multivariate analysis. Of note, the variance estimate contains a variance component that reflects within-model variability ($\mathrm{var}(\hat{\beta}_m)$), which equals the square of the standard error, and a variance component reflecting between-model variability ($(\hat{\beta}_m - \hat{\bar{\beta}})^2$) in estimate, which ensures that no overfitting occurs.

Our procedure boosts power if the SNP effect is concordant between traits, while retaining strongly discordant SNP effects if the model favors these. Model averaging offers several avenues for extension. One can constrain the SNP effects across multiple SNPs on the basis of biological knowledge of the relation between the SNPs and gene expression, or CpG methylation (analog to TWAS). Alternatively, it might be beneficial to average the AICc weights across regions of the genome. Model averaging can in principle accommodate any model for which the AICc information criterion can be expressed. These models should result in a vector of SNP effects (β) and an asymptotic variance for the SNP effects. In the current application, models per SNP are estimated in R by using the 'metafor' package, and models are averaged by using the 'AICcmodavg' package[42,43].

Simulations. We performed simulations to elucidate in which scenarios N-GWAMA and MA-GWAMA outperform univariate GWAMA, and when N-GWAMA outperforms MA-GWAMA, or the reverse. For each scenario, we simulated four heritable traits ($h_{SNP}^2 = 30\%$) affected by 80,000 SNPs. The genetic correlation among the four traits varied between 0.1 and 0.9. Using real genotypes and simulated effects, we sampled 80,000 causal SNPs for 80,000 individuals from the UK Biobank[12]. For each trait, 40,000 individuals were simulated, in which the sample overlap between the traits ranged from 0 to 25,000 individuals, to conduct univariate GWAMA, including 656,284 genotyped SNPs (minor allele frequency (MAF) >0.01). This method introduced partial sample overlap between the univariate GWAMAs. Next, we performed N-GWAMA and MA-GWAMA analyses and correlated the true SNP effects with the estimated SNP effects obtained from the univariate GWAMA, N-GWAMA, and MA-GWAMA.

We performed another four simulation scenarios in which we varied the SNP heritability together with the sample size. We kept the product of the h^2_{SNP} and the sample size constant to test whether at lower h^2_{SNP} and higher sample size (more in line with our empirical application) our findings concerning the relative power of univariate GWAMA, N-GWAMA, and MA-GWAMA would persist. The constant product of N X h^2_{SNP} implies that the expectation of the squared *z* statistic remains constant according to the expectation of the square *Z* statistics as formulated in Bulik-Sullivan et al.[2]:

$$E[\chi^2 | \ell_j] = \frac{Nh^2}{M}\ell_j + Na + 1$$

To validate MA-GWAMA, we simulated data for which the assumption of a unitary effect of the SNP on all traits was relaxed. We again simulated four traits that were affected by 80,000 SNPs. The SNP effects were perfectly correlated; however, we replaced true effects with zero in a way that guaranteed that 10,000 SNPs had a true effect on only one trait, 10,000 SNPs had a true effect on two

traits, and 10,000 SNPs had a true effect on three traits. On the basis of these effect sizes and genotypes, we simulated traits for 100,000 individuals and performed univariate GWAMA, N-GWAMA, and MA-GWAMA analyses as described above.

Polygenic prediction. To confirm the gain in power of our multivariate approaches, we performed polygenic score prediction (PRS) in two independent samples: (i) the Netherlands Twin Register (NTR)[23,44] and (ii) Understanding Society (UKHLS)[17]. We predicted the traits in the well-being spectrum (life satisfaction, positive affect, neuroticism, and depressive symptoms). In NTR, life satisfaction and positive-affect data were available in 9,143 and 6,836 genotyped participants, respectively. Life satisfaction was measured longitudinally by using the five-item Satisfaction with Life Scale[45]. Positive affect was also measured longitudinally by using four questions that were adapted from the Subjective Happiness Scale[46]. Neuroticism data were available for 8,527 genotyped participants. The Big Five personality traits (including neuroticism) were measured by using the NEO-FFI[47], a personality questionnaire consisting of five subscales: neuroticism, extraversion, openness, agreeableness, and conscientiousness. Depressive symptoms were obtained from the DSM-oriented Depression subscale of the age-appropriate survey from the ASEBA taxonomy[48] and were available for 7,898 participants.

UKHLS data were available for 9,944 participants. Life satisfaction was measured longitudinally (waves 1–6). Participants were asked how satisfied they were 'with life overall' and gave responses on a seven-point scale. Positive affect was also measured longitudinally (waves 1 and 4 only) by using The Warwick–Edinburgh Mental well-being scale (SWEMWBS)[49], a shortened version of WEMWBS. Neuroticism data were available for 8,198 genotyped participants from wave 3. The Big Five personality traits (including neuroticism) were measured by using The Big Five Inventory (BFI), a personality questionnaire consisting of five subscales: neuroticism, extraversion, openness, agreeableness, and conscientiousness. Depressive symptoms were measured longitudinally (waves 1–6) and obtained from The General Health Questionnaire (GHQ), available for 9,203 participants.

The weights used for the polygenic scores were based on the four univariate GWAMAs as well as our two versions of multivariate GWAMAs. Scores were based on the intersection of SNPs available in any of these GWAMAs and the prediction sample. Both in NTR and in UKHLS, SNPs were imputed to 1000 Genomes project March 2012 version 3 (ref. [50]). In NTR and UKHLS, 1,224,793 and 955,441 SNPs passed QC, respectively, and were used to construct polygenic scores. The traits were regressed on sex and age, as well as principal components included to correct for ancestry and the polygenic scores. Results can be found in Supplementary Table 12.

Summary-based transcriptome-wide association studies. Gene expression exhibits strong allelic heterogeneity[51], in which multiple SNPs local to the gene jointly influence gene expression levels. We aggregated SNP effects informed by their common effect on expression level of a gene (TWAS) or CpG methylation (MWAS), as proposed by Gusev et al.[26]. For TWAS, we used the RNA sequencing (RNA-seq) data from the BIOS Consortium[52]. The BIOS Consortium provides a data infrastructure hosting genetic (imputed SNPs), methylome (Illumina 450K array), transcriptome (RNA-seq), and phenotypic data on ~4,000 individuals from six Dutch biobanks, and a catalog with research output (see URLs). We used 3,344 whole-blood RNA-seq samples, measured with Illumina's Hiseq2000 (paired-end sequencing of 2×50 bp, >15 M read pairs per sample). Batch effects and the first 50 principal components without a GWAS hit were removed from the RNA-seq data, which was quantile-normal normalized for each gene. The corresponding genotype data that we used consisted of 881,977 unambiguous HapMap SNPs (MAF>5%, minor allele count>10, imputation info score>0.8). For eQTL analysis, linear regression on each SNP–gene pair closer than 250 kb was performed. FDR was based on ten permutations. Subject labels were permuted, and eQTL analysis was repeated. Subsequently, top associations per gene were counted and compared between permuted and observed data[53]. At an FDR of 5% ($P<1\times10^{-5}$) there were 13,870 genes with a significant eQTL. For each gene with a significant eQTL, a lasso model was fit in R with the function glmnet, with all SNPs closer than 250 kb to the gene as predictors and gene expression as outcome. For each gene, Lasso reduces the predictors to an optimal amount and provides the prediction model of gene expression (E) based on local SNPs S_1–S_N and lasso weights $q=q_1,q_2,\ldots,q_n$,:

$$E=\sum_{i=1}^{n} q_i S_i$$

On the basis of the prediction models, N-GWAMA summary statistics and LD based on the 1000 Genomes reference, TWAS was performed. That is, for each gene-prediction model containing SNPs S_1–S_N with weights $q=q_1,q_2,\ldots,q_n$, the corresponding GWAMA z scores $z=z_1,z_2,\ldots,z_n$ and LD, an n-by-n correlation matrix for eQTLs S_1–S_N, were used to construct the TWAS test statistic for each gene:

$$Z_{twas}=\frac{\sum_{i=1}^{n} q_i z_i}{\sqrt{q*LD*q}}$$

TWAS prediction models and LD matrices can be downloaded and can be used to test for association between the estimated cis (DNA) component of gene expression and the phenotype used for GWAS (see URLs).

Summary-based methylome-wide association studies. By applying the same approach for TWAS, MWAS was performed by using prediction models for each DNA-methylation CpG, with local SNPs as predictors. We used whole-blood methylation data from the BIOS Consortium[54]: 4,008 samples measured with Illumina 450K arrays. Methylation preprocessing was the same as that for RNA. Genotype and mQTL analysis procedures were the same as for eQTL analysis. At an FDR of 5% ($P<9.3\times10^{-5}$) there were 151,729 CpGs with a significant mQTL. For each CpG with a significant mQTL, we made a prediction model of methylation based on local SNPs (which is a weighted linear combination of SNPs). On the basis of the prediction models, MWAS statistics were generated (formula for TWAS). MWAS prediction models and corresponding LD matrices can be downloaded (see URLs).

Stratified LD-score regression. To determine whether specific genomic regions are enriched for genetic effects on the well-being spectrum, we used LD-score regression[2,3]. We were specifically interested in regions of the genome that are histone modified in a specific tissue.

We followed the exact procedure described by Finucane et al.[28], and estimated stratified LD-score regression for the 'baseline' model, which contains 53 categories. Additionally, we performed analyses by using cell-type-specific annotations for the four histone marks, corresponding to specific chemical modifications of the histone protein, which packages and orders the DNA molecule. Epigenetic modifications of histones, specifically histones bearing the marks H3K4me1, H3K4me3, H3K27ac, or H3K9ac, are associated with increased transcription. Each cell-type-specific annotation corresponds to a histone mark in a specific cell obtained from distinct human tissue, for example, H3K27ac in fetal brain cells, generating 220 combinations of histone modification by tissue. When generating estimates of enrichment for the 220 histone marks by tissue annotations, we controlled for overlap with the functional categories in the full baseline model but not for overlap with the 219 other cell-type-specific annotations. For the well-being spectrum, we ran LD-score regression on each of the 220 models (one for each histone by tissue combination) and ranked the histone by tissue annotations by P value derived from the Z values of the coefficient. Results are displayed in Supplementary Table 15.

Stratified LD-score regression of local gene expression across the human brain. We downloaded the normalized and quality-controlled gene expression measured in an anatomically comprehensive set of brain regions (see URLs). The data contained 3,707 measurements across six adult human brains (Hawrylycz et al.[29]). We computed differential gene expression for 48,154 probes mapping to 20,724 unique genes (probes that did not map to genes were omitted). We considered differential gene expression across 210 regions for which at least three measurements were available. Because Hawrylycz et al.[29] found little evidence for lateral difference in gene expression, regions in the left and right hemisphere were collapsed into a single region. For each gene in each region, a t test was performed, testing the difference in standardized expression between the region in question and all other brain regions. The top 10% of probes ranked in terms of t statistic per region were retained. The unique genes mapped to this set of probes were extracted (mapping ~2,900–3,500 genes to each region). The correlation between t statistics for the 48,154 probes identified fairly strong differential expression between the cortex, brainstem, and cerebellum and clustering of differential expression within these regions.

A partitioned LD score with respect to the genomic regions spanned by these genes (with gencode v19 used as a reference), and a 100-kb window around each gene was computed. The heritability of well-being was partitioned across the 54 baseline annotations[28] and each of the 210 brain regions (the regions were considered separately). The substantial differences in gene expression between gross anatomical brain regions (cerebellum, cortex, subcortical regions, and brainstem) dominated the results (Supplementary Table 16). We therefore proceeded to compute differential gene expression within the cerebellum, cortex, subcortical regions, and brainstem. In this analysis, we omitted the fiber bundles, because these are anatomically distinct from both the cortex and the subcortical regions yet not measured densely enough to warrant the computation of differential expression within these fiber bundle tissues. The procedure to compute differentially expressed genes was identical to the procedure used to compute differential expression across the whole brain but considered the gross anatomical regions separately. New LD scores were computed according to the local differential gene expression analyses (Supplementary Tables 17–20). All analyses were repeated by using height as a negative control trait. The genomic regions spanning genes differentially expressed in these 210 brain regions were not significantly enriched with SNP effects on height.

Stratified LD-score regression of single nuclei for seven types of neurons. We obtained the publicly available matrix of gene counts generated based on single nuclei from the prefrontal cortex and hippocampus of multiple human donors by

Habib et al.[34]. To compute differential enrichment, we deviated from the procedure outlined for regional brain expression, because the zero-inflated nature of single nuclei expression violates assumptions of the *t* test. The matrix contained counts for 32,111 genes measured in 14,964 nuclei. The nuclei were divided into seven types of neurons, two subtypes of astrocytes, oligodendrocytes, oligodendrocyte precursor cells, microglia, endothelial cells, and unclassified cells. We omitted genes for which the total count across cells was <150, or for which fewer than 30 cells had a count above 0, retaining 11,719 genes for analysis. For each gene, we computed the ratio of count per nuclei type over the total number of nuclei measured of the specific type (generating the average gene count in each nuclei type). Next, we computed the ratio of the average count per nuclei type over the average count of the gene across all nuclei (generating the nuclei-type-specific fold change in average expression). We then defined, for each nuclei type, the nuclei type specifically expressed set of genes as the 1,600 genes with the highest nuclei-type-specific fold change in average expression. For each of the gene sets, we constructed an LD score with respect to genes in the set to compute the gene-set-specific enrichment in h^2 in our multivariate GWAMA.

Our method to determine cell-type-specific expression purely relied on the relative (standardized) expression in one cell type over the (standardized) expression in others, whereas others have developed perhaps more sophisticated statistics to assess differential expression[55]. The test statistics developed by Finucane et al.[55] have never been applied to single-cell data, which have a skewed distribution of counts, and we, therefore, used a simpler metric. Application of both methods to the same gene expression dataset (Genotype-Tissue Expression (GTEx) project) and subsequent differential enrichment analysis for well-being yielded highly correlated enrichment estimates ($r = 0.83$).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability

N-GWAMA and MA-GWAMA software is available at: https://github.com/baselmans/multivariate_GWAMA/

Data availability

Summary Statistics excluding results from 23AndMe can be downloaded from https://surfdrive.surf.nl/files/index.php/s/Ow1qCDpFT421ZOO. The data transfer agreement with 23AndMe stipulates that we can publish effect sizes associated with 10,000 SNPs. These summary statistics can be downloaded from https://surfdrive.surf.nl/files/index.php/s/Ow1qCDpFT421ZOO. For 23AndMe dataset access, see https://research.23andme.com/dataset-access/. The Understanding Society data are distributed by the UK Data Service. The genome-wide scan data were analyzed and deposited by the Wellcome Trust Sanger Institute. Information on how to access the data can be found on the Understanding Society website at https://www.understandingsociety.ac.uk/. Genotype-trait data access for UKHLS is available by application to Metadac through http://www.metadac.ac.uk/.

References

40. Lee, D., Bigdeli, T. B., Riley, B. P., Fanous, A. H. & Bacanu, S. A. DIST: direct imputation of summary statistics for unmeasured SNPs. *Bioinformatics* **29**, 2925–2927 (2013).
41. Akaike, H. A. Bayesian extension of the minimum AIC procedure of autoregressive model fitting. *Biometrika* **66**, 237–242 (1979).
42. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48 (2009).
43. Mazerolle, M. J. *AICcmodavg: Model Selection and Multimodel Inference Based on (Q)AIC(c), R Package version 2.1-1* (R Foundation for Statistical Computing, 2017).
44. van Beijsterveldt, C. E. M. et al. The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. *Twin. Res. Hum. Genet.* **16**, 252–267 (2013).
45. Diener, E. D., Emmons, R. A., Larsen, R. J. & Griffin, S. The satisfaction with life scale. *J. Pers. Assess.* **49**, 71–75 (1985).
46. Lyubomirsky, S. & Lepper, H. S. A measure of subjective happiness: preliminary reliability and construct validation. *Soc. Indic. Res.* **46**, 137–155 (1999).
47. Costa, P. T. & McCrae, R. R. *The Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual* (Psychological Assessment Resources Inc., Odessa, 1992).
48. Achenbach, T. M. & Rescorla, L. *Manual for the ASEBA Adult Forms & Profiles* (University of Vermont, Burlington, Vermont, USA, 2003).
49. Tennant, R. et al. The Warwick-Edinburgh mental well-being scale (WEMWBS): Development and UK validation. *Health Qual. Life Outcomes* **5**, 1–13 (2007).
50. Altshuler, D. M. et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
51. Jansen, R. et al. Conditional eQTL analysis reveals allelic heterogeneity of gene expression. *Hum. Mol. Genet.* **26**, 1444–1451 (2017).
52. Zhernakova, D. V. et al. Identification of context-dependent expression quantitative trait loci in whole blood. *Nat. Genet.* **49**, 139–145 (2017).
53. Fehrmann, R. S. N. et al. Trans-eqtls reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the hla. *PLoS Genet.* **7**, e1002197 (2011).
54. Bonder, M. J. et al. Disease variants alter transcription factor levels and methylation of their binding sites. *Nat. Genet.* **49**, 131–138 (2017).
55. Finucane, H. K. et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).

nature research

Corresponding author(s): Prof. Meike Bartels

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No Software was used for data collection
Data analysis	GWAS was conducted using PLINK v1.9. Simulation of phenoytypes was conducting using GCTA v1.91 Construction of Polygenic Scores for out of sample prediction was conducted using PLINK v1.9 Genetic correlations were constructed using LD Score regression Stratified LD Score regression was conducted using LD Score regression N-weighted and model-averaging GWAMA was run using self-developed software and can be downloaded from Github: https://github.com/baselmans/multivariate_GWAMA

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Summary Statistics excluding results from 23AndMe can be downloaded from https://surfdrive.surf.nl/files/index.php/s/Ow1qCDpFT421ZOO. The data transfer agreement with 23AnMe stipulates that we can publish effect sizes associated with 10,000 SNPs. These summary statistics can be downloaded from https://surfdrive.surf.nl/files/index.php/s/Ow1qCDpFT421ZOO. For 23AndMe dataset access see: https://research.23andme.com/dataset-access/

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For each of the four phenotypes, we combine all publicly available summary statistics from new association analyses. Details are reported in Supplementary Figure 2 and Supplementary Table 3
Data exclusions	Analyses is restricted to samples with Caucasian ancestry
Replication	n/a: A broad range of simulation scenarios supports the validity of both our multivariate methods
Randomization	n/a: We did not used an experimental design
Blinding	n/a: We did not used an experimental design

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging