Investigating Causal Relationships between Major Depression and Chronic Pain using UK General-Population Datasets with Whole-Genome Genotyping

Keira Jacqueline Ann Johnston

MSc, BSc (Hons)

Submitted in fulfilment of the requirements for the

Degree of Doctor of Philosophy

University of Glasgow

Institute of Health and Wellbeing

College of Medical, Veterinary and Life Sciences

University of Glasgow

May 2021

Abstract

Chronic pain, considered here to be pain lasting 3 months or longer, imparts significant socioeconomic and public health burden around the globe. Chronic pain is associated with a wide range of conditions, illnesses, or injuries, and is categorised and investigated in many ways. Treatment and management of chronic pain is complicated by this heterogeneity, and by lack of full understanding of factors (including genetic) that influence vulnerability to developing chronic pain and biological mechanisms of chronic pain development. Major depression is commonly comorbid with chronic pain, and results of studies into potential causal direction between the two conditions are mixed. Due to symptom overlap and common comorbidity, it may be that cases of chronic pain are misclassified as major depression and vice versa. Understanding genetic factors that contribute to chronic pain vulnerability and development has the potential to improve treatment of both conditions, in addition to allowing for investigation of potential causal relationships and clinical heterogeneity.

Recently, the International Association for the Study of Pain released an updated definition of chronic pain and advocated for the study of chronic pain as a disease entity. Studying the genetics of chronic pain through genome wide association study of broad chronic pain traits, in line with this updated pain definition, may present a more tractable way to uncover common genetic variation associated with vulnerability to and mechanisms of development of chronic pain. This mode of study can also provide genome wide association study summary statistics for use in analyses that aim to investigate causality, genetic correlation and pleiotropy, and clinical heterogeneity in chronic pain and major depression.

The overall aim of this PhD project is therefore to explore causal relationships between chronic pain and MDD in large UK general-population cohorts with whole-genome genotyping data using a wide range of statistical genetic methods.

Data were obtained from two large UK cohorts with whole-genome genotyping. One, UK Biobank, is a cohort of 0.5 million participants recruited in middle age (40-79) with information on an extensive list of physical, behavioural and health related traits. Generation Scotland is a smaller (N ~ 22,000) Scottish cohort of participants recruited mainly through general practitioners in a family-based

manner, again with information of physical, health, and behavioural traits. Summary statistic data were also obtained from a 23andMe-Pfizer genome wide association study of chronic pain grade.

As part of this PhD the largest genome wide association study of any chronic pain trait to date was carried out in UK Biobank. Validation of the trait (multisite chronic pain) was carried out through polygenic risk score analysis in Generation Scotland, examining the relationship between this novel chronic pain trait and chronic pain grade. Genetic correlation analyses were used to explore the genetic overlap of multisite chronic pain and a range of traits of interest, including other chronic pain phenotypes such as chronic widespread pain and chronic pain grade, in addition to major depression. Gene-level analyses were carried out to investigate genes of interest associated with chronic pain and potentially relevant to mechanisms of chronic pain development. BUHMBOX analyses were performed to test for clinical heterogeneity in chronic pain with respect to major depression and vice versa in UK Biobank. Conditional false discovery rate analyses using 23andMe-Pfizer data were also used to explore pleiotropy in chronic pain grade and major depression and to highlight pleiotropic loci of interest. Mendelian randomisation analyses, including recent mendelian randomisation methods explicitly designed to account for extensive horizontal pleiotropy, were carried out to assess potential causal relationships between major depression and chronic pain grade, and between major depression and multisite chronic pain.

Results indicated multisite chronic pain was a polygenic, moderately heritable trait. Associated genes of interest implicated a strong central nervous system component, in addition to immune related genes. Conditional false discovery rate analysis highlighted loci of interest mapped to *LRFN5*, a gene involved in neuroinflammation, and that were associated with regulation of gene expression at this locus. Polygenic risk scoring analysis showed multisite chronic pain to be significantly associated with both chronic pain grade and chronic widespread pain, in addition to a multisite chronic pain-like trait in Generation Scotland, validating multisite chronic pain as a trait and indicating strong genetic overlap between widespread and non-widespread pain. Genetic correlation analysis showed significant genetic overlap between multisite chronic pain and mental health traits, markedly major depressive disorder, and depressive symptoms, but

a lower degree of genetic correlation with conditions associated with significant chronic pain such as rheumatoid arthritis, and no significant genetic correlation with inflammatory bowel diseases. BUHMBOX analyses showed no evidence of clinical heterogeneity in chronic pain with respect to major depression in UK Biobank or vice versa. Mendelian randomisation analyses showed no causal relationship between chronic pain grade and major depressive disorder, but a significant causal effect of multisite chronic pain on major depressive disorder.

In conclusion, I have shown that broad chronic pain traits such as multisite chronic pain present a powerful and tractable way to study mechanisms of, and factors contributing to vulnerability to, chronic pain development. Output from well-powered genome wide association studies can also be used to validate phenotypes, explore genetic overlap with traits of interest, and conduct causal analyses.

Table of Contents

Abstract 2
List of Tables 9
List of Figures 11
List of Equations 12
Publications 13
Additional Papers 14
Acknowledgements 16
Author's Declaration 17
Abbreviations 18
Chapter 1: Introduction 22
1.1 What is chronic pain? 22
1.1.1 Definitions 22
1.1.2 Measurement (Phenotyping) 23
1.1.3 Epidemiology of chronic pain 26
1.1.4 From Acute to Chronic Pain 27
1.1.5 Associations with Other Conditions 31
1.2 What is Major Depressive Disorder (MDD)? 35
1.2.1 Screening and Diagnosis of MDD 35
1.2.2 Epidemiology of MDD 37
1.3 Overlap between MDD and Chronic Pain 38
1.3.1 Comorbidity between MDD and Chronic Pain 38
1.3.2 Causal Relationships between MDD and Chronic Pain 39
1.3.3 Genetics of Complex Traits 41
1.3.4 Genetics of Chronic Pain and Chronic Pain Disorders 46
1.3.5 Genetics of MDD 49
1.4 Summary 50
1.5 Aims and Objectives 51
1.5.1 Overall Aim 51
1.5.2 Objectives 52
Chapter 2: Methodologies and Technical Information 53
2.1 Introduction 53
2.2 Methodologies 53
2.2.1 Genome-Wide Association Studies 53
2.2.2 Multiple-testing correction in a GWAS context 60

2.2.3 Conditional False Discovery Rate Analyses 60
2.2.4 BUHMBOX 62
2.2.5 Linkage-Disequilibrium Score Regression 64
2.2.6 Polygenic Risk Scoring 68
2.2.7 Mendelian Randomisation 69
2.3 Resources and Materials 80
2.3.1 FUMA and analyses therein (MAGMA, GTEx) 80
2.3.2 Cohort Profiles 84
2.3.3 Chronic Pain Phenotyping in Key Cohorts 87
2.3.4 Major Depression Phenotyping in UK Biobank 90
Chapter 3: Further Understanding Overlap of Chronic Pain and Depression: Pleiotropy and Clinical Heterogeneity 92
3.1 Introduction 92
3.2 Methods 93
3.2.1 Conditional False-Discovery Analysis of Chronic Pain Grade and Major Depressive Disorder 93
3.2.2 Further understanding the overlap of MDD and Chronic Pain 94
3.2.3 Clinical Heterogeneity in MDD and Chronic Pain 94
3.3 Results 97
3.3.1 cFDR: SNPs Associated with CPG and MDD 97
3.3.2 cFDR: Genomic Context of Trait-Associated SNPs 98
3.3.3 BUHMBOX: Whole-Group Pleiotropy in MDD and Chronic Pain in UK Biobank 102
3.3.5 Pleiotropic SNPs in *LRFN5* 103
3.4 Discussion 103
3.4.1 Pleiotropic Loci 103
3.4.2 Whole-group pleiotropy in MDD and chronic pain 104
Chapter 4 Common Genetic Variation Associated with Chronic Pain and Shared with Phenotypes of Interest 107
4.1 Introduction 107
4.2 Methods 108
4.2.1 Chronic Pain Phenotyping in UK Biobank 108
4.2.2 Genome-Wide Association Study of Multisite Chronic Pain 108
4.2.3 Linkage-Disequilibrium Score Regression 109
4.2.4 Phenotypic Correlations 110
4.3 Results 111
4.3.1 Description of Participants 111
4.3.2 Common genetic variants and genes associated with MCP 112

4.3.3 Genetic Correlations 114
4.4 Discussion 119
4.4.1 Genetic Correlations with MCP and Traits of Interest 119
4.4.2 Heritability and Polygenicity of Multisite Chronic Pain 125
4.4.3 Genes of Interest Associated with MCP 125
Chapter 5 Validation of Multisite Chronic Pain Phenotype 132
5.1 Introduction 132
5.2 Methods 132
5.2.1 Chronic Pain Phenotyping in Generation Scotland and UKB 132
5.2.2 Validation of MCP Polygenic Risk Score in Generation Scotland 134
5.2.3 Multisite Chronic Pain and Chronic Widespread Pain in UK Biobank: PRS Analysis 136
5.3 Results 136
5.3.1 MCP PRS Validation in Generation Scotland 136
5.3.2 Sex-Specific Associations between PRS and MCP in Generation Scotland 138
5.3.3 MCP and Chronic Widespread Pain in UK Biobank 140
5.4 Discussion 141
5.4.1 Multisite Chronic Pain and Chronic Widespread Pain 141
5.4.2 Validation of MCP PRS in an Independent Cohort 141
5.4.3 Sex Differences in PRS Associations 142
Chapter 6: Using Genetics to Assess Causal Relationships in Pain and MDD 144
6.1 Introduction 144
6.2 Methods 145
6.2.1 MR: Causal Relationship between Chronic Pain Grade and MDD 145
6.2.2 MR: Causal relationships between Multisite Chronic Pain and MDD 146
6.3 Results 147
6.3.1 Causal Relationships between Chronic Pain Grade and MDD 147
6.3.2 Causal Relationships between Multisite Chronic Pain and MDD 148
6.4 Discussion 152
6.4.1 Causal relationship between MCP and MDD 152
Chapter 7: General Discussion 155
7.1 History of Pain Theories 156
7.2 Evolutionary Perspectives of Pain 159
7.3 Multisite Chronic Pain in UK Biobank 160
7.3.1 Comparing MCP and Other Chronic Pain Phenotypes 160
7.3.2 Broad MDD Phenotyping Parallels 162

7.5 Causal Effect of Chronic Pain on MDD 162
7.6 Strengths & Limitations 163
7.6 Future Directions 165
7.6.1 Representative Cohorts 165
7.6.2 New Pain Data for UK Biobank 166
7.6.3 Alternative Approaches to Pleiotropy 167
7.6.4 Whole-Exome Data and Chronic Pain 168
7.6.5 Genomic Structural Equation Modelling Approaches 169
7.6.4 Affective Dysregulation and Pain 171
7.7 Overall Conclusions 172
Appendix 1: Genes Associated with Multisite Chronic Pain 174
Appendix 2: Phenotypic Correlation between Multisite Chronic Pain and Chronic Pain Grade in Generation Scotland 186
Appendix 3: Genetic Correlation between Tsepilov et al Phenotype GIP1 (Genetically Independent Phenotype 1) and Multisite Chronic Pain 188
References 190

List of Tables

Table 3. 1: Loci identified from cFDR analysis. 97
Table 3. 2: Output of 'rsnps' query 98
Table 3. 3: UCSC Genome Browser Results. 99
Table 3. 4: IGV eQTL Browser results. 101
Table 3. 5: BUHMBOX results for test of clinical heterogeneity in MDD cases in UK Biobank. 103
Table 3. 6: BUHMBOX results for test of clinical heterogeneity in chronic pain cases in UK Biobank. 103

Table 4. 1: Age, sex and MCP phenotype value of UK Biobank participants included in the MCP GWAS. 111
Table 4. 2: Genomic Risk Loci. 113
Table 4. 3: Genetic correlation results. 117
Table 4. 4 : Phenotypic correlations between MCP and traits of interest 118

Table 5. 1: Pain site options in Generation Scotland versus UK Biobank 133
Table 5. 2: Age and sex of participants included in CPG regression analyses... 134
Table 5. 3: Age and sex of participants included in MCP regression analyses. . 134
Table 5. 4: Summary of regression models with MCP as outcome. 135
Table 5. 5: Summary of regression models and sample sizes. 136
Table 5. 6: Results of the regression of MCP polygenic risk score on MCP in Generation Scotland, adjusted for age, sex and multidimensional scaling components 1-4.. 137
Table 5. 7: Results of the regression of MCP polygenic risk score on chronic pain grade in Generation Scotland, adjusted for age, sex and multidimensional scaling components 1-4. 137
Table 5. 8: Results for the regression of MCP polygenic risk score on MCP in Generation Scotland with inclusion of an interaction term (sex x PRS). 138
Table 5. 9: Results for the regression of MCP polygenic risk score on MCP in Generation Scotland in males only. 139
Table 5. 10: Results for the regression of MCP polygenic risk score on MCP in Generation Scotland in females only. 140
Table 5. 11 Summary of all four model key results. 140
Table 5. 12: Results of the regression of MCP polygenic risk score on chronic widespread pain in UK Biobank. 141

Table 6. 1: MR results with chronic pain grade as the exposure and MDD as the outcome across all three methods. 147
Table 6. 2: MR results with MDD as the exposure and chronic pain grade as the outcome across all three methods. 148
Table 6. 3: MR results for MR-RAPS analysis with MDD as the exposure and MCP as the outcome. 149

Table 6. 4: MR results for MR-RAPS analysis with MCP as the exposure and MDD as the outcome. 151

Table A1. 1: Genes found to be significantly ($p < 2.67 \times 10^{-6}$) associated with MCP in MAGMA gene-level analyses..... 180
Table A1. 2: Function of genes associated with MCP..... 185
Table A1. 3: MAGMA gene set analysis results (for curated gene sets i.e., MSigDB C2). 185

Table A2. 1: Phenotypic correlation between CPG and MCP. 187

Table A3. 1: Genetic correlation results. 188

List of Figures

Figure 1. 1: Pleiotropy. 44

Figure 2. 1: MR as a natural randomised control trial. 70
Figure 2. 2: MR assumptions. 71

Figure 3. 1. Single tissue eQTL lookups of rs11846556 102

Figure 4. 1 : MCP GWAS Manhattan plot 114
Figure 4. 2: Genetic correlations between MCP and a range of traits. 119
Figure 4. 3: Sources of bias in GWAS. 124

Figure 6. 1. Diagnostic plots of MR-RAPS analysis with MDD as exposure 150
Figure 6. 2: Diagnostic plots of MR-RAPS analysis with MCP as exposure and MDD as the outcome. 152

List of Equations

Equation 2. 1: Broad-sense heritability. 55
Equation 2. 2: Narrow-sense heritability. 55
Equation 2. 3: Proportion of variance explained. 57
Equation 2. 4: Conditional false discovery rate. 61
Equation 2. 5: Y matrix for BUHMBOX calculations. 63
Equation 2. 6: BUHMBOX test statistic. 63
Equation 2. 7: P value for BUHMBOX test statistic. 63
Equation 2. 8: Linkage disequilibrium estimate (r). 64
Equation 2. 9: Recombination rate. 65
Equation 2. 10: Amount of genetic variation tagged by variant j. 66
Equation 2. 11: Expected value for GWAS test statistic associated with variant j. 67
Equation 2. 12: Expected value of cross-trait product of GWAS z scores. 67
Equation 2. 13: Errors-in-variables regression. 76
Equation 2. 14: Log-likelihood function of the summary data. 76
Equation 2. 15: Profile score. 77
Equation 2. 16: Regression of phenotype Y on gene effects (gene-level MAGMA analysis). 81
Equation 2. 17: Transformation of gene p values to Z values for gene set analysis. 82
Equation 2. 18: Intercept-only regression (MAGMA gene set analysis, self-contained) 83
Equation 2. 19: MAGMA gene set analysis (competitive) 83

Equation 3. 1: Conditional false discovery rate. 94

Publications

Johnston, K. J. A., Adams, M. J., Nicholl, B. I., Ward, J., Strawbridge, R. J., Ferguson, A., Mcintosh, A. M., Bailey, M. E. S., & Smith, D. J. (2019). Genome-wide association study of multisite chronic pain in UK Biobank. PLoS Genetics, 15(6), 1–22. https://doi.org/10.1371/journal.pgen.1008164

Johnston, K. J. A., Adams, M. J., Nicholl, B. I., Ward, J., Strawbridge, R. J., McIntosh, A. M., Smith, D. J., & Bailey, M. E. S. (2019). Identification of novel common variants associated with chronic pain using conditional false discovery rate analysis with major depressive disorder and assessment of pleiotropic effects of LRFN5. Translational Psychiatry, 9(1). https://doi.org/10.1038/s41398-019-0613-4

Johnston, K. J. A., Ward, J., Ray, P. R., Adams, M. J., McIntosh, A. M., Smith, B. H., Strawbridge, R. J., Price, T. J., Smith, D. J., Nicholl, B. I., & Bailey, M. E. S. (2021). Sex-stratified genome-wide association study of multisite chronic pain in UK Biobank. PLOS Genetics, 17(4), e1009428. https://doi.org/10.1371/journal.pgen.1009428

Additional Papers

Ferguson, A., Lyall, L. M., Ward, J., Strawbridge, R. J., Cullen, B., Graham, N., Niedzwiedz, C. L., Johnston, K. J. A., MacKay, D., Biello, S. M., Pell, J. P., Cavanagh, J., McIntosh, A. M., Doherty, A., Bailey, M. E. S., Lyall, D. M., Wyse, C. A., & Smith, D. J. (2018). Genome-Wide Association Study of Circadian Rhythmicity in 71,500 UK Biobank Participants and Polygenic Association with Mood Instability. EBioMedicine, 35, 279–287. https://doi.org/10.1016/j.ebiom.2018.08.004

Morris, J., Bailey, M. E. S., Baldassarre, D., Cullen, B., de Faire, U., Ferguson, A., Gigante, B., Giral, P., Goel, A., Graham, N., Hamsten, A., Humphries, S. E., Johnston, K. J. A., Lyall, D. M., Lyall, L. M., Sennblad, B., Silveira, A., Smit, A. J., Tremoli, E., … Strawbridge, R. J. (2019). Genetic variation in CADM2 as a link between psychological traits and obesity. Scientific Reports, 9(1), 1–14. https://doi.org/10.1038/s41598-019-43861-9

Morris, J., Leung, S. S. Y., Bailey, M. E. S., Cullen, B., Ferguson, A., Graham, N., Johnston, K. J. A., Lyall, D. M., Lyall, L. M., Ward, J., Smith, D. J., & Strawbridge, R. J. (2020). Exploring the role of contactins across psychological, psychiatric and cardiometabolic traits within uk biobank. Genes, 11(11), 1–17. https://doi.org/10.3390/genes11111326

Strawbridge, R. J., Johnston, K. J. A., Bailey, M. E. S., Baldassarre, D., Cullen, B., Eriksson, P., deFaire, U., Ferguson, A., Gigante, B., Giral, P., Graham, N., Hamsten, A., Humphries, S. E., Kurl, S., Lyall, D. M., Lyall, L. M., Pell, J. P., Pirro, M., Savonen, K., … Smith, D. J. (2021). The overlap of genetic susceptibility to schizophrenia and cardiometabolic disease can be used to identify metabolically different groups of individuals. Scientific Reports, 11(1), 1–13. https://doi.org/10.1038/s41598-020-79964-x

Strawbridge, R. J., Ward, J., Bailey, M. E. S., Cullen, B., Ferguson, A., Graham, N., Johnston, K. J. A., Lyall, L. M., Pearsall, R., Pell, J., Shaw, R. J., Tank, R., Lyall, D. M., & Smith, D. J. (2020). Carotid intima-media thickness novel loci, sex-specific effects, and genetic correlations with obesity and glucometabolic

traits in UK Biobank. Arteriosclerosis, Thrombosis, and Vascular Biology, February, 446–461. https://doi.org/10.1161/ATVBAHA.119.313226

Strawbridge, R. J., Ward, J., Ferguson, A., Graham, N., Shaw, R. J., Cullen, B., Pearsall, R., Lyall, L. M., Johnston, K. J. A., Niedzwiedz, C. L., Pell, J. P., Mackay, D., Martin, J. L., Lyall, D. M., Bailey, M. E. S., & Smith, D. J. (2019). Identification of novel genome-wide associations for suicidality in UK Biobank, genetic correlation with psychiatric disorders and polygenic association with completed suicide. EBioMedicine, 41, 517–525. https://doi.org/10.1016/j.ebiom.2019.02.005

Ward, J., Lyall, L. M., Bethlehem, R. A. I., Ferguson, A., Strawbridge, R. J., Lyall, D. M., Cullen, B., Graham, N., Johnston, K. J. A., Bailey, M. E. S., Murray, G. K., & Smith, D. J. (2019). Novel genome-wide associations for anhedonia, genetic correlation with psychiatric disorders, and polygenic association with brain structure. Translational Psychiatry, 9(1). https://doi.org/10.1038/s41398-019-0635-y

Ward, J., Tunbridge, E. M., Sandor, C., Lyall, L. M., Ferguson, A., Strawbridge, R. J., Lyall, D. M., Cullen, B., Graham, N., Johnston, K. J. A., Webber, C., Escott-Price, V., O'Donovan, M., Pell, J. P., Bailey, M. E. S., Harrison, P. J., & Smith, D. J. (2019). The genomic basis of mood instability: identification of 46 loci in 363,705 UK Biobank participants, genetic correlation with psychiatric disorders, and association with gene expression and function. Molecular Psychiatry. https://doi.org/10.1038/s41380-019-0439-8

Acknowledgements

Completing a PhD during a global pandemic is not easy, and I am grateful that I am not, and was never, alone.

I would like to thank my supervisory team, Dr Barbara Nicholl, Professor Daniel Smith, Professor Andrew McIntosh, Dr Mark Bailey and Dr Mark Adams, along with the wider Smith research group at the Institute of Health and Wellbeing, Glasgow. Thank you to the Medical Research Council and the Universities of Edinburgh & Glasgow. Thank you to Lilybank Gardens for being a better workplace than I could have imagined.

Thank you to the Lyalls and Dr Rona Strawbridge, for their friendship/support/mentorship/life advice/four-legged office presence (delete as appropriate). Thank you to Julia Morris for hundreds of questionable lunchtime conversations in the basement coffee room, and to the dedicated Breakdown Coffee duo Rosie Brown & Stephen Wilkie for being the best impromptu support group. Thank you to Amy Ferguson, Rachana Tank, and the occupants past and present of the PhD office of Lilybank (a.k.a. the PhQueens, “comparison is the thief of joy”). Thank you to Frances Bell, Evan Fleischer & Emma Shorter for all the adventures. Thank you to Wayne Chang for being there. Finally, thank you to my sisters, Patricia, for an endless supply of cat photos, and Rose, for being the best possible pandemic flatmate.

I dedicate this thesis to my father, Desmond Johnston, although he jokes that being related to him should be listed under ‘limitations’.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, this thesis is the result of my own work. The contents of this thesis have not been submitted for any other degree at the University of Glasgow or any other institution.

Keira Jacqueline Ann Johnston

May 2021

Abbreviations

AD	Anderson-Darling
BMI	Body Mass Index
BUHMBOX	Breaking Up Heterogeneous Mixture Based on Cross (X) Locus Correlations
ccFDR	Conjunctional Conditional False Discovery Rate
cFDR	Conditional False Discovery Rate
CIP	Congenital Insensitivity to Pain
CNS	Central Nervous System
CNV	Copy Number Variant
CPG	Chronic Pain Grade
CRPS	Complex Regional Pain Syndrome
CWP	Chronic Widespread Pain
dbSNP	The Single Nucleotide Polymorphism Database
DNA	Deoxyribonucleic Acid
DSM-5	Diagnostic and Statistical Manual of Mental Disorders 5th Edition
EAF	Effect Allele Frequency
EFA	Exploratory Factor Analysis
eQTL	Expression Quantiative Trait Locus
FDR	False Discovery Rate
FUMA	Functional Mapping and Annotation of GWAS
GENCODE	Encyclopedia of genes and gene elements (part of ENCODE -

	ENCyclopedia of DNA Elements)
GRM	Genetic Relatedness Matrix
GS: SFHS	Generation Scotland: Scottish Family Health Study
GTEx	Genotype-Tissue Expression
GWAS	Genome Wide Association Study
HPA	Hypothalamic Pituitary Adrenal
HRQOL	Health-Related Quality of Life
HSAN	Hereditary Sensory and Autonomic Neuropathy
HSE	Health and Safety Executive
HWE	Hardy-Weinberg Equilibrium
IASP	International Association for the Study of Pain
ICD-10	International Classification of Diseases 10th Revision
ICD-11	International Classification of Diseases 11th Revision
ICD-9	International Classification of Diseases 9th Revision
IGV	Integrative Genomics Viewer
IV	Instrumental Variable
IVW	Inverse-Variance Weighted
LAVA	Local Analysis of coVariant Association
LD	Linkage Disequilibrium
LDSR	Linkage Disequilibrium Score Regression
LRR	Leucine-Rich Repeat
MAF	Minor Allele Frequency

MAGMA	Multi-Marker Analysis of Genomic Annotation
MCP	Multisite Chronic Pain
MCT2	Monocarboxylate transporter 2
MCT2	Monocarboxylate transporter 2
MDD	Major Depressive Disorder
MDS	Multidimensional Scaling Components
MR	Mendelian Randomisation
MR-RAPS	Mendelian Randomisation with Robust Adjusted Profile Score
MS	Multiple Sclerosis
NCBI	National Center for Biotechnology Information
NOME	No Measurement Error
OR	Odds Ratio
PC	Principal Component
PEPD	Paroxysmal Extreme Pain Disorder
PGC	Psychiatric Genomics Consortium
PHQ-9	Patient Health Questionnaire 9
PRS	Polygenic Risk Score
PTSD	Post-Traumatic Stress Disorder
QOF	Quality and Outcomes Framework
QQ	Quantile-Quantile
QST	Quantitative Sensory Testing
RCT	Randomised Control Trial

RNA-seq	Ribonucleic Acid Sequencing
RPKM	Reads Per Kilobase Million
SD	Standard Deviation
SDI	Sociodemographic Index
SLE	Systemic Lupus Erythematosus
SNP	Single Nucleotide Polymorphism
SW	Shapiro-Wilk
TMD	Temporomandibular Disorder
UCSC	University of California Santa Cruz
UK	United Kingdom
US	United States
USA	United States of America
WHO	World Health Organisation
YLDs	Years lived with disability

Chapter 1: Introduction

This chapter introduces chronic pain and major depression, discussing defining and diagnosing chronic pain and depression, the epidemiology of both conditions, comorbidity of the two conditions, and introduces key concepts in complex trait genetics.

1.1 What is chronic pain?

1.1.1 Definitions

Chronic pain was defined by the International Association for the Study of Pain (IASP) (Treede et al., 2019) until recently as pain persisting beyond the normal healing time, agreed to be 3 months. 'Normal healing time' can vary widely depending on the condition causing the pain and is hard to accurately ascertain, with no standard length of time agreed, e.g., between clinicians and researchers. Another problem with this definition is the fact that many disorders where chronic pain is a main symptom effectively never involve complete healing or are associated with continued tissue damage or degeneration; a good example of this is rheumatoid arthritis. Chronic pain can also be involved where there is no known pathology or damaged tissue, either existing or detectable from the outset of the chronic pain condition (e.g., fibromyalgia). These issues led to a somewhat arbitrary agreed window of 12 weeks as the standard cut off point, beyond which a pain is considered chronic or persistent. Recently, an IASP Task Force was instrumental in adding a code for chronic pain to the ICD-11 (the WHO International Classification of Diseases 11th edition), and for advocating that chronic pain is a disease entity in its own right (Nicholasa et al., 2019; Treede et al., 2019). The IASP definition of pain itself was also recently updated (July 2020) (Raja et al., 2020), to state that pain is defined as:

> "An unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage"

Six key notes accompany this definition:

- Pain is always a personal experience that is influenced to varying degrees by biological, psychological, and social factors.

- Pain and nociception are different phenomena. Pain cannot be inferred solely from activity in sensory neurons.
- Through their life experiences, individuals learn the concept of pain.
- A person's report of an experience as pain should be respected.
- Although pain usually serves an adaptive role, it may have adverse effects on function and social and psychological well-being.
- Verbal description is only one of several behaviours to express pain; inability to communicate does not negate the possibility that a human or a nonhuman animal experiences pain.

This definition and accompanying notes emphasise that nociception refers to the neural process by which noxious stimuli are encoded, whereas pain refers to the unpleasant emotional, sensory perception that is linked to actually or potentially-occurring tissue damage (Jaracz et al., 2016), that pain and nociception do not necessarily occur together (Baliki & Apkarian, 2015), and that pain is thought to be a "complex, perceptual" experience with a large affective component (Asmundson & Katz, 2009).

IASP terminology also includes mechanistic descriptors of pain, defining pain as nociceptive, neuropathic, or nociplastic (IASP, 2017a). Nociceptive pain is defined as that which "arises from actual or threatened damage to non-neural tissue and is due to the activation of nociceptors", neuropathic pain as "caused by a lesion or disease of the somatosensory nervous system", and nociplastic as "pain that arises from altered nociception despite no clear evidence of actual or threatened tissue damage causing the activation of peripheral nociceptors or evidence for disease or lesion of the somatosensory system causing the pain". Additionally, though not included in the IASP terminology, mixed pain states (presence of pain types fitting multiple mechanistic descriptors in a single individual or patient) are receiving increased attention (Freynhagen et al., 2019, 2020).

1.1.2 Measurement (Phenotyping)

Pain is a subjective experience, and chronic pain falls under the umbrella of symptom-based disorders: there are no scans or biological tests that can be used to decisively diagnose chronic pain. At present there are also no objective biomarkers available for use in diagnosing chronic pain (Mouraux & Iannetti,

2018; Reckziegel et al., 2019), presenting a significant barrier both in clinical treatment of pain and in pain research. There are quantitative methods for clinical assessment of pain including chronic pain, such as Quantitative Sensory Testing (QST), originally developed to assess somatosensory changes associated with neuropathic pain and involving application of various stimuli by a clinician (Backonja et al., 2009; G. Cruccu et al., 2010; Giorgio Cruccu & Truini, 2009; Geber et al., 2011; Peripheral Neuropathy Association., 1993). Somatosensory changes in non-neuropathic pain conditions can also be assessed using QST (Geber et al., 2011), and QST is often applied in the study of central sensitisation (see 1.1.4). Other experimental quantitative methods to assess pain and chronic pain include cutaneous biopsy, microneurography, functional and structural brain imaging, chemical neuroimaging, and pharmacological phenotyping (stratifying pain patients by drug response) (Fillingim et al., 2016; Martucci & Mackey, 2016) – these methods have varying utility and usage rates in a clinical setting, and may fail to capture subjective and psychological aspects of pain and chronic pain experience.

In the context of patients or individuals reporting their pain, questionnaire assessments delivered in person by researchers or medical professionals, or remotely via survey, that ask the individual or patient about aspects of pain experience, such as severity, frequency, duration, and resultant disability, are widely used. Unsurprisingly, this generates a great deal of heterogeneity within the category 'chronic pain'. Different questionnaire-based methods to assess chronic pain in patients are reviewed by Dansie and Turk and by Fillingim et al, and can be sorted into seven broad categories; unidimensional pain measures, measures of pain quality and location, pain interference and function (general measures), pain interference and function(specific diseases), HRQOL (Health-Related Quality of Life) measures, psychosocial measures, and finally observational pain assessment measures (Dansie & Turk, 2013). In addition, tools such as the chronic pain grade (CPG) questionnaire, derived by Von Korff and colleagues and validated by Smith et al several years later, span across categories to assess pain intensity, duration, resultant disability and impact on quality of life (Smith et al., 1997; Von Korff et al., 1992).

Several questionnaires for chronic pain assessment, such as the Brief Pain Inventory (Cleeland & Ryan, 1994), also include questions on site of chronic pain

on the body – most often assessed by asking the patient to shade areas on a pain drawing (Jensen & Karoly, 2001). Diagnosis of certain chronic pain conditions is also based on chronic pain location meeting requirements in terms of 'widespreadness' or presence in a minimum number of body quadrants and tender points – these conditions include fibromyalgia and chronic widespread pain (CWP) itself (distinct from its role as a cardinal symptom of fibromyalgia) (Clauw, 2014; Wolfe et al., 1990, 2011). CWP is defined as constant axial (pain confined to a certain area/ 'tender point') pain, in addition to pain in both the upper and lower body quadrants, and left and right side of the body (Burri et al., 2015; Wolfe et al., 1990).

Chronic pain may also be characterised based on probable causal or related injury or illness – neuropathic pain is caused by damage to the somatosensory nervous system (Colloca et al., 2017), and may be chronic in nature. However, neuropathic and non-neuropathic types of chronic pain may 'converge' over time, in terms of changes in the dorsal horn and dorsal root ganglion (DRG) (Xu & Yaksh, 2011). In addition, individuals may be diagnosed with neuropathic pain in complete absence of definite or clear lesions or nervous system damage (Finnerup et al., 2016), and the extent or severity of pain experienced may not match observable nervous system damage (Weir et al., 2019).

Cancer pain may also be chronic in nature, with causes of pain in individuals with cancer ranging widely. Cancer pain can be neuropathic (Mulvey et al., 2017; Stewart, 2014), pain classed as both neuropathic and non-neuropathic can co-occur due directly to tumour growth and activity, to surgical and/or pharmacological cancer treatment, or due to comorbid chronic pain conditions (Caraceni & Shkodra, 2019). Pain may not be related directly to cancer, and distinguishing between acute and chronic pain in the context of cancer is difficult, further complicating classification and treatment (Caraceni & Shkodra, 2019)

Measuring and characterising chronic pain both clinically and in the context of research is challenging, resulting in extensive heterogeneity among and within chronic pain phenotypes, with many chronic pain conditions often occurring together (Maixner et al., 2016). Individuals with chronic pain often receive at least one misdiagnosis (Hendler, 2016), and may also be given an inappropriate

psychiatric diagnosis, such as somatic symptom disorder (Katz et al., 2015). One recent systematic review concluded that there are “hardly two research groups that assess chronic pain in exactly the same manner” (Steingrímsdóttir et al., 2017).

1.1.3 Epidemiology of chronic pain

Chronic pain is estimated to affect approximately 20% of the adult population worldwide (Breivik et al., 2006; Goldberg & McGee, 2011; Gureje et al., 2008; Palmer et al., 2000; Santos-Eggimann et al., 2000; Von Korff et al., 2005), and prevalence can be much higher in certain population subgroups (e.g. amputees, where 85% are affected (Schug & Bruce, 2017)). Disorders involving chronic pain, including migraine, neck and back pain, low back pain and general musculoskeletal disorders, were amongst the top 10 global contributors to years lived with disability (YLDs) consistently from 1990-2017 (GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, 2016; James et al., 2018). Low back pain represented the leading cause of disability worldwide until very recently (replaced by major depressive disorder (MDD) (WHO, 2017)).

Chronic pain and chronic pain disorders are widely documented as being more prevalent in women than in men, often twice as common in women (Bartley & Fillingim, 2013; Fillingim, 2015; Fillingim et al., 2009; Hardt et al., 2008; Munce & Stewart, 2007; Rollman & Lautenbacher, 2001; Tsang et al., 2008). Low back pain also remains in the top three of YLD in both the highest and lowest SDI (sociodemographic index) quintiles (James et al., 2018). For example, there are stark contrasts between the rates of YLDs between high-SDI and low-SDI groups of individuals with low-back pain globally (a difference of approximately twice the level of YLDs per 100,000 higher for low-SDI compared to high SDI) (James et al., 2018). Overall, although chronic pain contributes to disability levels similarly across developed and developing countries, deprivation is associated with increased disability and less effective management for those with chronic pain (Bonathan et al., 2013; Dorner et al., 2011; Jackson et al., 2015; Mills et al., 2019; Poleshuck & Green, 2008; Yu et al., 2020).

Increased mortality may be associated with chronic pain phenotypes such as chronic widespread pain(both all-cause mortality and specific causes of death) (H. I. Andersson, 2009; Macfarlane et al., 2017). Chronic widespread pain is

defined as chronic pain in multiple sites of the body including both above and below the waist, on right and left body quadrants, and axially (Butler et al., 2016; F. Wolfe et al., 1990, 2011). This is distinct from multisite chronic pain (2.3.3.1.2), where chronic pain can be present at a few sites and not necessarily fulfilling quadrant, axial or above/below waist location requirements. The relationship between chronic widespread pain and mortality may be mediated by lifestyle factors associated with pain such as poor diet, reduced physical activity levels, smoking and high BMI (Macfarlane et al., 2017). Psychosocial factors, including depression, may also be involved in the relationship between chronic widespread pain and excess mortality (Da Silva et al., 2018).

1.1.4 From Acute to Chronic Pain

The mechanisms of chronic pain development are not fully known, but likely involve both central and peripheral nervous-system processes, the immune system, and genetic and environmental risk factors, including previous injury and psychological stress (reviewed by (Denk & Mcmahon, 2017)). The relationship between acute and chronic pain also tends to vary greatly: not every person who experiences serious injury or undergoes surgery goes on to develop chronic pain, and conversely, chronic pain may develop after seemingly innocuous procedures (Denk et al., 2014). Additionally, across a variety of chronic conditions associated with chronic pain, the degree of tissue damage is not necessarily correlated with the severity of pain experienced. This has been observed with endometriosis, where disease severity in terms of lesion size and type is generally not associated with increasing severity of chronic pelvic pain experienced (Stratton & Berkley, 2011; Vercellini et al., 2007). This poor correlation between tissue damage or extent of disease and chronic pain experienced is also seen in both osteoarthritis (Dieppe & Lohmander, 2005; Neogi, 2013; Trouvin & Perrot, 2018; Valdes et al., 2012) and rheumatoid arthritis (Meeus et al., 2012).

Significant peripheral neuropathy or central nervous system injury can also be present without subsequent development of chronic neuropathic pain (Colloca et al., 2017). In conditions involving widespread chronic pain such as fibromyalgia, complex regional pain syndrome (CRPS), and conditions such as irritable bowel syndrome and temporomandibular disorder (TMD), there may be an absence of

damaged or diseased tissue altogether, with the individual experiencing debilitating pain regardless (Cairns, 2010; C. Chang et al., 2019; Feng et al., 2012; Goebel, 2011; Jahan et al., 2012; Kosek et al., 2016; Sluka & Clauw, 2016; Verne & Zhou, 2011). This further supports viewing chronic pain as a disease entity as outlined in 1.1.1.

Central sensitisation is associated with the development and maintenance of chronic pain, with features of central sensitisation found across a range of chronic pain-associated conditions (Harte et al., 2018). Central sensitisation is defined by the IASP as increased responsiveness, to normal or sub-threshold afferent input, of nociceptor neurons in the CNS, resulting in hypersensitivity to stimuli and increased pain response (IASP, 2017; Ji et al., 2018). This phenomenon can only be observed directly when both input and output of the neural system are known e.g. through QST (see also 1.1.2), or indirectly through healthcare-professional administered assessment or questionnaire assessment of manifestations of central sensitisations i.e. allodynia (pain resulting from normally innocuous stimuli) or hyperalgesia (heightened sensitivity to pain). As well as being implicated in the transition from acute to chronic pain in general, central sensitisation has also been found to be a common occurrence across chronic pain diagnostic boundaries, from chronic pain at specific body sites such as the shoulder (Sanchis et al., 2015), or pelvis (Kaya et al., 2013), to a range of conditions associated with significant chronic pain, including endometriosis (P. Zheng et al., 2019), rheumatoid arthritis (Meeus et al., 2012), osteoarthritis (Lluch et al., 2014), temporomandibular disorders (La Touche et al., 2018) and fibromyalgia (Desmeules et al., 2003; Woolf, 2011). Although earlier definitions of central sensitisation state a requirement for initial noxious/ painful stimuli, recent study has highlighted that peripheral input (sustained or repeated application of noxious stimulus) may not be required – central sensitisation may result from changes in the CNS that are independent of peripheral input (Hains & Waxman, 2006; Latremoliere & Woolf, 2009; Yang et al., 2014), including dysfunction in endogenous pain control systems (Yarnitsky, 2015).

In addition to central sensitisation specifically, a range of other functional (changes to activity) and structural (changes to composition or appearance) changes in the brain and spinal cord are associated with the development and maintenance of chronic pain (Baliki et al., 2014; Baliki & Apkarian, 2015; Bliss et

al., 2016; Hashmi et al., 2013; Khoutorsky & Price, 2018; Mansour et al., 2013; Sheng et al., 2017). Structural changes such as synaptic spine density, cellular changes (both loss and gain) involving microglia and multiple neuron types, and remodelling of neuronal circuits that results in separation or bringing together of nociceptive and non-nociceptive neurons, have been linked to chronic pain development (Kuner & Flor, 2016; Mansour et al., 2013). Functional changes associated with chronic pain include synaptic plasticity in multiple different brain regions linked to pain such as the anterior cingulate cortex, thalamus, and dorsal horn of the spinal cord (Bliss et al., 2016), the periaqueductal grey (Yu et al., 2014), and more recently in visual networks (Shen et al., 2019).

Considering the above, the transition from acute to chronic pain may occur as follows: firstly, acute injury results in prolonged activation of peripheral nociceptors, namely Aδ- and C-fibres (Apkarian et al., 2005; Moehring et al., 2018). This prolonged activation can lead to neuroplastic changes in central as well as peripheral somatosensory circuits (Cichon et al., 2017; Zhuo, 2008), and changes in higher brain regions associated with emotion. One of the specific kinds of synaptic plasticity that may constitute these neuroplastic changes in the case of chronic pain development include increased glutamate release and increase in the postsynaptic response to glutamate in the spinal cord in the ascending pain pathway (the route of signal transmission from the periphery towards the CNS) (Kuner & Flor, 2016; Latremoliere & Woolf, 2009). The descending pain pathway (the downward route of nerves from the CNS to the periphery via the spinal cord) is also thought to be involved in chronic pain development, through modulation of spinal responses to noxious stimuli (E. P. Mills et al., 2018; Ossipov et al., 2014). In cases without underlying injury or tissue damage, this central sensitisation through neuroplastic changes is still thought to occur – instead of persistent engagement of ascending/descending pain circuits driving persistent experience of pain, pain circuitry outside of these pathways is affected during acute injury and contributes to pain experienced after the healing period. One example of such circuitry is the nucelus accumbens, where studies in rodents showed neuroplasticity associated with development of chronic pain (Chang et al., 2014; Ferris et al., 2019; Goffer et al., 2013). Another example is, in humans, structural changes in corticolimbic circuits (encompassing the prefrontal cortices, hippocampus and amygdala) have also

been found to predict transition to chronic pain (Baliki et al., 2012; Vachon-Presseau et al., 2016).

A range of social and psychological factors are also likely to be involved in the transition from acute to chronic pain, and the role of non-medical/ non-biological factors is increasingly recognised as important in chronic pain management. The biopsychosocial model (Bevers et al., 2016) of chronic pain outlines how psychological, social and biological factors interact to influence the development and course of chronic pain. Factors such as ethnicity, age and gender fall under the psychosocial label in addition to potentially being markers for biological factors linked to chronic pain development (Fillingim, 2017), and lifestyle or behavioural factors such as level of physical activity and cigarette smoking are also associated with risk of chronic pain development (Mills et al., 2019). Previous studies found that factors related to social support such as spousal negative reinforcement of pain behaviours were involved in chronic pain-related disability, and that an introverted personality and tendency towards catastrophizing were associated with increased chronic post-surgical pain (reviewed by (Katz & Seltzer, 2009)). Factors such as low mood and somatising tendency may also contribute to increased risk of developing chronic pain, and at the societal level psychosocial aspects of the workplace may also contribute to chronic pain development risk (Vargas-Prada & Coggon, 2015). A recent systematic review found that fear-avoidance beliefs and depression/ anxiety were both associated with transition from acute to chronic pain in a range of scenarios including post-surgical and non-specific widespread pain syndromes, but also that some studies found no link between psychosocial factors examined and pain chronicity (Hruschak & Cochran, 2018).

The imprecision hypothesis (Moseley & Vlaeyen, 2015) outlines the method by which biopsychosocial factors influence chronic pain development suggesting that a lack of precision in integrating multisensory information (physical, nociceptive, psychological, emotional) leads to chronic pain development through the painful response then generalizing to non-painful events.

Additionally, the functions of brain areas involved in nociception are not limited to pain processing: many are also involved in emotional regulation (Tracey, 2010; Tracey & Johns, 2010), including affective aspects of the pain experience (Peirs

& Seal, 2016; Schweinhardt & Bushnell, 2010). A recent systematic review found maladaptive emotional regulation in general to be linked to increased risk of chronic pain development (Koechlin et al., 2018).

Overall, research across multiple fields suggests that chronic pain conforms to the biopsychosocial model of disease. A complex array of genetic, medical, lifestyle, social and psychological factors are associated with and likely contribute to risk of developing chronic pain, and to pathology or mechanisms of chronic pain development. However, unifying qualities among chronic pain conditions exist across all three (biological, psychological, social) domains, and these similarities could aid understanding of chronic pain development in general and do so more powerfully in comparison to study of chronic pain within disease or diagnostic boundaries. Such similarities include absence of identifiable injury or cause of pain for many individuals with chronic pain, likely extensive CNS involvement in a wide range of chronic pain states and overlap with brain areas involved in emotion and affect.

1.1.5 Associations with Other Conditions

Individuals with certain traits and conditions experience chronic pain at significantly higher rates compared to the general population, and for some conditions and disorders chronic pain is a hallmark symptom. Conditions associated with chronic pain include obesity (Okifuji & Hare, 2015; Paley & Johnson, 2016), and high BMI more generally, with chronic pain incidence estimated to be ~68-254% higher in individuals classed as obese compared to individuals with a BMI of less than 30 kg/m^2 (Paley & Johnson, 2016). Higher BMI and increased body fat may influence chronic pain development through mechanical stress (Okifuji & Hare, 2015; Wearing et al., 2006), activity of molecules secreted from adipose tissue (Hauner, 2005; Urban & Little, 2018), and general inflammation (DeVon et al., 2014; Eichwald & Talbot, 2020).

Autoimmune disorders are also associated with chronic pain (Mifflin & Kerr, 2017; Phillips & Clauw, 2013). The immune system in general is also implicated in chronic pain development, including inflammatory responses in the brain and spinal cord (neuroinflammation) (Ren & Dubner, 2010). The complement system, part of the innate immune system, has also been found to play a part in synaptic pruning and neuronal connectivity during both development and as part of

neurodegenerative disease progression (Stephan et al., 2012). There is also significant communication between the nervous and immune systems in nociception and in sensitisation processes that can lead to chronic pain (Kwiatkowski & Mika, 2018; Pinho-Ribeiro et al., 2017). Though not classed as an autoimmune disease, another disorder with immune involvement, asthma, may also be associated with increased chronic pain risk – this may be due to musculoskeletal damage involved with severe coughing during asthma attacks or with postural changes associated with asthma (Lunardi et al., 2011), with higher opioid use associated with having asthma (Naik et al., 2019). Additionally, autoimmune conditions that can involve significant and chronic pain such as lupus have been found to be more common in those with asthma (Krishna et al., 2019), and pain has been found to be a significant comorbidity and generally more common in individuals with asthma compared to those without (Weatherburn et al., 2017).

Insomnia and sleep disturbance are also commonly experienced by those with chronic pain, with ~65% of those with chronic pain conditions also having clinical insomnia (Alföldi et al., 2014), rates which are 2-20x higher than those estimated for the general population (Roth, 2007; Singareddy et al., 2012; Y. Zhang et al., 2019). Reduced sleep duration and poor sleep quality may be a significant risk factors in development of subsequent chronic pain, in addition to potentially being caused by pain (Broberg et al., 2021; Haack et al., 2020; Jank et al., 2017; Sun et al., 2020). Opioid treatment of chronic pain can also negatively impact sleep (Ferini-Strambi, 2017; Tentindo et al., 2018). Improving sleep duration and quality has the potential to improve treatment outcomes for comorbid chronic pain, with individuals with chronic pain likely to experience increased pain sensitivity, lower mood, and higher levels of disability in comparison to individuals with chronic pain but without comorbid sleep issues (reviewed by Cheatle et al., 2016).

Neurological diseases, such as Parkinson's disease, are also associated with chronic pain (Borsook, 2012), as are migraine (Minen et al., 2016) and multiple sclerosis (MS) (Marrie et al., 2012). 30-95% of individuals with Parkinson's disease experience chronic pain (Broen et al., 2012; Buhmann et al., 2017; Valkovic et al., 2015), which can be related to rigidity, posture changes, reduced movement of the joints, and involuntary muscle contractions experienced as part of

Parkinson's, or a central pain syndrome which could be due to Parkinson's-related brain changes (Blanchet & Brefel-Courbon, 2018). Pain can also be classified in terms of whether it is thought to be directly related, indirectly related, or not related to Parkinson's disease in the individual, and further labelled in terms of whether this pain is experienced in the off or on-phase of the condition (Skogar & Lokk, 2016). Individuals with MS tend to experience pain and pain syndromes more often than the general population, with estimates of pain prevalence of ~30-80% (Drulovic et al., 2015; Foley et al., 2013; Heitmann et al., 2020; O'Connor et al., 2008; Solaro et al., 2013), and estimates of chronic pain prevalence more specifically ranging from ~40-50% (Ehde et al., 2003; Ferraro et al., 2018) to as high as 86% (Urits et al., 2019).

A wide range of psychiatric traits and disorders have been found to be associated with chronic pain. These include addiction and substance use disorders (Cheatle & Gallagher, 2006; Elman & Borsook, 2016; Speed et al., 2018), with 8-12% of those with chronic pain prescribed opioids going on to develop an opioid use disorder (reviewed by Speed et al., 2018), in contrast to 0.6% of the US population aged 12+ in general estimated to misuse analgesic medication (SAMHSA, 2018).

PTSD in both veterans and civilian populations is associated with higher rates of chronic pain (Akhtar et al., 2019; Dunn et al., 2011; Outcalt et al., 2015; Phifer et al., 2011; Shipherd et al., 2007). For example, a non-veteran sample attending pain clinic for treatment of chronic pain was found to have rates of PTSD over four times as high as that of the general US population (28% vs. ~6%) (Akhtar et al., 2019), and other studies found between 46%-66% of combat veterans seeking chronic pain treatment had PTSD (Dunn et al., 2011; Shipherd et al., 2007). A systematic review found consistent evidence that PTSD was associated with chronic pain (Fishbain et al., 2017).

In addition to PTSD, anxiety disorders in general are commonly comorbid with chronic pain (Asmundson & Katz, 2009; Gureje, 2008). 2012 Canadian Community Health Survey–Mental Health participants with chronic pain were found to have generalised anxiety disorder (GAD) up to 2.6x more often than in comparison to the entire cohort (Csupak et al., 2018), and World Mental Health Survey results

indicated that participants reporting chronic pain showed increased odds from 90-170% of having a comorbid anxiety disorder (Gureje, 2008).

Individuals with schizophrenia commonly experience chronic pain, and often have comorbid conditions associated with significant chronic pain (De Hert et al., 2011; Gabilondo et al., 2017; Smith, Langan, et al., 2013; Von Hausswolff-Juhlin et al., 2009). However systematic reviews found prevalence of pain with apparent medical cause to be lower amongst a sample of individuals with schizophrenia in comparison to the general population (Engels et al., 2014), or similar when compared to age and sex-matched controls (Stubbs et al., 2014). In contrast other studies, for example of cohorts of veterans, found schizophrenia to be associated with higher rates of chronic pain (in comparison to veterans without this psychiatric comorbidity) (Birgenheir et al., 2013). In addition, differences in pain perception and the integration and processing of sensory information (interoception) in those with schizophrenia, have been reported. One study showed participants with schizophrenia to have elevated sensitivity to acute pain and reduced sensitivity to prolonged pain in an experimental setting (Lévesque et al., 2012), though another study highlighted that such differences may be due to issues in expressing and reporting pain for individuals with schizophrenia, as opposed to nociception-related effects (Urban-Kowalczyk et al., 2015). Autism spectrum disorder and anorexia nervosa have also been associated with altered pain perception and interoception (Bär et al., 2015; Bischoff-Grethe et al., 2018; C. Clarke, 2015; Gu et al., 2018; Strigo et al., 2013), which may impact chronic pain prevalence and reporting in these specific populations. There is growing evidence that many autistic people also have significant joint hypermobility (Baeza-Velasco et al., 2018; Casanova et al., 2020; Csecs et al., 2020), often associated with chronic pain, and which may or may not be subthreshold to official Joint Hypermobility Syndrome (JHS) or Ehlers-Danlos (Castori et al., 2017) diagnosis.

Similarly to schizophrenia, living with bipolar disorder is associated with a range of serious and pain-associated chronic physical conditions (De Hert et al., 2011). In contrast to results from some studies of individuals with schizophrenia, those with bipolar disorder tend to experience chronic pain at rates higher than the general population (Nicholl et al., 2014; Stubbs et al., 2015), for example with a

relative risk for “clinically relevant pain” of 2.14 and of migraine specifically of 3.3 (Stubbs et al., 2015).

Chronic pain conditions are often commonly comorbid with one another (Maixner et al., 2016). Chronic pain syndromes involving specific body parts or areas (e.g. irritable bowel syndrome, low back pain) were found to be associated with one another (Kato et al., 2009), and chronic pain, including both abdominal and joint pain, is a common symptom for those with inflammatory bowel disease and is often not resolved even in the absence of active disease (Docherty et al., 2011; Norton et al., 2017). Arthritis and fibromyalgia have also been found to be associated with one another (Haliloglu et al., 2014). Neuropathic ocular pain has also been found to be associated with other chronic pain syndromes (Galor et al., 2016). Rheumatoid arthritis is associated with a wide variety of pain experiences, but pain is often the most significant and disabling symptom, even with well-managed inflammation (Walsh & McWilliams, 2014).

Explanatory factors connecting chronic pain and other disorders, including MDD, involve shared biological mechanisms, environmental factors, shared psychological aspects, or most likely a complex mixture of multiple genetic and non-genetic factors. There is extensive overlap not only between different chronic pain conditions, but also between chronic pain conditions, chronic pain experience in a general sense, and a diverse range of traits and conditions, many of which do not feature chronic pain as a core symptom. The focus of this thesis is aspects of the relationship between chronic pain and MDD specifically (see also 1.3.1).

1.2 What is Major Depressive Disorder (MDD)?

1.2.1 Screening and Diagnosis of MDD

Diagnoses of depression and of MDD are based on the self-report of symptoms, often in a primary care setting using self-report inventories where the individual completes a survey or questionnaire. Most depression rating scales fall under this umbrella, although some are completed by researchers (e.g. Hamilton Depression Rating Scale (Hamilton, 1960; Williams, 1988)). The most commonly used screening tools in a primary care setting for adults is the Patient Health Questionnaire-9 (PHQ-9) (Kroenke et al., 2001; Spitzer et al., 2000). The PHQ-

9can also be used in more specific populations such as post-partum and older adults, although more specialised screening tools such as the Edinburgh Post-Natal Depression Scale and the Geriatric Depression Scale are also available (reviewed by (Maurer et al., 2018; Sharp & Lipsky, 2002)). The PHQ-9 is also one of three measures of depression severity recommended by the UK general practice contract Quality and Outcomes Framework (QOF) (Kendrick et al., 2009).

If an individual is screened and scores positively for MDD, this diagnosis should then be confirmed using the Diagnostic and Statistical Manual of Mental Disorders (DSM), currently in its fifth edition (DSM-5). The DSM classification is used by researchers in mental health (Regier et al., 2013) and consists of lists of symptoms and threshold levels of endorsements of these symptoms for a positive diagnosis of a psychiatric disorder.

In order to meet the criteria for a DSM-5 diagnosis of MDD, an individual must have five or more symptoms from two lists of criteria (A and B), at least one of which must come from the A list; A: depressed mood, markedly diminished interest or pleasure in almost all activities, B: significant weight loss/gain or decrease/increase in appetite, insomnia or excessive sleep, psychomotor agitation or retardation, fatigue or loss of energy, feelings of worthlessness or excessive/inappropriate guilt, diminished concentration or indecisiveness and finally recurrent thoughts of death, suicidal ideation, plans or an attempt. There is also an ICD-10 equivalent for DSM-5 MDD, 'Major Depressive Episode', again with two lists of criteria (reviewed in McIntosh et al., 2019). For both DSM-5 and ICD-10 diagnoses both sets of criteria also require that the symptoms have lasted at least two weeks, that there is significant functional impairment, and that the disorder is not better accounted for by another condition.

Even use of the same 'instrument' to diagnose MDD (such as the DSM) can result in a wide range of symptom profiles being grouped into the same diagnostic category. The single diagnosis of MDD based on DSM-IV criteria can cover over 100 different and in some cases non-overlapping symptom combinations (Fried & Nesse, 2015a; Olbert et al., 2014; Zimmerman et al., 2015).

In addition, many large epidemiological studies of depression also use self-reported depression phenotypes (e.g., answering survey or questionnaire items as to whether participant has ever been diagnosed with depression by a doctor,

seen a psychologist or psychiatrist). These are often very different to methods used in primary care or other clinical settings. However, self-reported phenotypes can share significant overlap with clinical diagnoses of MDD, and the two together (MDD and self-reported depression) have been used as a single diagnostic group in some studies of MDD (McIntosh et al., 2019).

1.2.2 Epidemiology of MDD

An extensive review found that lifetime prevalence estimates of MDD from population surveys worldwide ranged from 1-19%, with prevalence higher in high-income versus low-income countries, and with a worldwide average prevalence of 11.1% and age of onset at 24 years old (Kessler & Bromet, 2013). Another study, using the World Mental Health Survey, found a global estimate of MDD prevalence to be 5.5-5.9% (Ferrari et al., 2013).

A study of US populations, using DSM-5 diagnoses, found the 12-month and lifetime prevalence of MDD to be 10.4% and 20.6% respectively (Hasin et al., 2018). A European estimate of 12-month MDD prevalence was found to be 5% (Ferrari et al., 2013). Prevalence of 12.2% has been inferred for Scotland from work in the Generation Scotland: Scottish Family Health Study (Fernandez-Pujals et al., 2015).

Similar to chronic pain, Kessler & Bromet also found that women were twice as likely to have MDD as men, and that this was consistent across different adult population samples around the world (Kessler & Bromet, 2013). Earlier work also found lifetime incidence of MDD to be almost twice as high in women compared to men (20% vs 12%) (Belmaker & Agam, 2008). Work involving the GS: SFHS also found higher prevalence in women than in men (15.8% versus 9.1%) (Fernandez-Pujals et al., 2015). This 2:1 ratio appears to vary with age (WHO, 2017), first emerging in adolescence and early adulthood (Avenevoli et al., 2015; Patton et al., 2008). In contrast to some studies suggesting convergence of male and female prevalence rates of major depression in older age in some populations (Forlani et al., 2014; Kuehner, 2017; Patten et al., 2016), this ratio does appear to persist into old age (Byers et al., 2010; Girgus et al., 2017; Luppa et al., 2012). Additionally, in pre-puberty males may be at greater risk of depression than females of the same age (Douglas & Scott, 2014).

The prevalence of MDD for those with chronic comorbidities is between three and seven times higher compared to those without. Comorbidities here refers to other health, including mental health, conditions experienced by the same individual simultaneously with MDD. Earlier work across the 48 contiguous states of the USA found an MDD prevalence of 16.2%, and that most lifetime and 12 month MDD cases (>70% in both categories) had comorbid psychiatric disorders (Kessler et al., 2003). Furthermore, the comorbidity of MDD with other psychiatric and substance use disorders was substantial in later work on a large US sample (Hasin et al., 2018). Risk of all-cause mortality was significantly increased in most common mental disorders, including in depression (Chesney et al., 2014).

1.3 Overlap between MDD and Chronic Pain

1.3.1 Comorbidity between MDD and Chronic Pain

MDD and chronic pain are often comorbid: chronic pain is found at higher rates than expected in those with MDD and vice versa, and this is true across a diverse range of populations around the world. Estimates quantifying the degree of comorbidity between MDD and chronic pain vary widely: one review found chronic pain in people with depression to range from 15-100%, and prevalence of depression in people with chronic pain from 1.5-100% (Bair et al., 2003). Another study found 65.7% of those with MDD had chronic pain, compared to 43.5% of those without MDD, and that chronic pain was more likely to be disabling in those with MDD (Arnow et al., 2006). In people with chronic pain 10.4% also met the criteria for MDD, compared to 4.5% of people without chronic pain who met the criteria for MDD (Arnow et al., 2006).

Analyses of the World Mental Health survey results found higher rates of mood disorders including MDD are found in those with chronic pain across a range of global populations, and these rates increase with number of pain sites (Gureje et al., 2008; Tsang et al., 2008). 66.3% of individuals with MDD reported chronic pain in a US study (compared to 49% of the whole sample) – if this was more stringently limited to chronic pain that led to medical consultation or medication use this resulted in 44.2% of MDD subjects with this level of pain (compared to 21.8% in the entire sample). 73.3% of individuals reporting a chronic painful physical condition also meet the criteria for MDD (Ohayon &

Schatzberg, 2010). An independent positive association was also found between depression/ anxiety and chronic pain in a study of a New Zealand population (Dominick et al., 2012). Chronic pain was more prevalent in individuals with MDD compared to those without a history of mood disorder (50.4% versus 38.2%) in a subset of UK Biobank (Nicholl et al., 2014), with a positive relationship seen between the number of sites of chronic pain and the risk of MDD. It was also found that unexplained painful physical symptoms, including chronic pain, are experienced by up to two thirds of patients with MDD (Jaracz et al., 2016). This comorbidity can negatively impact success in treatment and management for either disorder (Asmundson & Katz, 2009; Bair et al., 2003, 2008; Jaracz et al., 2016; Ohayon & Schatzberg, 2010).

1.3.2 Causal Relationships between MDD and Chronic Pain

Although comorbidity between chronic pain and MDD is high, the temporal nature of the relationship is not fully clear. Causality in relationships between MDD and chronic pain has been previously explored in both pre-clinical (non-human) and human samples, but with conflicting results. In mouse models of neuropathic pain, pain was found to have a causal effect on depressive behaviour in several studies, as was arthritis, IBS (in female mice and not males) (reviewed by (Li, 2015)). A general chronic-pain phenotype in Wistar-Kyoto mice was also found to exacerbate depression-like symptoms (reviewed by (Li, 2015)). Animal model studies investigating any potential causal effect of depression on pain, however, show less clear results, whereas studies assessing causal effects of pain on depression showed some consistency in results regardless of modality (the way pain/depression-like symptoms are measured) (reviewed by (Li, 2015)).

A range of cross-sectional and longitudinal studies in human populations tend to suggest that chronic pain has a causal effect on depression. A longitudinal study found that chronic pain in rheumatoid arthritis patients seemed to have a causal effect on development of depression (Brown, 1990). A later extensive review showed several studies where results suggest that pain causes depression (demonstrated through depression severity increasing with number of sites of pain), and that depression was not antecedent to pain but was a consequence. Studies were of a range of pain types, including cancer pain. They also highlight three studies of intermittent pain and depression, which showed depression to

be consequent to pain episodes. (Fishbain et al., 1997). Later studies also showed pain at baseline to be predictive of depression onset (Gureje et al., 2001), and that pain contributes to the risk of a first episode of depression (Gerrits et al., 2014).

Other studies show mixed results, suggest depression precedes chronic pain development, or indicate that the depression-pain relationship is reciprocal. In a study of US participants where pain and MDD were surveyed, pain occurred prior to the first depressive episode in 57.1% of cases, concurrently with a depressive episode in 14.3% of cases, and following a depressive episode in 24.3% of cases (Ohayon & Schatzberg, 2010). Other work suggests the pain-depression relationship to be bidirectional (Bair et al., 2003; Kroenke et al., 2011; Von Korff & Simon, 1996). Studies also link both pain and depression to HPA axis dysfunction (Blackburn-Munro, 2001; Hasler, 2010), or indicate that the shared genetic and environmental factors influencing MDD and chronic pain may act independently on either condition (Pinheiro et al., 2015). A study in paediatric chronic pain indicated onset of psychiatric disorders preceded chronic pain development (Tegethoff et al., 2015), and a study of adults (free from chronic pain at baseline) followed up for 24 months found depression to triple the incidence of chronic pain in later waves (Currie & Wang, 2005).

Studies investigating causal relationships between MDD and chronic pain in human populations vary widely in many respects, including the assessment of MDD and chronic pain, the kinds of chronic pain conditions and bodily sites investigated, in sample size and in other population characteristics. Cross-sectional studies most often do not or cannot explicitly test causality. In addition, even longitudinal studies with data collection over multiple time points may be subject to extensive confounding (Streeter et al., 2017) which may influence results. The use of a large general-population sample with genotyping data, such as the UK Biobank, can address these outstanding issues in investigating causal direction between depression and chronic pain development.

1.3.3 Genetics of Complex Traits

1.3.3.1 Common genetic variation and common traits and diseases

Many human traits, which can include physical characteristics such as height and weight, disease status, or personality and mental health related characteristics, have a genetic component; they cluster within families and are hereditary, and this is observable and quantifiable through twin and pedigree studies. In some cases, variation in phenotypic or trait value is due to a single mutation or disruption in a single gene, and inheritance patterns clearly show the dominant or recessive nature of the mutation underlying the trait – termed Mendelian inheritance in reference to Mendelian traits, first outlined by Gregor Mendel in his work in plant genetics (Mendel, 1866). One example of a Mendelian disease trait, where phenotypic variation can be mapped back to a single gene, is Huntington's disease. Here the causal variant is a CAG repeat expansion in the *huntington* gene inherited in an autosomal dominant fashion (Macdonald et al., 1993) and protein-coding changes drive trait variation (Botstein & Risch, 2003).

In other cases, traits show a genetic, heritable component, but patterns of inheritance are less clear. Rather than changes at a single gene resulting in corresponding changes to a single phenotype or trait, trait variation is influenced by many small-effect variants, the external environment, and interactions between these components. These complex disease traits also tend to be more common than traits or diseases that are associated with large detrimental effects at single genes – common genetic variation most likely contributes the largest proportion of variance to the phenotype, and this variation would not persist in the population at the frequency it does if it were extremely deleterious, due to natural selection. Genetic variants with large effects are virtually always rarer –these large-effect variants will have been subject to negative selection and therefore circulate at low frequency in the population.

Rare variants do not provide a sample pool large enough to test for association across the genome with sufficient power. The finer resolution of common genetic variation contributing to most complex traits is not fully understood. Variation in complex traits also appears to be often influenced by variants in non-coding regions of the genome (Li et al., 2016; Pickrell, 2014; Welter et al.,

2014), again in keeping with selective constraint ideas (i.e. most of the variation in the human genome that could potentially be associated with any trait is in non-coding regions (Hindorff et al., 2009)). Common genetic variation here refers to Single Nucleotide Polymorphisms (SNPs), single-base changes in DNA sequence, usually with minor allele frequency (MAF) of more than 5%, and not less than 1%. SNPs with an MAF of less than 1% are considered rare variants.

One example of a complex trait is human height, where several hundred SNPs have been found to be associated with height (Wood et al., 2014) and environmental factors such as nutrition also contribute. Disease traits can also be complex – complex diseases include Parkinson's disease, where both genetic and environmental factors are thought to contribute to the disease phenotype, and high blood pressure (hypertension), the pathological 'upper end' of a continuous complex trait phenotypic value spectrum (blood pressure), also influenced by many common genetic variants (Evangelou et al., 2018) and by environmental and lifestyle factors.

These two types of traits, (single gene) Mendelian and complex (or quantitative), are not necessarily as distinct as previously thought. Instead, Fisher's infinitesimal model of inheritance of quantitative traits, whereby an infinitely large number of genetic variants, in addition to environmental factors, contribute to phenotypic variation (Barton et al., 2017; Mather, 1964), and Mendelian trait inheritance can be thought of as opposite ends of a spectrum in terms of number of contributing genetic variants. Fisher's infinitesimal model unified competing schools of thought at the turn of the 20th century (biometricians versus Mendelian geneticists) to establish quantitative genetics as a research field (Nelson et al., 2013).

Another important model in considering complex traits, specifically disease traits, is the liability-threshold model. Generally speaking, a threshold model is any model where a threshold distinguishes ranges of values i.e. where behaviour predicted by the model (the outcome) varies in some way above or below a particular threshold. In genetics such models were first applied in studies of guinea pig polydactyly by Wright (Wright, 1934b, 1934a). In this work he outlined that although the trait in question was binary in nature (three-toed versus four-toed), the underlying genetic factors contributing to this could not be a

"singular factor" (Wright, 1934a), and additionally that environmental factors contributed to whether the animal's morphology was pushed over this "physiological threshold" (Wright, 1934a). He also theorised that particular guinea pig strains were much closer to the polydactyl threshold, due to increased genetic and/or environmental risk burden, despite appearing to be phenotypically identical to "normal" strains (Wright, 1934b). The modern disease liability-threshold model in complex trait genetics with regard to human disease is attributed to Falconer (Falconer, 1965, 1967), with disease traits as "threshold characters" and liability to developing disease described as a graded attribute which incorporates both innate and external contributors to increased risk of developing disease.

Disease liability-threshold models represent a way to incorporate both genetic and environmental contributions to disease-trait phenotypic variance for binary traits, where above a certain threshold of accumulated genetic and environmental risk factors the outcome varies significantly (i.e. disease is present, versus below the threshold, disease is not).

Common genetic variants can be tested for their association with traits of interest via Genome Wide Association Studies (GWASs) (Visscher et al., 2017), discussed in further detail in 2.2.1. MDD and chronic pain are both complex traits, with an inherited genetic component in addition to environmental factors, and interaction between genetic and environmental factors, contributing to phenotypic variance. Both traits can be examined within the disease liability-threshold model framework, e.g., in chronic pain the "physiological threshold" for diagnosis when chronic pain is considered a binary or threshold character may be reached with increasing genetic risk burden in combination with environmental factors (e.g. injury, surgery, or disease).

1.3.3.2 Pleiotropy and genetic correlation

Many hundreds or even thousands of common genetic variants contribute to variation in each complex trait, and the number of human complex traits is large but still finite. This means that there is significant genetic 'overlap' in terms of the genetic architecture of complex traits. Common genetic variants often contribute to variation in more than a single trait. These contributions to variation in more than one trait may be made independent of one another, so-

called ‘biological’ or ‘horizontal’ pleiotropy (Fig 1: a + b), or a variant may contribute to variation in a trait, which then itself contributes to variation in a second trait, termed ‘mediated’ or ‘vertical’ pleiotropy (Fig 1: c). Pleiotropy overall is extremely commonplace in human complex traits and diseases (Gratten & Visscher, 2016; Hackinger & Zeggini, 2017; Visscher & Yang, 2016; Watanabe et al., 2018), and complicates the investigation of causal relationships and mechanisms of disease development. Furthermore, ‘directional’ pleiotropy refers to when shared variants tend to be associated with the same direction of effect in both traits, making the average value across variants non-zero, and ‘balanced’ pleiotropy to when there are opposing directions of effect associated with shared variants, effectively cancelling each other out.

Genetic correlation and pleiotropy are closely linked. Two traits are genetically correlated when a significant proportion of associated genetic variation is shared between them, as can occur with a high enough degree of pleiotropy. Formally, genetic correlation r_g is the additive genetic covariance between two traits scaled by the geometric mean of the trait variances. Genetic correlation can inform on shared genetic influences contributing to variation in two compared traits, as well as in applications such as validation of measurement of a phenotype in one cohort by assessing its genetic correlation with the same phenotype in a separate cohort (which should approach 1, if the measurement is examining sufficiently similar trait constructs).

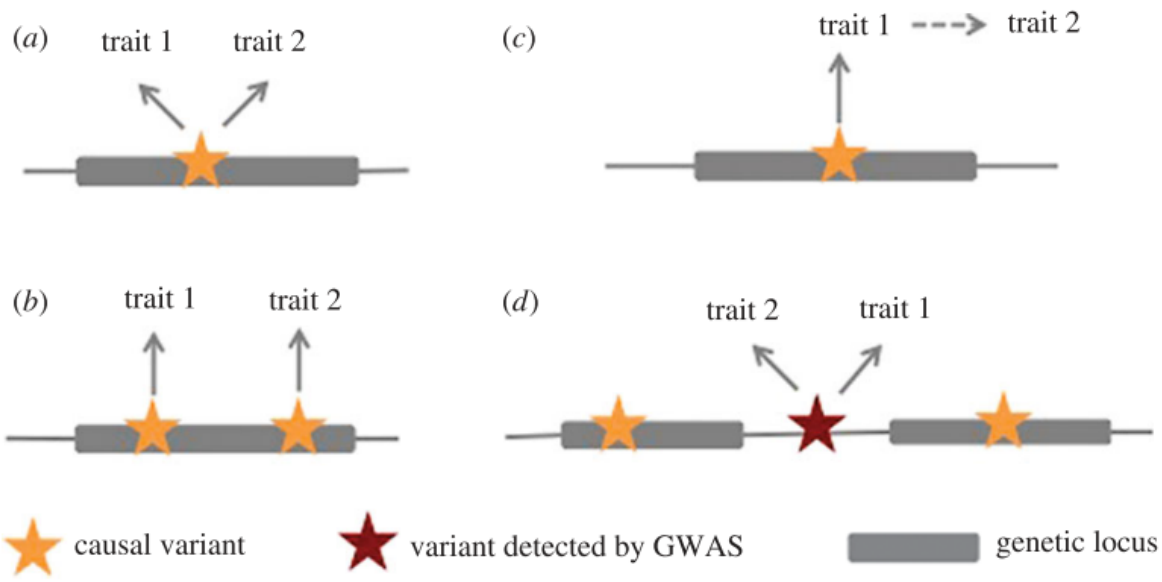

Figure 1. 1: Pleiotropy.

(a) and (b) show biological (horizontal) pleiotropy, with a single causal variant influencing more than 1 trait, or two causal variants at a single locus affecting two traits (and so a single locus affecting two traits) due to linkage disequilibrium. (c) shows mediated a.k.a vertical pleiotropy – variant influences trait which then has an effect/ associated with an effect on a 'subsequent' trait. Note that this type of pleiotropy is the cornerstone of MR. (d) shows spurious pleiotropy where the causal variants/ associated variants are in independent loci but are tagged by (i.e. in LD with) a single variant in both trait 1 and 2. Diagram from (Hackinger & Zeggini, 2017).

Heterogeneity, specifically clinical heterogeneity, is due to the misclassification of individuals into disease or phenotype categories. This misclassification can be due to shared risk factors, and overlap in symptom profiles, and error in measurement and assessment.

In psychiatric disorders error in measurement is introduced as diagnoses may overlap and are based on questionnaire assessment (i.e., there are no laboratory, biomarker or imaging tests to decisively deliver a psychiatric diagnosis). Depression or MDD may therefore be several distinct disorders (Cai et al., 2020; Schwabe et al., 2019), and in our own assessment of this we introduce heterogeneity via the structure of the questionnaires. For example, considering all DSM-5 depression symptoms to be of equal importance and of more importance than non-DSM symptoms results in a large number of non-overlapping symptom profiles being categorised under the same diagnostic label (Fried et al., 2016; Fried & Nesse, 2015b; Olbert et al., 2014). The situation for chronic pain is similar to that of major depression (see 1.1.2), again potentially leading to clinical heterogeneity because of error in measurement and assessment.

Partially overlapping symptom profiles of chronic pain and depression may also contribute to potential clinical heterogeneity in depression with respect to chronic pain and vice versa. For example, those with chronic pain commonly report fatigue (Van Damme et al., 2018), and fatigue or loss of energy is included as a non-core symptom of MDD in DSM definitions, complicating the diagnosis of depression in chronic pain patients (Knaster et al., 2016). Additionally, manifestation of depression or depressive symptoms in certain groups could be misclassified as chronic pain altogether – men often view depression symptoms, particularly physical or somatic symptoms, as an indicator of physical illness (Seidler et al., 2016). In general many people with depression seek treatment in primary care for somatic symptoms, including aches, pains, and fatigue (reviewed (Kapfhammer, 2006)). Chronic pain and MDD also share

many risk factors, e.g., being female, or of lower socioeconomic status, and are commonly comorbid, further contributing to increased likelihood of clinical heterogeneity in the two conditions.

Clinical heterogeneity can also be described as subgroup pleiotropy (Han et al., 2016), and if the general case is that e.g. MDD is misclassified as chronic pain, then that may be a major contributor to the observation of shared genetic factors between the two conditions, as opposed to this sharing of genetic factors being an indicator of true pleiotropy. It is therefore important to distinguish clinical heterogeneity, or subgroup pleiotropy, from true or whole-group pleiotropy to further understand the genetic architecture of both MDD and chronic pain. Confirming that misdiagnosis of chronic pain as MDD and vice versa is not a significant issue also has implications for examining causal relationships between the two conditions.

1.3.4 Genetics of Chronic Pain and Chronic Pain Disorders

Twin studies have also shown several chronic pain disorders to have a heritable component – a systematic review of a range of twin studies of chronic pain phenotypes found heritability to range from ~25% (irritable bowel syndrome) to 77% (in studies of headache including migraine) (Nielsen et al., 2012). Nielsen et al also note that heterogeneity and lack of pain intensity measurement, lack of assessing the pain itself, and use of dichotomous pain phenotyping may reduce power. This approach to measurement may also mean that genetics and resulting heritability estimates may be related to tissue pathology rather than pain processing or the chronic pain itself. Subsequent twin studies of chronic pain phenotypes have also indicated moderate heritability, including phenotypes such as low back pain, H^2 = 21-67% (P. H. Ferreira et al., 2013; Junqueira et al., 2014), and the number of sites (0-31) of chronic pain H^2 = 55-63% (Burri et al., 2018).

Non-family genetic studies of chronic pain have, to date, commonly been investigated using candidate gene and animal model-based approaches (Zorina-Lichtenwalter et al., 2016, 2017). Although not a focus of this thesis animal models of pain and chronic pain reviewed in greater detail by (Burma et al., 2017; Mogil et al., 2010). In addition to chronic pain as described and investigated in this thesis (as a complex trait), rare autosomal recessive genetic

diseases such as congenital insensitivity to pain (CIP) have been previously described (Golshani et al., 2014; Nagasako et al., 2003). CIP inhibits the ability to perceive any physical pain, and this difference in perception is present from birth. CIP is heterogeneous both in terms of clinical presentation and genetic mutations causing the disease, and mutations in genes including *NTRK1* (neurotrophic tyrosine kinase receptor type 1) (Kilic et al., 2009), *PRDM12* (PR domain zinc finger protein 12) (Chen et al., 2015) , and *SCN9A* (sodium channel voltage-gated type IX, alpha subunit) (Majeed et al., 2018; Peddareddygari et al., 2014), among many others, have been implicated. CIP may also be more specifically classified according to symptoms, genetic cause, and/or comorbidities such as intellectual disability, into one of five kinds of hereditary sensory and autonomic neuropathy (HSAN) (Houlden et al., 2006; Lafrenière et al., 2004; Minde et al., 2004; F. Zhao et al., 2020). More recently, a novel pseudogene microdeletion (in *FAAH,* fatty acid amide hydrolase) was found in a 66 year old British woman that conferred pain insensitivity, fast-healing wounds, and absence of anxiety, fear and depression (Habib et al., 2019). Two further rare genetic disease associated with mutations in *SCN9A* have also been described. One such condition is primary erythromelalgia, characterised by erythema (rashes), temperature changes (warmth) and episodes of burning pain in the extremities (Dabby, 2012; Fischer & Waxman, 2010; Mann et al., 2019; Tang et al., 2015). Disease inheritance in primary erythromelalgia is autosomal dominant, with gain-of-function mutations in the *SCN9A* gene (which encodes voltage-gated sodium channels) causing symptoms (Fischer & Waxman, 2010; Mann et al., 2019; Tang et al., 2015). Another autosomal dominant disease associated with extreme pain is paroxysmal extreme pain disorder (PEPD), caused by gain-of-function mutations in genes encoding voltage-gated sodium channels (Fertleman et al., 2006; Fischer & Waxman, 2010). PEPD manifests as episodic burning pain (of ocular, mandibular and rectal regions as opposed to extremities in erythromelalgia), which can be accompanied by non-epileptic seizures and slowed heart rate in addition to skin flushing (Fischer & Waxman, 2010).

Candidate gene studies, where variants within a gene region chosen a priori by researchers are tested for their association with a complex trait (Patnala et al., 2013; Tabor et al., 2002) (in contrast to single-gene Mendelian pain disorders described above), have also been carried out for chronic pain phenotypes.

However, candidate gene studies in general may be more likely to yield false-positive associations (Sullivan, 2007), and so candidate genes in the study of chronic pain, such as *COMT, SLC6A4, GCH1, OPRM1* and *ADRB2* (Mogil, 2012; Veluchamy et al., 2018), may not be associated with chronic pain in large-scale GWAS of chronic pain as a trait.

Certain chronic pain phenotypes such as CPG, a graded classification of chronic pain assessing pain severity, duration, resultant disability and impact on quality of life first constructed by von Korff (Von Korff et al., 1992) and colleagues and later validated by Smith et al (Smith et al., 1997), pain at specific bodily sites (e.g. low back pain), and specific chronic pain related conditions (e.g. migraine, temporo-mandibular joint disorder), have been shown to be complex traits with moderate heritability, and common genetic variation (SNP variation) has been found to contribute to variation in these traits (Hocking et al., 2012; McIntosh et al., 2016; Nicholl et al., 2011; M. J. Peters et al., 2013; Suri et al., 2018; Zorina-Lichtenwalter et al., 2016, 2017). However, as pain assessment and experience are so heterogeneous (Steingrímsdóttir et al., 2017; Vellucci, 2012), there are few large-scale genetic, particularly GWAS, studies of chronic pain as a phenotype in its own right (Nicholl et al., 2011; Tsepilov et al., 2020; Zorina-Lichtenwalter et al., 2016, 2017). However, GWAS studies of chronic low back pain and chronic pain in particular body sites have been previously carried out (Meng et al., 2020; Suri et al., 2018).

Due to the fact that common genetic variants across the genome are tested for their association with a complex trait in a GWAS, sufficient sample size is essential (Hong & Park, 2012). At an absolute minimum, this sufficient total sample size is estimated to be 2,000 individuals (Hong & Park, 2012), and in general number of variants discovered appears to reliably increase with increasing sample size (Visscher et al., 2017). GWAS of chronic pain likely requires extremely large sample sizes to find associated common genetic variants. For example, a genome-wide association study with a sample size of ~23,000 found no SNPs to be significantly associated with CPG (McIntosh et al., 2016), a GWAS meta-analysis of low-back pain in ~150,000 individuals showed only three trait-associated SNPs at genome-wide significance (Suri et al., 2018) and recent GWAS of a musculoskeletal pain phenotype with a sample size of ~190,000 found nine trait-associated loci (Tsepilov et al., 2020). Required

sample sizes for discovery of a large number of chronic pain associated SNPs are likely to approach the magnitude (~ 0.5 – 1 million) of those in recent large GWAS meta-analyses of MDD (Howard et al., 2019; Wray et al., 2018).

1.3.5 Genetics of MDD

As previously discussed with respect to chronic pain, candidate gene analyses have been problematic generally, with replication of gene-trait associations often inconsistent in subsequent studies. Wray et al (2012) systematically tested 180 previously highlighted potential candidate genes for MDD, and showed no significant findings (Wray et al., 2012). A more recent paper also investigated a range of historical MDD candidate genes, and again found "not much support" (Border et al., 2019). Linkage analysis findings are also non-overlapping with GWAS findings and likely assumptions of analyses were not robust (reviewed (McIntosh et al., 2019)). Twin studies indicate MDD has a significant genetic component, with heritability estimated at ~30-40% (Kendall et al., 2021; Polderman et al., 2015; Sullivan et al., 2000). Heritability estimates from other types of familial relationships including extended kinship constructs and varying familial-relationship dyads also produce results within a similar range (Fernandez-Pujals et al., 2015; also reviewed by Kendall et al., 2021).

As recently as 6 years ago an extensive review of the genetics of major depression asserted that no GWAS up to that point had found loci significantly associated with MDD, depression or for any traits genetically related to MDD (e.g. neuroticism) (Flint & Kendler, 2014). This paper also emphasised that candidate gene work in MDD had, for the most part, only revealed false positives.

As discussed above (1.3.3.1), MDD is a complex, quantitative trait, where the genetic architecture is highly polygenic, and many common variants of small effect contribute to variation in the trait. MDD is also more common and less heritable than e.g., schizophrenia, further complicating the search for trait-associated common genetic variation. This is emphasised by the fact that sample sizes of over 0.3 million individuals were required before more than one or two variants were found significantly associated with MDD (Hyde et al., 2016; Wray et al., 2018), and a recent GWAS meta-analysis including broader depression phenotypes, the largest to date with a sample size over 1 million participants, found more than 100 SNPs significantly associated with depression (Howard et al.,

2019). Additionally, a much more deeply phenotyped cohort of a more extreme depression phenotype of severe recurrent MDD (mean number of episodes being 5.6) (where contrasts in trait-associated alleles are likely to be much larger between cases and controls and so there is more power to detect them) showed only two significantly-associated SNPs (Cai et al., 2015).

Analysis of broader depression phenotypes (i.e., aside from clinician diagnosed MDD) has also been shown to be of value. In analyses of broad depression, ICD-9 or ICD-10 coded MDD and probable MDD in UK Biobank, Howard et al showed all three phenotypes to be highly genetically correlated (r_g = 0.85-0.87), genetically correlated with depression phenotypes from an independent study (r_g = 0.63-0.79), and that the broad depression phenotype was most highly genetically correlated (more so than either probable or ICD-coded MDD) with clinically defined MDD from an independent study (Howard, Adams, Shirali, et al., 2018; Howard et al., 2019).

GWAS findings of MDD thus far highlight the importance of the immune system, synaptic plasticity and neurogenesis, prefrontal brain regions and multiple types of neurotransmission (calcium, glutamate), as well as genetic correlation with a wide range of psychiatric, behavioural and physical/health traits including schizophrenia, bipolar disorder, neuroticism and BMI (Howard et al., 2019; Wray et al., 2018).

MDD has been found to be significantly genetically correlated with chronic pain. In family-based analyses of environmental and genetic risk for chronic pain, chronic pain grade (M Von Korff et al., 1992) was found to be genetically correlated with MDD at ρ = 0.53, indicating that just over half of the common genetic variation contributing to either disorder is shared. Positive genetic correlations were found between chronic pain at a range of body sites in UK Biobank, and between most of the different chronic pain-site phenotypes and MDD (r_g = ~0.3 – 0.5) (Meng et al., 2019).

1.4 Summary

MDD and chronic pain are commonly comorbid, and represent significant global socioeconomic and health burdens, both individually and when considered together. Mechanisms of chronic pain development, and drivers of differing

vulnerability to developing chronic pain, are not currently fully understood, but likely include biological and medical, and psychosocial factors, and complex interactions between these factors.

Both MDD and chronic pain as phenotypes represent very heterogeneous constructs, complicating the understanding of their aetiology. Chronic pain has recently been defined as a disease entity in its own right by an IASP taskforce and studying chronic pain as a complex trait may be a more tractable way to investigate chronic pain vulnerability and mechanisms in comparison to only studying conditions and disorders associated with significant chronic pain separately. This is comparable to recent large-scale analyses investigating MDD in terms of "broad depression" phenotypes. Common genetic variation associated with these two conditions, chronic pain as a disease and MDD, can be used to address outstanding questions on:

- Common genetic variation associated with chronic pain
- Pleiotropy – to investigate the degree to which common genetic variation is shared between chronic pain and MDD, and which genomic loci are involved
- Clinical heterogeneity – is it possible that depression is mis-diagnosed as chronic pain, and vice versa?
- Causal relationships between chronic pain and depression through Mendelian Randomisation analyses.

Further understanding both chronic pain and MDD through use of common genetic variant data also has the potential to shed light on aetiology and highlight potential new treatment options for both conditions.

1.5 Aims and Objectives

1.5.1 Overall Aim

The over-arching aim of this PhD project is to explore causal relationships between chronic pain and MDD in large UK general-population cohorts with whole-genome genotyping data using a wide range of statistical genetic methods.

1.5.2 Objectives

The overall aim will be achieved through investigations that set out to address 3 main objectives.

1. To uncover common genetic variation associated with chronic pain phenotypes
2. To investigate genetic correlation and pleiotropy between MDD and chronic pain
3. To test for clinical heterogeneity between MDD and chronic pain

Chapter 2: Methodologies and Technical Information

2.1 Introduction

This chapter introduces and outlines in detail different methods in statistical genetics used throughout this thesis. Certain key concepts involved in analyses, such as aspects of complex trait genetics, are also discussed in greater detail compared to their introduction in Chapter 1. Datasets, cohorts, and phenotypes which are used in multiple analyses and referred to in multiple results chapters are also described.

2.2 Methodologies

2.2.1 Genome-Wide Association Studies

Genome-wide association studies (GWASs) are a search for common genetic variation that is associated with a complex trait of interest. Methodologically, this involves many millions of regressions, where single-nucleotide polymorphism (SNP) genotype (i.e., allele complement) is a predictor or independent variable, and trait value (e.g., blood pressure, or case vs control status of a disease trait) is the outcome or dependent variable. Each regression tests whether genotype is associated with trait value. As outlined previously in Chapter 1: Genetics of Complex Traits, SNPs are single-base changes in the genomic DNA sequence, each making a very small contribution to the variation in a trait.

This common variation may 'tag' (be physically nearby and in LD with) a causal variant, whilst having statistical properties that allow for the surveying of the genome, and the general population, in this manner. These properties include the genetic variation being common, which means the sample size of each of the three genotypes (e.g., AA, AT, TT) is more likely to be sufficiently large to give enough power to test association when effect sizes are low.

GWASs were made possible by the advent of next generation sequencing methods (reviewed by Goodwin et al., 2016) and of SNP reference panels, with decades of work both sequencing point-mutation changes and mapping these genetic variants in the human genome involved. In the 1980s, Botstein and colleagues proposed restriction fragment length polymorphisms (RFLPs) be used as molecular markers in linkage studies, with the first RFLP map of the human

genome completed in 1987 (Botstein et al., 1980; also reviewed by Kruglyak, 2007). However, linkage studies are underpowered for the discovery of common loci associated with complex traits, and the need for association studies in non-family structured populations was increasingly recognised (reviewed by Kruglyak, 2007). The SNP consortium (Thorisson & Stein, 2003) and HapMap project (Belmont et al., 2005; International HapMap Consortium, 2003) were formed with the initial goal of providing a dense genome-wide map of SNPs for use as molecular markers in association studies. In addition to highlighting the need for association studies of complex traits, it was suggested linkage disequilibrium (see also 2.2.5) mapping (Lander, 1996) would also be necessary: in line with these previous theories, studies showed SNPs chosen as markers could not be uniformly spaced across the genome, nor could they be randomly chosen – LD mapping would be necessary to obtain a set of optimally informative markers (Carlson et al., 2004; Gabriel et al., 2002). In addition to the HapMap project, more recent endeavours such as the 1000 Genomes project (Auton et al., 2015) and Haplotype Reference Consortium (McCarthy et al., 2016) provide reference panels for a range of human populations, and can also be used for obtaining informative SNP marker sets. For example, UK Biobank (see 2.3.2.1) phasing and imputation was carried out using 1000 Genomes and Haplotype Reference Consortium data (Bycroft et al., 2018; Marchini, 2015).

It should also be noted that the contributions of non-common-SNP genetic variants to phenotypic variation in a trait are unmeasured in GWAS, and such genetic variants may also contribute to missing heritability (see next section, 2.2.1.1). Rare SNP variants (MAF < 1%) are not assayed in GWAS, and non-point mutations (mutations where alterations involve more than a single nucleotide), such as larger (2 or more bases) insertions and deletions, chromosomal rearrangements such as inversion and translocations, and copy number variants (CNVs) (J. M. Kidd et al., 2008; Lodish et al., 2016; Scherer et al., 2007) are also not investigated (A. J. Clarke & Cooper, 2010; Maher, 2008; Manolio et al., 2009; McCarroll, 2008).

2.2.1.1 The problem of missing heritability in GWASs of complex traits

Heritability is generally defined as the proportion of phenotypic variation in a trait which is due to genetic differences between individuals in a population.

These genetic differences include the effects of alleles in an additive sense, but also include potential inter- and within-loci effects (dominance and epistasis, respectively) - this broad-sense heritability can be calculated using Equation 2.1.

$$H^2 = \frac{V_A + V_D + V_E}{V_P} = \frac{V_G}{V_P}$$

Equation 2. 1: Broad-sense heritability.

In complex traits such as those measured using GWAS, many small-effect genetic variants likely contribute to this heritability. In a GWAS context, heritability in terms of the proportion of phenotypic variance explained by SNPs under an additive model of inheritance is usually calculated – a narrow-sense heritability (Equation 2.2). One method to do this is through linkage-disequilibrium score regression (LDSR, see 2.2.5), with rescaling of the regression slope to give the proportion of variation attributable to the SNPs used in score estimation (h^2_{SNP}).

$$h^2 = \frac{V_A}{V_p}$$

Equation 2. 2: Narrow-sense heritability.

Prior to GWASs, heritability in complex traits was most often estimated through twin and pedigree studies, comparing the phenotypic correlations between relatives where the shared proportion of the genome between them is known (e.g., parent-offspring pairs or trios, full and half-sib pairs, monozygotic versus dizygotic twins). For example, narrow-sense heritability can be estimated from the slope of the regression line between mid-parent phenotypic value and offspring phenotypic value (Visscher et al., 2008), or broad-sense heritability through comparing phenotypic correlations between different sets of relatives (Griffiths et al., 2000).

Missing heritability in GWAS is the heritability that cannot be explained by common SNPs assayed in the analyses (Timpson et al., 2018), or the much lower values of SNP-heritability from GWAS in comparison to estimates of heritability from twin and pedigree studies for the same traits (Maher, 2008; Manolio et al., 2009). For example, estimates of heritability for major depression from twin studies range from 30-40% (Sullivan et al., 2000), but estimates of heritability from GWASs are less than 10% (Howard et al., 2019).

One possible reason for missing heritability in GWASs may be due to not detecting all contributing genetic variation. This can be due to low power to find trait-associated variants, and as power increases with increasing sample size, we may see more of the phenotypic variation in complex traits explained by common genetic variation. Another potential reason we may not be detecting all contributing variants is due to genotyping – only a certain proportion of all SNPs are genotyped, and the contribution of rare variants (MAF < 1%) or other non-SNP variation (e.g., Copy Number Variants, CNVs) is not usually examined (A. J. Clarke & Cooper, 2010; Manolio et al., 2009; Marjoram et al., 2014). Furthermore, the majority of GWASs are carried out using European-ancestry samples – isolated populations and African populations may be enriched for unique variants and contain more genetic variation in general, respectively, and GWASs carried out in these populations may reveal previously undiscovered trait-associated variants (Manolio et al., 2009).

Heritability estimates from twin studies capture non-additive genetic contributions to phenotypic variation (i.e., estimates are of broad-sense heritability), whereas estimates of heritability derived from GWAS assess only additive contributions to phenotypic variation (narrow-sense heritability). For example, estimates from twin studies could be inflated due to common-environment effects (e.g., identical twins more likely to be similarly treated than non-identical twins and pairs of siblings) which generate a gene-by-environment interaction and inflate heritability estimates. Therefore, another way heritability may go missing is in the comparison of narrow-sense (GWAS) and broad-sense (twin or certain pedigree analyses) heritability estimates.

Broad-sense heritability can be higher than narrow sense due to both epistasis and dominance effects, but in addition to this 'legitimate' increase in phenotypic variation explained in comparison to narrow-sense heritability, broad-sense heritability in twin studies can be inflated due to confounding (Hemani et al., 2013). Specifically, variance is generated due to the confounding between common-environment effects and non-additive genetic effects (Evans et al., 2002). In a GWAS, the proportion of variation in phenotype explained can be formalised as the ratio of SNP-heritability (contribution of known variants to phenotypic variation) to total additive genetic contribution to variation in a trait (Equation 2.3) (Zuk et al., 2012).

$$\pi_{explained} = \frac{h^2_{known}}{h^2_{\text{all}}}$$

Equation 2. 3: Proportion of variance explained.

Rather than missing heritability being due to not discovering all possible trait-associated variants (i.e. the numerator is underestimated), the denominator (total narrow-sense heritability) may instead be overestimated (Zuk et al., 2012). While the numerator is directly estimated from the GWAS data, the denominator is estimated indirectly from population parameters, in a way that does not account for the effects of gene x gene interactions on heritability attributable to additive variation.

In contrast, it may be the case for most complex traits that non-additive genetic contribution to phenotypic variation is minimal (W. G. Hill et al., 2008). In this case, missing heritability cannot be explained solely by the comparison of narrow versus broad-sense estimates. It may never be possible to find this missing heritability: Barton argues that missing heritability is "to be expected" (Barton et al., 2017), as SNPs are not perfectly associated with causal alleles (J. Yang et al., 2010), so only the top tail of the distribution is obtainable even if all genomes in all people are assayed (Boyle et al., 2017).

2.2.1.2 Population stratification

Population stratification is the presence of systematic differences in allele frequencies between subpopulations in a population (e.g., human populations from different continents are subpopulations of the global population). This is due to non-random mating, which in turn can be caused by physical barriers to migration and admixture such as distance, or more subtle influences such as selective mating and related factors such as language and country boundaries. Genetic drift, a neutral process by which allele frequencies change with time, then occurs in this subpopulation, and stochasticity and the possible differences in the original 'split' subpopulations results in systematic allele frequency differences.

Stratification in the sample population means that any association between genotype and trait of interest may not be due to a genetic variant's association with the trait of interest (the fundamental question asked in GWAS), but instead

due to the variant segregating at higher or lower frequency in a certain subpopulation. These subpopulations may then be unequally represented between cases or controls, or distribution of quantitative trait values varies between subpopulations.

GWASs are mostly carried out on samples of white Europeans (Chaichoompu et al., 2020). This gap in the field is due to early GWASs being performed almost entirely on unrelated, relatively small, European ancestry samples (Mills & Rahal, 2019; Visscher et al., 2012, 2017): methodology and reference panels were developed with these populations in mind. There is comparatively vast genetic diversity and different haplotype block structures in non-European populations, particularly African populations (Ardlie et al., 2002; Calafell et al., 1998; Peterson et al., 2019; Richter et al., 2017; Rito et al., 2013, 2019; Rosenberg et al., 2002; Schlebusch et al., 2017; J. C. Stephens et al., 2001; Tishkoff et al., 2009). This genetic diversity makes it more difficult to build usable reference panels for these populations relative to white Europeans. Reference panels are needed for genotype imputation, amongst other functions, as part of GWASs.

Using large, admixed, and ancestrally diverse populations in general is also difficult in GWAS not only due to reference panel build issues, but due to population stratification on a larger scale in comparison to populations without admixture, and the issues this presents for standard GWAS analysis. Modelling the extent of fine-scale population structure in genetically diverse populations with a long evolutionary history, such as African populations, can be more computationally inefficient and complex in comparison to modelling population structure in white/ white-European populations. As previously mentioned, GWAS is a regression analysis – in standard regression data are assumed to have an identically and independently distributed property (i.e., all variables share the same underlying probability distribution and so are mutually independent of one another). As sample sizes increase and/or include participants of diverse and/or admixed ancestry, the chances of including related participants (either in terms of familial or ancestral relatedness) increases and this non-independence can generate spurious genetic association results (Peterson et al., 2019; Sul et al., 2018). Analyses in this thesis are carried out on primarily white study participants, partly due to the above considerations and due to the demographic

composition of UK Biobank, Generation Scotland and 23andMe Pfizer datasets (2.3.2) – this limitation is further discussed in 7.5.

Potential inflation of test statistic values, and associated false-positive results, due to population structure, can be mitigated via calculation of a genomic inflation factor, commonly λ (J. Yang, Weedon, et al., 2011). Lambda is largely a function of population stratification, which can be corrected for in a range of ways. Older methods include genomic control (Devlin et al., 2004; Devlin & Roeder, 1999), whereby every association test statistic (i.e. per-SNP) is adjusted by an overall genomic inflation factor – this may not be appropriate as some SNPs differ in terms of allele frequency across ancestral populations more than others – some results will be over-adjusted and others under-adjusted, resulting in loss of power overall. Another method is structured association, where samples are sorted into discrete subpopulation-based clusters, and evidence of association is then assessed on a by-cluster basis (Pritchard et al., 2002; Satten et al., 2002). This method is also flawed, this time due to issues with defining the number of clusters, and inability of the method to allow for membership of more than one cluster. The most widely-used approach for population stratification correction in GWAS is therefore genetic principal component (GPC) analysis-based methods such as EIGENSTRAT (Price et al., 2006), where no prior knowledge of population ancestry is required and underlying stratification is modelled empirically from the genetic data of the sample population. In addition, newer GWAS methods such as BOLT-LMM (Loh et al., 2015), take a Bayesian linear mixed-model approach in order to account for relatedness and cryptic population stratification in GWAS samples.

2.2.1.3 Relatedness and Population Stratification– BOLT-LMM

Linear mixed models allow for both fixed and random effects (https://stats.idre.ucla.edu/other/mult-pkg/introduction-to-linear-mixed-models/; Dean & Nielsen, 2007), effectively allowing for hierarchical structure within sample data. Hierarchical structure in sample data means there are 'levels' to the data e.g., individuals make up the sample, but an added level is that these individuals are related in family groups, or are students sampled from different classrooms, or individuals from different geographic locations. Observations per-individual are likely to be non-independent as these groupings

may make individuals more similar to one another within groups compared to between groups. To find a true estimate of the relationship between per-individual observations and an outcome, groupings must be taken into account in any model. Non-independence and correlation can also occur in genetic data, due to individuals belonging to groups i.e., presence of population stratification and/or cryptic relatedness within a sample.

One linear-mixed model approach used to account for population stratification and cryptic relatedness in GWAS samples is BOLT-LMM (Loh et al., 2015). In contrast to traditional GWAS where related individuals are removed from the sample and GPCs are added as covariates to the GWAS model, BOLT-LMM incorporates a genetic relatedness matrix (GRM) into the GWAS model, allowing related individuals to remain in the sample while still adjusting for stratification and relatedness.

Mixed model approaches in general are gaining traction in association studies particularly in terms of investigating non-infinitesimal traits (Loh et al., 2015; Sul et al., 2018), but BOLT-LMM is one of the most computationally efficient, and additionally allows for modelling of both infinitesimal and non-infinitesimal trait architectures directly.

2.2.2 Multiple-testing correction in a GWAS context

As a GWAS is essentially the process of running millions of regressions of SNP genotype on trait value, correction for multiple comparisons is vital. Due to linkage disequilibrium some genotyped variants are inherited together more often than expected by chance, making the number of independent tests smaller than the number of genotyped SNPs tested in the GWAS. The standard practice for multiple comparison correction in GWAS is Bonferroni correction, giving a genome-wide significance alpha value of 5×10^{-8} i.e., the nominal alpha value of 0.05 is divided by 1 million (an estimate of the number of independent tests).

2.2.3 Conditional False Discovery Rate Analyses

As explained in the previous section, multiple testing correction is of great importance in a GWAS context, with Bonferroni correction the standard practice. Bonferroni correction is generally considered very conservative, a quality which some argue may make it less than ideal in the context of a GWAS where the aim

is discovery of new trait-associated genetic variation rather than testing of a pre-set hypothesis, and the number of tests is extremely large. Bonferroni correction is a method for controlling the family-wise error rate – specifically, if we set a significance level according to this procedure, we specify the probability of concluding at least one false positive result out of the entire set of tests we are carrying out.

An alternative set of multiple-testing corrections fall instead under the umbrella of false discovery rate (FDR)-controlling procedures. The false discovery rate is the rate or proportion of type 1 errors amongst a set of tests – in contrast to Bonferroni correction, FDR-controlling procedures do not control the family-wise error rate (chance of at least one type 1 error amongst a set of tests), providing a less stringent, but more powerful approach. The tail-end FDR procedure is concerned with controlling the FDR at a pre-defined level, and deciding the maximum test statistic value from a list of ordered test statistic values which allows for this (Benjamini & Hochberg, 1995), which then becomes the new cut-off value for deciding significance. Local FDR reframes the FDR as a Bayesian posterior probability that the SNP in question is not associated with the disease or trait, given its association test statistic (usually a p value) (Benjamini & Hochberg, 1995; Storey, 2002). Conditional FDR then simply extends local FDR analysis and incorporates association data for a second, genetically correlated trait, to ask 'what is the posterior probability that the SNP in question is not associated with trait 1 given its association test statistics for both trait 1 and trait 2' (Equation 2.4). This is equivalent to adjusting each association test statistic (p value) for trait 1 by an empirical conditional probability value, which can be calculated by finding the proportion of instances where the two conditions $p_i \leq P_i$ *and* $p_j \leq P_j$ are true.

$$cFDR = \Pr\left(H_{0(i)} \middle| p_i \leq P_i, p_j \leq P_j\right) = \frac{p_i}{\Pr(p_i \leq P_i \,|\, p_j \leq P_j)}$$

Equation 2. 4: Conditional false discovery rate.

cFDR analyses have been used to find novel variants associated with schizophrenia, type 2 diabetes, Alzheimer disease, bipolar disorder and systolic blood pressure (Andreassen et al., 2014; Andreassen, Djurovic, et al., 2013; Andreassen, Thompson, et al., 2013; Wang et al., 2016). This therefore represents a promising and potentially more cost-effective method for identifying new SNPs associated with complex traits by maximising the utility of existing GWAS outputs.

2.2.4 BUHMBOX

BUHMBOX (Breaking Up Heterogeneous Mixture Based on cross(X)-locus correlations) is based on the principle that if clinical heterogeneity were present (a subset of disease A cases are mis-diagnosed disease B cases), disease B risk variant (allele) frequencies will be higher only within a subset of disease A cases (Han et al., 2016), and under 'true' or whole-group pleiotropy, disease B risk alleles will be found at higher allele frequencies in all disease A cases in the sample. In addition, under true whole-group pleiotropy, the expected correlation between risk allele dosages at different loci should be "consistently positive" (Han et al., 2016). These between-loci pairwise correlations are combined into a single BUHMBOX statistic, which tests for excessive positive correlations. This test, and thus its statistic, will be significant in the case of heterogeneity, and non-significant in the cases of whole-group pleiotropy (lack of true heterogeneity) or insufficient power.

The statistic itself is calculated in several steps. Genotype data in a sample of disease A cases and controls is assembled, along with information about SNPs associated with disease B (risk allele, risk allele frequency and effect size (measured as or converted to odds ratio)). A set of SNPs is compiled where all SNPs are associated with disease or trait A at $p < 10^{-4}$ and are pruned in controls by excluding SNPs with $r^2 > 0.1$. SNPs with an info score of < 0.8, MAF < 0.01 and HWE test p value of $< 10^{-6}$ are also excluded. Genetic principal components are regressed out from risk allele dosages to give residual dosages for each individual locus. Individuals without complete information on SNP rsID, risk allele and dosage, risk allele frequency and effect size are excluded. A correlation matrix, R, of residual risk-allele dosages in N cases of disease A is constructed, along with a correlation matrix R' of risk-allele dosages in N' controls. These

matrices are then used to calculate Y (Equation 2.5), a matrix where non-diagonal elements are z scores from delta correlations, where a delta correlation is the relative increase in correlation between risk allele dosages at different loci in cases compared to controls (Han et al., 2016).

$$Y = \sqrt{\frac{N * N'}{N + N'}}(R - R')$$

Equation 2. 5: Y matrix for BUHMBOX calculations.

The BUHMBOX statistic is then calculated according to Equation 2.6 using the matrix generated by Equation 2.5, where y_{ij} in Equation 2.6 is the element in Y row *I* column *j*. w_{ij} refers to a weighting function designed to maximise power, discussed in detail in the BUHMBOX method paper Supplementary Note (Han et al., 2016), and utilising risk allele frequency and allele-disease association OR values.

$$S_{BUHMBOX} = \frac{\sum_{i<j} w_{ij} y_{ij}}{\sum_{i<j} w_{ij}^2}$$

Equation 2. 6: BUHMBOX test statistic.

A p value is calculated using Equation 2.7, where φ is the cumulative density function of the standard normal distribution.

$$P_{BUHMBOX} = 1 - \varphi(S_{BUHMBOX})$$

Equation 2. 7: P value for BUHMBOX test statistic.

Population stratification and linkage disequilibrium could lead to a false positive result of the BUHMBOX test. Population stratification is addressed through

regressing out genetic principal components (GPCs), as part of the calculation of delta-correlations (Han et al., 2016). Linkage disequilibrium is adjusted for through LD pruning and examining delta-correlations (Han et al., 2016).

Insufficient power may lead to potential false insignificance or false negative result from BUHMBOX analysis. Insufficient power can result when the number of disease A cases is too small, heterogeneity proportion is too low, and the number of known risk alleles and/or their effect sizes are low. Through simulation Han et al showed that high power (approaching 100%) at moderate suspected true heterogeneity proportions (0.2) can be achieved if the number of risk loci used in analyses is 100 or greater, and when the number of individuals with the disease (case individuals) is greater than 2,000 (Han et al., 2016).

2.2.5 Linkage-Disequilibrium Score Regression

Linkage disequilibrium (LD) is a property of genetic variants, namely alleles at different loci, whereby they are inherited together more often than is expected by chance (Pritchard & Przeworki, 2001).

LD can be measured in a range of ways (Devlin & Risch, 1995), most commonly between pairs of genetic markers (and mostly using an r^2 estimate; see Equation 2.8 below which gives r. Note the numerator is equal to 'D', another common LD measure, and the denominator can be written $(p1p2q1q2)^{1/2}$). p1 is the frequency of allele 1 at a biallelic locus SNP 1, q1 is the frequency of allele 2 at SNP1, q2 is the frequency of allele 2 at SNP 2, and p2 is the frequency of allele 1 at SNP 2. r^2 can then be used to prune out SNPs correlated (in LD) at an undesirable level e.g., SNPs at $r^2 > 0.1$ (10%). Note that r^2 is a preferred measure of LD rather than D, as r^2 correctly accounts for differences in allele frequencies at loci being compared.

$$r = \frac{\pi_{11}\pi_{22} - \pi_{12}\pi_{21}}{(\pi_1 + \pi_2 + \pi_{+1}\pi_{+2})^{1/2}}$$

Equation 2. 8: Linkage disequilibrium estimate (r).

LD can be caused by a range of factors, including those in the evolutionary history of the population in which the variants segregate, as well as influences at the molecular level. Formally, r^2 is a function of a scaled recombination rate parameter, or ρ (Equation 2.9) (Pritchard & Przeworki, 2001).

$$\rho = 4N_e c$$

Equation 2. 9: Recombination rate.

Where *c* is the rate of recombination between the two markers and N_e is the effective population size. Genetic recombination refers to the rearrangement of DNA sequences and its consequences (Alberts, Johnson & Lewis, 2002; Carroll, 2001; Heyer et al., 2010). Where effective population size and recombination rate are relatively large, r^2 is inversely proportional to ρ i.e., LD between two markers decreases with increasing recombination. Recombination rate between markers tends to increase with increasing physical distance, and recombination rate in general varies across the human genome (Altshuler et al., 2010; Kong et al., 2002; Y. Liu et al., 2017; Myers et al., 2005; Stapley et al., 2017). Mutation rates also tend to vary across the human genome (Casane et al., 1997; Nachman & Crowell, 2000; Smith et al., 2002; K. H. Wolfe et al., 1989), and as high mutation rates break down LD between loci and nearby markers this also affects the degree of LD between markers. Also, at the molecular level, gene conversion can lead to decrease or breakdown in LD – gene conversion is the swapping of short sections of chromosomes between copies of a chromosome pair (i.e., is a form of non-reciprocal recombination), and effect on LD is equivalent to two recombination events in close proximity (and so acts on LD in similar fashion to recombination in general).

Higher-order influences such as genetic drift can also influence levels of LD – genetic drift is the change in allele frequencies from one generation to the next due to random sampling without replacement in a finite population (Kimura, 1954; Masel, 2011; S. Wright, 1937). In small (finite) populations drift can lead to general loss of haplotypes over time and lead to an increase in LD (Charlesworth, 2009; Star & Spencer, 2013). Conversely, rapid population grown

can reduce genetic drift and so reduce levels of LD. Admixture, or migration, can generate LD in a population as gene flow temporarily results in long haplotype blocks (Darvasi & Shifman, 2005; Smith & O'Brien, 2005). Natural selection acting on linked variants can also lead to increase in LD through genetic hitchhiking (Smith & Haigh, 1974), as a haplotype flanking an advantageous variant is swept to high frequency or potentially to fixation in a population – negative selection can remove regions linked to deleterious variants, also inflating LD.

Linkage disequilibrium score regression (LDSR) is a widely used method making use of linkage disequilibrium and its relationship with GWAS test statistics in order to quantify genetic correlation between traits, and to differentiate between population stratification and polygenicity in the inflation of association statistics in GWAS data estimated by the lambda value (Bulik-Sullivan et al., 2015).

The LD score is a measure of the amount of genetic variation tagged by variant *j* (Equation 2.10), calculated as the sum across *k* individuals included in the reference panel (see Supplementary Note in (Bulik-Sullivan et al., 2015) for full derivation).

$$l_j = \sum_k r_{jk}^2$$

Equation 2. 10: Amount of genetic variation tagged by variant j.

Variants in LD with a causal variant for a trait display an elevation in their test statistic in a GWAS, and this elevation is proportional to the degree of LD with the causal variant (Pritchard & Przeworki, 2001; J. Yang, Weedon, et al., 2011). Additionally, inflation in test statistics caused by population stratification does not correlate with LD (Devlin & Roeder, 1999; Lin & Sullivan, 2009; Voight & Pritchard, 2005). The expected value of the test statistic from GWAS for a variant *j* therefore depends on sample size (N), SNP-heritability (h^2), number of markers included in the calculation (M), the contribution of confounding biases such as cryptic relatedness and population stratification (a), and the LD score of the variant l_j.

$$E\left[x^2|l_j\right] = \frac{Nh^2 l_j}{M} + Na + 1$$

Equation 2. 11: Expected value for GWAS test statistic associated with variant j.

The intercept of the regression of the test statistic on LD score (Equation 2.11), minus one, provides a measure of test statistic inflation and an indication of whether this is due to stratification or polygenicity. Therefore, the closer the intercept value to 1, the lower the contribution of confounding bias (such as stratification) to inflation of GWAS test statistics, and an intercept value including 1 indicates no significant inflation in test statistics due to these influences.

To obtain a genetic correlation value for two traits, LDSR can be extended (cross-trait LDSR), replacing the chi-squared test statistic of a single study with the product of two z scores calculated from GWAS effect sizes (beta values) for two separate traits (Bulik-sullivan et al., 2015) (Equation 2.12). N_1 and N_2 refer to the sample size for each of the two traits being compared, M to the number of markers included in the calculation, N_s to the number of overlapping samples (individuals included in GWAS for both trait 1 and 2), $e_{g.}$, to genetic covariance between the two traits, e to phenotypic correlation among the N_s overlapping samples.

$$E\left[z_{1j} z_{2j} |\ l_j\right] = \frac{\sqrt{N_1 N_2} e_g}{M} l_j + \frac{e N_s}{\sqrt{N_1 N_2}}$$

Equation 2. 12: Expected value of cross-trait product of GWAS z scores.

A genetic covariance value between the two traits can then be obtained by regression of this z score product on LD score, and normalisation of this covariance value by square root of the product of the SNP-heritabilities for the corresponding studies gives a genetic correlation value between the two traits of

interest (Bulik-sullivan et al., 2015). As any sample overlap inflates the z score product and then influences the intercept (term to the right-hand side of '+' in Equation 2.12) rather than the slope value for the regression, genetic correlation as calculated by cross-trait LDSR is not biased by sample overlap.

2.2.6 Polygenic Risk Scoring

An individual's burden of risk (trait-associated) alleles can be quantified by calculating their polygenic risk score (PRS). This can be calculated as a simple sum of independent trait-associated and trait-increasing alleles present in the individual, determined from GWAS output, or as a weighted sum where each trait-associated SNP included in the score is weighted by its effect size value (Chatterjee et al., 2016; Dudbridge, 2013). Independence of trait-associated SNPs contributing to a PRS is ensured via LD-based pruning, which can be carried out using tools such as PLINK (S. Purcell et al., 2007). To avoid over-fitting and over-estimation of the predictive accuracy of PRSs, the cohort from which the score is constructed (i.e. discovery cohort/ sample or training data) should be independent from the cohort in which PRSs are calculated and analyses are performed (target sample) (S. W. Choi et al., 2020; Wray et al., 2013).

Significance thresholds to use in PRS construction have been contested – in the case of complex traits and considering the infinitesimal model (Barton et al., 2017), it may in fact be more powerful to include variants associated with the trait at much lower than traditional genome-wide significance thresholds (Wray et al., 2013), and this approach (using all SNPs whether significantly associated or otherwise) is commonly used in animal and plant breeding applications (Erbe et al., 2012; Hayes et al., 2009; Meuwissen et al., 2001). Several purpose-built statistical tools have been constructed for PRS analyses, one of which, PRSice (Euesden et al., 2015), calculates the 'optimum' PRS from a range of PRSs with varying variant inclusion thresholds based on maximising Nagelkerke R^2 value (a measure of predictive value of a model and a quantification of amount of variation explained).

PRSs can be included in regression models and used to validate GWAS results, by testing whether a PRS for a trait is significantly associated with that trait in an independent cohort. In this case population stratification must be accounted for by covarying for genetic principal components (similarly to GWAS analyses), or

other measures of substructure (underlying stratification) in genetic data such as Multidimensional Scaling (MDS) components. PRS analysis can also be used to investigate shared genetic factors between traits in a similar fashion to genetic correlation – should a PRS for one trait be significantly associated with a different trait in an independent cohort, this suggests shared genetic factors underlie the two traits. PRSs have been valuable in a clinical setting in some cases and may inform treatment or allow stratification of patients by genetic risk for disease – examples of diseases where this is true include coronary heart disease and certain cancers (reviewed in (Chatterjee et al., 2016; Torkamani et al., 2018). In psychiatry, the clinical utility of PRSs is less clear, but there may be potential for PRS use in diagnosing of individuals whose symptoms meet multiple diagnostic criteria (Ruderfer et al., 2018), and perhaps eventually for prediction of illness and to inform treatment, as seen for some physical diseases, although this is in its infancy, particularly for psychiatric disorders with lower heritability in comparison to more highly heritable disorders such as schizophrenia, such as MDD (Binder, 2019).

2.2.7 Mendelian Randomisation

The first outlining of the principles of Mendelian randomisation, the “natural randomised control trial” (Smith & Ebrahim, 2005; Smith & Hemani, 2014) (Fig 2.1) framework, is attributed to Katan (Katan, 1986). The basic premise of MR is that the causal effect of an exposure on an outcome can be estimated through division of the regression coefficients from the regression of the outcome on the instrument by the regression coefficient of the regression of the exposure on the instrument.

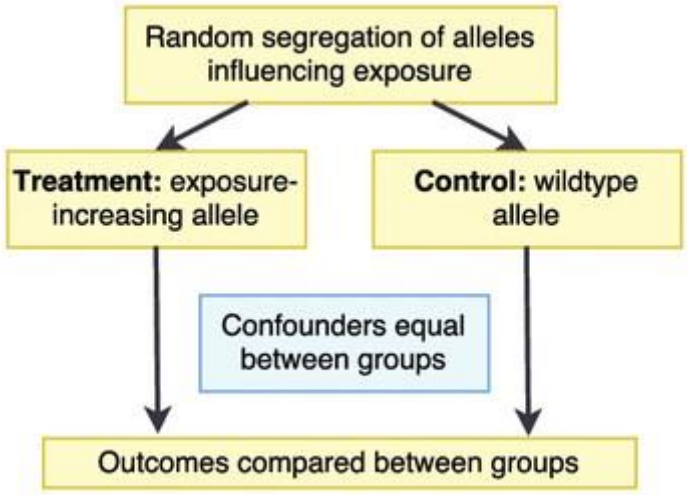

Figure 2. 1: MR as a natural randomised control trial.

Adapted from: https://www.researchgate.net/figure/Principles-and-assumptions-behind-Mendelian-randomization-A-Diagram-illustrating-the_fig1_325460102

In the early 2000s MR gained traction in the context of genetic and observational epidemiology (Brown, 2003; Keavney et al., 2006; Smith & Ebrahim, 2003). Even during these early stages, the potential problems with Mendelian Randomisation such as pleiotropy, gene-environment interactions, gene-gene interactions and population stratification were recognised (Thomas & Conti, 2001a). Also recognised was the potential insight MR could provide into causal relationships when only observational/ cross-sectional data were available. Use of aspects of the genotype as instrumental variables meant that reverse causation issues are avoided, as the genotype is generated prior to experience of both the exposure and the outcome, and germline genotype is unaltered by exposures and outcomes. Regression dilution bias, whereby errors in measurement of the independent variable cause the regression slope to be biased towards zero (Hutcheon et al., 2010), is also avoided as the genetic variants associated with the exposure tend to remain associated to the same degree throughout the life course (Smith & Ebrahim, 2004), mitigating random measurement error in measurement of the exposure variable. To a degree, issues with confounding can also be avoided if genetic variants used as instruments are unrelated to factors that confound exposure and outcome such as socioeconomic status (Lawlor et al., 2008) (see below for further discussion of this in the context of increasingly complex exposures).

However, this is only true if the instrument-outcome-exposure relationship adheres to a list of specific assumptions. This allows for investigation of the exposure-outcome relationship in a way that is analogous to a randomised control trial (RCT), as the 'participants' are 'dosed' with the exposure (measured via instrumental variable(s)) at conception, and this dosage is randomised according to Mendel's second law (the law of independent assortment) (Mendel, 1866). The assumptions that allow causal effect estimation are as follows (Lawlor et al., 2008) (Fig 2.2);

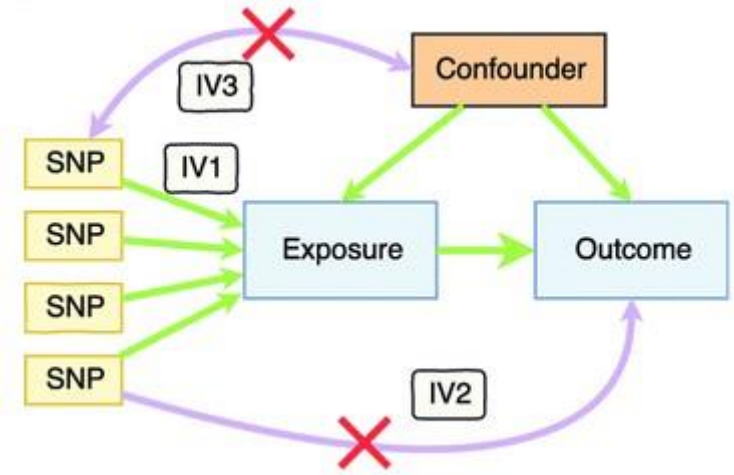

Figure 2. 2: MR assumptions.

IV1 = instrumental variable assumption 1, IV2 = instrumental variable assumption 2, IV3 = instrumental variable assumption 3. Adapted from: https://www.researchgate.net/figure/Principles-and-assumptions-behind-Mendelian-randomization-A-Diagram-illustrating-the_fig1_325460102

1. the instrument is associated with the exposure (IV1)
2. the instrument affects the outcome only via the exposure (IV2)
3. the instrument is not associated with any confounders of the exposure-outcome relationship. (IV3)

If there is only one IV, the simple ratio of regression coefficients described above can be used as-is, and this is the Wald ratio method. However, in most cases, there will be more than one IV. This is because in the cases of MR analysis of complex traits, instruments are genetic variants and are commonly chosen from GWAS (Burgess et al., 2017). Here is where the problems of population stratification, pleiotropy, gene-gene and gene-environment interactions, as envisioned in the early 2000s (Thomas & Conti, 2001b) become apparent. As the

number of IVs increases so does the chance that one or more will be associated with a second trait other than the exposure (pleiotropy). It also becomes more likely that an IV will be associated with another genetic factor (G) in addition to an environmental factor (E) – if G and E are independently distributed and the relationship between exposure and outcome is linear, this G x E interaction may not be a problem, but with non-independent distribution of G and E (where G is associated with likelihood of exposure to E) and non-linear relationships between exposure and outcome, this can result in both false positive and false negative results in MR analyses. If an IV is involved in a gene-gene interaction, this would produce similar results to G x E in MR analyses. Population stratification in GWAS introduces "distortion" of estimates of association between genetic variants and traits, which can then bias results of MR analyses through introduction of confounding between genetic variants and trait values.

There are two main branches of MR analysis, depending on whether the researcher has access to individual-level genetic information, or just to genome-wide association study (GWAS) summary statistics. In the former, associations between genetic variants (instruments) and exposures are measured in the same dataset as the measurement of instrument-outcome associations. In the latter, two independent GWAS summary statistic datasets are used, one for instrument-outcome association measurement, and one for instrument-exposure measurement. Discussion below will focus on two-sample MR (where two independent GWAS summary statistic datasets are used) which is used in analyses in this thesis, but one-sample MR (and the relative merits of one versus two-sample MR) is summarised in detail elsewhere (Haycock et al., 2016; Lawlor, 2016; Smith & Ebrahim, 2003).

In the context of MR, the derivation of instruments from GWAS can be problematic for three main reasons. Firstly, since associated variants are likely to have small effects on the variation of the exposure, they may be weak as instruments (and so may not meet assumption 1). This can result in weak instrument bias, which is when the causal estimate tends towards the confounded observed estimate between exposure and outcome. In some situations, inclusion of a greater number of instruments can increase power – this is not true if many, or all, of the instruments are weak, and is known as 'many weak instruments' bias (Bound et al., 1995).

The decision as to what constitutes robust association in an instrument is less straightforward for complex, heterogeneous exposures. APOE genotype and serum cholesterol have a clear relationship, and intuitively genotype makes a good 'stand-in' (instrument) for serum cholesterol level (Katan, 1986). For exposures such as BMI, employment status or MDD the genetic variants found to be associated with the traits through GWAS will have negligible effect sizes, and the 'threshold' for choosing certain variants rather than others is tricky to define. Genome-wide significance may be chosen as a threshold, but in the case of MDD this provides potentially over a hundred variants as the starting pool of instruments – and again each will only contribute to a tiny proportion of the variance in the exposure. Many algorithms exist to prioritise and rank GWAS SNP associations according to predicted functional consequence (de Leeuw et al., 2015; McLaren et al., 2010, 2016), but again the relationship between predicted functional consequence of a SNP-change and the end-point of variation in the exposure trait value is not clear and many variants may have relevant functional annotation (or conversely, it may be that none of the trait-associated variants have relevant functional annotation).

Secondly, pleiotropy is of great concern. Biological a.k.a. horizontal pleiotropy could mean that assumptions 2 and 3 are violated, as the variant may be associated indirectly with confounders and/or directly with the outcome (reviewed by Hemani, Bowden, & Smith, 2018). Furthermore, chances of pleiotropy are increased with increasing number of associated variants. Many hundreds of SNPs are associated with many complex traits at genome-wide significance, as sample sizes and variant-discovery power of GWASs increase. For example in a recent analysis of ~0.8 million individuals over 100 variants were found to be associated with MDD at genome-wide significance (Howard, Adams, Clarke, et al., 2018). It is impossible to empirically test MR assumptions 2 and 3 – not all possible confounders are known, and their possible association with instrument(s) is not assessed – it is likely, due to pleiotropy, that each instrument is associated with at least one confounder or the outcome.

Thirdly, there may be extensive measurement error in a GWAS, depending on factors such as sample size. Both the SNP-exposure association and the SNP-outcome association may be measured with considerable error, depending on

study size and design (particularly case-control versus continuous/ quantitative-trait association analyses) (Liao et al., 2014).

2.2.7.1 Pleiotropy in Mendelian Randomisation Analyses

As discussed directly above, pleiotropy may render an instrument invalid. In complex traits this is an issue because pleiotropy is widespread and likely to be unavoidable when choosing instruments (Sivakumaran et al., 2011; Solovieff et al., 2013; Timpson et al., 2018; Visscher & Yang, 2016). Considering this, MR methodology has been developed to account and correct for pleiotropy amongst instruments. Two methods for dealing with pleiotropic instruments in MR, Inverse-Variance Weighted MR and MR-Egger, are conceptually based upon dealing with heterogeneity in estimates derived from meta-analyses, and small-study bias in meta-analysis respectively. In contrast MR-RAPS is based on errors-in-variables regression models.

2.2.7.2 Inverse-variance weighted (IVW) MR

In IVW analyses the Wald ratio estimates of the causal effect of the exposure on the outcome, obtained for each individual instrument, are essentially weighted and combined in a fixed-effect meta-analysis model to obtain an overall estimate of causal effect of exposure on outcome. This can be visualised as a line of best-fit passing through causal estimates plotted on an instrument-outcome versus instrument-exposure coefficient plot: in IVW, this line is constrained to pass through the origin (Bowden et al., 2015). It is assumed that heterogeneity in causal estimate values across instruments is due to horizontal pleiotropy in at least one or more instruments, but could be due to a range of issues that lead to model assumptions not being met (Hemani et al., 2018).

An adaptation of Cochran's Q statistic can be used to quantify and statistically test the significance of this heterogeneity-indicated horizontal pleiotropy (Bowden et al., 2017, 2019; Burgess et al., 2013). Q, derived from the IVW estimate, should follow a chi-squared distribution with degrees of freedom equal to the number of SNP instruments minus 1, and significant departure indicates heterogeneity (and so potential horizontal pleiotropy). Additionally, a measure of instrument strength in IVW MR analyses is the F-statistic (Bowden et al., 2016, 2017, 2019; Burgess et al., 2011; Staiger & Stock, 1997).

2.2.7.3 MR-Egger

MR-Egger is another MR method, and is based on the fact that small-study bias in meta-analysis can be visualised by plotting the precision associated with study estimate against estimates themselves in a funnel plot (Egger et al., 1997). Unlike IVW MR analyses which indicates presence of general horizontal pleiotropy whether balanced or directional, MR-Egger detects directional horizontal pleiotropy specifically. Directional pleiotropy is when pleiotropic effects of genetic variants are not balanced around the null (zero), but tend to be in the same direction (trait increasing or trait decreasing) across different traits (Bowden et al., 2015).

In MR analysis, directional pleiotropy can be considered a kind of small-study bias, with each SNP instrument representing a 'study', with asymmetry in a plot of 'precision' (size of the association between instrument and exposure) against the causal estimate for that instrument indicating directional pleiotropy. If the intercept in MR-Egger is significantly different from zero, this indicates directional pleiotropy is present amongst instruments. Note that balanced horizontal pleiotropy (i.e., where effect direction is heterogeneous amongst individual estimates, effectively cancelling out overall) would not be detected. Analogous to the F-statistic in IVW analysis, a version of the I^2 statistic (Higgins et al., 2003) termed I^2_{GX} (Bowden et al., 2016) can be calculated in MR-Egger analysis to give an estimate of instrument strength. I^2_{GX} can range from 0 to 1 and quantifies the degree of bias (or dilution) of the causal estimate obtained from MR-Egger due to measurement error in SNP-exposure association values.

Overall, MR-Egger and/or IVW, or other MR analyses, can be done in tandem to further interrogate causal estimates obtained via MR, and attempt to identify and adjust for the presence of horizontal pleiotropy amongst instruments. Multiple approaches can be used to further understand the most prominent type of horizontal pleiotropy present e.g., IVW and MR-Egger to assess for presence of horizontal pleiotropy generally and directional pleiotropy, respectively.

2.2.7.4 MR-RAPS

An alternative methodology treats potential violations of IV assumptions an 'errors in variables regression' problem framework, in contrast to meta-

analysing effect estimates from individual instruments. Errors-in-variables models account for error in measurement of the independent variables (R. J. Carroll, 2006)– in standard regression (and so in IVW, MR-Egger and similar methods) it is assumed that independent variables are measured without error and these models therefore only account for errors in dependent (outcome) variables, and measurement error in SNP-exposure association is assessed separately through measuring e.g. I^2_{GX} and F. MR-RAPS (Robust Adjusted Profile Score) adjusts the profile likelihood of the summary data (Zhao et al., 2020).

The effect of an exposure on an outcome is modelled as an errors-in-variables regression (Equation 2.13).

$$\Gamma_j \approx \beta_0 \gamma_j$$

Equation 2. 13: Errors-in-variables regression.

Where Γ_j is the association between instrument *j* and the outcome, and γ_j is the association between instrument *j* and the exposure, and β_0 gives an estimate of the causal effect of exposure on outcome. For MR-RAPS analysis, first a log-likelihood function of the summary data is obtained (Equation 2.14). This is the natural log transformation of the likelihood function of the summary data, where the likelihood function measures goodness-of-fit of the errors-in-variables regression model given the values of model parameters.

$$l(\beta, \gamma_j \dots, \gamma_p) = -\frac{1}{2}\left[\sum_{j=1}^{\rho} \frac{\left(\hat{\gamma}_j - \gamma_j\right)^2}{\sigma_{x_j}^2} + \sum_{j=1}^{p} \frac{\left(\hat{\Gamma}_j - \gamma_j\beta\right)^2}{\sigma_{y_j}^2}\right]$$

Equation 2. 14: Log-likelihood function of the summary data.

'Profiling out' of nuisance parameters (γ_j) from the log-likelihood function gives the profile score (profile log-likelihood of β) (Equation 2.15) (Zhao et al., 2020).

$$l(\beta) = -\frac{1}{2}\sum_{j=1}^{\rho}\frac{\left(\hat{\Gamma}_j - \beta\hat{\gamma}_j\right)^2}{\sigma_{x_j}^2 B^2 + \sigma_{y_j}^2}$$

Equation 2. 15: Profile score.

A maximum likelihood estimator of β is given by $\hat{\beta} = \arg\max_{\beta} l(\beta)$. Briefly, maximum likelihood estimation is the estimation of model parameters of a function (here, regression of exposure on outcome) via maximizing the likelihood function (of β) given the data x so the data are most probable under the assumed statistical model.

Zhao et al showed the relationship between exposure and outcome deviates from the linear relationship described above due to systematic pleiotropy (almost all instruments show horizontal pleiotropy), and this can be modelled under a random-effects model.

When a profile score is calculated according to this model, it is biased (does not have mean zero at the true value). Inflation in the variance of Γ (due to systematic horizontal pleiotropy in instruments) is described by the unknown additive constant τ_0^2, and as a result the profile log-likelihood and one of the associated profile scores has a corresponding maximum likelihood estimator that is not statistically consistent. In order to correct for this bias (and so effectively model systematic pleiotropy), the profile score is modified ('adjusted') ((McCullagh & Tibshirani, 1990), see also Zhao et al., 2020 section 4.2).

In addition to systematic pleiotropy, idiosyncratic pleiotropy (horizontal pleiotropy of a single instrument or small subset of instruments) can mean even an adjusted profile score will not be able to deliver the best causal estimate – this idiosyncratic pleiotropy is indicated by outliers on diagnostic plots of the adjusted profile score estimator (QQ plots and leave-one-out versus instruments strength plots). To mitigate the effects of idiosyncratic pleiotropy on the adjusted profile score estimator of the causal estimate, the adjusted profile score can be made robust, through robust regression techniques first developed by Huber (Huber, 1964). This involves changing the l2-loss in the profile

likelihood to a robust loss function, either the Huber or the Tukey biweight loss function.

Overall, using a robust adjusted profile score to estimate the causal effect is based on a model that is most likely to match underlying instrument (SNP) biology i.e., widespread pleiotropy in variants associated with in complex traits. MR-RAPS allows estimation of a causal estimate in scenarios where both systematic pleiotropy (most or all instruments are pleiotropic) and idiosyncratic pleiotropy (a small subset or single instrument(s) are/ is pleiotropic) are present, and this can be explicitly modelled. Additional added benefits of MR-RAPS include the fact that inclusion of additional weak instruments (e.g., associated SNPs at less than genome-wide significance) can improve accuracy of the causal estimate, and that this type of in-depth statistical correction is usually only possible with access to individual-level data (and through MR-RAPS is possible with summary statistics).

In addition to IVW, MR-Egger and MR-RAPS described above and used in analyses described in later chapters of this thesis, a wide range of other MR approaches are also in common usage (Bowden et al., 2017; Burgess et al., 2017; Evans & Smith, 2015; Smith & Hemani, 2014; Zheng, Baird, et al., 2017), many also developed with respect to specific challenges of two-sample MR with multiple instruments derived from GWASs.

2.2.7.5 Summary Statistics & Methodological Issues in Two-Sample MR

If harmonisation of GWAS summary statistics for two-sample MR approaches is not carried out correctly, causal estimates can be wrong (reviewed by Hartwig, Davies, Hemani, & Smith, 2016). Harmonisation can be summarised as 4 main steps;

1. merging of the GWAS datasets (one for exposure, one for outcome)
2. choosing a subset of SNPs from the merged dataset
3. matching the effect alleles in exposure and outcome GWAS datasets
4. linkage-disequilibrium pruning

In step 1, SNPs must be present in both GWAS datasets and have no missing allele information. SNPs must also be reported on the same strand of DNA in both GWAS datasets. For example, a SNP may be read as A/G if reported on the

forward strand, and T/C if reported on the reverse strand. If there are strand discrepancies between GWAS datasets, labelling can be easily converted (e.g., A/G ➔ T/C).

In step 2 a subset is chosen via a significance threshold of the researcher's choosing – this may be genome-wide significance, nominal significance, or another threshold (in relation to the SNP-exposure *p*-value). It is also ensured that the effect allele (allele for which the association beta or OR is reported) is exposure-increasing. If the effect (beta) value is less than zero (not exposure increasing), effect allele is swapped with non-effect in the exposure GWAS dataset, and the beta value is multiplied by -1. Effect alleles are then matched between exposure and outcome datasets, 'flipping' alleles in the outcome dataset where necessary & possible (step 3).

Finally, LD pruning is carried out, r^2 threshold depending on the type of MR analysis to follow (e.g., a PLINK default r^2 threshold of 0.2 is acceptable for MR-Egger, but $r^2 < 0.01$ is required for MR-RAPS). This results in a set of independent instruments, ready for two-sample MR analyses.

Selection bias can be avoided by selecting instruments (based on, for example, *p*-value of association with the exposure) in a third, independent dataset. Selection bias occurs if genetic instruments influence the likelihood of taking part in a study or participating fully in follow-up. As an example, to mitigate this, if the MR analysis was to be carried out with BMI as an exposure and MDD as an outcome, two independent GWAS summary statistic datasets would be used, one for BMI and one for MDD, with a third independent dataset of an entirely separate GWAS of BMI used for instrument selection initially – if SNPs are associated with BMI in two independent GWASs, this suggests the association is true rather than driven by an association between SNP and likelihood of participating in a particular study. However, selection bias may be of less concern in large, general-population cohorts such as UK Biobank in comparison to cohorts where participants are recruited from hospital or general practice settings, or specifically based upon a condition of interest.

These methodological stumbling blocks relating to harmonisation errors can result in discordant results between two-sample MR analyses even when using the same GWAS datasets (Hartwig et al., 2016). In two independent MR analyses

of C-reactive protein and schizophrenia, Prins et al found a protective (negative) causal effect of CRP on schizophrenia (Prins et al., 2016), whereas Inoshita et al found a positive causal effect of CRP on schizophrenia, with the latter analysis since retracted as the results were likely biased due to harmonisation issues (Hartwig et al., 2016).

2.3 Resources and Materials

2.3.1 FUMA and analyses therein (MAGMA, GTEx)

FUMA is an integrative, open-access web platform used for the annotation, prioritisation, visualisation and interpretation of GWAS results , with GWAS summary statistics as input (Watanabe et al., 2017). A range of the available tools within FUMA have been utilised in analyses in this thesis; MAGMA and GTEx are of importance and summarised below.

2.3.1.1 MAGMA

In a GWAS, association between SNPs and a trait of interest is tested for. GWAS output (summary statistics) can be further characterised at the gene level, to investigate genes and gene 'sets' (functional groupings of genes) which are significantly associated with the trait of interest. In gene a.k.a. gene-based or gene-level testing, effects of variants are aggregated at the gene level – SNPs within the same gene have their test statistics combined to give a single p value for the test of the association of the trait with that gene. This method was inspired by pathway analyses in microarray data (Wang et al., 2007), and some of the first implementations involved adapting the GSEA (Gene Set Enrichment Analysis) algorithm (Subramanian et al., 2005), and adjustment for multiple testing is achieved through a permutation-based procedure (Subramanian et al., 2005; Wang et al., 2007).

Combining the test statistics of each SNP in a gene into a single gene-level test statistic (p value), was done by calculating a maximum statistic by Wang et al (Wang et al., 2007) (summing p values or the logarithms of p values), followed by permutation-based adjustment. Permutation testing approaches are used to adjust for multiple comparisons, and in contrast to Bonferroni or Benjamini-Hochberg where a family-wise or false discovery rate is controlled at a desired value, the underlying null distribution of the sample data is simulated by

resampling test statistics under the null hypothesis (Camargo et al., 2008; Conneely & Boehnke, 2007).

There may be issues with this overall approach if multiple LD blocks in a gene contain a SNP contributing significantly to variation in the trait (i.e. putatively causal), and although a permutation-based approach to multiple testing corrections keeps the Type 1 error rate effectively the same across genes of different size, there may be a loss of power for larger genes, and there is some evidence that a permutation-based approach is not accurate with increasing LD as "undue weight" is given to highly correlated markers (Moskvina et al., 2012). There are a range of methods for gene-level association testing (gene-based testing) and gene-set analyses (De Leeuw et al., 2016; Holmans et al., 2009; P. H. Lee et al., 2012; Lips et al., 2012) which aim to address these issues related to linkage disequilibrium and gene size, and one of the best-performing of such methods is MAGMA (Multi-marker Analysis of GenoMic Annotation) (de Leeuw et al., 2015; De Leeuw et al., 2016).

MAGMA gene-based testing or gene analysis uses a multiple linear principal components regression model to test for association between each gene and the trait of interest, and an F test is used to compute the gene-level p value. The SNP matrix for a gene, consisting of rows of participants and columns of SNP genotypes (i.e. each element in the matrix is a 0, 1 or 2) is projected onto that gene's genetic principal components (PCs), PCs with very small eigenvalues are removed, and then remaining PCs are used as predictors of the trait of interest in the linear regression model (de Leeuw et al., 2015) (Equation 2.16), where *Y* phenotype or trait value, α_{0g} is the intercept, X_g^* is the matrix of PCs, α_g is the vector of genetic effects for gene *g*, *W* an optional matrix of additional covariates and β_g the vector of covariate effects.

$$Y = \alpha_{0g}\vec{I} + X_g^*\alpha_g + W\beta_g + \varepsilon_g$$

Equation 2. 16: Regression of phenotype Y on gene effects (gene-level MAGMA analysis).

An F test then tests the null hypothesis of no association between the gene and the trait (all values in the vector α_g being zero for a gene). This method means that LD is fully accounted for, the model is flexible (allows for extra covariates and interaction terms without change to the underlying model), and computation time is much faster in comparison to permutation-based test methods.

There is some discussion in the literature as to defining gene boundaries in the context of gene-level association testing (reviewed (Wang et al., 2007), see also (Portin & Wilkins, 2017)), which can be done according to SNPs locations in relation to expression boundaries, coding regions and varying length of up/downstream sequence, and may or may not include SNPs correlated with SNPs mapped to gene locations. In MAGMA analyses, genes are defined by their transcription start and stop sites, as given by human genome reference builds & Entrez gene IDs, and SNPs are mapped to the gene if they are located within that interval (between start and stop site) – options also exist to add upstream and downstream extensions of this interval (de Leeuw et al., 2015).

As an extension of gene-based or gene-level association testing using MAGMA, trait-associated genes can be tested for membership of functional pathways (gene-set analysis). This is achieved by transforming the p value for each gene p_g (calculated during the gene-level analyses) into a Z value using Equation 2.17 below.

$$z_g = \phi^{-1}(1 - p_g)$$

Equation 2. 17: Transformation of gene p values to Z values for gene set analysis.

Where ϕ^{-1} is the probit function. This gives a variable *Z* containing all values of z_g. To ask whether all genes in a set *s* are associated with the trait (self-contained gene-set analysis), an intercept-only regression is carried out (Equation 2.18).

$$z_s = \beta_0 \vec{1} + \varepsilon_s$$

Equation 2. 18: Intercept-only regression (MAGMA gene set analysis, self-contained)

Competitive gene-set analysis tests whether genes in a set are more strongly associated with the trait than genes in another set. The regression used in self-contained gene-set analyses is expanded through use of a binary indicator variable S_s with elements s_g (with s_g = 1 for genes present in a set, and = 0 for those outside the set) (Equation 2.19).

$$Z = \beta_{0s} \vec{1} + S_s \beta_s + \varepsilon$$

Equation 2. 19: MAGMA gene set analysis (competitive)

The parameter β_s shows the difference in association between genes in the set and those not in the set and testing the null hypothesis β_s = 0 against the one-sided alternative β_s > 0 is equivalent to performing a one-sided two-sample t-test that compares mean association of genes in the set with the mean association of genes not in the set. Similarly, self-contained gene-set analyses is the same as carrying out a one-sided one-sample t-test, comparing the mean association value of genes in the gene set to 0.

Aggregating SNP-level statistics to the gene and gene-set level allows for an increase in power as fewer tests are performed overall. These types of analyses can also inform on potential functional impact of trait-associated SNP variation, by indicating loci for further investigation.

2.3.1.2 GTEx

The Genotype-Tissue Expression (GTEx) project is resource that enables study of relationships between genetic variation, gene expression, and other molecular phenotypes in a range of human tissues (Aguet et al., 2017; Ardlie et al., 2015). As of the GTEx v6 data freeze used within FUMA, the resource consists of data from over 7,000 cell and tissue samples from 449 donors (the most recent GTEx

release (v8) is made up of over 1,000 individuals). One of the key goals for the GTEx project was to identify eQTLs for all genes for a range of human tissue types, and over 150,000 cis-eQTLs have also been mapped using these data. This was done by calculating significance correlations between genotypes and gene expression levels by performing linear regression of genotype on quantile normalized gene-level expression values (following correction for technical covariates) using Matrix eQTL (Shabalin, 2012) – see also Ardlie et al Supplementary Information including Figure S8 (Ardlie et al., 2015). Gene-specific p values with correction for multiple testing (of multiple SNPs per gene) were calculated using a permutation-based approach. SNPs were mapped to genes if they were located within 1Mb of the transcription start site.

Normalized gene expression data (reads per kilobase per million) for 56, 320 genes in 53 tissues were taken from GTEx v6 for use in FUMA (Watanabe et al., 2017). These 56, 320 genes were filtered to include genes with an average RPKM per tissue greater than or equal to 1, in at least one tissue type, giving a set of 28, 520 genes, of which 22, 146 were mapped to entrez ID identifiers. Gene expression analysis using GTEx in FUMA is an extension of MAGMA gene-level analyses: “gene-property” analysis is performed using the average expression of genes per tissue type as a gene covariate, in order to test the (positive) relationship between genes highly expressed in a certain tissue, and genetic associations. Gene expression values are log2 transformed average RPKM per tissue type, after winsorized at 50 (based on GTEx RNA-seq data). FUMA tissue expression analysis is performed separately for 30 general tissue types and 53 specific tissue types.

The full GTEx data are accessible via dbGap, and certain subsets can be explored and visualised using the online GTEx portal [https://www.gtexportal.org/home/].

2.3.2 Cohort Profiles

2.3.2.1 UK Biobank

UK Biobank is a UK general-population cohort of 0.5 million individuals recruited in middle age (40-79 years) from 2006-2010, with ongoing follow-up assessments including imaging, repeated measures of baseline phenotypic measures and

linkage to health records and death registers (Sudlow et al., 2015). Many thousands of phenotypic measures, such as blood pressure, height and weight were recorded, along with whole-genome genotyping (Bycroft et al., 2018). A subset of the cohort also completed an online follow-up Thoughts and Feelings questionnaire, from which DSM-5-approximating psychiatric disorder phenotypes can be derived (Davis et al., 2020). In addition, and importantly for this PhD project, all 0.5 million UKB participants were also asked through touchscreen questionnaire about pain and duration of any pain, at a range of bodily sites, at baseline assessment. (Pain phenotyping is discussed in detail below and in relevant results chapters). Approved UKB projects with datasets used for analyses described in this thesis were 6553 and 7155.

A large proportion of the information collected during the UKB assessment centre visits or as part of online follow up, including questions on pain, is based on self-report, which may also represent a limitation in comparison to use of data collected during interview by a healthcare professional or with testing or sample collection. In general, information collected by self-selected participants in a self-report format can be subject to a range of biases such as confounding (the risk factor(s) being studied is correlated with an unmeasured risk factor), information bias (systematic measuring errors during data collection), and selection bias (the studied population is non-representative of the general population) (Janssens & Kraft, 2012).

More specific limitations include the fact that UKB participants tend to be wealthier and healthier than the general UK population, and are likely to be older, less likely to be obese, less likely to be physically inactive, and less likely to smoke and drink on a daily basis: the participation rate for UK Biobank was also 5.45% (Fry et al., 2017).This is in line with findings showing that research participants and those who purchase direct-to-consumer genetic tests tend to be non-representative of target or general populations (Klijs et al., 2015; Leitsalu et al., 2015; Prictor et al., 2018; Stamatakis et al., 2021). This ‘healthy volunteer effect’ can have adverse effects when estimating relative risk of lifestyle and environment exposures in the study of chronic disease (reviewed by Stamatakis et al., 2021). Additionally, UKB participants are also ethnically homogenous, with the majority being white (94.6%), again affecting the extent to which results of studies using UKB can be generalised to other populations.

Issues can also arise if the trait of interest is correlated with study participation – for example, in the study of chronic pain in non-clinical cohorts, results could be biased if participation is less likely for individuals with the most severe and/or disabling chronic pain or chronic pain conditions.

Despite the unrepresentativeness of UKB with respect to the UK general population, some analyses found that results of studies using the UKB resource are still largely generalisable to the UK population (Fry et al., 2017). Furthermore, lack of representation in general is not necessarily problematic, if this is considered during interpretation of study results (Rothman et al., 2013).

2.3.2.2 23andMe

23andMe is a private direct-to-consumer genetic testing company. Consumers in 50 countries worldwide including the US, UK and Canada may purchase saliva testing kits and receive information on ancestry and genetic predisposition to disease. Data is also used in research – 80% of 23andMe customers 'opt-in' for this, with each consumer contributing on average to 200 different studies. One such study was a GWAS of chronic pain grade carried out using 23andMe consumer genotyping data in collaboration with Pfizer. As discussed in the previous section research participants may be non-representative of wider target populations, and individuals purchasing direct-to-consumer genetic testing kits in particular tend to be white, have higher educational attainment, and have higher income (Gollust et al., 2017; J. S. Roberts et al., 2017).

The sample characteristics of the collaborative chronic pain grade GWAS carried out by 23andMe and Pfizer have been summarised by McIntosh et al (McIntosh et al., 2016). Validated pain questionnaires identical to those used in Generation Scotland: Scottish Family Health Study (GS: SFHS) were completed by more than 32,000 research participants, from which a sample of 23,332 unrelated white European ancestry participants was derived, consisting of 10, 780 pain cases (i.e., those with any chronic pain grade that was not zero) and 12, 552 controls. This GWAS was carried out with adjustment for age, sex, BMI, current and previous manual labour and the first five genetic principal components.

2.3.2.3 Generation Scotland: Scottish Family Health Study

The GS: SFHS is a genetic epidemiology study with a family-based recruitment process and structure, comprised of sociodemographic, clinical and DNA data from ~ 24,000 participants recruited across Scotland, aged 18-98 years old (B. H. Smith et al., 2006, 2013). Recruitment was carried out from 2006-2011 through identifying suitable potential participants registered at participating general medical practices, and the final cohort was 59% female. As in UK Biobank, the sample is in general healthier and wealthier than the general Scottish population, but nevertheless contains participants from a wide range of socioeconomic backgrounds and with a wide range of clinical features (Smith et al., 2013). The family-based structure and depth and breadth of phenotyping information allows for family-based genetic studies, for example into parent-of-origin effects, and for different forms of genetic studies such as investigating the role of rare alleles in health and disease.

2.3.3 Chronic Pain Phenotyping in Key Cohorts

2.3.3.1 Chronic Pain in UK Biobank

At baseline all UKB participants were asked about 'Pain type(s) experienced in the last month' (data field 6159). Participants could choose from seven non mutually exclusive body sites or 'pain all over the body' or could answer 'none of the above' or 'prefer not to answer'. If participants selected 'pain all over the body' they could not then select a specific site. Each body site, and the 'all over the body' option, had a corresponding question item where participants could answer if this pain had lasted for 3+ months or not – to which participants could respond 'yes', 'no', 'do not know' or 'prefer not to answer'.

As discussed in 1.1.2, those defined as having 'chronic pain' can be extremely heterogeneous groups of people, and chronic pain is measured and defined in a wide range of ways. It may be more powerful to consider chronic pain as a disease in its own right, as recently outlined in IASP taskforce discussions and recent ICD-11 coding additions of "chronic primary pain". Multisite chronic pain (MCP) is a derived quasi-quantitative trait, constructed in order to investigate chronic pain as a phenotype in its own right, and with consideration of the fact

that there are unlikely to be legitimate cut-off points between localised and widespread chronic pain (Kamaleri et al., 2008).

The trait value for MCP was derived from the number of sites at which the participant had experienced chronic pain for 3+ months (0-7). Those answering 'prefer not to answer' with regards to pain types experienced in the past month were removed from analyses. Those answering that they 'did not know' the duration of any pain were not labelled as having chronic pain at that site but were not excluded from analyses. In the GWAS analyses discussed in Chapter 4, those answering that they experienced pain all over the body, which lasted for 3+ months, were excluded from initial GWAS. Rationale behind this is further discussed in Chapter 4, but briefly: pain all over the body may represent an extreme phenotype presentation of MCP, but this may also represent a distinct phenotype in comparison to having a number of individual chronic pain sites (Gerhardt et al., 2016a; Viniol et al., 2013; Zadro et al., 2020), or even in comparison to participants selecting 7 individual sites of chronic pain (Nicholl et al., 2014). In addition, 'all over the body' does not necessarily follow linearly from an MCP trait value of 7, potentially representing an experience of pain without distinct sites that can be quantified and therefore complicating GWAS analyses (which are regression-based). The relationship between chronic pain all over the body and MCP is investigated in downstream analyses in Chapter 4 in order to address these issues.

A phenotype approximating chronic widespread pain can also be derived in UK Biobank, and consists of those labelled as having chronic pain all over the body i.e., those who answer 'All over the body' to pain types experienced in the past month, and answer that this pain has lasted for 3+ months.

2.3.3.2 Chronic Pain in Generation Scotland and 23andMe-Pfizer Sample

Chronic pain grade a validated chronic pain phenotype (B. H. Smith et al., 1997; M Von Korff et al., 1992) derived from questionnaire participation, was ascertained in both GS: SFHS and 23andMe-Pfizer sample. Chronic pain grade incorporates scores on both disability and pain severity, and trait value ranges from 0 (no chronic pain) to 4 (most severe and most disabling chronic pain) depending on both disability due to pain and pain intensity. These scores are calculated from the answers to seven questions, all of which besides question 4

("About how many days in the last six months have you been kept from your usual activities (work, school or housework) because of this pain?") are answered by giving a rating from 0-10. A rating of 0 represents 'no pain' for questions 1-3, 'no change' for questions 6-7, or 'no interference' for question 5, and ratings of 10 represent 'pain as bad as it could be' for questions 1-3, 'unable to carry on activities' for question 5 and 'extreme change' for questions 6-7. Pain intensity is then calculated as the mean of question 1 + question 2 + question3 multiplied by 10, and disability score as the mean of the sum of rating values for questions 5-7, multiplied by 10. Disability points are then calculated from the recoded disability score (0-29 = 0, 30-49 = 1, 50 – 69 = 2, >70 = 3) added to the recoded number of days value from question 4 (0-6 days = 0, 7-14 days = 1, 15-30 days = 2, >31 days = 3).

Chronic pain grade is then assigned based on both pain intensity and disability due to pain, as measured using disability points and pain intensity score. Chronic pain grade classification of 0 corresponds to disability points of 0 and pain intensity of 0, grade 1 to pain intensity of < 50 and disability points < 3, 2 to pain intensity greater than or equal to 50 and disability points < 3, grade 3 to disability points of 3 or 4, regardless of pain intensity, and grade 4 to disability points of 5+, again regardless of pain intensity.

Such a phenotype may be potentially problematic when trying to understand the mechanisms of chronic pain development, as increasing trait value is not only correlated with increased chronic pain severity but with how that pain affects interaction with the environment (disability due to pain). Disability due to pain is likely also influenced by a range of factors, some of which may constitute confounders of the relationship between pain and pain-related disability. For instance, low socioeconomic status is associated with both chronic pain and can contribute to disability related to chronic pain (reviewed by Mills et al., 2019), a relationship which can complicate a study to find genetic variation associated specifically with chronic pain. Furthermore, extremes in chronic pain grade trait value do not necessarily represent the most severe chronic pain, only the pain associated with greatest disability – although these factors would be expected to correlate with one another (Chiarotto et al., 2019), again higher disability points could be related to other, non-pain-severity factors that increase pain-related disability (i.e. theoretically individuals with the 'same' chronic pain but who

experience environments that are disabling to different extent can have differing chronic pain grade classifications). For example, an individual living in smaller, less accessible housing may be more likely to find pain has a greater impact on daily tasks compared to an individual with similarly severe pain who has access to resources that allow them to modify their living environment. A phenotype closely matching that of MCP in UK Biobank was also derived using GS: SFHS data. GS: SFHS participants were asked "are you currently troubled by any pain or discomfort?" as part of a chronic pain identification questionnaire – if answering 'yes', they could then choose from six specific bodily sites and 'other'. In contrast to UK Biobank there was no option to note whether pain at specific sites had lasted longer than 3 months, but participants were asked the single question "have you had this pain or discomfort for more than 3 months?" which could refer to one of, several of, or all their sites of pain and is not discernible from the data. Body site options also differ slightly between cohorts. Therefore, MCP in GS: SFHS can take a value from 0-6 sites of chronic pain. This assumes that answering yes to the question "Have you had this pain or discomfort more than 3 months" indicates that pain at every site indicated by the participant is chronic.

2.3.4 Major Depression Phenotyping in UK Biobank

A subset of UKB participants (N = 157, 366) fully completed the online follow-up 'thoughts and feelings' mental health questionnaire (Davis et al., 2020). This mental health questionnaire was designed by an expert working group and involved consultation with a patient group, and aims to make use of existing, validated measures. Though case classifications aim to replicate a psychiatric diagnosis, their delivery and reliance upon self-report means they can only be thought of as "likely" psychiatric diagnoses (Davis et al., 2020). Despite this, prevalence and patterns of association between demographic factors and other disorders were found to match expectations based on previous research and the expectations of the Health and Safety Executive (HSE) (Davis et al., 2020).

MDD and related phenotypes in UK Biobank was derived following protocol found at http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/MentalStatesDerivation.pdf?fbclid=IwAR1Bsy3hnKzC6uThVpcz8bkbzV9yH-

9dkp0gVCvOSuaV1CZcm1nu0p0qYII (UKB application 7155) and associated with (Smith et al., 2013). These MDD phenotypes are also discussed by Davis et al (Davis et al., 2020). Derived UKB MDD phenotypes are “single probable major depressive episode”, “probable recurrent major depression (moderate)” and “probable recurrent major depression (severe)”, and the latter two can be combined into “probable recurrent major depression”. A ‘ranked mood’ variable can then be constructed, where each participant has a value from 0-4, 0 indicating they did not meet criteria for any derived major depression or bipolar disorder phenotype, 1 indicating meeting criteria for single episode major depression, 2 for probable recurrent major depression, and 3 or 4 indicating having met criteria for either bipolar disorder I or II.

Chapter 3: Further Understanding Overlap of Chronic Pain and Depression: Pleiotropy and Clinical Heterogeneity

3.1 Introduction

Analyses undertaken in this chapter address objectives 1, 2 and 3: to investigate genetic correlation and pleiotropy between MDD and chronic pain, to uncover common genetic variation associated with chronic pain phenotypes, and to test for clinical heterogeneity between MDD and chronic pain. Some of the analyses described in this chapter have been published as part of an article in Translational Psychiatry (Johnston et al., 2019).

As previously described (2.2.3), cFDR analyses provide an alternative route to SNP discovery, making use of existing GWAS datasets and leveraging association with related conditions to boost discovery power. In the case of chronic pain, the association with mood disorders (commonly comorbid with chronic pain and chronic pain conditions) is of substantial interest, as improved understanding of the biological underpinnings of this overlap may provide ideas for the development of novel treatment strategies. cFDR analyses have been used to find novel variants and pleiotropic loci associated with schizophrenia, type 2 diabetes, Alzheimer disease, bipolar disorder and systolic blood pressure (Andreassen et al., 2014; Andreassen, Djurovic, et al., 2013; Andreassen, Thompson, et al., 2013; Wang et al., 2016).

Additionally, to date there is a relative lack of large, well-powered GWASs of chronic pain as a phenotype in its own right (1.3.4), and in those which have been carried out few variants have been found to be significantly associated with chronic pain. One such GWAS is the 23andMe-Pfizer GWAS of CPG (2.3.2.2), N = 23, 332 (McIntosh et al., 2016), where no SNPs were found significantly associated with CPG. cFDR therefore represents a promising and potentially more cost-effective method for identifying new SNPs associated with complex traits such as chronic pain, by repurposing existing GWAS outputs. Genetic correlation between traits can be driven by pleiotropy (1.3.3.2), and MDD and chronic pain grade have been previously found be genetically correlated at $\rho \sim 0.5$ (McIntosh et al., 2016). However, even if pleiotropy is detected it is unclear whether this is whole-group (so-called 'true' pleiotropy) or subgroup-driven

(clinical) heterogeneity (see also 2.2.4). Additionally, other types of analyses to check for clinical heterogeneity (e.g., polygenic risk scoring) are not robust in the face of pleiotropy. In order to test for presence of clinical heterogeneity and to distinguish this from whole-group pleiotropy in MDD and chronic pain in UK Biobank, BUHMBOX (Han et al., 2016) analysis was carried out.

3.2 Methods

3.2.1 Conditional False-Discovery Analysis of Chronic Pain Grade and Major Depressive Disorder

3.2.1.1 Phenotype Definition and Source Data

cFDR analyses require two independent GWAS summary statistic datasets. For chronic pain, summary statistics from a GWAS carried out collaboratively with Pfizer-23andMe Inc of CPG (McIntosh et al., 2016)) were used. This GWAS sample consisted of 23,332 unrelated white European ancestry participants was derived, consisting of 10, 780 pain cases (i.e., those with any chronic pain grade that was not zero) and 12, 552 controls.

For MDD, summary statistics from a recent case-control GWAS meta-analysis (Wray et al., 2018) were provided by the Psychiatric Genomics Consortium (PGC). After removal of data from 23andMe and UK Biobank participants, this gave a dataset originating from an analysis using 43 028 cases and 87 522 controls. Phenotype definitions, study population demographics and meta-analysis procedures for the MDD GWAS have been described previously (Wray et al., 2018).

3.2.1.2 Data Preparation and Linkage Disequilibrium Pruning & cFDR Analysis

A dataset of SNPs for which a *p*-value for association, chromosome position data and rsID were available in both MDD and CPG datasets was constructed. This was then LD pruned. Firstly, PLINK-format genotype data, for each SNP in the newly compiled CPG-MDD summary-statistic dataset, was extracted from the UK Biobank genotype data (approved application 6553). Pruning was carried out using command line PLINK (version 1.9) --indep-pairwise function. These parameters are as recommended for cFDR analyses (Andreassen, Djurovic, et al.,

2013; Liley & Wallace, 2015). This resulted in an LD-pruned dataset of 774 292 SNPs with association data available for both MDD and CPG. These SNPs were then taken forward and cFDR were calculated using equation X in 2.2.3, repeated below (Equation 1), following detailed formulae and derivations from Liley & Wallace (Liley & Wallace, 2015) using R (version 3.5.2). Conjunctional cFDR (ccFDR) values are then the highest cFDR value between the two cFDR analyses.

$$cFDR = \Pr\left(H_{0(i)} \middle| p_i \leq P_i, p_j \leq P_j\right) = \frac{p_i}{\Pr(p_i \leq P_i \,|\, p_j \leq P_j)}$$

Equation 3. 1: Conditional false discovery rate.

3.2.2 Further understanding the overlap of MDD and Chronic Pain

The genomic context for SNPs significantly associated with MDD, CPG or both was examined. The R package 'rsnps' was used to extract data from records in NCBI dbSNP (https://www.ncbi.nlm.nih.gov/snp). Genomic context for each SNP was examined in the UCSC Genome Browser (build GRCh38/hg38) (Kent et al., 2002), using a window of 0.5Mbp around each SNP and data from the GENCODE v24 track, validated or reviewed by either Refseq or SwissProt staff. Genes partially or fully contained within this window were noted. The presence of cis-eQTLs close to the significant SNPs was investigated using the IGV eQTL Browser (Aguet et al., 2017) web interface.

3.2.3 Clinical Heterogeneity in MDD and Chronic Pain

Briefly, BUHMBOX requires GWAS summary statistics for disease A obtained in a sample that is independent from the sample where disease B is measured (and vice versa). First, clinical heterogeneity in MDD cases was tested for, using CPG as the independent GWAS dataset, and secondly clinical heterogeneity in chronic pain cases was tested for using MDD as the independent GWAS dataset. BUHMBOX is fully described in 2.2.4.

3.2.3.1 Phenotypes & Data

Chronic pain is widely defined as pain persisting beyond the healing period (with the threshold duration of 3 months taken to be the healing period) (Greene, 2010; Merskey & Bogduk, 1994), and can be assessed using the CPG questionnaire (M Von Korff et al., 1992). On this scale 0 = no chronic pain (no pain that has persisted beyond 3 months), 4 = most severe chronic pain (pain persisting beyond 3 months which also fulfils specific criteria relating to impact on daily functioning, mood, and ability to work). In contrast to MCP, questions on pain duration and impact on quality of life are also incorporated into this chronic pain phenotype. A GWAS of CPG (0-4) using a linear regression model was carried out by 23andMe in collaboration with Pfizer (McIntosh et al., 2016), these summary data are used here.

Wray et al carried out a GWAS meta-analysis of MDD (Wray et al., 2018). Summary statistic data with UK Biobank and 23andMe participants removed were used in this analysis. Effect allele frequencies were obtained from the GWAS summary statistic dataset where UK Biobank individuals were not removed, downloadable from the PGC website– this is unlikely to bias results significantly and is acceptable for BUHMBOX analysis (Han et al., 2016), despite EAF calculations involving UK Biobank participant data.

With respect to phenotyping in UK Biobank for these analyses, being classed as having MDD approximates a DSM-5 diagnosis of MDD (see 2.3.4). Controls consist of those with no mood disorder, and those with bipolar type 1 or 2 are removed from the analyses, along with those answering ‘Prefer not to say’ or ‘Don’t know’ in components of the MDD phenotype. The number of MDD cases and controls prior to BUHMBOX quality control were 34,025 and 93,819, respectively.

During the baseline investigations, UK Biobank participants were asked via a touchscreen questionnaire about “pain types experienced in the last month” (field ID 6159), with possible answers: ‘None of the above’; ‘Prefer not to answer’; pain at seven different body sites (head, face, neck/shoulder, back, stomach/abdomen, hip, knee); or ‘all over the body’ (see Chapter 2 section: Chronic pain phenotyping in key cohorts). Those who answered that they had chronic pain at any site were classed as cases, and chronic pain at none of the sites as controls – those who answered ‘don’t know’ or ‘prefer not to answer’

were removed from analyses. The number of chronic pain cases prior to BUHMBOX quality control was N = 215, 383 and controls N = 279, 641.

3.2.3.2 Genotype Data

UK Biobank phenotyping, genotyping and quality control has been described in detail elsewhere (Bycroft et al., 2018; Sudlow et al., 2015). The first eight genetic principal components (pre-calculated and included as part of UK Biobank data) were used in BUHMBOX calculations.

3.2.3.3 BUHMBOX Procedure

SNPs associated with CPG at a p value of 10^{-4} or less, their effect allele frequencies and effect sizes (odds ratios, ORs) were compiled from the UKB genotyping data, as were MDD cases/controls UKB-IDs and genetic principal components. ORs as a measure of effect size for SNPs associated with CPG were obtained by taking the exponent of beta values for each SNP (in personal communication between Dr Nicholas Graham and Dr Buhm Han this transformation was deemed acceptable in BUHMBOX analysis, despite the CPG GWAS being non case-control).

BUHMBOX quality control steps (Han et al., 2016), other exclusions (see 'Phenotyping') and linkage-disequilibrium (LD) pruning using command-line PLINK 'indep-pairwise' function with recommended parameter settings of 50 kilobase window size, a step size of 5 SNPs and r^2 threshold of 0.1 (Han et al., 2016) was carried out. The resulting number of MDD cases = 3,455 and controls = 9,681, and the number of independent CPG-associated SNPs used in calculations was 156. 147 of these SNPs were imputed with mean average call rate (the imputation quality metric provided with the 23andMe GWAS data) of 0.99. BUHMBOX was carried out to obtain a BUHMBOX test statistic value. If the test statistic-associated z value is negative, the resultant P value is transformed via 1 – P value (BUHMBOX analysis performs a one-sided test only).

SNPs associated with MDD at a p value <= 10^{-4}, effect allele frequencies and effect sizes were compiled, along with chronic pain case and control UKB IDs and genetic principal components. BUHMBOX-specific quality control steps, phenotypic exclusions and LD pruning were carried out at previously described. The resulting number of chronic pain cases was 51, 494 and controls was 67, 857,

with 120 SNPs also taken forward for use in BUHMBOX calculations. 110 of these SNPs were imputed with mean imputation score of 0.96. BUHMBOX test-statistic z values were converted as needed as previously described.

3.3 Results

3.3.1 cFDR: SNPs Associated with CPG and MDD

cFDR analyses were carried out to investigate pleiotropic loci associated with both MDD and CPG. Eleven SNPs in total were found at cFDR ≤ 0.01 (Table 3.1), six of which, located on chromosomes 12 and 14, were associated with CPG and nine of which, located on chromosomes 1 and 14, were associated with MDD. Four of these 11 SNPs, all located within a 131 kilobase-pair region on chromosome 14, were found to be pleiotropic (ccFDR ≤0.01).

rsID	Position	Alleles	B (CPG)	p (CPG)	OR (MDD)	p (MDD)	cFDR (CPG)	cFDR (MDD)	ccFDR
rs4904790	14:42242623	C/T	-0.03	1.44×10^{-3}	1.042	1.37×10^{-5}	0.02	3.58E-03	0.02
rs1584317	14:42213816	C/G	0.028	4.48×10^{-3}	0.96	4.59×10^{-6}	0.029	3.57E-03	0.03
rs11846556	14:42183025	A/G	-0.037	1.11×10^{-4}	1.05	2.98×10^{-7}	5.57×10^{-4}	3.76E-05	5.57E-04
rs10131184	14:42166111	A/G	0.035	2.82×10^{-4}	0.95	2.53×10^{-8}	8.46×10^{-4}	7.10E-06	8.46E-04
rs8015100	14:42095232	A/T	-0.033	6.67×10^{-4}	1.06	1.50×10^{-9}	6.67×10^{-4}	9.27E-07	6.67E-04
rs11157241	14:42051771	C/T	0.035	2.91×10^{-4}	0.94	4.44×10^{-9}	5.83×10^{-4}	1.28E-06	5.83E-04
rs10138559	14:41975989	C/T	-0.02	0.03	1.042	1.04×10^{-6}	0.1	5.25E-03	0.1
rs10872954	14:41948768	A/G	-0.026	6.45×10^{-3}	1.04	7.68×10^{-6}	0.053	7.02E-03	0.053
rs149981001	12:60264802	C/T	0.2	6.24×10^{-8}	1.087	0.0181	1.06×10^{-3}	0.018	0.018
rs147573737	12:60231575	C/T	-0.2	2.09×10^{-7}	0.917	0.023	2.21×10^{-3}	0.023	0.023
rs35641559	1:73760104	C/T	0.02	0.03	0.961	2.08×10^{-6}	0.1	8.83E-03	0.11

Table 3. 1: Loci identified from cFDR analysis.

Position = position given as chromosome: base pair location. cFDR (CPG) = cFDR for CPG conditioning on MDD; B (CPG)/p (CPG) = effect size and p value from the CPG GWAS; cFDR (MDD) = cFDR for MDD conditioning on CPG; OR (MDD)/p (MDD) = effect size (odds ratio for the effect allele) and p value from the MDD GWAS. Alleles are given as effect allele/other; effect allele is defined as the allele for which association with the trait was tested in the original (CPG or MDD) GWAS. rsIDs for SNPs associated with both MDD and CPG (pleiotropic SNPs) (ccFDR < 0.01) are shown in bold.

3.3.2 cFDR: Genomic Context of Trait-Associated SNPs

The R package 'rsnps' and the UCSC Genome Browser were used to investigate genomic context of SNPs found to be of interest through cFDR analyses, including nearest genes to these loci (Tables 3.2, 3.3).

rsID	Chromosome	cFDR-Associated Trait	Gene(s)	Alleles	Major	Minor	MAF	AA
rs35641559	1	MDD	LOC105378800	C/T	T	C	0.4641	T
rs149981001	12	CPG	NA	C/T	C	T	0.0022	C
rs147573737	12	CPG	NA	C/T	T	C	0.0024	T
rs4904790	14	MDD	LRFN5	C/T	C	T	0.3175	C
rs1584317	14	MDD	LRFN5	C/G	G	C	0.2993	G
rs11846556	14	Both	LRFN5	A/G	A	G	0.3676	G
rs10131184	14	Both	LRFN5	A/G	G	A	0.239	G
rs8015100	14	Both	LRFN5	A/T	A	T	0.2546	A
rs11157241	14	Both	NA	C/T	T	C	0.2508	T
rs10138559	14	MDD	NA	C/T	C	T	0.4399	T
rs10872954	14	MDD	NA	A/G	A	G	0.4343	A

Table 3. 2: Output of 'rsnps' query

SNP ID (rsID), location, cFDR-associated trait, associated genes (Gene(s)), minor allele frequency (MAF) and ancestral allele (AA) are shown. 'NA' indicates no result for that query in that category.

rsID	Chromosome	cFDR-Associated Trait	Gene(s)
rs35641559	1	MDD	*LINC01360, LRRIQ3, FPGT, FPGT-TNNI3K*
rs149981001	12	CPG	*SLC16A7*
rs147573737	12	CPG	*SLC16A7*
rs4904790	14	MDD	*LRFN5*
rs1584317	14	MDD	*LRFN5*
rs11846556	14	Both	*LRFN5*
rs10131184	14	Both	*LRFN5*
rs8015100	14	Both	*LRFN5*
rs11157241	14	Both	*LRFN5*
rs10138559	14	MDD	*LRFN5*
rs10872954	14	MDD	*LRFN5*

Table 3. 3: UCSC Genome Browser Results.

3.3.2.1 CPG-Associated SNPs

No genes were found through an 'rsnps' query for cFDR-associated SNPs associated solely with CPG (Table 3.2). *SLC16A7,* which encodes monocarboxylate transporter 2 (MCT2), is located within 1Mbp of SNPs which were solely associated with CPG (Table 3.3). In the central nervous system, MCT2 is involved in high affinity, proton-coupled transport of metabolites (particularly lactate) into neurons and may play a role in neuronal uptake of energy substrates released by glia (Y. Itoh et al., 2003; Pellerin, 2003). MCT2 is localised to the post-synaptic compartment in many human neurons and may have a specialised role in synaptic functioning (Chiry et al., 2008; Pierre et al., 2002). Regulation of *SLC16A7* has also been linked to disorders of the brain: loss or under-expression has been associated with temporal-lobe epilepsy (Lauritzen et al., 2012) and it may be expressed and methylated at different levels in patients with psychosis versus controls (C. Chen et al., 2014).

3.3.2.2 MDD-Associated SNPs

A single SNP on chromosome 1 was solely associated with MDD and located within 1Mbp of *LRRIQ3* and *FPGT (*Table 3.3*). LRRIQ3* encodes leucine-rich repeat (LRR) and IQ motif containing protein 3, a calcium-channel component. LRR-domain containing proteins in general are involved in cell-cell communication,

including processes involved in innate immunity and neuronal development (Bella et al., 2008; Ng et al., 2011). *FPGT* encodes fucose-1-phosphate guanylyltransferase, a protein involved in the alternative (salvage) pathways of fucose metabolism (Becker & Lowe, 2003). Fucose metabolism is important in a variety of cell-cell communication and host-microbe interaction situations, but is also important during neuronal development (Becker & Lowe, 2003). Previous studies have found associations between variants in the *LRRIQ3* region, schizophrenia (Ripke et al., 2014), neurodevelopmental disorders (Reuter et al., 2017) and migraine (Gormley et al., 2016).

The pleiotropic *LRFN5* SNPs were all located just upstream of the 5'-most promoter, or within a large intron close to the 5'-end of the gene. The SNPs solely associated with MDD were also located within this intron or were located a little further upstream of the gene. The CPG-only SNPs were all located downstream of *SLC16A7*.

GTEx eQTL results revealed that some of these phenotype-associated SNPs were also associated with expression levels of nearby genes. A SNP associated only with MDD, on chromosome 1 (rs35641559), was found to be significantly associated with expression of a long non-coding RNA gene *LINC01360* in the testis (FDR < 0.05, Table 3.4). The MDD-associated and pleiotropic SNPs on chromosome 14 are all significantly associated with expression of *LRFN5* in a range of tissues, including brain, heart, adipose tissue and spleen. The CPG-associated SNPs on chromosome 12 were not significantly associated with expression of any gene in the eQTL database (Table 3.4).

rsID	Chrom	cFDR-Associated Trait	cis-eQTL Tissue Location(s)
rs35641559	1	MDD	testis
rs149981001	12	CPG	NA
rs147573737	12	CPG	NA
rs4904790	14	MDD	cerebellar hemisphere, cerebellum, transformed fibroblasts
rs1584317	14	MDD	transformed fibroblasts
rs11846556	14	Both	aorta, tibial artery, cerebellar hemisphere, cerebellum, transformed fibroblasts, spleen
rs10131184	14	Both	subcutaneous adipose, aorta, tibial artery, cerebellar hemisphere, cerebellum, thyroid, transformed fibroblasts, oesophagus muscularis, ovary, skin (lower leg, not sun-exposed), spleen
rs8015100	14	Both	omentum, aorta, coronary artery, cerebellum, cerebellar hemisphere, transformed fibroblasts, oesophagus muscularis, ovary, spleen, thyroid
rs11157241	14	Both	subcutaneous adipose, aorta, tibial artery, cerebellar hemisphere, cerebellum, transformed fibroblasts, oesophagus muscularis, skin (lower leg, not sun-exposed), spleen, thyroid
rs10138559	14	MDD	coronary artery, aorta, tibial artery, cerebellum, cerebellar hemisphere, transformed fibroblasts, oesophagus muscularis, spleen, thyroid
rs10872954	14	MDD	aorta, transformed fibroblasts, spleen

Table 3. 4: IGV eQTL Browser results.

The tissue location(s) of cis-eQTLs where a gene is significantly regulated by the queried SNP (rsID column) (FDR < 0.05) are listed, along with SNP ID (rsID), chromosomal location (Chrom) and cFDR-associated trait.

Single tissue eQTL lookups of rs11846556 (Figure 3.1) showed different trends in expression pattern of *LRFN5* with 0, 1 and 2 A alleles, with a trend towards increased expression in the cerebellum and cerebellar hemisphere associated with homozygosity for the A allele (Figure 3.1 A & B respectively), and decreased expression in tibial artery and transformed fibroblasts associated with increasing number of copies of the A allele at this SNP locus (Figure 3.1 C & D respectively).

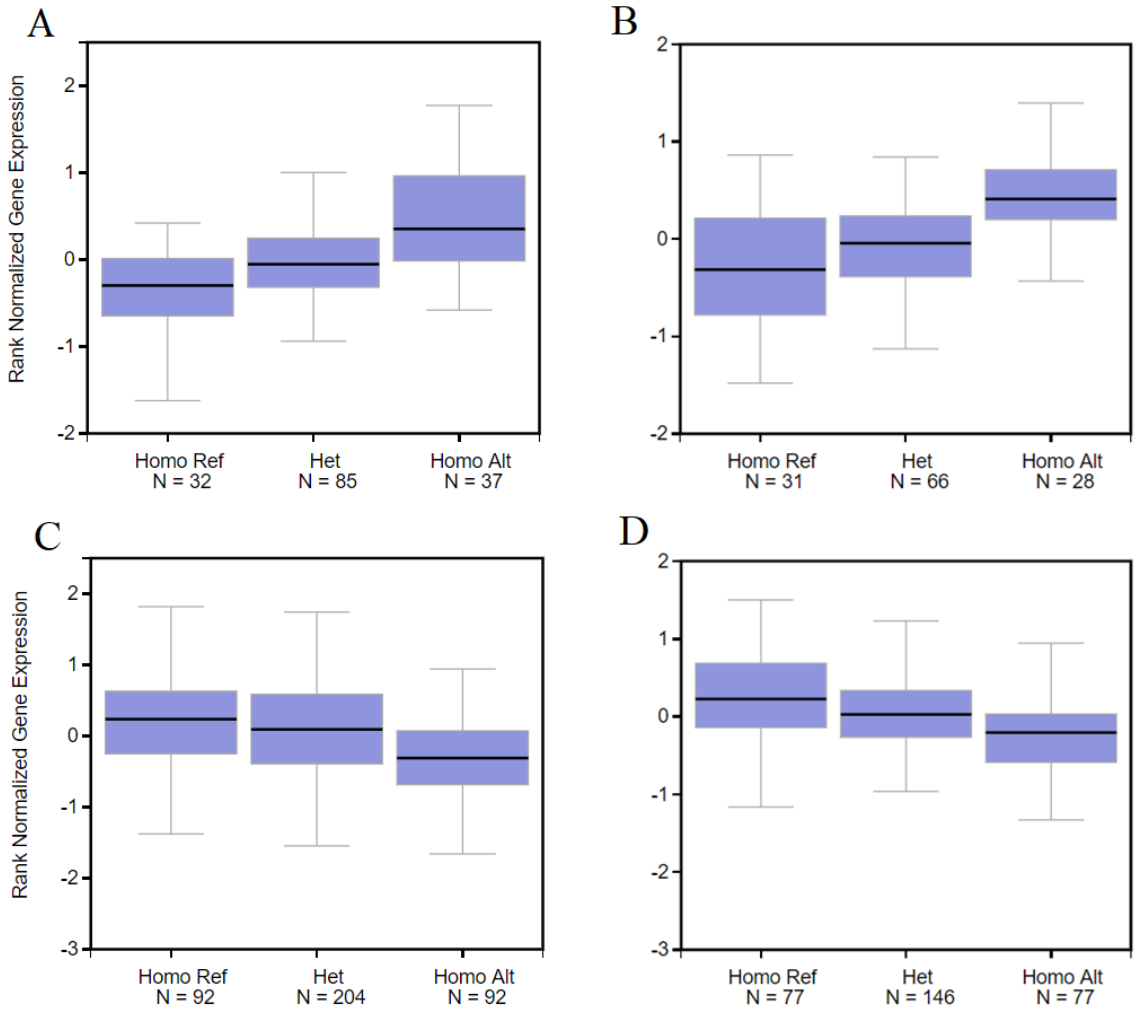

Figure 3. 1. Single tissue eQTL lookups of rs11846556

A = Cerebellum, B = cerebellar hemisphere, C = tibial artery, D = transformed fibroblasts. Homo Ref = homozygous for the reference allele (GG), Het = heterozygote (AG), Homo Alt = homozygous for the alternative allele (AA). Boxplots display minimum, maximum, median, 1st and 3rd quartile rank normalised gene expression values.

3.3.3 BUHMBOX: Whole-Group Pleiotropy in MDD and Chronic Pain in UK Biobank

BUHMBOX analyses were carried out to test for clinical heterogeneity in chronic pain, using chronic pain grade data, with respect to MDD and vice versa in UK Biobank. No evidence for clinically heterogeneity was found in either MDD or chronic pain cases. The BUHMBOX test statistic was insignificant at p = 0.277 (Table 3.5), indicating no clinical heterogeneity in terms of CPG-like MDD cases within MDD in UK Biobank.

p	p (adj)	log(p)	N	N cases	N controls	Z	N loci
0.723	0.277	-0.141	13,136	3,455	9,681	-0.592	156

Table 3. 5: BUHMBOX results for test of clinical heterogeneity in MDD cases in UK Biobank.

p = BUHMBOX p value, p (adj) = 1 – BUHMBOX p value, log(p) = log-transformed p value (base 10), N = total number of individuals included in analysis, N cases = number of case participants included in analysis, N controls = number of control participants included in analysis, Z = BUHMBOX test statistic Z score value, N loci = number of SNPs included in analysis.

The BUHMBOX test statistic was also non-significant at p = 0.29 (Table 3.6) in analyses of chronic pain, showing no clinical heterogeneity, or in other words no MDD-like chronic pain cases.

p	p (adj)	log(p)	N	N cases	N controls	Z	N loci
0.706	0.29	-0.151	119,351	51,494	67,857	-0.541	120

Table 3. 6: BUHMBOX results for test of clinical heterogeneity in chronic pain cases in UK Biobank.

p = BUHMBOX p value, p (adj) = 1 – BUHMBOX p value, log(p) = log-transformed p value (base 10), N = total number of individuals included in analysis, N cases = number of case individuals included in analysis, N controls = number of control individuals included in analysis, Z = BUHMBOX test statistic Z score value, N loci = number of SNPs included in analysis.

3.3.5 Pleiotropic SNPs in *LRFN5*

Conditional false discovery rate analyses, in addition to showing SNPs associated with chronic pain grade, indicated significant pleiotropy at the *LRFN5* locus.

3.4 Discussion

3.4.1 Pleiotropic Loci

LRFN5 encodes leucine-rich repeat (LRR) and fibronectin type 3 domain-containing protein 5. Proteins in the LRFN family span the plasma membrane, with extracellular domains thought to participate in cell-cell interactions necessary for both neuronal development (Morimura et al., 2006; Nam et al., 2011) and synapse formation (Choi et al., 2016). Lrfn5, along with another member of the Lrfn protein family, Lrfn2, may induce both inhibitory and excitatory presynaptic differentiation in nearby neuronal cells (Lin et al., 2018), a process that may play a critical role general brain development and function (Córdova-Palomera et al., 2016). This gene family is expressed primarily in the CNS. Polymorphic markers linked to *LRFN5* have been associated with progressive autism and familial schizophrenia (De Bruijn et al., 2010; Xu et al.,

2009). Neuroinflammation has also been linked to reduced expression of Lrfn5 protein (Y. Zhu et al., 2016).

Each of the four pleiotropic SNPs is associated with opposing directions of effect in MDD and CPG. This may be due to underlying differences in development of MDD and CPG related to brain structure and connectivity – e.g. the maintenance of CP has been theorised to involve neurogenesis and synaptic plasticity (Apkarian et al., 2011; Baliki et al., 2014; Vasic & Schmidt, 2017), and in contrast impaired neurogenesis has been associated with depression (Fang et al., 2018; Jacobs et al., 2000).

However, effect sizes compared are those in the original MDD and CPG GWASs – in this case the confidence intervals include zero (in CPG no SNPs were found to be significantly associated in the original GWAS). cFDR analyses using p values indicate pleiotropy in terms of significant cFDR-derived association, and new effect sizes are not estimated.

3.4.2 Whole-group pleiotropy in MDD and chronic pain

There was not significant evidence for misclassification of individuals with MDD as having chronic pain or vice versa, suggesting that genetic correlation and pleiotropy between MDD and chronic pain in this cohort is driven by whole-group pleiotropy. BUHMBOX is unable to distinguish between horizontal and vertical pleiotropy, so even though 'true' pleiotropy is indicated by these analyses, causal relationships cannot be explored. However, in analyses of rs11846556 genotype (a pleiotropic SNP) carried out in attempts to distinguish mediated from horizontal pleiotropy, it was found that pleiotropy between the two phenotypes may be mediated (i.e., vertical), at least in relation to the *LRFN5* locus.

3.4.2.1 BUHMBOX Power

Non-significant BUHMBOX results may have been due to insufficient power. Power to detect moderate heterogeneity (proportion of cohort who actually are genetically distinct), i.e. a true underlying heterogeneity proportion of $\pi = 0.2$, approaches 100% when the number of cases is greater than ~1,500, or when the number of risk SNPs is greater than 50 (see (Han et al., 2016) Figure 3.).

In these analyses the minimum number of cases is 3, 455, and minimum number of trait-associated SNPs (risk SNPs) is 120. However, the true subgroup heterogeneity proportion is unknown, though it may be acceptable to assume moderate heterogeneity in MDD, as clinical heterogeneity in MDD can be estimated at 25-30% as based on the typical/ atypical symptom profile framework (Penninx et al., 2013). Clinical heterogeneity in chronic pain is less easy to estimate, as unlike in MDD there is no agreed 'single' clinical diagnosis of chronic pain and a lot of study is on chronic pain disorders rather than of chronic pain as a disease in itself (see 1.1.2 and 1.3.3.2.1).

Therefore, these analyses may be underpowered due to low proportion of 'true' underlying heterogeneity, as heterogeneity proportion is unknown but estimated as moderate. In GWAS of CPG no SNPs were found to be associated with the trait at genome-wide significance, which may also mean BUHMBOX analyses are underpowered.

Future steps may include use of larger, more well-powered independent chronic pain and MDD GWASs. Repetition of BUHMBOX analyses using depressive symptoms as opposed to GWASs of MDD itself may also be of interest – it may be more likely that chronic pain is misclassified as a depressive symptom as opposed chronic pain being misdiagnosed as MDD.

Previous analyses using BUHMBOX use phenotypes such as serotypes of rheumatoid arthritis, which are distinct disorders with clear clinical differences, where participants or patients can be logically classified as a case or control. In contrast it may not be ideal to consider chronic pain as a case-control phenotype, and in addition to this, chronic pain phenotyping varies widely (see 1.1.2). Additionally, BUHMBOX is not agnostic: this analysis only tests for clinical heterogeneity with respect to a second phenotype chosen *a priori*. In other words, it is not possible to test for presence of any clinical heterogeneity in general within a phenotype.

Chapter 4 Common Genetic Variation Associated with Chronic Pain and Shared with Phenotypes of Interest

4.1 Introduction

This chapter specifically addresses objectives 1: To uncover common genetic variation associated with chronic pain phenotypes, and 2: To investigate genetic correlation and pleiotropy between MDD and chronic pain. Analyses carried out in this chapter have been published as part of an article in PLOS Genetics (Johnston, Adams, Nicholl, Ward, Strawbridge, Ferguson, et al., 2019).

As previously discussed, (1.1, 1.3.3, 1.3.4), chronic pain is a complex trait, and few large-scale genetic studies of chronic pain exist. Chapter 3 made use of one of these few large-scale GWAS studies of chronic pain (defined as chronic pain grade) along with existing MDD GWAS summary statistics to investigate pleiotropy, and to identify SNPs associated with Chronic Pain Grade. In contrast, in this chapter a new chronic pain phenotype, MCP, was defined in UK Biobank (see 2.3.2.1), and a GWAS was carried out to find common genetic variation (SNPs) associated with MCP. The summary statistics generated from this GWAS were then used to conduct linkage disequilibrium score regression analyses (LDSR) (2.2.5) examining genetic correlation between MCP and a range of other traits, including MDD.

It can be argued that to understand genetic variation that contributes to vulnerability to, development and maintenance of chronic pain it is more powerful to examine measures of chronic pain as complex neuropathological traits in themselves. This contrasts with GWAS of chronic pain in specific body sites, or of disorders and diseases where chronic pain is a major component such as fibromyalgia and migraine. This view of chronic pain as a disease entity is also in line with recent IASP definitions of Chronic Primary Pain for the ICD-11 (Nicholasa et al., 2019; Treede et al., 2019), and an IASP update on the definition of pain in general (see also 1.1.1). MCP represents a quasi-quantitative chronic pain phenotype, with the aim of examining underlying chronic pain on a continuous scale rather than by threshold, specific body site, or associated specific chronic pain disorder. In keeping with recent IASP publications and previous research on chronic pain cut off points, derivation of

the phenotype MCP aims to capture underlying vulnerability to the development of chronic pain and potential genetic factors associated with maintenance of chronic pain.

The range of phenotypes chosen for genetic correlation analyses was based on previous association evidence (see 1.1.5) and represents a small fraction of all possible correlations that could be calculated. Exploring the full constellation of genetic correlations between this new chronic pain phenotype and other traits and disorders is of interest, but somewhat beyond the scope of this thesis where the main genetic correlation-addressing objective is in relation to major depression.

4.2 Methods

4.2.1 Chronic Pain Phenotyping in UK Biobank

UK Biobank participants were asked via a touchscreen questionnaire at baseline about "pain types experienced in the last month" (field ID 6159), with possible answers: 'None of the above'; 'Prefer not to answer'; pain at seven different body sites (head, face, neck/shoulder, back, stomach/abdomen, hip, knee); or 'all over the body'.

MCP was defined as the sum of body sites at which chronic pain (at least 3 months duration) was recorded: 0 to 7 sites. Chronic pain phenotyping in UK Biobank is discussed further in 2.3.3.1. Those who answered that they had chronic pain 'all over the body' were excluded from the MCP GWAS, as were 10,000 randomly selected individuals reporting no chronic pain. These participants were used in a secondary GWAS of chronic widespread pain in this chapter, and as cases and controls, respectively, in subsequent polygenic risk score (PRS) analyses (see Chapter 5).

4.2.2 Genome-Wide Association Study of Multisite Chronic Pain

SNPs with an imputation quality score of less than 0.3, Minor Allele Frequency (MAF) < 0.01 and Hardy-Weinberg equilibrium (HWE) test $p < 10^{-6}$ were removed from the analyses. Participants whose self-reported sex did not match their genetically determined sex, those who had putative sex-chromosome aneuploidy, those considered outliers due to missing heterozygosity, those with more than 10%

missing genetic data and those who were not of self-reported white British ancestry were excluded from analyses. A list of such "poor quality" samples (due to these reasons) was derived by Bycroft et al (Bycroft et al., 2018) and was used here as part of genetic quality control. Briefly, putative sex chromosome aneuploidy was defined by visual inspection of scatterplots of mean log2 ratio (L2R) on X and Y chromosomes, and 652 UKB participants meet these criteria for putative sex-chromosome aneuploidy (Supplemental Information S 3.6 (Bycroft et al., 2018)). Samples with a population-structure-adjusted heterozygosity value above the mean heterozygosity value (0.1903) and missing rate greater than 0.05 as computed using PLINK '—miss' command were also flagged as potentially poor quality ((Bycroft et al., 2018); 968 such samples are listed in this paper's Supplemental Information S 3.5.3).

These exclusions are a standard part of GWAS analysis (Coleman et al., 2016; Marees et al., 2018), and represent indications of sample contaminations, genotyping error, inbreeding, markers under significant selection, or markers which are rare variants – these conditions would mean the statistical assumptions necessary for GWAS would be violated. These genetic quality control measures left a subset of "hard-called" PLINK-format genotypes (SNPs) of 615, 839 (see BOLT-LMM manual 5.1.2) on which the mixed model was built. GWAS was then carried out using BOLT-LMM under an infinitesimal model (see 2.2.1.3), adjusting for age, sex and chip (genotyping array). Relatedness and population stratification were adjusted for within the BOLT-LMM model via use of a Genetic Relatedness Matrix (GRM), and age was found to have a relationship with MCP conforming to linearity. Genomic risk loci were identified via the definition employed by FUMA (Watanabe et al., 2017).

4.2.3 Linkage-Disequilibrium Score Regression

Genetic correlations between MCP and 22 complex traits selected on the basis of epidemiological evidence or suspected relationship with chronic pain (1.1.5) were calculated using LDSR (Bulik-Sullivan et al., 2015), implemented either using the 'ldsc' package (Bulik-Sullivan et al., 2015) and downloaded publicly-available summary statistics and summary statistics from in-house analyses or using LD Hub (Zheng et al., 2017). LD Hub datasets from the categories Psychiatric, Personality, Autoimmune and Neurological were selected and

datasets with the attached warning note ‘Caution: using this data may yield less robust results due to minor departure from LD structure’ were excluded from the analyses. Where multiple GWAS datasets were available for the same trait, the one with the largest sample size and/or European ancestry was retained with priority given to European ancestry, for example, for PTSD multiple different ancestry groups were available and European was selected.

Genetic correlation between MCP and between von Korff chronic pain grade (see 1.1.2 & 2.3.2.2) was also calculated. A secondary GWAS of participants with chronic pain all over the body (termed chronic widespread pain, CWP) versus 10,000 chronic pain-free controls was also carried out (total N = 15,258), and these summary statistics were used in LDSR analyses to calculate genetic correlation between CWP and MCP.

P values for heritability are estimated according to formulae given by Altman & Bland (Altman & Bland, 2011).

4.2.4 Phenotypic Correlations

Phenotypic information on BMI was obtained from baseline measure of BMI (data field 21001). For MDD, anxiety, schizophrenia, autism spectrum disorder, anorexia nervosa and bipolar disorder, UK Biobank data field 20544 “mental health problems ever diagnosed by a professional” (part of the Thoughts and Feelings questionnaire in online follow-up) was used to derive a dichotomous phenotype value. For subjective well-being the UK Biobank data field 20459 “general happiness with own health”, a Likert-like self reported measure of subjective wellbeing where 1 = extremely happy and 6 = extremely unhappy, was used to derive a continuous measure of subjective well-being.

Phenotypic information for the traits rheumatoid arthritis, asthma, primary biliary cirrhosis/cholangitis, inflammatory bowel disease, Crohn’s disease, ulcerative colitis and Parkinson disease was derived from the UK Biobank data field 20002 “non-cancer illness codes, self-reported”. Phenotypic information for PTSD, a psychiatric cross-disorder phenotype, neuroticism, celiac disease, and depressive symptoms was not available within datasets associated with UKB projects to which access was available to for this PhD project.

For all available traits, those who answered, “prefer not to answer”, “do not know”, or who did not have complete phenotypic information on both traits to be used in the phenotypic correlation coefficient calculation were removed before estimation of phenotypic correlation.

For continuous traits, phenotypic correlation coefficients between the trait and MCP were calculated as Pearson’s rho. For dichotomous traits a special case of Pearson’s rho, the point biserial correlation coefficient, was calculated to give an estimate of phenotypic correlation. The point biserial correlation is appropriate for use in estimating correlation when one variable is dichotomous and the other continuous (Kornbrot, 2014; Sheskin, 2000). Note that sample size varied depending upon availability of variables (phenotypic information) in UKB datasets available for use in this PhD project, and due to the fact psychiatric trait phenotype information was derived from the Thoughts and Feelings online follow-up data which was only completed by a subset of the UK Biobank sample (max N = 157, 366) (see also 2.3.4).

4.3 Results

4.3.1 Description of Participants

A total of 387, 649 UK Biobank participants with a mean age of ~56 years old and 53.9% of whom were female were included in the MCP GWAS analysis (Table 4.1).

MCP	total N (%)	male N (%)	female N (%)	age (mean)
0	218622 (56.4)	105474 (48.2)	113148 (51.8)	56.71
1	92718 (23.92)	42734 (46.1)	49984 (53.9)	57.03
2	44612 (11.51)	18612 (41.7)	26000 (58.3)	57.29
3	20147 (5.2)	7771 (38.6)	12376 (61.4)	57.65
4	8289 (2.14)	2970 (35.8)	5319 (64.2)	57.48
5	2503 (0.65)	780 (31.2)	1723 (68.8)	56.53
6	652 (0.17)	181 (27.8)	471 (72.2)	56.2
7	106 (0.03)	34 (32.1)	72 (67.9)	56.17
total	387649	178556 (46.1)	209093 (53.9)	56.91

Table 4. 1: Age, sex and MCP phenotype value of UK Biobank participants included in the MCP GWAS.

MCP = MCP phenotype value (0 = no chronic pain).

4.3.2 Common genetic variants and genes associated with MCP

MCP was found to have a moderate SNP-heritability value (h^2_{SNP} = 10.2%). Genome-wide association analyses showed 76 SNPs were associated with MCP at a genome-wide significance p value threshold of 5×10^{-8}, spread across 39 genomic risk loci (Table 4.2). 143 genes were also found to be significantly associated with MCP in gene-level analyses (Appendix 1).

Genomic Risk Locus	rsID (Lead SNP)	Chr	Pos	GWAS p
1	rs10888692	1	50991473	5.30E-09
2	rs197422	1	112317512	2.00E-09
3	rs59898460	1	150493004	9.20E-12
4	rs12071912	1	243241614	5.30E-09
5	rs4852567	2	80703379	4.30E-08
6	rs7628207	3	49754970	8.40E-10
7	rs28428925	3	107294634	1.40E-09
8	rs6770476	3	136073920	9.40E-09
9	rs34811474	4	25408838	2.70E-11
10	rs13135092	4	103198082	1.50E-13
11	rs13136239	4	140908755	3.60E-08
12	rs6869446	5	65570607	9.50E-09
13	rs1976423	5	104042643	8.20E-09
14	rs17474406	5	122732342	2.40E-08
15	rs1946247	5	160836620	4.90E-08
16	rs11751591	6	33794215	2.70E-10
17	rs6907508	6	34592090	1.10E-08
18	rs6926377	6	145105354	7.90E-09
19	rs10259354	7	3487414	3.00E-08

20	rs7798894	7	21552995	1.60E-08
21	rs6966540	7	95727967	3.30E-08
22	rs12537376	7	114025053	1.70E-09
23	rs11786084	8	142651709	2.30E-08
24	rs10992729	9	96181075	1.10E-09
25	rs6478241	9	119252629	3.10E-09
26	9:140251458_G_A	9	140251458	5.30E-14
27	rs2183271	10	21957229	3.10E-08
28	rs11599236	10	106454672	3.30E-08
29	rs12765185	10	134977077	3.90E-08
30	rs61883178	11	16317779	2.00E-10
31	rs1443914	13	53917230	2.80E-11
32	rs12435797	14	73797669	3.70E-08
33	rs2006281	14	104327732	3.40E-08
34	rs2386584	15	91539572	2.80E-11
35	rs285026	16	77100089	1.90E-08
36	rs11871043	17	43172849	1.70E-09
37	rs11079993	17	50301552	5.70E-12
38	rs62098013	18	50863861	4.00E-11
39	rs2424248	20	19650324	3.70E-10

Table 4. 2: Genomic Risk Loci.

Chr = chromosome, pos = position (basepairs), GWAS p = p value for lead SNP association with MCP.

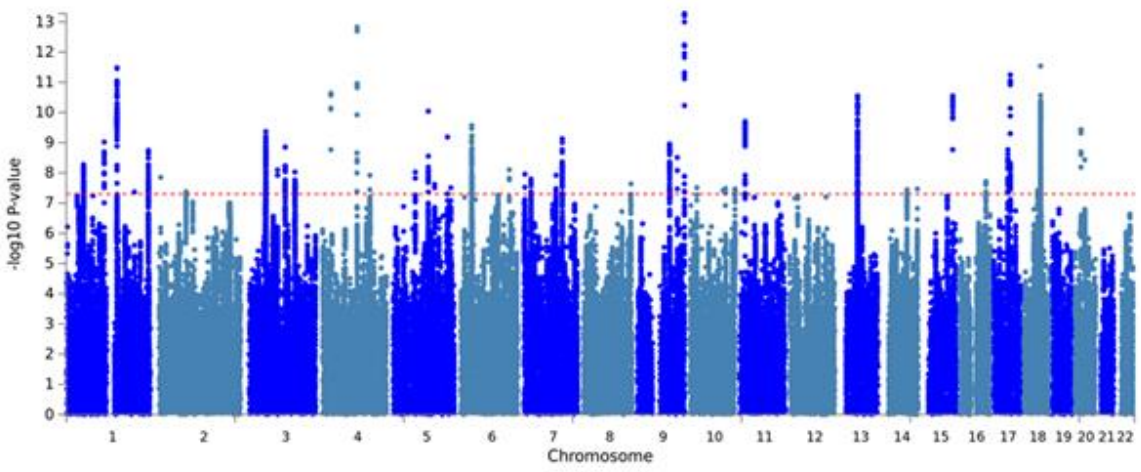

Figure 4. 1 : MCP GWAS Manhattan plot

-log10P-value = transformed p values (-log base 10) for SNP-trait association. The red dotted line indicates the significance threshold (~7 i.e. $p < 5 \times 10^{-8}$).

Genes associated with MCP were also found to be significantly enriched in biological pathways through MAGMA gene set analyses (see 2.3.1.1). These MsigDB (Molecular Signatures Database) (Liberzon et al., 2015) C2 (curated gene set) canonical pathways were *DCC*-mediated attractive signalling, PLC-ß-mediated events, BCR signalling and $\alpha_6\beta_4$ and $\alpha_6\beta_1$ integrin signalling (Appendix 1 Table A1.3). Function of these genes is also summarised in 4.4.3.3.

4.3.3 Genetic Correlations

MCP was significantly genetically correlated with a range of psychiatric disorders and phenotypes, notably MDD, depressive symptoms, anxiety, PTSD, and schizophrenia (Table 4.3, Figure 4.1). MCP was not found to be significantly genetically correlated with inflammatory bowel diseases (Crohn's disease, ulcerative colitis, inflammatory bowel disease) or other autoimmune diseases Celiac disease and systemic lupus erythematosus. The only psychiatric phenotype examined which was not significantly genetically correlated with MCP was bipolar disorder. Rheumatoid arthritis, an autoimmune disease associated with significant chronic pain (Walsh & McWilliams, 2014), was significantly genetically correlated with MCP, but with an r_g value of only 16% (Table 4.3). In contrast genetic overlap between MCP and MDD, depressive symptoms, and neuroticism ranged from 40-59% (Table 4.3).

There were no genome-wide significant SNP associations in the second chronic widespread pain (CWP) GWAS, likely due to the fact the sample size is too small

for significant power, but the summary statistics could still be taken forward for LDSR analysis – this is another route to assessing the relationship between chronic widespread pain and MCP, in contrast to the polygenic risk score analyses described in Chapter 5.

Trait	r_g	se	z	h^2	h^2p (FDR)	source	PMID	Category	p	p (FDR)
MDD	0.53	0.03	18.92	0.077	1.25×10^{-47}	PGC	29700475	psychiatric	7.68×10^{-80}	1.69×10^{-78}
Depressive symptoms	0.59	0.03	17.16	0.047	6.87×10^{-29}	ld_hub	27089181	psychiatric	5.63×10^{-66}	6.19×10^{-65}
BMI	0.31	0.02	15.69	0.138	5.42×10^{-59}	GIANT consortium	25673413	anthropometric	1.90×10^{-55}	1.39×10^{-54}
Neuroticism	0.4	0.03	11.9	0.089	3.66×10^{-26}	ld_hub	27089181	personality	1.24×10^{-32}	6.82×10^{-32}
Subjective well being	-0.36	0.04	-8.94	0.025	2.77×10^{-32}	ld_hub	27089181	psychiatric	3.78×10^{-19}	1.66×10^{-18}
Low Relative Amplitude	-0.3	0.05	-6.37	0.053	3.03×10^{-13}	In-house analysis	30120083	circadian	1.91×10^{-10}	7.00×10^{-10}
Rheumatoid Arthritis	0.16	0.03	4.7	0.160	7.41×10^{-8}	ld_hub	24390342	autoimmune	2.64×10^{-6}	8.30×10^{-6}
Anxiety (Case-Control)	0.49	0.11	4.53	0.081	0.00405	PGC	26754954	psychiatric	5.91×10^{-6}	1.63×10^{-5}
Schizophrenia	0.1	0.03	4.08	0.443	6.56×10^{-79}	PGC	25056061	psychiatric	4.50×10^{-5}	1.10×10^{-4}
Asthma	0.22	0.06	3.63	0.123	3.53×10^{-6}	ld_hub	17611496	autoimmune	3.00×10^{-4}	6.60×10^{-4}
PGC cross-disorder analysis	0.13	0.04	3.54	0.172	7.89×10^{-36}	ld_hub	23453885	psychiatric	4.00×10^{-4}	8.00×10^{-4}
PTSD (European Ancestry)	0.41	0.12	3.28	0.097	0.030855	PGC	28439101	psychiatric	0.001047	1.92×10^{-3}
Autism spectrum disorder	-0.1	0.04	-2.22	0.451	9.38×10^{-17}	ld_hub	NA	psychiatric	0.026	0.0443
Primary biliary cirrhosis	0.1	0.04	2.17	0.376	1.11×10^{-8}	ld_hub	26394269	autoimmune	0.03	0.047
Anorexia Nervosa	-0.06	0.03	-2.14	0.556	2.18×10^{-63}	ld_hub	24514567	psychiatric	0.032	0.0471
Inflammatory Bowel Disease (European Ancestry)	0.05	0.03	1.75	0.333	9.17×10^{-21}	ld_hub	26192919	autoimmune	0.08	0.1101
Celiac disease	-0.07	0.05	-1.49	0.314	2.50×10^{-10}	ld_hub	20190752	autoimmune	0.136	0.1756
Crohn's disease	0.04	0.03	1.35	0.504	2.65×10^{-17}	ld_hub	26192919	autoimmune	0.179	0.2125
Systemic lupus erythematosus	0.06	0.04	1.33	0.390	9.77×10^{-9}	ld_hub	26502338	autoimmune	0.184	0.2125
Ulcerative colitis	0.04	0.04	1.08	0.257	1.19×10^{-14}	ld_hub	26192919	autoimmune	0.281	0.3094
Bipolar disorder	-0.02	0.04	-0.66	0.436	5.51×10^{-29}	ld_hub	21926972	psychiatric	0.509	0.5329

Parkinson's disease	0.0	0.04	0.05	0.409	0.000761	ld_hub	19915575	neurological	0.961	0.9612

Table 4. 3: Genetic correlation results.

r_g = genetic correlation coefficient value, se = standard error of correlation value, z = z value, h^2 = SNP-heritability value, h^2 p (FDR) = p value (FDR-corrected) for SNP-heritability, source = source of GWAS summary statistics, PMID = PubMed ID of associated paper (if applicable), p = p value for genetic correlation coefficient, p(fdr) = FDR-corrected p value for genetic correlation coefficient.

Trait	r_g	r_p	p	N
MDD	0.53	0.15	$< 2 \times 10^{-16}$	155570
Depressive symptoms	0.59	NA	NA	NA
BMI	0.31	0.148	$< 2 \times 10^{-16}$	491364
Neuroticism	0.4	NA	NA	NA
Subjective well being	-0.36	0.26	$< 2 \times 10^{-16}$	155653
Low Relative Amplitude	-0.3	-3.13×10^{-4}	0.925	91077
Rheumatoid Arthritis	0.16	0.055	$< 2 \times 10^{-16}$	155570
Anxiety (Case-Control)	0.49	0.1039	$< 2 \times 10^{-16}$	155570
Schizophrenia	0.1	-0.0032	0.212	155570
Asthma	0.22	0.06	$< 2 \times 10^{-16}$	155570
PGC cross-disorder analysis	0.13	NA	NA	NA
PTSD (European Ancestry)	0.41	NA	NA	NA
Autism spectrum disorder	-0.1	0.0075	0.0033	155570
Primary biliary cirrhosis	0.1	0.007	0.0032	155570
Anorexia Nervosa	-0.06	0.018	4.35×10^{-13}	155570
Inflammatory Bowel Disease (European Ancestry)	0.05	0.007	0.0045	155570
Celiac disease	-0.07	NA	NA	NA
Crohn's disease	0.04	0.018	1.04×10^{-12}	155570
Systemic lupus erythematosus	0.06	0.02	1.73×10^{-18}	155570
Ulcerative colitis	0.04	0.014	1.44×10^{-8}	155570
Bipolar disorder	-0.02	0.015	9.02×10^{-9}	155570
Parkinson's disease	0.0	0.013	3.71×10^{-7}	155570

Table 4. 4 : Phenotypic correlations between MCP and traits of interest

r_p = phenotypic correlation coefficient (Pearson's rho/ point biserial correlation coefficient), r_g = genetic correlation coefficient for comparison, p = p value associated with phenotypic correlation coefficient, N = sample size for phenotypic correlation estimation.

The genetic correlation between von Korff chronic pain grade and MCP was large and significant at $r_g = -0.78$ ($p = 3.46 \times 10^{-13}$), but negative (Figure 4.1). The genetic correlation between chronic widespread pain and MCP was significant and positive at $r_g = 0.83$ ($p = 2.4 \times 10^{-54}$) (Figure 4.1).

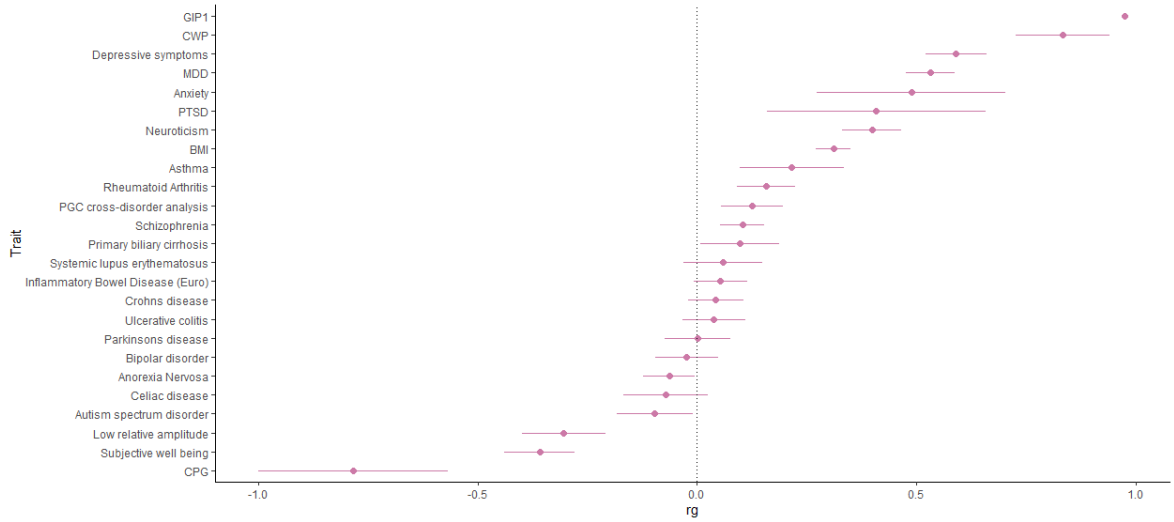

Figure 4. 2: Genetic correlations between MCP and a range of traits.

GIP1 = genetically independent phenotype 1 (see Appendix 3). CWP = chronic widespread pain. Error bars indicate 95% confidence interval (estimated as +/- 2 x standard error of the genetic correlation value r_g).

Overall, psychiatric phenotypes, particularly MDD and depressive symptoms, shared the largest and most statistically significant proportions of common genetic variation with MCP. Many conditions commonly associated with significant chronic pain, including inflammatory bowel diseases and systemic lupus erythematosus, showed no genetic overlap with MCP.

4.4 Discussion

4.4.1 Genetic Correlations with MCP and Traits of Interest

A range of traits of interest, either found previously to be associated with chronic pain in the literature, or with potential involvement of or association with chronic pain but with inconclusive evidence from past epidemiological studies, were chosen for LD-score regression analysis (see 1.1.5).

4.4.1.1 Psychiatric Traits

The psychiatric phenotype most significantly genetically correlated with MCP was MDD (r_g = 0.53) while the largest significant genetic correlation coefficient was for MCP and depressive symptoms (r_g = 0.59). This matches closely with a genetic correlation value for chronic pain grade and MDD found by McIntosh et al via a mixed-modelling approach (rho = 0.53). MCP was also positively genetically

correlated with neuroticism (r_g = 0.40), anxiety (r_g = 0.49), schizophrenia (r_g = 0.10), cross-disorder psychiatric phenotype (r_g = 0.13) and PTSD (r_g = 0.41). Significant negative genetic correlation was found between MCP and anorexia nervosa, autism spectrum disorder, and between MCP and subjective well-being. There was no significant genetic correlation between MCP and bipolar disorder. The genetic overlap between schizophrenia and MCP is somewhat in contrast to findings indicating people with schizophrenia tend to show less sensitivity to ongoing or chronic pain compared to the general population (1.1.5), and may indicate that these differences could be due to environmental factors, including difficulties in reporting pain for people with schizophrenia. A lack of significant genetic correlation between bipolar disorder and MCP may also indicate non-genetic factors drive the overlap in bipolar disorder and chronic pain. The genetic overlap between MCP and psychiatric disorders and traits, particularly MDD and depressive symptoms, emphasises the psychological and affective component to chronic pain.

4.4.1.2 Autoimmune Traits

Autoimmune disorders and disorders with a significant autoimmune component such as rheumatoid arthritis, asthma and primary biliary cholangitis showed positive genetic correlation with MCP. However, gastrointestinal autoimmune disorders ulcerative colitis and Crohn's Disease did not. This suggests that distinct genetic variation and mechanisms underlie chronic pain associated with these disorders compared to those outwith the digestive system. Pain related to inflammatory bowel diseases may represent something less 'chronic' and more 'on-going acute', as stricture, abscesses and partial or complete obstruction of the small bowel result in pain (Docherty et al., 2011). Structural and functional brain changes associated with the transition to chronic pain may also play a less central role in gastrointestinal autoimmune disorder-associated pain, due to potential for the enteric nervous system to act independently from the CNS, and the role of the gut-brain axis in chronic abdominal pain (Carabotti et al., 2015; Cryan & Dinan, 2012). In addition, Crohn's disease is associated with pain not only viscerally and in relation to disease exacerbation, strictures and abscesses, but also with pain in the joints (arthritis) and back, which for some individuals never goes away even with remission or successful management of active Crohn's disease (Norton et al., 2017). A GWAS of Crohn's or other inflammatory

bowel disease may therefore not capture genetic variation involved in pain and chronic pain, but instead significantly associated variation is related to disease activity and inflammation more specific to the digestive system.

There was no significant genetic correlation between MCP and systemic lupus erythematosus (SLE). This may be, again, because pain associated with SLE is complex and multifactorial, and can vary between individuals with the same SLE diagnosis (Waldheim et al., 2018). SLE involves multiple body systems – specific types of arthritis can be involved, neuropathic pain or headache syndromes are often experienced, and SLE is also associated with pericarditis and Raynaud's (Fava & Petri, 2019). Pain therefore likely depends at least in part on organ system and active disease, and SNPs associated with SLE will not necessarily be associated with chronic pain. Similarly to other conditions (Crohn's, osteoarthritis, rheumatoid arthritis), pain in SLE has been found not to "track with disease activity" (Fava & Petri, 2019).

4.4.1.3 Parkinson's Disease

Chronic pain is often reported in those with neurological diseases (Borsook, 2012), including Parkinson's disease (Ford, 2012; Simuni & Sethi, 2008), reaching prevalence of 60% in certain patient populations. However, no significant genetic correlation was found between Parkinson's disease and MCP ($p = 0.96$), suggesting other factors associated with Parkinson's disease may contribute to chronic pain, as opposed to shared genetic contributions to risk for both disorders. Some of the pain experienced by those with Parkinson's disease may also be distinct from an unexplained chronic pain or widespread chronic pain, and may be related to muscle rigidity, dystonia, reduced movement in the joints, changes in posture and associated radicular pain due to trapped nerves – this pain could be viewed as chronic in the sense that potentially causal contributory factors are chronic in nature but may be distinct from chronic pain without comorbid Parkinson's disease. Again, GWAS of Parkinson's disease may reveal SNPs associated with disease activity and progression, rather than with chronic pain experienced as part of Parkinson's disease.

4.4.1.4 Neurodevelopmental, Circadian and Anthropometric Traits

Significant negative genetic correlation was found between autism spectrum disorder and MCP. This negative genetic correlation between autism spectrum disorder and MCP may suggest differences observed between autistic people and the general population in terms of integrating bodily signals (interoception) lead to reduced experience of or reporting of chronic pain. Alternatively, increased prevalence of autism spectrum disorder diagnosis in men (Halladay et al., 2015), who tend to have chronic pain at reduced rates compared to women (see 1.2.2), may drive this negative genetic correlation value.

Significant negative genetic correlation was found for low relative amplitude and MCP, which is unexpected: low relative amplitude is a circadian rhythmicity and health phenotype that indicates poor circadian regulation, which is associated with a range of poor health outcomes. A PRS for low relative amplitude was significantly associated with mood instability, MDD and neuroticism (Ferguson et al., 2018). The fact that shared common genetic variation between MCP and low relative amplitude is associated with opposing directions of effect in these two disorders may indicate that the association between poor sleep and circadian rhythm and chronic pain is instead driven by other lifestyle factors, rather than shared genetic factors predisposing to increased risk for both chronic pain and low relative amplitude. There may also be a significant underrepresentation of people with chronic illness and chronic pain amongst the sub-sample of UK Biobank (N = 71, 500) who took part in activity monitoring, introducing potential bias into SNP effect value estimates.

Significant positive genetic correlation was found between BMI and MCP, indicating that a moderate proportion of variants are shared between MCP and BMI, and contribute to an increase in both BMI and chronic pain. This is in line with work linking increased BMI (obesity) and adiposity to immune activation and chronic inflammation, which play a key role in pain perception and development of chronic pain.

4.4.1.5 Chronic Widespread Pain

Chronic widespread pain and MCP were strongly positively correlated, but the genetic correlation value was significantly different to 1– this may be due to

small sample size of the CWP GWAS or may mean that these are subtly different phenotypes with a large genetic overlap. For example, Nicholl et al found that multisite pain was associated with MDD and bipolar disorder severity except at the 'extreme' – chronic pain 'all over the body' (= chronic widespread pain) was less strongly associated with MDD and bipolar disorder severity in comparison to chronic pain at 4-7 body sites. Other work also suggests that chronic widespread pain represents a distinct phenotype in comparison to chronic pain in general (Gerhardt et al., 2016b; Mansfield et al., 2017). However, it may be the true 'clinical reality' that there are no natural or logical cut-off points for localised versus widespread chronic pain (Kamaleri et al., 2008). Traits with lower genetic correlation values are commonly used as proxies for one another e.g. educational attainment as proxy for intelligence (r_g ~ 70%) (Savage et al., 2018), or current age as a proxy for life span (r_g ~40-70%) (K. M. Wright et al., 2019). It may therefore be acceptable to say r_g = 0.83 means that these GWASs capture genetic variation associated with the same underlying construct, and this definition of chronic widespread pain (chronic pain all over the body for 3+ months) could be added as the maximum trait value of MCP.

4.4.1.6 Chronic Pain Grade

Genetic correlation coefficient value between MCP and CPG was significant and large, but negative. This suggests most shared SNPs between the two conditions are associated with opposing direction of effect i.e., most SNPs are associated with an increase in trait value in MCP and decrease in trait value in CPG, or vice versa. Another important difference in the GWASs of the two chronic pain traits is that CPG was adjusted for both body-mass index (BMI) and an employment-related variable (manual labour). This difference may drive the unexpected differences in effect sign for MCP and CPG-associated variants, and a difference in trait-associated variants may additionally be generated because BMI is also heritable and is genetically correlated (as demonstrated above) with chronic pain.

Adjusting for genetically correlated traits, such as adjusting for BMI in a GWAS of chronic pain, can bias results (Aschard et al., 2015; Vansteelandt et al., 2009). If the relationship between a genetic variant (SNP), covariate of interest, and outcome (i.e. GWAS trait) is as shown in Figure 4.2 A below, then the adjusted

GWAS estimate of effect size represents the direct effect (or measures directly the magnitude of association) between SNP and outcome trait, and the unadjusted GWAS estimate gives a total effect (direct + indirect) of SNP on outcome trait – in every other situation or relationship type between SNP, covariate and outcome trait (Figure 4.2 B-D) the adjusted estimate may be biased (Aschard et al., 2015).

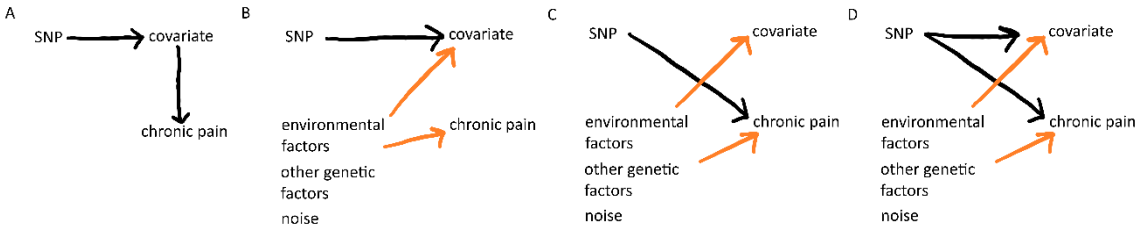

Figure 4. 3: Sources of bias in GWAS.

Based on Figure 1 (Aschard et al., 2015). Possible relationships between SNP, covariate and outcome variable in GWAS.

Aschard et al suggest that adjustment for environmental or demographic traits would not likely introduce bias as they do not have genetic associations – however, recent GWASs including of demographic traits such as socioeconomic status (Hill et al., 2019) indicate that genetic variation is in fact associated with traits previously considered environmental/ demographic and therefore 'safe' as GWAS covariates. Adjustment for manual labour may therefore also introduce bias into GWAS effect size estimates in the Pfizer-23andMe CPG GWAS.

As a sensitivity analysis, a second MCP GWAS was carried out, identical in every way to the main analyses except with adjustment for BMI – the genetic correlation results with CPG remained the same, suggesting adjustment for manual labour may contribute to the negative correlation value.

Departure of the genetic correlation value from 1 may also be due to differences in the trait concepts of CPG and MCP themselves – CPG takes account of disability and the impact of chronic pain on daily life and functioning (Smith et al., 1997; Von Korff et al., 1992), whereas MCP simply sums chronic pain sites. It has been theorised that phenotypic correlations between traits tend to reflect genetic correlations, to the extent that phenotypic correlations can be used as proxies of genetic correlation (Cheverud's conjecture) – if phenotypic correlation between CPG and MCP were found to be negative, a negative genetic correlation

as found here may not be unexpected after all. This was explored using Generation Scotland participant data (see Appendix 2), where it was found that phenotypic correlation between CPG and MCP was significant and positive (~0.3). This again suggests that there may be potential bias in the GWAS output of the 23andMe-Pfizer GWAS due to the inclusion of manual labour as a covariate, which may be the cause of unexpected negative genetic correlation between CPG and MCP.

4.4.2 Heritability and Polygenicity of Multisite Chronic Pain

MCP was found to be moderately heritable. This reduction in heritability value when comparing SNP-heritability (a narrow-sense heritability) with twin study derived estimates of heritability (a broad sense heritability measure) is to be expected (see 2.2.1.1). This heritability value is of similar magnitude to recent SNP-heritability estimates of MDD (8.9% (Howard et al., 2019)). Results also indicated a high degree of polygenicity, shown through MAGMA gene-level analysis.

4.4.3 Genes of Interest Associated with MCP

Genes found to be associated with MCP through MAGMA gene-level analyses suggested CNS involvement in chronic pain, with genes found to be involved in processes such as synaptic connectivity (*SDK1*) (Yamagata & Sanes, 2008) and glial-guided neuronal migration (*ASTN2*) (Wilson et al., 2010). Genes associated with MCP were also involved in Notch signalling pathway and implicated in neurogenesis and CNS plasticity (*NUMB, MAML3*) (Ables et al., 2011; Andersson et al., 2011; Kitagawa, 2015), in non-UK Biobank studies. Several MCP-associated genes were also involved in immune processes, cell cycle regulation, protein degradation, and apoptosis. A full list of genes associated with MCP is discussed in Appendix 1.

4.4.3.1 Associations with Other Chronic Pain Conditions

Five of the 143 genes significantly associated with MCP are also listed in the Human Pain Genetics Database (Meloto et al., 2018), an online repository documenting genetic contributors to chronic pain and chronic pain conditions. These genes, *ASTN2, SLC24A3, RABGAP1L, F2* and *FHL5* have been previously associated with migraine (Anttila et al., 2013; Gormley et al., 2016; Rodriguez-

Acevedo et al., 2015), a chronic pain condition where large and well-powered GWAS have been previously carried out. *DCC* and *SOX5* (which jointly functions with *SOX6* in chondrogenesis) have been associated with chronic back pain (Suri et al., 2018), *GABRB3* (encoding one of three beta subunits of the GABA A receptor along with *GABRB2*) has been associated with migraine and fibromyalgia (Zorina-Lichtenwalter et al., 2016). *AMIGO3, SLC39A8, ECM1, EXD3* and *FOXP2* have been associated with a musculoskeletal pain phenotype (Tsepilov et al., 2020) (see also Appendix 3) in addition to MCP.

Genes associated with chronic pain related phenotypes in previous candidate gene studies including *COMT, OPRM1, GCH1* and *BDNF* were not found to be associated with MCP – this could be due to the pitfalls of candidate gene studies generally, and is in keeping with general inconsistency/ lack of replication for candidate gene study findings (Mogil, 2012). For example, the association between *COMT* and individual differences in pain perception was originally found in studies of healthy individuals or fibromyalgia patients exposed to pain in an experimental setting, with relatively small sample sizes N ~ 29-202 (Diatchenko et al., 2005; Martínez-Jauand et al., 2013; Zubieta et al., 2003), which were likely not powerful enough for discovery of trait-associated common genetic variation. In addition, while these studies may indicate pain perception differences associated with *COMT* haplotypes or polymorphisms, pain perception, particularly in response to acute pain challenges delivered in an experimental setting, may not be equivalent to chronic pain.

None of the genes involved in CIP, erythromelalgia or PEPD (*SCN9A, FAAH, NTRK1, PRDM12)* were found to be associated with MCP. This suggests that CIP and chronic pain (MCP) are distinct, despite the role CIP-associated genes play in the perception of pain (as discovered through mutations leading to CIP), this is different to chronic pain.

4.4.3.2 Associations with Other Disorders

Several MCP-associated genes have been previously implicated in other traits and disorders, which was explored by manually searching GeneCards (Stelzer et al., 2016) and Online Mendelian Inheritance in Man (OMIM) (McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University) databases with gene names as search terms. Disorders such as Brugada Syndrome 9 and Spinal ataxia

19 & 22 (*KCND3*) (Duarri et al., 2012; Giudicessi et al., 2011; Y. C. Lee et al., 2012), systemic lupus erythematosus (SLE) (Y RNAs) (Kowalski & Krude, 2015), Joubert syndrome 31 and short-rib thoracic dysplasia 13 (*CEP120*) (Roosing et al., 2016), were found to be associated with genes which were also associated with MCP (relevant genes given in parentheses). Genes associated with MCP were also found to be associated with Amyotrophic lateral sclerosis (ALS) (*FAF1*) (Baron et al., 2014), Urbach-Wiethe disease (*ECM1*) (Hamada et al., 2003; Oyama et al., 2003), cohesinopathies associated with intellectual disability as well as Cornelia de Lange Syndrome (*STAG1*) (Lehalle et al., 2017; Liu & Krantz, 2009), split hand/ split foot malformation (*DYNC1I1*) (S. H. Roberts et al., 1991; Tayebi et al., 2014), and a wide range of cancers (*PRC1*) (J. Li et al., 2018). Schizophrenia (*GABRB2*) (Laroche et al., 2008; T. Li et al., 2013; Lo et al., 2007; Petryshen et al., 2005; Sanjuá et al., 2006; Tolosa et al., 2010; Yeung et al., 2018; Yin et al., 2018), intellectual disability and epilepsy (*GABRB2*) (Srivastava, Cohen, Pevsner, et al., 2014), and neuroleptic-induced tardive dyskinesia (*GABRB2*) (Inada et al., 2008) were also found to be associated with MCP-related genes.

These disorders can be roughly grouped according to pathogenic similarities, which short-rib thoracis dysplasias, electrodactyly and Urbach-Wiethe disease involving musculoskeletal and soft tissue malformations. Short-rib thoracic dysplasias are a group of autosomal recessive ciliopathies, associated with short ribs, abnormalities of the hip joint, and potential involvement of other organs and tissues (Schmidts et al., 2013). Split hand/ split foot malformation (electrodactyly) can be caused by many different mutations, and can be inherited singly or as a symptom of a congenital syndrome, with failure to maintain signalling from and typical development of the median apical ectodermal ridge (AER) (a structure at the distal end of limb buds coordinating limb development) identified as a main mechanism of pathogenesis (Tayebi et al., 2014). Urbach-Wiethe disease is a rare, autosomal recessive disorder characterised deposits of collagen in skin and soft tissues. Complications due to these collagen deposists can manifest as papules around the eyes and fingers, and calcification of brain tissue (most often basal ganglia) that can lead to seizures and cognitive changes (Parida et al., 2015).

Several genes associated with MCP were also found to be associated in neurodegenerative disorders with motor function involvement. Spinal ataxias 19

& 22 are rare, progressive, degenerative nervous system diseases associated with cerebellar atrophy and a range of motor coordination, balance and speech related symptoms (Duarri et al., 2012). Joubert syndrome is another condition affecting the cerebellum, and is also associated with motor and cognitive symptoms (Roosing et al., 2016; Uniprot). ALS is a neurodegenerative disorder affecting motor function, with extra-motor symptoms in up to half of cases such as changes in behaviour, executive dysfunction and problems with language (Masrori & Van Damme, 2020).

GABRB2, encodes the GABAA beta-2 subunit protein, a component of ionotropic (neurotransmitter-binding) GABAA receptors which form the major inhibitory system in the brain (Jacob et al., 2008). Dysregulation of this system has been suggested to play a key role in schizophrenia pathogenesis (Lichtshtein et al., 1978), and variants in this gene have since been associated with schizophrenia (Laroche et al., 2008; Lo et al., 2007; Yeung et al., 2018).

Brugada Syndrome 9 is a type of rare heart arrythmia disorder, associated with increased risk of sudden death (Gourraud et al., 2016). Pathology may be a result of sodium channel defects and either concurrent gain of function or loss of function in cardiac potassium or calcium channels, respectively, or of purely sodium-channel related defects (Gourraud et al., 2016).

SLE is an autoimmune disorder associated with significant chronic pain and potential involvement of a range of tissues and organs (Fava & Petri, 2019, see also 1.1.5). Y-RNAs are generally involved in maintenance of typical cell function, and form a part of autoantigen complexes found in serum from individuals with SLE (Driedonks & Nolte-T'Hoen, 2019). Extracellular vesicle exchange (involving these circulating RNAs) is generally important in immune-related processes including inflammation, immune suppression, and tumour micro environment establishment (Driedonks & Nolte-T'Hoen, 2019). Another gene involved in immune processes is *PRC1,* which encodes an evolutionarily conserved Polycomb group (PcG) protein. *PRC1* has been shown to be involved in cancer metastasis through immunosuppressive activities (Su et al., 2019), and this protein is also involved in epigenetic regulation of gene expression and resultant cell fate decisions (Schuettengruber et al., 2017).

Cohesinopathies are caused by mutations in genes coding for components of the cohesion complex, which guides sister chromatid segregation during cell division (Piché et al., 2019). A wide range of symptoms across a variety of associated disorders are associated with malfunctioning of the cohesion complex, including intellectual and growth delays.

Overall, it is difficult to extrapolate from these genes shared between MCP and other disorders in terms of any causal roles these genes may play in either disorder, or in relation to whether this genetic overlap drives any increased chronic pain seen in these disorders (if present). Conclusions as to the mechanisms of chronic pain development also cannot be drawn based on the putative roles these genes play in each disorder. As an extreme generalisation, genes associated with both MCP and the disorders discussed in this section seem to suggest involvement of the CNS, particularly the cerebellum, the immune system, and structural changes in organs and tissues. Dysfunction in the inhibitory system of the brain could also be associated with chronic pain.

4.4.3.3 Function of Genes Associated with MCP

Many genes associated with MCP are implicated in CNS development and functioning. For example, several genes associated with MCP were linked to synapse development and plasticity (*CTNNA2, CEP120, KNDC1, CA10, FOXP2, NRXN1, SLC4A10, LANCL1, SEMA3F*) development of the nervous system (e.g. *AMIGO3, NCAM1*), development of astrocytes (*UTRN*), and peripheral nerve myelination (*DAG1*) (Appendix 1 Table A1.2).

Several genes associated with MCP through MAGMA analyses have been linked to regulation of cell cycle progression, including DNA replication regulation and apoptotic processes. These included *STAG1,* involved in organisation of sister chromatids, genes associated with regulation of the cell cycle (e.g., *ANAPC4, PRC1, BOLL*) and several genes involved in apoptotic processes (e.g., *FAM120A, MON1B, SEMA3F*) (Appendix 1 Table A1.2).

Genes associated with MCP were also found to be involved in a range of immune-related processes, including neutrophil activation *(UTRN),* T-cell activation (*FYN, PABPC4*), and innate immune signalling (e.g., *TRAIP, ILF3, VPS33B*) (Appendix 1 Table A1.2),

Other genes associated with MCP were involved in a wide range of processes including DNA replication regulation (*PURG*), angiogenesis both specific to the brain (*BAI2*) and generally (*F2*), protein transport (e.g., *SORT1, TM9SF4*) degradation (*UBA7, PSMA5*), and repair (*PCMT1*). Several genes associated with MCP are implicated in regulation of gene transcription (e.g., *SMARCC1, ASXL3, AGO2*) and pre-mRNA processing (*PRPF3, PTBP1*) and mRNA processing (e.g., *SNRPC*). Other processes associated with genes found to be associated with MCP included cell development/ differentiation (e.g., *FYN, LEMD2*), and mitochondrial metabolism (e.g., *UQCC2*) and protein synthesis (e.g., *DHX30*). MCP-associated genes were also linked to roles in cell adhesion, migration, and outgrowth (e.g., *LAMB2, RHOA, AMIGO3*) (Appendix 1 Table A1.2),

4.4.3.4 Pathways Enriched for MCP-Associated Genes

PLC-ß-mediated events include immune signalling cascades (Bueno et al., 2006) and synapse formation (Hwang et al., 2005; Südhof, 2018), and disruption is associated with a wide range of conditions including schizophrenia, epilepsy, cancers and autoimmune disease (Yang et al., 2013). BCR signalling coordinates B cell development, and is key for various immune processes (Kurosaki, 2000; Liu et al., 2020). Integrin signalling generally mediates cell-cell adhesion, regulation of gene expression and cell growth, with $\alpha_6\beta_1$ and $\alpha_6\beta_4$ involved in maintaining tissue integrity in muscle, skin and kidney (Anderson et al., 2014).

DCC-mediated attractive signalling is involved in cell migration and motility, including processes such as neuronal haptotaxis (Meijers et al., 2020) and axon guidance (Torres-Berrío et al., 2020), in addition to its role in colorectal and other cancers as a (malfunctioning) tumour suppressor gene (Mehlen & Fearon, 2004). The protein product of *DCC* functions as a receptor, binding Netrin ligands (secreted ligands involved in regulation of axon guidance and migration in addition to roles during development of a wide range of other tissues (Larrieu-Lahargue et al., 2012)) – this signalling and its role as a cue for axon guidance is highly evolutionarily conserved (reviewed by Boyer & Gupton, 2018). The role of *DCC* in neuronal migration is key for brain, particularly cortical, development through coordination of newly born cortical neurons. Commissural axons (axons directed towards the ventral midline of the CNS) that express Dcc proteins on their surface are attracted to Netrin sources, and are repelled in the additional

presence of uncoordinated 5 (UNC5) (reviewed by Meijers et al., 2020). Dcc-Netrin binding has also been found in mouse studies to specifically repel GABAergic neurons from the ventricular zone of the ganglionic eminence (a temporary structure involved in cell and axon migration during foetal development) (Yamagishi et al., 2021). *DCC* may also play a role in synaptic plasticity the adult brain, with studies in rodents showing that deletion of this protein in neurons in the adult forebrain led to loss of long-term potentiation and negative impact on spatial and recognition memory (Horn et al., 2013), and a putative DCC ligand found to be highly expressed in neurogenic brain regions (Yamagishi et al., 2015). Emerging work in humans and rodents also links *DCC* to corpus callosum development through regulation of development of astroglia (Morcom et al., 2021).

DCC has also been associated with psychiatric phenotypes including mood instability (Ward et al., 2019), self-injurious behaviour (Campos et al., 2020), suicidality (Strawbridge et al., 2019), insomnia (Byrne et al., 2013), depression (Li et al., 2020), and to complex brain-related traits such as putamen volume (Satizabal et al., 2019) and intelligence (Savage et al., 2018).

Overall, results indicate MCP is a moderately heritable, polygenic trait, significantly genetically correlated with a range of traits and disorders – most markedly other chronic pain phenotypes and mood disorder phenotypes. Genetic correlation results in particular emphasise that GWAS findings from studies of chronic pain-associated conditions, rather than chronic pain itself, may capture genetic variation associated with disease specific processes rather than pain (as indicated by low/ moderate genetic correlation and in some cases no significant genetic correlation between MCP and conditions associated with significant chronic pain). Findings also suggest a key role for both nervous system and immune-related changes in the development and maintenance of chronic pain and implicate pathways such as *DCC*-mediated attractive signalling which have previously been found to be linked to nervous system development, cell proliferation, and a wide range of psychiatric phenotypes.

Chapter 5 Validation of Multisite Chronic Pain Phenotype

5.1 Introduction

Analyses were carried out to validate the MCP phenotype using PRSs, both in an independent general-population cohort, and in a subset of UK Biobank. This chapter specifically addresses objectives 1: to uncover common genetic variation associated with chronic pain phenotypes, and 2: To investigate genetic correlation and pleiotropy between MDD and chronic pain.

Both utility and validity of trait PRSs can be assessed by testing for association between the trait investigated in the original GWAS, and PRS in an independent cohort. In this chapter, this was achieved through constructing a MCP PRS for a subset of Generation Scotland participants and testing for association between PRS and an MCP-like phenotype within Generation Scotland, and for association between MCP PRS and CPG, a well-validated chronic pain phenotype.

PRS analyses can also be used to explore whether common genetic variation is shared between two different disorders or traits of interest. This type of PRS analysis was undertaken here to examine the relationship between MCP and CWP in UK Biobank a chronic pain phenotype that is potentially genetically distinct from localised chronic pain and from MCP (Kamaleri et al., 2008; Phillips & Clauw, 2011).

In addition, outlined in Chapter 1, chronic pain is more common in women than men. This could be due to a range of genetic and lifestyle factors. Therefore, it is of interest to investigate potential genetics-by-sex interactions in chronic pain, also achievable through PRS analyses.

5.2 Methods

5.2.1 Chronic Pain Phenotyping in Generation Scotland and UKB

Chronic pain phenotyping is similar between Generation Scotland (Smith et al., 2013) and UK Biobank (Sudlow et al., 2015) (see also 2.3.3), but with a few key differences, such as specific body sites used in the questionnaire (Table 5.1). An MCP phenotype was derived in both cohorts, but it was not possible for this phenotype to be identical, due to these differences in the types of questions on

pain that were asked in each of the two studies. CPG score can also be calculated for Generation Scotland participants. In general Generation Scotland participants were asked a greater range of questions on their pain, including questions on social and work-related impact of pain (a total of 24 pain-related question items are present in Generation Scotland, compared to effectively two in UK Biobank).

Generation Scotland	**UK Biobank**
Back	Back
Neck or Shoulder	Neck or Shoulder
Headache, facial or dental pain	Headache
Stomachache/ abdominal	Stomach/ abdominal
Arms, hands, hips, legs, feet (limbs)	Hip
Chest	Facial
Other	Knee
	All over the body

Table 5. 1: Pain site options in Generation Scotland versus UK Biobank

5.2.1.1 Chronic pain grade

CPG (2.3.3.2) phenotype value (0-4) was calculated for each Generation Scotland participant included in PRS analyses (N = 6, 080 total). This sample consists of a subset of Generation Scotland participants who were not related to one another (Generation Scotland was developed using a family-based recruitment structure, see Chapter 2.3.2.3) and who had complete information on CPG phenotype, genotyping data, age, sex, and multidimensional scaling components (MDS) available. The subset of unrelated participants was created by using the GCTA (J. Yang, Lee, et al., 2011) '--grm-cutoff' option to derive a set of individuals related at < 0.025 genetic covariance from the Generation Scotland GRM. Overlapping participants between Generation Scotland and UK Biobank were also removed from this subset.

CPG	Mean age	N male (%)	N total
0	49.44	1692 (64.75)	3708
1	52.92	579 (22.16)	1260
2	51.97	224 (8.57)	711
3	52.72	60 (2.30)	198
4	55.05	58 (2.22)	203
total	50.75	2613 (42.98)	6080

Table 5. 2: Age and sex of participants included in CPG regression analyses.

5.2.1.2 Multisite chronic pain

A chronic pain phenotype similar to UK Biobank MCP was derived in Generation Scotland (see 2.3.3.2.2)) and was calculated for unrelated Generation Scotland participants with complete genotype data and information on age, sex and MDS components. As before, any participants in Generation Scotland who were also participants in UK Biobank were also removed.

MCP	Mean age	N male (% male)	N total
0	49.41	1801 (63.46)	3898
1	53.01	513 (18.08)	1169
2	53.53	329 (11.59)	801
3	55.38	120 (4.23)	418
4	53.67	45 (1.59)	170
5	52.91	25 (0.88)	74
6	52.68	5 (0.18)	25
total	51.09	2838 (43.28)	6558

Table 5. 3: Age and sex of participants included in MCP regression analyses.

5.2.2 Validation of MCP Polygenic Risk Score in Generation Scotland

5.2.2.1 Polygenic Risk Scoring in Generation Scotland

A MCP PRS value was calculated for unrelated Generation Scotland participants, who had not participated in UK Biobank, and whose genetic data passed quality

control checks. This was carried out using PRSice (Euesden et al., 2015) (see also 2.2.6)– the target phenotype was MCP derived in Generation Scotland as explained above. The best PRS was calculated by PRSice as one consisting of SNPs associated with MCP at a GWAS p < 0.4, and this PRS was standardised by PRS z value throughout. A weighted PRS is unit-less, and using a standardised score in this way aids interpretation of regression output i.e., for every 1-SD PRS increase, MCP phenotype value increased by X.

5.2.2.2 Regression Analyses

The relationship between MCP PRS and MCP phenotype in unrelated Generation Scotland participants with complete data on PRS, age, sex, MCP phenotype value and Multidimensional Scaling (MDS) components 1-4 (N = 6, 558) was then examined via linear regression (adjusting for age, sex and MDS components). MDS components are included to account for population stratification between UK Biobank and Generation Scotland (see Chapter 2.2.1.2 & Chapter 2.2.6). Four regression models in total were run (model formulae and sample size summarised Table 5.4).

Model	**Formula**	**N**
Initial	MCP~ Age + Sex + PRS + C1 + C2 + C3 + C4	6558
Sex Interaction	MCP~ Age + Sex*PRS + C1 + C2 + C3 + C4	6568
Sex-stratified: Male	MCP~ Age + PRS + C1 + C2 + C3 + C4	2838
Sex-stratified: Female	MCP~ Age + PRS + C1 + C2 + C3 + C4	3720

Table 5. 4: Summary of regression models with MCP as outcome.

C1-4 = MDS components 1-4, MCP = Multisite Chronic Pain trait value (0-6), PRS = Polygenic Risk Score

The relationship between MCP PRS and CPG phenotype in unrelated GS participants with complete data on PRS, age, sex, CPG phenotype value and Multidimensional Scaling (MDS) components 1-4 (N = 6, 080) was then examined via linear regression. (adjusting for age, sex and MDS components). Four regression models in total were run (model formulae and sample size summarised Table 5.5).

Model	Formula	N
Initial	CPG~ Age + Sex + PRS + C1 + C2 + C3 + C4	6080
Sex Interaction	CPG~ Age + Sex*PRS + C1 + C2 + C3 + C4	6080
Sex-stratified: Male	CPG~ Age + PRS + C1 + C2 + C3 + C4	2613
Sex-stratified: Female	CPG~ Age + PRS + C1 + C2 + C3 + C4	3467

Table 5. 5: Summary of regression models and sample sizes.

C1-4 = MDS components 1-4, CPG = Chronic Pain Grade, PRS = Polygenic Risk Score

5.2.3 Multisite Chronic Pain and Chronic Widespread Pain in UK Biobank: PRS Analysis

An MCP PRS was calculated for individuals who reported chronic pain all over the body in UK Biobank (excluded from previous GWAS analyses described in Chapter 5, N = 6, 815) and in 10, 000 randomly selected UKB participants who reported no chronic pain at any site or all over the body (also excluded from previous MCP GWAS analyses). The PRS was calculated using SNPs associated with MCP at $p < 0.01$, weighting by MCP GWAS effect size (GWAS beta) for each SNP. PRS score was standardised by standard deviation (SD) to give a z-PRS. A weighted PRS is unit-less, and using a standardised score in this way aids interpretation of regression output i.e., for every 1-SD PRS increase, CPG score increased by X.

5.2.3.1 Regression Analyses

The association between MCP PRS and CWP status was investigated using logistic regression, adjusting for age, sex, genotyping array and the first eight UK Biobank genetic PCs (to account for potential population stratification when comparing the two subsets of UK Biobank participants).

5.3 Results

5.3.1 MCP PRS Validation in Generation Scotland

PRS analyses undertaken to validate the MCP phenotype showed that MCP PRS was significantly ($p < 0.05$) associated with both MCP and CPG in Generation Scotland (Table 5.6: $p = 8 \times 10^{-32}$, Table 5.7: 2.87×10^{-23}, respectively). Every 1-SD increase in PRS value was associated with a 0.17-site increase in MCP phenotype value, and with a 0.13 increase in CPG phenotype value.

	Estimate	**SE**	**t**	**p**
(Intercept)	0.31	0.07	4.23	2.40×10^{-5}
PRS	0.17	0.01	11.80	8.08×10^{-32}
age	0.01	0.00	11.76	1.30×10^{-31}
sexM	-0.25	0.03	-8.60	9.81×10^{-18}
C1	-14.75	10.17	-1.45	0.15
C2	-25.06	11.34	-2.21	0.03
C3	-4.94	1.85	-2.67	0.01
C4	3.32	2.49	1.33	0.18

Table 5. 6: Results of the regression of MCP polygenic risk score on MCP in Generation Scotland, adjusted for age, sex and multidimensional scaling components 1-4..

Estimate = regression coefficient value, SE = standard error of regression coefficient value, t = t-statistic value, p = p value. Default level for factor variable 'Sex' is set to female (F). PRS refers to standardised (z) PRS.

	Estimate	SE	t	p
(Intercept)	0.32	0.07	4.82	1.45×10^{-6}
PRS	0.13	0.01	9.98	2.87×10^{-23}
age	0.01	0.00	9.48	3.57×10^{-21}
sexM	-0.21	0.03	-8.14	4.72×10^{-16}
C1	-5.89	9.13	-0.65	0.52
C2	-20.28	10.12	-2.00	0.05
C3	-3.07	1.64	-1.87	0.06
C4	2.08	2.21	0.94	0.35

Table 5. 7: Results of the regression of MCP polygenic risk score on chronic pain grade in Generation Scotland, adjusted for age, sex and multidimensional scaling components 1-4.

Estimate = regression coefficient value, SE = standard error of regression coefficient value, t = t-statistic value, p = p value. Default level for factor variable 'Sex' is set to female (F). PRS refers to standardised (z) PRS.

5.3.2 Sex-Specific Associations between PRS and MCP in Generation Scotland

There was a significant interaction between sex and PRS (Table 5.11: p = 0.002), and in sex-stratified regression analyses the association between PRS and MCP phenotype value was markedly higher in females than in males (Table 5.11: 0.21 vs 0.12, respectively). This significance survives Bonferroni correction of significance threshold (all p values for PRS terms in models << p_{bonf} = 0.0125).

	Estimate	**SE**	**t**	**p**
(Intercept)	0.31	0.07	4.19	2.82×10^{-5}
PRS	0.21	0.02	11.01	5.96×10^{-28}
sexM	-0.25	0.03	-8.67	5.36×10^{-18}
age	0.01	0.00	11.77	1.16×10^{-31}
C1	-14.43	10.17	-1.42	0.156
C2	-24.71	11.33	-2.18	0.029
C3	-4.90	1.85	-2.65	0.008
C4	3.15	2.49	1.26	0.206
PRS: sexM	-0.09	0.03	-3.12	0.002

Table 5. 8: Results for the regression of MCP polygenic risk score on MCP in Generation Scotland with inclusion of an interaction term (sex x PRS).

Estimate = regression coefficient value, SE = standard error of regression coefficient value, t = t-statistic value, p = p value. Default level for factor variable 'Sex' is set to female (F). PRS refers to standardised (z) PRS.

	Estimate	**SE**	**t**	**p**
(Intercept)	0.26	0.10	2.61	0.0092
PRS	0.12	0.02	5.87	5.00×10^{-9}
age	0.01	0.00	7.06	2.02×10^{-12}
C1	-30.96	14.08	-2.20	0.028
C2	-32.63	15.62	-2.09	0.037
C3	-2.36	2.51	-0.94	0.347
C4	1.02	3.45	0.30	0.767

Table 5. 9: Results for the regression of MCP polygenic risk score on MCP in Generation Scotland in males only.

Estimate = regression coefficient value, SE = standard error of regression coefficient value, t = t-statistic value, p = p value. PRS refers to standardised (z) PRS.

	Estimate	SE	t	p
(Intercept)	0.15	0.10	1.42	0.15
PRS	0.21	0.02	10.29	1.67×10^{-24}
age	0.01	0.00	9.50	3.79×10^{-21}
C1	-2.50	14.28	-0.18	0.86
C2	-18.92	15.98	-1.18	0.24
C3	-6.94	2.64	-2.63	0.01
C4	4.85	3.50	1.38	0.17

Table 5. 10: Results for the regression of MCP polygenic risk score on MCP in Generation Scotland in females only.

Estimate = regression coefficient value, SE = standard error of regression coefficient value, t = t-statistic value, p = p value. PRS refers to standardised (z) PRS.

Model	PRS	PRS*Sex	Significant
Initial	0.17	NA	Yes
Sex Interaction	0.21	-0.09	Yes, Yes
Sex-stratified: Male	0.12	NA	Yes
Sex-stratified: Female	0.21	NA	Yes

Table 5. 11 Summary of all four model key results.

*PRS = coefficient value for PRS term in regression models, PRS*Sex = coefficient value for sex-PRS interaction term (where applicable). Significant = p value for coefficient < 0.0125 (Bonferroni-corrected by number of regression models run in total (4)). PRS refers to standardised (z) PRS.*

5.3.3 MCP and Chronic Widespread Pain in UK Biobank

PRS analyses carried out to assess the relationship between CWP and MCP, as well as to partially validate the MCP phenotype, indicated that genetic risk for MCP (MCP PRS) was significantly associated with having chronic widespread pain, with every 1-standard deviation (SD) increased in PRS associated with a 63% increase in the odds of having chronic widespread pain (Table 5.12: OR = 1.63, p = 1.45×10^{-109}).

Term	Estimate	SE (Estimate)	Z	P	OR
(Intercept)	-61.418	2.763	-22.23	1.90×10^{-109}	2.12×10^{-27}
Age	0.016	0.002	7.45	9.25×10^{-14}	1.02
Sex	-0.488	0.035	-14.01	5.56×10^{-45}	0.61
PRS	0.488	0.022	22.24	1.45×10^{-109}	1.63

Table 5. 12: Results of the regression of MCP polygenic risk score on chronic widespread pain in UK Biobank.

Regression beta coefficient values (Estimate), odds ratios (OR), and P values. The reference level for 'sex' is set to female, PRS = z-polygenic risk score.

5.4 Discussion

5.4.1 Multisite Chronic Pain and Chronic Widespread Pain

Clinical syndromes involving chronic pain all over the body such as fibromyalgia, and chronic widespread pain itself, may represent the upper end of a spectrum of centralisation of pain, or the extreme of a chronic pain state (Phillips & Clauw, 2011). It has also been suggested that there are not "natural cut-off points" when it comes to chronic widespread pain versus localised chronic pain (Kamaleri et al., 2008). In support of this, MCP PRS was significantly associated with chronic widespread pain, indicating that chronic widespread pain be the upper end of a spectrum of increasingly widespread chronic pain, as previously suggested (Kamaleri et al., 2008; Phillips & Clauw, 2011), and that there are likely to be genetic variants that predispose both to MCP and to CWP.

5.4.2 Validation of MCP PRS in an Independent Cohort

Polygenic risk for MCP was significantly associated with both increasing MCP trait value and increasing chronic pain grade trait value in an independent cohort. This indicates that SNP associations discovered in UK Biobank are not limited to this specific cohort, and instead capture variation in chronic pain more generally. In addition, the significant association of MCP PRS with CPG is encouraging as CPG represents a validated chronic pain phenotype, again indicating that the

GWAS of MCP in UK Biobank was capturing genetic variation contributing generally to chronic pain.

5.4.3 Sex Differences in PRS Associations

An MCP PRS was associated with MCP in Generation Scotland in both men and women, but the size of the association was greater in women (β = 0.21, if a 95% CI is taken to be 0.21 +/- 0.04 this is significantly larger than the size of the male PRS-MCP association β = 0.12). There was also significant PRS-by-sex interaction. These results may indicate a genetic contribution to sex differences in prevalence of chronic pain, and were further examined in the published journal article associated with analyses in this chapter (Johnston et al., 2021).

Sex as a biological variable has a range of effects on how the genome functions and therefore on resulting phenotypic trait values (Bernabeu et al., 2021; Khramtsova et al., 2019; Rawlik et al., 2016). These effects can be mediated by sex differences in DNA methylation (Ge et al., 2017; Gilks et al., 2014; Hall et al., 2014; McCormick et al., 2017; Rahmioglu et al., 2015), sex differences in gene expression (Quinn & Cidlowski, 2016; X. Xu et al., 2012) and in eQTL effects (Kukurba et al., 2016; Yao et al., 2013), and varying levels and actions of hormones (Gomez-Santos et al., 2011; Kósa et al., 2009). Sex-specific pleiotropy, whereby genetic variants are associated with multiple traits but these relationships differ according to sex, can also contribute to sex differences in complex traits (Mitra et al., 2016; Rahmioglu et al., 2015), including chronic pain. Environmental factors strongly correlated with sex can also contribute to sex differences in complex trait phenotypic values.

Sex-differential gene expression has been observed in populations of sensory neurons (Mecklenburg et al., 2020) including within the dorsal root ganglion (K. Stephens et al., 2018) and tibial nerve (Ray et al., 2019).

In rodents, it has been found that different immune cells mediate mechanical pain hypersensitivity depending on sex, and that this relationship is affected by the action of testosterone (Mapplebeck et al., 2016; Sorge et al., 2015; Sorge & Totsch, 2017). In humans, improvement in chronic pain symptoms associated with some chronic pain conditions (particularly autoimmune conditions such as rheumatoid arthritis and MS) has been observed during pregnancy (Adams Waldorf & Nelson, 2008; Krause & Makol, 2016; Ray-Griffith et al., 2018; Varytė

et al., 2020). Again, this relationship is thought to be associated with varying sex hormone levels and their downstream effects on immune cell populations such as T helper cells (Craft et al., 2004; Pozzilli & Pugliatti, 2015). The relationship between pain perception and varying sex hormone levels during the human menstrual cycle, and between pain symptoms in chronic pain conditions and the menstrual cycle, is not fully understood (reviewed by Iacovides et al., 2015).

A range of diverse non-genetic (environmental) factors associated with female gender may also contribute to differences in chronic pain prevalence between sexes. Pain and chronic pain are more common in people reporting intimate partner/ domestic violence (Alhalal et al., 2018; Craner et al., 2020; Wuest et al., 2008), who also tend to be women (World Health Organization, 2021). Willingness to seek medical treatment, which is generally higher in women than men (Höhn et al., 2020; Thompson et al., 2016), may also contribute to higher prevalence estimates for chronic pain in women, as women are more likely to seek treatment for pain (K. D. S. Ferreira & Speciali, 2015)– however some evidence is mixed (Hunt et al., 2011) and this relationship may depend on pain type and be generally variable. Women may also be more likely to use maladaptive coping strategies for chronic pain (El-Shormilisy et al., 2015), potentially increasing any time period spent in pain and contributing to higher chronic pain prevalence. Adverse childhood experiences (ACEs) such as parental conflict, poverty, and psychological, physical or sexual abuse are also associated with higher rates of chronic pain in adulthood (Edwards et al., 2016; Groenewald et al., 2020), and some types of ACEs have been found to be more commonly experienced by women (Bellis et al., 2014; CDC, 2019), and in other cases greater variation and complexity in ACEs has been reported by women compared to men (Haahr-Pedersen et al., 2020).

Chapter 6: Using Genetics to Assess Causal Relationships in Pain and MDD

6.1 Introduction

Analyses undertaken in this chapter address the over-arching aim of this PhD project: to investigate causal relationships between MDD and chronic pain. Although the relationship between MDD and chronic pain in terms of comorbidity and genetic overlap is well documented, it is unclear as to whether a causal aspect to the relationship between MDD and chronic pain exists (see 1.3.2). In the absence of longitudinal data collection (that is specifically designed to control for many potential confounding factors in any underlying causal relationship between MDD and chronic pain, e.g., BMI, socioeconomic deprivation), it can be difficult to examine causality. One potential method to determine causality in exposures where randomised control trials (RCTs) are inappropriate for ethical, feasibility and financial reasons, and where longitudinal data is scarce or non-existent, is Mendelian Randomisation (MR) (reviewed by (Zheng, Baird, et al., 2017), see also 2.2.7). Using MR can mean that bias in causal estimates due to both reverse causality and confounding are in theory avoided.

The first section of this chapter (6.2.1), compares three different MR methods; Inverse-Variance Weighted (IVW), MR-Egger, and MR with Robust Adjusted Profile Score (MR-RAPS). Subsequent analyses focus solely on MR-RAPS (6.2.2) – this is because MR-RAPS is likely to be the most appropriate method when some degree of horizontal pleiotropy is almost certain to exist between the two traits. This is the case with depression and chronic pain, shown both by analyses in this thesis (Chapters 3, 4 and 5) showing significant genetic correlation between the two conditions, and in prior literature (see 1.1.5, 1.3.1). Differences in MR approaches and respective advantages and limitations are discussed in further detail in 2.2.7.

The next section of this chapter (6.2.3) examines the relationship between SNP genotype rs1186556 and chronic pain, and between SNP genotype and depression, in UK Biobank. Rs1186556 was found to be pleiotropic in cFDR analyses of MDD and CPG (Chapter 3), and mapped to *LRFN5*, a locus previously linked to

neuroinflammation and involved in neuronal development and synapse formation (3.3.5).

6.2 Methods

6.2.1 MR: Causal Relationship between Chronic Pain Grade and MDD

6.2.1.1 GWAS Datasets & Data Preparation

GWAS summary statistics from the 23andMe-Pfizer CPG GWAS were used. This GWAS was carried out on 23, 301 unrelated individuals of European descent (see 2.3.2.2 and 2.3.3.2.1). Harmonisation was carried out (see 2.2.7.5), with an r^2 threshold of 0.01 chosen throughout. LD between instruments was checked using PLINK command --r2 --ld-window-r2 0.01, checking against UK Biobank genotyping data as a reference. This command delivers a list of pairwise correlation values (measured in r^2) between instrument SNPs, for any value of $r^2 > 0.01$. This stringent threshold was used to make MR-RAPS (Zhao et al., 2020) valid. Where LD was found ($r^2 > 0.01$), the SNP with the lowest GWAS p value amongst the inter-correlated group was reserved. These reserved SNPs plus the SNPs that were not found to be in LD comprised the list of instruments for MR. SNPs associated with the exposure (here, CPG) at $p < 10^{-5}$ and satisfying the other criteria were taken forward into the MR analysis. Note that this p value threshold is one order of magnitude more conservative than recommended for MR-RAPS (where recommended threshold is $p < 10^{-4}$), to account for the fact that CPG-associated SNP effect sizes were calculated in a sample that was not independent of the source GWAS for the exposure trait (CPG). After the above steps were completed, 25 SNPs were taken forward to assess the causal effect of CPG on MDD.

MR-Egger, IVW and MR-RAPS were carried out, and Q and I^2_{GX} values were calculated.

A large GWAS meta-analysis was carried out by Wray et al (Wray et al., 2018), and data from UK Biobank and 23andMe participants were removed from those results, to give GWAS summary statistics for a cohort consisting of 43, 028 and 87, 522 MDD cases and controls, respectively. Harmonisation and pruning steps were carried out as described in the preceding section, leaving 44 independent

SNPs associated with MDD at $p < 10^{-5}$ to be used as instruments to assess the causal effect of MDD on CPG.

6.2.1.2 Mendelian Randomisation Analysis

MR-Egger (2.2.7.3), IVW (2.2.7.2) and MR-RAPS (2.2.7.4) were carried out using summary statistic datasets prepared as described in previous sections, first to estimate the causal effect of CPG on MDD, then to estimate the causal effect of MDD on CPG. Q and I_{GX}^2 values were calculated to assess pleiotropy in SNP instruments.

IVW MR involves fixed-effect meta-analysis of the Wald ratio causal effect estimates for each SNP. MR-Egger analysis provides a causal estimate through treating each instrument (SNP) as a study in a meta-analysis, where the overall effect estimate (causal effect estimate) is given by the slope of the MR-Egger regression – in contrast to IVW analysis the intercept of this regression is not constrained to pass through the origin. Directional horizontal pleiotropy was also tested for through testing whether the intercept of this regression was significantly different to zero. Both IVW and MR-Egger MR analyses were carried out using code written in R (versions 3.5.3 – 3.6.0), partially based on code templates distributed as part of the 2018 Mendelian Randomisation workshop (University of Bristol).

MR-RAPS was carried out using the package 'mr-raps' in R (version 3.6.0) (Zhao et al., 2020).

6.2.2 MR: Causal relationships between Multisite Chronic Pain and MDD

6.2.2.1 GWAS Datasets & Data Preparation

Summary statistics for the GWAS carried out in Chapter 4 in UKB on MCP were used to derive instruments for MCP, a second chronic pain phenotype distinct from CPG. After harmonisation and pruning steps were completed as above, this resulted in 200 independent SNP instruments associated with MCP at $p < 10^{-5}$ for use in investigating potential causal effects of MCP on MDD.

Wray et al summary statistics were used and pruning and harmonisation steps followed as previously outlined. This resulted in 99 independent SNPs associated with MDD at $p < 10^{-5}$ to be used as instruments in estimating the causal effect of

MDD on MCP. Note that the number of instruments for MDD varies between the MCP<->MDD and CPG<->MDD analyses as pruning and harmonisation is carried out with respect to the outcome variable GWAS data in each case.

6.2.2.2 Mendelian Randomisation Analysis

MR-RAPS (2.2.7.4) analyses were carried out, first to estimate the causal effect of MCP on MDD, then to estimate the causal effect of MDD on MCP.

6.3 Results

6.3.1 Causal Relationships between Chronic Pain Grade and MDD

MR analyses investigating a potentially causal effect of CPG on MDD found no significant causal effect after FDR correction for multiple testing (Table 6.1, p value = 0.08). With respect to MR-Egger, where the causal effect was non-significant even prior to FDR correction, this may be due to causal dilution caused by violation of the NOME (No Measurement Error) assumption, as indicated by the I^2_{GX} value of < 0.9.

Method	β	SE	β/ SE	p	p (FDR)	Q	p (Q)	I^2	τ^2	I^2_{GX}	model type	loss function
IVW	0.068	0.03	2.218	0.036	0.054	21.75	0.594	0	0	NA	NA	NA
MR-Egger	0.0805	0.044	1.83	0.08	0.08	NA	NA	NA	NA	0.84	NA	NA
MR-Egger (Intercept)	-0.0015	0.004	-0.395	0.7	NA	NA	NA	NA	NA	NA	NA	NA
MR-RAPS	0.07	0.034	2.0588235	0.035	0.054	NA	NA	NA	NA	NA	simple	L2

Table 6. 1: MR results with chronic pain grade as the exposure and MDD as the outcome across all three methods.

β = causal estimate, SE = standard error of causal estimate, β/ SE = z value, p = p value, p (FDR) = FDR-adjusted p value, Q = Cochran's Q value, p (Q) = p value for Cochran's Q, I^2 = heterogeneity estimate (IVW), τ^2 = among-study variance (heterogeneity estimate, IVW), I^2_{gx} = heterogeneity indicator (MR-Egger).

No significant directional or balanced horizontal pleiotropy was detected in these analyses, as indicated by the non-significant intercept value in MR-Egger

(Table 6.1 p = 0.7), and the non-significant Q value for IVW MR (Table 6.1 p = 0.59), respectively. This is corroborated in MR-RAPS analysis, as no significant over-dispersion (indicating widespread horizontal pleiotropy amongst instruments) or idiosyncratic pleiotropy (horizontal pleiotropy in a small subset of instruments) was detected – simple, non-robust (L2 loss function) regression was best-fitting.

In MR analyses investigating potential causal effect of MDD on CPG, again no significant causal effect was found across any of the four MR methods. IVW results indicated no horizontal pleiotropy among instruments (Table 6.2 p (Q) > 0.05), and MR-Egger intercept results suggest no direction pleiotropy (Table 6.2 p > 0.05), but MR-Egger results suggest significant violation of the NOME assumption (Table 6.2 I^2_{GX} <<0.9). The fact that the best-fitting MR-RAPS model was one with a Huber loss function indicates idiosyncratic pleiotropy among instruments – a small subset of instruments for MDD were horizontally pleiotropic.

Method	β	SE	β/ SE	p	p (FDR)	Q	p (Q)	I^2	τ^2	I^2_{GX}	model type	loss function
IVW	-0.022	0.03	-0.74	0.46	0.69	45.9	0.35	6.3	0.003	NA	NA	NA
MR-Egger	0.078	0.063	1.24	0.22	0.66	NA	NA	NA	NA	0.44	NA	NA
MR-Egger (Intercept)	-0.007	0.004	-1.86	0.07	NA	NA	NA	NA	NA	NA	NA	NA
MR-RAPS	-0.012	0.032	-0.375	0.712	0.712	NA	NA	NA	NA	NA	simple, robust	Huber

Table 6. 2: MR results with MDD as the exposure and chronic pain grade as the outcome across all three methods.

β = causal estimate, SE = standard error of causal estimate, β/ SE = z value, p = p value, p (FDR) = FDR-adjusted p value, Q = Cochran's Q value, p (Q) = p value for Cochran's Q, I^2 = heterogeneity estimate (IVW), τ^2 = among-study variance (heterogeneity estimate, IVW), I^2_{gx} = heterogeneity indicator (MR-Egger).

6.3.2 Causal Relationships between Multisite Chronic Pain and MDD

MR-RAPS analysis was performed to investigate causal relationships between MDD and MCP, first with MDD as the exposure and MCP as the outcome. QQ plots, leave-one out versus t-value plots and Anderson-Darling/ Shapiro-Wilk test p

values indicated that models without dispersion were best-fitting (Table 6.3 rows 1-3, pAD > 0.05, pSW > 0.05). Effects of outliers (idiosyncratic pleiotropy) are not ameliorated in models with dispersion despite robust regression (Figure 6.1: D, E, F right-hand panels). The best-fitting model with greatest amelioration of pleiotropy was one without over-dispersion and with a Tukey loss function (Table 6.3: row 3, Figure 6.1: C).

overdispersion	Loss function	β	SE (β)	p (β)	p (AD)	p (SW)	τ	p (τ)	C.F
FALSE	L2	0.0117	0.0052	0.0241	0.9375	5.34E-01	NA	NA	A
FALSE	Huber	0.0153	0.0054	0.0042	0.9285	5.23E-01	NA	NA	B
FALSE	**Tukey**	**0.0185**	**0.0054**	**0.0006**	**0.9230**	**5.18E-01**	**NA**	**NA**	C
TRUE	L2	-0.0096	0.0132	0.4671	0.0080	1.76E-03	1.61E-04	0.0470	D
TRUE	Huber	-0.0056	0.0126	0.6556	0.0087	2.11E-03	1.30E-04	0.0677	E
TRUE	Tukey	-0.0065	0.0137	0.6330	0.0055	9.03E-04	1.67E-04	0.0627	F

Table 6. 3: MR results for MR-RAPS analysis with MDD as the exposure and MCP as the outcome.

β refers to the causal effect, SE (β) and p (β) to the standard error and p value of β, p (AD) to the Anderson-Darling test of normality p value, p (SW) to the Shapiro-Wilk test of normality p value, τ to the over-dispersion statistic size and p (τ) to the p value. C.F = corresponding QQ plot panel for the model. p (τ) was calculated from the tau estimate and its standard error (Altman & Bland, 2011). The row of the table corresponding to the regression model found to be best-fitting is in bold.

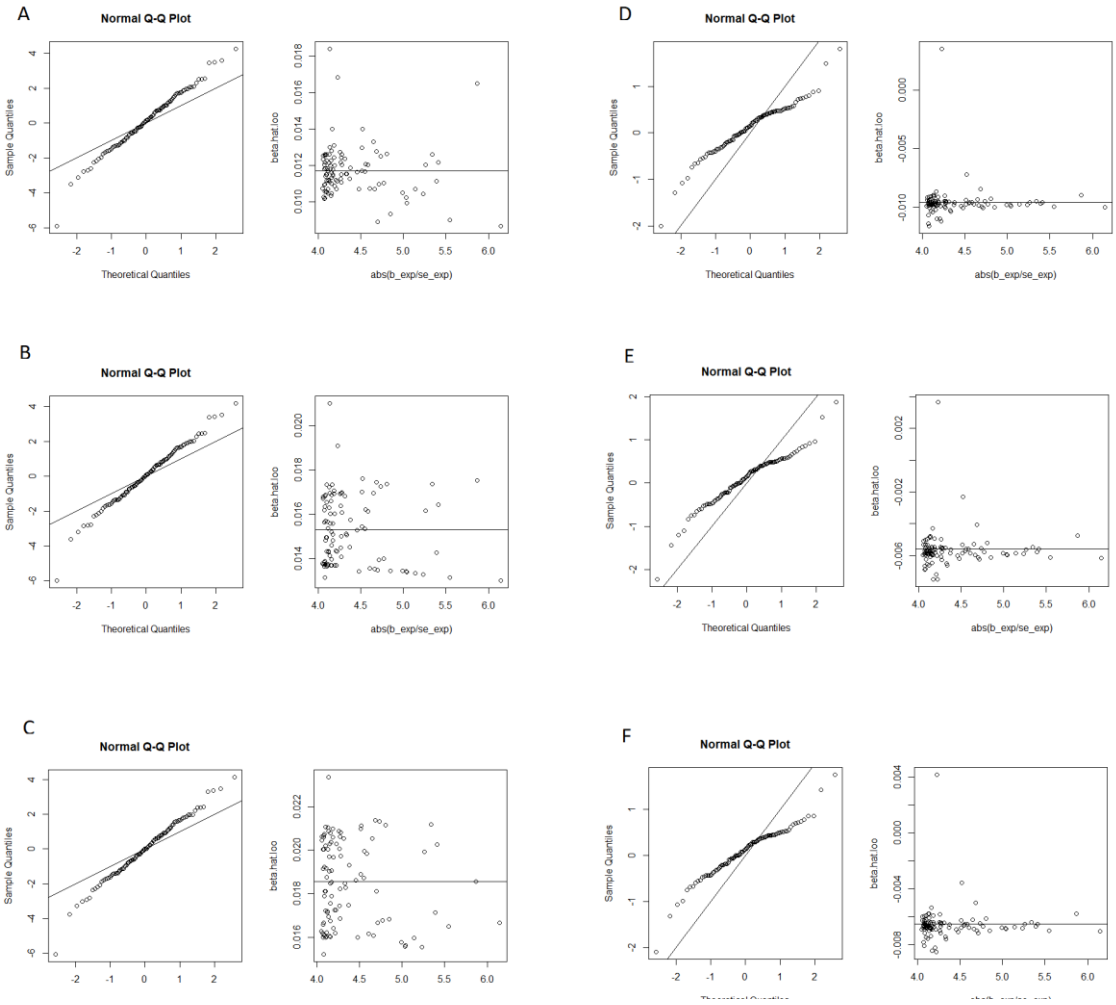

Figure 6. 1. Diagnostic plots of MR-RAPS analysis with MDD as exposure

Quantile-quantile plots (left-hand panels, 'normal Q-Q plot') and leave-one-out β versus t-value plots (right-hand panels) for each of the six models fitted during MR-RAPS analyses (A-F) are shown. Beta.hat.loo = leave-one-out β value estimate, abs(b_exp/se_exp) = absolute β value divided by standard error (t value). Each point represents a SNP instrument.

These results indicate idiosyncratic pleiotropy (pleiotropy in some but not all instruments). The causal effect of MDD on MCP is positive and significant at beta = 0.019 and p = 0.0006, but the diagnostic plots show a 'swapping' of sign for the causal estimate (Figure 6.1), suggesting that there is not a truly significant causal effect of MDD on MCP.

MR-RAPS analyses were then carried out with MCP as the exposure and MDD as the outcome. Models with dispersion are a better fit than those without (Figure 6.2: A, B, C vs D, E, F, Table 6.4: rows 4-6, pAD > 0.05, pSW > 0.05, pτ << 0.05). This indicates that effectively all instruments are horizontally pleiotropic (affecting MDD through pathways other than via MCP). The causal effect of MCP on MDD is positive and significant at beta = 0.16 and p = 0.047.

overdispersion	Loss function	β	SE (β)	p (β)	p (AD)	p (SW)	τ	p (τ)	C.F
FALSE	L2	0.1714	0.0605	0.0046	0.4256	0.0853	NA	NA	A
FALSE	Huber	0.1815	0.0621	0.0034	0.4247	0.0835	NA	NA	B
FALSE	Tukey	0.2097	0.0621	0.0007	0.4221	0.0784	NA	NA	C
TRUE	L2	0.1201	0.0790	0.1286	0.8374	0.2853	9.81E-05	2.43E-03	D
TRUE	Huber	0.1446	0.0801	0.0712	0.8289	0.2724	9.18E-05	5.13E-03	E
TRUE	**Tukey**	**0.1578**	**0.0795**	**0.0471**	**0.8236**	**0.2641**	**8.77E-05**	**7.09E-03**	F

Table 6. 4: MR results for MR-RAPS analysis with MCP as the exposure and MDD as the outcome.

β refers to the causal effect, SE (β) and p (β) to the standard error and p value of β, p (AD) to the Anderson-Darling test of normality p value, p (SW) to the Shapiro-Wilk test of normality p value, τ to the over-dispersion statistic size and p (τ) to the p value. p (τ) was calculated from the τ estimate and its standard error (Altman & Bland, 2011)The row of the table corresponding to the regression model found to be of best fit is in bold.

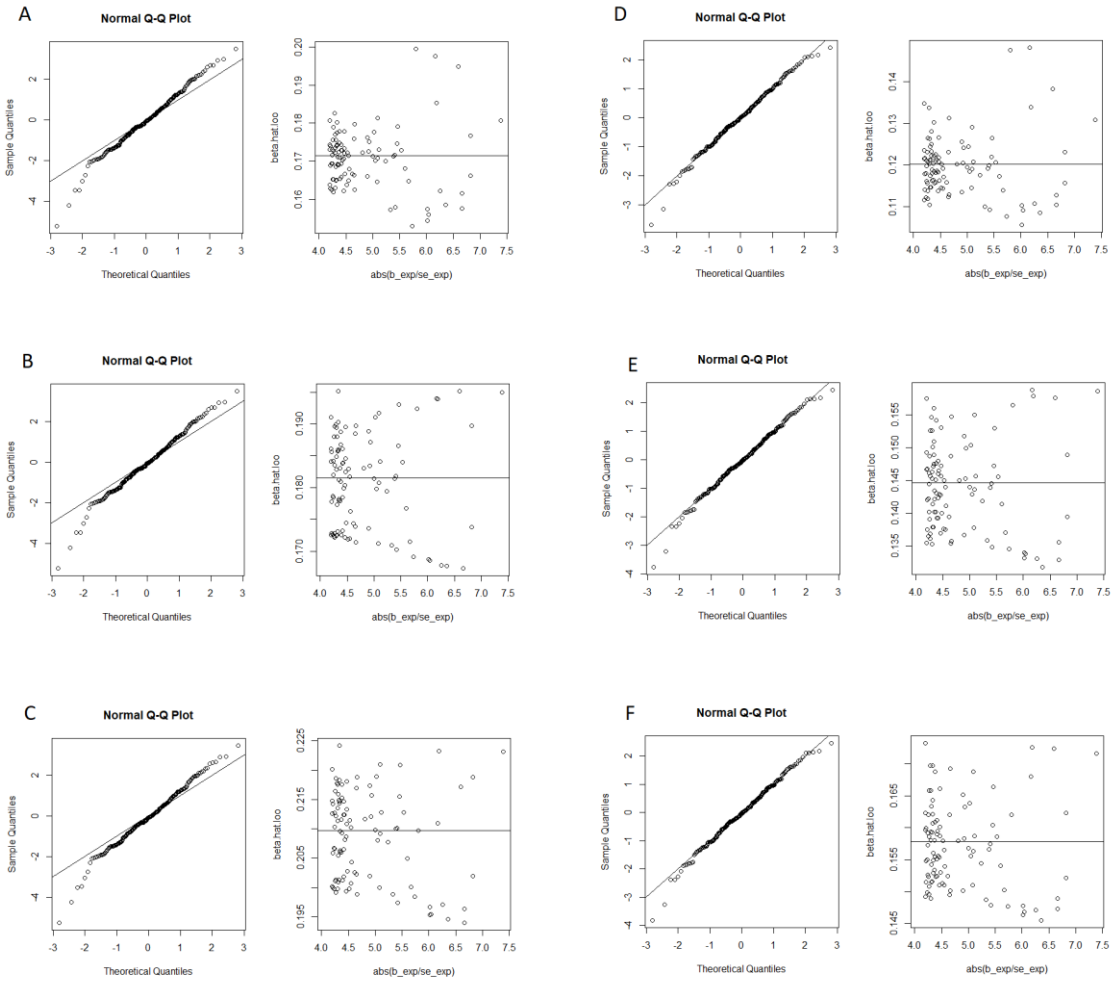

Figure 6. 2: Diagnostic plots of MR-RAPS analysis with MCP as exposure and MDD as the outcome.

Quantile-quantile plots (left-hand panels, 'normal Q-Q plot') and leave-one-out β versus t-value plots (right-hand panels) for each of the six models fitted during MR-RAPS analyses (A-F) are shown. Beta.hat.loo = leave-one-out β value estimate, abs(b_exp/se_exp) = absolute β value divided by standard error (t value). Each point represents a SNP instrument.

Overall, the results of this analysis suggest a causal effect of MCP on MDD.

6.4 Discussion

6.4.1 Causal relationship between MCP and MDD

Several MR analyses indicated no causal effect of CPG on MDD or vice versa, but MR-RAPS suggested MCP has a causal effect on MDD.

The finding of no causal effect of CPG on MDD could be due to the comparatively smaller sample size of the 23andMe-Pfizer CPG GWAS, an order of magnitude smaller than the MCP GWAS, reducing power. It is also possible that potential bias introduced into the CPG GWAS output through adjustment for traits which

are likely to be genetically correlated with CPG (BMI and manual labour) may affect causal analyses (see 4.4.1.6).

Differences between CPG and MCP as traits may also contribute to potential differences in observed causal estimates. Although the environment has a large influence on MCP (heritability, i.e., additive genetic component of trait variation, is only ~10%), environmental components of chronic pain experience are perhaps more explicitly measured with CPG. For example, interference in daily activities (e.g., item 4, how many days in the last 6 months have you been kept from your usual activities because of this pain?), is assessed as part of CPG. The individual's interaction with their environment and the relationship between this interaction and any chronic pain is, in contrast, not directly measured or used to calculate the number of chronic pain sites on the body (i.e., MCP).

MR-RAPS results for the causal effect of MCP on MDD also showed that effectively all instruments were horizontally pleiotropic – their effect on MDD is not exclusively through MCP and is instead dispersed or diluted out through other MCP-correlated traits. This emphasises that despite evidence of a significant causal role of chronic pain in MDD, many interrelated factors are involved in the relationship between MDD and chronic pain, and the relationship between the two phenotypes is complex. However, lack of significant causal effect of MDD on MCP can be interpreted as evidence that despite high comorbidity and genetic overlap, MDD does not directly cause chronic pain.

Heterogeneity in MDD may mean that, although unexplained physical symptoms including chronic pain and headaches are experienced by some people with an MDD diagnosis, any potential causal effect of some specific subtypes of depression on MCP is obscured as the MDD GWAS included MDD cases overall, many of whom likely experience no chronic pain or unexplained physical symptoms. It is perhaps then even more significant that significant (if small) causal effect of MCP on MDD was found, given that MCP is also a heterogeneous trait. The translational impact of the results of these analyses is, however, minimal. MR studies of modifiable exposures or medications can have a direct translational impact as the causal effect of discrete, actionable changes (e.g. smoking cessation or taking a medication) can be ascertained – for example the impact of PSCK9 inhibitors on risk for type 2 diabetes was investigated in this

manner and has potential direct clinical implications (Schmidt et al., 2017). In contrast, in this section the causal impact of chronic pain (a heterogeneous, complex-trait exposure that is perhaps not modifiable in the way lifestyle factors such as cigarette smoking are) on MDD (and vice versa) is estimated.

Chapter 7: General Discussion

This chapter is designed to bring together an overview of results from each of the previous chapters, along with discussion of strengths, limitations, and future directions of work carried out in this thesis. The over-arching aim of this thesis was to investigate causal relationships between major depression and chronic pain. Three main research objectives were addressed to achieve this aim; to uncover common genetic variation associated with chronic pain phenotypes, to investigate genetic correlation and pleiotropy between MDD and chronic pain, and finally to test for clinical heterogeneity between MDD and chronic pain.

Novel common genetic variation associated with chronic pain phenotypes Chronic Pain Grade (CPG) was discovered first through leveraging genetic overlap with MDD (Chapter 3), then through genome wide association analysis with a novel chronic pain phenotype (Multisite Chronic Pain, MCP) (Chapter 4). Eleven novel SNPs were found to be associated with CPG, and 76 with MCP. MCP was found to be a moderately heritable, polygenic trait, with associated genes suggesting a strong central nervous system component to MCP development.

Genetic correlation and pleiotropy was examined through analyses in Chapter 3 which identified pleiotropic loci of interest (associated with effects in both MDD and chronic pain), and through estimating genetic correlation between MCP and MDD (Chapter 4). Analyses in Chapter 3 also tested for clinical heterogeneity in MDD with respect to CPG and vice versa, using GWAS summary statistics for each trait. *LRFN5,* a gene involved in cell-cell adhesion in the CNS and previously implicated in neuroinflammation and major depression, was found to be pleiotropic with respect to MDD and CPG. Significant genetic correlation was observed between MCP and a range of traits and disorders, most notably mood and psychiatric traits, with low-moderate or non-significant genetic overlap observed between MCP and conditions involving significant chronic pain (e.g., IBDs and rheumatoid arthritis). No significant evidence for clinical heterogeneity in MDD or in CPG was found in UK Biobank.

Analyses undertaken in Chapter 5 examined whether genetic risk for MCP was associated with CPG, and with chronic widespread pain, where it was found that genetic risk for MCP was significantly associated with both CPG and chronic widespread pain, and with MCP in an independent cohort. Finally, GWAS

summary statistics were taken forward for causal analyses in Chapter 6, where a significant causal effect of MCP on MDD was observed.

7.1 History of Pain Theories

Understanding pain and chronic pain has been a key philosophical as well as scientific question throughout human history. Ancient theories on pain include those formulated by Plato, such as intensity theory (Zeyl & Sattler, 2019), i.e. that pain is an "emotion" and occurs when a stimulus is sufficiently intense and long-lasting. In the 1600s Descartes posited Cartesian dualism, suggesting pain is a result of physical injury or of psychological (emotional) injury, but never both and neither can influence the other. However, he described the mind ('soul') and body as intertwined and connected the concept of pain to a soul, with the soul of pain in the pineal gland. Descartes also described 'fibres' that could transmit pain messages to the brain (Benini & Deleo, 1999; Moayedi & Davis, 2013). Later, in the 1800s Bell outlined specificity theory, assigning types of sensations to particular pathways, and also described brain as a complex structure with different components (Bell & Shaw, 1868;Bell, 1811)

Also during this period Muller suggested that different sensations, including pain, are due to activity at different receptors (reviewed by Perl, 2007). The work of von Frey also assigned different sensations to separate and specific receptors, with four separate modalities of somatosensory system activity (pain, cold, heat and touch) – different small areas of skin were linked to different types of sensation, and Von Frey observed relationships between types of neural structure (histologically defined) and these small skin areas (specificity theory) (reviewed by Moayedi & Davis, 2013; Rey, 1995). However, contrasting theories around the same time such as those of Erb argued that pain was a result of a stimulus being sufficiently intense to evoke painful sensation through activity at receptors that usually were involved in other, non-painful sensation (i.e. the opposite of specificity theory/ intensity theory as suggested by Plato, and in opposition to Muller) (reviewed by Moayedi & Davis, 2013; Perl, 2007). In the 1900s 'pattern theory' attributed to Nafe stated that there are no separate or specific receptors, and instead different types of sensations lead to different sequences or patterns of signals being transmitted to the brain.

The first theory to incorporate endogenous (non-peripheral) modulation of pain signals was that of Wall & Melzack in 1965: gate control theory. Here, stimuli

have to pass through 3 locations in the spinal cord before reaching the brain, with the substantia gelatinosa in the dorsal horn acting as a gate that modulates these signals – if a signal reaches significant intensity, this is transmitted to the brain (the gate 'opens'), with additional control mechanism in cortical regions of the brain. This theory incorporated psychological and cognitive aspects in pain in addition to explaining clinical observations of non-noxious activity (e.g. rubbing an injured area) attenuating the level of pain felt, and the wide variation seen in the relationships between stimuli and resultant pain (Katz & Rosenbloom, 2015; Melzack & Wall, 1965; Mendell, 2014).

Later, in 1990, Melzack also described the idea of a neuromatrix, whereby the CNS is responsible for producing painful sensation, not the periphery (though the periphery can influence this sensation, in line with previous gate-control ideas) (Melzack, 1990, 2001). The 'neuromatrix' consists of multiple CNS locations that work in concert to produce a "neurosignature", leading to pain. Peripheral information can influence this neurosignature but cannot lead to production of a neurosignature in isolation: this theory recognised influence of cognitive and emotional (but not social) factors in pain, and the fact that pain (particularly chronic pain syndromes) can be experienced in the absence of, or related only to minimal, direct sensory input. In this context mechanistic descriptors of pain (nociceptive, neuropathic, nociplastic, see also 1.1.1) may therefore describe ways in which the neurosignature is modulated, rather than discrete and causal descriptions for pain experiences.

The biopsychosocial model of disease (Bevers et al., 2016) as applied to chronic pain (Fillingim, 2015, 2017, see also 1.1.4), can be thought of as uniting these three broad categories of theories on pain (specificity theory, intensity theory, and ideas of pain as a kind of emotion), and additionally emphasises multifactorial contributions, including environmental (e.g. social) factor contributions, to chronic pain development.

Results of analyses in this thesis are in line with these theories – derivation of a broad chronic pain trait (MCP) where sensory input associated with inflammation or injury is not directly measured as a component of the phenotype ties in with ideas of the neuromatrix (with a neurosignature and chronic pain modulated by but not wholly produced by nociceptive or other peripheral input). Furthermore, several disease traits which commonly involve significant chronic pain were not

genetically correlated with MCP, including Crohn's disease, lupus and rheumatoid arthritis (4.4.1.2). This suggests GWAS in these traits does not capture underlying mechanisms of chronic pain development, but rather genetic variation associated with more specific disease characteristics (which have a variable relationship with pain experienced). However, it should be noted that comparatively smaller sample sizes of some GWAS datasets (e.g., 5, 956 Crohn's disease cases and 21, 770 controls) (Liu et al., 2015) may mean power to detect significant genetic correlation is reduced. This smaller sample size is directly related in many cases to lower prevalence of many chronic pain conditions – again taking Crohn's disease as an example, prevalence in cases per 100,000 people in Europe is estimated to be 1.5-213 (Burisch et al., 2013), whereas for MDD (where highly significant genetic correlation with MCP was observed and the GWAS sample size was much larger) this figure is ~12,000 (see 1.2.2). It may also be the case that genetic correlation on a more local level is present but not detectable between MCP and chronic pain conditions (see 7.6.3). Additionally, phenotyping chronic pain in addition to chronic pain-associated condition status may be more informative in finding genetic predictors of chronic pain, as across a range of conditions disease severity, disease activity and tissue damage are not necessarily reflected in severity of pain experienced.

Several MCP-associated genes (particularly *DCC*) are involved in neuronal migration (see 4.4.3), outlined as a key component of the formation and alteration of the neuromatrix across the life course (Melzack, 1990). Pathways enriched for MCP-associated genes (see 4.4.3) included *DCC*-mediated attractive signalling, involved in cell motility and migration including in neural cell haptotaxis and synapse formation, and PLC-ß-mediated signalling, also involved in synapse formation.

Results of GWAS analyses also indicate that chronic pain follows the biopsychosocial model of disease: ~10% of trait variation is attributed to common genetic (SNP) variation, with therefore ~90% attributed potentially to non-genetic (environmental) factors in addition to other types of genetic variation not assayed in GWAS. Genetic correlations between MCP and other traits of interest indicated significant and large overlap in psychiatric traits, such as MDD and depressive symptoms, emphasising an affective component of chronic pain in addition to biological and social components. Considering chronic

pain as a disease or trait in this way is also in line with recent definitions of Chronic Primary Pain for the ICD-11 (see 1.1.1).

7.2 Evolutionary Perspectives of Pain

Acute pain is considered to be adaptive, as it warns against damage and danger (De C Williams, 2019) – in humans one demonstration of this is severe injury and even limb loss in individuals with congenital insensitivity to pain (see also 1.3.4). If chronic pain is related to a high enough degree to acute pain, this may explain why chronic pain as a trait is maintained in the population – as an unavoidable side effect of acute pain where any deleterious effect is outweighed by the adaptive benefit of acute pain. However, the relationship between acute and chronic pain is less straightforward, and most likely not strong enough to explain maintenance of chronic pain. This is demonstrated by the fact that severe and debilitating chronic pain can result from initial injuries where the experience of acute pain is minimal e.g. CRPS (C. Chang et al., 2019; F. & Chandan G., 2014), and that chronic pain is thought to be often absent in wild non-human animals (though this may be due to the fact this phenomenon is under-studied) (De C Williams, 2019).

Other traits apart from acute pain may be both highly correlated with chronic pain and highly adaptive, driving the maintenance of chronic pain as a trait in the population in the face of natural selection. For example, neural plasticity is thought to be involved in chronic pain development, but is also an adaptive (or rather essential) trait in general brain development and function (Mateos-Aparicio & Rodríguez-Moreno, 2019)– the selective disadvantage conferred by chronic pain would not outweigh the adaptive role of neural plasticity. It could also be the case that chronic pain itself is adaptive and so maintained through positive selection – studies in laboratory animals have shown advantages for predator avoidance associated with nociceptive sensitisation after injury (Crook et al., 2014; Lister et al., 2020).

At the genetic rather than trait level, many hundreds of genetic variants contribute a small amount to variation in complex traits (see 1.3.3.1), and pleiotropy is widespread – results of analyses in this thesis indicate that chronic pain (MCP) is a highly polygenic complex trait (4.3.2, 4.3.3, 4.4.2), and pleiotropic variants were found specifically to contribute to both MDD and chronic pain (3.4.2). The vast majority of MCP-associated variants are therefore

likely to be selectively neutral (have a small effect size, and circulate in the population as common SNPs), or exert effects (however small) on traits that are adaptive, and so are maintained in the population.

Another important point to note is that traits persist in a population not necessarily because they are adaptive (or strongly associated with an adaptive trait) at all: they are just not sufficiently deleterious to be removed by purifying selection. It is also important to note that maladaptive in a clinical sense is not the same as maladaptive in evolutionary terms. Traits may be selectively neutral, particularly traits like chronic pain where trait onset is often established after the reproductive period (Macfarlane, 2016), does not have severe enough fitness effects (in terms of negative impact on reproduction) to be subject to negative selection, or both. Additionally, natural selection, particularly if not extremely strong purifying selection, acts over long, multigenerational timespans. It may be the case that chronic pain played an adaptive role in the context of more ancient human environments (Walters, 2019), and this has recently stopped being the case – again lack of evidence of chronic pain in wild animals (but presence in domesticated animals) could be taken to support this theory (De C Williams, 2019).

7.3 Multisite Chronic Pain in UK Biobank

7.3.1 Comparing MCP and Other Chronic Pain Phenotypes

Studying chronic pain as a disease entity or phenotype in its own right may present a more tractable way to uncover genetic factors involved in development of and vulnerability to chronic pain. An attempt to do this was made through derivation of the MCP phenotype in UK Biobank.

MCP was found to be a moderately heritable, polygenic complex trait, similar to MDD. Genes associated supported a view of chronic pain as a disorder with a significant central nervous system component, implicating neuroinflammation and neuronal plasticity. Interestingly, so-called ‘classic’ chronic pain genes such as *COMT* was not found to be associated with MCP. Similarly to early candidate gene studies of MDD, this may be due to *COMT’s* association with individual variation in pain perception being specific to those cohorts, or an artefact of reduced power (associated with small sample sizes and with methodological issues in candidate gene analysis). It may be the case that more general genetic

factors influencing chronic pain susceptibility or development are likely to be found in GWAS presented in this thesis, and that GWASs of chronic pain conditions are more likely to show genetic variation associated with more specific disease or condition-associated processes.

The relationship between MCP and other more 'general' i.e., non-disease or chronic pain condition-associated chronic pain traits was also explored. Chronic widespread pain could be considered the 'upper end' or more extreme presentation chronic pain (Kamaleri et al., 2008; Phillips & Clauw, 2011), and this was supported by genetic correlation results between MCP and 'chronic pain all over the body', $r_g = 0.83$. MCP-PRS was also significantly associated with MCP in an independent cohort (Generation Scotland), and with having CWP in UK Biobank. In addition to demonstrating genetic overlap with these different chronic pain phenotypes, these results further support view of MCP as a useful chronic pain phenotype and suggest associations between this phenotype and common genetic variation are not specific to UK Biobank. CPG and MCP were also found to be phenotypically correlated in an independent cohort, further legitimising MCP as a chronic pain phenotype. Finally, extremely high genetic correlation was seen between a chronic pain phenotype derived in a data-driven manner by Tsepilov et al in UK Biobank, and MCP (Appendix 3). This suggests the assumptions made in defining MCP as a trait (namely equivalence between genetic predictors of musculoskeletal and non-musculoskeletal chronic pain (Tsepilov et al., 2020)) are in fact acceptable, and again indicating MCP represents a valid broad chronic pain phenotype derived from a basic pain questionnaire.

Key differences between MCP and CPG include that a larger sample size was possible for the MCP GWAS, and there was no adjustment for correlated traits. This increased power to find trait-associated genetic variants. Environmental aspects of chronic pain experience may be more explicitly captured in CPG compared to in measuring number of chronic pain sites (as is done with MCP) – this aspect may explain differences in MCP vs CPG such as the unexpected negative genetic correlation between the two traits. This could also have an impact on causal estimates between CPG and MDD.

7.3.2 Broad MDD Phenotyping Parallels

MCP can be considered a broad chronic pain phenotype, as more detailed information on pain (such as impact on daily functioning or on mood) was not collected at baseline by the original pain questionnaire, and GWAS of MCP does not consider or adjust for participants having chronic pain conditions, or for related traits of interest such as BMI.

Similar 'broad' phenotyping has been shown to be useful for investigating the genetics of depression and MDD (1.3.5), meaning questionnaire ascertained probable MDD or other depression phenotypes share significant enough genetic overlap with 'narrow' (detailed or specific clinician diagnosed) MDD phenotypes to be informative. However, fully dissecting the heterogeneity of MDD, and of chronic pain, will require both broad and narrow phenotyping approaches (which can be non-mutually exclusive, as suggested for the field of MDD genetics research) (Cai et al., 2020). There is opportunity to carry out both approaches on a large scale in UK Biobank with the release of the new chronic pain questionnaire data, a more in-depth follow-up assessment of pain experience (see 7.5.2).

MDD and chronic pain share overlapping symptom profiles, can respond to the same pharmacological and psychological treatments, and are significantly genetically correlated. Although this means the two conditions are more likely to be misclassified (misdiagnosed) as one another, BUHMBOX analyses indicated that the two conditions are distinct with respect to one another, and the relationship between them represents true, biological pleiotropy. In addition to this finding being of general interest, this also aids in interpretation of causal analyses results – if misclassification is the underlying reason for genetic correlation, the question being asked in MR inadvertently becomes "does X cause X" to some degree, rather than "does X cause Y".

7.5 Causal Effect of Chronic Pain on MDD

Results indicate a significant causal effect of chronic pain on MDD, but not the reverse. Attempts at triangulation were made via use of different Mendelian randomisation approaches. MR-RAPS indicated that effectively all instruments

were horizontally pleiotropic (associated with MDD through pathways other than via MCP). The size of the causal effect is relatively small (beta = 0.16) – again suggesting involvement of non-additive-genetic factors and gene-by-environment interactions may be involved in the high degree of comorbidity between depression and chronic pain.

The potential introduction of bias in effect sizes through adjustment for manual labour in the 23andMe-Pfizer chronic pain grade GWAS could also have affected MR results. Sensitivity analyses to check for this kind of bias were not possible as summary statistics for an unadjusted CPG GWAS do not exist (see also Appendix 2).

MR-RAPS results show that something specific to chronic pain is causal for MDD, and not just non-genetic factors closely associated with chronic pain. Results of these causal analyses indicate that chronic pain itself contributes to development of MDD. Although studies on causal relationships between pain and depression often show mixed results (see 1.3.2), including that depression has a causal effect on chronic pain, MR-RAPS results here indicated that MDD did not cause chronic pain, suggesting chronic pain subsequent to depression is not a direct result of the depression itself. Causal analyses also suggest multifactorial pathways leading to MDD from pain, detected as widespread horizontal pleiotropy – chronic pain itself has an independent, causal effect, but many other factors (genetic and environmental) are likely involved.

7.6 Strengths & Limitations

A main limitation to analyses in this thesis, and to GWAS analyses in general (Mills & Rahal, 2019, 2020), is that cohorts primarily consist of white European-ancestry participants (e.g. GS: SFHS: 99% white). While this is approximately representative of a Scottish population (Smith et al., 2006, 2013), this demographic make-up is non-representative in terms of ethnicity of global or even UK-wide populations, and SNP-trait associations may not be generalisable to populations with different ancestry, including admixed samples. In addition, ethnicity itself also acts to confound associations between SNPs and traits of interest – if magnitude of trait values or disease prevalence vary between ethnic groups, this may generate spurious SNP-trait associations (Medina-Gomez et al., 2015). This confounding occurs even though race or ethnicity as biological

constructs cannot be quantified in a meaningful way. In other words this confounding is epidemiological rather than genetic, and is different from population structure ("AAPA Statement on Biological Aspects of Race," 1996; Blackburn, 2000; Yudell et al., 2016).

Another potential limitation is that extensive heterogeneity in MCP as a trait construct is not fully explored in, and is beyond the scope of, this thesis. As discussed above and in the previous section7.3.2, despite advances made in MDD research using broad phenotyping, more detailed phenotyping is also of value, and examining both detailed and broad chronic pain phenotypes in a non-mutually exclusive fashion would be of interest. Heterogeneity in terms of clinical heterogeneity in chronic pain is only examined with respect to MDD and vice versa (a methodological constraint with BUHMBOX analysis). With respect to causal analyses, again as chronic pain measured as MCP is a broad trait construct, this is not really modifiable in the way e.g., cigarette smoking is, meaning MR results are difficult to interpret or involve in e.g., treatment guidelines in chronic pain.

As previously discussed, (4.4.1.6, see also Appendix 2), use of summary statistics adjusted for traits that are genetically correlated with the main trait of interest, such as manual labour and BMI in the case of CPG, has the potential to bias GWAS results and subsequent analyses involving GWAS outputs. This is a possibility with 23andMe-Pfizer CPG GWAS outputs, and may have affected genetic correlation and causal analyses.

A key strength is that analyses undertaken in this thesis represent the largest GWAS of a chronic pain trait to date, giving insight into potential mechanisms of chronic pain development. A novel chronic pain trait based on recent re-definition of chronic pain emphasising its independence from nociception and detectable biological causes, was derived, providing genetic research in keeping with shifting paradigms in defining chronic pain. This way of exploring chronic pain as a 'broad', complex trait phenotype is similar to recent study of broad MDD phenotypes and represents a potentially powerful route to understanding genetic contribution to chronic pain. This view of chronic pain is also in line with theories on chronic pain as a complex trait following the biopsychosocial model of disease. In addition, analyses in this thesis employ a varied range of statistical

genetics techniques to answer research objectives, examining common genetic variation, genetic correlation, and pleiotropy from multiple angles – several techniques (BUHMBOX, cFDR, MR-RAPS) have never before been applied in research into MDD and chronic pain.

7.6 Future Directions

7.6.1 Representative Cohorts

Recent advances for successful GWAS analysis in ancestrally diverse populations (Peterson et al., 2019), including in admixed populations, such as Tractor (Atkinson et al., 2021) could be used to investigate genetic variation contributing to chronic pain, including in the entirety of UK Biobank (as opposed to the white British sub-sample), addressing a key limitation of ancestrally homogenous cohort use. Tractor is a statistical framework and associated software package that allows for the inclusion of admixed individuals in GWAS, achieved through leveraging local ancestry in contrast to traditional GWAS where population stratification is adjusted for using e.g. genetic principal components, which represent a broader estimate of admixture (Atkinson et al., 2021).

The model used within Tractor allows for inclusion of terms that estimate SNP-trait association within different ancestry categories (Equation 7.1), where *b* values are effect estimates, X_1 represents the number of haplotypes of the index ancestry at the locus in question for each individual, X_2 is the number of copies of the risk allele coming from the first ancestry, X_3 is the copies coming from the second ancestry, and X_4 - X_k are additional covariates such as age, sex, and an estimate of global (rather than local) ancestry (Atkinson et al., 2021).

$$Logit[Y] = b_0 + b_1X_1 + \ b_2X_2 + b_3X_3 \ldots + b_kX_k$$

Equation 7. 1: Tractor association model

Tractor analysis can therefore boost SNP discovery power and have a downstream effect on utility of PRS in less ancestrally homogenous populations. This would be achieved by allowing calculation of ancestry-specific SNP effect sizes which then contribute to weighting in PRS calculation. Another potential benefit of the inclusion of mixed-ancestry participants in GWAS is the ability to use admixture based fine-mapping approaches for discovery of causal variants. Fine-mapping refers to analysis of trait-associated genetic loci, found through

GWAS, in order to determine which genetic variants within these regions are putative causal variants for the trait (Schaid et al., 2018). One fine-mapping approach is use of admixture (trans-ethnic fine-mapping) whereby meta-analysis is carried out across GWASs of the same trait across ancestrally diverse populations. This allows for pinpointing putative causal variants as patterns of LD vary between populations – signals that persist despite variation in LD block structure indicate potential causal variants (Y. R. Li & Keating, 2014). Tractor analysis specifically can also improve location of causal variants within GWAS results due to improved power to find causal variants in non-European populations (Atkinson et al., 2021). In addition to application of newer GWAS methods for diverse and admixed populations being applied to the full UK Biobank sample, more ethnically and ancestrally diverse cohorts such as All of Us (All of Us Research Program Investigators, 2019), could also be used in GWAS analyses of chronic pain conditions and chronic pain phenotypes. All of Us is a large general-population biobank research program funded by the National Institutes of Health (NIH) in the USA, with recruitment ongoing and a commitment to recruiting a diverse participant pool. The program aims to recruit 1 million participants and to contain genotyping data in addition to information on a wide range of traits and conditions of public health interest.

7.6.2 New Pain Data for UK Biobank

Other emerging datasets could also be used to further investigate the genetics of MCP, in addition to allowing for derivation of more detailed chronic pain phenotypes. In particular, the recent release of the new UK Biobank pain questionnaire data (UKB Data Showcase category 154: Experience of Pain) gives the opportunity to derive narrower chronic pain phenotypes, due to comparative increased detail of questioning compared to the baseline UK Biobank pain questionnaire. Whereas the baseline UK Biobank touchscreen questionnaire on pain effectively contained just two questions, on pain site and the duration of pain at that site, the new UK Biobank pain questionnaire, an online follow-up questionnaire completed by ~167,000 participants, contains ten sections. Questions asked of participants include ones on the location (an extended number of site options in comparison to the baseline pain questionnaire), nature, and impact of pain, in addition to sections separately asking about medical

conditions, depression, fatigue, and health outcomes that may be relevant to pain. Separate sections also exist on neuropathic pain and headache.

The new UK Biobank pain questionnaire also provides opportunity for potentially very large longitudinal studies of pain and depression – baseline pain data were collected at recruitment from 2006-2010, whereas the new pain questionnaire was sent to participants initially in 2017 (UK Biobank, 2020). With more detailed information on timing, duration, and associated symptoms and pain, genetic similarity, and differences between chronic pain with certain characteristics could also be explored. However, there may be issues with trying to carry out longitudinal studies as data collection (for the new pain questionnaire) was carried out in instances that were relatively far apart in time, and on a subset of participants rather than the entire UK Biobank sample of 0.5 million (Caruana et al., 2015). Another potentially interesting analysis would be comparison of 'depression-in-pain', derived from mood information in the new UKB pain questionnaire, versus MDD with and without comorbid chronic pain.

Outside of UK Biobank, output from analyses of conditions associated with chronic pain, such as Ehlers-Danlos syndrome (Forghani, 2019), which have previously not been investigated at scale and/or using GWAS are expected in coming years. The relationship between these phenotypes and MCP in terms of genetic overlap and causal effect would also be of interest.

7.6.3 Alternative Approaches to Pleiotropy

As previously discussed, the clinical heterogeneity of chronic pain, and of MCP specifically, cannot be fully explored using just BUHMBOX. This method can only examine clinical heterogeneity in a trait with respect to a second, defined trait or condition – there is no scope for an agnostic investigation of clinical heterogeneity. Additional data-driven approaches, such as cluster analysis as previously used to explore clinical heterogeneity in a range of traits and diseases (Guo et al., 2017; Mu et al., 2017; Nagel et al., 2018) could be used to characterise heterogeneity within the broad MCP trait construct, particularly in conjunction with new pain questionnaire data.

Emerging methods could also be used to interrogate location-specific pleiotropy in MDD and chronic pain and act as an extension of cFDR analyses. An example of

such a method is LAVA (Local Analysis of coVariant Association) a recently developed framework for calculating local genetic correlation (Werme et al., 2021). Most current genetic correlation analysis methods, including LDSR, measure a global average of r_g across the genome – this may not allow detection of heterogeneous genetic correlation relationships where r_g varies between genomic regions and where global r_g is non-significant, and does not give an idea of specific pleiotropic loci of interest. In contrast, LAVA allows for estimating location-specific r_g, and similarly to cFDR analyses undertaken in this thesis, could be used to investigate shared loci in MCP and MDD.

7.6.4 Whole-Exome Data and Chronic Pain

Whole-exome data, recently released for UK Biobank, could also be used to explore exome regions associated with chronic pain. As previously discussed, (1.3.3.1, 2.2.1.1), rare variants are excluded from GWAS analyses, but studying their association with traits of interest is possible through aggregating them at the gene or exon level. The exome refers to sections of the genome containing protein coding sections (exons), which comprise a small fraction of the total genome (Dunham et al., 2012) but are where most rare variants of large effect are thought to reside. Although most of the genetic variation associated with complex traits such as MCP and chronic pain in general is likely to be common and of small effect (i.e., SNP) (1.3.3.1), rarer variants of large effect in exons could also contribute to trait variation (2.2.1.1).

Whole-exome association studies, whereby larger-effect and rarer genetic variation is tested for association with traits or conditions of interest, have been used in a clinical setting to determine genetic causes or contributory factors to a range of genetic disorders, including both single-gene and complex disorders (Rabbani et al., 2014; Retterer et al., 2016; Srivastava, Cohen, Vernon, et al., 2014). These include heterogeneous monogenic disorders such as hearing loss, movement disorders (study N ranging from 9-270) and rare subtypes of diabetes (N = 1) (reviewed by Rabbani et al., 2014), and determination of specific pathological phenotype such as abnormality of the nervous system or mitochondrial dysfunction (with successful diagnoses in ~30% of N = 3, 040 cases) (Retterer et al., 2016).

In addition to these applications, whole-exome association studies have the potential to shed light on rare genetic variation contributing to common, complex traits using large cohorts (Cirulli et al., 2020; Povysil et al., 2019). The missing heritability (see also 2.2.1.1) in complex traits could be partially explained by the contribution of rare variants to phenotypic variance remaining unmeasured (Eyre-Walker, 2010) (as these variants are present at much lower allele frequency in the population due to negative selection pressures, and so are unmeasured as part of GWAS). Rare variants have been found to contribute to complex traits such as height (Marouli et al., 2017), autism spectrum disorder (De Rubeis et al., 2014; Wilfert et al., 2021), and schizophrenia (Fromer et al., 2014; S. M. Purcell et al., 2014), in studies with sample sizes ranging from 600 to 450,000, indicating exome data available in the UK Biobank (N ~ 50,000 with an eventual goal of sequencing the exomes of 450,000 participants) could also be used to investigate rare variant associations with complex traits such as chronic pain.

With respect to chronic pain more specifically, in their large-scale analyses of UK Biobank exome data Cirulli et al found rare variants within the genes *TET3, PTPRR, PHLDB1, TSPYl4, IQCM, ACTN2* to be associated at a suggestive level ($p < 10^{-3}$) with back pain for 3+ months (N_{case} = 6819, $N_{control}$ =2981), within *MMS19* with hip pain for 3+ months (N_{case} = 3378, $N_{control}$ =953), within *TNS3, ZNF347, HEATR6, RBL1 and FAM17A1* with stomach or abdominal pain for 3+ months (N_{case} = 1631, $N_{control}$ =1401), and within *PRG4, NLRC5, ITGAE, PLEKHA6 and EIF2AK4* with knee pain for 3+ months (N_{case} = 6798, $N_{control}$ = 1710). Rare variants were also found to be associated ($p < 10^{-3}$) with neck/shoulder pain for 3+ months, implicating genes *LARP7, LRRC7, OTOG, MEI1, GDF1, RSPH1* and *ZNF462* (N_{case} = 6009, $N_{control}$ =2656). Although for the most part these associations are suggestive ($p > 10^{-6}$), they again suggest exome association studies could be of interest in the study of chronic pain.

7.6.5 Genomic Structural Equation Modelling Approaches

A potentially powerful way to investigate genetic variation contributing to chronic pain development and maintenance, in contrast to deriving a 'general' or broad chronic pain trait such as MCP, could be to use various applications of GenomicSEM (Grotzinger et al., 2019). GenomicSEM (genomic structural equation modelling), is a flexible framework that allows for studying multivariate genetic

architecture of groups of related traits through fitting network models, taking GWAS summary statistics as input.

GenomicSEM applications of interest for the study of chronic pain include common factor GWAS approaches. Existing GWAS output for chronic pain conditions could be used in this way to find common genetic variation shared across chronic pain conditions e.g., from large-scale GWASs of Crohn's disease, rheumatoid arthritis, and lupus. This is a similar approach to recent work attempting to uncover variation associated with a 'p factor' associated with psychiatric traits (Grotzinger et al., 2019). It would also be of interest to compare the output of this analysis with that of the MCP GWAS, to further explore the extent to which MCP represents genetic variation associated more generally with chronic pain.

Another genomicSEM application of interest could be 'GWAS by subtraction' (Demange et al., 2021). GWAS by subtraction involves 'subtracting' the genetic influence on a trait from each SNPs association with a second trait – the remaining SNP-trait association values represent a new GWAS of an unmeasured trait of interest. These kinds of analyses could be used to highlight genetic variation captured by chronic pain condition GWASs that is independent of chronic pain itself, and which may be more specific to disease processes. For example, if the genetic influence on MCP were subtracted from that on e.g., rheumatoid arthritis, remaining SNP-arthritis association may highlight loci with a more specific role in disease progression. A separate and potentially useful way to use GWAS by subtraction could also be 'unadjustment' of GWAS where adjustment for particular covariates has likely introduced bias into estimation of SNP-trait associations e.g., subtraction of the relationship a SNP has with manual labour from its relationship with CPG-adjusted-for-manual-labour, if original CPG datasets and raw data are not available for reanalysis for legal and data protection reasons.

Exploratory factor analysis (EFA) could also be used in conjunction with existing chronic pain condition GWAS outputs, and MCP GWAS output, to explore the relationship between MCP and chronic pain conditions generally, and to further characterise heterogeneity in the MCP trait construct e.g., do certain kinds of chronic pain condition cluster with MCP more so than others? EFA could also

present a way to investigate the genetic contributions to broad factors associated with chronic pain and affective spectrum disorders.

Affective spectrum disorders is an umbrella term describing a group of mood conditions including MDD, GAD, and PTSD among others, which are also commonly comorbid with chronic pain generally in addition to chronic pain conditions including fibromyalgia, migraine and IBS (Gardner & Boles, 2011; Hudson & Pope, 1994). As significant genetic overlap between chronic pain and mood disorders, including MDD but also phenotypes such as neuroticism and PTSD, has been demonstrated (4.3.3), it would be interesting to further explore the affective spectrum disorder group using GWAS outputs and genomicSEM.

7.6.4 Affective Dysregulation and Pain

Another potentially interesting area linking chronic pain and psychopathology may be aspects of emotional self-regulation. As previously mentioned, comorbidity in chronic pain and multiple affective spectrum disorders suggests affect and emotion play a key role in the development of physical and psychiatric distress. Results both from the literature and from analyses undertaken in this thesis, demonstrating significant genetic correlation between MCP and mood disorders and mood-related phenotypes in particular (1.1.5, 4.3.3), are in agreement with this.

Although a negative genetic correlation was seen between MCP and autism spectrum disorder, the relationship between neurodevelopmental disorders such as autism spectrum disorder and chronic pain warrants further investigation. In addition to physical comorbidities in autism spectrum disorder that potentially contribute to chronic pain (1.1.5), certain emotion-related personality constructs associated with autism spectrum disorders, such as alexithymia, could be involved. Alexithymia refers to difficulty identifying and expressing emotions and was originally described in studies of patients with a range of psychosomatic conditions (Goerlich, 2018; Poquérusse et al., 2018). Alexithymia involves confusing bodily sensation and emotion – many people scoring highly on alexithymia scales may only be able to describe emotion in terms of bodily sensation. Although alexithymia shows high degree of overlap with autism spectrum disorder it is not universal or a core component (Kinnaird et al., 2019): alexithymic traits are also common in people with neurodegenerative disease,

depression, and eating disorders, and in neurotypical relatives of autistic people (De Berardis et al., 2017; Martino et al., 2020; Poquérusse et al., 2018). Alexithymia is also thought to contribute to emotional dysregulation, and has been found to be more common in those with chronic pain (Aaron et al., 2019).

7.7 Overall Conclusions

In addressing the objective of uncovering common genetic variation associated with chronic pain phenotypes, it was demonstrated that examining chronic pain as a broad phenotype is a powerful way to look at chronic pain.

The objective of investigating pleiotropy and genetic correlation was achieved through use of well-powered MCP GWAS results. Significant genetic overlap was found between MCP and a range of traits of interest, including MDD ($r_g = 0.53$). This overlap highlighted an affective component to chronic pain and indicated that genetic variation associated with chronic pain conditions like IBDs may not be associated with pain or pain experience, but instead with disease processes more specifically – this is in line with recent work highlighting the value of studying chronic pain as a disease rather than a symptom (see 1.1.1).

The comorbidity between chronic pain and psychiatric, mood and neurodevelopmental phenotypes has been documented to varying degrees, but in some cases, results are mixed and/ or may only hold in specific non-general-population cohorts (see 1.1.5). The genetic correlation between MDD and chronic pain (MCP) was quantified in analyses in this thesis (4.3.3), and results specifically informed subsequent causality analyses. Horizontal pleiotropy in MR as indicated in genetic correlation and cFDR results, for example, has the potential to bias causal estimates. Clinical heterogeneity can also be thought of as a type of pleiotropy. Therefore, further characterising pleiotropy with respect to MDD and chronic pain was also a key part of this thesis and of investigating causal relationships between MDD and chronic pain. This was achieved through BUHMBOX analyses, and through use of MR methods that quantified and adjusted for horizontal pleiotropy amongst SNP instruments. A significant causal effect roughly equivalent to each +1 increase in MCP trait value resulting in 17% increase in the odds of having MDD (OR = 1.17) was found.

Mixed results are also seen in studies of causal relationships between MDD and chronic pain. Genetic correlation and MR analyses in this thesis contribute to

answering both of these outstanding questions, quantifying genetic overlap and size and direction of the causal effect of chronic pain on MDD

Genes found to be associated with MCP (N = 143, see 4.4.3, Appendix 1) highlighted neural system development and functioning, immune processes, and cell cycle regulation as broad functional categories important for chronic pain development and maintenance. This is in agreement with past studies indicating the importance of neural plasticity in chronic pain development, and indicating changes in brain structure and function associated with chronic pain development and maintenance (see 1.1.4).

Overall, results of analyses completed as part of this thesis emphasise the existence of genetic variation shared across chronic pain conditions regardless of putative cause or mechanistic description. Large-scale investigations of a broad chronic pain trait were shown to be a powerful way to explore this genetic variation. Viewing chronic pain a complex disease trait that follows the biopsychosocial model of disease draws on both historical theories of pain and chronic pain development, and is also in line with recent work of IASP taskforces to redefine pain.

Appendix 1: Genes Associated with Multisite Chronic Pain

A total of 143 genes were found to be significantly associated with MCP using MAGMA gene level analyses (see 2.3.1.1). Significance thresholds for association are Bonferroni-adjusted (0.05 divided by number of genes tested (18, 670)) = 2.67×10^{-6}. Genes of interest associated with MCP, their association with other traits and disorders, and functional roles are also discussed in the published article that resulted from these analyses (Johnston et al., 2019).

As 143 genes total were found to be significantly associated with MCP, it was not feasible to discuss them in the same level of detail as with *DCC* (see 4.4.3.4). *DCC* is described as it was the most significantly associated gene of the 143, has also been implicated in other psychiatric and brain-structure related phenotypes, and the pathway showing most significant enrichment of MCP-associated genes was found to be *DCC*-mediated attractive signalling.

CHR	NSNPS	ZSTAT	P	SYMBOL
1	34	5.3574	4.22×10^{-8}	BAI2
1	30	4.9792	3.19×10^{-7}	PABPC4
1	417	4.6593	1.59×10^{-6}	DNAJC6
1	32	4.7804	8.75×10^{-7}	TRMT13
1	119	4.5989	2.12×10^{-6}	SORT1
1	41	4.7137	1.22×10^{-6}	PSMA5
1	236	4.8592	5.89×10^{-7}	FAM212B
1	8	4.6565	1.61×10^{-6}	C1orf51
1	50	6.1047	5.15×10^{-10}	MRPS21
1	77	5.7623	4.15×10^{-9}	PRPF3
1	311	5.6039	1.05×10^{-8}	RPRD2
1	57	6.1773	3.26×10^{-10}	TARS2
1	12	6.3192	1.31×10^{-10}	ECM1
1	192	4.6001	2.11×10^{-6}	GATAD2B
1	12	4.577	2.36×10^{-6}	CRTC2
1	214	4.8049	7.74×10^{-7}	NUP210L
1	1673	5.1844	1.08×10^{-7}	RABGAP1L
1	373	5.1296	1.45×10^{-7}	FAM129A
1	157	5.3923	3.48×10^{-8}	CEP170
1	411	5.0361	2.38×10^{-7}	SDCCAG8
1	1883	4.6993	1.31×10^{-6}	KIF26B
2	3196	5.0683	2.01×10^{-7}	NRXN1
2	21	4.6699	1.51×10^{-6}	VAMP5
2	1078	5.086	1.83×10^{-7}	SLC4A10
2	153	4.6202	1.92×10^{-6}	RFTN2
2	94	4.5783	2.34×10^{-6}	AC011997.1

2	67	4.6412	1.73×10^{-6}	BOLL
2	103	5.3326	4.84×10^{-8}	LANCL1
2	448	5.8441	2.55×10^{-9}	CPS1
2	4060	5.7201	5.32×10^{-9}	ERBB4
2	659	5.1788	1.12×10^{-7}	SPHKAP
3	208	4.8881	5.09×10^{-7}	SMARCC1
3	35	4.9732	3.29×10^{-7}	DHX30
3	12	4.8195	7.20×10^{-7}	LAMB2
3	3	4.7316	1.11×10^{-6}	CCDC71
3	19	5.1965	1.02×10^{-7}	C3orf84
3	70	5.1723	1.16×10^{-7}	CCDC36
3	4	4.8776	5.37×10^{-7}	RP11-3B7.1
3	86	5.2367	8.17×10^{-8}	RHOA
3	7	5.8232	2.89×10^{-9}	TCTA
3	97	4.9452	3.80×10^{-7}	DAG1
3	150	5.9396	1.43×10^{-9}	BSN
3	7	5.3003	5.78×10^{-8}	MST1
3	47	5.9543	1.31×10^{-9}	RNF123
3	3	5.7329	4.94×10^{-9}	AMIGO3
3	3	5.7329	4.94×10^{-9}	GMPPB
3	88	5.4909	2.00×10^{-8}	IP6K1
3	16	5.6163	9.75×10^{-9}	CDHR4
3	7	4.8532	6.07×10^{-7}	UBA7
3	38	5.3125	5.41×10^{-8}	TRAIP
3	16	5.6583	7.64×10^{-9}	CAMKV
3	25	4.9781	3.21×10^{-7}	MST1R
3	12	4.892	4.99×10^{-7}	CTD-2330K9.3

3	27	6.3576	1.02×10^{-10}	MON1A
3	238	5.6467	8.18×10^{-9}	RBM6
3	26	6.0831	5.89×10^{-10}	RBM5
3	58	5.992	1.04×10^{-9}	SEMA3F
3	2	4.6616	1.57×10^{-6}	GNAT1
3	226	4.9441	3.82×10^{-7}	EIF4E3
3	4594	5.784	3.65×10^{-9}	ROBO2
3	491	5.8122	3.08×10^{-9}	BBX
3	50	5.7467	4.55×10^{-9}	MSL2
3	119	4.8309	6.80×10^{-7}	PCCB
3	590	5.5475	1.45×10^{-8}	STAG1
3	21	4.9664	3.41×10^{-7}	PSMD2
4	215	4.6644	1.55×10^{-6}	GRK4
4	987	7.3313	1.14×10^{-13}	MAML3
5	620	5.0251	2.52×10^{-7}	FAM172A
5	616	6.6132	1.88×10^{-11}	GABRB2
5	2066	4.7086	1.25×10^{-6}	TENM2
5	60	4.632	1.81×10^{-6}	NPM1
6	61	5.8246	2.86×10^{-9}	UQCC2
6	100	5.988	1.06×10^{-9}	IP6K3
6	64	4.5588	2.57×10^{-6}	LEMD2
6	140	4.6315	1.82×10^{-6}	PACSIN1
6	200	5.839	2.63×10^{-9}	C6orf106
6	45	5.6501	8.02×10^{-9}	SNRPC
6	233	5.5866	1.16×10^{-8}	UHRF1BP1
6	156	4.6163	1.95×10^{-6}	PXT1
6	132	4.848	6.24×10^{-7}	FHL5

6	140	5.5273	1.63×10^{-8}	LIN28B
6	544	6.0603	6.79×10^{-10}	FYN
6	1493	5.0295	2.46×10^{-7}	LAMA2
6	20	4.8934	4.95×10^{-7}	GINM1
6	109	4.8721	5.52×10^{-7}	KATNA1
6	89	4.6412	1.73×10^{-6}	LATS1
6	44	4.7059	1.26×10^{-6}	NUP43
6	154	4.6245	1.88×10^{-6}	PCMT1
7	3444	7.3968	6.97×10^{-14}	SDK1
7	245	5.4101	3.15×10^{-8}	SP4
7	381	5.3545	4.29×10^{-8}	GRM3
7	242	5.9014	1.80×10^{-9}	SLC25A13
7	792	6.6565	1.40×10^{-11}	FOXP2
8	52	4.9155	4.43×10^{-7}	PURG
8	291	4.6758	1.46×10^{-6}	AGO2
8	514	4.8893	5.06×10^{-7}	PTK2
9	301	6.1799	3.21×10^{-10}	FAM120A
9	399	5.3574	4.22×10^{-8}	PHF2
9	2845	6.4635	5.12×10^{-11}	ASTN2
9	103	4.5993	2.12×10^{-6}	GOLGA1
9	297	4.8088	7.59×10^{-7}	SCAI
9	94	5.409	3.17×10^{-8}	DNM1
9	262	6.3722	9.31×10^{-11}	EXD3
10	847	4.6285	1.84×10^{-6}	NEBL
10	160	6.0382	7.79×10^{-10}	MLLT10
10	74	4.575	2.38×10^{-6}	ZRANB1
10	364	5.3329	4.83×10^{-8}	JAKMIP3

11	49	4.639	1.75×10^{-6}	F2
11	139	4.7341	1.10×10^{-6}	CKAP5
11	36	5.074	1.95×10^{-7}	TSKU
11	835	4.8551	6.02×10^{-7}	NCAM1
12	430	4.9129	4.49×10^{-7}	RERG
12	604	4.6857	1.40×10^{-6}	PTPRO
12	32	4.868	5.64×10^{-7}	ERBB3
13	65	4.7009	1.30×10^{-6}	OLFM4
13	83	5.0493	2.22×10^{-7}	EFNB2
14	385	4.6906	1.36×10^{-6}	NUMB
14	39	4.6872	1.38×10^{-6}	ZFYVE21
14	232	4.997	2.91×10^{-7}	PPP1R13B
15	68	5.9866	1.07×10^{-9}	VPS33B
16	8933	4.7735	9.05×10^{-7}	RBFOX1
16	31	4.862	5.81×10^{-7}	MARVELD3
16	68	4.9134	4.48×10^{-7}	ATXN1L
16	155	4.9441	3.82×10^{-7}	IST1
16	62	5.0294	2.46×10^{-7}	ZNF821
17	83	5.5391	1.52×10^{-8}	DCAKD
17	154	5.8159	3.01×10^{-9}	NMT1
17	12	4.7923	8.24×10^{-7}	HEXIM2
18	344	5.3098	5.49×10^{-8}	ASXL3
18	4053	8.2342	9.04×10^{-17}	DCC
18	637	4.5548	2.62×10^{-6}	TCF4
19	46	4.7804	8.75×10^{-7}	PTBP1
19	93	4.7278	1.14×10^{-6}	SLC44A2
19	46	5.1026	1.67×10^{-7}	ILF3

19	46	5.6931	6.24×10^{-9}	ATP13A1
19	30	4.9853	3.09×10^{-7}	ZNF101
20	1440	5.3386	4.68×10^{-8}	SLC24A3
20	99	5.0374	2.36×10^{-7}	TM9SF4
20	67	4.9893	3.03×10^{-7}	KIF3B
20	86	4.9567	3.59×10^{-7}	ASXL1
20	201	5.4406	2.66×10^{-8}	C20orf112
20	348	4.562	2.53×10^{-6}	ZBTB46
22	357	4.972	3.31×10^{-7}	TCF20

Table A1. 1: Genes found to be significantly ($p < 2.67 \times 10^{-6}$) associated with MCP in MAGMA gene-level analyses.

CHR = chromosome, NSNPS = number of SNPs in the GWAS data that were annotated to the gene, ZSTAT = gene Z-value, P = p value for association between gene and MCP.

General Function	Description	Gene Symbol(s)	Reference(s)
Nervous system development & function	Astrocyte development	UTRN	(Sogos et al., 2002)
	Neuronal cell function	DYNC1I1	(Goldstein & Yang, 2000)
		SOX6	(Kurtsdotter et al., 2017)
	Presynaptic cytoskeleton organisation	BSN	(Frank et al., 2010)
	establishment of nervous system connectivity	TENM2	(Rebolledo-Jaramillo & Ziegler, 2018)
	neurite formation, neuron morphogenesis	PACSIN1	(Mondal et al., 2020)
	Glutamatergic neurotransmission & memory	GRM3	(De Quervain & Papassotiropoulos, 2006)
	Nervous system development	NCAM1	(Paratcha et al., 2003)
		EFNB2	(Cramer & Miko, 2016)
		ATXN1L	(Didonna et al., 2020)
		TCF4	(Mesman et al., 2020)
		BBX	(T. L. Chen et al., 2014)
		PTK2	(X. R. Ren et al., 2004)
		ERBB3	(Britsch et al., 1998)
		SORT1	(Nykjaer et al., 2004)
		FYN	(Yaka et al., 2002; Zamoyska et al., 2003)
		RHOA	(K. Y. Wu et al., 2005)
		DAG1	(K. M. Wright et al., 2012)
		AMIGO3	(Kuja-Panula et al., 2003)
		ROBO2	(T. Kidd et al., 1998)
	Synapse development and plasticity	CTNNA2	(Zhong et al., 2016)
		CEP120	(Guerrier & Polleux, 2007)
		KNDC1	(Hayashi et al., 2017)
		CA10	(Sterky et al., 2017)
		FOXP2	(Vernes et al., 2011)
		NRXN1	(Araç et al., 2007; Missler et al., 2003)
		SLC4A10	(Gurnett et al., 2008)
		LANCL1	(W. Zhang et al., 2009)
		SEMA3F	(Nakayama et al., 2018)
		LAMB2	(Hunter et al., 1989; Nishimune et al., 2004)

	Peripheral nerve myelination	DAG1	(Masaki & Matsumura, 2010; Saito et al., 2003)
Cell cycle progression	DNA proof-reading	EXD3	(Bębenek & Ziuzia-Graczyk, 2018)
	Regulation	ANAPC4	(J. M. Peters, 2006)
		PRC1	(J. Li et al., 2018; Shrestha et al., 2012; C. Zhu & Jiang, 2005)
		BOLL	(Kang et al., 2015)
		LATS1	(Furth & Aylon, 2017)
	Sister chromatid organisation	STAG1	(van der Lelij et al., 2017)
		LEMD2	(von Appen et al., 2020)
		KATNA1	(McNally et al., 2000)
		CKAP5	(Schneider et al., 2017)
		KIF3B	(Zhou et al., 2019)
	Apoptosis	FAM120A	(Tanaka et al., 2009)
		MON1B	(Kinchen & Ravichandran, 2010)
		FAM129A	(H. Ji et al., 2012)
		DHX30	(Bosco et al., 2020)
		FAF1	(Menges et al., 2009)
		SEMA3F	(Nakayama et al., 2018)
		PTK2	(Kurenova et al., 2004)
		PTPRO	(Motiwala et al., 2004)
		OLFM4	(Anholt, 2014)
		PPP1R13B	(Samuels-Lev et al., 2001)
	Synapsis	CCDC36	(Stanzione et al., 2017)
	Cell proliferation	NPM1	(Okuda et al., 2000)
		PTK2	(X. R. Ren et al., 2004)
		F2	(Danckwardt et al., 2006)
		ERBB3	(Holbro et al., 2003)
		KIF3B	(Zhou et al., 2019)
	Cytokinesis	IST1	(Renvoise et al., 2010)
	DNA replication regulation	PURG	(Johnson et al., 2013)
Immune-related	Neutrophil activation	UTRN	(Cerecedo et al., 2010)
	T cell activation	PABPC4	(H. Yang et al., 1995)

		FYN	(Sharma et al., 2016; Zamoyska et al., 2003)
	B-cell antigen receptor-mediated signalling	RFTN2	(Saeki et al., 2003)
	innate immune signalling	TRAIP	(M. Zhang et al., 2012)
		MST1R	(Sakamoto et al., 1997)
		ILF3	(Pfeifer et al., 2008)
		VPS33B	(Akbar et al., 2016)
		OLFM4	(Wenli Liu, Yan, et al., 2010)
	immune surveillance	NCAM1	(Van Acker et al., 2017)
Other	Brain-specific inhibition of angiogenesis	BAI2	(Okajima et al., 2010)
	Angiogenesis	F2	(Danckwardt et al., 2006)
	Thrombosis	SLC44A2	(Bennett et al., 2020)
	Heat shock protein	DNAJC6	(Alderson et al., 2016)
	tRNA processing	TRMT13	(Towns & Begley, 2012)
		TARS2	(Lightowlers et al., 2015)
	Protein transport	RABGAP1L	(T. Itoh et al., 2006)
		NUP43	(Cronshaw et al., 2002)
		VPS33B	(Ambrosio & Di Pietro, 2019)
		TM9SF4	(Vernay et al., 2018)
	protein degradation	PSMA5	(Tomko & Hochstrasser, 2013)
		UBA7	(H. Li et al., 2018)
		PSMD2	(Rock et al., 1994)
	Inhibitor of serine/threonine-protein kinase	PAK4	(Vadlamudi & Kumar, 2003)
		FAM212B	(Y. Y. Liu et al., 2019)
	mitochondrial protein synthesis	MRPS21	(Kenmochi et al., 2001)
		DHX30	(Bosco et al., 2020)
	protein repair	PCMT1	(DeVry & Clarke, 1999; Tsai & Clarke, 1994)
	mitochondrial metabolism	PCCB	(Chapman et al., 2018; Ugarte et al., 1999)
		UQCC2	(Tucker et al., 2013)
		SLC25A13	(Convertini et al., 2019)
	pre-mRNA processing	PRPF3	(Heng et al., 1998; Martínez-Gimeno et al., 2003)
		PTBP1	(Vuong et al., 2016)
	Regulation of gene transcription	CRTC2	(Cheng & Saltiel, 2006)

		SMARCC1	(He et al., 2020)
		GATAD2B	(Willemsen et al., 2013)
		MLLT10	(Ogoh et al., 2017)
		HEXIM2	(Yik et al., 2005)
		ASXL3	(Katoh & Katoh, 2004)
		ZNF101	(Bellefroid et al., 1993)
		TCF20	(Sanz et al., 1995)
		MSL2	(L. Wu et al., 2011)
		AGO2	(Hansen et al., 2011)
	Adipogenesis	ASXL1	(Park et al., 2011)
	mRNA processing	EIF4E3	(Joshi et al., 2004)
		SNRPC	(Du & Rosbash, 2002)
		ILF3	(Pfeifer et al., 2008)
	Cell development/ differentiation	UHRF1BP1	(El Baroudi et al., 2017; Unoki et al., 2004)
		LEMD2	(M. D. Huber et al., 2009)
	Organelle transport	KIF26B	(Miki et al., 2001)
	Myogenesis	VAMP5	(Zeng et al., 1998)
		UQCC2	(Feichtinger et al., 2017)
	Urea cycle	CPS1	(Martínez et al., 2010)
	Cell adhesion, migration, outgrowth	RHOA	(Valderrama et al., 2006)
		DAG1	(Morikawa et al., 2017)
		MST1R	(Ghigna et al., 2005)
		LAMA2	(Vuolteenaho et al., 1994)
		SCAI	(Brandt et al., 2009)
		OLFM4	(Wenli Liu, Lee, et al., 2010)
		EFNB2	(F. Zhu et al., 2020)
		ZFYVE21	(Nagano et al., 2010)
		PTK2	(Chan et al., 2009; Hsia et al., 2003)
	Membrane trafficking	MON1A	(Bagley et al., 2012)
	Cardiac myofibril assembly	NEBL	(Moncman & Wang, 2002)
	Manganese transport	ATP13A1	(Anagianni & Tuschl, 2019; Farley, 2012)
	Potassium-dependent sodium/calcium	SLC24A3	(Kraev et al., 2001)

	exchange		
	DNA damage response signalling	SDCCAG8	(Chaki et al., 2012)

Table A1. 2: Function of genes associated with MCP.

This table (Table A1.2) is designed to give a brief overview of gene function (note also that some of these categories are non-mutually exclusive). For example, *PAK4* (RAC1 Activated Kinase 4) is listed under its main associated function (as an inhibitor of serine/threonine-protein kinase), but the protein encoded by this gene is also implicated in cell motility, proliferation, and angiogenesis. The full picture of gene function, gene interaction, and potential health and disease related effects associated with every gene found to be associated with MCP is in many cases not fully known and is beyond the scope of this thesis.

GeneSet	**p value**	**genes**
REACTOME_DCC_MEDIATED_ATTRACTIVE_SIGNALING	5.10×10^{-5}	DCC, NCK1
REACTOME_PLC_BETA_MEDIATED_EVENTS	9.85×10^{-5}	PRKAR2A, GNAT1, ITPR3
SIG_BCR_SIGNALING_PATHWAY	1.41×10^{-4}	PPP1R13B, DAG1, ITPR3
PID_A6B1_A6B4_INTEGRIN_PATHWAY	1.41×10^{-4}	LAMB2, MST1, MST1R

Table A1. 3: MAGMA gene set analysis results (for curated gene sets i.e., MSigDB C2).

GeneSet = Canonical MSigDB pathway enriched for MCP-associated genes. p value = p value for MAGMA gene set analysis test. Genes = MCP-associated genes.

Appendix 2: Phenotypic Correlation between Multisite Chronic Pain and Chronic Pain Grade in Generation Scotland

Cheverud's conjecture posits that phenotypic and genetic correlations are likely to be similar in both direction and size, and phenotypic correlations can therefore be used as proxies for genetic correlations between traits (Cheverud, 1988; Sodini et al., 2018). Evidence to support Cheverud's conjecture has been found in plants, animals, and recently in humans across a large number of traits using UK Biobank (Sodini et al., 2018).

Despite differences between the two phenotypes, MCP and CPG, such as assessment of disability due to pain in CPG and not MCP, intuitively one could expect that MCP and CPG are likely to be positively genetically and phenotypically correlated, at least to some degree. Therefore, the negative genetic correlation was described as unexpected. Potential reasons for this negative genetic correlation are discussed in 4.4.1.6, but an alternate explanation is that a negative genetic correlation may, counterintuitively, be expected between CPG and MCP, if there were a negative phenotypic correlation between these phenotypes. To investigate this possibility, phenotypic correlations were calculated between MCP and CPG in a cohort where both phenotypes can be derived (Generation Scotland).

Both Pearson's rho and Kendall's tau were calculated, and p value significance thresholds Bonferroni adjusted (0.05/2). CPG can be treated as a continuous variable, and indeed was in previous PRS analyses in this thesis, and in 23andMe-Pfizer GWAS, but is technically an ordinal construct hence calculation of Kendall's tau as sensitivity analysis. N = 7, 574 GS participants had complete phenotype data for both CPG and MCP (mean age 50.9 years, 37% male), and their data were used in this analysis. Defining CPG and the Generation Scotland version of MCP is described in detail in 2.3.3.2.

Test	Correlation coefficient value (CI)	p value	Type
Pearson	0.317 (0.296-0.337)	$< 2.2 \times 10^{-16}$	Parametric
Kendall	0.246 (NA)	$< 2.2 \times 10^{-16}$	Non-Parametric

Table A2. 1: Phenotypic correlation between CPG and MCP.

CI = confidence interval (where applicable).

Results suggest genetic correlation should be expected to be positive between CPG and MCP – as discussed in 4.4.1.6, it is therefore likely that the covariates adjusted for in the 23andMe-Pfizer GWAS may be involved in the unexpectedly negative genetic correlation between the two phenotypes. i.e., the "true" underlying relationship is that, on average, allelic effects for pleiotropic variants are in the same direction in CPG and MCP but adjusting for manual labour in particular in the 23andMe-Pfizer GWAS has obscured that in subsequent genetic correlation analyses.

Aschard et al also provide derivation of a Wald test for specific use to test for bias in GWAS caused by adjustment for covariates – as there is not an unadjusted GWAS of CPG available for comparison, these analyses could not be carried out, but would be of interest for formally testing whether adjustment for manual labour led to bias in the CPG GWAS output.

Appendix 3: Genetic Correlation between Tsepilov et al Phenotype GIP1 (Genetically Independent Phenotype 1) and Multisite Chronic Pain

Tsepilov et al generated several genetically independent phenotypes (GIPs) related to musculoskeletal chronic pain at four bodily locations in UK Biobank (Tsepilov et al., 2020). GIP1 represents the "leading" GIP and explains 78.4% of the genetic variance in the musculoskeletal chronic pain traits and is genetically correlated with many psychiatric traits to a similar degree as MCP. GIP1 is also described as the most stable and most heritable of the GIPs derived and investigated. GIP1 shows enrichment with multiple nervous-system related terms, and GWAS of GIP1 indicates some overlapping genes that were also found to be associated with MCP.

Tsepilov et al argue that a GWAS of MCP relies on the assumption that there is equivalence between genetic predictors of musculoskeletal pain conditions and non-musculoskeletal pain conditions, an assumption that may be too strong. However, similarities between the trait construct GIP1, explaining the majority of genetic variance in the examined musculoskeletal pain traits in UKB, and MCP, indicates that it may be an acceptable assumption that genetic predictors are shared between diverse types of pain conditions. It is also of note that the first GIP for a wider range of transformed pain traits (i.e., all but one of the pain site options in UKB, some which are not likely to be musculoskeletal e.g., stomach/ abdominal pain) is almost genetically equivalent to 'musculoskeletal' GIP1 r_g = 0.99.

To investigate the genetic overlap between MCP and GIP1, LDSR was carried out using GIP1 GWAS summary statistics downloaded from https://zenodo.org/record/3797553 [13/12/2020] and the summary statistics from the MCP GWAS (Chapter 4)

Trait 1	Trait 2	r_g	se	z	p
MCP	GIP1	0.9753	0.003	327.0843	<<0.001

Table A3. 1: Genetic correlation results.

r_g = genetic correlation coefficient, se = standard error of genetic correlation coefficient.

Results indicate that GIP1 and MCP are highly genetically correlated (r_g = 0.98). This is significantly lower than r_g = 1 (0.98 + 2 x SE < 1), but as discussed previously (4.4.1.5) genetic correlation values of this magnitude can indicate that these phenotypes are measuring the same underlying trait construct. Differences between MCP and GIP1 in terms of associated genes and other downstream results in the GIP1 GWAS analysis could therefore be due to differences in power (N for the Tsepilov et al discovery cohort is roughly 100,000 fewer participants than the MCP GWAS sample size). Although ~20% of genetic variance in musculoskeletal traits is not explained by GIP1, the large genetic overlap between MCP and GIP1 indicates that genetic predictors of a biopsychological component to chronic pain may be shared across a diverse set of chronic pain conditions. Furthermore, this remaining proportion of genetic variance not attributed to GIP1 may in fact be related to disease and tissue-specific elements of chronic pain conditions, rather than being informative on the development of chronic pain itself.

References

AAPA statement on biological aspects of race. (1996). In *American Journal of Physical Anthropology* (Vol. 101, Issue 4, pp. 569–570). https://doi.org/10.1002/ajpa.1331010408

Aaron, R. V., Fisher, E. A., De La Vega, R., Lumley, M. A., & Palermo, T. M. (2019). Alexithymia in individuals with chronic pain and its relation to pain intensity, physical interference, depression, and anxiety: A systematic review and meta-analysis. *Pain*, *160*(5), 994–1006. https://doi.org/10.1097/j.pain.0000000000001487

Ables, J. L., Breunig, J. J., Eisch, A. J., & Rakic, P. (2011). Not(ch) just development: Notch signalling in the adult brain. *Nature Reviews Neuroscience*, *12*(5), 269–283. https://doi.org/10.1038/nrn3024

Adams Waldorf, K. M., & Nelson, J. L. (2008). Autoimmune disease during pregnancy and the microchimerism legacy of pregnancy. *Immunological Investigations*, *37*(5–6), 631–644. https://doi.org/10.1080/08820130802205886

Aguet, F., Brown, A. A., Castel, S. E., Davis, J. R., He, Y., Jo, B., Mohammadi, P., Park, Y. S., Parsana, P., Segrè, A. V., Strober, B. J., Zappala, Z., Cummings, B. B., Gelfand, E. T., Hadley, K., Huang, K. H., Lek, M., Li, X., Nedzel, J. L., … Zhu, J. (2017). Genetic effects on gene expression across human tissues. *Nature*, *550*(7675), 204–213. https://doi.org/10.1038/nature24277

Akbar, M. A., Mandraju, R., Tracy, C., Hu, W., Pasare, C., & Krämer, H. (2016). ARC Syndrome-linked Vps33B protein is required for inflammatory endosomal maturation and signal termination. *Immunity*, *45*(2), 267–279. https://doi.org/doi:10.1016/j.immuni.2016.07.010

Akhtar, E., Ballew, A. T., Orr, W. N., Mayorga, A., & Khan, T. W. (2019). The Prevalence of Post-Traumatic Stress Disorder Symptoms in Chronic Pain Patients in a Tertiary Care Setting: A Cross-Sectional Study. *Psychosomatics*, *60*(3), 255–262. https://doi.org/10.1016/j.psym.2018.07.012

Alberts B, Johnson A, Lewis J, et al. (2002). General Recombination. In *Molecular Biology of the Cell* (4th ed.). Garland Science. https://www.ncbi.nlm.nih.gov/books/NBK26898/

Alderson, T. R. R., Kim, J. H. H., & Markley, J. L. L. (2016). Dynamical Structures of Hsp70 and Hsp70-Hsp40 Complexes. *Structure*, *24*(7), 1014–1030. https://doi.org/10.1016/j.str.2016.05.011

Alföldi, P., Wiklund, T., & Gerdle, B. (2014). Comorbid insomnia in patients with chronic pain: A study based on the Swedish quality registry for pain rehabilitation (SQRP). *Disability and Rehabilitation*, *36*(20), 1661–1669. https://doi.org/10.3109/09638288.2013.864712

Alhalal, E., Ford-Gilboe, M., Wong, C., & Albuhairan, F. (2018). Factors mediating the impacts of child abuse and intimate partner violence on chronic pain: A cross-sectional study 11 Medical and Health Sciences 1117 Public Health and Health Services 17 Psychology and Cognitive Sciences 1701 Psychology. *BMC Women's Health*, *18*(1), 1–15. https://doi.org/10.1186/s12905-018-0642-9

Altman, D. G., & Bland, J. M. (2011). Statistics notes: How to obtain the P value from a confidence interval. *BMJ (Online)*, *343*(7825), 1–2. https://doi.org/10.1136/bmj.d2304

Altshuler, D. L., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Collins, F. S., De La Vega, F. M., Donnelly, P., Egholm, M., Flicek, P., Gabriel, S. B., Gibbs, R. A., Knoppers, B. M., Lander, E. S., Lehrach, H., Mardis, E. R., McVean, G. A., Nickerson, D. A., … Peterson, J. L. (2010). A map of human genome variation from population-scale sequencing. *Nature*, *467*(7319), 1061–1073. https://doi.org/10.1038/nature09534

Ambrosio, A. L., & Di Pietro, S. M. (2019). Mechanism of platelet a-granule biogenesis: Study of

cargo transport and the VPS33B-VPS16B complex in a model system. *Blood Advances*, *3*(17), 2617–2626. https://doi.org/10.1182/bloodadvances.2018028969

Anagianni, S., & Tuschl, K. (2019). Genetic Disorders of Manganese Metabolism. *Current Neurology and Neuroscience Reports*, *19*(6). https://doi.org/10.1007/s11910-019-0942-y

Anderson, L. R., Owens, T. W., & Naylor, M. J. (2014). Structural and mechanical functions of integrins. *Biophysical Reviews*, *6*(2), 203–213. https://doi.org/10.1007/s12551-013-0124-0

Andersson, E. R., Sandberg, R., & Lendahl, U. (2011). Notch signaling: simplicity in design, versatility in function. *Development*, *138*(17), 3593–3612. https://doi.org/10.1242/dev.063610

Andersson, H. I. (2009). Increased mortality among individuals with chronic widespread pain relates to lifestyle factors: A prospective population-based study. *Disability and Rehabilitation*, *31*(24), 1980–1987. https://doi.org/10.3109/09638280902874154

Andreassen, O. A., Djurovic, S., Thompson, W. K., Schork, A. J., Kendler, K. S., O'Donovan, M. C., Rujescu, D., Werge, T., Van De Bunt, M., Morris, A. P., McCarthy, M. I., Roddey, J. C., McEvoy, L. K., Desikan, R. S., & Dale, A. M. (2013). Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *American Journal of Human Genetics*, *92*(2), 197–209. https://doi.org/10.1016/j.ajhg.2013.01.001

Andreassen, O. A., McEvoy, L. K., Thompson, W. K., Wang, Y., Reppe, S., Schork, A. J., Zuber, V., Barrett-Connor, E., Gautvik, K., Aukrust, P., Karlsen, T. H., Djurovic, S., Desikan, R. S., & Dale, A. M. (2014). Identifying common genetic variants in blood pressure due to polygenic pleiotropy with associated phenotypes. *Hypertension*, *63*(4), 819–826. https://doi.org/10.1161/HYPERTENSIONAHA.113.02077

Andreassen, O. A., Thompson, W. K., Schork, A. J., Ripke, S., Mattingsdal, M., Kelsoe, J. R., Kendler, K. S., O'Donovan, M. C., Rujescu, D., Werge, T., Sklar, P., Roddey, J. C., Chen, C. H., McEvoy, L., Desikan, R. S., Djurovic, S., & Dale, A. M. (2013). Improved Detection of Common Variants Associated with Schizophrenia and Bipolar Disorder Using Pleiotropy-Informed Conditional False Discovery Rate. *PLoS Genetics*, *9*(4). https://doi.org/10.1371/journal.pgen.1003455

Anholt, R. R. H. (2014). Olfactomedin proteins: Central players in development and disease. *Frontiers in Cell and Developmental Biology*, *2*(FEB), 1–10. https://doi.org/10.3389/fcell.2014.00006

Anttila, V., Winsvold, B. S., Gormley, P., Kurth, T., Bettella, F., McMahon, G., Kallela, M., Malik, R., De Vries, B., Terwindt, G., Medland, S. E., Todt, U., McArdle, W. L., Quaye, L., Koiranen, M., Ikram, M., Lehtimäki, T., Stam, A. H., Ligthart, L., … Palotie, A. (2013). Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nature Genetics*, *45*(8), 912–917. https://doi.org/10.1038/ng.2676

Apkarian, A. V., Bushnell, M. C., Treede, R. D., & Zubieta, J. K. (2005). Human brain mechanisms of pain perception and regulation in health and disease. *European Journal of Pain*, *9*(4), 463. https://doi.org/10.1016/j.ejpain.2004.11.001

Apkarian, A. V., Hashmi, J. A., & Baliki, M. N. (2011). Pain and the brain: Specificity and plasticity of the brain in clinical chronic pain. *Pain*, *152*(SUPPL.3), S49–S64. https://doi.org/10.1016/j.pain.2010.11.010

Araç, D., Boucard, A. A., Özkan, E., Strop, P., Newell, E., Südhof, T. C., & Brunger, A. T. (2007). Structures of Neuroligin-1 and the Neuroligin-1/Neurexin-1β Complex Reveal Specific Protein-Protein and Protein-Ca2+ Interactions. *Neuron*, *56*(6), 992–1003. https://doi.org/10.1016/j.neuron.2007.12.002

Ardlie, K. G., DeLuca, D. S., Segrè, A. V., Sullivan, T. J., Young, T. R., Gelfand, E. T., Trowbridge, C. A.,

Maller, J. B., Tukiainen, T., Lek, M., Ward, L. D., Kheradpour, P., Iriarte, B., Meng, Y., Palmer, C. D., Esko, T., Winckler, W., Hirschhorn, J. N., Kellis, M., … Lockhart. (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*, *348*(6235), 648–660. https://doi.org/10.1126/science.1262110

Ardlie, K. G., Kruglyak, L., & Seielstad, M. (2002). Patterns of linkage disequilibrium in the human genome. *Nature Reviews Genetics*, *3*(4), 299–309. https://doi.org/10.1038/nrg777

Arnow, B. A., Hunkeler, E. M., Blasey, C. M., Lee, J., Constantino, M. J., Fireman, B., Kraemer, H. C., Dea, R., Robinson, R., & Hayward, C. (2006). Comorbid depression, chronic pain, and disability in primary care. *Psychosomatic Medicine*, *68*(2), 262–268. https://doi.org/10.1097/01.psy.0000204851.15499.fc

Aschard, H., Vilhjálmsson, B. J., Joshi, A. D., Price, A. L., & Kraft, P. (2015). Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *American Journal of Human Genetics*, *96*(2), 329–339. https://doi.org/10.1016/j.ajhg.2014.12.021

Asmundson, G. J. G., & Katz, J. (2009). Understanding the co-occurrence of anxiety disorders and chronic pain: State-of-the-art. *Depression and Anxiety*, *26*(10), 888–901. https://doi.org/10.1002/da.20600

Atkinson, E. G., Maihofer, A. X., Kanai, M., Martin, A. R., Karczewski, K. J., Santoro, M. L., Ulirsch, J. C., Kamatani, Y., Okada, Y., Finucane, H. K., Koenen, K. C., Nievergelt, C. M., Daly, M. J., & Neale, B. M. (2021). Tractor uses local ancestry to enable the inclusion of admixed individuals in GWAS and to boost power. *Nature Genetics*, *53*(February). https://doi.org/10.1038/s41588-020-00766-y

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., Lander, E. S., Lee, C., Lehrach, H., … Schloss, J. A. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. https://doi.org/10.1038/nature15393

Avenevoli, S., Swendsen, J., He, J. P., Burstein, M., & Merikangas, K. R. (2015). Major Depression in the National Comorbidity Survey–Adolescent Supplement: Prevalence, Correlates, and Treatment. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(1), 37-44.e2. https://doi.org/10.1016/j.jaac.2014.10.010

Backonja, M. M., Walk, D., Edwards, R. R., Sehgal, N., Moeller-Bertram, T., Wasan, A., Irving, G., Argoff, C., & Wallace, M. (2009). Quantitative sensory testing in measurement of neuropathic pain phenomena and other sensory abnormalities. *Clinical Journal of Pain*, *25*(7), 641–647. https://doi.org/10.1097/AJP.0b013e3181a68c7e

Baeza-Velasco, C., Cohen, D., Hamonet, C., Vlamynck, E., Diaz, L., Cravero, C., Cappe, E., & Guinchat, V. (2018). Autism, Joint Hypermobility-Related Disorders and Pain. *Frontiers in Psychiatry*, *9*(December), 1–8. https://doi.org/10.3389/fpsyt.2018.00656

Bagley, D. C., Paradkar, P. N., Kaplan, J., & Ward, D. M. (2012). Mon1a protein acts in trafficking through the secretory apparatus. *Journal of Biological Chemistry*, *287*(30), 25577–25588. https://doi.org/10.1074/jbc.M112.354043

Bair, M. J., Robinson, R. L., Katon, W., & Kroenke, K. (2003). Depression and Pain Comorbidity. *Archives of Internal Medicine*, *163*(20), 2433. https://doi.org/10.1001/archinte.163.20.2433

Bair, M. J., Wu, J., Damush, T. M., Sutherland, J. M., & Kroenke, K. (2008). Association of depression and anxiety alone and in combination with chronic musculoskeletal pain in primary care patients. *Psychosomatic Medicine*, *70*(8), 890–897. https://doi.org/10.1097/PSY.0b013e318185c510

Baliki, M. N., & Apkarian, A. V. (2015). Nociception, Pain, Negative Moods, and Behavior Selection.

Neuron, *87*(3), 474–491. https://doi.org/10.1016/j.neuron.2015.06.005

Baliki, M. N., Mansour, A. R., Baria, A. T., & Apkarian, A. V. (2014). Functional reorganization of the default mode network across chronic pain conditions. *PLoS ONE*, *9*(9). https://doi.org/10.1371/journal.pone.0106133

Baliki, M. N., Petre, B., Torbey, S., Herrmann, K. M., Huang, L., Schnitzer, T. J., Fields, H. L., & Apkarian, A. V. (2012). Corticostriatal functional connectivity predicts transition to chronic back pain. *Nature Neuroscience*, *15*(8), 1117–1119. https://doi.org/10.1038/nn.3153

Bär, K. J., de la Cruz, F., Berger, S., Schultz, C. C., & Wagner, G. (2015). Structural and functional differences in the cingulate cortex relate to disease severity in anorexia nervosa. *Journal of Psychiatry and Neuroscience*, *40*(4), 269–279. https://doi.org/10.1503/jpn.140193

Baron, Y., Pedrioli, P. G., Tyagi, K., Johnson, C., Wood, N. T., Fountaine, D., Wightman, M., & Alexandru, G. (2014). VAPB/ALS8 interacts with FFAT-like proteins including the p97 cofactor FAF1 and the ASNA1 ATPase. *BMC Biology*, *12*(1), 1–20. https://doi.org/10.1186/1741-7007-12-39

Bartley, E. J., & Fillingim, R. B. (2013). Sex differences in pain: A brief review of clinical and experimental findings. *British Journal of Anaesthesia*, *111*(1), 52–58. https://doi.org/10.1093/bja/aet127

Barton, N. H., Etheridge, A. M., & Véber, A. (2017). The infinitesimal model: Definition, derivation, and implications. *Theoretical Population Biology*, *118*, 50–73. https://doi.org/10.1016/j.tpb.2017.06.001

Bębenek, A., & Ziuzia-Graczyk, I. (2018). Fidelity of DNA replication—a matter of proofreading. *Current Genetics*, *64*(5), 985–996. https://doi.org/10.1007/s00294-018-0820-1

Becker, D. J., & Lowe, J. B. (2003). Fucose: Biosynthesis and biological function in mammals. *Glycobiology*, *13*(7). https://doi.org/10.1093/glycob/cwg054

Bell, C, & Shaw, A. (1868). Reprint of the "Idea of a New Anatomy of the Brain." *Journal of Anatomy and Physiology*, *3*(Pt 1), 147–182. http://www.ncbi.nlm.nih.gov/pubmed/17230788%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1318665

Bell, Charles. (1811). Idea of a new anatomy of the brain: submitted for the observations of his friends. In *Readings in the history of psychology.* Strahan and Preston. https://doi.org/10.1037/11304-015

Bella, J., Hindle, K. L., McEwan, P. A., & Lovell, S. C. (2008). The leucine-rich repeat structure. *Cellular and Molecular Life Sciences*, *65*(15), 2307–2333. https://doi.org/10.1007/s00018-008-8019-0

Bellefroid, E. J., Marine, J. C., Ried, T., Lecocq, P. J., Riviere, M., Amemiya, C., Poncelet, D. A., Coulie, P. G., De Jong, P., Szpirer, C., Ward, D. C., & Martial, J. A. (1993). Clustered organization of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells. *EMBO Journal*, *12*(4), 1363–1374. https://doi.org/10.1002/j.1460-2075.1993.tb05781.x

Bellis, M. A., Hughes, K., Leckenby, N., Perkins, C., & Lowey, H. (2014). National household survey of adverse childhood experiences and their relationship with resilience to health-harming behaviors in England. *BMC Medicine*, *12*(1). https://doi.org/10.1186/1741-7015-12-72

Belmaker, R. ., & Agam, G. (2008). Major depressive disorder. *The New England Journal of Medicine*, *358*, 55–68. https://doi.org/10.1007/978-1-4939-2528-5_5

Belmont, J. W., Boudreau, A., Leal, S. M., Hardenbol, P., Pasternak, S., Wheeler, D. A., Willis, T. D., Yu, F., Yang, H., Gao, Y., Hu, H., Hu, W., Li, C., Lin, W., Liu, S., Pan, H., Tang, X., Wang, J.,

Wang, W., … Stewart, J. (2005). A haplotype map of the human genome. *Nature*, *437*(7063), 1299–1320. https://doi.org/10.1038/nature04226

Benini, A., & Deleo, J. A. (1999). ´ Descartes ' Physiology of Pain. *Spine*, *24*(20), 2115–2119.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. In *Journal of the Royal Statistical Society* (Vol. 57, Issue 1, pp. 289–300). https://doi.org/10.2307/2346101

Bennett, J. A., Mastrangelo, M. A., Ture, S. K., Smith, C. O., Loelius, S. G., Berg, R. A., Shi, X., Burke, R. M., Spinelli, S. L., Cameron, S. J., Carey, T. E., Brookes, P. S., Gerszten, R. E., Sabater-Lleal, M., de Vries, P. S., Huffman, J. E., Smith, N. L., Morrell, C. N., & Lowenstein, C. J. (2020). The choline transporter Slc44a2 controls platelet activation and thrombosis by regulating mitochondrial function. *Nature Communications*, *11*(1), 1–9. https://doi.org/10.1038/s41467-020-17254-w

Bernabeu, E., Canela-xandri, O., Rawlik, K., Talenti, A., Prendergast, J., & Tenesa, A. (2021). Sex differences in genetic architecture in the UK Biobank. *Nature Genetics*, *53*(September). https://doi.org/10.1038/s41588-021-00912-0

Bevers, K., Watts, L., Kishino, N. D., & Gatchel, R. J. (2016). The Biopsychosocial model of the assessment, prevention, and treatment of chronic pain. *US Neurology*, *12*(2), 98–104. https://doi.org/10.17925/USN.2016.12.02.98

Binder, E. B. (2019). Polygenic risk scores in schizophrenia: Ready for the real world? *American Journal of Psychiatry*, *176*(10), 783–784. https://doi.org/10.1176/appi.ajp.2019.19080825

Birgenheir, D. G., Ilgen, M. A., Bohnert, A. S. B., Abraham, K. M., Bowersox, N. W., Austin, K., & Kilbourne, A. M. (2013). Pain conditions among veterans with schizophrenia or bipolar disorder. *General Hospital Psychiatry*, *35*(5), 480–484. https://doi.org/10.1016/j.genhosppsych.2013.03.019

Bischoff-Grethe, A., Wierenga, C. E., Berner, L. A., Simmons, A. N., Bailer, U., Paulus, M. P., & Kaye, W. H. (2018). Neural hypersensitivity to pleasant touch in women remitted from anorexia nervosa. *Translational Psychiatry*, *8*(1). https://doi.org/10.1038/s41398-018-0218-3

Blackburn-Munro, G. B.-M. and R. E. (2001). Review Article Chronic Pain , Chronic Stress and Depression : Coincidence or Consequence ? *Journal of Neuroendocinology*, *13*, 1009–1024. https://doi.org/10.1046/j.0007-1331.2001.00727.x

Blackburn, D. G. (2000). Why race is not a biological concept. In *Race and Racism in Theory and Practice* (Issue January 1998, pp. 3–26).

Blanchet, P. J., & Brefel-Courbon, C. (2018). Chronic pain and pain processing in Parkinson's disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *87*(October), 200–206. https://doi.org/10.1016/j.pnpbp.2017.10.010

Bliss, T. V. P., Collingridge, G. L., Kaang, B. K., & Zhuo, M. (2016). Synaptic plasticity in the anterior cingulate cortex in acute and chronic pain. *Nature Reviews Neuroscience*, *17*(8), 485–496. https://doi.org/10.1038/nrn.2016.68

Bonathan, C., Hearn, L., & Williams, A. C. de C. (2013). Socioeconomic status and the course and consequences of chronic pain. *Pain Management*, *3*(3), 159–162. https://doi.org/10.2217/pmt.13.18

Border, R., Johnson, E. C., Evans, L. M., Smolen, A., Berley, N., Sullivan, P. F., & Keller, M. C. (2019). No support for historical candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples. *American Journal of Psychiatry*, *176*(5), 376–387. https://doi.org/10.1176/appi.ajp.2018.18070881

Borsook, D. (2012). Neurological diseases and pain. *Brain*, *135*(2), 320–344.

https://doi.org/10.1093/brain/awr271

Bosco, B., Rossi, A., Rizzotto, D., Giorgetta, S., Perzolli, A., Bonollo, F., Gaucherot, A., Catez, F., Diaz, J.-J., Dassi, E., & Inga, A. (2020). DHX30 coordinates cytoplasmic translation and mitochondrial function contributing to cancer cell survival. *BioArxiv*. https://doi.org/10.1101/2020.07.13.196709

Botstein, D., & Risch, N. (2003). *Discovering genotypes underlying human phenotypes : past successes for mendelian disease , future approaches*. *33*(march). https://doi.org/10.1038/ng1090

Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *Am J Hum Gen*, *32*, 314–331. papers2://publication/uuid/0B80518E-A22B-41F3-BE43-171F51007E42

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous E. *Journal of the American Statistical Association*, *90*(430), 443–450.

Bowden, J., Del Greco M, F., Minelli, C., Davey Smith, G., Sheehan, N., & Thompson, J. (2017). A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Statistics in Medicine*, *36*(11), 1783–1802. https://doi.org/10.1002/sim.7221

Bowden, J., Del Greco M, F., Minelli, C., Zhao, Q., Lawlor, D. A., Sheehan, N. A., Thompson, J., & Davey Smith, G. (2019). Improving the accuracy of two-sample summary-data Mendelian randomization: Moving beyond the NOME assumption. *International Journal of Epidemiology*, *48*(3), 728–742. https://doi.org/10.1093/ije/dyy258

Bowden, J., Fabiola Del Greco, M., Minelli, C., Smith, G. D., Sheehan, N. A., & Thompson, J. R. (2016). Assessing the suitability of summary data for two-sample mendelian randomization analyses using MR-Egger regression: The role of the I2statistic. *International Journal of Epidemiology*, *45*(6), 1961–1974. https://doi.org/10.1093/ije/dyw220

Bowden, J., Smith, G. D., & Burgess, S. (2015). Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *International Journal of Epidemiology*, *44*(2), 512–525. https://doi.org/10.1093/ije/dyv080

Boyer, N. P., & Gupton, S. L. (2018). Revisiting netrin-1: One who guides (Axons). *Frontiers in Cellular Neuroscience*, *12*(July), 1–18. https://doi.org/10.3389/fncel.2018.00221

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*, *169*(7), 1177–1186. https://doi.org/10.1016/j.cell.2017.05.038

Brandt, D. T., Baarlink, C., Kitzing, T. M., Kremmer, E., Ivaska, J., Nollau, P., & Grosse, R. (2009). SCAI acts as a suppressor of cancer cell invasion through the transcriptional control of β1-integrin. *Nature Cell Biology*, *11*(5), 557–568. https://doi.org/10.1038/ncb1862

Breivik, H., Collett, B., Ventafridda, V., Cohen, R., & Gallacher, D. (2006). Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment. *European Journal of Pain*, *10*(4), 287–333. https://doi.org/10.1016/j.ejpain.2005.06.009

Britsch, S., Li, L., Kirchhoff, S., Theuring, F., Brinkmann, V., Birchmeier, C., & Riethmacher, D. (1998). The ErbB2 and ErbB3 receptors and their ligand, neuregulin-1, are essential for development of the sympathetic nervous system. *Genes and Development*, *12*(12), 1825–1836. https://doi.org/10.1101/gad.12.12.1825

Broberg, M., Karjalainen, J., FinnGen, & Ollila, H. M. (2021). Mendelian randomization highlights insomnia as a risk factor for pain diagnoses. *Sleep*, *44*(7), 1–8.

https://doi.org/10.1093/sleep/zsab025

Broen, M. P. G., Braaksma, M. M., Patijn, J., & Weber, W. E. J. (2012). Prevalence of pain in Parkinson's disease: A systematic review using the modified QUADAS tool. *Movement Disorders*, *27*(4), 480–484. https://doi.org/10.1002/mds.24054

Brown, D. R. (2003). *Mendelian randomisation : a new spin or real progress ? For personal use . Only reproduce with permission from The Lancet .* 930–931.

Brown, G. (1990). A Causal Analysis of Chronic Pain and Depression. *J Abnorm Psychol*, *99*(2), 127–137. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=paovfta&NEWS=N&AN=00004468-199005000-00002

Bueno, C., Lemke, C. D., Criado, G., Baroja, M. L., Ferguson, S. S. G., Rahman, A. K. M. N. U., Tsoukas, C. D., McCormick, J. K., & Madrenas, J. (2006). Bacterial Superantigens Bypass Lck-Dependent T Cell Receptor Signaling by Activating a Gα11-Dependent, PLC-β-Mediated Pathway. *Immunity*, *25*(1), 67–78. https://doi.org/10.1016/j.immuni.2006.04.012

Buhmann, C., Wrobel, N., Grashorn, W., Fruendt, O., Wesemann, K., Diedrich, S., & Bingel, U. (2017). Pain in Parkinson disease: a cross-sectional survey of its prevalence, specifics, and therapy. *Journal of Neurology*, *264*(4), 758–769. https://doi.org/10.1007/s00415-017-8426-y

Bulik-sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Case, T., Consortium, C., Duncan, L., Perry, J. R. B., Patterson, N., Robinson, E. B., Daly, M. J., Price, A. L., & Neale, B. M. (2015). An atlas of genetic correlations across human diseases and traits. *Nature Publishing Group*, *47*(11), 1236–1241. https://doi.org/10.1038/ng.3406

Bulik-Sullivan, B., Loh, P. R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., Daly, M. J., Price, A. L., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K. H., Holmans, P. A., Lee, P., Collier, D. A., Huang, H., Pers, T. H., Agartz, I., Agerbo, E., … O'Donovan, M. C. (2015). LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, *47*(3), 291–295. https://doi.org/10.1038/ng.3211

Burgess, S., Butterworth, A., & Thompson, S. G. (2013). Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic Epidemiology*, *37*(7), 658–665. https://doi.org/10.1002/gepi.21758

Burgess, S., Small, D. S., & Thompson, S. G. (2017). A review of instrumental variable estimators for Mendelian randomization. *Statistical Methods in Medical Research*, *26*(5), 2333–2355. https://doi.org/10.1177/0962280215597579

Burgess, S., Thompson, S. G., & CRP, C. G. C. (2011). Avoiding bias from weak instruments in Mendelian randomization studies. *International Journal of Epidemiology*, *40*, 755–764. https://doi.org/10.1093/ije/dyr036

Burisch, J., Jess, T., Martinato, M., & Lakatos, P. L. (2013). The burden of inflammatory bowel disease in Europe. *Journal of Crohn's and Colitis*, *7*(4), 322–337. https://doi.org/10.1016/j.crohns.2013.01.010

Burma, N. E., Leduc-Pessah, H., Fan, C. Y., & Trang, T. (2017). Animal models of chronic pain: Advances and challenges for clinical translation. *Journal of Neuroscience Research*, *95*(6), 1242–1256. https://doi.org/10.1002/jnr.23768

Burri, A., Ogata, S., Rice, D., & Williams, F. M. K. (2018). Twelve-year follow-up of chronic pain in twins: Changes in environmental and genetic influence over time. *European Journal of Pain (United Kingdom)*, *22*(8), 1439–1447. https://doi.org/10.1002/ejp.1233

Burri, A., Ogata, S., Vehof, J., & Williams, F. (2015). Chronic widespread pain: clinical comorbidities and psychological correlates. *Pain*, *156*(8), 1458–1464.

https://doi.org/10.1097/j.pain.0000000000000182

Butler, S., Landmark, T., Glette, M., Borchgrevink, P., & Woodhouse, A. (2016). Chronic widespread pain - The need for a standard definition. *Pain*, *157*(3), 541–543. https://doi.org/10.1097/j.pain.0000000000000417

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P., & Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203–209. https://doi.org/10.1038/s41586-018-0579-z

Byers, A. L., Yaffe, K., Covinsky, K. E., Friedman, M. B., & Bruce, M. L. (2010). High Occurrence of Mood and Anxiety Disorders Among Older Adults. *Archives of General Psychiatry*, *67*(5), 489. https://doi.org/10.1001/archgenpsychiatry.2010.35

Byrne, E. M., Gehrman, P. R., Medland, S. E., Nyholt, D. R., Heath, A. C., Madden, P. A. F., Hickie, I. B., Van Duijn, C. M., Henders, A. K., Montgomery, G. W., Martin, N. G., & Wray, N. R. (2013). A genome-wide association study of sleep habits and insomnia. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, *162*(5), 439–451. https://doi.org/10.1002/ajmg.b.32168

Cai, N., Bigdeli, T. B., Kretzschmar, W., Li, Y., Liang, J., Song, L., Hu, J., Li, Q., Jin, W., Hu, Z., Wang, G., Wang, L., Qian, P., Liu, Y., Jiang, T., Lu, Y., Zhang, X., Yin, Y., Li, Y., … Flint, J. (2015). Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature*, *523*(7562), 588–591. https://doi.org/10.1038/nature14659

Cai, N., Choi, K. W., & Fried, E. I. (2020). Reviewing the genetics of heterogeneity in depression: operationalizations, manifestations and etiologies. *Human Molecular Genetics*, *29*(R1), R10–R18. https://doi.org/10.1093/hmg/ddaa115

Cairns, B. E. (2010). Pathophysiology of TMD pain - basic mechanisms and their implications for pharmacotherapy. *Journal of Oral Rehabilitation*, *37*(6), 391–410. https://doi.org/10.1111/j.1365-2842.2010.02074.x

Calafell, F., Shuster, A., Speed, W. C., Kidd, J. R., & Kidd, K. K. (1998). Short tandem repeat polymorphism evolution in humans. *European Journal of Human Genetics*, *6*(1), 38–49. https://doi.org/10.1038/sj.ejhg.5200151

Camargo, A., Azuaje, F., Wang, H., & Zheng, H. (2008). Permutation - Based statistical tests for multiple hypotheses. *Source Code for Biology and Medicine*, *3*(Dcm), 1–8. https://doi.org/10.1186/1751-0473-3-15

Campos, A. I., Verweij, K. J. H., Statham, D. J., Madden, P. A. F., Maciejewski, D. F., Davis, K. A. S., John, A., Hotopf, M., Heath, A. C., Martin, N. G., & Rentería, M. E. (2020). Genetic aetiology of self-harm ideation and behaviour. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-66737-9

Carabotti, M., Scirocco, A., Antonietta, M., & Severi, C. (2015). The gut-brain axis : interactions between enteric microbiota , central and enteric nervous systems. *Ann Gastroenterol*, *28*(2), 203–209. https://doi.org/10.1038/ajgsup.2012.3

Caraceni, A., & Shkodra, M. (2019). Cancer pain assessment and classification. *Cancers*, *11*(4). https://doi.org/10.3390/cancers11040510

Carlson, C. S., Eberle, M. A., Rieder, M. J., Yi, Q., Kruglyak, L., & Nickerson, D. A. (2004). Selecting a Maximally Informative Set of Single-Nucleotide Polymorphisms for Association Analyses Using Linkage Disequilibrium. *American Journal of Human Genetics*, *74*(1), 106–120. https://doi.org/10.1086/381000

Carroll, D. (2001). Genetic Recombination. In *Brenner's Encyclopedia of Genetics: Second Edition* (pp. 277–280). https://doi.org/10.1016/B978-0-12-374984-0.00627-6

Carroll, R. J. (2006). *Measurement error in nonlinear models: A modern perspective* (2nd ed.). CRC.

Caruana, E. J., Roman, M., Hernández-Sánchez, J., & Solli, P. (2015). Longitudinal studies. *Journal of Thoracic Disease*, *7*(11), E537–E540. https://doi.org/10.3978/j.issn.2072-1439.2015.10.63

Casane, D., Boissinot, S., Chang, B. H. J., Shimmin, L. C., & Li, W. H. (1997). Mutation pattern variation among regions of the primate genome. *Journal of Molecular Evolution*, *45*(3), 216–226. https://doi.org/10.1007/PL00006223

Casanova, E. L., Baeza-velasco, C., Buchanan, C. B., & Casanova, M. F. (2020). Relationship between eds and autism. *Journal of Personalized Medicine*, *10*(260), 1–21. https://doi.org/doi:10.3390/jpm10040260

Castori, M., Tinkle, B., Levy, H., Grahame, R., Malfait, F., & Hakim, A. (2017). A framework for the classification of joint hypermobility and related conditions. *American Journal of Medical Genetics, Part C: Seminars in Medical Genetics*, *175*(1), 148–157. https://doi.org/10.1002/ajmg.c.31539

CDC. (2019). *Adverse Childhood Experiences (ACEs)*. Vital Signs. https://www.cdc.gov/vitalsigns/aces/index.html

Cerecedo, D., Cisneros, B., Gómez, P., & Galván, I. J. (2010). Distribution of dystrophin- and utrophin-associated protein complexes during activation of human neutrophils. *Experimental Hematology*, *38*(8), 618–628. https://doi.org/10.1016/j.exphem.2010.04.010

Chaichoompu, K., Abegaz, F., Cavadas, B., Fernandes, V., Müller-Myhsok, B., Pereira, L., & Van Steen, K. (2020). A different view on fine-scale population structure in Western African populations. *Human Genetics*, *139*(1), 45–59. https://doi.org/10.1007/s00439-019-02069-7

Chaki, M., Airik, R., Ghosh, A. K., Giles, R. H., Chen, R., Slaats, G. G., Wang, H., Hurd, T. W., Zhou, W., Cluckey, A., Gee, H. Y., Ramaswami, G., Hong, C. J., Hamilton, B. A., Červenka, I., Ganji, R. S., Bryja, V., Arts, H. H., Van Reeuwijk, J., … Hildebrandt, F. (2012). Exome capture reveals ZNF423 and CEP164 mutations, linking renal ciliopathies to DNA damage response signaling. *Cell*, *150*(3), 533–548. https://doi.org/10.1016/j.cell.2012.06.028

Chan, K. T., Cortesio, C. L., & Huttenlocher, A. (2009). Fak alters invadopodia and focal adhesion composition and dynamics to regulate breast cancer invasion. *Journal of Cell Biology*, *185*(2), 357–370. https://doi.org/10.1083/jcb.200809110

Chang, C., McDonnell, P., & Gershwin, M. E. (2019). Complex regional pain syndrome – False hopes and miscommunications. *Autoimmunity Reviews*, *18*(3), 270–278. https://doi.org/10.1016/j.autrev.2018.10.003

Chang, P. C., Pollema-Mays, S. L., Centeno, M. V., Procissi, D., Contini, M., Baria, A. T., Martina, M., & Apkarian, A. V. (2014). Role of nucleus accumbens in neuropathic pain: Linked multi-scale evidence in the rat transitioning to neuropathic pain. *Pain*, *155*(6), 1128–1139. https://doi.org/10.1016/j.pain.2014.02.019

Chapman, K. A., Ostrovsky, J., Rao, M., Dingley, S. D., Polyak, E., Yudkoff, M., Xiao, R., Bennett, M. J., & Falk, M. J. (2018). Propionyl-CoA carboxylase pcca-1 and pccb-1 gene deletions in Caenorhabditis elegans globally impair mitochondrial energy metabolism. *Journal of Inherited Metabolic Disease*, *41*(2), 157–168. https://doi.org/10.1007/s10545-017-0111-x

Charlesworth, B. (2009). Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*(3), 195–205. https://doi.org/10.1038/nrg2526

Chatterjee, N., Shi, J., & García-Closas, M. (2016). Developing and evaluating polygenic risk

prediction models for stratified disease prevention. *Nature Reviews Genetics*, *17*(7), 392–406. https://doi.org/10.1038/nrg.2016.27

Cheatle, M. D., Foster, S., Pinkett, A., Lesneski, M., Qu, D., & Dhingra, L. (2016). Assessing and Managing Sleep Disturbance in Patients with Chronic Pain. *Anesthesiology Clinics*, *34*(2), 379–393. https://doi.org/10.1016/j.anclin.2016.01.007

Cheatle, M. D., & Gallagher, R. M. (2006). Chronic pain and comorbid mood and substance use disorders: A biopsychosocial treatment approach. *Current Psychiatry Reports*, *8*(5), 371–376. https://doi.org/10.1007/s11920-006-0038-7

Chen, C., Zhang, C., Cheng, L., Reilly, J. L., Bishop, J. R., Sweeney, J. A., Chen, H. Y., Gershon, E. S., & Liu, C. (2014). Correlation between DNA methylation and gene expression in the brains of patients with bipolar disorder and schizophrenia. *Bipolar Disorders*, *16*(8), 790–799. https://doi.org/10.1111/bdi.12255

Chen, T. L., Zhou, L., Yuan, Y., Fang, Y., Guo, Y., Huang, H. Z., Zhou, Q., & Lv, X. Y. (2014). Characterization of Bbx, a member of a novel subfamily of the HMG-box superfamily together with Cic. *Development Genes and Evolution*, *224*(4–6), 261–268. https://doi.org/10.1007/s00427-014-0476-x

Chen, Y. C., Auer-Grumbach, M., Matsukawa, S., Zitzelsberger, M., Themistocleous, A. C., Strom, T. M., Samara, C., Moore, A. W., Cho, L. T. Y., Young, G. T., Weiss, C., Schabhüttl, M., Stucka, R., Schmid, A. B., Parman, Y., Graul-Neumann, L., Heinritz, W., Passarge, E., Watson, R. M., … Senderek, J. (2015). Transcriptional regulator PRDM12 is essential for human pain perception. *Nature Genetics*, *47*(7), 803–808. https://doi.org/10.1038/ng.3308

Cheng, A., & Saltiel, A. R. (2006). More TORC for the gluconeogenic engine. *BioEssays*, *28*(3), 231–234. https://doi.org/10.1002/bies.20375

Chesney, E., Goodwin, G. M., & Fazel, S. (2014). Risks of all-cause and suicide mortality in mental disorders: A meta-review. *World Psychiatry*, *13*(2), 153–160. https://doi.org/10.1002/wps.20128

Cheverud, J. M. (1988). A Comparison of Genetic and Phenotypic Correlations. *Evolution*, *42*(5), 958–968.

Chiarotto, A., Maxwell, L. J., Ostelo, R. W., Boers, M., Tugwell, P., & Terwee, C. B. (2019). Measurement Properties of Visual Analogue Scale, Numeric Rating Scale, and Pain Severity Subscale of the Brief Pain Inventory in Patients With Low Back Pain: A Systematic Review. *Journal of Pain*, *20*(3), 245–263. https://doi.org/10.1016/j.jpain.2018.07.009

Chiry, O., Fishbein, W. N., Merezhinskaya, N., Clarke, S., Galuske, R., Magistretti, P. J., & Pellerin, L. (2008). Distribution of the monocarboxylate transporter MCT2 in human cerebral cortex: An immunohistochemical study. *Brain Research*, *1226*, 61–69. https://doi.org/10.1016/j.brainres.2008.06.025

Choi, S. W., Mak, T. S. H., & O'Reilly, P. F. (2020). Tutorial: a guide to performing polygenic risk score analyses. *Nature Protocols*, *15*(9), 2759–2772. https://doi.org/10.1038/s41596-020-0353-1

Choi, Y., Nam, J., Whitcomb, D. J., Song, Y. S., Kim, D., Jeon, S., Um, J. W., Lee, S. G., Woo, J., Kwon, S. K., Li, Y., Mah, W., Kim, H. M., Ko, J., Cho, K., & Kim, E. (2016). SALM5 trans-synaptically interacts with LAR-RPTPs in a splicing-dependent manner to regulate synapse development. *Scientific Reports*, *6*(December 2015), 1–12. https://doi.org/10.1038/srep26676

Cichon, J., Blanck, T. J. J., Gan, W. B., & Yang, G. (2017). Activation of cortical somatostatin interneurons prevents the development of neuropathic pain. *Nature Neuroscience*, *20*(8), 1122–1132. https://doi.org/10.1038/nn.4595

Cirulli, E. T., White, S., Read, R. W., Elhanan, G., Metcalf, W. J., Tanudjaja, F., Fath, D. M., Sandoval, E., Isaksson, M., Schlauch, K. A., Grzymski, J. J., Lu, J. T., & Washington, N. L. (2020). Genome-wide rare variant analysis for thousands of phenotypes in over 70,000 exomes from two cohorts. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-14288-y

Clarke, A. J., & Cooper, D. N. (2010). GWAS: heritability missing in action? *European Journal of Human Genetics : EJHG*, *18*(8), 859–861. https://doi.org/10.1038/ejhg.2010.35

Clarke, C. (2015). Autism Spectrum Disorder and Amplified Pain. *Case Reports in Psychiatry*, *2015*, 1–4. https://doi.org/10.1155/2015/930874

Clauw, D. J. (2014). Fibromyalgia A Clinical Review. *JAMA*, *311*(15), 1547–1555. https://doi.org/10.1001/jama.2014.3266

Cleeland, C. S., & Ryan, K. M. (1994). Pain assessment: Global use of the Brief Pain Inventory. *Annals of the Academy of Medicine, Singapore*, *23*(2), 129–138.

Coleman, J. R. I., Euesden, J., Patel, H., Folarin, A. A., Newhouse, S., & Breen, G. (2016). Quality control, imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray. *Briefings in Functional Genomics*, *15*(4), 298–304. https://doi.org/10.1093/bfgp/elv037

Colloca, L., Ludman, T., Bouhassira, D., Baron, R., Dickenson, A. H., Yarnitsky, D., Freeman, R., Truini, A., Attal, N., Finnerup, N. B., Eccleston, C., Kalso, E., Bennett, D. L. H., Dworkin, R. H., & Raja, S. N. (2017). Neuropathic Pain. *Nat Rev Dis Primers*, *3*(Imi), 1–45. https://doi.org/10.1038/nrdp.2017.2.Neuropathic

Conneely, K. N., & Boehnke, M. (2007). So many correlated tests, so little time! Rapid adjustment of P values for multiple correlated tests. *American Journal of Human Genetics*, *81*(6), 1158–1168. https://doi.org/10.1086/522036

Convertini, P., Todisco, S., De Santis, F., Pappalardo, I., Iacobazzi, D., Morelli, M. A. C., Fondufe-Mittendorf, Y. N., Martelli, G., Palmieri, F., & Infantino, V. (2019). Transcriptional regulation factors of the human mitochondrial aspartate/glutamate carrier gene, isoform 2 (SLC25A13): USF1 as basal factor and FOXA2 as activator in liver cells. *International Journal of Molecular Sciences*, *20*(8), 1–16. https://doi.org/10.3390/ijms20081888

Córdova-Palomera, A., Tornador, C., Falcón, C., Bargalló, N., Brambilla, P., Crespo-Facorro, B., Deco, G., & Faanás, L. (2016). Environmental factors linked to depression vulnerability are associated with altered cerebellar resting-state synchronization. *Scientific Reports*, *6*(March), 1–11. https://doi.org/10.1038/srep37384

Craft, R. M., Mogil, J. S., & Maria Aloisi, A. (2004). Sex differences in pain and analgesia: The role of gonadal hormones. *European Journal of Pain*, *8*(5), 397–411. https://doi.org/10.1016/j.ejpain.2004.01.003

Cramer, K. S., & Miko, I. J. (2016). Eph-ephrin signaling in nervous system development. *F1000Research*, *5*, 1–8. https://doi.org/10.12688/f1000research.7417.1

Craner, J. R., Lake, E. S., Bancroft, K. E., & Hanson, K. M. (2020). Partner abuse among treatment-seeking individuals with chronic pain: Prevalence, characteristics, and association with pain-related outcomes. *Pain Medicine (United States)*, *21*(11), 2789–2798. https://doi.org/10.1093/PM/PNAA126

Cronshaw, J. M., Krutchinsky, A. N., Zhang, W., Chait, B. T., & Matunis, M. L. J. (2002). Proteomic analysis of the mammalian nuclear pore complex. *Journal of Cell Biology*, *158*(5), 915–927. https://doi.org/10.1083/jcb.200206106

Crook, R. J., Dickson, K., Hanlon, R. T., & Walters, E. T. (2014). Nociceptive sensitization reduces predation risk. *Current Biology*, *24*(10), 1121–1125.

https://doi.org/10.1016/j.cub.2014.03.043

Cruccu, G., Sommer, C., Anand, P., Attal, N., Baron, R., Garcia-Larrea, L., Haanpaa, M., Jensen, T. S., Serra, J., & Treede, R. D. (2010). EFNS guidelines on neuropathic pain assessment: Revised 2009. *European Journal of Neurology*, *17*(8), 1010–1018. https://doi.org/10.1111/j.1468-1331.2010.02969.x

Cruccu, Giorgio, & Truini, A. (2009). Tools for Assessing Neuropathic Pain. *PLoS Medicine*, *6*(4), 2–6. https://doi.org/10.1371/journal.pmed.1000047

Cryan, J. F., & Dinan, T. G. (2012). Mind-altering microorganisms: The impact of the gut microbiota on brain and behaviour. *Nature Reviews Neuroscience*, *13*(10), 701–712. https://doi.org/10.1038/nrn3346

Csecs, J. L. L., Iodice, V., Rae, C. L., Brooke, A., Dowell, N. G., Prowse, F., Themelis, K., Hugo, D., & Eccles, J. A. (2020). Increased rate of joint hypermobility in autism and related neurodevelopmental conditions is linked to dysautonomia and pain. *MedRxiv*.

Csupak, B., Sommer, J. L., Jacobsohn, E., & El-Gabalawy, R. (2018). A population-based examination of the co-occurrence and functional correlates of chronic pain and generalized anxiety disorder. *Journal of Anxiety Disorders*, *56*(February), 74–80. https://doi.org/10.1016/j.janxdis.2018.04.005

Currie, S. R., & Wang, J. (2005). More data on major depression as an antecedent risk factor for first onset of chronic back pain. *Psychological Medicine*, *35*(9), 1275–1282. https://doi.org/10.1017/S0033291705004952

Da Silva, J. A. P., Geenen, R., & Jacobs, J. W. G. (2018). Chronic widespread pain and increased mortality: Biopsychosocial interconnections. *Annals of the Rheumatic Diseases*, *77*(6), 790–792. https://doi.org/10.1136/annrheumdis-2017-211893

Dabby, R. (2012). Pain disorders and erythromelalgia caused by voltage-gated sodium channel mutations. *Current Neurology and Neuroscience Reports*, *12*(1), 76–83. https://doi.org/10.1007/s11910-011-0233-8

Danckwardt, S., Hartmann, K., Gehring, N. H., Hentze, M. W., & Kulozik, A. E. (2006). 3' End processing of the prothrombin mRNA in thrombophilia. *Acta Haematologica*, *115*(3–4), 192–197. https://doi.org/10.1159/000090934

Dansie, E. J., & Turk, D. C. (2013). *Assessment of patients with chronic pain*. *111*(1), 19–25. https://doi.org/10.1093/bja/aet124

Darvasi, A., & Shifman, S. (2005). The beauty of admixture. *Nature Genetics*, *37*(2), 118–119. https://doi.org/10.1038/ng0205-118

Davis, K. A. S., Coleman, J. R. I., Adams, M., Allen, N., Breen, G., Cullen, B., Dickens, C., Fox, E., Graham, N., Holliday, J., Howard, L. M., John, A., Lee, W., McCabe, R., McIntosh, A., Pearsall, R., Smith, D. J., Sudlow, C., Ward, J., … Hotopf, M. (2020). Mental health in UK Biobank – development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. *BJPsych Open*, *6*(2), 1–8. https://doi.org/10.1192/bjo.2019.100

De Berardis, D., Fornaro, M., Orsolini, L., Valchera, A., Carano, A., Vellante, F., Perna, G., Serafini, G., Gonda, X., Pompili, M., Martinotti, G., & Di Giannantonio, M. (2017). Alexithymia and suicide risk in psychiatric disorders: A mini-review. *Frontiers in Psychiatry*, *8*(AUG), 148. https://doi.org/10.3389/fpsyt.2017.00148

De Bruijn, D. R. H., Van Dijk, A. H. A., Pfundt, R., Hoischen, A., Merkx, G. F. M., Gradek, G. A., Lybæk, H., Stray-Pedersen, A., Brunner, H. G., & Houge, G. (2010). Severe progressive autism associated with two de novo changes: A 2.6-Mb 2q31.1 deletion and a balanced

t(14;21)(q21.1;p11.2) translocation with long-range epigenetic silencing of LRFN5 expression. *Molecular Syndromology*, *1*(1), 46–57. https://doi.org/10.1159/000280290

De C Williams, A. C. (2019). Persistence of pain in humans and other mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1785). https://doi.org/10.1098/rstb.2019.0276

De Hert, M., Correll, C. U., Bobes, J., Cetkovich-Bakmas, M., Cohen, D. A. N., Asai, I., Detraux, J., Gautam, S., Möller, H. J., Ndetei, D. M., Newcomer, J. W., Uwakwe, R., & Leucht, S. (2011). Physical illness in patients with severe mental disorders. I. Prevalence, impact of medications and disparities in health care. *World Psychiatry*, *10*(1), 52–77. https://doi.org/10.1002/j.2051-5545.2011.tb00014.x

de Leeuw, C. A., Mooij, J. M., Heskes, T., & Posthuma, D. (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Computational Biology*, *11*(4), 1–19. https://doi.org/10.1371/journal.pcbi.1004219

De Leeuw, C. A., Neale, B. M., Heskes, T., & Posthuma, D. (2016). The statistical properties of gene-set analysis. *Nature Reviews Genetics*, *17*(6), 353–364. https://doi.org/10.1038/nrg.2016.29

De Quervain, D. J. F., & Papassotiropoulos, A. (2006). Identification of a genetic cluster influencing memory performance and hippocampal activity in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(11), 4270–4274. https://doi.org/10.1073/pnas.0510212103

De Rubeis, S., He, X., Goldberg, A. P., Poultney, C. S., Samocha, K., Cicek, A. E., Kou, Y., Liu, L., Fromer, M., Walker, S., Singh, T., Klei, L., Kosmicki, J., Fu, S. C., Aleksic, B., Biscaldi, M., Bolton, P. F., Brownfeld, J. M., Cai, J., … Buxbaum, J. D. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*, *515*(7526), 209–215. https://doi.org/10.1038/nature13772

Dean, C. B., & Nielsen, J. D. (2007). Generalized linear mixed models: A review and some extensions. *Lifetime Data Analysis*, *13*(4), 497–512. https://doi.org/10.1007/s10985-007-9065-x

Demange, P. A., Malanchini, M., Mallard, T. T., Biroli, P., Cox, S. R., Grotzinger, A. D., Tucker-Drob, E. M., Abdellaoui, A., Arseneault, L., van Bergen, E., Boomsma, D. I., Caspi, A., Corcoran, D. L., Domingue, B. W., Harris, K. M., Ip, H. F., Mitchell, C., Moffitt, T. E., Poulton, R., … Nivard, M. G. (2021). Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nature Genetics*, *53*(1), 35–44. https://doi.org/10.1038/s41588-020-00754-2

Denk, F., & Mcmahon, S. B. (2017). Neurobiological basis for pain vulnerability : why me ? *Pain*, *158*, 108–114.

Denk, F., McMahon, S. B., & Tracey, I. (2014). Pain vulnerability: A neurobiological perspective. *Nature Neuroscience*, *17*(2), 192–200. https://doi.org/10.1038/nn.3628

Desmeules, J. A., Cedraschi, C., Rapiti, E., Baumgartner, E., Finckh, A., Cohen, P., Dayer, P., & Vischer, T. L. (2003). Neurophysiologic evidence for a central sensitization in patients with fibromyalgia. *Arthritis and Rheumatism*, *48*(5), 1420–1429. https://doi.org/10.1002/art.10893

Devlin, B., Bacanu, S. A., & Roeder, K. (2004). Genomic control to the extreme. *Nature Genetics*, *36*(11), 1129–1131. https://doi.org/10.1038/ng1104-1129

Devlin, B., & Risch, N. (1995). Linkage Disequilibrium Measures for Fine-Scale Mapping : A Comparison A Comparison of Linkage Disequilibrium Measures for Fine-Scale Mapping. *Genomics*, *29*, 311–322. https://doi.org/10.1159/000154430

Devlin, B., & Roeder, K. (1999). Genomic control for association studies. *Biometrics*, *55*(4), 997–1004. https://doi.org/10.1111/j.0006-341X.1999.00997.x

DeVon, H. A., Piano, M. R., Rosenfeld, A. G., & Hoppensteadt, D. A. (2014). The association of pain with protein inflammatory biomarkers: A review of the literature. *Nursing Research*, *63*(1), 51–62. https://doi.org/10.1097/NNR.0000000000000013

DeVry, C. G., & Clarke, S. (1999). Polymorphic forms of the protein L-isoaspartate (D-aspartate) O-methyltransferase involved in the repair of age-damaged proteins. *Journal of Human Genetics*, *44*(5), 275–288. https://doi.org/10.1007/s100380050161

Diatchenko, L., Slade, G. D., Nackley, A. G., Bhalang, K., Sigurdsson, A., Belfer, I., Goldman, D., Xu, K., Shabalina, S. A., Shagin, D., Max, M. B., Makarov, S. S., & Maixner, W. (2005). Genetic basis for individual variations in pain perception and the development of a chronic pain condition. *Human Molecular Genetics*, *14*(1), 135–143. https://doi.org/10.1093/hmg/ddi013

Didonna, A., Puig, E. C., Ma, Q., Matsunaga, A., Ho, B., Caillier, S. J., Shams, H., Lee, N., Hauser, S. L., Tan, Q., Zamvil, S. S., & Oksenberg, J. R. (2020). Ataxin-1 regulates B cell function and the severity of autoimmune experimental encephalomyelitis. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(38), 23742–23750. https://doi.org/10.1073/pnas.2003798117

Dieppe, P. A., & Lohmander, L. S. (2005). Pathogenesis and management of pain in osteoarthritis. *Lancet*, *365*(9463), 965–973. https://doi.org/10.1016/S0140-6736(05)71086-2

Docherty, M. J., Jones, R. C. W., & Wallace, M. S. (2011). Managing Pain in Inflammatory Bowel Disease. *Gastroenterology & Hepatology*, *7*(9), 592–601. https://doi.org/10.1103/PhysRevA.85.032324

Dominick, C. H., Blyth, F. M., & Nicholas, M. K. (2012). Unpacking the burden: Understanding the relationships between chronic pain and comorbidity in the general population. *Pain*, *153*(2), 293–304. https://doi.org/10.1016/j.pain.2011.09.018

Dorner, T. E., Muckenhuber, J., Stronegger, W. J., Rsky, É., Gustorff, B., & Freidl, W. (2011). The impact of socio-economic status on pain and the perception of disability due to pain. *European Journal of Pain*, *15*(1), 103–109. https://doi.org/10.1016/j.ejpain.2010.05.013

Douglas, J., & Scott, J. (2014). A systematic review of gender-specific rates of unipolar and bipolar disorders in community studies of pre-pubertal children. *Bipolar Disorders*, *16*(1), 5–15. https://doi.org/10.1111/bdi.12155

Driedonks, T. A. P., & Nolte-T'Hoen, E. N. M. (2019). Circulating Y-RNAs in extracellular vesicles and ribonucleoprotein complexes; Implications for the immune system. *Frontiers in Immunology*, *10*(JAN), 1–15. https://doi.org/10.3389/fimmu.2018.03164

Drulovic, J., Basic-Kes, V., Grgic, S., Vojinovic, S., Dincic, E., Toncev, G., Kezic, M. G., Kisic-Tepavcevic, D., Dujmovic, I., Mesaros, S., Miletic-Drakulic, S., & Pekmezovic, T. (2015). The Prevalence of Pain in Adults with Multiple Sclerosis: A Multicenter Cross-Sectional Survey. *Pain Medicine (United States)*, *16*(8), 1597–1602. https://doi.org/10.1111/pme.12731

Du, H., & Rosbash, M. (2002). The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing. *Nature*, *419*(6902), 86–90. https://doi.org/10.1038/nature00947

Duarri, A., Jezierska, J., Fokkens, M., Meijer, M., Schelhaas, H. J., Den Dunnen, W. F. A., Van Dijk, F., Verschuuren-Bemelmans, C., Hageman, G., Van De Vlies, P., Küsters, B., Van De Warrenburg, B. P., Kremer, B., Wijmenga, C., Sinke, R. J., Swertz, M. A., Kampinga, H. H., Boddeke, E., & Verbeek, D. S. (2012). Mutations in potassium channel KCND3 cause spinocerebellar ataxia type 19. *Annals of Neurology*, *72*(6), 870–880. https://doi.org/10.1002/ana.23700

Dudbridge, F. (2013). Power and Predictive Accuracy of Polygenic Risk Scores. *PLoS Genetics*, *9*(3).

https://doi.org/10.1371/journal.pgen.1003348

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., Frietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B. R., Landt, S. G., Lee, B. K., Pauli, F., Rosenbloom, K. R., Sabo, P., Safi, A., Sanyal, A., … Lochovsky, L. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74. https://doi.org/10.1038/nature11247

Dunn, A. S., Julian, T., Formolo, L. R., Green, B. N., & Chicoine, D. R. (2011). Preliminary analysis of posttraumatic stress disorder screening within specialty clinic setting for OIF/OEF veterans seeking care for neck or back pain. *The Journal of Rehabilitation Research and Development*, *48*(5), 493. https://doi.org/10.1682/JRRD.2010.05.0104

Edwards, R. R., Dworkin, R. H., Sullivan, M. D., Turk, D. C., & Wasan, A. D. (2016). The Role of Psychosocial Processes in the Development and Maintenance of Chronic Pain. *Journal of Pain*, *17*(9), T70–T92. https://doi.org/10.1016/j.jpain.2016.01.001

Egger, M., Davey, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple , graphical test. *BMJ*, *315*, 629. https://doi.org/https://doi.org/10.1136/bmj.315.7109.629

Ehde, D. M., Gibbons, L. E., Chwastiak, L., Bombardier, C. H., Sullivan, M. D., & Kraft, G. H. (2003). *C hro nic pain in a large co mmunity sample o f perso ns with multiple sclero sis*. 605–611.

Eichwald, T., & Talbot, S. (2020). Neuro-Immunity Controls Obesity-Induced Pain. *Frontiers in Human Neuroscience*, *14*(June), 1–11. https://doi.org/10.3389/fnhum.2020.00181

El-Shormilisy, N., Strong, J., & Meredith, P. J. (2015). Associations among gender, coping patterns and functioning for individuals with chronic pain: A systematic review. *Pain Research and Management*, *20*(1), 48–55. https://doi.org/10.1155/2015/490610

El Baroudi, M., Machiels, J. P., & Schmitz, S. (2017). Expression of SESN1, UHRF1BP1, and miR-377-3p as prognostic markers in mutated TP53 squamous cell carcinoma of the head and neck. *Cancer Biology and Therapy*, *18*(10), 775–782. https://doi.org/10.1080/15384047.2017.1373212

Elman, I., & Borsook, D. (2016). Common Brain Mechanisms of Chronic Pain and Addiction. *Neuron*, *89*(1), 11–36. https://doi.org/10.1016/j.neuron.2015.11.027

Engels, G., Francke, A. L., Van Meijel, B., Douma, J. G., De Kam, H., Wesselink, W., Houtjes, W., & Scherder, E. J. A. (2014). Clinical pain in schizophrenia: A systematic review. *Journal of Pain*, *15*(5), 457–467. https://doi.org/10.1016/j.jpain.2013.11.005

Erbe, M., Hayes, B. J., Matukumalli, L. K., Goswami, S., Bowman, P. J., Reich, C. M., Mason, B. A., & Goddard, M. E. (2012). Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *Journal of Dairy Science*, *95*(7), 4114–4129. https://doi.org/10.3168/jds.2011-5019

Euesden, J., Lewis, C. M., & Reilly, P. F. O. (2015). *Genome analysis PRSice : Polygenic Risk Score software*. *31*(December 2014), 1466–1468. https://doi.org/10.1093/bioinformatics/btu848

Evangelou, E., Warren, H. R., Mosen-Ansorena, D., Mifsud, B., Pazoki, R., Gao, H., Ntritsos, G., Dimou, N., Cabrera, C. P., Karaman, I., Ng, F. L., Evangelou, M., Witkowska, K., Tzanis, E., Hellwege, J. N., Giri, A., Velez Edwards, D. R., Sun, Y. V., Cho, K., … Caulfield, M. J. (2018). Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nature Genetics*, *50*(10), 1412–1425. https://doi.org/10.1038/s41588-018-0205-x

Evans, D. M., & Davey Smith, G. (2015). Mendelian Randomization: New Applications in the Coming Age of Hypothesis-Free Causality. *Annual Review of Genomics and Human Genetics*,

16(1), 327–350. https://doi.org/10.1146/annurev-genom-090314-050016

Evans, D. M., Gillespie, N. A., & Martin, N. G. (2002). Biometrical genetics. *Biological Psychology*, *61*(1–2), 33–51. https://doi.org/10.1016/S0301-0511(02)00051-0

Eyre-Walker, A. (2010). Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proceedings of the National Academy of Sciences*, *107*(suppl_1), 1752–1756. https://doi.org/10.1073/pnas.0906182107

F., E., & Chandan G., R. (2014). Venipuncture-induced complex regional pain syndrome: A case report and review of the literature. *Case Reports in Medicine*, *2014*, 10–13. https://doi.org/10.1155/2014/613921

Falconer, D. . (1965). The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Annals of Human Genetics*, *29*(1), 51–76. https://doi.org/10.1111/j.1469-1809.1965.tb00500.x

Falconer, D. . (1967). The inheritance of liability to diseases with variable age of onset, with particular reference to diabetes mellitus. *Annals of Human Genetics*, *31*(1), 1–20. https://doi.org/10.1111/j.1469-1809.1967.tb02015.x

Fang, J., Demic, S., & Cheng, S. (2018). The reduction of adult neurogenesis in depression impairs the retrieval of new as well as remote episodic memory. *PLoS ONE*, *13*(6), 1–23. https://doi.org/10.1371/journal.pone.0198406

Farley, R. A. (2012). Active ion transport by atp-driven ion pumps. In *Cell Physiology Source Book* (Fourth Edi, pp. 167–177). Elsevier Inc. https://doi.org/10.1016/B978-0-12-387738-3.00012-3

Fava, A., & Petri, M. (2019). Systemic lupus erythematosus: Diagnosis and clinical management. *Journal of Autoimmunity*, *96*(September 2018), 1–13. https://doi.org/10.1016/j.jaut.2018.11.001

Feichtinger, R. G., Brunner-Krainz, M., Alhaddad, B., Wortmann, S. B., Kovacs-Nagy, R., Stojakovic, T., Erwa, W., Resch, B., Windischhofer, W., Verheyen, S., Uhrig, S., Windpassinger, C., Locker, F., Makowski, C., Strom, T. M., Meitinger, T., Prokisch, H., Sperl, W., Haack, T. B., & Mayr, J. A. (2017). Combined Respiratory Chain Deficiency and UQCC2 Mutations in Neonatal Encephalomyopathy: Defective Supercomplex Assembly in Complex III Deficiencies. *Oxidative Medicine and Cellular Longevity*, *2017*. https://doi.org/10.1155/2017/7202589

Feng, B., La, J. H., Schwartz, E. S., & Gebhart, G. F. (2012). Neural and neuro-immune mechanisms of visceral hypersensitivity in irritable bowel syndrome. *American Journal of Physiology - Gastrointestinal and Liver Physiology*, *302*(10). https://doi.org/10.1152/ajpgi.00542.2011

Ferguson, A., Lyall, L. M., Ward, J., Strawbridge, R. J., Cullen, B., Graham, N., Niedzwiedz, C. L., Johnston, K. J. A., MacKay, D., Biello, S. M., Pell, J. P., Cavanagh, J., McIntosh, A. M., Doherty, A., Bailey, M. E. S., Lyall, D. M., Wyse, C. A., & Smith, D. J. (2018). Genome-Wide Association Study of Circadian Rhythmicity in 71,500 UK Biobank Participants and Polygenic Association with Mood Instability. *EBioMedicine*, *35*, 279–287. https://doi.org/10.1016/j.ebiom.2018.08.004

Ferini-Strambi, L. (2017). Neuropathic Pain and Sleep: A Review. *Pain and Therapy*, *6*(S1), 19–23. https://doi.org/10.1007/s40122-017-0089-y

Fernandez-Pujals, A. M., Adams, M. J., Thomson, P., McKechanie, A. G., Blackwood, D. H. R., Smith, B. H., Dominiczak, A. F., Morris, A. D., Matthews, K., Campbell, A., Linksted, P., Haley, C. S., Deary, I. J., Porteous, D. J., MacIntyre, D. J., & McIntosh, A. M. (2015). Epidemiology and heritability of major depressive disorder, stratified by age of onset, sex, and illness course in generation Scotland: Scottish family health study (GS: SFHS). *PLoS ONE*, *10*(11), 1–18. https://doi.org/10.1371/journal.pone.0142197

Ferrari, A. J., Somerville, A. J., Baxter, A. J., Norman, R., Patten, S. B., Vos, T., & Whiteford, H. A. (2013). Global variation in the prevalence and incidence of major depressive disorder: A systematic review of the epidemiological literature. *Psychological Medicine*, *43*(3), 471–481. https://doi.org/10.1017/S0033291712001511

Ferraro, D., Plantone, D., Morselli, F., Dallari, G., Simone, A. M., Vitetta, F., Sola, P., Primiano, G., Nociti, V., Pardini, M., Mirabella, M., & Vollono, C. (2018). Systematic assessment and characterization of chronic pain in multiple sclerosis patients. *Neurological Sciences*, *39*(3), 445–453. https://doi.org/10.1007/s10072-017-3217-x

Ferreira, K. D. S., & Speciali, J. G. (2015). Epidemiology of chronic pain in the office of a pain specialist neurologist. *Arquivos de Neuro-Psiquiatria*, *73*(7), 582–585. https://doi.org/10.1590/0004-282X20150064

Ferreira, P. H., Beckenkamp, P., Maher, C. G., Hopper, J. L., & Ferreira, M. L. (2013). Nature or nurture in low back pain? Results of a systematic review of studies based on twin samples. *European Journal of Pain (United Kingdom)*, *17*(7), 957–971. https://doi.org/10.1002/j.1532-2149.2012.00277.x

Ferris, C. F., Nodine, S., Pottala, T., Cai, X., Knox, T. M., Fofana, F. H., Kim, S., Kulkarni, P., Crystal, J. D., & Hohmann, A. G. (2019). Alterations in brain neurocircuitry following treatment with the chemotherapeutic agent paclitaxel in rats. *Neurobiology of Pain*, *6*, 100034. https://doi.org/10.1016/j.ynpai.2019.100034

Fertleman, C. R., Baker, M. D., Parker, K. A., Moffatt, S., Elmslie, F. V., Abrahamsen, B., Ostman, J., Klugbauer, N., Wood, J. N., Gardiner, R. M., & Rees, M. (2006). SCN9A Mutations in Paroxysmal Extreme Pain Disorder: Allelic Variants Underlie Distinct Channel Defects and Phenotypes. *Neuron*, *52*(5), 767–774. https://doi.org/10.1016/j.neuron.2006.10.006

Fillingim, R. B. (2015). Biopsychosocial contributions to sex differences in pain. *BJOG: An International Journal of Obstetrics and Gynaecology*, *122*(6), 769. https://doi.org/10.1111/1471-0528.13337

Fillingim, R. B. (2017). Individual differences in pain. *Pain*, *158*(4), S11–S18. https://doi.org/10.1097/j.pain.0000000000000775

Fillingim, R. B., King, C. D., Ribeiro-Dasilva, M. C., Rahim-Williams, B., & Riley, J. L. (2009). Sex, Gender, and Pain: A Review of Recent Clinical and Experimental Findings. *Journal of Pain*, *10*(5), 447–485. https://doi.org/10.1016/j.jpain.2008.12.001

Fillingim, R. B., Loeser, J. D., Baron, R., & Edwards, R. R. (2016). Assessment of Chronic Pain: Domains, Methods, and Mechanisms. *Journal of Pain*, *17*(9), T10–T20. https://doi.org/10.1016/j.jpain.2015.08.010

Finnerup, N. B., Haroutounian, S., Kamerman, P., Baron, R., Bennett, D. L. H., Bouhassira, D., Cruccu, G., Freeman, R., Hansson, P., Nurmikko, T., Raja, S. N., Rice, A. S. C., Serra, J., Smith, B. H., Treede, R.-D., & Jensen, T. S. (2016). Neuropathic pain: an updated grading system for research and clinical practice. *Pain*, *157*(8), 1599–1606. https://doi.org/10.1017/CBO9780511841378.007

Fischer, T. Z., & Waxman, S. G. (2010). Familial pain syndromes from mutations of the Nav1.7 sodium channel. *Annals of the New York Academy of Sciences*, *1184*, 196–207. https://doi.org/10.1111/j.1749-6632.2009.05110.x

Fishbain, D. A., Cutler, R., Rosomoff, H. L., & Rosomoff, R. S. (1997). Chronic pain-associated depression: Antecedent or consequence of chronic pain? A review. *Clinical Journal of Pain*, *13*(2), 116–137. https://doi.org/10.1097/00002508-199706000-00006

Fishbain, D. A., Pulikal, A., Lewis, J. E., & Gao, J. (2017). Chronic Pain Types Differ in Their Reported Prevalence of Post -Traumatic Stress Disorder (PTSD) and There Is Consistent Evidence That

Chronic Pain Is Associated with PTSD: An Evidence-Based Structured Systematic Review. *Pain Medicine (Malden, Mass.)*, *18*(4), 711–735. https://doi.org/10.1093/pm/pnw065

Flint, J., & Kendler, K. S. (2014). The Genetics of Major Depression. *Neuron*, *81*, 484–503. https://doi.org/10.1016/j.neuron.2014.01.027

Foley, P. L., Vesterinen, H. M., Laird, B. J., Sena, E. S., Colvin, L. A., Chandran, S., MacLeod, M. R., & Fallon, M. T. (2013). Prevalence and natural history of pain in adults with multiple sclerosis: Systematic review and meta-analysis. *Pain*, *154*(5), 632–642. https://doi.org/10.1016/j.pain.2012.12.002

Ford, B. (2012). Pain in Parkinson's disease. *Parkinson's Disease, Second Edition*, *25*(July 2009), 225–234. https://doi.org/10.1201/b12948

Forghani, I. (2019). Updates in clinical and genetics aspects of hypermobile ehlers danlos syndrome. *Balkan Medical Journal*, *36*(1), 12–16. https://doi.org/10.4274/balkanmedj.2018.1113

Forlani, C., Morri, M., Ferrari, B., Dalmonte, E., Menchetti, M., De Ronchi, D., & Atti, A. R. (2014). Prevalence and gender differences in late-life depression: a population-based study . In *The American journal of geriatric psychiatry : official journal of the American Association for Geriatric Psychiatry U6 - ctx_ver=Z39.88-2004&ctx_enc=info%3Aofi%2Fenc%3AUTF-8&rfr_id=info%3Asid%2Fsummon.serialssolutions.com&rft_val_fmt=info%3Aofi%2Ffmt%3* (Vol. 22, Issue 4, p. 370). http://glasgow.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwpV1bS-QwFA6ji76J7rrel7wvlabpZSr4IOOMy4KwoO6DL5LmMnSZaYfFxb_vOUknzQ4oii9laGhgki8n30nO-Q4hPDmNoxWbIDPGs0JUQx6boYR9g2VDYzR4SED_Y43n-5PLfDzKf95n48Fgqc4Dllc0rVVOPAU26dKrzkJp7Q9POryDacck2ndMvO8UXsB

Frank, T., Rutherford, M. A., Strenzke, N., Neef, A., Pangršič, T., Khimich, D., Fetjova, A., Gundelfinger, E. D., Liberman, M. C., Harke, B., Bryan, K. E., Lee, A., Egner, A., Riedel, D., & Moser, T. (2010). Bassoon and the synaptic ribbon organize Ca2+ channels and vesicles to add release sites and promote refilling. *Neuron*, *68*(4), 724–738. https://doi.org/10.1016/j.neuron.2010.10.027

Freynhagen, R., Parada, H. A., Calderon-Ospina, C. A., Chen, J., Rakhmawati Emril, D., Fernández-Villacorta, F. J., Franco, H., Ho, K. Y., Lara-Solares, A., Li, C. C. F., Mimenza Alvarado, A., Nimmaanrat, S., Dolma Santos, M., & Ciampi de Andrade, D. (2019). Current understanding of the mixed pain concept: a brief narrative review. *Current Medical Research and Opinion*, *35*(6), 1011–1018. https://doi.org/10.1080/03007995.2018.1552042

Freynhagen, R., Rey, R., & Argoff, C. (2020). When to consider "mixed pain"? The right questions can make a difference! *Current Medical Research and Opinion*, *36*(12), 2037–2046. https://doi.org/10.1080/03007995.2020.1832058

Fried, E. I., Epskamp, S., Nesse, R. M., Tuerlinckx, F., & Borsboom, D. (2016). What are "good" depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders*, *189*, 314–320. https://doi.org/10.1016/j.jad.2015.09.005

Fried, E. I., & Nesse, R. M. (2015a). Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders*, *172*, 96–102. https://doi.org/10.1016/j.jad.2014.10.010.Depression

Fried, E. I., & Nesse, R. M. (2015b). Depression sum-scores don't add up: Why analyzing specific depression symptoms is essential. *BMC Medicine*, *13*(1), 1–11. https://doi.org/10.1186/s12916-015-0325-4

Fromer, M., Pocklington, A. J., Kavanagh, D. H., Williams, H. J., Dwyer, S., Gormley, P., Georgieva,

L., Rees, E., Palta, P., Ruderfer, D. M., Carrera, N., Humphreys, I., Johnson, J. S., Roussos, P., Barker, D. D., Banks, E., Milanova, V., Grant, S. G., Hannon, E., … O'Donovan, M. C. (2014). De novo mutations in schizophrenia implicate synaptic networks. *Nature*, *506*(7487), 179–184. https://doi.org/10.1038/nature12929

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., & Allen, N. E. (2017a). Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with Those of the General Population. *American Journal of Epidemiology*, *186*(9). https://doi.org/10.1093/aje/kwx246

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., & Allen, N. E. (2017b). Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with Those of the General Population. *American Journal of Epidemiology*, *186*(9), 1026–1034. https://doi.org/10.1093/aje/kwx246

Furth, N., & Aylon, Y. (2017). The LATS1 and LATS2 tumor suppressors: Beyond the hippo pathway. *Cell Death and Differentiation*, *24*(9), 1488–1501. https://doi.org/10.1038/cdd.2017.99

Gabilondo, A., Alonso-Moran, E., Nuño-Solinis, R., Orueta, J. F., & Iruin, A. (2017). Comorbidities with chronic physical conditions and gender profiles of illness in schizophrenia. Results from PREST, a new health dataset. *Journal of Psychosomatic Research*, *93*, 102–109. https://doi.org/10.1016/j.jpsychores.2016.12.011

Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., Liu-Cordero, S. N., Rotimi, C., Adeyemo, A., Cooper, R., Ward, R., Lander, E. S., Daly, M. J., & Altshuler, D. (2002). The structure of haplotype blocks in the human genome. *Science*, *296*(5576), 2225–2229. https://doi.org/10.1126/science.1069424

Galor, A., Covington, D., Levitt, A. E., McManus, K. T., Seiden, B., Felix, E. R., Kalangara, J., Feuer, W., Patin, D. J., Martin, E. R., Sarantopoulos, K. D., & Levitt, R. C. (2016). Neuropathic ocular pain due to dry eye is associated with multiple comorbid chronic pain syndromes. *Journal of Pain*, *17*(3), 310–318. https://doi.org/10.1016/j.jpain.2015.10.019

Gardner, A., & Boles, R. G. (2011). Beyond the serotonin hypothesis: Mitochondria, inflammation and neurodegeneration in major depression and affective spectrum disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(3), 730–743. https://doi.org/10.1016/j.pnpbp.2010.07.030

GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*, *388*, 1545–1602. https://doi.org/10.1016/S0140-6736(16)31678-6

Ge, T., Chen, C. Y., Neale, B. M., Sabuncu, M. R., & Smoller, J. W. (2017). Phenome-wide heritability analysis of the UK Biobank. *PLoS Genetics*, *13*(4), 1–21. https://doi.org/10.1371/journal.pgen.1006711

Geber, C., Klein, T., Azad, S., Birklein, F., Gierthmühlen, J., Huge, V., Lauchart, M., Nitzsche, D., Stengel, M., Valet, M., Baron, R., Maier, C., Tölle, T., & Treede, R. D. (2011). Test-retest and interobserver reliability of quantitative sensory testing according to the protocol of the German Research Network on Neuropathic Pain (DFNS): A multi-centre study. *Pain*, *152*(3), 548–556. https://doi.org/10.1016/j.pain.2010.11.013

Gerhardt, A., Eich, W., Janke, S., Leisner, S., Treede, R. D., & Tesarz, J. (2016a). Chronic Widespread Back Pain is Distinct from Chronic Local Back Pain. *Clinical Journal of Pain*, *32*(7), 568–579. https://doi.org/10.1097/AJP.0000000000000300

Gerhardt, A., Eich, W., Janke, S., Leisner, S., Treede, R., & Tesarz, J. (2016b). Chronic Widespread

Back Pain is Distinct From Chronic Local Back Pain. *The Clinical Journal of Pain*, *32*(7), 568–579. https://doi.org/10.1097/AJP.0000000000000300

Gerrits, M. M. J. G., Van Oppen, P., Van Marwijk, H. W. J., Penninx, B. W. J. H., & Van Der Horst, H. E. (2014). Pain and the onset of depressive and anxiety disorders. *Pain*, *155*(1), 53–59. https://doi.org/10.1016/j.pain.2013.09.005

Ghigna, C., Giordano, S., Shen, H., Benvenuto, F., Castiglioni, F., Comoglio, P. M., Green, M. R., Riva, S., & Biamonti, G. (2005). Cell motility is controlled by SF2/ASF through alternative splicing of the Ron protooncogene. *Molecular Cell*, *20*(6), 881–890. https://doi.org/10.1016/j.molcel.2005.10.026

Gilks, W. P., Abbott, J. K., & Morrow, E. H. (2014). Sex differences in disease genetics: Evidence, evolution, and detection. *Trends in Genetics*, *30*(10), 453–463. https://doi.org/10.1016/j.tig.2014.08.006

Girgus, J. S., Yang, K., & Ferri, C. V. (2017). The gender difference in depression: Are elderly women at greater risk for depression than elderly men? *Geriatrics (Switzerland)*, *2*(4). https://doi.org/10.3390/geriatrics2040035

Giudicessi, J. R., Ye, D., Tester, D. J., Crotti, L., Mugione, A., Nesterenko, V. V., Albertson, R. M., Antzelevitch, C., Schwartz, P. J., & Ackerman, M. J. (2011). Transient outward current (Ito) gain-of-function mutations in the KCND3-encoded Kv4.3 potassium channel and Brugada syndrome. *Heart Rhythm*, *8*(7), 1024–1032. https://doi.org/10.1016/j.hrthm.2011.02.021

Goebel, A. (2011). Complex regional pain syndrome in adults. *Rheumatology*, *50*(10), 1739–1750. https://doi.org/10.1093/rheumatology/ker202

Goerlich, K. S. (2018). The multifaceted nature of alexithymia - A neuroscientific perspective. *Frontiers in Psychology*, *9*(AUG), 1–7. https://doi.org/10.3389/fpsyg.2018.01614

Goffer, Y., Xu, D., Eberle, S. E., D'amour, J., Lee, M., Tukey, D., Froemke, R. C., Ziff, E. B., & Wang, J. (2013). Calcium-permeable AMPA receptors in the nucleus accumbens regulate depression-like behaviors in the chronic neuropathic pain state. *Journal of Neuroscience*, *33*(48), 19034–19044. https://doi.org/10.1523/JNEUROSCI.2454-13.2013

Goldberg, D. S., & McGee, S. J. (2011). Pain as a global public health priority. *BMC Public Health*, *11*(1), 770. https://doi.org/10.1186/1471-2458-11-770

Goldstein, L. S. B., & Yang, Z. (2000). Microtubule-Based Transport Systems In Neurons: The Roles of Kinesins and Dyneins. *Annual Review of Neuroscience*, *23*, 39–71.

Gollust, S. E., Gray, S. W., Carere, D. A., Koenig, B. A., Lehmann, L. S., McGuire, A. M. Y. L., Sharp, R. R., Spector-Bagdady, K., Wang, N., Green, robert C., & Roberts, J. S. (2017). Consumer Perspectives on Access to Direct-to-Consumer Genetic Testing: Role of Demographic Factors and the Testing Experience. *Milbank Quarterly*, *95*(2), 291–318. https://doi.org/10.1111/1468-0009.12262

Golshani, A. E., Kamdar, A. A., Spence, S. C., & Beckmann, N. M. (2014). Congenital indifference to pain: An illustrated case report and literature review. *Journal of Radiology Case Reports*, *8*(8), 16–23. https://doi.org/10.3941/jrcr.v8i8.2194

Gomez-Santos, C., Hernandez-Morante, J. J., Margareto, J., Larrarte, E., Formiguera, X., Martínez, C. M., & Garaulet, M. (2011). Profile of adipose tissue gene expression in premenopausal and postmenopausal women: Site-specific differences. *Menopause*, *18*(6), 675–684. https://doi.org/10.1097/gme.0b013e31820641da

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*(6), 333–351. https://doi.org/10.1038/nrg.2016.49

Gormley, P., Anttila, V., Winsvold, B. S., Palta, P., Esko, T., Hinds, D. A., Buring, J. E., Schürks, M., Ridker, P. M., Hrafnsdottir, M. G., Artto, V., Kaunisto, M., Vepsäläinen, S., Malik, R., Heath, A. C., Madden, P. A. F., Martin, N. G., Montgomery, G. W., Kurki, M. I., … Schreiber, S. (2016). Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nature Genetics*, *48*(8), 16–18. https://doi.org/10.1038/ng.3598

Gourraud, J. B., Barc, J., Thollet, A., Le Scouarnec, S., Le Marec, H., Schott, J. J., Redon, R., & Probst, V. (2016). The Brugada Syndrome: A Rare Arrhythmia Disorder with Complex Inheritance. *Frontiers in Cardiovascular Medicine*, *3*(April), 1–11. https://doi.org/10.3389/fcvm.2016.00009

Gratten, J., & Visscher, P. M. (2016). Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. *Genome Medicine*, *8*(1), 78. https://doi.org/10.1186/s13073-016-0332-x

Greene, S. A. (2010). Chronic Pain: Pathophysiology and Treatment Implications. *Topics in Companion Animal Medicine*, *25*(1), 5–9. https://doi.org/10.1053/j.tcam.2009.10.009

Griffiths, A., Miller, J., Suzuki, D., & Al, E. (2000). Quantifying heritability. In *An Introduction to Genetic Analysis* (7th ed.). W. H. Freeman. https://www.ncbi.nlm.nih.gov/books/NBK21866/

Groenewald, C. B., Murray, C. B., & Palermo, T. M. (2020). Adverse childhood experiences and chronic pain among children and adolescents in the united states. *Pain Reports*, *5*(5), 1–7. https://doi.org/10.1097/PR9.0000000000000839

Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., Ip, H. F., Marioni, R. E., McIntosh, A. M., Deary, I. J., Koellinger, P. D., Harden, K. P., Nivard, M. G., & Tucker-Drob, E. M. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour*, *3*(5), 513–525. https://doi.org/10.1038/s41562-019-0566-x

Gu, X., Zhou, T. J., Anagnostou, E., Soorya, L., Kolevzon, A., Hof, P. R., & Fan, J. (2018). Heightened brain response to pain anticipation in high-functioning adults with autism spectrum disorder. *European Journal of Neuroscience*, *47*(6), 592–601. https://doi.org/10.1111/ejn.13598

Guerrier, S., & Polleux, F. (2007). The Ups and Downs of Neural Progenitors: Cep120 and TACCs Control Interkinetic Nuclear Migration. *Neuron*, *56*(1), 1–3. https://doi.org/10.1016/j.neuron.2007.09.019

Guo, Q., Lu, X., Gao, Y., Zhang, J., Yan, B., Su, D., Song, A., Zhao, X., & Wang, G. (2017). Cluster analysis: A new approach for identification of underlying risk factors for coronary artery disease in essential hypertensive patients. *Scientific Reports*, *7*(November 2016), 1–7. https://doi.org/10.1038/srep43965

Gureje, O. (2008). Comorbidity of pain and anxiety disorders. *Current Psychiatry Reports*, *10*(4), 318–322. https://doi.org/10.1007/s11920-008-0051-0

Gureje, O., Simon, G. E., & Von Korff, M. (2001). A cross-national study of the course of persistent pain in primary care. *Pain*, *92*(1–2), 195–200. https://doi.org/10.1016/S0304-3959(00)00483-8

Gureje, O., Von Korff, M., Kola, L., Demyttenaere, K., He, Y., Posada-Villa, J., Lepine, J. P., Angermeyer, M. C., Levinson, D., de Girolamo, G., Iwata, N., Karam, A., Luiz Guimaraes Borges, G., de Graaf, R., Browne, M. O., Stein, D. J., Haro, J. M., Bromet, E. J., Kessler, R. C., & Alonso, J. (2008). The relation between multiple pains and mental disorders: Results from the World Mental Health Surveys. *Pain*, *135*(1–2), 82–91. https://doi.org/10.1016/j.pain.2007.05.005

Gurnett, C. A., Veile, R., Zempel, J., Blackburn, L., Lovett, M., & Bowcock, A. (2008). Disruption of Sodium Bicarbonate Transporter SLC4A10 in a Patient With Complex Partial Epilepsy and

Mental Retardation. *Archives of Neurology*, *65*(4), 550–553.

Haack, M., Simpson, N., Sethna, N., Kaur, S., & Mullington, J. (2020). Sleep deficiency and chronic pain: potential underlying mechanisms and clinical implications. *Neuropsychopharmacology*, *45*(1), 205–216. https://doi.org/10.1038/s41386-019-0439-z

Haahr-Pedersen, I., Perera, C., Hyland, P., Vallières, F., Murphy, D., Hansen, M., Spitz, P., Hansen, P., & Cloitre, M. (2020). Females have more complex patterns of childhood adversity: implications for mental, social, and emotional outcomes in adulthood. *European Journal of Psychotraumatology*, *11*(1). https://doi.org/10.1080/20008198.2019.1708618

Habib, A. M., Okorokov, A. L., Hill, M. N., Bras, J. T., Lee, M. C., Li, S., Gossage, S. J., van Drimmelen, M., Morena, M., Houlden, H., Ramirez, J. D., Bennett, D. L. H., Srivastava, D., & Cox, J. J. (2019). Microdeletion in a FAAH pseudogene identified in a patient with high anandamide concentrations and pain insensitivity. *British Journal of Anaesthesia*, *123*(2), e249–e253. https://doi.org/10.1016/j.bja.2019.02.019

Hackinger, S., & Zeggini, E. (2017). *Statistical methods to detect pleiotropy in human complex traits*. https://doi.org/10.1098/rsob.170125

Hains, B. C., & Waxman, S. G. (2006). Activated microglia contribute to the maintenance of chronic pain after spinal cord injury. *Journal of Neuroscience*, *26*(16), 4308–4317. https://doi.org/10.1523/JNEUROSCI.0003-06.2006

Haliloglu, S., Carlioglu, A., Akdeniz, D., Karaaslan, Y., & Kosar, A. (2014). Fibromyalgia in patients with other rheumatic diseases: Prevalence and relationship with disease activity. *Rheumatology International*, *34*(9), 1275–1280. https://doi.org/10.1007/s00296-014-2972-8

Hall, E., Volkov, P., Dayeh, T., Esguerra, J. L. o. S., Salö, S., Eliasson, L., Rönn, T., Bacos, K., & Ling, C. (2014). Sex differences in the genome-wide DNA methylation pattern and impact on gene expression, microRNA levels and insulin secretion in human pancreatic islets. *Genome Biology*, *15*(12), 522. https://doi.org/10.1186/s13059-014-0522-z

Halladay, A. K., Bishop, S., Constantino, J. N., Daniels, A. M., Koenig, K., Palmer, K., Messinger, D., Pelphrey, K., Sanders, S. J., Singer, A. T., Taylor, J. L., & Szatmari, P. (2015). Sex and gender differences in autism spectrum disorder: Summarizing evidence gaps and identifying emerging areas of priority. *Molecular Autism*, *6*(1), 1–5. https://doi.org/10.1186/s13229-015-0019-y

Hamada, T., Wessagowit, V., South, A. P., Ashton, G. H. S., Chan, I., Oyama, N., Siriwattana, A., Jewhasuchin, P., Charuwichitratana, S., Thappa, D. M., Lenane, P., Krafchik, B., Kulthanan, K., Shimizu, H., Kaya, T. I., Erdal, M. E., Paradisi, M., Paller, A. S., Seishima, M., … McGrath, J. A. (2003). Extracellular matrix protein 1 gene (ECM1) mutations in lipoid proteinosis and genotype-phenotype correlation. *Journal of Investigative Dermatology*, *120*(3), 345–350. https://doi.org/10.1046/j.1523-1747.2003.12073.x

Hamilton, M. (1960). A Rating Scale for Depresison. *Journal of Neurology, Neurosurgery and Psychiatry*, *23*, 56. https://doi.org/10.1192/bjp.bp.107.040485

Han, B., Pouget, J. G., Slowikowski, K., Stahl, E., Lee, C. H., Diogo, D., Hu, X., Park, Y. R., Kim, E., Gregersen, P. K., Dahlqvist, S. R., Worthington, J., Martin, J., Eyre, S., Klareskog, L., Huizinga, T., Chen, W.-M., Onengut-Gumuscu, S., Rich, S. S., … Raychaudhuri, S. (2016). A method to decipher pleiotropy by detecting underlying heterogeneity driven by hidden subgroups applied to autoimmune and neuropsychiatric diseases. *Nature Genetics*, *48*(7), 803–810. https://doi.org/10.1038/ng.3572

Hansen, T. B., Wiklund, E. D., Bramsen, J. B., Villadsen, S. B., Statham, A. L., Clark, S. J., & Kjems, J. (2011). MiRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO Journal*, *30*(21), 4414–4422. https://doi.org/10.1038/emboj.2011.359

Hardt, J., Jacobsen, C., Goldberg, J., Nickel, R., & Buchwald, D. (2008). Prevalence of chronic pain in a representative sample in the United States. *Pain Medicine*, *9*(7), 803–812. https://doi.org/10.1111/j.1526-4637.2008.00425.x

Harte, S. E., Harris, R. E., & Clauw, D. J. (2018). The neurobiology of central sensitization. *Journal of Applied Biobehavioral Research*, *23*(2). https://doi.org/10.1111/jabr.12137

Hartwig, F. P., Davies, N. M., Hemani, G., & Smith, G. D. (2016). Counterfactual causation: Avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *International Journal of Epidemiology*, *45*(6), 1717–1726. https://doi.org/10.1093/ije/dyx028

Hashmi, J. A., Baliki, M. N., Huang, L., Baria, A. T., Torbey, S., Hermann, K. M., Schnitzer, T. J., & Apkarian, A. V. (2013). Shape shifting pain: Chronification of back pain shifts brain representation from nociceptive to emotional circuits. *Brain*, *136*(9), 2751–2768. https://doi.org/10.1093/brain/awt211

Hasin, D. S., Sarvet, A. L., Meyers, J. L., Saha, T. D., Ruan, W. J., Stohl, M., & Grant, B. F. (2018). Epidemiology of adult DSM-5 major depressive disorder and its specifiers in the United States. *JAMA Psychiatry*, *75*(4), 336–346. https://doi.org/10.1001/jamapsychiatry.2017.4602

Hasler, G. (2010). Pathophysiology of depression: do we have any solid evidence of interest to clinicians? *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, *9*(3), 155–161. https://doi.org/10.1002/j.2051-5545.2010.tb00298.x

Hauner, H. (2005). Secretory factors from human adipose tissue and their functional role. *Proceedings of the Nutrition Society*, *64*(2), 163–169. https://doi.org/10.1079/pns2005428

Hayashi, K., Furuya, A., Sakamaki, Y., Akagi, T., Shinoda, Y., Sadakata, T., Hashikawa, T., Shimizu, K., Minami, H., Sano, Y., Nakayama, M., & Furuichi, T. (2017). The brain-specific RasGEF very-KIND is required for normal dendritic growth in cerebellar granule cells and proper motor coordination. *PLoS ONE*, *12*(3), 1–15. https://doi.org/10.1371/journal.pone.0173175

Haycock, P. C., Burgess, S., Wade, K. H., Bowden, J., Relton, C., & Smith, G. D. (2016). *Best (but oft-forgotten) practices : the design , analysis , and interpretation of Mendelian randomization studies*. https://doi.org/10.3945/ajcn.115.118216.INTRODUCTION

Hayes, B. J., Bowman, P. J., Chamberlain, A. J., & Goddard, M. E. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science*, *92*(2), 433–443. https://doi.org/10.3168/jds.2008-1646

He, S., Wu, Z., Tian, Y., Yu, Z., Yu, J., Wang, X., Li, J., Liu, B., & Xu, Y. (2020). Structure of nucleosome-bound human BAF complex. *Science*, *367*(6480), 875–881. https://doi.org/10.1126/science.aaz9761

Heitmann, H., Haller, B., Tiemann, L., Mühlau, M., Berthele, A., Tölle, T. R., Salmen, A., Ambrosius, B., Bayas, A., Asseyer, S., Hartung, H. P., Heesen, C., Stangel, M., Wildemann, B., Haars, S., Groppa, S., Luessi, F., Kümpfel, T., Nischwitz, S., … Ploner, M. (2020). Longitudinal prevalence and determinants of pain in multiple sclerosis: results from the German National Multiple Sclerosis Cohort study. *Pain*, *161*(4), 787–796. https://doi.org/10.1097/j.pain.0000000000001767

Hemani, G., Bowden, J., & Smith, G. D. (2018). Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Human Molecular Genetics*, *27*(2), 195–208. https://doi.org/10.1093/hmg/ddy163

Hemani, G., Knott, S., & Haley, C. (2013). An Evolutionary Perspective on Epistasis and the Missing Heritability. *PLoS Genetics*, *9*(2). https://doi.org/10.1371/journal.pgen.1003295

Hendler, N. (2016). Why Chronic Pain Patients are Misdiagnosed 40 to 80% of the Time? *Journal on Recent Advances in Pain*, *2*(3), 94–98. https://doi.org/10.5005/jp-journals-10046-0051

Heng, H. H. Q., Wang, A., & Hu, J. (1998). Mapping of the human HPRP3 and HPRP4 genes encoding U4/U6-associated splicing factors to chromosomes 1q21.1 and 9q31-q33. *Genomics*, *48*(2), 273–275. https://doi.org/10.1006/geno.1997.5181

Heyer, W.-D., Ehmsen, K. T., & Liu, J. (2010). Regulation of Homologous Recombination in Eukaryotes. *Annual Review of Genetics*, *44*(1), 113–139. https://doi.org/10.1146/annurev-genet-051710-150955

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*(7414), 557–560.

Hill, W. D., Davies, N. M., Ritchie, S. J., Skene, N. G., Bryois, J., Bell, S., Di Angelantonio, E., Roberts, D. J., Xueyi, S., Davies, G., Liewald, D. C. M., Porteous, D. J., Hayward, C., Butterworth, A. S., McIntosh, A. M., Gale, C. R., & Deary, I. J. (2019). Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nature Communications*, *10*(1), 1–16. https://doi.org/10.1038/s41467-019-13585-5

Hill, W. G., Goddard, M. E., & Visscher, P. M. (2008). Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics*, *4*(2). https://doi.org/10.1371/journal.pgen.1000008

Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S., & Manolio, T. A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(23), 9362–9367. https://doi.org/10.1073/pnas.0903103106

Hocking, L. J., Morris, A. D., Dominiczak, A. F., Porteous, D. J., & Smith, B. H. (2012). Heritability of chronic pain in 2195 extended families. *European Journal of Pain (United Kingdom)*, *16*(7), 1053–1063. https://doi.org/10.1002/j.1532-2149.2011.00095.x

Höhn, A., Gampe, J., Lindahl-Jacobsen, R., Christensen, K., & Oksuyzan, A. (2020). Do men avoid seeking medical advice? A register-based analysis of gender-specific changes in primary healthcare use after first hospitalisation at ages 60+ in Denmark. *Journal of Epidemiology and Community Health*, *74*(7), 573–579. https://doi.org/10.1136/jech-2019-213435

Holbro, T., Beerli, R. R., Maurer, F., Koziczak, M., Barbas, C. F., & Hynes, N. E. (2003). The ErbB2/ErbB3 heterodimer functions as an oncogenic unit: ErbB2 requires ErbB3 to drive breast tumor cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 8933–8938. https://doi.org/10.1073/pnas.1537685100

Holmans, P., Green, E. K., Pahwa, J. S., Ferreira, M. A. R., Purcell, S. M., Sklar, P., Owen, M. J., O'Donovan, M. C., & Craddock, N. (2009). Gene Ontology Analysis of GWA Study Data Sets Provides Insights into the Biology of Bipolar Disorder. *American Journal of Human Genetics*, *85*(1), 13–24. https://doi.org/10.1016/j.ajhg.2009.05.011

Hong, E. P., & Park, J. W. (2012). Sample Size and Statistical Power Calculation in Genetic Association Studies. *Genomics and Informatics*, *10*(2), 117–122.

Horn, K. E., Glasgow, S. D., Gobert, D., Bull, S. J., Luk, T., Girgis, J., Tremblay, M. E., McEachern, D., Bouchard, J. F., Haber, M., Hamel, E., Krimpenfort, P., Murai, K. K., Berns, A., Doucet, G., Chapman, C. A., Ruthazer, E. S., & Kennedy, T. E. (2013). DCC Expression by Neurons Regulates Synaptic Plasticity in the Adult Brain. *Cell Reports*, *3*(1), 173–185. https://doi.org/10.1016/j.celrep.2012.12.005

Houlden, H., King, R., Blake, J., Groves, M., Love, S., Woodward, C., Hammans, S., Nicoll, J., Lennox, G., O'Donovan, D. G., Gabriel, C., Thomas, P. K., & Reilly, M. M. (2006). Clinical, pathological and genetic characterization of hereditary sensory and autonomic neuropathy type I (HSAN I). *Brain*, *129*(2), 411–425. https://doi.org/10.1093/brain/awh712

Howard, D. M., Adams, M. J., Clarke, T., Hafferty, J. D., Gibson, J., Shirali, M., Coleman, J. R. I.,

Hagenaars, S. P., Ward, J., Eleanor, M., Alloza, C., Shen, X., Barbu, M. C., Xu, E. Y., Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I. J., & Hemani, G. (2018). *Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions*.

Howard, D. M., Adams, M. J., Clarke, T. K., Hafferty, J. D., Gibson, J., Shirali, M., Coleman, J. R. I., Hagenaars, S. P., Ward, J., Wigmore, E. M., Alloza, C., Shen, X., Barbu, M. C., Xu, E. Y., Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I. J., … McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, *22*(3), 343–352. https://doi.org/10.1038/s41593-018-0326-7

Howard, D. M., Adams, M. J., Shirali, M., Clarke, T. K., Marioni, R. E., Davies, G., Coleman, J. R. I., Alloza, C., Shen, X., Barbu, M. C., Wigmore, E. M., Gibson, J., Hagenaars, S. P., Lewis, C. M., Ward, J., Smith, D. J., Sullivan, P. F., Haley, C. S., Breen, G., … McIntosh, A. M. (2018). Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nature Communications*, *9*(1), 1–10. https://doi.org/10.1038/s41467-018-03819-3

Hruschak, V., & Cochran, G. (2018). Psychosocial predictors in the transition from acute to chronic pain: a systematic review. *Psychology, Health and Medicine*, *23*(10), 1151–1167. https://doi.org/10.1080/13548506.2018.1446097

Hsia, D. A., Mitra, S. K., Hauck, C. R., Streblow, D. N., Nelson, J. A., Ilic, D., Huang, S., Li, E., Nemerow, G. R., Leng, J., Spencer, K. S. R., Cheresh, D. A., & Schlaepfer, D. D. (2003). Differential regulation of cell motility and invasion by FAK. *Journal of Cell Biology*, *160*(5), 753–767. https://doi.org/10.1083/jcb.200212114

Huber, M. D., Guan, T., & Gerace, L. (2009). Overlapping Functions of Nuclear Envelope Proteins NET25 (Lem2) and Emerin in Regulation of Extracellular Signal-Regulated Kinase Signaling in Myoblast Differentiation. *Molecular and Cellular Biology*, *29*(21), 5718–5728. https://doi.org/10.1128/mcb.00270-09

Huber, P. J. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, *35*(1), 73–101.

Hudson, J. I., & Pope, H. G. (1994). The concept of affective spectrum disorder: relationship to fibromyalgia and other syndromes of chronic fatigue and chronic muscle pain. *Bailliere's Clinical Rheumatology*, *8*(4), 839–856. https://doi.org/10.1016/S0950-3579(05)80051-2

Hunt, K., Adamson, J., Hewitt, C., & Nazareth, I. (2011). Do women consult more than men? A review of gender and consultation for back pain and headache. *Journal of Health Services Research and Policy*, *16*(2), 108–117. https://doi.org/10.1258/jhsrp.2010.009131

Hunter, D. D., Shah, V., Merlie, J. P., & Sanes, J. R. (1989). A laminin-like adhesive protein concentrated in the synaptic cleft of the neuromuscular junction. *Nature*, *338*(6212), 229–234. https://doi.org/10.1038/338229a0

Hutcheon, J. A., Chiolero, A., & Hanley, J. A. (2010). Random measurement error and regression dilution bias. *BMJ*, *340*, c2289. https://doi.org/10.1136/bmj.c2289

Hwang, J. I., Hyeon, S. K., Jae, R. L., Kim, E., Sung, H. R., & Suh, P. G. (2005). The interaction of phospholipase C-β3 with Shank2 regulates mGluR-mediated calcium signal. *Journal of Biological Chemistry*, *280*(13), 12467–12473. https://doi.org/10.1074/jbc.M410740200

Hyde, C. L., Nagle, M. W., Tian, C., Chen, X., Paciga, S. A., Wendland, J. R., Tung, J. Y., Hinds, D. A., Perlis, R. H., & Winslow, A. R. (2016). Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nature Genetics*, *48*(9), 1031–1036. https://doi.org/10.1038/ng.3623

Iacovides, S., Avidon, I., & Baker, F. C. (2015). Does pain vary across the menstrual cycle? A review. *European Journal of Pain (United Kingdom)*, *19*(10), 1389–1405. https://doi.org/10.1002/ejp.714

IASP. (2017a). *IASP Pain Terminology*. https://www.iasp-pain.org/resources/terminology/

IASP. (2017b). *Sensitization*. IASP Pain Terminology. https://www.iasp-pain.org/Education/Content.aspx?ItemNumber=1698#Sensitization

Inada, T., Koga, M., Ishiguro, H., … Y. H.-P., & 2008, U. (2008). Pathway-based association analysis of genome-wide screening data suggest that genes associated with the γ-aminobutyric acid receptor signaling pathway are. *Pharmacogenetics and Genomics*, *18*(4).

International HapMap Consortium. (2003). International HapMap Consortium. The International HapMap Project. *Nature*, *426*(6968), 789–796. http://www.hapmap.

Investigators, T. A. of U. R. P. (2019). The "All of Us" Research Program. *The New England Journal of Medicine*, *381*(7), 668–676.

Itoh, T., Satoh, M., Kanno, E., & Fukuda, M. (2006). Screening for target Rabs of TBC (Tre-2/Bub2/Cdc16) domain-containing proteins based on their Rab-binding activity. *Genes to Cells*, *11*(9), 1023–1037. https://doi.org/10.1111/j.1365-2443.2006.00997.x

Itoh, Y., Esaki, T., Shimoji, K., Cook, M., Law, M. J., Kaufman, E., & Sokoloff, L. (2003). Dichloroacetate effects on glucose and lactate oxidation by neurons and astroglia in vitro and on glucose utilization by brain in vivo. *Proceedings of the National Academy of Sciences*, *100*(8), 4879–4884. https://doi.org/10.1073/pnas.0831078100

Jackson, T., Thomas, S., Stabile, V., Han, X., Shotwell, M., & McQueen, K. (2015). Prevalence of chronic pain in low-income and middle-income countries: a systematic review and meta-analysis. *The Lancet*, *385*, S10. https://doi.org/10.1016/s0140-6736(15)60805-4

Jacob, T. C., Moss, S. J., & Jurd, R. (2008). GABA(A) receptor trafficking and its role in the dynamic modulation of neuronal inhibition. *Nature Reviews. Neuroscience*, *9*(5), 331–343. https://doi.org/10.1038/nrn2370

Jacobs, B., van Praag, H., & Gage, F. (2000). Adult brain neurogenesis and psychiatry: a novel theory of depression. *Molecular Psychiatry*, *5*, 262–269.

Jahan, F., Nanji, K., Qidwai, W., & Qasim, R. (2012). Fibromyalgia syndrome: An overview of pathophysiology, diagnosis and management. *Oman Medical Journal*, *27*(3), 192–195. https://doi.org/10.5001/omj.2012.44

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, Z., Abera, S. F., Abil, O. Z., Abraha, H. N., Abu-Raddad, L. J., Abu-Rmeileh, N. M. E., Accrombessi, M. M. K., … Murray, C. J. L. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 1789–1858. https://doi.org/10.1016/S0140-6736(18)32279-7

Jank, R., Gallee, A., Boeckle, M., Fiegl, S., & Pieh, C. (2017). Chronic Pain and Sleep Disorders in Primary Care. *Pain Research and Treatment*, *2017*, 34–37. https://doi.org/10.1155/2017/9081802

Janssens, A. C. J. W., & Kraft, P. (2012). Research Conducted Using Data Obtained through Online Communities: Ethical Implications of Methodological Limitations. *PLoS Medicine*, *9*(10), 10–13. https://doi.org/10.1371/journal.pmed.1001328

Jaracz, J., Gattner, K., Jaracz, K., & Górna, K. (2016). Unexplained Painful Physical Symptoms in Patients with Major Depressive Disorder: Prevalence, Pathophysiology and Management.

CNS Drugs, *30*(4), 293–304. https://doi.org/10.1007/s40263-016-0328-5

Jensen, M. P., & Karoly, P. (2001). Self-report scales and procedures for assessing pain in adults. In *Handbook of pain assessment, 2nd ed.* (pp. 15–34). The Guilford Press.

Ji, H., Ding, Z., Hawke, D., Xing, D., Jiang, B. H., Mills, G. B., & Lu, Z. (2012). AKT-dependent phosphorylation of Niban regulates nucleophosmin- and MDM2-mediated p53 stability and cell apoptosis. *EMBO Reports*, *13*(6), 554–560. https://doi.org/10.1038/embor.2012.53

Ji, R.-R., Nackley, A., Huh, Y., Terrando, N., & Maixner, W. (2018). *Neuroinflammation and Central Sensitization in Chronic and Widespread Pain*. *2*, 343–366.

Johnson, E. M., Daniel, D. C., & Gordon, J. (2013). The pur protein family: Genetic and structural features in development and disease. *Journal of Cellular Physiology*, *228*(5), 930–937. https://doi.org/10.1002/jcp.24237

Johnston, K. J. A., Adams, M. J., Nicholl, B. I., Ward, J., Strawbridge, R. J., Ferguson, A., Mcintosh, A. M., Bailey, M. E. S., & Smith, D. J. (2019). Genome-wide association study of multisite chronic pain in UK Biobank. *PLoS Genetics*, *15*(6), 1–22.

Johnston, K. J. A., Adams, M. J., Nicholl, B. I., Ward, J., Strawbridge, R. J., McIntosh, A. M., Smith, D. J., & Bailey, M. E. S. (2019). Identification of novel common variants associated with chronic pain using conditional false discovery rate analysis with major depressive disorder and assessment of pleiotropic effects of LRFN5. *Translational Psychiatry*, *9*(1). https://doi.org/10.1038/s41398-019-0613-4

Johnston, K. J. A., Ward, J., Ray, P. R., Adams, M. J., McIntosh, A. M., Smith, B. H., Strawbridge, R. J., Price, T. J., Smith, D. J., Nicholl, B. I., & Bailey, M. E. S. (2021). Sex-stratified genome-wide association study of multisite chronic pain in UK Biobank. *PLOS Genetics*, *17*(4), e1009428. https://doi.org/10.1371/journal.pgen.1009428

Joshi, B., Cameron, A., & Jagus, R. (2004). Characterization of mammalian eIF4E-family members. *European Journal of Biochemistry*, *271*(11), 2189–2203. https://doi.org/10.1111/j.1432-1033.2004.04149.x

Junqueira, D. R. G., Ferreira, M. L., Refshauge, K., Maher, C. G., Hopper, J. L., Hancock, M., Carvalho, M. G., & Ferreira, P. H. (2014). Heritability and lifestyle factors in chronic low back pain: Results of the Australian Twin Low Back Pain Study (The AUTBACK study). *European Journal of Pain (United Kingdom)*, *18*(10), 1410–1418. https://doi.org/10.1002/ejp.506

Kamaleri, Y., Natvig, B., Ihlebaek, C. M., Benth, J. S., & Bruusgaard, D. (2008). Number of pain sites is associated with demographic, lifestyle, and health-related factors in the general population. *European Journal of Pain*, *12*(6), 742–748. https://doi.org/10.1016/j.ejpain.2007.11.005

Kang, K. J., Pyo, J. H., Ryu, K. J., Kim, S. J., Ha, J. M., Choi, K., Hong, S. N., Min, B. H., Chang, D. K., Son, H. J., Rhee, P. L., Kim, J. J., & Kim, Y. H. (2015). Oncogenic Role of BOLL in Colorectal Cancer. *Digestive Diseases and Sciences*, *60*(6), 1663–1673. https://doi.org/10.1007/s10620-015-3533-z

Kapfhammer, H. P. (2006). Somatic symptoms in depression. *Dialogues in Clinical Neuroscience*, *8*(2), 227–239. https://doi.org/10.31887/dcns.2006.8.2/hpkapfhammer

Katan, M. (1986). Apolipoprotein E isoforms, serum cholesterol, and cancer. *Lancet*, *327*(8479), 507–508.

Kato, K., Sullivan, P. F., Evengård, B., & Pedersen, N. L. (2009). A population-based twin study of functional somatic syndromes. *Psychological Medicine*, *39*(3), 497–505. https://doi.org/10.1017/S0033291708003784

Katoh, M., & Katoh, M. (2004). Identification and characterization of ASXL3 gene in silico . In

International journal of oncology (Vol. 24, Issue 6, p. 1617). http://glasgow.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwpV1dS8MwFA1OUXwRv78lf6Cja9OsVXwYc0NEn7aJ-DKSNNGha4dzCP56b5Kua4uC4kspCS1tzs3l3CT3XIR8r-46FZ8gqB97hAVSBdwNOY-UUnGsPMqZ4oLpDd7uFe206c1j0FlUlgPPy5LUKCfWgU3a9KrzorT2v0GHNoBdJ9H-Afj8pdAA9wA_XMEA4Po

Katz, J., & Rosenbloom, B. N. (2015). The golden anniversary of Melzack and Wall's gate control theory of pain: Celebrating 50 years of pain research and management. *Pain Research and Management*, *20*(6), 285–286. https://doi.org/10.1155/2015/865487

Katz, J., Rosenbloom, B. N., & Fashler, S. (2015). Chronic pain, psychopathology, and DSM-5 somatic symptom disorder. *Canadian Journal of Psychiatry*, *60*(4), 160–167. https://doi.org/10.1177/070674371506000402

Katz, J., & Seltzer, Z. (2009). Transition from acute to chronic postsurgical pain: risk factors and protective fa. *Expert Review of Neurotherapeutics*, *9*(5), 723–744. https://doi.org/10.1586/ern.09.20

Kaya, S., Hermans, L., & Meeus, M. (2013). Central sensitization in chronic pelvic pain. *Pain Physician*, 291–308. http://www.painphysicianjournal.com/current/pdf?article=MTkzNg%3D%3D&journal=76

Keavney, B., Danesh, J., Parish, S., Palmer, A., Clark, S., Youngman, L., Lathrop, M., Peto, R., Collins, R., & Studies, I. (2006). *Fibrinogen and coronary heart disease : test of causality by ' Mendelian randomization .' July*, 935–943. https://doi.org/10.1093/ije/dyl114

Kendall, K. M., Van Assche, E., Andlauer, T. F. M., Choi, K. W., Luykx, J. J., Schulte, E. C., & Lu, Y. (2021). The genetic basis of major depression. *Psychological Medicine*, 1–14. https://doi.org/doi:10.1017/S0033291721000441

Kendrick, T., Dowrick, C., McBride, A., Howe, A., Clarke, P., Maisey, S., MichaelMoore, & Smith, P. W. (2009). Management of depression in UK general practice in relation to scores on depression severity questionnaires: Analysis of medical record data. *BMJ (Online)*, *338*(7697), 1–8. https://doi.org/10.1136/bmj.b750

Kenmochi, N., Suzuki, T., Uechi, T., Magoori, M., Kuniba, M., Higa, S., Watanabe, K., & Tanaka, T. (2001). The human mitochondrial ribosomal protein genes: Mapping of 54 genes to the chromosomes and implications for human disorders. *Genomics*, *77*(1–2), 65–70. https://doi.org/10.1006/geno.2001.6622

Kent, W. J., Sugnet, C. W., Furey, T. S., & Roskin, K. M. (2002). The Human Genome Browser at UCSC. *Genome Research*, *12*, 996–1006. https://doi.org/10.1101/gr.229102.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., Rush, A. J., Walters, E. E., & Wang, P. S. (2003). The epidemiology of major depressive disorder. *The Journal of the American Medical Association*, *289*(23), 3095–3105. https://doi.org/10.1001/jama.289.23.3095

Kessler, R. C., & Bromet, E. J. (2013). The Epidemiology of Depression Across Cultures. *Annual Review of Public Health*, *34*(1), 119–138. https://doi.org/10.1146/annurev-publhealth-031912-114409

Khoutorsky, A., & Price, T. J. (2018). Translational Control Mechanisms in Persistent Pain. *Trends in Neurosciences*, *41*(2), 100–114. https://doi.org/10.1016/j.tins.2017.11.006

Khramtsova, E. A., Davis, L. K., & Stranger, B. E. (2019). The role of sex in the genomics of human complex traits. *Nature Reviews Genetics*, *20*(3), 173–190. https://doi.org/10.1038/s41576-018-0083-1

Kidd, J. M., Cooper, G. M., Donahue, W. F., Hayden, H. S., Sampas, N., Graves, T., Hansen, N., Teague, B., Alkan, C., Antonacci, F., Haugen, E., Zerr, T., Yamada, N. A., Tsang, P., Newman, T. L., Tüzün, E., Cheng, Z., Ebling, H. M., Tusneem, N., … Eichler, E. E. (2008). Mapping and sequencing of structural variation from eight human genomes. *Nature*, *453*(7191), 56–64. https://doi.org/10.1038/nature06862

Kidd, T., Brose, K., Mitchell, K. J., Fetter, R. D., Tessier-Lavigne, M., Goodman, C. S., & Tear, G. (1998). Roundabout controls axon crossing of the CNS midline and defines a novel subfamily of evolutionarily conserved guidance receptors. *Cell*, *92*(2), 205–215. https://doi.org/10.1016/S0092-8674(00)80915-0

Kilic, S. S., Ozturk, R., Sarisozen, B., Rotthier, A., Baets, J., & Timmerman, V. (2009). Humoral immunodeficiency in congenital insensitivity to pain with anhidrosis. *Neurogenetics*, *10*(2), 161–165. https://doi.org/10.1007/s10048-008-0165-x

Kimura, M. (1954). Solution of a process of tandom genetic drift with a continuous model. *Proceedings of the National Academy of Sciences*, *2*(2), 144–150.

Kinchen, J. M., & Ravichandran, K. S. (2010). Identification of two evolutionarily conserved genes regulating processing of engulfed apoptotic cells. *Nature*, *464*(7289), 778–782. https://doi.org/10.1038/nature08853

Kinnaird, E., Stewart, C., & Tchanturia, K. (2019). Investigating alexithymia in autism: A systematic review and meta-analysis. *European Psychiatry*, *55*, 80–89. https://doi.org/10.1016/j.eurpsy.2018.09.004

Kitagawa, M. (2015). Notch signalling in the nucleus: Roles of Mastermind-like (MAML) transcriptional coactivators. *Journal of Biochemistry*, *159*(3), 287–294. https://doi.org/10.1093/jb/mvv123

Klijs, B., Scholtens, S., Mandemakers, J. J., Snieder, H., Stolk, R. P., & Smidt, N. (2015). Representativeness of the LifeLines cohort study. *PLoS ONE*, *10*(9), 1–12. https://doi.org/10.1371/journal.pone.0137203

Knaster, P., Estlander, A. M., Karlsson, H., Kaprio, J., & Kalso, E. (2016). Diagnosing depression in chronic pain patients: DSM-IV major depressive Disorder vs. Beck depression Inventory (BDI). *PLoS ONE*, *11*(3), 1–9. https://doi.org/10.1371/journal.pone.0151982

Koechlin, H., Coakley, R., Schechter, N., Werner, C., & Kossowsky, J. (2018). The role of emotion regulation in chronic pain: A systematic literature review. *Journal of Psychosomatic Research*, *107*(February), 38–45. https://doi.org/10.1016/j.jpsychores.2018.02.002

Kong, A., Gudbjartsson, D. F., Sainz, J., Jonsdottir, G. M., Gudjonsson, S. A., Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G., Shlien, A., Palsson, S. T., Frigge, M. L., Thorgeirsson, T. E., Gulcher, J. R., & Stefansson, K. (2002). A high-resolution recombination map of the human genome. *Nature Genetics*, *31*(3), 241–247. https://doi.org/10.1038/ng917

Kornbrot, D. (2014). Point Biserial Correlation. *Wiley StatsRef: Statistics Reference Online*, *2*, 1–3. https://doi.org/10.4135/9781412983907.n1452

Kósa, J. P., Balla, B., Speer, G., Kiss, J., Borsy, A., Podani, J., Takács, I., Lazáry, Á., Nagy, Z., Bácsi, K., Orosz, L., & Lakatos, P. (2009). Effect of menopause on gene expression pattern in bone tissue of nonosteoporotic women. *Menopause*, *16*(2), 367–377. https://doi.org/10.1097/gme.0b013e318188b260

Kosek, E., Cohen, M., Baron, R., Gebhart, G. F., Mico, J. A., Rice, A. S. C., Rief, W., & Sluka, A. K. (2016). Do we need a third mechanistic descriptor for chronic pain states? *Pain*, *157*(7), 1382–1386. https://doi.org/10.1097/j.pain.0000000000000507

Kowalski, M. P., & Krude, T. (2015). Functional roles of non-coding Y RNAs. *International Journal of Biochemistry and Cell Biology*, *66*, 20–29. https://doi.org/10.1016/j.biocel.2015.07.003

Kraev, A., Quednau, B. D., Leach, S., Li, X. F., Dong, H., Winkfein, R., Perizzolo, M., Cai, X., Yang, R. M., Philipson, K. D., & Lytton, J. (2001). Molecular Cloning of a Third Member of the Potassium-dependent Sodium-Calcium Exchanger Gene Family, NCKX3. *Journal of Biological Chemistry*, *276*(25), 23161–23172. https://doi.org/10.1074/jbc.M102314200

Krause, M. L., & Makol, A. (2016). Management of rheumatoid arthritis during pregnancy: Challenges and solutions. *Open Access Rheumatology: Research and Reviews*, *8*, 23–36. https://doi.org/10.2147/OARRR.S85340

Krishna, M. T., Subramanian, A., Adderley, N. J., Zemedikun, D. T., Gkoutos, G. V., & Nirantharakumar, K. (2019). Allergic diseases and long-term risk of autoimmune disorders: Longitudinal cohort study and cluster analysis. *European Respiratory Journal*, *54*(5). https://doi.org/10.1183/13993003.00476-2019

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kroenke, K., Wu, J., Bair, M. J., Krebs, E. E., Damush, T. M., & Tu, W. (2011). Reciprocal relationship between pain and depression: A 12-month longitudinal analysis in primary care. *Journal of Pain*, *12*(9), 964–973. https://doi.org/10.1016/j.jpain.2011.03.003

Kruglyak, L. (2007). The road to genome-wide association studies. *The Statistics of Gene Mapping*, 279–305. https://doi.org/10.1007/978-0-387-49686-3_12

Kuehner, C. (2017). Why is depression more common among women than among men? *The Lancet Psychiatry*, *4*(2), 146–158. https://doi.org/10.1016/S2215-0366(16)30263-2

Kuja-Panula, J., Kiiltomäki, M., Yamashiro, T., Rouhiainen, A., & Rauvala, H. (2003). AMIGO, a transmembrane protein implicated in axon tract development, defines a novel protein family with leucine-rich repeats. *Journal of Cell Biology*, *160*(6), 963–973. https://doi.org/10.1083/jcb.200209074

Kukurba, K. R., Parsana, P., Balliu, B., Smith, K. S., Zappala, Z., Knowles, D. A., Favé, M. J., Davis, J. R., Li, X., Zhu, X., Potash, J. B., Weissman, M. M., Shi, J., Kundaje, A., Levinson, D. F., Awadalla, P., Mostafavi, S., Battle, A., & Montgomery, S. B. (2016). Impact of the X chromosome and sex on regulatory variation. *Genome Research*, *26*(6), 768–777. https://doi.org/10.1101/gr.197897.115

Kuner, R., & Flor, H. (2016). Structural plasticity and reorganisation in chronic pain. *Nature Reviews Neuroscience*, *18*(1), 20–30. https://doi.org/10.1038/nrn.2016.162

Kurenova, E., Xu, L.-H., Yang, X., Baldwin, A. S., Craven, R. J., Hanks, S. K., Liu, Z., & Cance, W. G. (2004). Focal Adhesion Kinase Suppresses Apoptosis by Binding to the Death Domain of Receptor-Interacting Protein. *Molecular and Cellular Biology*, *24*(10), 4361–4371. https://doi.org/10.1128/mcb.24.10.4361-4371.2004

Kurosaki, T. (2000). Functional dissection of BCR signaling pathways. *Current Opinion in Immunology*, *12*(3), 276–281. https://doi.org/10.1016/S0952-7915(00)00087-X

Kurtsdotter, I., Topcic, D., Karlén, A., Singla, B., Hagey, D. W., Bergsland, M., Siesjö, P., Nistér, M., Carlson, J. W., Lefebvre, V., Persson, O., Holmberg, J., & Muhr, J. (2017). SOX5/6/21 prevent oncogene-driven transformation of brain stem cells. *Cancer Research*, *77*(18), 4985–4997. https://doi.org/10.1158/0008-5472.CAN-17-0704

Kwiatkowski, K., & Mika, J. (2018). The importance of chemokines in neuropathic pain development and opioid analgesic potency. *Pharmacological Reports*, *70*(4), 821–830.

https://doi.org/10.1016/j.pharep.2018.01.006

La Touche, R., Paris-Alemany, A., Hidalgo-Pérez, A., López-de-Uralde-Villanueva, I., Angulo-Diaz-Parreño, S., & Muñoz-García, D. (2018). Evidence for Central Sensitization in Patients with Temporomandibular Disorders: A Systematic Review and Meta-analysis of Observational Studies. *Pain Practice*, *18*(3), 388–409. https://doi.org/10.1111/papr.12604

Lafrenière, R. G., MacDonald, M. L. E., Dubé, M. P., MacFarlane, J., O'Driscoll, M., Brais, B., Meilleur, S., Brinkman, R. R., Dadivas, O., Pape, T., Platon, C., Radomski, C., Risler, J., Thompson, J., Guerra-Escobio, A. M., Davar, G., Breakefield, X. O., Pimstone, S. N., Green, R., … Samuels, M. E. (2004). Identification of a Novel Gene (HSN2) Causing Hereditary Sensory and Autonomic Neuropathy Type II through the Study of Canadian Genetic Isolates. *American Journal of Human Genetics*, *74*(5), 1064–1073. https://doi.org/10.1086/420795

Lander, E. S. (1996). The new genomics: Global views of biology. *Science*, *274*(5287), 536–539. https://doi.org/10.1126/science.274.5287.536

Laroche, F., Ramoz, N., Leroy, S., Fortin, C., Rousselot-Paillet, B., Philippe, A., Colleaux, L., Bresson, J.-L., Mogenet, A., Golse, B., Mouren-Simeoni, M.-C., Gorwood, P., Galli, T., Simonneau, M., Krebs, M.-O., & Robel, L. (2008). Polymorphisms of coding trinucleotide repeats of homeogenes in neurodevelopmental psychiatric disorders. *Psychiatric Genetics*, *18*(6), 295–301. https://doi.org/10.1097/YPG.0b013e3283060fa5

Larrieu-Lahargue, F., Thomas, K. R., & Li, D. Y. (2012). Netrin Ligands and Receptors: Lessons From Neurons to the Endothelium. *Trends in Cardiovascular Medicine*, *22*(2), 44–47. https://doi.org/10.1016/j.tcm.2012.06.010

Latremoliere, A., & Woolf, C. J. (2009). Central Sensitization: A Generator of Pain Hypersensitivity by Central Neural Plasticity. *Journal of Pain*, *10*(9), 895–926. https://doi.org/10.1016/j.jpain.2009.06.012

Lauritzen, F., Heuser, K., de Lanerolle, N., Lee, T., Spencer, D., Kim, J., Gjedde, A., Eid, T., & Bergersena, L. (2012). Redistribution of monocarboxylate transporter 2 on the surface of astrocytes in the human epileptogenic hippocampus. *Glia*, *60*(7), 1172–1181. https://doi.org/10.3174/ajnr.A1256.Functional

Lawlor, D. A. (2016). *Commentary : Two-sample Mendelian randomization : opportunities and challenges Z : randomize Z : randomize Z : Genetic Z : Genetic*. 908–915. https://doi.org/10.1093/ije/dyw127

Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N., & Smith, G. D. (2008). Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Statistics in Medicine*, *27*(8), 1133–1163. https://doi.org/10.1002/sim.3034

Lee, P. H., O'dushlaine, C., Thomas, B., & Purcell, S. M. (2012). INRICH: Interval-based enrichment analysis for genome-wide association studies. *Bioinformatics*, *28*(13), 1797–1799. https://doi.org/10.1093/bioinformatics/bts191

Lee, Y. C., Durr, A., Majczenko, K., Huang, Y. H., Liu, Y. C., Lien, C. C., Tsai, P. C., Ichikawa, Y., Goto, J., Monin, M. L., Li, J. Z., Chung, M. Y., Mundwiller, E., Shakkottai, V., Liu, T. T., Tesson, C., Lu, Y. C., Brice, A., Tsuji, S., … Soong, B. W. (2012). Mutations in KCND3 cause spinocerebellar ataxia type 22. *Annals of Neurology*, *72*(6), 859–869. https://doi.org/10.1002/ana.23701 [doi]

Lehalle, D., Mosca-Boidron, A. L., Begtrup, A., Boute-Benejean, O., Charles, P., Cho, M. T., Clarkson, A., Devinsky, O., Duffourd, Y., Duplomb-Jego, L., Gérard, B., Jacquette, A., Kuentz, P., Masurel-Paulet, A., McDougall, C., Moutton, S., Olivié, H., Park, S. M., Rauch, A., … Faivre, L. (2017). STAG1 mutations cause a novel cohesinopathy characterised by unspecific syndromic intellectual disability. *Journal of Medical Genetics*, *54*(7), 479–488. https://doi.org/10.1136/jmedgenet-2016-104468

Leitsalu, L., Haller, T., Esko, T., Tammesoo, M. L., Alavere, H., Snieder, H., Perola, M., Ng, P. C., Mägi, R., Milani, L., Fischer, K., & Metspalu, A. (2015). Cohort profile: Estonian biobank of the Estonian genome center, university of Tartu. *International Journal of Epidemiology*, *44*(4), 1137–1147. https://doi.org/10.1093/ije/dyt268

Lévesque, M., Potvin, S., Marchand, S., Stip, E., Grignon, S., Pierre, L., Lipp, O., & Goffaux, P. (2012). Pain Perception in Schizophrenia: Evidence of a Specific Pain Response Profile. *Pain Medicine (United States)*, *13*(12), 1571–1579. https://doi.org/10.1111/j.1526-4637.2012.01505.x

Li, H. J., Qu, N., Hui, L., Cai, X., Zhang, C. Y., Zhong, B. L., Zhang, S. F., Chen, J., Xia, B., Wang, L., Jia, Q. F., Li, W., Chang, H., Xiao, X., Li, M., & Li, Y. (2020). Further confirmation of netrin 1 receptor (DCC) as a depression risk gene via integrations of multi-omics data. *Translational Psychiatry*, *10*(1). https://doi.org/10.1038/s41398-020-0777-y

Li, H., Xing, L., Zhao, N., Wang, J., & Zheng, N. (2018). Furosine induced apoptosis by the regulation of STAT1/STAT2 and UBA7/UBE2L6 genes in HepG2 cells. *International Journal of Molecular Sciences*, *19*(6), 1–13. https://doi.org/10.3390/ijms19061629

Li, J., Dallmayer, M., Kirchner, T., Musa, J., & Grünewald, T. G. P. (2018). PRC1: Linking Cytokinesis, Chromosomal Instability, and Cancer Evolution. *Trends in Cancer*, *4*(1), 59–73. https://doi.org/10.1016/j.trecan.2017.11.002

Li, J. X. (2015). Pain and depression comorbidity: A preclinical perspective. *Behavioural Brain Research*, *276*, 92–98. https://doi.org/10.1016/j.bbr.2014.04.042

Li, T., Zeng, Z., Zhao, Q., Wang, T., Huang, K., Li, J., Li, Y., Liu, J., Wei, Z., Wang, Y., Feng, G., He, L., & Shi, Y. (2013). *FoxP2* is significantly associated with schizophrenia and major depression in the Chinese Han Population. *The World Journal of Biological Psychiatry*, *14*(2), 146–150. https://doi.org/10.3109/15622975.2011.615860

Li, Y. I., Geijn, B. Van De, Raj, A., Knowles, D. A., Petti, A. A., Golan, D., Gilad, Y., & Pritchard, J. K. (2016). RNA splicing is a primary link between genetic variation and disease. *Science*, *352*(6285), 600–605.

Li, Y. R., & Keating, B. J. (2014). Trans-ethnic genome-wide association studies: Advantages and challenges of mapping in diverse populations. *Genome Medicine*, *6*(10), 1–14. https://doi.org/10.1186/s13073-014-0091-5

Liao, J., Li, X., Wong, T., Wang, J. J., Khor, C. C., Tai, E. S., Aung, T., Teo, Y., & Cheng, C. (2014). *Impact of Measurement Error on Testing Genetic Association with Quantitative Traits*. *9*(1), 1–9. https://doi.org/10.1371/journal.pone.0087044

Liberzon, A., Birger, C., Ghandi, M., Jill, P., Tamayo, P., Jolla, L., & Jolla, L. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Systems*, *1*(6), 417–425. https://doi.org/10.1016/j.cels.2015.12.004.The

Lichtshtein, D., Dobkin, J., Ebstein, R. P., Biederman, J., Rimon, R., & Belmaker, R. H. (1978). Gamma-aminobutyric acid (GABA) in the CSF of schizophrenic patients before and after neuroleptic treatment. *British Journal of Psychiatry*, *132*(2), 145–148. https://doi.org/10.1192/bjp.132.2.145

Lightowlers, R. N., Taylor, R. W., & Turnbull, D. M. (2015). Mutations causing mitochondrial disease: What is new and what challenges remain? *Science*, *349*(6255), 1494–1499. https://doi.org/10.1126/science.aac7516

Liley, J., & Wallace, C. (2015). A pleiotropy-informed Bayesian false discovery rate adapted to a shared control design finds new disease associations from GWAS summary statistics. *PLoS Genetics*, *11*(2), e1004926. https://doi.org/10.1371/journal.pgen.1004926

Lin, D. Y., & Sullivan, P. F. (2009). Meta-Analysis of Genome-wide Association Studies with Overlapping Subjects. *American Journal of Human Genetics*, *85*(6), 862–872. https://doi.org/10.1016/j.ajhg.2009.11.001

Lin, Z., Liu, J., Ding, H., Xu, F., & Liu, H. (2018). Structural basis of SALM5-induced PTPδ dimerization for synaptic differentiation. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-017-02414-2

Lips, E. S., Cornelisse, L. N., Toonen, R. F., Min, J. L., Hultman, C. M., Holmans, P. A., O'Donovan, M. C., Purcell, S. M., Smit, A. B., Verhage, M., Sullivan, P. F., Visscher, P. M., & Posthuma, D. (2012). Functional gene group analysis identifies synaptic gene groups as risk factor for schizophrenia. *Molecular Psychiatry*, *17*(10), 996–1006. https://doi.org/10.1038/mp.2011.117

Lister, K. C., Bouchard, S. M., Markova, T., Aternali, A., Denecli, P., Pimentel, S. D., Majeed, M., Austin, J. S., Amanda, A. C., & Mogil, J. S. (2020). Chronic pain produces hypervigilance to predator odor in mice. *Current Biology*, *30*(15), R866–R867. https://doi.org/10.1016/j.cub.2020.06.025

Liu, J., & Krantz, I. D. (2009). Cornelia de Lange syndrome, cohesin, and beyond. *Clinical Genetics*, *76*(4), 303–314. https://doi.org/10.1111/j.1399-0004.2009.01271.x

Liu, J. Z., Van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., Takahashi, A., Ripke, S., Lee, J. C., Jostins, L., Shah, T., Abedian, S., Cheon, J. H., Cho, J., Daryani, N. E., Franke, L., Fuyuno, Y., Hart, A., Juyal, R. C., Juyal, G., … Weersma, R. K. (2015). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature Genetics*, *47*(9), 979–986. https://doi.org/10.1038/ng.3359

Liu, Wanli, Tolar, P., Song, W., & Kim, T. J. (2020). Editorial: BCR Signaling and B Cell Activation. *Nature Immunology*, *11*(45), 1–2. https://doi.org/10.1038/ni1480

Liu, Wenli, Lee, H. W., Liu, Y., Wang, R., & Rodgers, G. P. (2010). Olfactomedin 4 is a novel target gene of retinoic acids and 5-aza-2'-deoxycytidine involved in human myeloid leukemia cell growth, differentiation, and apoptosis. *Blood*, *116*(23), 4938–4947. https://doi.org/10.1182/blood-2009-10-246439

Liu, Wenli, Yan, M., Liu, Y., Wang, R., Li, C., Deng, C., Singh, A., Coleman, W. G., & Rodgers, G. P. (2010). Olfactomedin 4 down-regulates innate immunity against Helicobacter pylori infection. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(24), 11056–11061. https://doi.org/10.1073/pnas.1001269107

Liu, Y., Sarkar, A., Kheradpour, P., Ernst, J., & Kellis, M. (2017). Evidence of reduced recombination rate in human regulatory domains. *Genome Biology*, *18*(1), 1–11. https://doi.org/10.1186/s13059-017-1308-x

Liu, Y. Y., Tanikawa, C., Ueda, K., & Matsuda, K. (2019). INKA2, a novel p53 target that interacts with the serine/threonine kinase PAK4. *International Journal of Oncology*, *54*(6), 1907–1920. https://doi.org/10.3892/ijo.2019.4786

Lluch, E., Torres, R., Nijs, J., & Van Oosterwijck, J. (2014). Evidence for central sensitization in patients with osteoarthritis pain: A systematic literature review. *European Journal of Pain (United Kingdom)*, *18*(10), 1367–1375. https://doi.org/10.1002/j.1532-2149.2014.499.x

Lo, W.-S., Harano, M., Gawlik, M., Yu, Z., Chen, J., Pun, F. W., Tong, K.-L., Zhao, C., Ng, S.-K., Tsang, S.-Y., Uchimura, N., Stober, G., & Xue, H. (2007). GABRB2 Association with Schizophrenia: Commonalities and Differences Between Ethnic Groups and Clinical Subtypes. *Biological Psychiatry*, *61*(5), 653–660. https://doi.org/10.1016/J.BIOPSYCH.2006.05.003

Lodish, H. F., Berk, A., Kaiser, C. A., Krieger, M., Bretscher, A., Ploegh, H. L., Amon, A., & Martin, K. C. (2016). *Molecular Cell Biology* (Eighth). W.H. Freeman-Macmillan Learning.

https://go.exlibris.link/8HWT4TXr

Loh, P. R., Tucker, G., Bulik-Sullivan, B. K., Vilhjálmsson, B. J., Finucane, H. K., Salem, R. M., Chasman, D. I., Ridker, P. M., Neale, B. M., Berger, B., Patterson, N., & Price, A. L. (2015). Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nature Genetics*, *47*(3), 284–290. https://doi.org/10.1038/ng.3190

Lunardi, A. C., Marques Da Silva, C. C. B., Rodrigues Mendes, F. A., Marques, A. P., Stelmach, R., & Fernandes Carvalho, C. R. (2011). Musculoskeletal dysfunction and pain in adults with asthma. *Journal of Asthma*, *48*(1), 105–110. https://doi.org/10.3109/02770903.2010.520229

Luppa, M., Sikorski, C., Luck, T., Ehreke, L., Konnopka, A., Wiese, B., Weyerer, S., König, H. H., & Riedel-Heller, S. G. (2012). Age- and gender-specific prevalence of depression in latest-life - Systematic review and meta-analysis. *Journal of Affective Disorders*, *136*(3), 212–221. https://doi.org/10.1016/j.jad.2010.11.033

Macdonald, M. E., Ambrose, C. M., Duyao, M. P., Myers, R. H., Lin, C., Srinidhi, L., Barnes, G., Taylor, S. A., James, M., Groat, N., Macfarlane, H., Jenkins, B., Anderson, M. A., Wexler, N. S., Gusella, J. F., Riba-ramirer, L., Shah, M., Stanton, V. P., Strobel, S. A., … Harper, P. S. (1993). *A Novel Gene Containing a Trinucleotide That Is Expanded and Unstable on Huntington ' s Disease Chromosomes*. *72*, 971–983.

Macfarlane, G. J. (2016). The epidemiology of chronic pain. *Pain*, *157*(10), 2158–2159. https://doi.org/10.1097/j.pain.0000000000000676

Macfarlane, G. J., Barnish, M. S., & Jones, G. T. (2017). Persons with chronic widespread pain experience excess mortality: longitudinal results from UK Biobank and meta-analysis. *Annals of the Rheumatic Diseases*. https://doi.org/10.1136/annrheumdis-2017-211476

Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature*, *456*(7218), 18–21. https://doi.org/10.1038/456018a

Maixner, W., Fillingim, R. B., Williams, D. A., Smith, S. B., & Slade, G. D. (2016). Overlapping Chronic Pain Conditions: Implications for Diagnosis and Classification. *Journal of Pain*, *17*(9), T93–T107. https://doi.org/10.1016/j.jpain.2016.06.002

Majeed, M. H., Ubaidulhaq, M., Rugnath, A., & Eriator, I. K. E. (2018). Extreme ends of pain sensitivity in SCN9A mutation variants: Case report and literature review. *Innovations in Clinical Neuroscience*, *15*(11–12), 33–35.

Mann, N., King, T., & Murphy, R. (2019). Review of primary and secondary erythromelalgia. *Clinical and Experimental Dermatology*, *44*(5), 477–482. https://doi.org/10.1111/ced.13891

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., Mccarthy, M. I., Ramos, E. M., Cardon, L. R., Chakravarti, A., Cho, J. H., Guttmacher, A. E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C. N., Slatkin, M., Valle, D., Whittemore, A. S., … Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature*, *461*(7265), 747–753. https://doi.org/10.1038/nature08494

Mansfield, K. E., Sim, J., Croft, P., & Jordan, K. P. (2017). *Identifying patients with chronic widespread pain in primary care*. *158*(Box 1), 110–119.

Mansour, A. R., Baliki, M. N., Huang, L., Torbey, S., Herrmann, K. M., Schnitzer, T. J., & Vania Apkarian, A. (2013). Brain white matter structural properties predict transition to chronic pain. *Pain*, *154*(10), 2160–2168. https://doi.org/10.1016/j.pain.2013.06.044

Mapplebeck, J. C. S., Beggs, S., & Salter, M. W. (2016). Sex differences in pain : a tale of two immune cells. *Pain*, *157*, 2–6.

Marchini, J. (2015). *UK Biobank Phasing and Imputation Documentation Version 1.2*. UK Biobank. https://doi.org/10.3788/gzxb20114002.0217

Marees, A. T., de Kluiver, H., Stringer, S., Vorspan, F., Curis, E., Marie-Claire, C., & Derks, E. M. (2018). A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *International Journal of Methods in Psychiatric Research*, *27*(2), 1–10. https://doi.org/10.1002/mpr.1608

Marjoram, P., Zubair, A., & Nuzhdin, S. V. (2014). Post-GWAS: where next? More samples, more SNPs or more biology? *Heredity*, *112*(1), 79–88. https://doi.org/10.1038/hdy.2013.52

Marouli, E., Graff, M., Medina-Gomez, C., Lo, K. S., Wood, A. R., Kjaer, T. R., Fine, R. S., Lu, Y., Schurmann, C., Highland, H. M., Rüeger, S., Thorleifsson, G., Justice, A. E., Lamparter, D., Stirrups, K. E., Turcot, V., Young, K. L., Winkler, T. W., Esko, T., … Lettre, G. (2017). Rare and low-frequency coding variants alter human adult height. *Nature*, *542*(7640), 186–190. https://doi.org/10.1038/nature21039

Marrie, R. A., Yu, B. N., Leung, S., Elliott, L., Warren, S., Wolfson, C., Tremlett, H., Blanchard, J., & Fisk, J. D. (2012). The incidence and prevalence of fibromyalgia are higher in multiple sclerosis than the general population: A population-based study. *Multiple Sclerosis and Related Disorders*, *1*(4), 162–167. https://doi.org/10.1016/j.msard.2012.06.001

Martínez-Gimeno, M., José Gamundi, M., Hernan, I., Maseras, M., Millá, E., Ayuso, C., García-Sandoval, B., Beneyto, M., Vilela, C., Baiget, M., Antiñolo, G., & Carballo, M. (2003). Mutations in the pre-mRNA splicing-factor genes PRPF3, PRPF8, and PRPF31 in Spanish families with autosomal dominant retinitis pigmentosa. *Investigative Ophthalmology and Visual Science*, *44*(5), 2171–2177. https://doi.org/10.1167/iovs.02-0871

Martínez-Jauand, M., Sitges, C., Rodríguez, V., Picornell, A., Ramon, M., Buskila, D., & Montoya, P. (2013). Pain sensitivity in fibromyalgia is associated with catechol-O- methyltransferase (COMT) gene. *European Journal of Pain (United Kingdom)*, *17*(1), 16–27. https://doi.org/10.1002/j.1532-2149.2012.00153.x

Martínez, A. I., Pérez-Arellano, I., Pekkala, S., Barcelona, B., & Cervera, J. (2010). Genetic, structural and biochemical basis of carbamoyl phosphate synthetase 1 deficiency. *Molecular Genetics and Metabolism*, *101*(4), 311–323. https://doi.org/10.1016/j.ymgme.2010.08.002

Martino, G., Caputo, A., Schwarz, P., Bellone, F., Fries, W., Quattropani, M. C., & Vicario, C. M. (2020). Alexithymia and inflammatory bowel disease: A systematic review. *Frontiers in Psychology*, *11*(August), 1–9. https://doi.org/10.3389/fpsyg.2020.01763

Martucci, K. T., & Mackey, S. C. (2016). Imaging Pain. *Anesthesiology Clinics*, *34*(2), 255–269. https://doi.org/10.1016/j.anclin.2016.01.001

Masaki, T., & Matsumura, K. (2010). Biological role of dystroglycan in schwann cell function and its implications in peripheral nervous system diseases. *Journal of Biomedicine and Biotechnology*, *2010*. https://doi.org/10.1155/2010/740403

Masel, J. (2011). Genetic drift. *Current Biology*, *21*(20), R837–R838. https://doi.org/10.1016/j.cub.2011.08.007

Masrori, P., & Van Damme, P. (2020). Amyotrophic lateral sclerosis: a clinical review. *European Journal of Neurology*, *27*(10), 1918–1929. https://doi.org/10.1111/ene.14393

Mateos-Aparicio, P., & Rodríguez-Moreno, A. (2019). The impact of studying brain plasticity. *Frontiers in Cellular Neuroscience*, *13*(February), 1–5. https://doi.org/10.3389/fncel.2019.00066

Mather, K. (1964). R . A . Fisher ' s Work in Genetics. *Biometrics*, *20*(2), 330–342.

Maurer, D. M., Raymond, T. J., & Davis, B. N. (2018). Depression : Screening and Diagnosis. *American Family Physician*, *98*(October), 509.

McCarroll, S. A. (2008). Extending genome-wide association studies to copy-number variation.

Human Molecular Genetics, *17*(R2), 135–142. https://doi.org/10.1093/hmg/ddn282

McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., Kang, H. M., Fuchsberger, C., Danecek, P., Sharp, K., Luo, Y., Sidore, C., Kwong, A., Timpson, N., Koskinen, S., Vrieze, S., Scott, L. J., Zhang, H., Mahajan, A., … Marchini, J. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics*, *48*(10), 1279–1283. https://doi.org/10.1038/ng.3643

McCormick, H., Young, P. E., Hur, S. S. J., Booher, K., Chung, H., Cropley, J. E., Giannoulatou, E., & Suter, C. M. (2017). Isogenic mice exhibit sexually-dimorphic DNA methylation patterns across multiple tissues. *BMC Genomics*, *18*(1), 1–9. https://doi.org/10.1186/s12864-017-4350-x

McCullagh, P., & Tibshirani, R. (1990). A simple method for the adjustment of profile likelihoods. *Journal of the Royal Statistical Society. Series B (Methodological)*, *52*(2), 325–344.

McIntosh, A. M., Hall, L. S., Zeng, Y., Adams, M. J., Gibson, J., Wigmore, E., Hagenaars, S. P., Davies, G., Fernandez-Pujals, A. M., Campbell, A. I., Clarke, T. K., Hayward, C., Haley, C. S., Porteous, D. J., Deary, I. J., Smith, D. J., Nicholl, B. I., Hinds, D. A., Jones, A. V., … Hocking, L. J. (2016). Genetic and Environmental Risk for Chronic Pain and the Contribution of Risk Variants for Major Depressive Disorder: A Family-Based Mixed-Model Analysis. *PLoS Medicine*, *13*(8), 1–17. https://doi.org/10.1371/journal.pmed.1002090

McIntosh, A. M., Sullivan, P. F., & Lewis, C. M. (2019). Uncovering the Genetic Architecture of Major Depression. *Neuron*, *102*(1), 91–103. https://doi.org/10.1016/j.neuron.2019.03.022

McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, M. (n.d.). *Online Mendelian Inheritance in Man, OMIM*. Retrieved October 7, 2020, from omim.org

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biology*, *17*(1), 1–14. https://doi.org/10.1186/s13059-016-0974-4

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., & Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*, *26*(16), 2069–2070. https://doi.org/10.1093/bioinformatics/btq330

McNally, K. P., Bazirgan, O. A., & McNally, F. J. (2000). Two domains of p80 katanin regulate microtubule severing and spindle pole targeting by p60 katanin. *Journal of Cell Science*, *113*(9), 1623–1633.

Mecklenburg, J., Zou, Y., Wangzhou, A., Garcia, D., Lai, Z., Tumanov, A. V., Dussor, G., Price, T. J., & Akopian, A. N. (2020). Transcriptomic sex differences in sensory neuronal populations of mice. *Scientific Reports*, *10*(1), 1–18. https://doi.org/10.1038/s41598-020-72285-z

Medina-Gomez, C., Felix, J. F., Estrada, K., Peters, M. J., Herrera, L., Kruithof, C. J., Duijts, L., Hofman, A., van Duijn, C. M., Uitterlinden, A. G., Jaddoe, V. W. V., & Rivadeneira, F. (2015). Challenges in conducting genome-wide association studies in highly admixed multi-ethnic populations: the Generation R Study. *European Journal of Epidemiology*, *30*(4), 317–330. https://doi.org/10.1007/s10654-015-9998-4

Meeus, M., Vervisch, S., De Clerck, L. S., Moorkens, G., Hans, G., & Nijs, J. (2012). Central sensitization in patients with rheumatoid arthritis: A systematic literature review. *Seminars in Arthritis and Rheumatism*, *41*(4), 556–567. https://doi.org/10.1016/j.semarthrit.2011.08.001

Mehlen, P., & Fearon, E. R. (2004). Role of the dependence receptor DCC in colorectal cancer pathogenesis. *Journal of Clinical Oncology*, *22*(16), 3420–3428. https://doi.org/10.1200/JCO.2004.02.019

Meijers, R., Smock, R. G., Zhang, Y., & Wang, J. H. (2020). Netrin Synergizes Signaling and Adhesion through DCC. *Trends in Biochemical Sciences*, *45*(1), 6–12. https://doi.org/10.1016/j.tibs.2019.10.005

Meloto, C. B., Benavides, R., Lichtenwalter, R. N., Wen, X., Tugarinov, N., Zorina-Lichtenwalter, K., Chabot-Doré, A. J., Piltonen, M. H., Cattaneo, S., Verma, V., Klares, R., Khoury, S., Parisien, M., & Diatchenko, L. (2018). Human pain genetics database: A resource dedicated to human pain genetics research. In *Pain* (Vol. 159, Issue 4). https://doi.org/10.1097/j.pain.0000000000001135

Melzack, R. (1990). Phantom limbs and the concept of a neuromatrix. *Trends in Neurosciences*, *13*(3), 88–92. https://doi.org/10.1016/0166-2236(90)90179-E

Melzack, R. (2001). Pain and the Neuromatrix in the Brain. *Journal of Dental Education*, *65*(12), 1378–1382. https://doi.org/10.1002/j.0022-0337.2001.65.12.tb03497.x

Melzack, R., & Wall, P. D. (1965). Pain Mechanisms: A New Theory. *Science*, *150*(3699), 971–979. https://doi.org/10.1016/S0262-4079(07)61565-6

Mendel, G. (1866). Versuche über Plflanzen-hybriden. *Verhandlungen Des Naturforschenden Ver-Eines in Brünn, Bd. IV Für Das Jahr 1865*.

Mendell, L. M. (2014). Constructing and deconstructing the gate theory of pain. *Pain*, *155*(2), 210–216. https://doi.org/10.1016/j.pain.2013.12.010

Meng, W., Adams, M. J., Reel, P., Rajendrakumar, A., Huang, Y., Deary, I. J., Palmer, C. N. A., McIntosh, A. M., & Smith, B. H. (2019). Genetic correlations between pain phenotypes and depression and neuroticism. *European Journal of Human Genetics*, 358–366. https://doi.org/10.1038/s41431-019-0530-2

Meng, W., Adams, M. J., Reel, P., Rajendrakumar, A., Huang, Y., Deary, I. J., Palmer, C. N. A., McIntosh, A. M., & Smith, B. H. (2020). Genetic correlations between pain phenotypes and depression and neuroticism. *European Journal of Human Genetics*, *28*(3), 358–366. https://doi.org/10.1038/s41431-019-0530-2

Menges, C. W., Altomare, D. A., & Testa, J. R. (2009). FAS-associated factor 1 (FAF1): Diverse functions and implications for oncogenesis. *Cell Cycle*, *8*(16), 2528–2534. https://doi.org/10.4161/cc.8.16.9280

Merskey, H., & Bogduk, N. (1994). Classification of Chronic Pain. In *IASP Pain Terminology*. https://doi.org/10.1002/ana.20394

Mesman, S., Bakker, R., & Smidt, M. P. (2020). Tcf4 is required for correct brain development during embryogenesis. *Molecular and Cellular Neuroscience*, *106*(April), 103502. https://doi.org/10.1016/j.mcn.2020.103502

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829.

Mifflin, K. A., & Kerr, B. J. (2017). Pain in autoimmune disorders. *Journal of Neuroscience Research*, *95*(6), 1282–1294. https://doi.org/10.1002/jnr.23844

Miki, H., Setou, M., Kaneshiro, K., & Hirokawa, N. (2001). All kinesin superfamily protein, KIF, genes in mouse and human. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(13), 7004–7011. https://doi.org/10.1073/pnas.111145398

Mills, E. P., Di Pietro, F., Alshelh, Z., Peck, C. C., Murray, G. M., Vickers, E. R., & Henderson, L. A. (2018). Brainstem pain-control circuitry connectivity in chronic neuropathic pain. *Journal of Neuroscience*, *38*(2), 465–473. https://doi.org/10.1523/JNEUROSCI.1647-17.2017

Mills, M. C., & Rahal, C. (2019). A scientometric review of genome-wide association studies.

Communications Biology, *2*(1). https://doi.org/10.1038/s42003-018-0261-x

Mills, M. C., & Rahal, C. (2020). The GWAS Diversity Monitor tracks diversity by disease in real time. *Nature Genetics*, *52*(3), 242–243. https://doi.org/10.1038/s41588-020-0580-y

Mills, S. E. E., Nicolson, K. P., & Smith, B. H. (2019). Chronic pain: a review of its epidemiology and associated factors in population-based studies. *British Journal of Anaesthesia*, *123*(2), e273–e283. https://doi.org/10.1016/j.bja.2019.03.023

Minde, J., Toolanen, G., Andersson, T., Nennesmo, I., Remahl, I. N., Svensson, O., & Solders, G. (2004). Familial insensitivity to pain (HSAN V) and a mutation in the NGFB gene. A neurophysiological and pathological study. *Muscle and Nerve*, *30*(6), 752–760. https://doi.org/10.1002/mus.20172

Minen, M. T., De Dhaem, O. B., Van Diest, A. K., Powers, S., Schwedt, T. J., Lipton, R., & Silbersweig, D. (2016). Migraine and its psychiatric comorbidities. *Journal of Neurology, Neurosurgery and Psychiatry*, *87*(7), 741–749. https://doi.org/10.1136/jnnp-2015-312233

Missler, M., Zhang, W., Rohlmann, A., Kattenstroth, G., Hammer, R. E., Gottmann, K., & Sudhof, T. C. (2003). a -Neurexins couple Ca 2 1 channels to synaptic vesicle exocytosis. *Nature*, *423*, 939–948. https://doi.org/doi:10.1038/nature01755

Mitra, I., Tsang, K., Ladd-Acosta, C., Croen, L. A., Aldinger, K. A., Hendren, R. L., Traglia, M., Lavillaureix, A., Zaitlen, N., Oldham, M. C., Levitt, P., Nelson, S., Amaral, D. G., Herz-Picciotto, I., Fallin, M. D., & Weiss, L. A. (2016). Pleiotropic Mechanisms Indicated for Sex Differences in Autism. *PLoS Genetics*, *12*(11), 1–27. https://doi.org/10.1371/journal.pgen.1006425

Moayedi, M., & Davis, K. D. (2013). Theories of pain: From specificity to gate control. *Journal of Neurophysiology*, *109*(1), 5–12. https://doi.org/10.1152/jn.00457.2012

Moehring, F., Halder, P., Seal, R. P., & Stucky, C. L. (2018). Uncovering the Cells and Circuits of Touch in Normal and Pathological Settings. *Neuron*, *100*(2), 349–360. https://doi.org/10.1016/j.neuron.2018.10.019

Mogil, J. S. (2012). Pain genetics: Past, present and future. *Trends in Genetics*, *28*(6), 258–266. https://doi.org/10.1016/j.tig.2012.02.004

Mogil, J. S., Davis, K. D., & Derbyshire, S. W. (2010). The necessity of animal models in pain research. *Pain*, *151*(1), 12–17. https://doi.org/10.1016/j.pain.2010.07.015

Moncman, C. L., & Wang, K. (2002). Targeted disruption of nebulette protein expression alters cardiac myofibril assembly and function. *Experimental Cell Research*, *273*(2), 204–218. https://doi.org/10.1006/excr.2001.5423

Mondal, B., Jin, H., Kallappagoudar, S., Sedkov, Y., Martinez, T., Sentmanat, M. F., Poet, G. J., Li, C., Fan, Y., Pruett-Miller, S. M., & Herz, H. M. (2020). The histone deacetylase complex midac regulates a neurodevelopmental gene expression program to control neurite outgrowth. *ELife*, *9*, 1–29. https://doi.org/10.7554/eLife.57519

Morcom, L., Gobius, I., Marsh, A. P. L., Suárez, R., Lim, J. W. C., Bridges, C., Ye, Y., Fenlon, L. R., Zagar, Y., Douglass, A. M., Donahoo, A. L. S., Fothergill, T., Shaikh, S., Kozulin, P., Edwards, T. J., Cooper, H. M., Consortium, I. R. C., Sherr, E. H., Chédotal, A., … Richards, L. J. (2021). Dcc regulates astroglial development essential for telencephalic morphogenesis and corpus callosum formation. *ELife*, *10*, 1–35. https://doi.org/10.7554/ELIFE.61769

Morikawa, Y., Heallen, T., Leach, J., Xiao, Y., & Martin, J. F. (2017). Dystrophin-glycoprotein complex sequesters Yap to inhibit cardiomyocyte proliferation. *Nature*, *547*(7662), 227–231. https://doi.org/10.1038/nature22979

Morimura, N., Inoue, T., Katayama, K. ichi, & Aruga, J. (2006). Comparative analysis of structure, expression and PSD95-binding capacity of Lrfn, a novel family of neuronal transmembrane

proteins. *Gene*, *380*(2), 72–83. https://doi.org/10.1016/j.gene.2006.05.014

Moseley, G. L., & Vlaeyen, J. W. S. (2015). Beyond nociception: The imprecision hypothesis of chronic pain. *Pain*, *156*(1), 35–38. https://doi.org/10.1016/j.pain.0000000000000014

Moskvina, V., Schmidt, K. M., Vedernikov, A., Owen, M. J., Craddock, N., Holmans, P., & O'Donovan, M. C. (2012). Permutation-based approaches do not adequately allow for linkage disequilibrium in gene-wide multi-locus association analysis. *European Journal of Human Genetics*, *20*(8), 890–896. https://doi.org/10.1038/ejhg.2012.8

Motiwala, T., Kutay, H., Ghoshal, K., Bai, S., Seimiya, H., Tsuruo, T., Suster, S., Morrison, C., & Jacob, S. T. (2004). Protein tyrosine phosphatase receptor-type O (PTPRO) exhibits characteristics of a candidate tumor suppressor in human lung cancer. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(38), 13844–13849. https://doi.org/10.1073/pnas.0405451101

Mouraux, A., & Iannetti, G. D. (2018). The search for pain biomarkers in the human brain. *Brain*, *141*(12), 3290–3307. https://doi.org/10.1093/brain/awy281

Mu, J., Chaudhuri, K. R., Bielza, C., de Pedro-Cuesta, J., Larrañaga, P., & Martinez-Martin, P. (2017). Parkinson's disease subtypes identified from cluster analysis of motor and non-motor symptoms. *Frontiers in Aging Neuroscience*, *9*(SEP), 1–10. https://doi.org/10.3389/fnagi.2017.00301

Mulvey, M. R., Boland, E. G., Bouhassira, D., Freynhagen, R., Hardy, J., Hjermstad, M. J., Mercadante, S., Pérez, C., & Bennett, M. I. (2017). Neuropathic pain in cancer: Systematic review, performance of screening tools and analysis of symptom profiles. *British Journal of Anaesthesia*, *119*(4), 765–774. https://doi.org/10.1093/bja/aex175

Munce, S. E. P., & Stewart, D. E. (2007). Gender differences in depression and chronic pain conditions in a national epidemiologic survey. *Psychosomatics*, *48*(5), 394–399. https://doi.org/10.1176/appi.psy.48.5.394

Myers, S., Bottolo, L., Freeman, C., McVean, G., & Donnelly, P. (2005). Genetics: A fine-scale map of recombination rates and hotspots across the human genome. *Science*, *310*(5746), 321–324. https://doi.org/10.1126/science.1117196

Nachman, M. W., & Crowell, S. L. (2000). Estimate of the mutation rate per nucleotide in humans. *Genetics*, *156*(1), 297–304.

Nagano, M., Hoshino, D., Sakamoto, T., Kawasaki, N., Koshikawa, N., & Seiki, M. (2010). ZF21 protein regulates cell adhesion and motility. *Journal of Biological Chemistry*, *285*(27), 21013–21022. https://doi.org/10.1074/jbc.M110.106443

Nagasako, E. M., Oaklander, A. L., & Dworkin, R. H. (2003). Congenital insensitivity to pain: An update. *Pain*, *101*(3), 213–219. https://doi.org/10.1016/S0304-3959(02)00482-7

Nagel, M., Watanabe, K., Stringer, S., Posthuma, D., & Van Der Sluis, S. (2018). Item-level analyses reveal genetic heterogeneity in neuroticism. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-03242-8

Naik, R., Goodrich, G., Manian, D. V., Goyal, A., Rotella, K., Al-Shaikhly, T., & Joks, R. (2019). Association of Prescription Opioid Use for Chronic Pain Syndromes with Asthma and Allergy. *Journal of Allergy and Clinical Immunology*, *143*(2), AB223. https://doi.org/10.1016/j.jaci.2018.12.682

Nakayama, H., Kusumoto, C., Nakahara, M., Fujiwara, A., & Higashiyama, S. (2018). Semaphorin 3F and netrin-1: The novel function as a regulator of tumor microenvironment. *Frontiers in Physiology*, *9*(November), 1–11. https://doi.org/10.3389/fphys.2018.01662

Nam, J., Mah, W., & Kim, E. (2011). The SALM/Lrfn family of leucine-rich repeat-containing cell

adhesion molecules. *Seminars in Cell and Developmental Biology*, *22*(5), 492–498. https://doi.org/10.1016/j.semcdb.2011.06.005

Nelson, R. M., Pettersson, M. E., & Carlborg, Ö. (2013). A century after Fisher: Time for a new paradigm in quantitative genetics. *Trends in Genetics*, *29*(12), 669–676. https://doi.org/10.1016/j.tig.2013.09.006

Neogi, T. (2013). The epidemiology and impact of pain in osteoarthritis. *Osteoarthritis and Cartilage*, *21*(9), 1145–1153. https://doi.org/10.1016/j.joca.2013.03.018

Ng, A. C. Y., Eisenberg, J. M., Heath, R. J. W., Huett, A., Robinson, C. M., Nau, G. J., & Xavier, R. J. (2011). Human leucine-rich repeat proteins: a genome-wide bioinformatic categorization and functional analysis in innate immunity. *Proceedings of the National Academy of Sciences*, *108*(Supplement_1), 4631–4638. https://doi.org/10.1073/pnas.1000093107

Nicholasa, M., Vlaeyenb, J. W. S., Riefe, W., Barkee, A., Azizf, Q., Benolielg, R., Cohenh, M., Eversi, S., Maria Adele Giamberardinoj, A. G., Korwisie, B., Perrotl, S., Svenssonm, P., Wango, S.-J., Treedeq, R.-D., & Pain, T. I. T. for the C. of C. (2019). The IASP classification of chronic pain for ICD-11: Chronic primary pain. *Pain*, *160*(1), 53–59. https://doi.org/10.1097/j.pain.0000000000001365

Nicholl, B. I., Holliday, K. L., MacFarlane, G. J., Thomson, W., Davies, K. A., O'Neill, T. W., Bartfai, G., Boonen, S., Casanueva, F. F., Finn, J. D., Forti, G., Giwercman, A., Huhtaniemi, I. T., Kula, K., Punab, M., Silman, A. J., Vanderschueren, D., Wu, F. C. W., & McBeth, J. (2011). Association of HTR2A polymorphisms with chronic widespread pain and the extent of musculoskeletal pain: Results from two population-based cohorts. *Arthritis and Rheumatism*, *63*(3), 810–818. https://doi.org/10.1002/art.30185

Nicholl, B. I., Mackay, D., Cullen, B., Martin, D. J., Ul-Haq, Z., Mair, F. S., Evans, J., McIntosh, A. M., Gallagher, J., Roberts, B., Deary, I. J., Pell, J. P., & Smith, D. J. (2014). Chronic multisite pain in major depression and bipolar disorder: cross-sectional study of 149,611 participants in UK Biobank. *BMC Psychiatry*. https://doi.org/10.1186/s12888-014-0350-4

Nielsen, C., Knudsen, G., & Steingrímsdóttir, Ó. (2012). Twin studies of pain. *Clinical Genetics*, *82*(4), 331–340. https://doi.org/10.1111/j.1399-0004.2012.01938.x

Nishimune, H., Sanes, J. R., & Carlson, S. S. (2004). A synaptic laminin-calcium channel interaction organizes active zones in motor nerve terminals. *Nature*, *432*(7017), 580–587. https://doi.org/10.1038/nature03112

Norton, C., Czuber-Dochan, W., Artom, M., Sweeney, L., & Hart, A. (2017). Systematic review: interventions for abdominal pain management in inflammatory bowel disease. *Alimentary Pharmacology and Therapeutics*, *46*(2), 115–125. https://doi.org/10.1111/apt.14108

Nykjaer, A., Lee, R., Teng, K. K., Jansen, P., Madsen, P., Morten S. Nielsen, C. J., Kliemannel, M., Schwarz, E., Willnow, T. E., Hempstead, B. L., Petersen, C. M., & 1Department. (2004). Sortilin is essential for proNGF- induced neuronal cell death. *Nature*, *427*, 843–848. https://doi.org/10.1038/nature02289

O'Connor, A. B., Schwid, S. R., Herrmann, D. N., Markman, J. D., & Dworkin, R. H. (2008). Pain associated with multiple sclerosis: Systematic review and proposed classification. *Pain*, *137*(1), 96–111. https://doi.org/10.1016/j.pain.2007.08.024

Ogoh, H., Yamagata, K., Nakao, T., Sandell, L. L., Yamamoto, A., Yamashita, A., Tanga, N., Suzuki, M., Abe, T., Kitabayashi, I., Watanabe, T., & Sakai, D. (2017). Mllt10 knockout mouse model reveals critical role of Af10-dependent H3K79 methylation in midfacial development. *Scientific Reports*, *7*(1), 1–18. https://doi.org/10.1038/s41598-017-11745-5

Ohayon, M. M., & Schatzberg, A. F. (2010). Chronic pain and major depressive disorder in the general population. *Journal of Psychiatric Research*, *44*(7), 454–461.

https://doi.org/10.1016/j.jpsychires.2009.10.013

Okajima, D., Kudo, G., & Yokota, H. (2010). Brain-specific angiogenesis inhibitor 2 (BAI2) may be activated by proteolytic processing. *Journal of Receptors and Signal Transduction*, *30*(3), 143–153. https://doi.org/10.3109/10799891003671139

Okifuji, A., & Hare, B. D. (2015). The association between chronic pain and obesity. *Journal of Pain Research*, *8*, 399–408. https://doi.org/10.2147/JPR.S55598

Okuda, M., Horn, H. F., Tarapore, P., Tokuyama, Y., Smulian, A. G., Chan, P. K., Knudsen, E. S., Hofmann, I. A., Snyder, J. D., Bove, K. E., & Fukasawa, K. (2000). Nucleophosmin/B23 is a target of CDK2/cyclin E in centrosome duplication. *Cell*, *103*(1), 127–140. https://doi.org/10.1016/S0092-8674(00)00093-3

Olbert, C. M., Gala, G. J., & Tupler, L. A. (2014). Quantifying heterogeneity attributable to polythetic diagnostic criteria: Theoretical framework and empirical application. *Journal of Abnormal Psychology*, *123*(2), 452–462. https://doi.org/10.1037/a0036068

Ossipov, M. H., Morimura, K., & Porreca, F. (2014). Descending pain modulation and chronification of pain. *Current Opinion in Supportive and Palliative Care*, *8*(2), 143–151. https://doi.org/10.1097/SPC.0000000000000055

Outcalt, S. D., Kroenke, K., Krebs, E. E., Chumbler, N. R., Wu, J., Yu, Z., & Bair, M. J. (2015). Chronic pain and comorbid mental health conditions: independent associations of posttraumatic stress disorder and depression with pain, disability, and quality of life. *Journal of Behavioral Medicine*, *38*(3), 535–543. https://doi.org/10.1007/s10865-015-9628-3

Oyama, N., Chan, I., Neill, S. M., Hamada, T., South, A. P., Wessagowit, V., Wojnarowska, F., D'Cruz, D., Hughes, G. J., Black, M. M., & McGrath, J. A. (2003). Autoantibodies to extracellular matrix protein 1 in lichen sclerosus. *Lancet*, *362*(9378), 118–123. https://doi.org/10.1016/S0140-6736(03)13863-9

Paley, C. A., & Johnson, M. I. (2016). Physical activity to reduce systemic inflammation associated with chronic pain and obesity a narrative review. *Clinical Journal of Pain*, *32*(4), 365–370. https://doi.org/10.1097/AJP.0000000000000258

Palmer, K. T., Walsh, K., Bendall, H., Cooper, C., & Coggon, D. (2000). Back pain in Britain: Comparison of two prevalence surveys at an interval of 10 years. *British Medical Journal*, *320*(7249), 1577–1578.

Paratcha, G., Ledda, F., & Ibáñez, C. F. (2003). The neural cell adhesion molecule NCAM is an alternative signaling receptor for GDNF family ligands. *Cell*, *113*(7), 867–879. https://doi.org/10.1016/S0092-8674(03)00435-5

Parida, J. R., Misra, D. P., & Agarwal, V. (2015). Urbach-Wiethe syndrome. *BMJ Case Reports*, *2015*, 1–3. https://doi.org/10.1136/bcr-2015-212443

Park, U. H., Yoon, S. K., Park, T., Kim, E. J., & Um, S. J. (2011). Additional sex comb-like (ASXL) proteins 1 and 2 play opposite roles in adipogenesis via reciprocal regulation of peroxisome proliferator-activated receptor γ. *Journal of Biological Chemistry*, *286*(2), 1354–1363. https://doi.org/10.1074/jbc.M110.177816

Patnala, R., Clements, J., & Batra, J. (2013). Candidate gene association studies: A comprehensive guide to useful in silico tools. *BMC Genetics*, *14*. https://doi.org/10.1186/1471-2156-14-39

Patten, S., Williams, J., Lavorato, D., Wang, J., Bulloch, A., & Sajobi, T. (2016). The association between major depression prevalence and sex becomes weaker with age. *Social Psychiatry and Psychiatric Epidemiology*, *51*. https://doi.org/10.1007/s00127-015-1166-3

Patton, G. C., Olsson, C., Bond, L., Toumbourou, J. W., Carlin, J. B., Hemphill, S. A., & Catalano, R. F. (2008). Predicting female depression across puberty: A two-nation longitudinal study.

Journal of the American Academy of Child and Adolescent Psychiatry, *47*(12), 1424–1432. https://doi.org/10.1097/CHI.0b013e3181886ebe

Peddareddygari, L. R., Oberoi, K., & Grewal, R. P. (2014). Congenital Insensitivity to Pain: A Case Report and Review of the Literature. *Case Reports in Neurological Medicine*, *2014*, 1–4. https://doi.org/10.1155/2014/141953

Peirs, C., & Seal, R. P. (2016). Neural circuits for pain: Recent advances and current views. *Science*, *354*(6312), 578–583. https://doi.org/10.1083/jcb.200511140

Pellerin, L. (2003). Lactate as a pivotal element in neuron-glia metabolic cooperation. *Neurochemistry International*, *43*(4–5), 331–338. https://doi.org/10.1016/S0197-0186(03)00020-2

Penninx, B. W. J. H., Milaneschi, Y., Lamers, F., & Vogelzangs, N. (2013). Understanding the somatic consequences of depression: Biological mechanisms and the role of depression symptom profile. *BMC Medicine*, *11*(1), 1. https://doi.org/10.1186/1741-7015-11-129

Peripheral Neuropathy Association. (1993). Quantitative sensory testing: a consensus report from the Peripheral Neuropathy Association. *Neurology*, *43*(5), 1050–1052.

Perl, E. R. (2007). Ideas about pain, a historical review. *Nature Reviews Neuroscience*, *8*(January), 71–80.

Peters, J. M. (2006). The anaphase promoting complex/cyclosome: A machine designed to destroy. *Nature Reviews Molecular Cell Biology*, *7*(9), 644–656. https://doi.org/10.1038/nrm1988

Peters, M. J., Broer, L., Willemen, H. L. D. M., Eiriksdottir, G., Hocking, L. J., Holliday, K. L., Horan, M. A., Meulenbelt, I., Neogi, T., Popham, M., Schmidt, C. O., Soni, A., Valdes, A. M., Amin, N., Dennison, E. M., Eijkelkamp, N., Harris, T. B., Hart, D. J., Hofman, A., … van Meurs, J. B. J. (2013). Genome-wide association study meta-analysis of chronic widespread pain: evidence for involvement of the 5p15.2 region. *Annals of the Rheumatic Diseases*. https://doi.org/10.1136/annrheumdis-2012-201742

Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C. Y., Popejoy, A. B., Periyasamy, S., Lam, M., Iyegbe, C., Strawbridge, R. J., Brick, L., Carey, C. E., Martin, A. R., Meyers, J. L., Su, J., Chen, J., Edwards, A. C., Kalungi, A., Koen, N., Majara, L., … Duncan, L. E. (2019). Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell*, *179*(3), 589–603. https://doi.org/10.1016/j.cell.2019.08.051

Petryshen, T. L., Middleton, F. A., Tahl, A. R., Rockwell, G. N., Purcell, S., Aldinger, K. A., Kirby, A., Morley, C. P., McGann, L., Gentile, K. L., Waggoner, S. G., Medeiros, H. M., Carvalho, C., Macedo, A., Albus, M., Maier, W., Trixler, M., Eichhammer, P., Schwab, S. G., … Sklar, P. (2005). Genetic investigation of chromosome 5q GABAA receptor subunit genes in schizophrenia. *Molecular Psychiatry*, *10*(12), 1074–1088. https://doi.org/10.1038/sj.mp.4001739

Pfeifer, I., Elsby, R., Fernandez, M., Faria, P. A., Nussenzveig, D. R., Lossos, I. S., Fontoura, B. M. A., Martin, W. D., & Barber, G. N. (2008). NFAR-1 and -2 modulate translation and are required for efficient host defense. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(11), 4173–4178. https://doi.org/10.1073/pnas.0711222105

Phifer, J., Skelton, K., Weiss, T., Schwartz, A. C., Wingo, A., Gillespie, C. F., Sands, L. A., Sayyar, S., Bradley, B., Jovanovic, T., & Ressler, K. J. (2011). Pain symptomatology and pain medication use in civilian PTSD. *Pain*, *152*(10), 2233–2240. https://doi.org/10.1016/j.pain.2011.04.019

Phillips, K., & Clauw, D. J. (2011). Central pain mechanisms in chronic pain states - Maybe it is all in their head. *Best Practice and Research: Clinical Rheumatology*, *25*(2), 141–154. https://doi.org/10.1016/j.berh.2011.02.005

Phillips, K., & Clauw, D. J. (2013). Central pain mechanisms in the rheumatic diseases: Future directions. *Arthritis and Rheumatism*, *65*(2), 291–302. https://doi.org/10.1002/art.37739

Piché, J., Van Vliet, P. P., Pucéat, M., & Andelfinger, G. (2019). The expanding phenotypes of cohesinopathies: one ring to rule them all! *Cell Cycle*, *18*(21), 2828–2848. https://doi.org/10.1080/15384101.2019.1658476

Pickrell, J. K. (2014). Joint Analysis of Functional Genomic Data and Genome-wide Association Studies of 18 Human Traits. *The American Journal of Human Genetics*, *94*(4), 559–573. https://doi.org/10.1016/j.ajhg.2014.03.004

Pierre, K., Magistretti, P. J., & Pellerin, L. (2002). MCT2 is a major neuronal monocarboxylate transporter in the adult mouse brain. *J Cereb Blood Flow Metab*, *22*(5), 586–595. https://doi.org/10.1097/00004647-200205000-00010

Pinheiro, M. B., Ferreira, M. L., Refshauge, K., Colodro-Conde, L., Carrillo, E., Hopper, J. L., Ordoñana, J. R., & Ferreira, P. H. (2015). Genetics and the environment affect the relationship between depression and low back pain: A co-twin control study of Spanish twins. *Pain*, *156*(3), 496–503. https://doi.org/10.1097/01.j.pain.0000460330.56256.25

Pinho-Ribeiro, F. A., Verri, W. A., & Chiu, I. M. (2017). Nociceptor Sensory Neuron–Immune Interactions in Pain and Inflammation. *Trends in Immunology*, *38*(1), 5–19. https://doi.org/10.1016/j.it.2016.10.001

Polderman, T. J. C., Benyamin, B., De Leeuw, C. A., Sullivan, P. F., Van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*(7), 702–709. https://doi.org/10.1038/ng.3285

Poleshuck, E. L., & Green, C. R. (2008). Socioeconomic disadvantage and pain. *Pain*, *136*(3), 235–238. https://doi.org/10.1016/j.pain.2008.04.003

Poquérusse, J., Pastore, L., Dellantonio, S., & Esposito, G. (2018). Alexithymia and autism spectrum disorder: A complex relationship. *Frontiers in Psychology*, *9*(JUL), 1–10. https://doi.org/10.3389/fpsyg.2018.01196

Portin, P., & Wilkins, A. (2017). The Evolving Definition of the Term "Gene." *Genetics*, *205*, 1353–1364. https://doi.org/10.1525/bio.2009.59.11.3

Povysil, G., Petrovski, S., Hostyk, J., Aggarwal, V., Allen, A. S., & Goldstein, D. B. (2019). Rare-variant collapsing analyses for complex traits: guidelines and applications. *Nature Reviews Genetics*, *20*(12), 747–759. https://doi.org/10.1038/s41576-019-0177-4

Pozzilli, C., & Pugliatti, M. (2015). An overview of pregnancy-related issues in patients with multiple sclerosis. *European Journal of Neurology*, *22*, 34–39. https://doi.org/10.1111/ene.12797

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. https://doi.org/10.1038/ng1847

Prictor, M., Teare, H. J. A., & Kaye, J. (2018). Equitable Participation in Biobanks: The Risks and Benefits of a "Dynamic Consent" Approach. *Frontiers in Public Health*, *6*(September), 1–6. https://doi.org/10.3389/fpubh.2018.00253

Prins, B. P., Abbasi, A., Wong, A., Vaez, A., Nolte, I., Franceschini, N., Stuart, P. E., Achury, J. G., Mistry, V., Bradfield, J. P., Valdes, A. M., Bras, J., Shatunov, A., Consortium, P., Stroke, I., Consortium, G., Sclerosis, S., Oa, T., Consortium, D., … Alizadeh, B. Z. (2016). *Investigating the Causal Relationship of C-Reactive Protein with 32 Complex Somatic and Psychiatric Outcomes : A Large-Scale Cross-Consortium Mendelian Randomization Study*. 1–29. https://doi.org/10.1371/journal.pmed.1001976

Pritchard, J. K., & Przeworki, M. (2001). Linkage Disequilibrium in Humans: Models and Data. *The American Journal of Human Genetics*, *69*(1), 124–137. https://doi.org/10.1086/321272

Pritchard, J. K., Stephens, M., Rosenberg, N. A., & Donnelly, P. (2002). Association Mapping in Structured Populations. *The American Journal of Human Genetics*, *67*(1), 170–181. https://doi.org/10.1086/302959

Purcell, S. M., Moran, J. L., Fromer, M., Ruderfer, D., Solovieff, N., Roussos, P., O'Dushlaine, C., Chambert, K., Bergen, S. E., Kähler, A., Duncan, L., Stahl, E., Genovese, G., Fernández, E., Collins, M. O., Komiyama, N. H., Choudhary, J. S., Magnusson, P. K. E., Banks, E., … Sklar, P. (2014). A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*, *506*(7487), 185–190. https://doi.org/10.1038/nature12975

Purcell, S., Neale, B., Todd-brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., Bakker, P. I. W. De, Daly, M. J., & Sham, P. C. (2007). *REPORT PLINK : A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses*. *81*(September), 559–575. https://doi.org/10.1086/519795

Quinn, M. A., & Cidlowski, J. A. (2016). Endogenous hepatic glucocorticoid receptor signaling coordinates sex-biased inflammatory gene expression. *FASEB Journal*, *30*(2), 971–982. https://doi.org/10.1096/fj.15-278309

Rabbani, B., Tekin, M., & Mahdieh, N. (2014). The promise of whole-exome sequencing in medical genetics. *Journal of Human Genetics*, *59*(1), 5–15. https://doi.org/10.1038/jhg.2013.114

Rahmioglu, N., MacGregor, S., Drong, A. W., Hedman, Å. K., Harris, H. R., Randall, J. C., Prokopenko, I., Nyholt, D. R., Morris, A. P., Montgomery, G. W., Missmer, S. A., Lindgren, C. M., & Zondervan, K. T. (2015). Genome-wide enrichment analysis between endometriosis and obesity-related traits reveals novel susceptibility loci. *Human Molecular Genetics*, *24*(4), 1185–1199. https://doi.org/10.1093/hmg/ddu516

Raja, S. N., Carr, D. B., Cohen, M., Finnerup, N. B., Flor, H., Gibson, S., Keefe, F. J., Mogil, J. S., Ringkamp, M., Sluka, K. A., Song, X.-J., Stevens, B., Sullivan, M. D., Tutelmano, P. R., & Vader, K. (2020). The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. *PAIN*, *161*(9), 1976–1982. https://doi.org/10.1097/pr9.0000000000000643

Rawlik, K., Canela-Xandri, O., & Tenesa, A. (2016). Evidence for sex-specific genetic architectures across a spectrum of human complex traits. *Genome Biology*, *17*(1), 1–8. https://doi.org/10.1186/s13059-016-1025-x

Ray-Griffith, S. L., Wendel, M. P., Stowe, Z. N., & Magann, E. F. (2018). Chronic pain during pregnancy: a review of the literature. *International Journal of Women's Health*, *10*, 153–164. http://www.ncbi.nlm.nih.gov/pubmed/29692634%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5901203

Ray, P. R., Khan, J., Wangzhou, A., Tavares-Ferreira, D., Akopian, A. N., Dussor, G., & Price, T. J. (2019). Transcriptome analysis of the human tibial nerve identifies sexually dimorphic expression of genes involved in pain, inflammation, and neuro-immunity. *Frontiers in Molecular Neuroscience*, *12*(March). https://doi.org/10.3389/fnmol.2019.00037

Rebolledo-Jaramillo, B., & Ziegler, A. (2018). Teneurins: An Integrative Molecular, Functional, and Biomedical Overview of Their Role in Cancer. *Frontiers in Neuroscience*, *12*(December), 1–20. https://doi.org/10.3389/fnins.2018.00937

Reckziegel, D., Vachon-Presseau, E., Petre, B., Schnitzer, T. J., Baliki, M. N., & Apkarian, A. V. (2019). Deconstructing biomarkers for chronic pain: context- and hypothesis-dependent biomarker types in relation to chronic pain. *Pain*, *160*(Suppl 1), S37–S48. https://doi.org/10.1097/j.pain.0000000000001529

Regier, D. A., Kuhl, E. A., & Kupfer, D. J. (2013). The DSM-5: Classification and criteria changes. *World Psychiatry*, *12*(2), 92–98. https://doi.org/10.1002/wps.20050

Ren, K., & Dubner, R. (2010). Interactions between the immune and nervous systems in pain. *Nature Medicine*, *16*(11), 1267–1276. https://doi.org/10.1038/nm.2234

Ren, X. R., Ming, G. L., Xie, Y., Hong, Y., Sun, D. M., Zhao, Z. Q., Feng, Z., Wang, Q., Shim, S., Chen, Z. F., Song, H. J., Mei, L., & Xiong, W. C. (2004). Focal adhesion kinase in netrin-1 signaling. *Nature Neuroscience*, *7*(11), 1204–1212. https://doi.org/10.1038/nn1330

Renvoise, B., Parker, R. L., Yang, D., Bakowska, J. C., Hurley, J. H., & Blackstone, C. (2010). SPG20 Protein Spartin Is Recruited to Midbodies by ESCRT-III Protein Ist1 and Participates in Cytokinesis. *Molecular Biology of the Cell*, *20*, 4524–4530. https://doi.org/10.1091/mbc.E09

Retterer, K., Juusola, J., Cho, M. T., Vitazka, P., Millan, F., Gibellini, F., Vertino-Bell, A., Smaoui, N., Neidich, J., Monaghan, K. G., McKnight, D., Bai, R., Suchy, S., Friedman, B., Tahiliani, J., Pineda-Alvarez, D., Richard, G., Brandt, T., Haverfield, E., … Bale, S. (2016). Clinical application of whole-exome sequencing across clinical indications. *Genetics in Medicine*, *18*(7), 696–704. https://doi.org/10.1038/gim.2015.148

Reuter, M. S., Tawamie, H., Buchert, R., Gebril, O. H., Froukh, T., Thiel, C., Uebe, S., Ekici, A. B., Krumbiegel, M., Zweier, C., Hoyer, J., Eberlein, K., Bauer, J., Scheller, U., Strom, T. M., Hoffjan, S., Abdelraouf, E. R., Meguid, N. A., Abboud, A., … Jamra, R. A. (2017). Diagnostic yield and novel candidate genes by exome sequencing in 152 consanguineous families with neurodevelopmental disorders. *JAMA Psychiatry*, *74*(3), 293–299. https://doi.org/10.1001/jamapsychiatry.2016.3798

Rey, R. (1995). *The History of Pain*. Harvard University Press. https://go.exlibris.link/rLt9TN6s

Richter, D., Grün, R., Joannes-Boyau, R., Steele, T. E., Amani, F., Rué, M., Fernandes, P., Raynal, J. P., Geraads, D., Ben-Ncer, A., Hublin, J. J., & McPherron, S. P. (2017). The age of the hominin fossils from Jebel Irhoud, Morocco, and the origins of the Middle Stone Age. *Nature*, *546*(7657), 293–296. https://doi.org/10.1038/nature22335

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K. H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., Collier, D. A., Huang, H., Pers, T. H., Agartz, I., Agerbo, E., Albus, M., Alexander, M., Amin, F., Bacanu, S. A., Begemann, M., Belliveau, R. A., … O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*(7510), 421–427. https://doi.org/10.1038/nature13595

Rito, T., Richards, M. B., Fernandes, V., Alshamali, F., Cerny, V., Pereira, L., & Soares, P. (2013). The first modern human dispersals across Africa. *PLoS ONE*, *8*(11), 1–16. https://doi.org/10.1371/journal.pone.0080031

Rito, T., Vieira, D., Silva, M., Conde-Sousa, E., Pereira, L., Mellars, P., Richards, M. B., & Soares, P. (2019). A dispersal of Homo sapiens from southern to eastern Africa immediately preceded the out-of-Africa migration. *Scientific Reports*, *9*(1), 1–10. https://doi.org/10.1038/s41598-019-41176-3

Roberts, J. S., Gornick, M. C., Carere, D. A., Uhlmann, W. R., Ruffin, M. T., & Green, R. C. (2017). Direct-to-consumer genetic testing: User motivations, decision making, and perceived utility of results. *Public Health Genomics*, *20*(1), 36–45. https://doi.org/10.1159/000455006

Roberts, S. H., Hughes, H. E., Davies, S. J., & Meredith, A. L. (1991). Bilateral split hand and split foot malformation in a boy with a de novo interstitial deletion of 7q21.3. *Journal of Medical Genetics*, *28*, 479–481.

Rock, K. L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., & Goldberg, A. L. (1994). Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell*, *78*(5), 761–771.

https://doi.org/10.1016/S0092-8674(94)90462-6

Rodriguez-Acevedo, A. J., Ferreira, M. A., Benton, M. C., Carless, M. A., Goring, H. H., Curran, J. E., Blangero, J., Lea, R. A., & Griffiths, L. R. (2015). Common polygenic variation contributes to risk of migraine in the Norfolk Island population. *Human Genetics*, *134*(10), 1079–1087. https://doi.org/10.1007/s00439-015-1587-9

Rollman, G. B., & Lautenbacher, S. (2001). Sex differences in musculoskeletal pain. *Clinical Journal of Pain*, *17*(1), 20–24. https://doi.org/10.1097/00002508-200103000-00004

Roosing, S., Romani, M., Isrie, M., Rosti, R. O., Micalizzi, A., Musaev, D., Mazza, T., Al-gazali, L., Altunoglu, U., Boltshauser, E., D'Arrigo, S., Keersmaecker, B. De, Kayserili, H., Brandenberger, S., Kraoua, I., Mark, P. R., McKanna, T., Keirsbilck, J. Van, Moerman, P., … Valente, E. M. (2016). Mutations in cep120 cause joubert syndrome as well as complex ciliopathy phenotypes. *Journal of Medical Genetics*, *53*(9), 608–615. https://doi.org/10.1136/jmedgenet-2016-103832

Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A., & Feldman, M. W. (2002). Genetic Structure of Human Populations. *Science*, *298*(December), 2381–2385.

Roth, T. (2007). Insomnia: Definition, prevalence, etiology, and consequences. *Journal of Clinical Sleep Medicine*, *3*(5 SUPPL.), 3–6. https://doi.org/10.5664/jcsm.26929

Rothman, K. J., Gallacher, J. E. J., & Hatch, E. E. (2013). Why representativeness should be avoided. *International Journal of Epidemiology*, *42*(4), 1012–1014. https://doi.org/10.1093/ije/dys223

Ruderfer, D. M., Ripke, S., McQuillin, A., Boocock, J., Stahl, E. A., Pavlides, J. M. W., Mullins, N., Charney, A. W., Ori, A. P. S., Loohuis, L. M. O., Domenici, E., Di Florio, A., Papiol, S., Kalman, J. L., Trubetskoy, V., Adolfsson, R., Agartz, I., Agerbo, E., Akil, H., … Kendler, K. S. (2018). Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell*, *173*(7), 1705-1715.e16. https://doi.org/10.1016/j.cell.2018.05.046

Saeki, K., Miura, Y., Aki, D., Kurosaki, T., & Yoshimura, A. (2003). The B cell-specific major raft protein, Raftlin, is necessary for the integrity of lipid raft and BCR signal transduction. *EMBO Journal*, *22*(12), 3015–3026. https://doi.org/10.1093/emboj/cdg293

Saito, F., Moore, S. A., Barresi, R., Henry, M. D., Messing, A., Ross-Barta, S. E., Cohn, R. D., Williamson, R. A., Sluka, K. A., Sherman, D. L., Brophy, P. J., Schmelzer, J. D., Low, P. A., Wrabetz, L., Feltri, M. L., & Campbell, K. P. (2003). Unique role of dystroglycan in peripheral nerve myelination, nodal structure, and sodium channel stabilization. *Neuron*, *38*(5), 747–758. https://doi.org/10.1016/S0896-6273(03)00301-5

Sakamoto, O., Iwama, A., Amitani, R., Takehara, T., Yamaguchi, N., Yamamoto, T., Masuyama, K., Yamanaka, T., Ando, M., & Suda, T. (1997). Role of macrophage-stimulating protein and its receptor, RON tyrosine kinase, in ciliary motility. *Journal of Clinical Investigation*, *99*(4), 701–709. https://doi.org/10.1172/JCI119214

SAMHSA. (2018). Key substance use and mental health indicators in the United States: Results from the 2015 National Survey on Drug Use and Health. Publication No. SMA 16-4984, NSDUH Series H-51. *Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration*, 1–97.

Samuels-Lev, Y., O'Connor, D. J., Bergamaschi, D., Trigiante, G., Hsieh, J. K., Zhong, S., Campargue, I., Naumovski, L., Crook, T., & Lu, X. (2001). ASPP proteins specifically stimulate the apoptotic function of p53. *Molecular Cell*, *8*(4), 781–794. https://doi.org/10.1016/S1097-2765(01)00367-7

Sanchis, M. N., Lluch, E., Nijs, J., Struyf, F., & Kangasperko, M. (2015). The role of central sensitization in shoulder pain: A systematic literature review. *Seminars in Arthritis and*

Rheumatism, *44*(6), 710–716. https://doi.org/10.1016/j.semarthrit.2014.11.002

Sanjuá, J., Tolosa, A., Gonzá Lez A, J. C., Aguilar, E. J., Pé Rez-Tur, J., Ná Jera B , María, C., Moltó, D., & De Frutos, R. (2006). Association between FOXP2 polymorphisms and schizophrenia with auditory hallucinations. In *Psychiatric Genetics* (Vol. 16).

Santos-Eggimann, B., Wietlisbach, V., Rickenbach, M., Paccaud, F., & Gutzwiller, F. (2000). One-year prevalence of low back pain in two Swiss regions. Estimates from the population participating in the 1992-1993 MONICA project. *Spine*, *25*(19), 2473–2479. https://doi.org/10.1097/00007632-200010010-00009

Sanz, L., Moscat, J., & Diaz-Meco, M. T. (1995). Molecular characterization of a novel transcription factor that controls stromelysin expression. *Molecular and Cellular Biology*, *15*(6), 3164–3170. https://doi.org/10.1128/mcb.15.6.3164

Satizabal, C. L., Adams, H. H. H., Hibar, D. P., White, C. C., Knol, M. J., Stein, J. L., Scholz, M., Sargurupremraj, M., Jahanshad, N., Roshchupkin, G. V., Smith, A. V., Bis, J. C., Jian, X., Luciano, M., Hofer, E., Teumer, A., van der Lee, S. J., Yang, J., Yanek, L. R., … Ikram, M. A. (2019). Genetic architecture of subcortical brain structures in 38,851 individuals. *Nature Genetics*, *51*(11), 1624–1636. https://doi.org/10.1038/s41588-019-0511-y

Satten, G. A., Flanders, W. D., & Yang, Q. (2002). Accounting for Unmeasured Population Substructure in Case-Control Studies of Genetic Association Using a Novel Latent-Class Model. *The American Journal of Human Genetics*, *68*(2), 466–477. https://doi.org/10.1086/318195

Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., De Leeuw, C. A., Nagel, M., Awasthi, S., Barr, P. B., Coleman, J. R. I., Grasby, K. L., Hammerschlag, A. R., Kaminski, J. A., Karlsson, R., Krapohl, E., Lam, M., Nygaard, M., Reynolds, C. A., Trampush, J. W., … Posthuma, D. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*, *50*(7), 912–919. https://doi.org/10.1038/s41588-018-0152-6

Schaid, D. J., Chen, W., & Larson, N. B. (2018). From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nature Reviews Genetics*, *19*(8), 491–504. https://doi.org/10.1038/s41576-018-0016-z

Scherer, S. W., Lee, C., Birney, E., Altshuler, D. M., Eichler, E. E., Carter, N. P., Hurles, M. E., & Feuk, L. (2007). Challenges and standards in integrating surveys of structural variation. *Nature Genetics*, *39*(7S), S7–S15. https://doi.org/10.1038/ng2093

Schlebusch, C. M., Malmström, H., Günther, T., Sjödin, P., Coutinho, A., Edlund, H., Munters, A. R., Vicente, M., Steyn, M., Soodyall, H., Lombard, M., & Jakobsson, M. (2017). Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. *Science*, *358*(6363), 652–655. https://doi.org/10.1126/science.aao6266

Schmidt, A. F., Swerdlow, D. I., Holmes, M. V., Patel, R. S., Fairhurst-Hunter, Z., Lyall, D. M., Hartwig, F. P., Horta, B. L., Hyppönen, E., Power, C., Moldovan, M., van Iperen, E., Hovingh, G. K., Demuth, I., Norman, K., Steinhagen-Thiessen, E., Demuth, J., Bertram, L., Liu, T., … Sattar, N. (2017). PCSK9 genetic variants and risk of type 2 diabetes: a mendelian randomisation study. *The Lancet Diabetes and Endocrinology*, *5*(2), 97–105. https://doi.org/10.1016/S2213-8587(16)30396-5

Schmidts, M., Vodopiutz, J., Christou-Savina, S., Cortés, C. R., McInerney-Leo, A. M., Emes, R. D., Arts, H. H., Tüysüz, B., D'Silva, J., Leo, P. J., Giles, T. C., Oud, M. M., Harris, J. A., Koopmans, M., Marshall, M., Elçioglu, N., Kuechler, A., Bockenhauer, D., Moore, A. T., … Mitchison, H. M. (2013). Mutations in the gene encoding IFT dynein complex component WDR34 cause jeune asphyxiating thoracic dystrophy. *American Journal of Human Genetics*, *93*(5), 932–944. https://doi.org/10.1016/j.ajhg.2013.10.003

Schneider, M. A., Christopoulos, P., Muley, T., Warth, A., Klingmueller, U., Thomas, M., Herth, F. J. F., Dienemann, H., Mueller, N. S., Theis, F., & Meister, M. (2017). AURKA, DLGAP5, TPX2, KIF11 and CKAP5: Five specific mitosis-associated genes correlate with poor prognosis for non-small cell lung cancer patients. *International Journal of Oncology*, *50*(2), 365–372. https://doi.org/10.3892/ijo.2017.3834

Schuettengruber, B., Bourbon, H. M., Di Croce, L., & Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell*, *171*(1), 34–57. https://doi.org/10.1016/j.cell.2017.08.002

Schug, S. A., & Bruce, J. (2017). *Risk stratification for the development of chronic postsurgical pain*. *2*.

Schwabe, I., Milaneschi, Y., Gerring, Z., Sullivan, P. F., Schulte, E., Suppli, N. P., Thorp, J. G., Derks, E. M., & Middeldorp, C. M. (2019). Unraveling the genetic architecture of major depressive disorder: merits and pitfalls of the approaches used in genome-wide association studies. *Psychological Medicine*, 1–11. https://doi.org/10.1017/S0033291719002502

Schweinhardt, P., & Bushnell, M. C. (2010). Pain imaging in health and disease — how far have we come ? *Journal of Clinical Investigation*, *120*(11), 3788–3797. https://doi.org/10.1172/JCI43498.3788

Seidler, Z. E., Dawes, A. J., Rice, S. M., Oliffe, J. L., & Dhillon, H. M. (2016). The role of masculinity in men's help-seeking for depression: A systematic review. *Clinical Psychology Review*, *49*, 106–118. https://doi.org/10.1016/j.cpr.2016.09.002

Shabalin, A. A. (2012). Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. *Bioinformatics*, *28*(10), 1353–1358. https://doi.org/10.1093/bioinformatics/bts163

Sharma, N., Akhade, A. S., & Qadri, A. (2016). Src kinases central to T-cell receptor signaling regulate TLR-activated innate immune responses from human T cells. *Innate Immunity*, *22*(3), 238–244. https://doi.org/10.1177/1753425916632305

Sharp, L. K., & Lipsky, M. S. (2002). Screening for depression across the lifespan: A review of measures for use in primary care settings. *American Family Physician*, *66*(6), 1001–1008.

Shen, W., Tu, Y., Gollub, R. L., Ortiz, A., Napadow, V., Yu, S., Wilson, G., Park, J., Lang, C., Jung, M., Gerber, J., Mawla, I., Chan, S., Wasan, A. D., Edwards, R. R., Kaptchuk, T., Li, S., & Rosen, B. (2019). NeuroImage : Clinical Visual network alterations in brain functional connectivity in chronic low back pain : A resting state functional connectivity and machine learning study. *NeuroImage: Clinical*, *22*(March), 101775. https://doi.org/10.1016/j.nicl.2019.101775

Sheng, J., Liu, S., Wang, Y., Cui, R., & Zhang, X. (2017). The Link between Depression and Chronic Pain: Neural Mechanisms in the Brain. *Neural Plasticity*, *2017*. https://doi.org/10.1155/2017/9724371

Sheskin, D. J. (2000). *Parametric and non parametric statistical procedures* (2nd ed.). Chapman & Hall.

Shipherd, J. C., Keyes, M., Jovanovic, T., Ready, D. J., Baltzell, D., Worley, V., Gordon-Brown, V., Hayslett, C., & Duncan, E. (2007). Veterans seeking treatment for posttraumatic stress disorder: What about comorbid chronic pain? *Journal of Rehabilitation Research and Development*, *44*(2), 153–165. https://doi.org/10.1682/JRRD.2006.06.0065

Shrestha, S., Wilmeth, L. J., Eyer, J., & Shuster, C. B. (2012). PRC1 controls spindle polarization and recruitment of cytokinetic factors during monopolar cytokinesis. *Molecular Biology of the Cell*, *23*(7), 1196–1207. https://doi.org/10.1091/mbc.E11-12-1008

Simuni, T., & Sethi, K. (2008). Nonmotor manifestations of Parkinson's disease. *Annals of Neurology*, *64*(SUPPL. 2), 65–80. https://doi.org/10.1002/ana.21505

Singareddy, R., Vgontzas, A. N., Fernandez-Mendoza, J., Liao, D., Calhoun, S., Shaffer, M. L., & Bixler, E. O. (2012). Risk factors for incident chronic insomnia: A general population prospective study. *Sleep Medicine*, *13*(4), 346–353. https://doi.org/10.1016/j.sleep.2011.10.033

Sivakumaran, S., Agakov, F., Theodoratou, E., Prendergast, J. G., Zgaga, L., Manolio, T., Rudan, I., McKeigue, P., Wilson, J. F., & Campbell, H. (2011). Abundant pleiotropy in human complex diseases and traits. *American Journal of Human Genetics*, *89*(5), 607–618. https://doi.org/10.1016/j.ajhg.2011.10.004

Skogar, O., & Lokk, J. (2016). Pain management in patients with Parkinson's disease: Challenges and solutions. *Journal of Multidisciplinary Healthcare*, *9*, 469–479. https://doi.org/10.2147/JMDH.S105857

Sluka, K. A., & Clauw, D. J. (2016). Neurobiology of fibromyalgia and chronic widespread pain. *Neuroscience*, *338*, 114–129. https://doi.org/10.1016/j.neuroscience.2016.06.006

Smith, B. H., Campbell, A., Linksted, P., Fitzpatrick, B., Jackson, C., Kerr, S. M., Deary, I. J., MacIntyre, D. J., Campbell, H., McGilchrist, M., Hocking, L. J., Wisely, L., Ford, I., Lindsay, R. S., Morton, R., Palmer, C. N. A., Dominiczak, A. F., Porteous, D. J., & Morris, A. D. (2013). Cohort profile: Generation scotland: Scottish family health study (GS: SFHS). The study, its participants and their potential for genetic research on health and illness. *International Journal of Epidemiology*. https://doi.org/10.1093/ije/dys084

Smith, B. H., Campbell, H., Blackwood, D., Connell, J., Connor, M., Deary, I. J., Dominiczak, A. F., Fitzpatrick, B., Ford, I., Jackson, C., Haddow, G., Kerr, S., Lindsay, R., McGilchrist, M., Morton, R., Murray, G., Palmer, C. N. A., Pell, J. P., Ralston, S. H., … Morris, A. D. (2006). Generation Scotland: The Scottish Family Health Study; a new resource for researching genes and heritability. *BMC Medical Genetics*, *7*, 1–9. https://doi.org/10.1186/1471-2350-7-74

Smith, B. H., Penny, K. I., Purves, A. M., Munro, C., Wilson, B., Grimshaw, W. J., Chambers, W. A., & Smith, W. C. (1997). The chronic pain grade questionnaire: Validation and reliability in postal research. *Pain*, *71*(2), 141–147. https://doi.org/10.1016/S0304-3959(97)03347-2

Smith, D. J., Langan, J., McLean, G., Guthrie, B., & Mercer, S. W. (2013). Schizophrenia is associated with excess multiple physical-health comorbidities but low levels of recorded cardiovascular disease in primary care: Cross-sectional study. *BMJ Open*, *3*(4), 1–9. https://doi.org/10.1136/bmjopen-2013-002808

Smith, D. J., Nicholl, B. I., Cullen, B., Martin, D., Ul-Haq, Z., Evans, J., Gill, J. M. R., Roberts, B., Gallacher, J., Mackay, D., Hotopf, M., Deary, I., Craddock, N., & Pell, J. P. (2013). Prevalence and characteristics of probable major depression and bipolar disorder within UK Biobank: Cross-sectional study of 172,751 participants. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0075362

Smith, George D., & Ebrahim, S. (2003). "Mendelian randomization": Can genetic epidemiology contribute to understanding environmental determinants of disease? *International Journal of Epidemiology*, *32*(1), 1–22. https://doi.org/10.1093/ije/dyg070

Smith, George Davey, & Ebrahim, S. (2003). 30th Thomas Francis Jr memorial lecture: 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *International Journal of Epidemiology*, *32*, 1–22. https://doi.org/10.1093/ije/dyg070

Smith, George Davey, & Ebrahim, S. (2004). *Mendelian randomization : prospects , potentials , and limitations*. *33*(1), 30–42. https://doi.org/10.1093/ije/dyh132

Smith, George Davey, & Ebrahim, S. (2005). What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ*, *330*, 1076–1079.

Smith, George Davey, & Hemani, G. (2014a). Mendelian randomization: Geneticanchors for causal inference in epidemiological studies. *Human Molecular Genetics*, *23*(R1), 89–98. https://doi.org/10.1093/hmg/ddu328

Smith, George Davey, & Hemani, G. (2014b). Mendelian randomization: Geneticanchorsfor causal inference in epidemiological studies. *Human Molecular Genetics*, *23*(R1), 89–98. https://doi.org/10.1093/hmg/ddu328

Smith, J. M., & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetical Research*, *23*(1), 23–35.

Smith, M. W., & O'Brien, S. J. (2005). Mapping by admixture linkage disequilibrium: Advances, limitations and guidelines. *Nature Reviews Genetics*, *6*(8), 623–632. https://doi.org/10.1038/nrg1657

Smith, N. G. C., Webster, M. T., & Ellegren, H. (2002). Deterministic mutation rate variation in the human genome. *Genome Research*, *12*(9), 1350–1356. https://doi.org/10.1101/gr.220502

Sodini, S. M., Kemper, K. E., Wray, N. R., & Trzaskowski, M. (2018). Comparison of genotypic and phenotypic correlations: Cheverud's conjecture in humans. *Genetics*, *209*(3), 941–948. https://doi.org/10.1534/genetics.117.300630

Sogos, V., Curto, M., Reali, C., & Gremo, F. (2002). Developmentally regulated expression and localization of dystrophin and utrophin in the human fetal brain. *Mechanisms of Ageing and Development*, *123*(5), 455–462. https://doi.org/10.1016/S0047-6374(01)00360-8

Solaro, C., Trabucco, E., & Messmer Uccelli, M. (2013). Pain and multiple sclerosis: Pathophysiology and treatment topical collection on demyelinating disorders. *Current Neurology and Neuroscience Reports*, *13*(1). https://doi.org/10.1007/s11910-012-0320-5

Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M., & Smoller, J. W. (2013). Pleiotropy in complex traits: challenges and strategies. *Nature Reviews Genetics*, *14*(7), 483–495. https://doi.org/10.1038/nrg3461

Sorge, R. E., Mapplebeck, J. C. S., Rosen, S., Beggs, S., Taves, S., Alexander, J. K., Martin, L. J., Austin, J. S., Sotocinal, S. G., Chen, D., Yang, M., Shi, X. Q., Huang, H., Pillon, N. J., Bilan, P. J., Tu, Y., Klip, A., Ji, R. R., Zhang, J., … Mogil, J. S. (2015). Different immune cells mediate mechanical pain hypersensitivity in male and female mice. *Nature Neuroscience*, *18*(8), 1081–1083. https://doi.org/10.1038/nn.4053

Sorge, R. E., & Totsch, S. K. (2017). Sex Differences in Pain. *Journal of Neuroscience Research*, *95*(6), 1271–1281. https://doi.org/10.1002/jnr.23841

Speed, T. J., Parekh, V., Coe, W., & Antoine, D. (2018). Comorbid chronic pain and opioid use disorder: literature review and potential treatment innovations. *International Review of Psychiatry*, *30*(5), 136–146. https://doi.org/10.1080/09540261.2018.1514369

Spitzer, R. L., Kroenke, K., & Williams, J. B. W. (2000). Validation and utility of a self-report version of PRIME-MD. *Primary Care Companion to the Journal of Clinical Psychiatry*, *2*(1), 31.

Srivastava, S., Cohen, J., Pevsner, J., Aradhya, S., McKnight, D., Butler, E., Johnston, M., & Fatemi, A. (2014). A novel variant in GABRB2 associated with intellectual disability and epilepsy. *American Journal of Medical Genetics*, *164A*(11), 2914–2921. https://doi.org/10.1002/ajmg.a.36714

Srivastava, S., Cohen, J. S., Vernon, H., Barañano, K., McClellan, R., Jamal, L., Naidu, S. B., & Fatemi, A. (2014). Clinical whole exome sequencing in child neurology practice. *Annals of Neurology*, *76*(4), 473–483. https://doi.org/10.1002/ana.24251

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, *65*, 557–586.

Stamatakis, E., Owen, K. B., Shepherd, L., Drayton, B., Hamer, M., & Bauman, A. E. (2021). Is Cohort Representativeness Passé ? Poststratified Associations of Lifestyle Risk Factors with Mortality in the UK Biobank. *Epidemiology*, *32*(2), 179–188. https://doi.org/10.1097/EDE.0000000000001316

Stanzione, M., Baumann, M., Papanikos, F., Dereli, I., Lange, J., Ramlal, A., Tränkner, D., Shibuya, H., Massy, B. de, Watanabe, Y., Jasin, M., Keeney, S., & Tóth, A. (2017). Meiotic DNA break formation requires the unsynapsed chromosome axis-binding protein IHO1 (CCDC36) in mice. *Physiology & Behavior*, *176*(1), 139–148. https://doi.org/10.1038/ncb3417.Meiotic

Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W., & Smadja, C. M. (2017). Variation in recombination frequency and distribution across eukaryotes: Patterns and processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736). https://doi.org/10.1098/rstb.2016.0455

Star, B., & Spencer, H. G. (2013). Effects of genetic drift and gene flow on the selective maintenance of genetic variation. *Genetics*, *194*(1), 235–244. https://doi.org/10.1534/genetics.113.149781

Steingrímsdóttir, Ó. A., Landmark, T., Macfarlane, G. J., & Nielsen, C. S. (2017). Defining chronic pain in epidemiological studies: A systematic review and meta-analysis. *Pain*, *158*(11), 2092–2107. https://doi.org/10.1097/j.pain.0000000000001009

Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Iny Stein, T., Nudel, R., Lieder, I., Mazor, Y., Kaplan, S., Dahary, D., Warshawsky, D., Guan-Golan, Y., Kohn, A., Rappaport, N., Safran, M., & Lancet, D. (2016). The GeneCards suite: From gene data mining to disease genome sequence analyses. *Current Protocols in Bioinformatics*, *2016*(June), 1.30.1-1.30.33. https://doi.org/10.1002/cpbi.5

Stephan, A. H., Barres, B. A., & Stevens, B. (2012). The complement system: An unexpected role in synaptic pruning during development and disease. *Annual Review of Neuroscience*, *35*, 369–389. https://doi.org/10.1146/annurev-neuro-061010-113810

Stephens, J. C., Julie A. Schneider, Debra A. Tanguay, Julie Choi, Tara Acharya, Scott E. Stanley, Ruhong Jiang, Chad J. Messer, Anne Chew, Jin-Hua Han, Jicheng Duan, Janet L. Carr, Min Seob Lee, Beena Koshy, A. Madan Kumar, Ge Zhang, William R. Newell, Andreas Windemuth, Chuanbo Xu, Theodore S. Kalbfleisch, Sandra L. Shaner, … Gerald F. Vovis. (2001). Haplotype variation and haplotype linkage disequilibrium in 313 human genes. *Science*, *293*(5529), 489–493. https://doi.org/10.1126/science.1059431

Stephens, K., Zhou, W., Ji, Z., He, S., Ji, J., Guan, Y., & Taverna, S. (2018). Sex differences in gene regulation in the dorsal root ganglion after nerve injury. *The Journal of Pain*, *19*(3), S100. https://doi.org/10.1016/j.jpain.2017.12.225

Sterky, F. H., Trotter, J. H., Lee, S.-J., Recktenwald, C. V., Xiao Du, B. Z., Zhou, P., Schwenk, J., Fakler, B., & Südhof, T. C. (2017). Carbonic anhydrase-related protein CA10 is an evolutionarily conserved pan-neurexin ligand. *Proceedings of the National Academy of Sciences*, *114*(14), E2984–E2984. https://doi.org/10.1073/pnas.1703198114

Stewart, J. (2014). The challenges of cancer pain assessment. *Ulster Medical Journal*, *83*(1), 44–46.

Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *64*(3), 479–498. https://doi.org/10.1111/1467-9868.00346

Stratton, P., & Berkley, K. J. (2011). Chronic pelvic pain and endometriosis: Translational evidence of the relationship and implications. *Human Reproduction Update*, *17*(3), 327–346. https://doi.org/10.1093/humupd/dmq050

Strawbridge, R. J., Ward, J., Ferguson, A., Graham, N., Shaw, R. J., Cullen, B., Pearsall, R., Lyall, L.

M., Johnston, K. J. A., Niedzwiedz, C. L., Pell, J. P., Mackay, D., Martin, J. L., Lyall, D. M., Bailey, M. E. S., & Smith, D. J. (2019). Identification of novel genome-wide associations for suicidality in UK Biobank, genetic correlation with psychiatric disorders and polygenic association with completed suicide. *EBioMedicine*, *41*, 517–525. https://doi.org/10.1016/j.ebiom.2019.02.005

Streeter, A. J., Lin, N. X., Crathorne, L., Haasova, M., Hyde, C., Melzer, D., & Henley, W. E. (2017). Adjusting for unmeasured confounding in nonrandomized longitudinal studies: a methodological review. *Journal of Clinical Epidemiology*, *87*, 23–34. https://doi.org/10.1016/j.jclinepi.2017.04.022

Strigo, I. A., Matthews, S. C., Simmons, A. N., Oberndorfer, T., Klabunde, M., Reinhardt, L. E., & Kaye, W. H. (2013). Altered insula activation during pain anticipation in individuals recovered from anorexia nervosa: Evidence of interoceptive dysregulation. *International Journal of Eating Disorders*, *46*(1), 23–33. https://doi.org/10.1002/eat.22045

Stubbs, B., Eggermont, L., Mitchell, A. J., De Hert, M., Correll, C. U., Soundy, A., Rosenbaum, S., & Vancampfort, D. (2015). The prevalence of pain in bipolar disorder: A systematic review and large-scale meta-analysis. *Acta Psychiatrica Scandinavica*, *131*(2), 75–88. https://doi.org/10.1111/acps.12325

Stubbs, B., Mitchell, A. J., De Hert, M., Correll, C. U., Soundy, A., Stroobants, M., & Vancampfort, D. (2014). The prevalence and moderators of clinical pain in people with schizophrenia: A systematic review and large scale meta-analysis. *Schizophrenia Research*, *160*(1–3), 1–8. https://doi.org/10.1016/j.schres.2014.10.017

Su, W., Han, H. H., Wang, Y., Zhang, B., Zhou, B., Cheng, Y., Rumandla, A., Gurrapu, S., Chakraborty, G., Su, J., Yang, G., Liang, X., Wang, G., Rosen, N., Scher, H. I., Ouerfelli, O., & Giancotti, F. G. (2019). The Polycomb Repressor Complex 1 Drives Double-Negative Prostate Cancer Metastasis by Coordinating Stemness and Immune Suppression. *Cancer Cell*, *36*(2), 139-155.e10. https://doi.org/10.1016/j.ccell.2019.06.009

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102

Südhof, T. C. (2018). Towards an Understanding of Synapse Formation. *Neuron*. https://doi.org/10.1016/j.neuron.2018.09.040

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., & Collins, R. (2015). UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Medicine*. https://doi.org/10.1371/journal.pmed.1001779

Sul, J. H., Martin, L. S., & Eskin, E. (2018). Population structure in genetic studies: Confounding factors and mixed models. *PLoS Genetics*, *14*(12), 1–22. https://doi.org/10.1371/journal.pgen.1007309

Sullivan, P. F. (2007). Spurious Genetic Associations. *Biological Psychiatry*, *61*(10), 1121–1126. https://doi.org/10.1016/j.biopsych.2006.11.010

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic Epidemiology of Major Depression: Review and Meta-Analysis. *American Journal of Psychiatry*, *157*(10), 1552–1562. https://doi.org/10.1176/appi.ajp.157.10.1552

Sun, J., Yan, W., Zhang, X. N., Lin, X., Li, H., Gong, Y. M., Zhu, X. M., Zheng, Y. B., Guo, X. Y., Ma, Y.

D., Liu, Z. Y., Liu, L., Gao, J. H., Vitiello, M. V., Chang, S. H., Liu, X. G., & Lu, L. (2020). Polygenic evidence and overlapped brain functional connectivities for the association between chronic pain and sleep disturbance. *Translational Psychiatry*, *10*(1). https://doi.org/10.1038/s41398-020-00941-z

Suri, P., Palmer, M. R., Tsepilov, Y. A., Freidin, M. B., Boer, C. G., Yau, M. S., Evans, D. S., Gelemanovic, A., Bartz, T. M., Nethander, M., Arbeeva, L., Karssen, L., Neogi, T., Campbell, A., Mellstrom, D., Ohlsson, C., Lauc, G., Psaty, B. M., Karlsson, M. K., … Williams, F. M. K. (2018). Genome-wide meta-analysis of 158,000 individuals of European ancestry identifies three loci associated with chronic back pain. *PLoS Genetics*, *48*(7), 1–23. https://doi.org/10.1038/ng.3582.

Tabor, H. K., Risch, N. J., & Myers, R. M. (2002). Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nature Reviews Genetics*, *3*, 1–7.

Tanaka, M., Sasaki, K., Kamata, R., Hoshino, Y., Yanagihara, K., & Sakai, R. (2009). A Novel RNA-Binding Protein, Ossa/C9orf10, Regulates Activity of Src Kinases To Protect Cells from Oxidative Stress-Induced Apoptosis. *Molecular and Cellular Biology*, *29*(2), 402–413. https://doi.org/10.1128/MCB.01035-08

Tang, Z., Chen, Z., Tang, B., & Jiang, H. (2015). Primary erythromelalgia: A review. *Orphanet Journal of Rare Diseases*, *10*(1), 1–11. https://doi.org/10.1186/s13023-015-0347-1

Tayebi, N., Jamsheer, A., Flöttmann, R., Sowinska-Seidler, A., Doelken, S. C., Oehl-Jaschkowitz, B., Hülsemann, W., Habenicht, R., Klopocki, E., Mundlos, S., & Spielmann, M. (2014). Deletions of exons with regulatory activity at the DYNC1I1 locus are associated with split-hand/split-foot malformation: Array CGH screening of 134 unrelated families. *Orphanet Journal of Rare Diseases*, *9*(1), 1–9. https://doi.org/10.1186/s13023-014-0108-6

Tegethoff, M., Belardi, A., Stalujanis, E., & Meinlschmidt, G. (2015). Comorbidity of Mental Disorders and Chronic Pain: Chronology of Onset in Adolescents of a National Representative Cohort. *The Journal of Pain*, *16*(10).

Tentindo, G. S., Fishman, S. M., Li, C. S., Wang, Q., & Brass, S. D. (2018). The prevalence and awareness of sleep apnea in patients suffering chronic pain: An assessment using the STOP-Bang sleep apnea questionnaire. *Nature and Science of Sleep*, *10*, 217–224. https://doi.org/10.2147/NSS.S167658

Thomas, D. C., & Conti, D. V. (2001a). *Commentary : The concept of ' Mendelian randomization .'* 21–25. https://doi.org/10.1093/ije/dyh048

Thomas, D. C., & Conti, D. V. (2001b). Commentary : The concept of ' Mendelian randomization .' *International Journal of Epidemiology*, *33*, 21–25. https://doi.org/10.1093/ije/dyh048

Thompson, A. E., Anisimowicz, Y., Miedema, B., Hogg, W., Wodchis, W. P., & Aubrey-Bassler, K. (2016). The influence of gender and other patient characteristics on health care-seeking behaviour: A QUALICOPC study. *BMC Family Practice*, *17*(1), 1–7. https://doi.org/10.1186/s12875-016-0440-0

Thorisson, G. A., & Stein, L. D. (2003). The SNP consortium website: Past, present and future. *Nucleic Acids Research*, *31*(1), 124–127. https://doi.org/10.1093/nar/gkg052

Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J., & Richards, J. B. (2018). Genetic architecture: The shape of the genetic contribution to human traits and disease. *Nature Reviews Genetics*, *19*(2), 110–124. https://doi.org/10.1038/nrg.2017.101

Tishkoff, S. A., Reed, F. A., Friedlaender, F. R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J. B., Awomoyi, A. A., Bodo, J. M., Doumbo, O., Ibrahim, M., Juma, A. T., Kotze, M. J., Lema, G., Moore, J. H., Mortensen, H., Nyambo, T. B., Omar, S. A., Powell, K., … Williams, S. M. (2009). The genetic structure and history of Africans and African Americans. *Science*, *324*(5930),

1035–1044. https://doi.org/10.1126/science.1172257

Tolosa, A., Sanjuán, J., Dagnall, A. M., Moltó, M. D., Herrero, N., & de Frutos, R. (2010). FOXP2 gene and language impairment in schizophrenia: association and epigenetic studies. *BMC Medical Genetics*, *11*(1), 114. https://doi.org/10.1186/1471-2350-11-114

Tomko, R. J., & Hochstrasser, M. (2013). Molecular architecture and assembly of the eukaryotic proteasome. *Annual Review of Biochemistry*, *82*, 415–445. https://doi.org/10.1146/annurev-biochem-060410-150257

Torkamani, A., Wineinger, N. E., & Topol, E. J. (2018). The personal and clinical utility of polygenic risk scores. *Nature Reviews Genetics*, *19*(9), 581–590. https://doi.org/10.1038/s41576-018-0018-x

Torres-Berrío, A., Hernandez, G., Nestler, E. J., & Flores, C. (2020). The Netrin-1/DCC Guidance Cue Pathway as a Molecular Target in Depression: Translational Evidence. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2020.04.025

Towns, W. L., & Begley, T. J. (2012). Transfer RNA methytransferases and their corresponding modifications in budding yeast and humans: Activities, predications, and potential roles in human health. *DNA and Cell Biology*, *31*(4), 434–454. https://doi.org/10.1089/dna.2011.1437

Tracey, I. (2010). Getting the pain you expect: Mechanisms of placebo, nocebo and reappraisal effects in humans. *Nature Medicine*, *16*(11), 1277–1283. https://doi.org/10.1038/nm.2229

Tracey, I., & Johns, E. (2010). The pain matrix: Reloaded or reborn as we image tonic pain using arterial spin labelling. *Pain*, *148*(3), 359–360. https://doi.org/10.1016/j.pain.2009.11.009

Treede, R. D., Rief, W., Barke, A., Aziz, Q., Bennett, M. I., Benoliel, R., Cohen, M., Evers, S., Finnerup, N. B., First, M. B., Giamberardino, M. A., Kaasa, S., Korwisi, B., Kosek, E., Lavand'Homme, P., Nicholas, M., Perrot, S., Scholz, J., Schug, S., … Wang, S. J. (2019). Chronic pain as a symptom or a disease: The IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). *Pain*, *160*(1), 19–27. https://doi.org/10.1097/j.pain.0000000000001384

Trouvin, A., & Perrot, S. (2018). Pain in osteoarthritis . Implications for optimal management. *Joint Bone Spine*, *85*(4), 429–434. https://doi.org/10.1016/j.jbspin.2017.08.002

Tsai, W., & Clarke, S. (1994). Amino Acid Polymorphisms of the Human L-Isoaspartyl/D-Aspartyl Methyltransferase Inolved in Protein Repair. *Biochemical and Biophysical Research Communications*, *203*(1), 491–497.

Tsang, A., Von Korff, M., Lee, S., Alonso, J., Karam, E., Angermeyer, M. C., Borges, G. L. G., Bromet, E. J., de Girolamo, G., de Graaf, R., Gureje, O., Lepine, J. P., Haro, J. M., Levinson, D., Oakley Browne, M. A., Posada-Villa, J., Seedat, S., & Watanabe, M. (2008). Common Chronic Pain Conditions in Developed and Developing Countries: Gender and Age Differences and Comorbidity With Depression-Anxiety Disorders. *Journal of Pain*, *9*(10), 883–891. https://doi.org/10.1016/j.jpain.2008.05.005

Tsepilov, Y. A., Freidin, M. B., Shadrina, A. S., Sharapov, S. Z., Elgaeva, E. E., Zundert, J. van, Karssen, L., Suri, P., Williams, F. M. K., & Aulchenko, Y. S. (2020). Analysis of genetically independent phenotypes identifies shared genetic factors associated with chronic musculoskeletal pain conditions. *Communications Biology*, *3*(1), 1–13. https://doi.org/10.1038/s42003-020-1051-9

Tucker, E. J., Wanschers, B. F. J., Szklarczyk, R., Mountford, H. S., Wijeyeratne, X. W., van den Brand, M. A. M., Leenders, A. M., Rodenburg, R. J., Reljić, B., Compton, A. G., Frazier, A. E., Bruno, D. L., Christodoulou, J., Endo, H., Ryan, M. T., Nijtmans, L. G., Huynen, M. A., & Thorburn, D. R. (2013). Mutations in the UQCC1-Interacting Protein, UQCC2, Cause Human

Complex III Deficiency Associated with Perturbed Cytochrome b Protein Expression. *PLoS Genetics*, *9*(12). https://doi.org/10.1371/journal.pgen.1004034

Ugarte, M., Pérez-Cerdá, C., Rodríguez-Pombo, P., Desviat, L. R., Pérez, B., Richard, E., Muro, S., Campeau, E., Ohura, T., & Gravel, R. A. (1999). Overview of mutations in the PCCA and PCCB genes causing propionic acidemia. *Human Mutation*, *14*(4), 275–282. https://doi.org/10.1002/(SICI)1098-1004(199910)14:4<275::AID-HUMU1>3.0.CO;2-N

UK Biobank. (2020). *UK Biobank Pain web questionnaire* (Issue December).

Uniprot. (n.d.). *Disease - Joubert Syndrome 31*. Retrieved September 24, 2021, from https://www.uniprot.org/diseases/DI-05136

Unoki, M., Nishidate, T., & Nakamura, Y. (2004). ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. *Oncogene*, *23*(46), 7601–7610. https://doi.org/10.1038/sj.onc.1208053

Urban-Kowalczyk, M., Pigońska, J., & Śmigielski, J. (2015). Pain perception in schizophrenia: Influence of neuropeptides, cognitive disorders, and negative symptoms. *Neuropsychiatric Disease and Treatment*, *11*, 2023–2030. https://doi.org/10.2147/NDT.S87666

Urban, H., & Little, C. B. (2018). The role of fat and inflammation in the pathogenesis and management of osteoarthritis. *Rheumatology (United Kingdom)*, *57*(February), iv10–iv21. https://doi.org/10.1093/rheumatology/kex399

Urits, I., Adamian, L., Fiocchi, J., Hoyt, D., Ernst, C., Kaye, A. D., & Viswanath, O. (2019). Advances in the Understanding and Management of Chronic Pain in Multiple Sclerosis: a Comprehensive Review. *Current Pain and Headache Reports*, *23*(8). https://doi.org/10.1007/s11916-019-0800-2

Vachon-Presseau, E., Tétreault, P., Petre, B., Huang, L., Berger, S. E., Torbey, S., Baria, A. T., Mansour, A. R., Hashmi, J. A., Griffith, J. W., Comasco, E., Schnitzer, T. J., Baliki, M. N., & Apkarian, A. V. (2016). Corticolimbic anatomical characteristics predetermine risk for chronic pain. *Brain*, *139*(7), 1958–1970. https://doi.org/10.1093/brain/aww100

Vadlamudi, R. K., & Kumar, R. (2003). P21-Activated Kinases in Human Cancer. *Cancer and Metastasis Reviews*, *22*(4), 385–393. https://doi.org/10.1023/A:1023729130497

Valderrama, F., Cordeiro, J. V., Schleich, S., Frischknecht, F., & Way, M. (2006). Vaccinia Virus–Induced Cell Motility Requires F11L-Mediated Inhibition of RhoA Signaling Ferran. *Science*, *311*(January), 377–381.

Valdes, A. M., Doherty, S. A., Zhang, W., Muir, K. R., Maciewicz, R. A., & Doherty, M. (2012). Inverse relationship between preoperative radiographic severity and postoperative pain in patients with osteoarthritis who have undergone total joint arthroplasty. *Seminars in Arthritis and Rheumatism*, *41*(4), 568–575. https://doi.org/10.1016/j.semarthrit.2011.07.002

Valkovic, P., Minar, M., Singliarova, H., Harsany, J., Hanakova, M., Martinkova, J., Benetin, J., & LeDoux, M. S. (2015). Pain in Parkinson's disease: A cross-sectional study of its prevalence, types, and relationship to depression and quality of life. *PLoS ONE*, *10*(8), 1–11. https://doi.org/10.1371/journal.pone.0136541

Van Acker, H. H., Capsomidis, A., Smits, E. L., & Van Tendeloo, V. F. (2017). CD56 in the immune system: More than a marker for cytotoxicity? *Frontiers in Immunology*, *8*(JUL), 1–9. https://doi.org/10.3389/fimmu.2017.00892

Van Damme, S., Becker, S., & Van Der Linden, D. (2018). Tired of pain? Toward a better understanding of fatigue in chronic pain. *Pain*, *159*(1), 7–10. https://doi.org/10.1097/j.pain.0000000000001054

van der Lelij, P., Lieb, S., Jude, J., Wutz, G., Santos, C. P., Falkenberg, K., Schlattl, A., Ban, J.,

Schwentner, R., Hoffmann, T., Kovar, H., Real, F. X., Waldman, T., Pearson, M. A., Kraut, N., Peters, J. M., Zuber, J., & Petronczki, M. (2017). Synthetic lethality between the cohesin subunits STAG1 and STAG2 in diverse cancer contexts. *ELife*, *6*, 1–15. https://doi.org/10.7554/eLife.26980

Vansteelandt, S., Goetgeluk, S., Lutz, S., Waldman, I., Lyon, H., Schadt, E. E., Weiss, S. T., & Lange, C. (2009). On the adjustment for covariates in genetic association analysis: A novel, simple principle to infer direct causal effects. *Genetic Epidemiology*, *33*(5), 394–405. https://doi.org/10.1002/gepi.20393

Vargas-Prada, S., & Coggon, D. (2015). Psychological and psychosocial determinants of musculoskeletal pain and associated disability. *Best Practice and Research: Clinical Rheumatology*, *29*(3), 374–390. https://doi.org/10.1016/j.berh.2015.03.003

Varytė, G., Zakarevičienė, J., Ramašauskaitė, D., Laužikienė, D., & Arlauskienė, A. (2020). Pregnancy and Multiple Sclerosis: An Update on the Disease Modifying Treatment Strategy and a Review of Pregnancy's Impact on Disease Activity. *Medicina (Kaunas, Lithuania)*, *56*(2). https://doi.org/10.3390/medicina56020049

Vasic, V., & Schmidt, M. H. H. (2017). Resilience and vulnerability to pain and inflammation in the hippocampus. *International Journal of Molecular Sciences*, *18*(4). https://doi.org/10.3390/ijms18040739

Vellucci, R. (2012). Heterogeneity of chronic pain. In *Clinical Drug Investigation*. https://doi.org/10.2165/11630030-000000000-00000

Veluchamy, A., Hébert, H. L., Meng, W., Palmer, C. N. A., & Smith, B. H. (2018). Systematic review and meta-analysis of genetic risk factors for neuropathic pain. *Pain*, *159*(5), 825–848. https://doi.org/10.1097/j.pain.0000000000001164

Vercellini, P., Fedele, L., Aimi, G., Pietropaolo, G., Consonni, D., & Crosignani, P. G. (2007). Association between endometriosis stage, lesion type, patient characteristics and severity of pelvic pain symptoms: A multivariate analysis of over 1000 patients. *Human Reproduction*, *22*(1), 266–271. https://doi.org/10.1093/humrep/del339

Vernay, A., Lamrabet, O., Perrin, J., & Cosson, P. (2018). TM9SF4 levels determine sorting of transmembrane domains in the early secretory pathway. *Journal of Cell Science*, *131*(21), 1–9. https://doi.org/10.1242/jcs.220830

Verne, G. N., & Zhou, Q. (2011). New insights into visceral hypersensitivity —clinical implications in IBS. *Nature Reviews Gastroenterology & Hepatology*, *8*(6), 349–355. https://doi.org/10.1038/nrgastro.2011.83.New

Vernes, S. C., Oliver, P. L., Spiteri, E., Lockstone, H. E., Puliyadi, R., Taylor, J. M., Ho, J., Mombereau, C., Brewer, A., Lowy, E., Nicod, J., Groszer, M., Baban, D., Sahgal, N., Cazier, J.-B., Ragoussis, J., Davies, K. E., Geschwind, D. H., & Fisher, S. E. (2011). Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. *PLoS Genetics*, *7*(7), e1002145. https://doi.org/10.1371/journal.pgen.1002145

Viniol, A., Jegan, N., Leonhardt, C., Brugger, M., Strauch, K., Barth, J., Baum, E., & Becker, A. (2013). Differences between patients with chronic widespread pain and local chronic low back pain in primary care - A comparative cross-sectional analysis. *BMC Musculoskeletal Disorders*, *14*. https://doi.org/10.1186/1471-2474-14-351

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *American Journal of Human Genetics*, *90*(1), 7–24. https://doi.org/10.1016/j.ajhg.2011.11.029

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era - Concepts and misconceptions. *Nature Reviews Genetics*, *9*(4), 255–266. https://doi.org/10.1038/nrg2322

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *American Journal of Human Genetics*, *101*(1), 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005

Visscher, P. M., & Yang, J. (2016). A plethora of pleiotropy across complex traits. *Nature Publishing Group*, *48*(7), 707–708. https://doi.org/10.1038/ng.3604

Voight, B. F., & Pritchard, J. K. (2005). Confounding from cryptic relatedness in case-control association studies. *PLoS Genetics*, *1*(3). https://doi.org/10.1371/journal.pgen.0010032

von Appen, A., LaJoie, D., Johnson, I. E., Trnka, M. J., Pick, S. M., Burlingame, A. L., Ullman, K. S., & Frost, A. (2020). LEM2 phase separation promotes ESCRT-mediated nuclear envelope reformation. *Nature*, *582*(7810), 115–118. https://doi.org/10.1038/s41586-020-2232-x

Von Hausswolff-Juhlin, Y., Bjartveit, M., Lindström, E., & Jones, P. (2009). Schizophrenia and physical health problems. *Acta Psychiatrica Scandinavica*, *119*(SUPPL. 438), 15–21. https://doi.org/10.1111/j.1600-0447.2008.01309.x

Von Korff, M, Ormel, J., Keefe, F. J., & Dworkin, S. F. (1992). Grading the severity of chronic pain. *Pain*, *50*(1092), 133–149. https://doi.org/10.1016/0304-3959(92)90154-4

Von Korff, M, & Simon, G. (1996). The Relationship Between Pain and Depression. *British Journal of Psychiatry*, *168*, 101–108. https://doi.org/10.14219/jada.archive.2009.0101

Von Korff, Michael, Crane, P., Lane, M., Miglioretti, D. L., Simon, G., Saunders, K., Stang, P., Brandenburg, N., & Kessler, R. (2005). Chronic spinal pain and physical-mental comorbidity in the United States: Results from the national comorbidity survey replication. *Pain*, *113*(3), 331–339. https://doi.org/10.1016/j.pain.2004.11.010

Vuolteenaho, R., Nissinen, M., Sainio, K., Byers, M., Eddy, R., Hirvonen, H., Shows, T. B., Sariola, H., Engvall, E., & Tryggvason, K. (1994). Human laminin M chain (merosin): Complete primary structure, chromosomal assignment, and expression of the M and A chain in human fetal tissues. *Journal of Cell Biology*, *124*(3), 381–394. https://doi.org/10.1083/jcb.124.3.381

Vuong, J. K., Lin, C. H., Zhang, M., Chen, L., Black, D. L., & Zheng, S. (2016). PTBP1 and PTBP2 Serve Both Specific and Redundant Functions in Neuronal Pre-mRNA Splicing. *Cell Reports*, *17*(10), 2766–2775. https://doi.org/10.1016/j.celrep.2016.11.034

Waldheim, E., Ajeganova, S., Bergman, S., Frostegård, J., & Welin, E. (2018). Variation in pain related to systemic lupus erythematosus (SLE): a 7-year follow-up study. *Clinical Rheumatology*, *37*(7), 1825–1834. https://doi.org/10.1007/s10067-018-4079-1

Walsh, D. A., & McWilliams, D. F. (2014). Mechanisms, impact and management of pain in rheumatoid arthritis. *Nature Reviews Rheumatology*, *10*(10), 581–592. https://doi.org/10.1038/nrrheum.2014.64

Walters, E. T. (2019). Adaptive mechanisms driving maladaptive pain: How chronic ongoing activity in primary nociceptors can enhance evolutionary fitness after severe injury. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1785). https://doi.org/10.1098/rstb.2019.0277

Wang, K., Li, M., & Bucan, M. (2007). Pathway-based approaches for analysis of genomewide association studies. *American Journal of Human Genetics*, *81*(6), 1278–1283. https://doi.org/10.1086/522374

Wang, Y., Thompson, W. K., Schork, A. J., Holland, D., Chen, C. H., Bettella, F., Desikan, R. S., Li, W., Witoelar, A., Zuber, V., Devor, A., Nöthen, M. M., Rietschel, M., Chen, Q., Werge, T., Cichon, S., Weinberger, D. R., Djurovic, S., O'Donovan, M., … Dale, A. M. (2016). Leveraging Genomic Annotations and Pleiotropic Enrichment for Improved Replication Rates in Schizophrenia GWAS. *PLoS Genetics*, *12*(1), 1–22. https://doi.org/10.1371/journal.pgen.1005803

Ward, J., Tunbridge, E. M., Sandor, C., Lyall, L. M., Ferguson, A., Strawbridge, R. J., Lyall, D. M., Cullen, B., Graham, N., Johnston, K. J. A., Webber, C., Escott-Price, V., O'Donovan, M., Pell, J. P., Bailey, M. E. S., Harrison, P. J., & Smith, D. J. (2019). The genomic basis of mood instability: identification of 46 loci in 363,705 UK Biobank participants, genetic correlation with psychiatric disorders, and association with gene expression and function. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-019-0439-8

Watanabe, K., Stringer, S., Frei, O., Mirkov, M. U., Polderman, T. J. C., Sluis, S. van der, Andreassen, O. A., Neale, B. M., & Posthuma, D. (2018). A global overview of pleiotropy and genetic architecture in complex traits. *BioRxiv*, 500090. https://doi.org/10.1101/500090

Watanabe, K., Taskesen, E., Van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, *8*(1), 1–10. https://doi.org/10.1038/s41467-017-01261-5

Wearing, S. C., Hennig, E. M., Byrne, N. M., Steele, J. R., & Hills, A. P. (2006). Musculoskeletal disorders associated with obesity: A biomechanical perspective. *Obesity Reviews*, *7*(3), 239–250. https://doi.org/10.1111/j.1467-789X.2006.00251.x

Weatherburn, C. J., Guthrie, B., Mercer, S. W., & Morales, D. R. (2017). Comorbidities in adults with asthma: Population-based cross-sectional analysis of 1.4 million adults in Scotland. *Clinical and Experimental Allergy*, *47*(10), 1246–1252. https://doi.org/10.1111/cea.12971

Weir, G. A., Chesler, E. J., Finnerup, N. B., Calvo, M., Davies, A. J., He, H. L., Levitt, R. C., Smith, B. H., Neely, G. G., Costigan, M., & Bennett, D. L. (2019). *Review The Genetics of Neuropathic Pain from Model Organisms to Clinical Application*. 637–653. https://doi.org/10.1016/j.neuron.2019.09.018

Welter, D., Macarthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L., & Parkinson, H. (2014). *The NHGRI GWAS Catalog , a curated resource of SNP-trait associations*. *42*(December 2013), 1001–1006. https://doi.org/10.1093/nar/gkt1229

Werme, J., Van Der Sluis, S., Posthuma, D., & De Leeuw, C. A. (2021). LAVA: An integrated framework for local genetic correlation analysis. *BioRxiv*, 2020.12.31.424652. https://doi.org/10.1101/2020.12.31.424652

WHO. (2017). Depression and other common mental disorders: global health estimates. *World Health Organization*, 1–24. https://doi.org/CC BY-NC-SA 3.0 IGO

Wilfert, A. B., Turner, T. N., Murali, S. C., Hsieh, P., Sulovari, A., Wang, T., Coe, B. P., Guo, H., Hoekzema, K., Bakken, T. E., Winterkorn, L. H., Evani, U. S., Byrska-Bishop, M., Earl, R. K., Bernier, R. A., Zhou, X., Feliciano, P., Hall, J., Astrovskaya, I., … Eichler, E. E. (2021). Recent ultra-rare inherited variants implicate new autism candidate risk genes. *Nature Genetics*, *53*(8), 1125–1134. https://doi.org/10.1038/s41588-021-00899-8

Willemsen, M. H., Nijhof, B., Fenckova, M., Nillesen, W. M., Bongers, E. M. H. F., Castells-Nobau, A., Asztalos, L., Viragh, E., van Bon, B. W. M., Tezel, E., Veltman, J. A., Brunner, H. G., de Vries, B. B. A., de Ligt, J., Yntema, H. G., van Bokhoven, H., Isidor, B., Caignec, C. Le, Lorino, E., … Kleefstra, T. (2013). GATAD2B loss-of-function mutations cause a recognisable syndrome with intellectual disability and are associated with learning deficits and synaptic undergrowth in Drosophila. *Journal of Medical Genetics*, *50*(8), 507–514. https://doi.org/10.1136/jmedgenet-2012-101490

Williams, J. B. W. (1988). A Structured Interview Guide for the Hamilton Depression Rating Scale. *Archives of General Psychiatry*, *45*(8), 742–747. https://doi.org/10.1001/archpsyc.1988.01800320058007

Wilson, P. M., Fryer, R. H., Fang, Y., & Hatten, M. E. (2010). Astn2, A Novel Member of the

Astrotactin Gene Family, Regulates the Trafficking of ASTN1 during Glial-Guided Neuronal Migration. *Journal of Neuroscience*, *30*(25), 8529–8540. https://doi.org/10.1523/JNEUROSCI.0032-10.2010

Wolf, S. (1977). The Way In and the Way Out. Francois Magendie, Charles Bell and the Roots of the Spinal Nerves. *Bulletin of the New York Academy of Medicine*, *53*, 754–757.

Wolfe, F., Clauw, D. J., Fitzcharles, M. A., Goldenberg, D. L., Häuser, W., Katz, R. S., Mease, P., Russell, A. S., Russell, I. J., & Winfield, J. B. (2011). Fibromyalgia criteria and severity scales for clinical and epidemiological studies: A modification of the ACR preliminary diagnostic criteria for fibromyalgia. *Journal of Rheumatology*, *38*(6), 1113–1122. https://doi.org/10.3899/jrheum.100594

Wolfe, F., Smythe, H. A., Yunus, M. B., Bennett, R. M., Bombardier, C., Goldenberg, D. L., Tugwell, P., Campbell, S. M., Abeles, M., Clark, P., Fam, A. G., Farber, S. J., Fiechtner, J. J., Michael Franklin, C., Gatter, R. A., Hamaty, D., Lessard, J., Lichtbroun, A. S., Masi, A. T., … Sheon, R. P. (1990). The american college of rheumatology 1990 criteria for the classification of fibromyalgia. *Arthritis & Rheumatism*, *33*(2), 160–172. https://doi.org/10.1002/art.1780330203

Wolfe, K. H., Sharp, P. M., & Lit, W. (1989). Mutation rates differ among regions of the mammalian genome. *Nature*, *337*(January), 1–3.

Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., Chu, A. Y., Estrada, K., Luan, J., Kutalik, Z., Amin, N., Buchkovich, M. L., Croteau-Chonka, D. C., Day, F. R., Duan, Y., Fall, T., Fehrmann, R., Ferreira, T., Jackson, A. U., … Frayling, T. M. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*, *46*(11), 1173–1186. https://doi.org/10.1038/ng.3097

Woolf, C. J. (2011). Central sensitization: Implications for the diagnosis and treatment of pain. *Pain*, *152*(SUPPL.3), S2–S15. https://doi.org/10.1016/j.pain.2010.09.030

World Health Organization. (2021). *Violence Against Women*. https://www.who.int/news-room/fact-sheets/detail/violence-against-women

Wray, N. R., Pergadia, M. L., Blackwood, D. H. R., Penninx, B. W. J. H., Gordon, S. D., Nyholt, D. R., Ripke, S., MacIntyre, D. J., McGhee, K. A., MacLean, A. W., Smit, J. H., Hottenga, J. J., Willemsen, G., Middeldorp, C. M., De Geus, E. J. C., Lewis, C. M., McGuffin, P., Hickie, I. B., Van Den Oord, E. J. C. G., … Sullivan, P. F. (2012). Genome-wide association study of major depressive disorder: New results, meta-analysis, and lessons learned. *Molecular Psychiatry*, *17*(1), 36–48. https://doi.org/10.1038/mp.2010.109

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M. J., Agerbo, E., Air, T. M., Andlauer, T. M. F., Bacanu, S.-A., Bækvad-Hansen, M., Beekman, A. F. T., Bigdeli, T. B., Binder, E. B., Blackwood, D. R. H., Bryois, J., Buttenschøn, H. N., Bybjerg-Grauholm, J., … Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, *50*(May), 167577. https://doi.org/10.1038/s41588-018-0090-3

Wray, N. R., Yang, J., Hayes, B. J., Price, A. L., Goddard, M. E., & Visscher, P. M. (2013). Pitfalls of predicting complex traits from SNPs. *Nature Reviews Genetics*, *14*(7), 507–515. https://doi.org/10.1038/nrg3457

Wright, K. M., Lyon, K. A., Leung, H., Leahy, D. J., Ma, L., & Ginty, D. D. (2012). Dystroglycan Organizes Axon Guidance Cue Localization and Axonal Pathfinding. *Neuron*, *76*(5), 931–944. https://doi.org/10.1016/j.neuron.2012.10.009

Wright, K. M., Rand, K. A., Kermany, A., Noto, K., Curtis, D., Garrigan, D., Slinkov, D., Dorfman, I.,

Granka, J. M., Byrnes, J., Myres, N., Ball, C. A., & Ruby, J. G. (2019). A prospective analysis of genetic variants associated with human lifespan. *G3: Genes, Genomes, Genetics*, *9*(9), 2863–2878. https://doi.org/10.1534/g3.119.400448

Wright, S. (1934a). an Analysis of Variability in Number of Digits in an Inbred Strain of Guinea Pigs. *Genetics*, *19*(6), 506–536. https://doi.org/10.1093/genetics/19.6.506

Wright, S. (1934b). the Results of Crosses Between Inbred Strains of Guinea Pigs, Differing in Number of Digits. *Genetics*, *19*(6), 537–551. https://doi.org/10.1093/genetics/19.6.537

Wright, S. (1937). The distribution of gene frequencies in populations. *Science*, *85*(2212), 504. https://doi.org/10.1126/science.85.2212.504

Wu, K. Y., Hengst, U., Cox, L. J., Macosko, E. Z., Jeromin, A., Urquhart, E. R., & Jaffrey, S. R. (2005). Local translation of RhoA regulates growth cone collapse. *Nature*, *436*(7053), 1020–1024. https://doi.org/10.1038/nature03885

Wu, L., Zee, B. M., Wang, Y., Garcia, B. A., & Dou, Y. (2011). The RING Finger Protein MSL2 in the MOF Complex Is an E3 Ubiquitin Ligase for H2B K34 and Is Involved in Crosstalk with H3 K4 and K79 Methylation. *Molecular Cell*, *43*(1), 132–144. https://doi.org/10.1016/j.molcel.2011.05.015

Wuest, J., Merritt-Gray, M., Ford-Gilboe, M., Lent, B., Varcoe, C., & Campbell, J. C. (2008). Chronic Pain in Women Survivors of Intimate Partner Violence. *Journal of Pain*, *9*(11), 1049–1057. https://doi.org/10.1016/j.jpain.2008.06.009

Xu, B., Woodroffe, A., Rodriguez-Murillo, L., Roos, J. L., van Rensburg, E. J., Abecasis, G. R., Gogos, J. A., & Karayiorgou, M. (2009). Elucidating the genetic architecture of familial schizophrenia using rare copy number variant and linkage scans. *Proceedings of the National Academy of Sciences*, *106*(39), 16746–16751. https://doi.org/10.1073/pnas.0908584106

Xu, Q., & Yaksh, T. L. (2011). A brief comparison of the pathophysiology of inflammatory versus neuropathic pain. *Current Opinion in Anaesthesiology*, *24*(4), 400–407. https://doi.org/10.1097/ACO.0b013e32834871df

Xu, X., Coats, J. K., Yang, C. F., Wang, A., Ahmed, O. M., Alvarado, M., Izumi, T., & Shah, N. M. (2012). Modular genetic control of sexually dimorphic behaviors. *Cell*, *148*(3), 596–607. https://doi.org/10.1016/j.cell.2011.12.018

Yaka, R., Thornton, C., Vagts, A. J., Phamluong, K., Bonci, A., & Ron, D. (2002). NMDA receptor function is regulated by the inhibitory scaffolding protein, RACK1. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(8), 5710–5715. https://doi.org/10.1073/pnas.062046299

Yamagata, M., & Sanes, J. R. (2008). Dscam and Sidekick proteins direct lamina-specific synaptic connections in vertebrate retina. *Nature*, *451*(7177), 465–469. https://doi.org/10.1038/nature06469

Yamagishi, S., Bando, Y., & Sato, K. (2021). Involvement of Netrins and Their Receptors in Neuronal Migration in the Cerebral Cortex. *Frontiers in Cell and Developmental Biology*, *8*(January), 1–12. https://doi.org/10.3389/fcell.2020.590009

Yamagishi, S., Yamada, K., Sawada, M., Nakano, S., Mori, N., Sawamoto, K., & Sato, K. (2015). Netrin-5 is highly expressed in neurogenic regions of the adult brain. *Frontiers in Cellular Neuroscience*, *9*(APR), 1–9. https://doi.org/10.3389/fncel.2015.00146

Yang, H., Duckett, C. S., & Lindsten, T. (1995). iPABP, an inducible poly(A)-binding protein detected in activated human T cells. *Molecular and Cellular Biology*, *15*(12), 6770–6776. https://doi.org/10.1128/mcb.15.12.6770

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A.,

Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., & Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*(7), 565–569. https://doi.org/10.1038/ng.608

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, *88*(1), 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011

Yang, J., Weedon, M. N., Purcell, S., Lettre, G., Estrada, K., Willer, C. J., Smith, A. V, Ingelsson, E., Connell, J. R. O., Mangino, M., Ma, R., Madden, P. A., Heath, A. C., Nyholt, D. R., Martin, N. G., Montgomery, G. W., Frayling, T. M., Hirschhorn, J. N., Mccarthy, M. I., & Goddard, M. E. (2011). *Genomic inflation factors under polygenic inheritance*. *July 2010*, 807–812. https://doi.org/10.1038/ejhg.2011.39

Yang, Q., Wu, Z., Hadden, J. K., Odem, M. A., Zuo, Y., Crook, R. J., Frost, J. A., & Walters, E. T. (2014). Persistent pain after spinal cord injury is maintained by primary afferent activity. *Journal of Neuroscience*, *34*(32), 10765–10769. https://doi.org/10.1523/JNEUROSCI.5316-13.2014

Yang, Y. R., Follo, M. Y., Cocco, L., & Suh, P. G. (2013). The physiological roles of primary phospholipase C. *Advances in Biological Regulation*, *53*(3), 232–241. https://doi.org/10.1016/j.jbior.2013.08.003

Yao, C., Joehanes, R., Johnson, A. D., Huan, T., Esko, T., Ying, S., Freedman, J. E., Murabito, J., Lunetta, K. L., Metspalu, A., Munson, P. J., & Levy, D. (2013). Sex- and age-interacting eQTLs in human complex diseases. *Human Molecular Genetics*, *23*(7), 1947–1956. https://doi.org/10.1093/hmg/ddt582

Yarnitsky, D. (2015). Role of endogenous pain modulation in chronic pain. *Pain*, *156*(4), 24–31.

Yeung, R. K., Xiang, Z. H., Tsang, S. Y., Li, R., Ho, T. Y. C., Li, Q., Hui, C. K., Sham, P. C., Qiao, M. Q., & Xue, H. (2018). Gabrb2-knockout mice displayed schizophrenia-like and comorbid phenotypes with interneuron-astrocyte-microglia dysregulation. *Translational Psychiatry*, *8*(1). https://doi.org/10.1038/s41398-018-0176-9

Yik, J. H. N., Chen, R., Pezda, A. C., & Zhou, Q. (2005). Compensatory contributions of HEXIM1 and HEXIM2 in maintaining the balance of active and inactive positive transcription elongation factor b complexes for control of transcription. *Journal of Biological Chemistry*, *280*(16), 16368–16376. https://doi.org/10.1074/jbc.M500912200

Yin, J., Jia, N., Liu, Y., Jin, C., Zhang, F., Yu, S., Wang, J., & Yuan, J. (2018). No association between FOXP2 rs10447760 and schizophrenia in a replication study of the Chinese Han population. *Psychiatric Genetics*, 1. https://doi.org/10.1097/YPG.0000000000000193

Yu, R., Gollub, R. L., Spaeth, R., Napadow, V., Wasan, A., & Kong, J. (2014). Disrupted functional connectivity of the periaqueductal gray in chronic low back pain. *NeuroImage: Clinical*, *6*, 100–108. https://doi.org/10.1016/j.nicl.2014.08.019

Yu, T., Wei, Z., Xu, T., Gamber, M., Han, J., Jiang, Y., Li, J., Yang, D., & Sun, W. (2020). The association between the socioeconomic status and body pain: A nationwide survey. *Medicine (United States)*, *99*(12). https://doi.org/10.1097/MD.0000000000019454

Yudell, M., Roberts, D., DeSalle, R., & Tishkoff, S. (2016). Science and society: Taking race out of human genetics. *Science*, *351*(6273), 564–565. https://doi.org/10.1126/science.aac4951

Zadro, J. R., Nilsen, T. I. L., Shirley, D., Amorim, A. B., Ferreira, P. H., Lier, R., & Mork, P. J. (2020). Parental Multisite Chronic Pain and the Risk of Adult Offspring Developing Additional Chronic Pain Sites: Family-Linkage Data From the Norwegian HUNT Study. *Journal of Pain*, *21*(9–10), 968–978. https://doi.org/10.1016/j.jpain.2019.12.007

Zamoyska, R., Basson, A., Filby, A., Legname, G., Lovatt, M., & Seddon, B. (2003). The influence of the src-family kinases, Lck and Fyn, on T cell differentiation, survival and activation. *Immunological Reviews*, *191*, 107–118. https://doi.org/10.1034/j.1600-065X.2003.00015.x

Zeng, Q., Subramaniam, V. N., Wong, S. H., Tang, B. L., Parton, R. G., Rea, S., James, D. E., & Hong, W. (1998). A novel synaptobrevin/VAMP homologous protein (VAMP5) is increased during in vitro myogenesis and present in the plasma membrane. *Molecular Biology of the Cell*, *9*(9), 2423–2437. https://doi.org/10.1091/mbc.9.9.2423

Zeyl, D., & Sattler, B. (2019). *Plato's Timaeus*. The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/plato-timaeus/

Zhang, M., Wang, L., Zhao, X., Zhao, K., Meng, H., Zhao, W., & Gao, C. (2012). TRAF-interacting protein (TRIP) negatively regulates IFN-β production and antiviral response by promoting proteasomal degradation of TANK-binding kinase 1. *Journal of Experimental Medicine*, *209*(10), 1703–1711. https://doi.org/10.1084/jem.20120024

Zhang, W., Wang, L., Liu, Y., Xu, J., Zhu, G., Cang, H., Li, X., Bartlam, M., Hensley, K., Li, G., Rao, Z., & Zhang, X. C. (2009). Structure of human lanthionine synthetase C-like protein 1 and its interaction with Eps8 and glutathione. *Genes and Development*, *23*(12), 1387–1392. https://doi.org/10.1101/gad.1789209

Zhang, Y., Ren, R., Lei, F., Zhou, J., Zhang, J., Wing, Y. K., Sanford, L. D., & Tang, X. (2019). Worldwide and regional prevalence rates of co-occurrence of insomnia and insomnia symptoms with obstructive sleep apnea: A systematic review and meta-analysis. *Sleep Medicine Reviews*, *45*, 1–17. https://doi.org/10.1016/j.smrv.2019.01.004

Zhao, F., Mao, B., Geng, X., Ren, X., Wang, Y., Guan, Y., Li, S., Li, L., Zhang, S., You, Y., Cao, Y., Yang, T., & Zhao, X. (2020). Molecular genetic analysis in 21 Chinese families with congenital insensitivity to pain with or without anhidrosis. *European Journal of Neurology*, *27*(8), 1697–1705. https://doi.org/10.1111/ene.14234

Zhao, Q., Wang, J., Hemani, G., Bowden, J., & Small, D. S. (2020). Statistical inference in two-sample summary-data mendelian randomization using robust adjusted profile score. *The Annals of Statistics*, *48*(3), 1742–1769. https://doi.org/https://doi.org/10.1214/19-AOS1866

Zheng, J., Baird, D., Borges, M.-C., Bowden, J., Hemani, G., Haycock, P., Evans, D. M., & Smith, G. D. (2017). Recent Developments in Mendelian Randomization Studies. *Current Epidemiology Reports*, *4*(4), 330–345. https://doi.org/10.1007/s40471-017-0128-6

Zheng, J., Erzurumluoglu, A. M., Elsworth, B. L., Kemp, J. P., Howe, L., Haycock, P. C., Hemani, G., Tansey, K., Laurin, C., Pourcain, B. S., Warrington, N. M., Finucane, H. K., Price, A. L., Bulik-Sullivan, B. K., Anttila, V., Paternoster, L., Gaunt, T. R., Evans, D. M., & Neale, B. M. (2017). LD Hub: A centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics*, *33*(2), 272–279. https://doi.org/10.1093/bioinformatics/btw613

Zheng, P., Zhang, W., Leng, J., & Lang, J. (2019). Research on central sensitization of endometriosis-associated pain: A systematic review of the literature. *Journal of Pain Research*, *12*, 1447–1456. https://doi.org/10.2147/JPR.S197667

Zhong, X., Drgonova, J., Li, C., Uhl, G. R., & Medicine, G. (2016). Human cell adhesion molecules: annotated functional subtypes and overrepresentation of addiction-associated genes. *Annals of the New York Academy of Sciences*, *1349*(1), 83–95. https://doi.org/10.1111/nyas.12776.Human

Zhou, L., Ouyang, L., Chen, K., & Wang, X. (2019). Research progress on KIF3B and related diseases. *Annals of Translational Medicine*, *7*(18), 492–492. https://doi.org/10.21037/atm.2019.08.47

Zhu, C., & Jiang, W. (2005). Cell cycle-dependent translocation of PRC1 on the spindle by Kif4 is essential for midzone formation and cytokinesis. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(2), 343–348. https://doi.org/10.1073/pnas.0408438102

Zhu, F., Dai, S. N., Xu, D. L., Hou, C. Q., Liu, T. T., Chen, Q. Y., Wu, J. L., & Miao, Y. (2020). EFNB2 facilitates cell proliferation, migration, and invasion in pancreatic ductal adenocarcinoma via the p53/p21 pathway and EMT. *Biomedicine and Pharmacotherapy*, *125*(November 2019), 109972. https://doi.org/10.1016/j.biopha.2020.109972

Zhu, Y., Yao, S., Augustine, M. M., Xu, H., Wang, J., Sun, J., Broadwater, M., Ruff, W., Luo, L., Zhu, G., Tamada, K., & Chen, L. (2016). Neuron-specific SALM5 limits inflammation in the CNS via its interaction with HVEM. *Science Advances*, *2*(4), e1500637. https://doi.org/10.1126/sciadv.1500637

Zhuo, M. (2008). Cortical excitation and chronic pain. *Trends in Neurosciences*, *31*(4), 199–207. https://doi.org/10.1016/j.tins.2008.01.003

Zimmerman, M., Ellison, W., Young, D., Chelminski, I., & Dalrymple, K. (2015). How many different ways do patients meet the diagnostic criteria for major depressive disorder? *Comprehensive Psychiatry*, *56*, 29–34. https://doi.org/10.1016/j.comppsych.2014.09.007

Zorina-Lichtenwalter, K., Meloto, C. B., Khoury, S., & Diatchenko, L. (2016). Genetic predictors of human chronic pain conditions. *Neuroscience*, *338*, 36–62. https://doi.org/10.1016/j.neuroscience.2016.04.041

Zorina-Lichtenwalter, K., Parisien, M., & Diatchenko, L. (2017). Genetic studies of human neuropathic pain conditions. *Pain*, *159*(3), 1. https://doi.org/10.1097/j.pain.0000000000001099

Zubieta, J. K., Heitzeg, M. M., Smith, Y. R., Bueller, J. A., Xu, K., Xu, Y., Koeppe, R. A., Stohler, C. S., & Goldmans, D. (2003). COMT val158 genotype affects μ-opioid neurotransmitter responses to a pain stressor. *Science*, *299*(5610), 1240–1243. https://doi.org/10.1126/science.1078546

Zuk, O., Hechter, E., Sunyaev, S. R., & Lander, E. S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences*, *109*(4), 1193–1198. https://doi.org/10.1073/pnas.1119675109