Genetic architecture of \textit{DCC} and influence on psychological, psychiatric and cardiometabolic traits in multiple ancestry groups in UK Biobank

\textbf{Rona Strawbridge (✉ rona.strawbridge@glasgow.ac.uk)}
University of Glasgow \hspace{1cm} \url{https://orcid.org/0000-0001-8506-3585}

\textbf{Lewis Forsyth}
\textbf{Breda Cullen}
School of Health and Wellbeing, University of Glasgow \url{https://orcid.org/0000-0002-7259-9505}

\textbf{Nicholas Graham}
\textbf{Donald Lyall}
\textbf{Laura Lyall}
\textbf{Jill Pell}
\textbf{Joey Ward}
joey.ward@glasgow.ac.uk

\textbf{Daniel Smith}
Edinburgh University, Scotland

\textbf{Article}

\textbf{Keywords:}

\textbf{Posted Date:} October 19th, 2022

\textbf{DOI:} \url{https://doi.org/10.21203/rs.3.rs-2158925/v1}

\textbf{License:} ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.
\url{Read Full License}
Abstract

People with severe mental illness have a higher risk of cardiometabolic disease than the general population. Traditionally attributed to sociodemographic and behavioural factors and medication effects, recent genetic studies have provided evidence of shared biological mechanisms underlying mental illness and cardiometabolic disease. This study aimed to determine whether signals in the DCC locus, implicated in cardiometabolic and psychiatric conditions, were shared with, or distinct. Using the UK Biobank cohort, we systematically assessed the impact of genetic variation in the DCC (deleted in colorectal carcinoma) locus on traits related to cardiometabolic and psychiatric conditions in unrelated “white British” participants (N = 402837). Logistic or linear regression were applied assuming an additive genetic model and adjusting for age, sex, genotyping chip and population structure (eight genetic principal components). Bonferroni correction for the number of independent SNPs within the locus was applied. Conditional analyses (including lead variants as covariates) and trans-ancestry analyses were used to investigate linkage disequilibrium between signals. Significant associations were observed between DCC variants and smoking, anhedonia, body mass index (BMI), neuroticism and mood instability, with multiple conditionally-independent signals being identified for the latter three traits. Conditional analyses and linkage disequilibrium structure suggested signals for smoking and BMI were distinct from each other and the mood traits, whilst individual mood traits were inter-related in a complex manner. Genetic variation in the DCC locus had distinct effects on BMI, smoking and mood traits, and therefore is unlikely to contribute to shared mechanisms underpinning mental and cardiometabolic traits.

Introduction

A link is well-established between mental health traits (MHTs) and cardiometabolic traits (CMTs) in epidemiological studies, with the presence of severe mental disorders resulting in a reduction in life expectancy of 15 years in women and 20 years in men \(^1\), with a large proportion of deaths being attributed to CMT \(^2\). Indeed, individuals with serious mental illness have estimated 3.6 times higher likelihood of developing a cardiometabolic disease \(^3\). Though this association is clear, there exists very little knowledge about the mechanisms. Key contributors to this association may be lifestyle factors such as exercise, diet and drug use \(^4\) with additional links being drawn between the use of treatments for psychosis and increased body mass index (BMI) \(^5\). However, contemporary genetic studies suggest shared biological mechanisms underlying this association \(^6\)–\(^13\). A number of loci have been identified wherein genetic variation is pleiotropic for both cardiometabolic and mental health associations \(^6\), with many more loci implicated in both MHT and CMT.

Deleted in Colorectal Carcinoma (DCC) is a transmembrane receptor, which transmits signals involved in axonal development. Genetic variation in this locus has been implicated (through genome-wide association studies, GWAS) in significant number of phenotypes relevant to MHT and CMT including brain volume \(^14\), depression \(^15\), neuroticism \(^16\) and glucose homeostasis \(^17\), although there has been no systematic analysis of whether these associations overlap or are distinct. Until recently, GWAS reported
only the lead variants for each trait. Therefore it was impossible to compare effects of a locus across traits. Today, summary statistics are available for all variants analysed in GWAS, however differences in study recruitment and analytic design hinders appropriate cross-trait comparisons. The UK Biobank (UKB) provides new opportunities for appropriate comparisons between traits, with phenotypic and genetic data being available on CMTs and MHTs for ~0.5M participants in large sample sizes with consistent recruitment, statistical modelling, and data handling.

We set out to systematically investigate the DCC locus: for association with a wide range of psychological, psychiatric and cardiometabolic traits; to describe the genetic architecture underlying these associations and to explore mechanisms by which variants might have their effects.

Materials And Methods

Study description

UK Biobank recruited ~500 000 individuals at 22 centres across the UK, between 2006–2010, and has been described in detail elsewhere[18–20]. Blood was sampled and stored for genetic analysis. Participants underwent a physical examination and completed extensive questionnaires on lifestyle, personal and family history of disease. Baseline questionnaire data provided information on current smoking (data field #20116, current smokers vs former/non-smokers), risk-taking behaviour (#2040, “do you consider yourself to be someone who takes risks?”), mood instability (#1920, “does your mood often go up and down?”) and anhedonia (#2060, “over the past two weeks, how often have you had little interest or pleasure in doing things?”) Controls were those who responded “not at all”, with other responses being considered cases. Neuroticism (#20127) was assessed using the Eysenck Personality Questionnaire (Revised Short Form) which consisted of 12 yes/no (coded 0/1) questions (including #1920), which were summed, BMI was calculated from baseline height and weight measurements (#21001). Waist:hip ratio adjusted for BMI (WHRadjBMI) was calculated in a sex-and ancestry-specific manner from recorded waist (#48), hip (#49) and BMI (#21001) measurements as per Shungin et al[21]. WHRadjBMI values for men and women were analysed together by ancestry group. Systolic and diastolic blood pressure were recorded (average of two measurements) and adjusted for effects of anti-hypertensive medication where appropriate (as per Ehret et al, SBP + 15 mmHg and DBP + 10 mmHg if using anti-hypertensive medication,[22]). Probable type 2 diabetes (T2D) was defined as per Eastwood et al[23]. Ischemic heart disease (IHD, heart attack/angina) and stroke were assessed from self-report of a diagnosis (#6510). Venous thromboembolism was self-reported (deep-vein thrombosis and/or pulmonary embolism, #6152). A subset of UK Biobank participants completed an online mental health questionnaire (6–10 years after baseline)[24], enabling assessment of probable lifetime generalised anxiety disorder (GAD), bipolar disorder (BD) and major depressive disorder (MDD)[25]. Participants responding “don’t know” or “prefer not to say” to any question were excluded from analyses (<5%). Ancestry groups were broadly defined ancestry groups as per Eastwood et al[23] (see Supplemental Methods).
This work was approved under the generic ethical approval for UK Biobank studies granted by the NHS National Research Ethics Service (approval letter dated 29 June 2021, Ref 21/NW/0157). This project used data from UK Biobank applications 6553 (PI. RJS).

**Genetic data**

Extracted DNA was genotyped using the Affymetrix UK BiLEVE Axiom or UK Biobank Axiom platforms \(^{18}\), and imputed to the 1000 Genomes and Haplotype Reference Consortium reference panels \(^{18}\). Standard genetic quality control of pre and post-imputation data was conducted by the central UK Biobank team. Genetic data was used to verify an individual’s categorisation in the “white British” ancestry subset \(^{18}\).

**SNP selection**

The DCC locus was defined as the coding region of DCC ± 400kb (UCSC genome browser (https://genome-euro.ucsc.edu/cgi-bin/hgGateway?redirect=manual&source=genome.ucsc.edu), build 37). Genetic variants in this region were filtered for minor allele frequency > 1%, by ancestry group.

**Genetic analyses**

Individuals of self-reported “white British” ancestry make up the majority of the cohort. Therefore, genetic analyses were initially restricted to unrelated individuals of self-reported “white British” ancestry. Subsequently, analyses of associated phenotypes with all variants with MAF > 1% by ancestry group were conducted in the additional ancestry groups (as defined by Eastwood et al \(^{23}\)), white (non-British) European, South Asian, African-Caribbean and mixed ancestry groups \(^{23}\).

For the primary analyses in “white British” individuals, genetic analyses of 7 161 variants in the DCC locus were conducted using Plink 1.07 \(^{26}\), assuming an additive genetic model. For continuous and binary variables, linear and logistic regression were applied, respectively. All models were adjusted for age, sex, genotyping chip and population structure (eight genetic principal components), except WHRadjBMI. Analyses of IHD and Stroke were further adjusted for current smoking, anti-hypertensive and lipid-lowering medication. The covariates above were incorporated into the calculation of WHRadjBMI, which was performed separately by sex and ancestry group, therefore no covariates were used for genetic analyses of WHRadjBMI. Sensitivity analyses are described in the Supplemental Methods.

Secondary analyses in European, South Asian, African-Caribbean and mixed ancestry groups included 7 123, 7 720, 10 439 and 8 689 variants respectively.

The standard threshold for significance in a GWAS is \(P < 5 \times 10^{-5}\) (Bonferroni correction for 1 million tests). As this study is focusing on the DCC locus only, and because of prior evidence implicating the DCC locus in MHT and CMT, this threshold would be unnecessarily conservative. Therefore, Bonferroni correction for multiple testing was applied, with adjustment for the number of independent variants in the DCC locus for the “white British” ancestry sample. This was calculated using Plink 1.07 \(^{26}\) and the pairwise independence test (using default settings, see Supplemental Methods). For the “white British”
ancestry sample, of 7,161 SNPs in the DCC locus 1,419 were independent thus significance was set at p < 3.52x10^{-5} (0.05/1419). Whilst it is likely that the number of independent variants would vary between the ancestry groups due to differing LD structures, the same significance threshold was used for all ancestry groups.

**Meta-analyses**

To assess trans-ethnic consistency, inverse variance weighted (based on Beta and se) meta-analyses of ancestry-specific results for each phenotype, including all variants analysed for each ancestry, was conducted using METAL \(^{27}\) (See Supplemental Methods). Odd ratios (OR) were converted to beta coefficients for this analysis, as METAL is capable of handling quantitative but not binary phenotypes. Population stratification was controlled for in the ancestry-specific analyses, not in the meta-analyses. The significance threshold remained at p < 5.37x10^{-5}.

**Genetic architecture of DCC**

The genetic architecture of significant SNPs within the locus was assessed using Haploview \(^{28}\) to visualise linkage disequilibrium (LD) blocks, separately for each ancestry group. In addition, conditional analyses (including lead SNPs as covariates) were employed to assess the number of conditionally-independent or secondary signals for each significant trait (trait1 ~ age, sex, population structure, genotyping chip and lead SNP for trait1) and whether signals for each trait were distinct (trait1 ~ age, sex, population structure, genotyping chip and lead SNP for trait2).

**Follow-up analyses**

The Variant Effect Predictor (VEP) \(^{29}\) was used to assess the impact of all variants meeting the threshold for statistical significance (in any analyses). Genotype-specific effects on expression quantitative trait loci (eQTLs) were identified in two ways: firstly, lead variants and those with potential functional effects, and any SNPs in high LD (\(r^2 > 0.75\) in Europeans), were assessed for effects on expression of DCC or other nearby gene using the LDExpress (https://ldlink.nci.nih.gov/?tab=ldexpress); secondly, all eQTLs for DCC were identified using the GTEx resource \(^{30}\) and compared to the genetic association analyses.

**Comparison with literature**

The GWAS catalogue (https://www.ebi.ac.uk/gwas, 20210210) was used to identify variants in the DCC locus previously reported to be associated with a behavioural, cardiometabolic or psychiatric trait. Where possible, previously reported associations were compared to those observed here.

**Results**

The characteristics of the cohort, by ancestry group, are presented in Table 1.
Individual trait analyses in “white British” ancestry individuals

The significant associations between the DCC locus and phenotypes analysed are summarised in Table 2 and Fig. 1.

Significant associations were identified for BMI (42 SNPs, lead rs5824977, Table 1, Fig. 1A), Smoking (31 SNPs, rs12608052, Table 1, Fig. 1B), GAD (42 SNPs, 18:50678953-GA (proxy rs7229097), Table 1, Fig. 1C), MDD (86 SNPs, rs62098013, Table 1, Fig. 1D), mood instability (883 SNPs, rs17411061, Table 1, Fig. 1E), neuroticism score (1005 SNPs, rs8099145, Table 1, Fig. 1F) and anhedonia (1228 SNPs, rs11660938, Table 1, Fig. 1G). All associations demonstrated effect sizes in line with those normally observed for complex traits (ie. fairly small). The regional plots for most of these traits are suggestive of multiple signals, therefore conditional analyses were conducted. No significant associations were identified for risk-taking behaviour, BD, SBP, DBP, ISH, stroke, VTE, T2D or WHRadjBMI.

Secondary (or conditionally-independent) signals were defined as variants passing the threshold for significant associations and the Beta or OR being minimally reduced (a difference of \( \leq 0.05 \)) when including the lead variant as a covariate. Smoking, GAD, MDD and anhedonia demonstrated no significant secondary signals. For BMI (Figure S1A) a second, conditionally-independent signal was identified (rs7230034-T, Beta\(_{\text{discovery}}\)-0.076 (0.017) -> Beta\(_{\text{conditional}}\)-0.072 (0.017), Table 1, Figure S1B, Table S1). Further conditional analyses including both primary and secondary lead SNPs (rs5824977 and rs7230034) rendered the locus non-significant (Figure S1C, Table S1). Mood instability (Figure S1D) demonstrated a second, conditionally-independent signal (rs8096647-T, 1.02 (1.01–1.03) -> 1.02 (1.01–1.03), Table 1, Figure S1E, Table S2). Further conditioning on both signals removed the signal (Figure S1F). For neuroticism score (Figure S1G), a second, conditionally-independent signal was identified (rs28698732-A, -0.055 (0.011)-> -0.060 (0.011), Figure S1H, Table S3). Further conditioning on both signals removed the signal (Figure S1I, Table S3).

Given that GAD, MDD, mood instability, neuroticism score and anhedonia are positively related phenotypes, it was reassuring that the same effect direction was observed for all of their lead SNPs across all traits (Table 2).

Sensitivity analyses demonstrated that significant associations with MDD were not being driven by inclusion of individuals with GAD (Table S4) and associations with GAD were not driven by inclusion of individuals with MDD (Table S5), GAD and MDD are highly comorbid, so overlapping samples are inevitable, however these analyses suggest that the associations with GAD and MDD are not solely driven by inclusion in both analyses of a co-morbid GAD-MDD subsample. For associations with mood instability (Table S6), neuroticism score (Table S7) or anhedonia (Table S8), effect sizes were comparable, although in some cases associations were no longer significant (which is likely due to the reduced sample size in the sensitivity analyses). Therefore, associations between the DCC locus and mood-related traits are not being driven by individuals with mental illness.
Cross-trait analyses in “white British” ancestry individuals

To determine whether the signals for MHTs and CMTs overlapped or were distinct, conditional analyses including the lead SNP from the other traits were undertaken.

For BMI, including the other traits lead signals had a negligible effect on the association (Figure S2, Beta range −0.081−0.086). Similarly, the signal for smoking was unchanged after adjustment for the lead SNPs of the other traits (Figure S3, OR range 0.97−0.98).

For GAD, adjustment for BMI or smoking lead SNPs had no effect on the association (Figure S4A-C, OR 1.08). In contrast, the associations of DCC genetic variants with GAD were non-significant after adjustment for MDD, mood instability, neuroticism score or anhedonia lead SNPs (Figure S4D-G, ORs 1.08−1.09). The associations with MDD demonstrated the same null effect when adjusting for the BMI or smoking signal (Figure S5A-C, OR 0.99) and non-significant associations after adjustment for psychiatric-related traits (Figure S5D-G, ORs 0.98−0.99).

For mood instability, again the signal was conditionally-independent from that of BMI or smoking (Figure S6A-C, OR 1.03). Adjustment for GAD or MDD reduced but did not remove the association (Figure S6D-E, ORs 1.03−1.05), whereas adjustment for neuroticism score or anhedonia rendered the associations non-significant (Figure S6F-G, ORs 1.03−1.06). Associations with neuroticism score were unaffected by adjustment for BMI or smoking (Figure S7A-C Beta −0.058). Adjustment for GAD or MDD reduced the significance of the associations (Figure S7D-E, Betas 0.051 - -0.062), whereas adjustment for neuroticism score or anhedonia rendered the associations non-significant (Figure S7F-G, Betas 0.037 - -0.038). Associations with anhedonia were unchanged by conditioning on the BMI or smoking lead SNPs (Figure S8A-C, OR 1.05). Adjustment for GAD and MDD reduced but did not remove the association (Figure S8D-E, ORs 1.05−1.06), whereas adjustment for mood instability or neuroticism score rendered the associations non-significant (Figure S8F-G, OR 1.04 for both).

These results indicated that the BMI and smoking loci were distinct from each other and from the mood traits, but that the signals for mood traits were interrelated.

Associations in European, South Asian, African-Caribbean and Mixed ancestry individuals

Secondary analyses were conducted in European, South Asian, African Caribbean and mixed ancestry groups. Sample sizes for these ancestry groups were much smaller than the “white British” ancestry subset, in particular for self-reported GAD or MDD where there was insufficient power for analyses. Therefore, trans-ancestry analyses and meta-analyses focused on baseline BMI, smoking, mood instability, neuroticism score and anhedonia. In European ancestry samples, a significant association was identified for BMI (rs2339638, Figure S9A, Table S9). In the mixed ancestry sample, an association was observed for anhedonia (four SNPs, rs7232267, Figure S9B, Table S9). No significant associations
were observed in the African-Caribbean or south Asian samples. As these were the smallest samples (N < 10,000), this is perhaps unsurprising.

**Meta-analyses across multiple ancestry groups**

To investigate whether genetic effects of the \textit{DCC} locus on BMI, smoking, mood instability, neuroticism score and anhedonia were consistent across ancestry groups, meta-analyses were conducted (irrespective of significance in the individual ancestry groups). For BMI, 115 SNPs reached significance (Table 3 and Figure S10A), with 74% having low heterogeneity (I² < 25, Figure S10B). 1029 SNPs were significant in the meta-analysis of neuroticism score (Figure S10C), 23% of which had low heterogeneity (Figure S10D) and 67% had low or moderate heterogeneity (I² < 50). Whilst the majority of these variants were associated exclusively with BMI or neuroticism, there was one variant, rs11872713 (Table S10), significantly associated with both traits with low heterogeneity (I² = 0) for BMI but moderate heterogeneity (I² = 44) for neuroticism. The allele associated with reduced BMI was also associated with reduced neuroticism. No significant associations were observed in the meta-analyses of smoking, mood instability or anhedonia.

**Linkage disequilibrium analyses**

Linkage disequilibrium analyses (Fig. 2 and Figure S11) confirmed that the “white British” lead SNPs for BMI and smoking were rarely coinherited with each other (maximum LD r² = 0.13) or the signals for mood-related traits. Two observations stood out regarding the mood-related traits in “white British” ancestry samples (Fig. 2): firstly, a handful of SNPs constituted a single signal that influenced MDD, anhedonia, mood instability and GAD and neuroticism (unsurprising given the phenotypic relationships between these traits); secondly, for neuroticism and mood instability, there were additional conditionally-independent signals which were distinct from the lead mood traits signal and each other.

The signal for BMI in Europeans had minimal LD with the other lead SNPs (max r² = 0.09).

In the trans-ancestry meta-analysis, the BMI lead SNP (rs5824977) was the same as that for the “white British” analysis, which was unsurprising given the predominance of the “white British” sample. However, it appeared that the signal was consistent across ancestry groups, as demonstrated by heterogeneity Isq = 0 (Table 3). This was in contrast with the meta-analysis of neuroticism, where the lead SNP in the meta-analysis (rs7230285) was in high LD with that for the “white British” analysis (r² = 0.91), but high heterogeneity (Isq = 60%), which suggests ancestry group differences. That there were SNPs with low heterogeneity (rs1943107, Isq = 0) indicates that there were some genetic effects that were consistent across ancestry groups, as well as ancestry-specific effects.

The SNP (rs11872713) with effects on both neuroticism and BMI in the meta-analysis demonstrates varying degrees of LD with the (low heterogeneity) meta-analysis neuroticism signal, with low LD in African-Caribbean (r² = 0.13), “white British” (r² = 0.23) and European (r² = 0.25) ancestry groups or moderate LD in south Asian (r² = 0.28) or mixed (r² = 0.37) ancestry groups. How this should be interpreted is unclear.
Predicted functional effects and genotype-specific gene expression patterns

Of 1497 unique SNPs significant in any analysis, only rs2229080 (associated with anhedonia in the “white British” analysis, G allele, OR = 0.98) was predicted to have a functional effect by VEP, with the G allele resulting in a missense transcript that was tolerated or benign (according to SIFT and PolyPhen respectively). Rs2229080 (chr18_52906232_C_G_b38) demonstrated genotype-specific gene expression (eQTL) in the GTEx dataset, in nerve tissue for DCC (G allele associated with higher DCC mRNA levels, Fig. 3A) and LINC01917, the latter of which is a long non-coding RNA of unknown function with little expression outside of the testis. It is worth noting that rs2229080-G was the minor allele in European and African ancestry but not in south Asian ancestry (in both UK Biobank and dbSNP (https://www.ncbi.nlm.nih.gov/snp/rs2229080)). Outside of the testis, DCC was predominantly expressed in the brain tissue using GTEx (Fig. 3B), with little expression in the main metabolic tissues. This combined with SNP effects on neuroticism score and BMI could suggest that effects on BMI are subsequent to those on neuroticism. None of the lead SNPs, or high-LD proxies from single ancestry analyses or meta-analyses (Table 3, Table S10) demonstrated genotype-specific gene expression patterns.

Considering expression of the DCC gene in GTEx, of 1334 identified eQTLs (Table S11), 63% were identified in nerve tissue and the strongest 395 effects were detected in nerve tissue. Sixty-five SNPs with eQTLs also had significant association with MDD, mood instability, anhedonia or neuroticism score (Table S11). Where eQTLs were observed in both the adrenal gland and nerve tissue, the same allele had the opposite effects on mRNA levels of DCC (positive in one, negative in the other). The allele associated with increased risk of anhedonia (and consistent effect on one or more of mood instability, neuroticism score and/or MDD) was consistently associated with reduced DCC mRNA levels.

Comparison with literature

The DCC locus has previously been implicated (through GWAS) in many traits including behavioural, psychiatric and cardiometabolic diseases (Table S12). The effect directions from our study were compared to those published where there was sufficient information and loosely comparable traits (Table S13). Further detail is provided in the Supplemental Results.

Discussion

Here we present a systematic assessment of genetic variation in the DCC locus for impact on a wide range of MHT and CMT. In a very large sample of “white British” ancestry individuals, we identified significantly associated signals for current smoking, BMI, anhedonia, neuroticism, mood instability, MDD and GAD. Additional analyses demonstrated that: BMI and mood traits had multiple conditionally-independent signals; BMI, smoking and mood traits constituted distinct signals; some of the BMI and
mood trait signals appeared to be relevant across ancestry groups; genetic variation influences mood-related traits through expression of \textit{DCC} in the brain.

The \textit{DCC} (deleted in colorectal cancer) locus has been implicated in many mood-related and psychiatric traits, therefore most functional analyses of \textit{DCC} have investigated neuronal development or tumour progression. Evidence for metabolic traits is restricted to family studies which have described \textit{DCC} variants in idiopathic hypogonadotropic hypogonadism \(^1\), which includes changes in both mood and weight. In mice, homozygous knockout of \textit{dcc} is neonatal lethal, whilst heterozygous knockout has no impact on growth, metabolic, cardiovascular or neurological systems (https://www.mousephenotype.org/data/genes/MGI:94869#phenotypes-section). These results do not confirm our findings however it is possible that mood traits such as anhedonia, mood instability and neuroticism might require a stressor in addition to genetic predisposition for presentation. In addition, if increased BMI is secondary to mood traits, it might not be evident in a mouse model under controlled feeding conditions.

Our finding that there was a \textit{DCC} signal shared by a number of mood-related traits as well as additional conditionally-independent signals (which may be trait- and ancestry-specific) suggests that haplotype analyses of this region in diverse ancestries and with a wide range of phenotypes is required to better understand the complexity of this locus. BMI is complex, but can be considered a behavioural trait (through food preferences, feeding and exercise patterns). The results presented here suggest that BMI and smoking could be secondary to mood traits, given lack of evidence for direct effects of \textit{DCC} or neighbouring genes in metabolic tissues. Alternatively, Couvy-Duschesne et al \(^2\) demonstrated that relationships between brain size measures and depression were rendered null when adjusting for BMI (using mainly the same data as was included in this study). Genetic variation in this locus has been associated \(^3\) with the same regions as were analysed by Couvy-Duschesne \textit{et al}, suggesting that this locus could act through effects on brain size.

Rare variation in \textit{DCC} was investigated by Backman \textit{et al} using whole exome sequencing (in UKB), for traits including those investigated here, but no associations were identified \(^4\). Strict multiple testing correction could mean some true associations might not have been reported by Backman \textit{et al}. However, it appears that current evidence provides more support for influence of common than rare genetic variation in the \textit{DCC} locus on complex MHT and CMT.

\textbf{Limitations}

UKB is not truly representative of the general population (skewed towards lower deprivation than average) \(^5\), however this is unlikely to invalidate our findings. As UKB represents the healthy end of the health to disease spectrum, the range of phenotypes (for example BMI or blood pressure measurements) is smaller than for the general population. Similarly, as severe mental or physical illness is likely to be a barrier to participation in the UKB, the cases here are less ill/less different from the controls than if severe cases were included. Thus, these results might be an under-estimate of true effects. We acknowledge that these
results might not be generalisable outside of the UK population. Selecting phenotypes for secondary analyses in additional ancestry groups, based on results from the “white British” ancestry subset is a further limitation, however restricted power in the smaller ancestry groups could render such analyses uninformative due to low N. Including genetic data for non-European ancestry individuals imputed to reference panels that perform better for European individuals is not ideal. These limitations mean that whilst what we present is likely true, is does not represent a complete picture of the genetic architecture of the DCC locus in non-European individuals. Furthermore, we cannot exclude the possibility that genetic variants are correlated with or interact with covariates in an ancestry-dependent manner.

Conclusions

This study demonstrates the complexity of the DCC locus, with distinct signals influencing BMI, smoking and mood-related traits, with some traits having trans-ancestry and ancestry-specific signals. Future assessment of the DCC locus should consider multiple signals, for example using haplotype analysis. We cannot exclude DCC contributing to shared biological mechanisms underlying MHT and CMT, but current evidence is more suggestive of effects on BMI being secondary to those on mood-related traits.

Declarations

Acknowledgements

The authors thank all participants and staff of the UK Biobank study. This work uses data provided by patients and collected by the NHS as part of their care and support. The UK Biobank was established by the Wellcome Trust, Medical Research Council, Department of Health, Scottish Government and Northwest Regional Development Agency. UK Biobank has also had funding from the Welsh Assembly Government and the British Heart Foundation. Data collection was funded by UK Biobank. This project was completed using UK Biobank applications 6533 (Pl. DJS) and 1755 (Pl. JPP).

RJS is supported by a University of Glasgow LKAS fellowship and a UKRI Innovation-HDR-UK Fellowship (MR/S003061/1). DJS acknowledges the support of a Lister Prize Fellowship (173096) and MRC Mental Health Data Pathfinder Award (MC_PC_17217).

Conflict of Interest

The authors have no conflicts of interest to declare.

Supplementary information is available at MP’s website

References


Tables

Tables 1-3 are available in the supplementary files section.

Figures
Figure 1

Regional plots of white British ancestry analyses of A) BMI, B) current smoking, C) GAD, D) MDD, E) mood instability, F) neuroticism score, G) anhedonia. Of note, the lead SNP for GAD, 18:50678953_GA_G is not present in the Locuszoom reference panel so rs7229097 is plotted instead. Colours represent LD with the lead SNP, whilst grey indicates lack of LD information. The yellow horizontal line indicates threshold for significance.
Figure 2

Linkage disequilibrium of associated SNPs and the respective phenotypes and analyses. The plot gives the LD between SNPs in a random selection of 10,000 unrelated white British ancestry individuals. Colours and values of LD are given as R2.
Figure 3

mRNA expression of DCC, A) according to rs2229080 (chr18_52906232_C_G_b38) genotype and B) across all tissues in the GTEx database.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- DCCSupplementallInformation20220719.docx
- Table120221012.xlsx
- Table220221012.xlsx
- Table320221012.xlsx
- SFigure1020221012.jpg
- SFigure1120221012.jpg
- SFigure120221012a.jpg