Original research

Multitrait genetic association analysis identifies 50 new risk loci for gastro-oesophageal reflux, seven new loci for Barrett's oesophagus and provides insights into clinical heterogeneity in reflux diagnosis

Jue-Sheng Ong [ORCID],[1] Jiyuan An,[2] Xikun Han,[1] Matthew H Law,[1,3] Priyanka Nandakumar,[4] 23andMe Research team, Esophageal cancer consortium, Johannes Schumacher,[5] Ines Gockel,[6] Anne Bohmer,[7] Janusz Jankowski,[8,9] Claire Palles,[10] Catherine M Olsen,[11,12] Rachel E Neale,[11] Rebecca Fitzgerald,[13] Aaron P Thrift [ORCID],[14] Thomas L Vaughan,[15] Matthew F Buas,[16] David A Hinds,[4] Puya Gharahkhani,[1] Bradley J Kendall [ORCID],[12,17] Stuart MacGregor [ORCID][1]

ABSTRACT

Objective Gastro-oesophageal reflux disease (GERD) has heterogeneous aetiology primarily attributable to its symptom-based definitions. GERD genome-wide association studies (GWASs) have shown strong genetic overlaps with established risk factors such as obesity and depression. We hypothesised that the shared genetic architecture between GERD and these risk factors can be leveraged to (1) identify new GERD and Barrett's oesophagus (BE) risk loci and (2) explore potentially heterogeneous pathways leading to GERD and oesophageal complications.
Design We applied multitrait GWAS models combining GERD (78 707 cases; 288 734 controls) and genetically correlated traits including education attainment, depression and body mass index. We also used multitrait analysis to identify BE risk loci. Top hits were replicated in 23andMe (462 753 GERD cases, 24 099 BE cases, 1 484 025 controls). We additionally dissected the GERD loci into obesity-driven and depression-driven subgroups. These subgroups were investigated to determine how they relate to tissue-specific gene expression and to risk of serious oesophageal disease (BE and/or oesophageal adenocarcinoma, EA).
Results We identified 88 loci associated with GERD, with 59 replicating in 23andMe after multiple testing corrections. Our BE analysis identified seven novel loci. Additionally we showed that only the obesity-driven GERD loci (but not the depression-driven loci) were associated with genes enriched in oesophageal tissues and successfully predicted BE/EA.
Conclusion Our multitrait model identified many novel risk loci for GERD and BE. We present strong evidence for a genetic underpinning of disease heterogeneity in GERD and show that GERD loci associated with depressive symptoms are not strong predictors of BE/EA relative to obesity-driven GERD loci.

INTRODUCTION

Gastro-oesophageal reflux disease (GERD) is a complex condition with different consensus

Significance of this study

What is already known on this subject?
► Gastro-oesophageal reflux disease (GERD) is a common but heterogeneous disorder, which increases one's risk of developing Barrett's oesophagus (BE) and oesophageal adenocarcinoma (EA). Previous studies have identified 25 genetic loci, which confer risk for a broad GERD phenotype and have shown genetic overlaps with risk factors such as obesity and depression.

What are the new findings?
► By leveraging genetic data on BE, GERD and its risk factors in a multitrait model, we identified 88 risk loci for GERD and seven novel loci for BE. Stratifying the GERD loci by association on body mass index and depression revealed differences in gene expression across tissue types. Only loci falling in an obesity-driven GERD subgroup predicted the risk of BE/EA.

How might it impact on clinical practice in the foreseeable future?
► This study markedly increases our knowledge of the specific genes conferring GERD and BE risk. One subset of the GERD genes acts along a depression-driven GERD axis (and typically does not lead to downstream alterations of BE/EA risk), while another subset demarcates an obesity-driven GERD axis, whereby a person's genetic risk of GERD is likely to translate to a subsequent increase in risk of BE and EA.

statements using a variety of clinical, endoscopic and physiological criteria to define the disease.[1–3] Although in a narrow sense, GERD can be considered the result of pathological oesophageal acid exposure[4], the heterogeneous definition of GERD may help explain the range of traits that have been

associated with GERD risk including obesity and psychological traits.[5][6] GERD has risen in prevalence in many western countries in recent years and it now accounts for a substantial proportion of direct healthcare costs for digestive diseases.[7] GERD is the major risk factor for Barrett's oesophagus (BE), a precancerous condition in which GERD-induced erosions of the distal oesophageal squamous epithelium are replaced by metaplastic columnar epithelium.[8] Most oesophageal adenocarcinomas (EA), a lethal malignancy with poor survival, arise from underlying BE tissue.[9]

Recent studies have sought to better understand GERD and its downstream consequences (BE, EA) by identifying the genes influencing risk of GERD. For most complex conditions, the most successful study design has been the genome-wide association study (GWAS), with large sample size and a key feature in successful studies. Initial GWAS for GERD failed to identify any risk genes,[10][11] but more recent approaches using a broad phenotype definition to maximise sample size identified 25 risk loci.[12] The recent GWAS also revealed strong genetic relationships between GERD and a variety of traits.[12] For example, there was a strong genetic correlation between depression and GERD, likely reflecting the psychological component previously shown to be associated with GERD.[13] Another genetically correlated trait was obesity, an expected result given being overweight is a major risk factor for GERD.[14]

In this study, we took advantage of the genetic overlap between GERD and other traits[12] to improve our power for gene discovery on GERD and BE. Because some of the traits most strongly correlated to GERD have been subjected to GWAS on a very large scale, there is potential for a large increase in gene mapping power if a multitrait approach is taken. We, hence, employed a multitrait approach using GERD and BE, alongside three traits, which had large sample sizes and high correlation with GERD: obesity, depression and educational attainment (EDU). We then replicated the GERD and BE loci in a large replication cohort and conducted gene enrichment and transcriptome-wide association analyses. Finally, using BE as a clinical endpoint indicative

of pathological erosive chronic acid reflux (rather than forms of symptomatic reflux associated with psychological traits such as depression), we then assessed the relevance of subsets of the GERD loci to serious oesophageal disease.

METHODS

Detailed methods are provided in online supplemental materials.

Overview of methods

Leveraging the strong genetic correlation between GERD and related traits, we applied a multitrait GWAS model combining GWASs for body mass index (BMI), major depressive disorder (MDD), education attainment and GERD (and BE), to identify more susceptibility loci for GERD and BE (figure 1). The brief description for each input GWAS in the multitrait GWAS analysis and the equivalent effect sample size is shown in online supplemental table ST1. Candidate loci for GERD and BE achieving genome-wide significance (p<5e-8) were sent for replication in the independent 23andMe cohort (462 753 cases; 1 127 474 controls). Findings from these GWAS analyses were followed up with transcriptome-wide association studies (TWAS) and tissue enrichment analyses. TWAS analysis allows us to infer if there is a relationship between predicted gene-expression levels and GERD/BE risk. Tissue enrichment analyses allow us to assess whether the relevant GERD/BE-associated genes showed differential enrichment across 44 human tissues including oesophageal-related tissues. We finally applied a simple heuristic to dissect aetiological heterogeneity in GERD by separating GERD risk loci into obesity-driven and depression-driven categories; we then assessed these categories for differences in predicted gene expression in various tissues and for their ability to predict BE/EA susceptibility.

RESULTS

Multitrait analysis reveals 88 risk loci for GERD and 17 for BE with strong evidence of replication

In our multitrait GWAS combining GERD with BMI, MDD and EDU, we identified 88 GERD risk loci, where one or
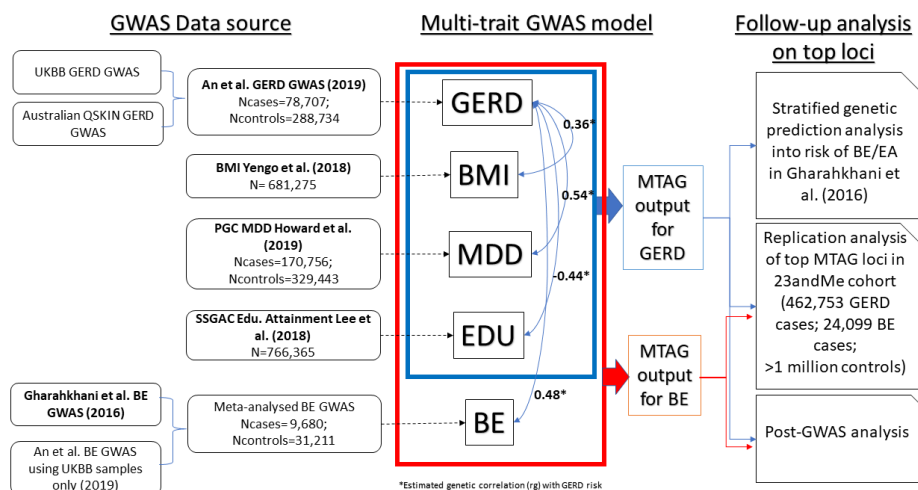
Figure 1 Schematic diagram describing overall study approach of the multi-trait GWAS analysis for GERD and BE susceptibility. GWAS data obtained from published studies are shown in bold. BE,Barrett's oesophagus; EA,esophageal adenocarcinoma; EDU, Education Attainment; GERD,gastro-oesophageal reflux disease; GWAS, genome-wide association study; MDD, major depressive disorder; MTAG,multitrait analysis of GWAS; PGC, Psychiatric Genomics Consortium; QSKIN, Queensland Sun and Health Study; SSGAC, Social Science Genetics Association Consortium; UKBB, UK Biobank. Traits within the blue/red boxes are traits selected for the multi-trait GWAS analysis for GERD (blue) and BE (red). The trait BE was not modelled in the MTAG model for GERD, to avoid sample overlap bias in the genetic prediction analysis for GERD into the BE/EA datasets. Asterisk (*) highlights genetic correlation estimates for each trait against GERD (shown by the blue arrows) obtained from previous An et al[12] findings.
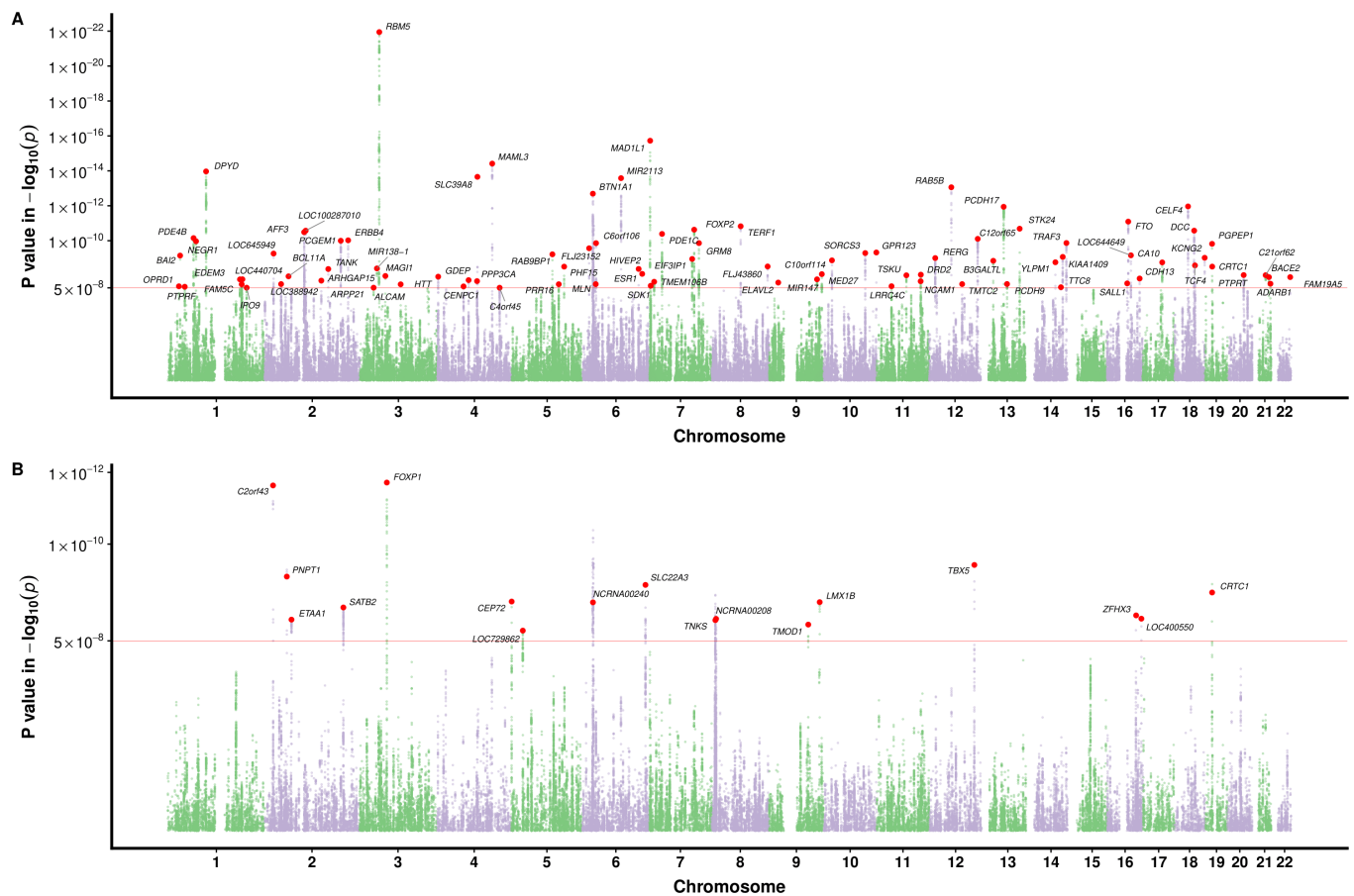
Figure 2 Manhattan plot for the multi-trait meta-analysed GWAS of GERD and BE. (A) for GERD; (B) for BE. The red dotted line indicates the genome-wide significance threshold at $p=5\times10^{-8}$. Single nucleotide polymorphisms (SNPs) with p values <0.01 were not shown. Loci with GERD association exceeding this threshold are deemed as a genome-wide GERD loci and followed up with the replication analysis in 23andMe. Each of the genome-wide loci is denoted with a red dot followed by the name of the relevant gene mapped through location proximity and expression quantiative trait loci (eQTL) information. BE,Barrett's oesophagus; GERD, gastro-oesophageal reflux disease; GWAS, genome-wide association study.

more single nucleotide polymorphisms (SNPs) exceeded genome-wide significance (5×10^{-8}). There was no evidence for inflation of the GWAS test statistic (LD-Score intercept <1, QQ-plot in online supplemental figure 1). The association estimates for each of the 88 lead SNPs are presented in online supplemental table ST2. To benchmark the power improvement resulting from our multitrait approach, we compared the average χ^2 test statistic for GERD from the multitrait GWAS (1.72) with that from the standard univariate GWAS (1.44); the increase corresponds to an estimated increase in GERD effective sample size of 64% (estimated via ratio of genomic inflation: $(1.72–1)/(1.44–1)$).[12]
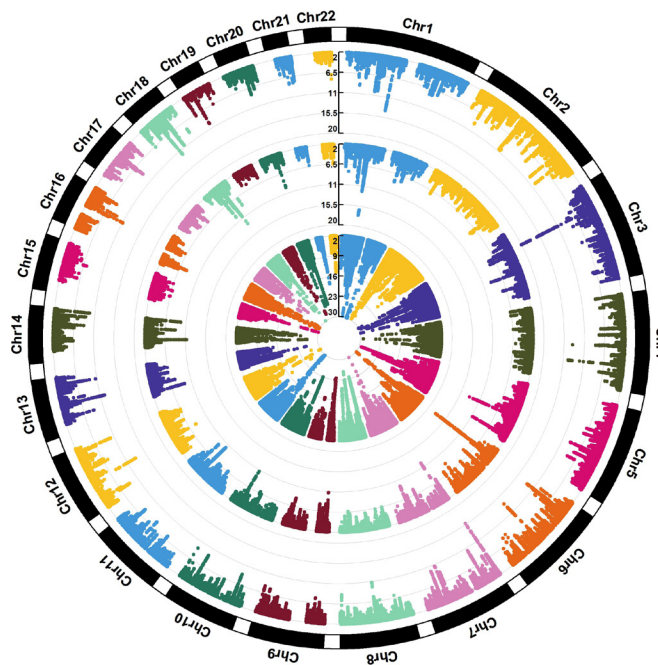
The GWAS Manhattan plot for the multitrait GERD GWAS along with the circular Manhattan plot showing the contribution of the BMI and depression associations to the multitrait GERD GWAS are shown in figures 2A and 3. The associations between independent genome-wide significant GERD loci in the multi-trait GWAS model (multitrait analysis of GWAS (MTAG)) against BMI, EDU and MDD are tabulated in online supplemental table ST3. The linkage disequilibrium among SNPs in each of the 88 GERD loci is displayed via locusZoom plots in online supplemental figure 2). The vast majority (78 of 88) of the GERD–MTAG associations replicated at an uncorrected p<0.05 level in the independent GERD validation sample from the 23andMe cohort, with 59 significant at p<0.05 after Bonferroni correction

(ie, $p<8.4\times10^{-4}$) for 88 tests (online supplemental table ST2). The effect sizes for GERD in the MTAG and 23andMe replication analyses were highly consistent ($r^2=0.85$) (figure 4A).

Our MTAG GWAS analysis focusing on BE discovered 17 BE susceptibility loci achieving genome-wide significance. Of these, 14 replicated at $p<2.9\times10^{-3}$ (0.05/17) in the independent 23andMe BE case–control cohort (online supplemental table ST4; QQ-plot in online supplemental figure 1). Relative to Gharahkhani et al,[15], seven of the loci were novel (rs2861695, rs10080150, rs10039754, rs622217, rs11792928, rs739414 and rs7187365). The effect sizes for BE in the MTAG and 23andMe replication analyses were highly consistent (figure 4B). Regional locusZoom plots for these novel loci are in online supplemental figure 3.

To assess whether our MTAG results were robust against issues of phenotypic heterogeneity, due to the use of self-report and inference through medication data that are diagnostically less reliable, we repeated our multitrait GWAS analyses using only GERD defined through the International Classification of Disease, tenth version (ICD-10) code (for GERD MTAG) and pathologically confirmed BE diagnosis (for BE MTAG). Differences in GERD and BE definition made no meaningful difference to the results, with strong correlations ($r^2>0.99$) observed between the SNP effect sizes of the original and revised MTAG models for both the GERD and BE analyses (online supplemental
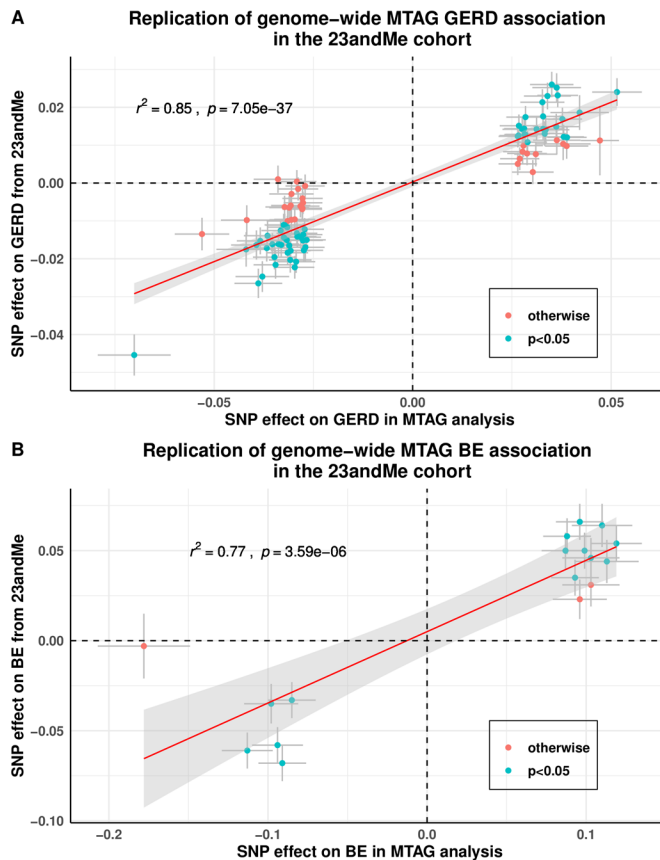
Figure 3 The circular Manhattan plot for BMI, MDD and the multitrait (MTAG) GERD GWAS. The trait from the innermost circle belongs to the (GIANT+UKB) Yengo *et al*. BMI GWAS[39], followed by the Howard *et al*. (excluding 23andMe) MDD GWAS[40] and the result from our multi-trait GERD GWAS. The plot illustrates different patterns of colocalisation between GERD and MDD/BMI for some regions, contributing to the genetic heterogeneity within GERD. BMI, body mass index; GERD, gastro-oesophageal reflux disease; GWAS, genome-wide association study; MDD, major depressive disorder; MTAG, multitrait analysis of GWAS.

Figure 4 Replication of GERD and BE association for the genome-wide significant loci from MTAG analysis on GERD and BE in the 23andMe cohort. A refers to the findings for GERD; (B) for BE. Data points that are shaded in blue are those that have a Bonferroni corrected p value below 0.05/88 for GERD and 0.05/17 for BE. Both the x- and y-axes represent log(OR) for GERD/BE and points are plotted with error bars representing one SE. Most of our GERD and BE loci showed strong evidence of being replicated, although the estimated effect size for GERD/BE in the 23andMe cohort were on average smaller than those estimated in the MTAG analysis (slope ~0.5 for both traits). BE, Barrett's oesophagus; GERD, gastro-oesophageal reflux disease; MTAG, multitrait analysis of genome-wide association studies, SNP, single nucleotide polymorphism.

figures 4 and 5). Moreover, results from the MTAG BE model excluding MDD and EDU were not meaningfully different from the BE+GERD+BMI+MDD+EDU model (online supplemental table ST5), suggesting that at least for these BE risk loci, the associated SNPs act on BE risk either via BMI or via pathways to GERD, which are not shown to be robustly related to traits such as MDD.

Transcriptome-wide association analyses reveal more than 200 genes associated with GERD and 49 with BE

Our metaXcan TWAS for GERD identified 37 significant genes when using expression quatitative trait loci (eQTL) information specifically derived from oesophageal muscularis tissue. Similarly, TWAS on gastro-oesophageal junction and oesophageal mucosa tissue identified 19 and 37 genes, respectively (online supplemental table ST6). As well as conducting TWAS on the three oesophageal tissues above, we also employed a TWAS approach using a much broader range of tissues types in multiXcan; this identified 212 significant genes after multiple testing corrections (see online supplemental table ST7).

A multiXcan TWAS for BE identified 49 significant genes using a weighted SNP-eQTL association across multiple tissues (online supplemental table ST8). For a TWAS focusing solely on each of the three oesophageal-related tissues, we identified fewer significant genes (<10 per tissue, online supplemental table ST9). Among the genes identified via TWAS, 31 genes were associated with both GERD and BE (online supplemental figure 6).

Tissue enrichment analysis shows enrichment of GERD-associated genes in brain tissues, but limited findings for BE genes

A tissue enrichment analysis based on the GERD genes mapped via the integrative web-based platform FUMA (available at http://fuma.ctglab.nl) for functional annotation revealed strong evidence for differential expression among brain tissues, but with no clear signal in oesophageal tissues (online supplemental figure 7). For BE-associated genes, moderate enrichment was seen in brain Cerebellar hemisphere, brain cerebellum, colon, bladder and oesophageal tissues (both in the gastro-oesophageal junction and the oesophagus muscularis), although the p values became nonsignificant after correction for multiple comparisons.

PheWAS analysis on novel hits for BE shows pleiotropic associations with other complex traits

As expected given our study design, and consistent with previous findings, the majority of GERD loci showed

pleiotropic associations with a variety of anthropometric traits, education level, behavioural traits (including smoking).[12] Several GERD genes were also associated with sarcoidosis, sleep duration and proxies of physical activity (online supplemental figure 8). A brief summary of our look-up for GERD loci driven by obesity and depression is in online supplemental table ST10.

For BE, again our study design leverages multiple input traits and, hence, pleiotropic associations are expected. PheWAS plots for the seven novel BE loci against traits available in the Open-Target platform are provided in online supplemental figure 9. The BE-associated SNP rs10080150 is associated with risk of ulcerative colitis[16], and rs7187365 and rs10039754 are associated with diaphragmatic hernia. Numerous BE SNPs (rs622217, rs10080150, rs11792928) showed pleiotropic associations with cardiovascular traits.

Genetic heterogeneity among the obesity-driven/depression-driven GERD subgroups showing differential patterns of enrichment on oesophageal tissues and genetic prediction onto BE/EA

Of the 88 genome-wide GERD loci, we categorised 27 as depression-driven, 46 as obesity-driven GERD loci and 15 as indeterminate/unclear based on their relative contribution to GERD in the multitrait model. We then sought to validate these categories using two separate approaches.

In our first approach, we compared the categories in a tissue enrichment-based analysis. The mapped genes for SNPs from each category using eQTL information and proximity

are shown in online supplemental tables ST11 and ST12.[17] While the regulation of gene expression was shown to be predominantly observed in brain tissues for the complete set of 88 genome-wide significant genetic loci, there were clear differences in the trend of gene expression across tissue types in the stratified obesity and depression driven subgroups (figure 5). Results for the directional (upward and downward separately) differential regulation for both subgroups are presented in online supplemental tables ST13 and ST14; directional regulation plots and gene expression heatmaps in online supplemental figures 10-13. Genes mapped in the depression GERD subgroup remain predominantly expressed in brain tissues. In contrast, for genes in the obesity GERD subgroup as well as there being expression in brain tissues, a clear signal can be seen in oesophageal tissues (online supplemental figure 10).

In our second approach, we compared the SNP effect sizes for GERD with those from an independent BE/EA GWAS. Based on the 88 GERD susceptibility loci from the MTAG analysis, each one unit (log(OR)) increase in risk of BE/EA was associated with a 0.6 ($p=4.2\times10^{-10}$) unit (log(OR)) increase in GERD susceptibility (figure 6; summary data for BE/EA for the 88 GERD SNPs available in online supplemental table ST15). However, the Cochran Q statistics for the derived overall estimate suggested substantial evidence for effect size heterogeneity among GERD SNPs (p heterogeneity <0.001), indicating the potential for multiple biological mechanisms in action (online supplemental table ST16). We, hence, followed up with an association analysis for BE/EA based on the two

Comparison of differential gene expression regulation (two-sided) across various human tissues for categorised GERD loci subgroups
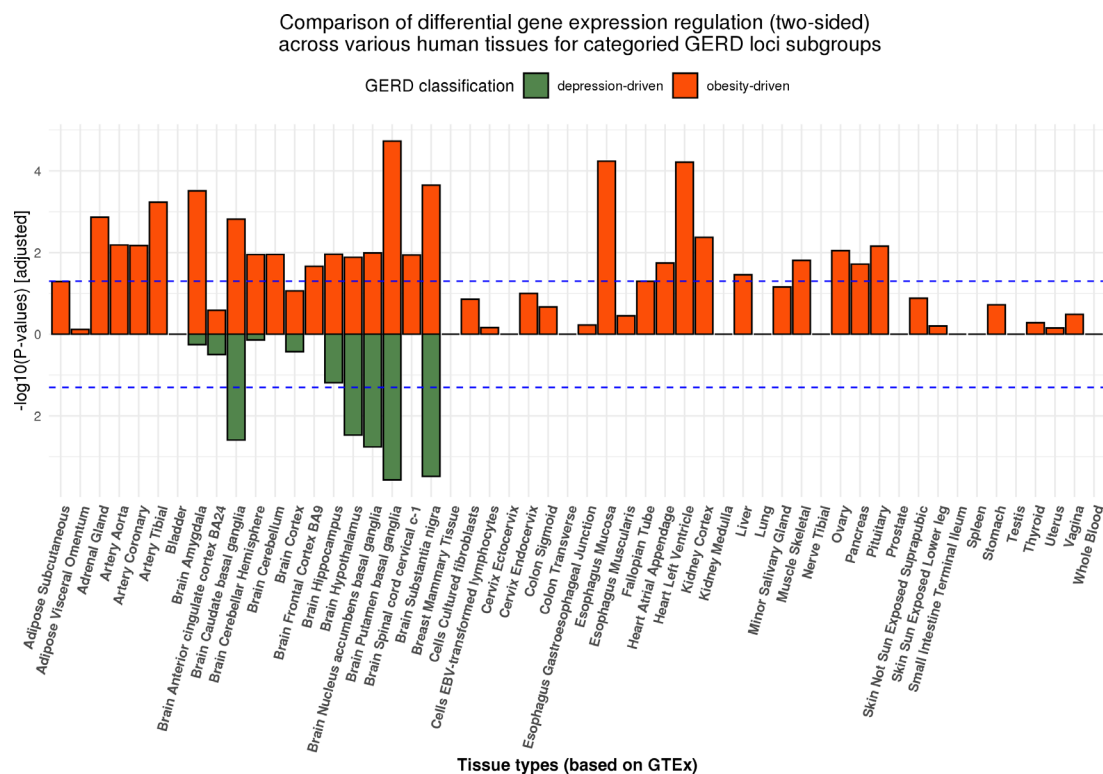
Figure 5 MAGMA gene-based tissue enrichment analysis for categorised functional and non-functional GERD gene sets on 53 human tissues. The dotted line represents the Bonferroni corrected significance threshold. While both the obesity-driven and neuropsychiatric-driven GERD genes were differentially expressed in brain tissues, only the regulation of gene expression in oesophageal tissues were detected from the obesity GERD gene set. Another observation is that the pattern of regulation is more pleiotropic across tissues for the obesity GERD set, consistent with the complex architecture of adiposity. The results for up-ward and down-ward regulation of gene expression can be viewed in online supplemental materials. GERD, gastro-oesophageal reflux disease

Genetic association between genome-wide significant MTAG GERD loci and Barrett Esophagus/Esophageal Adenocarcinoma combined
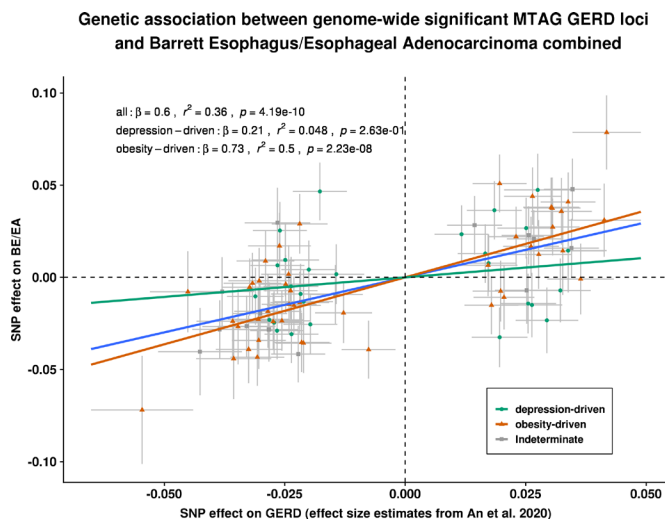
Figure 6 Estimation of the genetic association between GERD and BE/EA, stratified by genetic GERD subtypes. Each of the slopes represents the estimated magnitude of association between per unit increase in log(OR) of GERD on log(OR) of BE/OA, for all GERD loci (line in blue), obesity-driven GERD loci (line and points in green), and neuropsychiatric-driven GERD loci (line and points in orange) using an inverse variance weighted regression model. Cochran Q statistics indicated strong heterogeneity in the overall all-GERD estimate (p<0.001). In the stratified analysis, it becomes apparent that the overall estimate was primarily driven by the much stronger effect sizes from the obesity-driven GERD loci subset. BE, Barrett's oesophagus (or Barrett's esophagus); EA, oesophageal adenocarcinoma (or esophageal adenocarcinoma); GERD, gastro-oesophageal reflux disease; MTAG, multitraitanalysis of GWAS; SNP, single nucleotide polymorphism.

GERD subgroups. For GERD SNPs shown to be driven by obesity, the magnitude of association with BE/EA was 0.75 (p=$3.9×10^{-10}$) per unit log(OR) increase in GERD. In contrast, for GERD SNPs driven by neuropsychiatric factors correlated with depression, the magnitude of association with BE/EA was only 0.22 (p>0.05). The estimated effect size on BE/EA between the groups was substantially different (p value for difference=0.02). The inferences drawn for BE and EA taken separately were similar to the combined BE/EA findings (online supplemental figures 14 and 15). Importantly, adjusting for BMI did not completely attenuate the genetic relationship between obesity-driven GERD and BE/EA (online supplemental table ST17).

DISCUSSION

Leveraging the known genetic correlation between GERD diagnosis with depression, EDU and anthropometric traits, we identified 88 independent GERD loci, with more than two-thirds showing clear evidence of replication in the independent 23andMe cohort. Applying a similar multitrait model, we also discovered seven additional risk loci for BE that have not been previously reported. To the best of our knowledge, this is also the first study to have presented clear evidence for the existence of genetic subgroups within the complex and heterogeneous GERD (genetic) architecture.

Obesity and depression are known GERD risk factors; here, we first show that they can be combined to identify more loci for a broadly defined GERD phenotype, where we greatly expanded the number of GERD risk loci over and above previous GWAS findings[12] through a multitrait model. More than two-thirds of

the MTAG GERD loci replicated in 23andMe suggest that despite heterogeneous phenotyping being used, our findings represent genuine reflux loci. These results are also in good accord with our previous work showing high genetic concordance between these self-reported and clinical phenotypes. Second, we show that among the genome-wide significant GERD loci, a subgroup acts predominantly via effects on obesity while another acts via effects on depression. This is broadly consistent with previous observations that endoscopic findings[18] in many (38%) patients with typical GERD symptoms are normal; in such cases, psychological factors may drive symptom reporting.[19–23] We further showed that genes involved in the obesity-driven GERD subgroup were differentially expressed in oesophageal and adipose tissues—which was not evident from previous analyses evaluating GERD as a single, homogeneous disease.[12 15]

It was previously shown that psychosomatic factors are linked with functional oesophageal disorders through the gut-brain axis, which may alter a person's emotional status, perception of pain and influence the self-reporting of GERD symptoms and severity.[24 25] Previous studies have shown that among the broad class of individuals affected by GERD, those with functional oesophageal disorders typically have normal oesophageal acid exposure[26], and this may lead to them having lower risk of downstream BE/EA compared with other patients with GERD. Despite many studies showing strong association between depression with GERD (notably, the nonerosive reflux disease subtype),[27] studies evaluating psychosomatic factors and BE/EA susceptibility had been more limited,[28 29] likely due to sample size constraints. While there have been no large-scale observational studies to date to clarify any link between depression and BE, the previously documented weak correlation[30] between GERD symptom scores (which correlates strongly with presence of a functional oesophageal disorder) and severity of esophagitis, would imply that psychosomatic factors are unlikely to drive progression towards BE. Here, our results are in keeping with this as they show that the depression-driven GERD loci, which may predispose a person towards a functional oesophageal disorder rather than a disorder with abnormal oesophageal acid exposure, have no effect on BE/EA risk. In contrast, we show that GERD risk loci acting via the obesity-driven pathway do have an effect on BE/EA risk (figure 6).

PheWAS results reveal that these novel loci are enriched with anthropometric traits and cardiovascular-related risk factors (online supplemental figure 9), perhaps as a result of a pleiotropic effect on obesity and cardiovascular disease. Given our results showing very weak association between depression-driven GERD and BE/EA, we did not expect to see pleiotropy between the BE-associated SNPs and depression; while this was broadly true, one of the novel BE loci (rs2861695) showed moderate evidence of association with depression (p<6.7e-4).[31] Moreover, dropping proxies of socioeconomic status (EDU) and depression (MDD) from the model did not change our GWAS findings for BE (online supplemental table ST5). While our analyses showed a potential shared pathway between the obesity-driven GERD loci and BE, due to limited power, very few loci achieved genome-wide significance (ie, p<5e-8) across both traits. However, most of the BE loci did show suggestive evidence (ie, p<1e-5) of an association with GERD (online supplemental table ST18).

Two of the novel MTAG GERD loci, rs9636202 (nearest gene *PGPEP1*) and rs7206608 (nearest gene *CDH13*), were previously shown to be associated with periodontitis.[32 33] Apart from actual pH-monitoring, periodontitis is also reported as one of the risk factors that are more commonly observed among cases with erosive GERD.[34 35] However, whether the potential role of

PGPEP1 and *CDH13* on reflux esophagitis and poor oral health represents a common effect on (1) dietary behaviour or (2) changes in adiposity (resulting in genetic associations with BMI) remains unclear. Our gene-based PheWAS identified substantial overlap of GERD-associated genes with those involved in sarcoidosis, consistent with previous findings estimating one-third of patients with sarcoidosis being diagnosed with GERD-related symptoms and irritable bowel syndrome.[36] Although fine-mapping is required to unambiguously link SNPs to genes, genes associated with several of the loci identified here (*CRTC1*, *CDK2*, *PDE4B*, *DPYD*, *PDE1C*) are targeted by drugs that are currently on trial for oesophageal and digestive system cancers.[12] Despite the fact that we have not undertaken the substantial task of unambiguously fine-mapping each associated GERD SNP to the relevant gene(s), many of the genes we identify here may be promising drug targets for further research.

Our observation of a greater magnitude of association between obesity-driven GERD loci and BE/EA is not surprising, given prior evidence[37] showing that genetically predicted obesity was associated with increased risk of BE and EA. To evaluate whether obesity is a complete mediator between GERD and BE, we examined the genetic effect of obesity-driven GERD on BE adjusting for BMI and observed only partial attenuation (IVW-log(OR) from 0.726 (p=2.23e-8) to 0.513 (p=0.001)) of the GERD effect on BE (online supplemental table ST17). This implies that the relationship between GERD and BE cannot be solely explained by a common direct effect of obesity. Formal Mendelian randomisation (MR) analyses have been attempted previously[38]; however, interpreting the results from such studies is difficult given we have shown that there are at least two pathways leading to GERD (one acting via obesity associated SNPs, another acting via depression SNPs).

This study has several limitations. First, a proportion of UKB GERD cases was determined through use of heartburn-related medication[12]; however, as noted in our previous study, the genetic correlation between GERD inferred through medication, self-report and clinical diagnosis was very high. Similarly, our validation cohort (23andMe) adopted a broad GERD definition and diagnosis of BE were self-reported as opposed to being histologically verified. Overall, our approach of using a large GWAS sample size with a broad phenotype was much more effective in identifying novel loci than previous smaller studies using more well-defined phenotypes.[10 11] Our heuristic for making the GERD obesity versus depression-driven subgroup classifications is limited by the accuracy of GWAS data, focused on only 88 SNPs, and, for ease of implementation, we used the BMI and depression GWAS results to group SNPs. Furthermore, while the combined effect of the obesity-driven GERD genome-wide significant loci (assembled in a genetic risk score) was robustly associated with BE/EA risk, the predictions from this genetic risk score lack precision to reliably predict BE/EA in clinical settings. Finally, some SNPs may have pleiotropic effects on both obesity and on psychiatric traits or could simply reflect other pleiotropic pathways not mediated by these factors (eg, mucosal tight junctions, inflammation and repair, oesophageal and gastric motility). Modelling these risk factors which are independent of the obesity and depression axis in our multitrait model could conceivably increase power to detect additional loci for both GERD and BE. GWAS studies on these additional traits are currently limited in size but investigating this will be worthwhile in future work.

In conclusion, using a multitrait framework, we greatly expanded the number of genome-wide significant genetic loci for GERD and also increased the number of BE loci by seven. Here, we present strong evidence for a genetic underpinning of disease heterogeneity in GERD, where we showed that GERD loci associated with depressive symptoms are not strong predictors of BE/EA relative to obesity-driven GERD loci.

Author affiliations
¹Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia
²School of Biology & Environmental Science, Queensland University of Technology, Brisbane, Queensland, Australia
³Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Queensland, Australia
⁴23andMe, Mountain View, California, USA
⁵Institute of Human Genetics, Philipps University of Marburg, Marburg, Germany
⁶Department of Visceral, Transplant, Thoracic and Vascular Surgery, University Hospital Leipzig, Leipzig, Germany
⁷Institute of Human Genetics, University of Bonn, Bonn, Germany
⁸Centre for Medicine and Health Sciences, University of United Arab Emirates, Al Ain, Abu Dhabi, UAE
⁹UCL Medical School, University College London, London, UK
¹⁰Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK
¹¹Department of Population Health, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia
¹²Faculty of Medicine, The University of Queensland, Herston, Queensland, Australia
¹³MRC Cancer Center, University of Cambridge, Cambridge, UK
¹⁴Department of Medicine, and Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, Texas, USA
¹⁵Department of Epidemiology, University of Washington School of Public Health, Seattle, Washington, USA
¹⁶Roswell Park Comprehensive Cancer Center, Buffalo, New York, USA
¹⁷Department of Gastroenterology and Hepatology, Princess Alexandra Hospital, Woolloongabba, Queensland, Australia

ORCID iDs
Jue-Sheng Ong http://orcid.org/0000-0002-6062-710X
Aaron P Thrift http://orcid.org/0000-0002-0084-5308
Bradley J Kendall http://orcid.org/0000-0002-6471-2918
Stuart MacGregor http://orcid.org/0000-0001-6731-8142

REFERENCES

1 Vakil N, van Zanten SV, Kahrilas P, et al. The Montreal definition and classification of gastroesophageal reflux disease: a global evidence-based consensus. *Am J Gastroenterol* 2006;101:1900–20.
2 Aziz Q, Fass R, Gyawali CP, et al. Functional esophageal disorders. *Gastroenterology* 2016. doi:10.1053/j.gastro.2016.02.012. [Epub ahead of print: 15 Feb 2016].
3 Gyawali CP, Kahrilas PJ, Savarino E, et al. Modern diagnosis of GERD: the Lyon consensus. *Gut* 2018;67:1351–62.
4 Katzka DA, Pandolfino JE, Kahrilas PJ. Phenotypes of gastroesophageal reflux disease: where Rome, Lyon, and Montreal meet. *Clin Gastroenterol Hepatol* 2020;18:767–76.
5 Locke GR, Talley NJ, Fett SL, et al. Risk factors associated with symptoms of gastroesophageal reflux. *Am J Med* 1999;106:642–9.
6 Argyrou A, Legaki E, Koutserimpas C, et al. Risk factors for gastroesophageal reflux disease and analysis of genetic contributors. *World J Clin Cases* 2018;6:176–82.
7 Gawron AJ, French DD, Pandolfino JE, et al. Economic evaluations of gastroesophageal reflux disease medical management. *Pharmacoeconomics* 2014;32:745–58.
8 Whiteman DC, Kendall BJ. Barrett's oesophagus: epidemiology, diagnosis and clinical management. *Med J Aust* 2016;205:317–24.
9 Tan MC, Mansour N, White DL, et al. Systematic review with meta-analysis: prevalence of prior and concurrent Barrett's oesophagus in oesophageal adenocarcinoma patients. *Aliment Pharmacol Ther* 2020;52:20–36.
10 Bonfiglio F, Hysi PG, Ek W, et al. A meta-analysis of reflux genome-wide association studies in 6750 Northern Europeans from the general population. *Neurogastroenterol Motil* 2017;29. doi:
11 Gharakhani P, Tung J, Hinds D, et al. Chronic gastroesophageal reflux disease shares genetic background with esophageal adenocarcinoma and Barrett's esophagus. *Hum Mol Genet* 2016;25:828–35.
12 An J, Gharakhani P, Law MH, et al. Gastroesophageal reflux GWAS identifies risk loci that also associate with subsequent severe esophageal diseases. *Nat Commun* 2019;10:4219.
13 Kim SY, Kim H-J, Lim H, et al. Bidirectional association between gastroesophageal reflux disease and depression: two different nested case-control studies using a national sample cohort. *Sci Rep* 2018;8:11748.
14 Chang P, Friedenberg F. Obesity and GERD. *Gastroenterol Clin North Am* 2014;43:161–73.
15 Gharakhani P, Fitzgerald RC, Vaughan TL, et al. Genome-wide association studies in oesophageal adenocarcinoma and Barrett's oesophagus: a large-scale meta-analysis. *Lancet Oncol* 2016;17:1363–73.
16 Jostins L, Ripke S, Weersma RK, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119–24.
17 GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580–5.
18 Dent J, Vakil N, Jones R, et al. Accuracy of the diagnosis of GORD by questionnaire, physicians and a trial of proton pump inhibitor treatment: the diamond study. *Gut* 2010;59:714–21.
19 Willert RP, Delaney C, Kelly K, et al. Exploring the neurophysiological basis of chest wall allodynia induced by experimental oesophageal acidification - evidence of central sensitization. *Neurogastroenterol Motil* 2007;19:270–8.
20 Kessing BF, Bredenoord AJ, Saleh CMG, et al. Effects of anxiety and depression in patients with gastroesophageal reflux disease. *Clin Gastroenterol Hepatol* 2015;13:1089–95.
21 On ZX, Grant J, Shi Z, et al. The association between gastroesophageal reflux disease with sleep quality, depression, and anxiety in a cohort study of Australian men. *J Gastroenterol Hepatol* 2017;32:1170–7.
22 Fass R, Naliboff BD, Fass SS, et al. The effect of auditory stress on perception of intraesophageal acid in patients with gastroesophageal reflux disease. *Gastroenterology* 2008;134:696–705.
23 Patel D, Fass R, Vaezi M. Untangling Nonerosive reflux disease from functional heartburn. *Clin Gastroenterol Hepatol* 2020;30434–1.
24 Van Oudenhove L, Crowell MD, Drossman DA, et al. Biopsychosocial aspects of functional gastrointestinal disorders. *Gastroenterology* 2016;00218–3.
25 Johnston BT, Lewis SA, Collins JS, et al. Acid perception in gastro-oesophageal reflux disease is dependent on psychosocial factors. *Scand J Gastroenterol* 1995;30:1–5.
26 Chen C-L, Hsu P-I. Current advances in the diagnosis and treatment of nonerosive reflux disease. *Gastroenterol Res Pract* 2013;2013:1–8.
27 Yang X-J, Jiang H-M, Hou X-H, et al. Anxiety and depression in patients with gastroesophageal reflux disease and their effect on quality of life. *World J Gastroenterol* 2015;21:4302–9.
28 Sharma N, Ho KY. Risk factors for Barrett's oesophagus. *Gastrointest Tumors* 2016;3:103–8.
29 Hartono JL, Mahadeva S, Goh K-L. Anxiety and depression in various functional gastrointestinal disorders: do differences exist? *J Dig Dis* 2012;13:252–7.
30 Sharma VK. Role of endoscopy in GERD. *Gastroenterol Clin North Am* 2014;43:39–46.
31 Nagel M, Jansen PR, Stringer S, et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet* 2018;50:920–7.

32 da Silva MK, de Carvalho ACG, Alves EHP, *et al*. Genetic factors and the risk of periodontitis development: findings from a systematic review composed of 13 studies of meta-analysis with 71,531 participants. *Int J Dent* 2017;2017:1–9.

33 Shungin D, Haworth S, Divaris K, *et al*. Genome-wide analysis of dental caries and periodontitis combining clinical and self-reported data. *Nat Commun* 2019;10:1–13.

34 Watanabe M, Nakatani E, Yoshikawa H, *et al*. Oral soft tissue disorders are associated with gastroesophageal reflux disease: retrospective study. *BMC Gastroenterol* 2017;17:92.

35 Song JY, Kim HH, Cho EJ, *et al*. The relationship between gastroesophageal reflux disease and chronic periodontitis. *Gut Liver* 2014;8:35–40.

36 Reynolds HY. Sarcoidosis: impact of other illnesses on the presentation and management of multi-organ disease. *Lung* 2002;180:281–99.

37 Thrift AP, Shaheen NJ, Gammon MD, *et al*. Obesity and risk of esophageal adenocarcinoma and Barrett's esophagus: a Mendelian randomization study. *J Natl Cancer Inst* 2014;106:dju252.

38 Wu Y, Murray GK, Byrne EM, *et al*. GWAS of peptic ulcer disease implicates Helicobacter pylori infection, other gastrointestinal disorders and depression. *Nat Commun* 2021;12:1146.

39 Yengo L, Sidorenko J, Kemper KE, *et al*. Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum Mol Genet* 2018;27:3641–9.

40 Howard DM, Adams MJ, Shirali M, *et al*. Genome-Wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat Commun* 2018;9:1470.