

1 **Genome-Wide Association Study Identifies Novel Genetic Variants Associated with**  
2 **Widespread Pain in the UK Biobank (N=172,230)**

3 Qi Pan<sup>1</sup>, Tengda Cai<sup>1</sup>, Yiwen Tao<sup>1</sup>, Roger Compte<sup>2</sup>, Maryam Kazemi Naeini<sup>2</sup>, Mainul Haque<sup>3</sup>,  
4 Tania Dottorini<sup>4</sup>, Frances M.K. Williams<sup>2</sup>, Weihua Meng<sup>1,5,6\*</sup>

5 1 Nottingham Ningbo China Beacons of Excellence Research and Innovation Institute,  
6 University of Nottingham Ningbo China, Ningbo, China, 315100

7 2 Department Twin Research and Genetic Epidemiology, School of Life Course Science, King's  
8 College London, Westminster Bridge Road, London, UK, SE1 7EH

9 3 School of Mathematical Sciences, University of Nottingham Ningbo China, Ningbo, China,  
10 315100

11 4 School of Veterinary Medicine and Science, University of Nottingham, Nottingham, UK, LE12  
12 5RD

13 5 Division of Population Health and Genomics, Ninewells Hospital and Medical School,  
14 University of Dundee, Dundee, UK, DD2 4BF

15 6 Center for Public Health, Faculty of Medicine, Health and Life Sciences, School of Medicine,  
16 Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, UK, BT12 6BA

17 Corresponding Author: Dr Weihua Meng

18 Address: Division of Population Health and Genomics, Ninewells Hospital and Medical School,  
19 University of Dundee, Dundee, UK, DD2 4BF; Nottingham Ningbo China Beacons of  
20 Excellence Research and Innovation Institute, University of Nottingham Ningbo China, Ningbo,

21 China, 315100. Tel: +44 1382 383419; +86 574 88188632. Email: [w.meng@dundee.ac.uk](mailto:w.meng@dundee.ac.uk);

22 [weihua.meng@nottingham.edu.cn](mailto:weihua.meng@nottingham.edu.cn). ORCID: 0000-0001-5388-8494

23 ORCID codes

24 Qi Pan: 0000-0003-0135-9122

25 Tengda Cai: 0000-0003-1617-506X

26 Yiwen Tao: 0009-0000-0102-4205

27 Roger Compte: 0000-0002-7034-1092

28 Maryam Kazemi Naeini: 0000-0002-4655-0716

29 Mainul Haque: 0000-0003-1836-8076

30 Tania Dottorini: 0000-0001-7733-906X

31 Frances M.K. Williams: 0000-0002-2998-2744

32 Weihua Meng: 0000-0001-5388-8494

33

34

35

36

37

38

39

40

41 **Abstract**

42 **Objectives:** Widespread pain is a hallmark characteristic of fibromyalgia, commonly affecting  
43 older individuals. This study aimed to identify novel genetic variants associated with widespread  
44 pain by utilizing the extensive UK Biobank dataset.

45 **Methods:** We conducted a primary genome-wide association study (GWAS) using a novel  
46 definition of widespread pain, defined as pain experienced all over the body during the past  
47 month. Sex-stratified GWAS analysis approach was also performed to analyze the impact of sex  
48 on widespread pain.

49 **Results:** The primary GWAS identified one novel significant genetic locus (rs34691025,  $p =$   
50  $1.76 \times 10^{-8}$ ) on chromosome 5q13.2 within the *ARHGEF28* gene and several loci that  
51 approached genome-wide significance. The sex-stratified GWAS outputs revealed biological  
52 difference widespread pain between males and females, with a novel locus identified in the  
53 female-specific analysis within the *LRMDA* gene on chromosome 10. Genetic Correlation  
54 analysis demonstrated significant genetic correlations between widespread pain and other  
55 phenotypes, including joint disorders and spondylosis. The PheWAS revealed associations  
56 between the significant genetic variants with hearing disorders and cardiovascular diseases.

57 **Conclusions:** Our study advances the understanding of the genetic factors contributing to  
58 widespread pain, highlighting notable differences between males and females and identifying a  
59 novel genetic locus associated with this condition.

60 **Keywords:** widespread pain; UK Biobank; genetic correlations; genome-wide association study,  
61 phenome-wide association study

62

63

## 64 **Introduction**

65 Widespread pain, commonly characterized by pain experienced in various parts of the body  
66 simultaneously, stands as a notable concern in clinical practice [1]. Unlike localized pain, which  
67 is confined to a specific region, widespread pain manifests across multiple regions, often without  
68 an apparent localized source. This type of pain is frequently observed in conditions such as  
69 fibromyalgia, rheumatoid arthritis, and certain neurological disorders, presenting a challenge in  
70 both diagnosis and treatment [2].

71 A commonly accepted definition of widespread pain is chronic widespread pain (CWP), which is  
72 characterized by the presence of widespread pain persisting for at least three months. A  
73 substantial body of epidemiological research suggests that the incidence and prevalence of  
74 CWP are significant, although these figures vary across different studies and countries. A  
75 systematic review indicates that the prevalence of CWP in the general population is  
76 approximately 10.6%, while in the UK, it is estimated to be 14.2% [3, 4]. The likelihood and risk  
77 factors of CWP increase with age, and females are more susceptible than males [4]. Beyond  
78 age and gender, CWP is associated with various risk factors, including sleep disorders,  
79 headaches and other types of pain, depression, and illness behavior [5]. The escalating issue of  
80 population aging has exacerbated the financial burden of CWP, leading to significant social  
81 costs. In the US, the direct and indirect annual expenses associated with CWP per patient are  
82 estimated to be approximately \$12,428 [6]. Although the definition of CWP differs slightly from  
83 our focus on widespread pain, it is still meaningful to discuss the current epidemiological and  
84 genetic studies of CWP, as this information could reflect the nature of widespread pain.

85 The genetic basis of CWP has been the subject of extensive research, focusing particularly on  
86 its genetic architecture. Twin studies have played a pivotal role in dissecting the genetic and  
87 environmental contributions to CWP, with estimates indicating that genetic factors account for  
88 approximately 48–52% of the variance in CWP occurrence, underscoring a significant genetic

89 component [7]. In a comprehensive GWAS meta-analysis that amalgamated data from 14  
90 studies, a notable locus was identified on chromosome 5, positioned intergenically between  
91 *CCT5* and *FAM173B* [8]. Furthermore, a recent study by Rahman et al. utilizing the UK Biobank  
92 data on CWP identified the *RNF123* locus as being associated with CWP [9]. This association  
93 was specifically observed in cases reporting pain all over the body and persisting for over three  
94 months.

95 Given the extensive research on CWP, our study shifts the focus to widespread pain  
96 experienced within the last month, a criterion less explored in the literature. This unique  
97 definition allows for a more precise investigation of the genetic underpinnings of short-term  
98 widespread pain, which may differ from those of chronic conditions. The aim of this study is to  
99 identify novel genetic variants linked to widespread pain in the UK Biobank by conducting the  
100 GWAS research. We employ a unique definition that considers pain experienced all over the  
101 body during the past month for the GWAS and utilize three independent replication datasets  
102 (FibroGene, FinnGen and TwinsUK). Sex-stratified GWAS analysis approach is firstly  
103 conducted to analyze the impact of sex on widespread pain. Through this research, we hope to  
104 identify new loci associated with widespread pain, thereby uncovering novel genetic  
105 mechanisms and offering fresh insights for the treatment of this condition.

106

## 107 **Methods**

### 108 **Cohorts' information**

109 The UK Biobank was employed as the discovery cohort in our study. This cohort consists of  
110 over 500,000 participants aged 40-69 from England, Scotland, and Wales, recruited between  
111 2006 and 2011. They underwent clinical examinations, completed comprehensive  
112 questionnaires, and provided DNA samples with informed consent for research use. The study

113 received ethical approval from the UK's National Health Service National Research Ethics  
114 Service (reference 11/NW/0382). Further details can be found at [www.ukbiobank.ac.uk](http://www.ukbiobank.ac.uk).

115 The DNA extraction and quality control (QC) processes were standardized, with detailed  
116 methodologies available at [https://biobank.ctsu.ox.ac.uk/crystal/ukb/docs/genotyping\\_sample  
117 \\_workflow.pdf](https://biobank.ctsu.ox.ac.uk/crystal/ukb/docs/genotyping_sample_workflow.pdf). The Wellcome Trust Centre for Human Genetics at Oxford University oversaw  
118 the standardized QC for genotyping results, with a comprehensive description at [http://biobank  
119 .ctsu.ox.ac.uk/crystal/refer.cgi?id=155580](http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=155580). In July 2017, the UK Biobank released genetic data,  
120 including genotyped and imputed genotypes, from 501,708 samples to authorized researchers.  
121 Bycroft et al. provided a detailed description of the QC procedures for imputation [10].

122 In this study, a specific pain-related question developed by the UK Biobank was utilized: "In the  
123 last month, have you experienced any of the following that interfered with your usual activities?"  
124 Participants could select from the following options: (1) Headache; (2) Facial pain; (3) Neck or  
125 shoulder pain; (4) Back pain; (5) Stomach or abdominal pain; (6) Hip pain; (7) Knee pain; (8)  
126 Pain all over the body; (9) None of the above; (10) Prefer not to say (UK Biobank Questionnaire  
127 field ID: 6159). Multiple selections were allowed. Cases of widespread pain were identified by  
128 the selection of "Pain all over the body," regardless of other choices. Controls were those who  
129 chose "None of the above." The analysis excluded data from individuals not of white British  
130 descent (UK Biobank Questionnaire field ID: 21000) and those with self-reported diagnoses of  
131 rheumatoid arthritis, polymyalgia rheumatica, unspecified arthritis, systemic lupus  
132 erythematosus, ankylosing spondylitis, and myopathy (UK Biobank Questionnaire field ID:  
133 20002).

134 In the replication phase, we employed publicly available summary statistics for the phenotype of  
135 fibromyalgia as classified by ICD-10 in the FinnGen dataset, as well as summary statistics from  
136 the FibroGene fibromyalgia study and the CWP study within TwinsUK. The FinnGen dataset  
137 comprised 3,761 cases diagnosed with fibromyalgia and 375,928 controls without the diagnosis.

138 The GWAS for the FinnGen cohort was conducted according to the protocols of FinnGen  
139 GWAS round 9, with details available at <https://finngen.gitbook.io/documentation> [11]. Detailed  
140 information on sample phenotyping, genotyping, and the GWAS methodology applied to the  
141 FinnGen sample can be found at <https://risteys.finnngen.fi/> [11]. The FibroGene dataset included  
142 905 fibromyalgia cases recruited via advertisements in Fibromyalgia Action UK charity  
143 resources, and 14,213 controls from the general population program of the UK BioResource  
144 project. Details on genotyping and GWAS methodology for the FibroGene dataset can be  
145 referenced at [https://bioresource.nih.ac.uk/media/nlybpdcd/uk\\_axiom\\_biobank\\_genotyping\\_arr](https://bioresource.nih.ac.uk/media/nlybpdcd/uk_axiom_biobank_genotyping_arrays_datasheet.pdf)  
146 [ays\\_datasheet.pdf](https://bioresource.nih.ac.uk/media/nlybpdcd/uk_axiom_biobank_genotyping_arrays_datasheet.pdf) [12]. The TwinsUK GWAS study focusing on CWP comprised 903 individuals  
147 with CWP and 2,369 control subjects The TwinsUK cohort, established as a registry of volunteer  
148 same-sex twins across the United Kingdom, encompasses a broad age range from 16 to 98  
149 years. Initiated in 1992, the cohort has been continuously gathering a wealth of data and  
150 biological materials from its 14,274 registered twins. The ethical integrity of the study is upheld  
151 through approval from the Research Ethics Committee at Guy's and St. Thomas' NHS  
152 Foundation Trust. A significant aspect of the TwinsUK cohort is its repository of genetic  
153 information from 6,921 participants. This genetic data has undergone quality control,  
154 documented by Moayyeri et al [13]. For SNPs missing in the replication datasets, we employed  
155 a strategy based on linkage disequilibrium and physical distance to identify and use the most  
156 appropriate SNP for replication. Additionally, it is important to note that the replication  $p$ -values  
157 have not been adjusted for multiple testing across the 14 genes tested for replication.

## 158 **GWAS and Statistical analysis**

159 Our study comprised three independent GWAS. The primary GWAS was conducted with the  
160 objective of understanding the genetic foundation of widespread pain across the entire dataset.  
161 The secondary GWAS was designed to assess the influence of gender on widespread pain,  
162 with the analysis specifically focused on conducting separate GWAS for each sex. In this study,

163 the Genome-wide Complex Trait Analysis (GCTA, v1.94.1) software (available at  
164 <https://yanglab.westlake.edu.cn/software/gcta/#Overview>) was employed as the principal tool  
165 for executing GWAS [14]. Our research utilized the fastGWA function within GCTA, a mixed  
166 linear model association tool, for the GWAS analyses. Standard QC procedures were  
167 implemented, which involved the exclusion of single nucleotide polymorphisms (SNPs) with  
168 INFO scores below 0.3, SNPs with minor allele frequencies lower than 0.5%, or SNPs failing the  
169 Hardy-Weinberg tests ( $p < 10^{-6}$ ). Furthermore, SNPs located on the X and Y chromosomes, as  
170 well as mitochondrial SNPs, were excluded from the analysis. The association tests were  
171 conducted using the fastGWA function, with adjustments for age, sex, body mass index (BMI),  
172 and eight population principal components. A  $\chi^2$  test was utilized to assess gender differences  
173 between the cases and controls, while the comparison of age and BMI was performed using an  
174 independent t-test, employing R v4.2.2. R v4.2.2 was also used to select the data of white  
175 British descent in UK Biobank data and divide the cases and controls of widespread pain. A *p*-  
176 value less than  $5 \times 10^{-8}$  was considered indicative of genome-wide association significance.  
177 Additionally, GCTA was also employed to compute the narrow-sense heritability.

#### 178 **GWAS-associated analysis by FUMA and LDSC**

179 The FUMA web application served as the primary annotation tool throughout widespread pain  
180 study [15]. Additionally, we generated both Manhattan and Q-Q plots using this application. For  
181 regional visualization, we employed LocusZoom (<http://locuszoom.org/>) [16]. Notably, FUMA  
182 facilitated three key types of analyses, gene analysis, gene-set analysis, and tissue expression  
183 analysis. In the gene analysis, the summary statistics of SNPs were aggregated to the level of  
184 entire genes, allowing for the assessment of associations between genes and the phenotype  
185 under investigation. In the gene-set analysis, specific groups of genes sharing common  
186 biological, functional, or other characteristics were collectively tested, providing insights into the  
187 involvement of distinct biological pathways or cellular functions in the genetic basis of the



188 observed phenotype. The tissue expression analysis was conducted using data from GTEx  
189 (<https://www.gtexportal.org/home>), which has been integrated into the FUMA platform. For this  
190 analysis, the average gene expression per tissue type was used as a gene covariate, enabling  
191 the examination of relationships between gene expression in specific tissue types and the  
192 genetic associations with widespread pain, the focus of our study. We generated two gene  
193 expression heatmaps for genes identified through positional mapping. These heatmaps  
194 represent the average of normalized expression values, employing data from GTEx version 8  
195 across 54 and 30 general tissue types. Tissue specificity was assessed by examining the  
196 differentially expressed genes (DEG) for each tissue type. Furthermore, an enrichment test for  
197 DEGs was performed for both sets of tissue types in GTEx version 8, focusing specifically on  
198 genes identified through positional mapping rather than utilizing the full distribution of SNP p-  
199 values in MAGMA tissue expression analysis.

200 To explore potential genetic correlations between widespread pain and various other  
201 phenotypes, we assessed the genetic associations of widespread pain with 1,396  
202 characteristics from the UK Biobank through the Complex Trait Virtual Lab (CTG-VL)  
203 (<https://genoma.io/>). CTG-VL is an open-source platform that integrates available GWAS  
204 datasets to enable the estimation of genetic correlations for complex traits. Subsequently, to  
205 further investigate any potential genetic differences between the sexes, we utilized linkage  
206 disequilibrium score regression (LDSC) via LDSCv1.0.1 (available at  
207 <https://github.com/bulik/ldsc>) to examine the genetic correlation of widespread pain between  
208 males and females. LDSC utilizes the principle of linkage disequilibrium, which refers to the  
209 non-random association of alleles at different loci within a population [17].

210 **Expression quantitative trait loci (eQTL), chromatin interaction analysis and positional**  
211 **mapping**

212 Expression quantitative trait loci (eQTL) have become instrumental in elucidating the regulatory  
213 mechanisms of variants identified through GWAS. Cis-eQTL notably impact gene expression by  
214 interacting with variants that are situated in proximity (within 1 Mb) to the gene [18]. In  
215 eukaryotic cells, the genome is compactly organized within the micron-sized nucleus, with  
216 chromatin serving as the fundamental structural unit. This organization is essential, arranging  
217 the genome into a three-dimensional structure that is pivotal for processes such as DNA  
218 replication, DNA damage repair, gene transcription, and other crucial biological functions [19].  
219 Positional mapping within our study was conducted using a maximum distance of 10 kb. The  
220 integration of cis-eQTL analysis, chromatin interaction analysis, and positional mapping in our  
221 gene mapping approach offers a comprehensive perspective on the genomic architecture  
222 underpinning these findings.

### 223 **Phenome-wide association analysis (PheWAS)**

224 Our study implemented a Phenome-Wide Association Analysis (PheWAS) to investigate the  
225 relationships between significant SNP associations, their corresponding genes, and multiple  
226 phenotypes. The primary objectives of this analysis were to corroborate the associations  
227 identified in our GWAS with pain phenotypes and to uncover new associations between genetic  
228 variants linked to widespread pain and other phenotypes. The PheWAS was conducted utilizing  
229 an expansive dataset comprising 4,756 GWAS summary statistics available on the GWAS  
230 ATLAS platform <https://atlas.ctglab.nl/PheWAS> [20]. For this analysis, only SNPs exhibiting  
231 associations with  $p$ -values lower than 0.05 were included. Furthermore, adjustments for multiple  
232 comparisons were performed using the Bonferroni correction method.

233

## 234 **Results**

### 235 **GWAS results**

236 During the initial assessment phase (2006-2010) of the UK Biobank study, a total of 501,708  
237 participants were administered a pain questionnaire. In our research, we excluded non-white  
238 British participants and samples that failed to meet QC criteria. In the primary GWAS, we  
239 identified 4,617 cases (comprising 1,885 males and 2,732 females) and 167,613 controls  
240 (consisting of 80,228 males and 87,385 females) for analysis. The study utilized 11,165,459  
241 SNPs for the GWAS examination. The secondary phase of the study involved sex-stratified  
242 GWAS analyses, where, following identical QC procedures, the female cohort (90,117 samples)  
243 consisted of 2,732 cases and 87,385 controls, while the male cohort (82,113 samples) included  
244 1,885 cases and 80,228 controls. Table 1 provides a comprehensive summary of the clinical  
245 characteristics of cases and controls in the primary GWAS. Table 2 presents a detailed  
246 overview of the cases and controls in the secondary sex-stratified GWAS.

247 In our study, we identified one novel and distinct SNP cluster on chromosome 5 that exhibited  
248 significant associations with widespread pain, achieving genome-wide significance ( $p < 5 \times 10^{-8}$ ),  
249 as delineated in Figure 1. Additionally, we identified several significant loci proximal to the  
250 genome-wide significant threshold, which are listed in Table 3. This table enumerates a total of  
251 14 SNPs, encompassing both significant and marginally significant SNPs, with each  
252 representing the most statistically significant association within its respective locus. A detailed  
253 enumeration of all significantly associated SNPs from this GWAS is furnished in Supplementary  
254 Table 1. Notably, the significant association was observed in the SNP cluster situated within the  
255 *ARHGEF28* gene on chromosome 5q13.2, with a  $p$ -value of  $1.76 \times 10^{-8}$  for rs34691025. The  
256 regional plot illustrating the most significant locus in *ARHGEF28* is provided in Figure 2.  
257 Moreover, the Q-Q plot of the GWAS during the discovery phase is presented in Figure 3. The  
258 SNP-based heritability for widespread pain was estimated to be 0.18, with a standard error of  
259 0.02.

260 We tested the significant and independent SNPs identified in our discovery analysis, as well as  
261 other SNPs proximal to the genome-wide significance threshold, for replication in the FinnGen,  
262 FibroGene, and Twins UK cohorts. Within the TwinsUK cohort, specifically regarding CWP, we  
263 observed weak replication for rs761429057 in the *TMCC1* gene ( $p = 0.017$ ) among the 14 SNPs  
264 from the discovery cohort. However, in the context of fibromyalgia within the FibroGene and  
265 FinnGen cohorts, none of the 14 SNPs demonstrated replication in either cohort ( $p > 0.05$ ). The  
266  $p$ -values of the associations of these 14 independent SNPs from the discovery stage were  
267 extracted from FinnGen, FibroGene and Twins UK cohorts are presented in Table 3.

268 In the sex-stratified GWAS analyses, the male-specific GWAS did not identify any single locus  
269 of genome-wide significance associated with widespread pain. Conversely, the female-specific  
270 GWAS revealed one significant locus associated with widespread pain, which differed from the  
271 findings of the primary GWAS. This novel locus was identified within the *LRMDA* gene on  
272 chromosome 10, with a  $p = 4.04 \times 10^{-8}$  for rs137998089. Detailed information regarding these  
273 findings is provided in Table 4, and the corresponding Manhattan plot and regional plot are  
274 depicted in Figure 4 and 5, respectively.

### 275 **Gene, gene-set and tissue expression analysis**

276 In our gene analysis, we mapped SNPs within genes to a total of 18,115 protein-coding genes.  
277 As a result, we established a genome-wide significance threshold of  $p < 0.05 / 18,115 = 2.76 \times$   
278  $10^{-6}$ . At this stage, no genes met the criteria for significance. In our analysis of gene sets, we  
279 evaluated 17,109 gene sets in total. The threshold for statistical significance on a genome-wide  
280 scale was thus set at  $p < 0.05 / 17,109 = 2.92 \times 10^{-6}$ . However, none of the gene sets reached  
281 statistical significance. The top ten gene sets identified in this analysis are listed in  
282 Supplementary Table 2.

283 In the tissue expression analysis, we identified 30 general tissue types and 53 specific tissue  
284 types. However, none of the tissues exhibited a significant association. Further details and  
285 visualization of these results are available in Supplementary Figure 1 and 2.

286 In the gene expression heatmaps presented in Figure 6, the gene *ARHGEF28* exhibits higher  
287 expression levels in the nerve, oral, and adrenal gland tissues compared to other tissue types.  
288 Tissue specificity testing did not yield significant differences in the DEG across both the 54 and  
289 30 tissue types analyzed in GTEx version 8. Detailed visualizations of these findings are  
290 included in Supplementary Figure 3.

### 291 **Genetic correlation analysis by LDSC**

292 In our study examining the relationship between widespread pain and various pain phenotypes,  
293 as well as the genetic correlation between widespread pain in females and males, we  
294 discovered several significant genetic correlations, as depicted in Figure 7. After correcting for  
295 multiple testing, the three most notable genetic correlations ( $r_g$ ) were found in the following  
296 categories: other joint disorders, not elsewhere classified ( $r_g = 0.94$ ,  $p = 6.20 \times 10^{-6}$ ),  
297 spondylosis ( $r_g = 0.93$ ,  $p = 7.73 \times 10^{-6}$ ), and other diseases of the anus and rectum ( $r_g = 0.89$ ,  $p$   
298  $= 1.95 \times 10^{-6}$ ). The genetic correlation between widespread pain in females and males was  
299 determined to be  $r_g = 0.80$  ( $p = 7.36 \times 10^{-92}$ ). The comprehensive results are provided in  
300 Supplementary Table 3.

### 301 **Expression quantitative trait loci (eQTL), chromatin interaction analysis and positional** 302 **mapping**

303 From the cis-eQTL analysis, SNPs rs35588745 and rs71636094 were found to be significantly  
304 associated with the tissue type Cells\_Cultured\_fibroblasts, achieving a False Discovery Rate  
305 (FDR) of less than 0.05. The results of the chromatin interactions and cis-eQTL analyses are  
306 illustrated in the circus plot presented in Figure 8. This plot shows the genes that interact with

307 *ARHGEF28*, including *IQGAP2*, *GCTN4*, *FAM169A*, *NSA2*, *GFM2*, *HEXB*, *ENC1*, *UTP14*, and  
308 *AC008387.1*. Additionally, the positional mapping analysis did not identify significant results.

### 309 **PheWAS**

310 PheWAS were conducted using the GWAS ATLAS platform to explore phenotypes associated  
311 with independently significant SNPs (rs34691025) and significantly associated genes  
312 (*ARHGEF28*). SNP rs34691025 exhibited significant associations with the frequency of  
313 tenseness/restlessness in the last two weeks ( $p = 1.5 \times 10^{-4}$ ), the mode of anisotropy in the  
314 retrolenticular part of the internal capsule ( $p = 1.70 \times 10^{-6}$ ), and long-standing illness, disability,  
315 or infirmity ( $p = 3.60 \times 10^{-12}$ ). The gene *ARHGEF28* was associated with hearing  
316 difficulty/problems ( $p = 1.09 \times 10^{-23}$ ), high blood pressure ( $p = 3.42 \times 10^{-7}$ ), and hypertension ( $p$   
317  $= 2.26 \times 10^{-5}$ ). All significant traits that passed the Bonferroni correction are listed in  
318 Supplementary Table 4. The plots illustrating phenotypes associated with the SNP and gene are  
319 visualized in Figure 9 and 10, respectively.

320

### 321 **Discussion**

322 In this GWAS of widespread pain, conducted using the comprehensive UK Biobank dataset, we  
323 identified one novel significant genetic locus and several loci that marginally reached the  
324 genome-wide significance level associated with self-reported widespread pain during the past  
325 month. In the sex-stratified GWAS, we observed differences in the prevalence of genetic  
326 variants of the widespread pain between males and females. Furthermore, our post-GWAS  
327 analysis revealed significant genetic correlations between widespread pain and other  
328 phenotypes. Additionally, a PheWAS was conducted to further explore the genetic mechanisms  
329 underlying widespread pain. In this study, we utilized the generic pain questionnaire from the  
330 UK Biobank, a valuable screening instrument, to investigate the potential genetic underpinnings

331 of heterogeneous pain phenotypes, including widespread pain. Employing the UK Biobank  
332 offers researchers the advantage of mitigating potential issues related to reduced power due to  
333 phenotype heterogeneity. This is achieved by leveraging substantial sample sizes, thereby  
334 enhancing the clarity of statistical results amidst potential noise.

335 In our primary GWAS, we have analyzed and subsequently identified one novel locus that was  
336 associated with widespread pain. This significant locus was situated within the *ARHGEF28*  
337 gene on chromosome 5q13.2. This particular locus exhibited a significant  $p$ -value of  $1.76 \times 10^{-8}$   
338 for the marker rs34691025. This locus spans an extensive 11 kb within the intronic region of the  
339 *ARHGEF28* gene and encompasses as many as 9 genome-wide significant SNPs, as detailed  
340 in Supplementary Table 1. The *ARHGEF28* gene (Rho Guanine Nucleotide Exchange Factor 28)  
341 is a member of the Rho guanine nucleotide exchange factor family. It plays a pivotal role in  
342 regulating Rho GTPases involved in cytoskeletal reorganization and cell migration, may  
343 influence pain perception and sensitivity through its effects on neuronal plasticity and  
344 inflammatory processes. The cytoskeletal dynamics in neurons, particularly in nociceptive  
345 pathways, play a vital role in the development and maintenance of pain states, with alterations  
346 in these processes leading to changes in the structural and functional plasticity of pain pathways,  
347 thereby impacting pain perception. [21, 22]. Furthermore, *ARHGEF28*'s involvement in immune  
348 cell signaling pathways suggests a mechanism by which genetic predispositions could influence  
349 the development of chronic pain through immune-mediated inflammation [23, 24]. The link  
350 between the immune system and pain has been increasingly recognized, with several studies  
351 highlighting how immune responses contribute to the sensitization of the central nervous system,  
352 leading to pain conditions [25, 26]. The role of *ARHGEF28* in modulating immune cell functions,  
353 including those of mast cells and macrophages, which are key players in inflammation and pain,  
354 further supports its potential involvement in pain phenotypes. This dual role in neuronal and  
355 immune cell function underscores the complexity of the genetic underpinnings of pain and the

356 need for multidisciplinary approaches to fully understand these mechanisms. The association of  
357 *ARHGEF28* with widespread pain underscores the importance of genetic factors in the  
358 development of pain disorders and highlights new potential pathways for therapeutic  
359 intervention. Further research is needed to explore the functional impact of *ARHGEF28*  
360 variations on neuronal and immune cell functions and to understand how these effects  
361 contribute to the manifestation of widespread pain. A study identified a heterozygous mutation in  
362 amyotrophic lateral sclerosis patients that is predicted to generate a premature truncated  
363 *ARHGEF28* gene product, which could provide a possible link between this gene and  
364 widespread pain [27]. Moreover, a previous GWAS research has reported a significant  
365 association between the *ARHGEF28* gene and hearing problems, suggesting a potential new  
366 pathway linking widespread pain and auditory issues [28].

367 We further identified a cluster of SNPs within the *FAF1* gene on chromosome 1, displaying  
368 marginal genome-wide significance, with rs17387024 exhibiting a minimal  $p$ -value of  $7.87 \times 10^{-8}$ .  
369 The *FAF1* gene, formally known as Fas Associated Factor 1, is a protein-coding gene. The  
370 association between the *FAF1* gene and widespread pain may be attributed to its involvement in  
371 various biological processes relevant to pain perception and modulation. *FAF1* has been  
372 implicated in apoptosis, inflammation, and neuronal signaling, all of which are crucial in the  
373 pathogenesis of pain [29]. Studies have shown that *FAF1* is involved in the NF- $\kappa$ B signaling  
374 pathway, which plays a significant role in inflammation and pain [30, 31]. Additionally, *FAF1* has  
375 been associated with neuronal death and dysfunction, suggesting a potential link to neuropathic  
376 pain [32]. The interaction of *FAF1* with other proteins involved in pain pathways, such as *TRPV1*,  
377 further supports its role in pain perception [33]. Several previous GWAS have identified the  
378 *FAF1* gene as being associated with multisite chronic pain and general pain, providing a  
379 possible basis for its link with widespread pain [34-36]. However, further research is still  
380 necessary to fully elucidate the exact mechanism by which *FAF1* influences widespread pain.



381 In the sex-stratified GWAS, we discovered a novel locus unique to the female-specific analysis,  
382 located within the *LRMDA* gene on chromosome 10. The top SNP, rs5779595, exhibited a  
383 significant  $p$ -value of  $4.04 \times 10^{-8}$ . The *LRMDA* gene, also known as leucine-rich melanocyte  
384 differentiation-associated protein, has been implicated in some biological processes that could  
385 underlie its relationship with pain. Notably, *LRMDA* is involved in melanocyte differentiation and  
386 pigmentation, which might suggest modulating nociceptive pathways [37]. Furthermore, several  
387 GWAS studies have shown that the *LRMDA* gene is associated with diabetes, a condition that  
388 can potentially cause widespread pain and fibromyalgia [38, 39]. This association provides a  
389 potential link between the *LRMDA* gene and widespread pain, suggesting that the gene may  
390 play a role in the underlying mechanisms of the widespread pain condition. The genetic  
391 correlation for widespread pain between males and females was also calculated ( $r_g = 0.80$ ),  
392 indicating a potential disparity in the genetic underpinnings of widespread pain across genders.  
393 This divergence is likely attributable to the significant influence of gender differences on the  
394 heterogeneity of pain [40]. These findings highlight the necessity of accounting for biological  
395 differences between males and females in research on the genetics of pain.

396 Building on our research findings, there are significant genetic correlations between widespread  
397 pain and several other phenotypes, suggesting that shared genetic factors underpinning these  
398 conditions. The top three phenotypes most genetically correlated with widespread pain were  
399 other joint disorders, not elsewhere classified ( $r_g = 0.94$ ,  $p = 6.20 \times 10^{-6}$ ), spondylosis ( $r_g = 0.93$ ,  
400  $p = 7.73 \times 10^{-6}$ ), and other diseases of the anus and rectum ( $r_g = 0.89$ ,  $p = 1.95 \times 10^{-6}$ ). Our  
401 study provides evidence of the complex interplay of genetic determinants in the manifestation of  
402 pain, supporting previous research findings. Specifically, several prior analyses have identified  
403 an association between CWP and joint diseases, and some studies suggest that fibromyalgia  
404 may share common mechanisms with disorders of the anus and rectum [41, 42]. In the  
405 PheWAS, the results reveal that the top SNP (rs34691025) and corresponding gene

406 (*ARHGEF28*) are significantly associated with hearing disorders and cardiovascular diseases.  
407 This finding may provide additional evidence that these disorders could be significant  
408 contributors to widespread pain, or conversely, that widespread pain may influence the  
409 development of these conditions. The shared genetic determinants, correlations with other pain  
410 phenotypes, and associations with various medical conditions and medications highlight the  
411 necessity of adopting a holistic approach to comprehensively understanding and addressing  
412 widespread pain.

413 In our research, we employed a novel operational definition to delineate cases of widespread  
414 pain, characterizing them as individuals experiencing pain throughout their body during the last  
415 month. This approach stands in contrast to previous GWAS that defined cases based on the  
416 presence of pain all over the body lasting for more than three months [9]. Notably, the CWP  
417 cases in these studies numbered 6,914, a criterion that differs from our own. The adoption of  
418 this alternative definition was driven by the consideration that a one-month duration might  
419 capture a broader spectrum of pain phenotypes, potentially including both acute and chronic  
420 manifestations. By focusing on a one-month timeframe, we intend to provide a distinct and  
421 meaningful delineation of widespread pain cases, offering a new perspective on the genetic  
422 underpinnings of pain phenotypes. Our choice of the definition is motivated by the goal of  
423 enhancing the understanding of genetic factors associated with widespread pain.

424 Our research provides initial insights into the genetic underpinnings of widespread pain.  
425 However, it encounters certain limitations, particularly in the validation phase of our results.  
426 During the replication phase, we utilized three public datasets of CWP and fibromyalgia, yet we  
427 did not achieve significant replication results for our study. This outcome could potentially be  
428 attributed to differences in the definitions of widespread pain, CWP, and fibromyalgia.  
429 Additionally, the limited number of participants in the replication cohorts might have contributed  
430 to this challenge. In the female-specific GWAS, rs5779595 becomes the top and only SNP

431 identified within the locus, pointing to the potential spuriousness of this association and  
432 underscoring the need for further exploration and study. Furthermore, it is essential to discuss  
433 the methodological approach adopted in our study. We characterized widespread pain cases  
434 and controls based on the responses provided by UK Biobank participants to a specific question  
435 regarding pain experienced over the last month. However, a critical observation is that this  
436 question did not delve into nuanced details such as the severity or frequency of the pain.  
437 Consequently, the phenotype definition derived from these responses should be perceived and  
438 interpreted as being broadly defined. This broad definition may have implications for the  
439 specificity and sensitivity of our findings, as it potentially includes a heterogeneous group of  
440 individuals with varying pain experiences. Future research may need to incorporate more  
441 detailed questionnaires and larger sample sizes to capture the complexity of widespread pain  
442 and provide a more comprehensive understanding of its genetic basis. Additionally, exploring  
443 alternative phenotype definitions that account for pain severity and frequency could enhance the  
444 precision of genetic associations and contribute to a deeper understanding of the mechanisms  
445 underlying widespread pain.

446

## 447 **Conclusion**

448 In conclusion, our primary GWAS of widespread pain, utilizing the extensive UK Biobank  
449 dataset, successfully identified one novel genetic locus and several loci that reached genome-  
450 wide significance. The sex-stratified GWAS outputs revealed gender-specific variants  
451 associated with of widespread pain between males and females. Furthermore, our post-GWAS  
452 analysis demonstrated substantial genetic correlations and shared mechanisms between  
453 widespread pain and other phenotypes. This study advances our understanding of the genetic  
454 factors contributing to widespread pain.

455

456 **Declarations**

457 **Funding**

458 This study was mainly funded by the Pioneer and Leading Goose R&D Program of Zhejiang  
459 Province 2023 with reference number 2023C04049 and Ningbo International Collaboration  
460 Program 2023 with reference number 2023H025. TwinsUK is funded by the Wellcome Trust,  
461 Medical Research Council, Versus Arthritis, European Union Horizon 2020, Chronic Disease  
462 Research Foundation (CDRF), Zoe Ltd and the National Institute for Health Research (NIHR)  
463 Clinical Research Network (CRN) and Biomedical Research Centre based at Guy's and St  
464 Thomas' NHS Foundation Trust in partnership with King's College London.

465 **Data availability**

466 The summary statistics of the UK Biobank results on widespread pain can be accessed upon  
467 publication. Any other data relevant to the study that are not included in the article or its  
468 supplementary materials are available from the authors upon reasonable request.

469 **Acknowledgements**

470 The authors extend their sincere gratitude to the participants of the UK Biobank, FinnGen,  
471 FibroGene, and TwinsUK cohorts for their invaluable contributions of genetic and phenotypic  
472 data. This study will adhere to all ethical guidelines and data protection protocols of the UK  
473 Biobank. This research has been conducted using the UK Biobank Resource under Application  
474 Number 89386.

475 **Competing interests**

476 The authors declare no competing interests.

477 **Authors' contributions**

478 QP drafted the paper and performed the UK Biobank GWAS analysis. TC and YT contributed to  
479 data formatting. MH and TD provided comments to the paper. RC, MKN and FMKW conducted  
480 the replication study and provided comments to the paper. WM organized the project and  
481 provided comments.

482 **Corresponding authors**

483 Correspondence to Weihua Meng

484 **Consent for publication**

485 All authors have consent for publication.

486 **Ethics approval and consent to participate**

487 This study was approved by the Ethics Committee of the University of Nottingham, Ningbo,  
488 China. All authors have consent for participation. The TwinsUK Study was approved by London-  
489 Westminster Research Ethics Committee (REC reference EC04/015), and Guy's and St Thomas'  
490 NHS Foundation Trust Research and Development (R&D). The TwinsUK Biobank was  
491 approved by the HRA - Liverpool East Research Ethics Committee (REC reference  
492 19/NW/0187), IRAS ID 258513. All participants provide written, informed consent.

493

494 **References**

495

- 496 1. Butler S, Landmark T, Glette M, Borchgrevink P, Woodhouse A: **Chronic widespread pain—the**  
497 **need for a standard definition.** *PAIN* 2016, **157**(3):541-543.
- 498 2. Wolfe F, Butler SH, Fitzcharles M, Häuser W, Katz RL, Mease PJ, Rasker JJ, Russell AS, Russell IJ,  
499 Walitt B: **Revised chronic widespread pain criteria: development from and integration with**  
500 **fibromyalgia criteria.** *Scandinavian Journal of Pain* 2020, **20**(1):77-86.
- 501 3. Fayaz A, Croft P, Langford RM, Donaldson LJ, Jones GT: **Prevalence of chronic pain in the UK: a**  
502 **systematic review and meta-analysis of population studies.** *BMJ Open* 2016, **6**(6):e010364.

- 503 4. Mansfield KE, Sim J, Jordan JL, Jordan KP: **A systematic review and meta-analysis of the**  
504 **prevalence of chronic widespread pain in the general population.** *Pain* 2016, **157**(1):55-64.
- 505 5. Creed F: **A review of the incidence and risk factors for fibromyalgia and chronic widespread**  
506 **pain in population-based studies.** *PAIN* 2020, **161**(6):1169-1176.
- 507 6. Schaefer C, Mann R, Masters ET, Cappelleri JC, Daniel SR, Zlateva G, McElroy HJ, Chandran AB,  
508 Adams EH, Assaf AR *et al*: **The Comparative Burden of Chronic Widespread Pain and**  
509 **Fibromyalgia in the United States.** *Pain Practice* 2016, **16**(5):565-579.
- 510 7. Kato K, Sullivan PF, Evengård B, Pedersen NL: **Importance of genetic influences on chronic**  
511 **widespread pain.** *Arthritis Rheum* 2006, **54**(5):1682-1686.
- 512 8. Peters MJ, Broer L, Willems HL, Eiriksdottir G, Hocking LJ, Holliday KL, Horan MA, Meulenbelt I,  
513 Neogi T, Popham M *et al*: **Genome-wide association study meta-analysis of chronic widespread**  
514 **pain: evidence for involvement of the 5p15.2 region.** *Ann Rheum Dis* 2013, **72**(3):427-436.
- 515 9. Rahman MS, Winsvold BS, Chavez Chavez SO, Børte S, Tsepilov YA, Sharapov SZ, Aulchenko YS,  
516 Hagen K, Fors EA, Hveem K *et al*: **Genome-wide association study identifies RNF123 locus as**  
517 **associated with chronic widespread musculoskeletal pain.** *Ann Rheum Dis* 2021, **80**(9):1227-  
518 1235.
- 519 10. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O,  
520 O'Connell J *et al*: **The UK Biobank resource with deep phenotyping and genomic data.** *Nature*  
521 2018, **562**(7726):203-209.
- 522 11. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, Reeve MP, Laivuori H,  
523 Aavikko M, Kaunisto MA *et al*: **FinnGen provides genetic insights from a well-phenotyped**  
524 **isolated population.** *Nature* 2023, **613**(7944):508-518.
- 525 12. Eslam M, Hashem AM, Romero-Gomez M, Berg T, Dore GJ, Mangia A, Chan HLY, Irving WL,  
526 Sheridan D, Abate ML *et al*: **FibroGENE: A gene-based model for staging liver fibrosis.** *J Hepatol*  
527 2016, **64**(2):390-398.
- 528 13. Moayyeri A, Hammond CJ, Hart DJ, Spector TD: **The UK Adult Twin Registry (TwinsUK Resource).**  
529 *Twin Res Hum Genet* 2013, **16**(1):144-149.
- 530 14. Yang J, Lee SH, Goddard ME, Visscher PM: **GCTA: a tool for genome-wide complex trait analysis.**  
531 *Am J Hum Genet* 2011, **88**(1):76-82.
- 532 15. Watanabe K, Taskesen E, van Bochoven A, Posthuma D: **Functional mapping and annotation of**  
533 **genetic associations with FUMA.** *Nat Commun* 2017, **8**(1):1826.
- 534 16. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer  
535 CJ: **LocusZoom: regional visualization of genome-wide association scan results.** *Bioinformatics*  
536 2010, **26**(18):2336-2337.
- 537 17. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N, Daly MJ, Price AL, Neale BM:  
538 **LD Score regression distinguishes confounding from polygenicity in genome-wide association**  
539 **studies.** *Nat Genet* 2015, **47**(3):291-295.
- 540 18. Võsa U, Claringbould A, Westra HJ, Bonder MJ, Deelen P, Zeng B, Kirsten H, Saha A, Kreuzhuber  
541 R, Yazar S *et al*: **Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and**  
542 **polygenic scores that regulate blood gene expression.** *Nat Genet* 2021, **53**(9):1300-1310.
- 543 19. Li G, Sun T, Chang H, Cai L, Hong P, Zhou Q: **Chromatin Interaction Analysis with Updated ChIA-**  
544 **PET Tool (V3).** *Genes (Basel)* 2019, **10**(7).
- 545 20. Watanabe K, Stringer S, Frei O, Umičević Mirkov M, de Leeuw C, Polderman TJC, van der Sluis S,  
546 Andreassen OA, Neale BM, Posthuma D: **A global overview of pleiotropy and genetic**  
547 **architecture in complex traits.** *Nat Genet* 2019, **51**(9):1339-1348.
- 548 21. St John Smith E: **Advances in understanding nociception and neuropathic pain.** *J Neurol* 2018,  
549 **265**(2):231-238.

- 550 22. Sit ST, Manser E: **Rho GTPases and their role in organizing the actin cytoskeleton.** *J Cell Sci* 2011,  
551 **124**(Pt 5):679-683.
- 552 23. Song Y, Lin F, Ye CH, Huang H, Li X, Yao X, Xu Y, Wang C: **Rare, low-frequency and common**  
553 **coding variants of ARHGEF28 gene and their association with sporadic amyotrophic lateral**  
554 **sclerosis.** *Neurobiol Aging* 2020, **87**:138.e131-138.e136.
- 555 24. Malafoglia V, Ilari S, Vitiello L, Tenti M, Balzani E, Muscoli C, Raffaeli W, Bonci A: **The Interplay**  
556 **between Chronic Pain, Opioids, and the Immune System.** *The Neuroscientist* 2022, **28**(6):613-  
557 627.
- 558 25. Raoof R, Willemen HLDM, Eijkelkamp N: **Divergent roles of immune cells and their mediators in**  
559 **pain.** *Rheumatology* 2017, **57**(3):429-440.
- 560 26. Vergne-Salle P, Bertin P: **Chronic pain and neuroinflammation.** *Joint Bone Spine* 2021,  
561 **88**(6):105222.
- 562 27. Droppelmann CA, Wang J, Campos-Melo D, Keller B, Volkening K, Hegele RA, Strong MJ:  
563 **Detection of a novel frameshift mutation and regions with homozygosis within ARHGEF28**  
564 **gene in familial amyotrophic lateral sclerosis.** *Amyotroph Lateral Scler Frontotemporal Degener*  
565 2013, **14**(5-6):444-451.
- 566 28. Ivarsdottir EV, Holm H, Benonisdottir S, Olafsdottir T, Sveinbjornsson G, Thorleifsson G,  
567 Eggertsson HP, Halldorsson GH, Hjorleifsson KE, Melsted P *et al*: **The genetic architecture of**  
568 **age-related hearing impairment revealed by genome-wide association analysis.** *Commun Biol*  
569 2021, **4**(1):706.
- 570 29. Ryu SW, Lee SJ, Park MY, Jun JI, Jung YK, Kim E: **Fas-associated factor 1, FAF1, is a member of**  
571 **Fas death-inducing signaling complex.** *J Biol Chem* 2003, **278**(26):24003-24010.
- 572 30. Ahmed AS, Berg S, Alkass K, Druid H, Hart DA, Svensson CI, Kosek E: **NF- $\kappa$ B-Associated Pain-**  
573 **Related Neuropeptide Expression in Patients with Degenerative Disc Disease.** *Int J Mol Sci*  
574 2019, **20**(3).
- 575 31. Park MY, Jang HD, Lee SY, Lee KJ, Kim E: **Fas-associated factor-1 inhibits nuclear factor-kappaB**  
576 **(NF-kappaB) activity by interfering with nuclear translocation of the RelA (p65) subunit of NF-**  
577 **kappaB.** *J Biol Chem* 2004, **279**(4):2544-2549.
- 578 32. Park G, Kim B-S, Kim E: **A novel function of FAF1, which induces dopaminergic neuronal death**  
579 **through cell-to-cell transmission.** *Cell Communication and Signaling* 2020, **18**(1):133.
- 580 33. Kim S, Kang C, Shin CY, Hwang SW, Yang YD, Shim WS, Park MY, Kim E, Kim M, Kim BM *et al*:  
581 **TRPV1 recapitulates native capsaicin receptor in sensory neurons in association with Fas-**  
582 **associated factor 1.** *J Neurosci* 2006, **26**(9):2403-2412.
- 583 34. Johnston KJA, Adams MJ, Nicholl BI, Ward J, Strawbridge RJ, Ferguson A, McIntosh AM, Bailey  
584 MES, Smith DJ: **Genome-wide association study of multisite chronic pain in UK Biobank.** *PLoS*  
585 *Genet* 2019, **15**(6):e1008164.
- 586 35. Johnston KJA, Ward J, Ray PR, Adams MJ, McIntosh AM, Smith BH, Strawbridge RJ, Price TJ,  
587 Smith DJ, Nicholl BI *et al*: **Sex-stratified genome-wide association study of multisite chronic**  
588 **pain in UK Biobank.** *PLoS Genet* 2021, **17**(4):e1009428.
- 589 36. Mocchi E, Ward K, Perry JA, Starkweather A, Stone LS, Schabrun SM, Renn C, Dorsey SG, Ament  
590 SA: **Genome wide association joint analysis reveals 99 risk loci for pain susceptibility and**  
591 **pleiotropic relationships with psychiatric, metabolic, and immunological traits.** *PLoS Genet*  
592 2023, **19**(10):e1010977.
- 593 37. Fernández A, Hayashi M, Garrido G, Montero A, Guardia A, Suzuki T, Montoliu L: **Genetics of**  
594 **non-syndromic and syndromic oculocutaneous albinism in human and mouse.** *Pigment Cell*  
595 *Melanoma Res* 2021, **34**(4):786-799.
- 596 38. Vujkovic M, Keaton JM, Lynch JA, Miller DR, Zhou J, Tcheandjieu C, Huffman JE, Assimes TL,  
597 Lorenz K, Zhu X *et al*: **Discovery of 318 new risk loci for type 2 diabetes and related vascular**

- 598 **outcomes among 1.4 million participants in a multi-ancestry meta-analysis.** *Nat Genet* 2020,  
599 **52(7):680-691.**
- 600 39. Ishigaki K, Akiyama M, Kanai M, Takahashi A, Kawakami E, Sugishita H, Sakaue S, Matoba N, Low  
601 S-K, Okada Y *et al*: **Large-scale genome-wide association study in a Japanese population**  
602 **identifies novel susceptibility loci across different diseases.** *Nature Genetics* 2020, **52(7):669-**  
603 **679.**
- 604 40. Gao Q, Liu MQ, Li JX, Wang Y, Zhang Y, Zhu H: **Sex differences in stress-induced hyperalgesia**  
605 **and its mechanisms.** *J Neurosci Res* 2024, **102(1):e25266.**
- 606 41. Mayer TG, Towns BL, Neblett R, Theodore BR, Gatchel RJ: **Chronic Widespread Pain in Patients**  
607 **With Occupational Spinal Disorders: Prevalence, Psychiatric Comorbidity, and Association**  
608 **With Outcomes.** *Spine* 2008, **33(17):1889-1897.**
- 609 42. Caldarella MP, Giamberardino MA, Sacco F, Affaitati G, Milano A, Lerza R, Balatsinou C, Laterza F,  
610 Pierdomenico SD, Cuccurullo F *et al*: **Sensitivity Disturbances in Patients with Irritable Bowel**  
611 **Syndrome and Fibromyalgia.** *Official journal of the American College of Gastroenterology | ACG*  
612 2006, **101(12):2782-2789.**

613

#### 614 **Figure legends:**

- 615 **Fig.1** The Manhattan plot of the GWAS analysis on widespread pain (N = 172,230)  
616 The dashed grey line indicates the cut-off  $p$ -value of  $5 \times 10^{-8}$
- 617 **Fig.2** The regional plots of locus in *ARHGEF28* region
- 618 **Fig.3** The Q-Q plot of the GWAS analysis on widespread pain (N = 172,230)
- 619 **Fig.4** The Manhattan plot of the female-specific GWAS analysis on widespread pain (N =  
620 90,117)  
621 The dashed grey line indicates the cut-off  $p$ -value of  $5 \times 10^{-8}$
- 622 **Fig.5** The regional plots of locus in *LRMDA* region
- 623 **Fig.6** Gene expression heatmaps across 54 tissue types and 30 general tissue  
624 Categories from GTEx Version 8
- 625 **Fig.7** Genetic correlation results for widespread pain using LDSC on CTG-VL. For a full  
626 list, see Supplementary Table 3
- 627 **Fig.8** Circos plot illustrating chromatin interactions and eQTLs on chromosome 5
- 628 **Fig.9** PheWAS results for SNP rs34691025 within primary widespread pain GWAS
- 629 **Fig.10** PheWAS results for gene *ARHGEF28* within primary widespread pain GWAS



**Table 1** Clinical characteristics of widespread pain cases and controls in the UK Biobank

Covariates	Cases	Controls	<i>p</i> -value
Sex (male: female)	1,885:2,732	80,228:87,385	<0.001
Age (years)	56.7 (7.67)	56.9 (7.95)	0.55
BMI (kg/m <sup>3</sup> )	29.6 (5.77)	26.7 (4.28)	<0.001

A  $\chi^2$  test was used to test the difference of gender frequency between cases and controls and an independent t test was used for other covariates.

Continuous covariates were presented as mean (standard deviation)

BMI represents body mass index.

**Table 2** Clinical characteristics of widespread pain cases and controls in the UK Biobank (sex-stratified GWAS)

GWAS Analysis	Covariates	Cases	Controls	<i>p</i> -value
Male-specific GWAS	Age (years)	57.3 (7.60)	57.2 (8.04)	0.054
	BMI (kg/m <sup>3</sup> )	29.6 (5.11)	27.3 (3.90)	<0.001
Female-specific GWAS	Age (years)	57.2 (7.58)	56.8 (7.87)	<0.05
	BMI (kg/m <sup>3</sup> )	29.8 (6.36)	26.2 (4.57)	<0.001

Total number of females: 90,117; Total number of males: 82,113. An independent t test was used for other covariates.

Continuous covariates were presented as mean (standard deviation)

BMI represents body mass index.

**Table 3** The Novel Significant SNP and SNPs Approaching Significance Identified by GWAS on Widespread Pain

rsID	Chr	SNP position	Gene /Nearest Gene	Effect allele	Discovery Stage (UK Biobank)				Replication Stage (Henkel et al. and FinnGen)		
					Non effective allele	Frequency the effect allele	Beta	p-value	p-value (beta) in FibroGene Cohort	p-value (beta) in Twins UK cohort	p-value (beta) in Finn Gen
rs34691025	5	73180281	ARHGEF28	A	G	0.798	0.004	1.76x10 <sup>-8</sup>	0.34 (0.06)	0.90 (0.002)	0.79 (0.01)
rs11959702	5	73178796	ARHGEF28	G	T	0.798	0.004	1.77x10 <sup>-8</sup>	0.35 (0.06)	0.87 (0.002)	0.79 (0.01)
rs377632600	5	73180608	ARHGEF28	T	G	0.801	0.004	2.74x10 <sup>-8</sup>	0.34 (0.10)	0.91 (0.002)	0.78 (0.01)
rs729272	5	73185504	ARHGEF28	A	G	0.798	0.004	2.89x10 <sup>-8</sup>	0.55 (0.04)	0.87 (-0.002)	0.79 (0.01)
rs6453030	5	73187503	ARHGEF28	G	C	0.798	0.004	3.30x10 <sup>-8</sup>	0.57 (0.04)	0.82 (0.003)	0.85 (0.01)
rs117537726	18	10297679	AP005209.2	C	A	0.968	-0.008	6.05x10 <sup>-8</sup>	0.55 (-0.09)	0.53 (-0.017)	0.39 (-0.07)
rs17387024	1	50993178	FAF1	T	C	0.852	-0.004	7.84x10 <sup>-8</sup>	0.86 (-0.01)	0.85 (-0.003)	0.69 (0.02)
rs72831687	6	16092360	AL021407.2	G	A	0.992	-0.016	1.61x10 <sup>-7</sup>	0.51 (0.21)	0.87 (0.002)	0.72 (0.04)
rs761429057	3	129474948	TMCC1	AAAAAA AAAAAG	A	0.953	-0.008	3.01x10 <sup>-7</sup>	0.36 (-2.29)	0.015 (0.044)	0.47 (0.52)
rs538146040	15	93131238	AC091544.2	G	A	0.993	-0.018	3.58x10 <sup>-7</sup>	0.87 (0.07)	0.61 (0.006)	0.70 (-0.25)
rs206969	12	120835893	AC003982.1	C	T	0.774	-0.004	4.00x10 <sup>-7</sup>	0.80 (0.02)	0.91 (-0.001)	0.18 (-0.05)
rs113280333	3	69242581	FRMD4B	C	T	0.983	-0.011	4.53x10 <sup>-7</sup>	0.78 (0.06)	0.34 (0.0016)	0.26 (-0.07)
rs148472276	2	3215642	EIPR1	C	A	0.994	-0.018	5.31x10 <sup>-7</sup>	0.41 (-0.29)	0.56 (0.014)	0.58 (0.17)
rs4782029	16	17248016	XYLT1	A	C	0.087	0.005	7.04x10 <sup>-7</sup>	0.17 (-0.12)	0.46 (0.012)	0.23 (0.06)

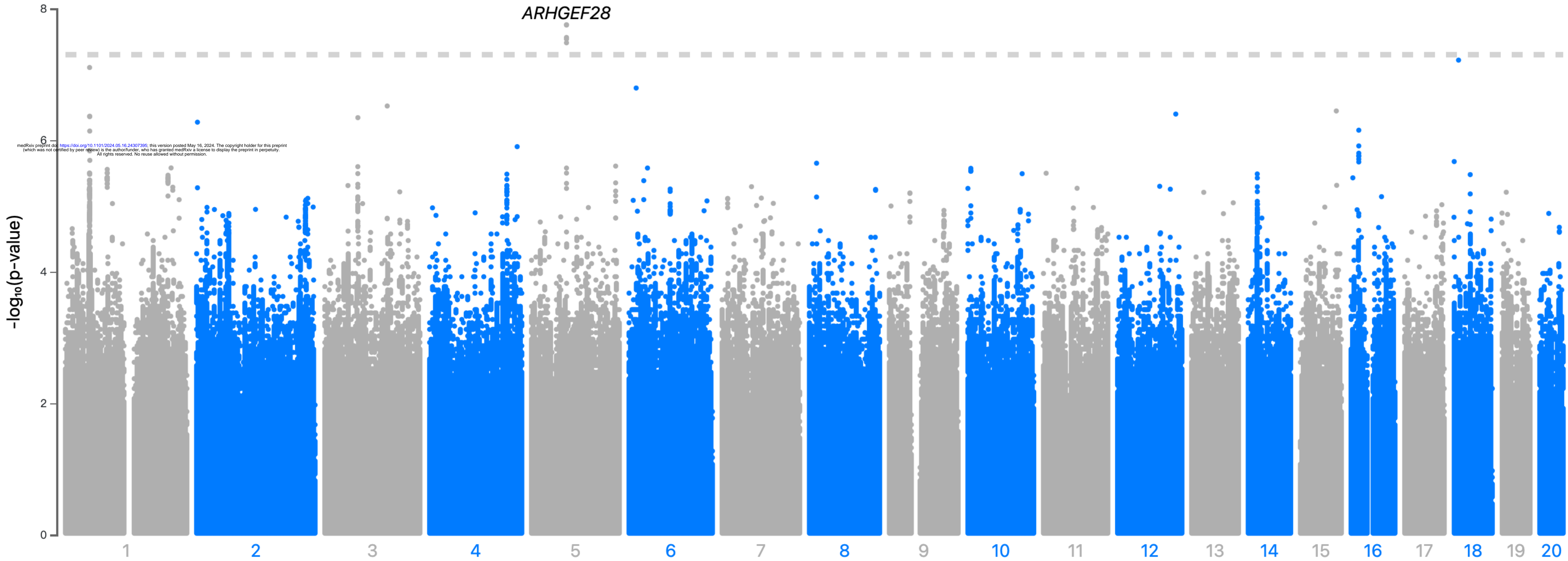
Chr: chromosome

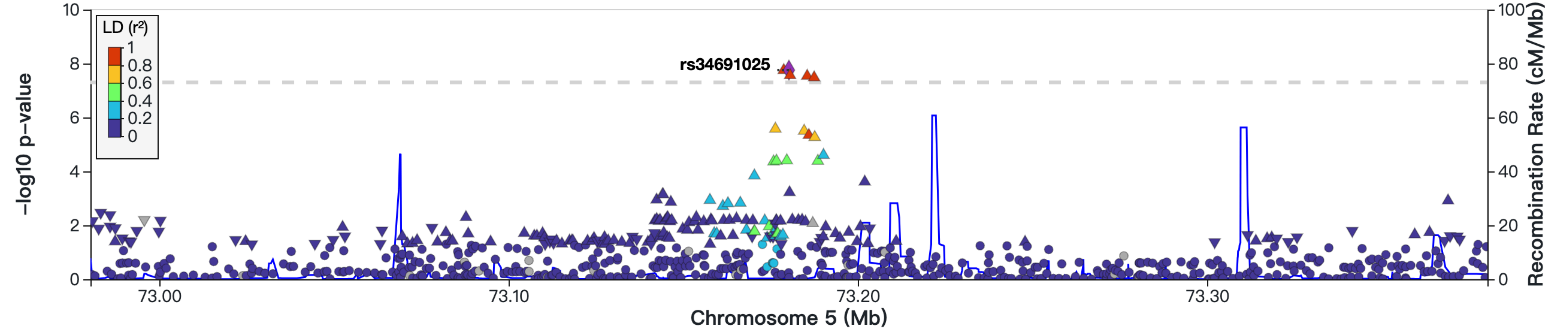
**Table 4** The locus and top SNP identified by the female-specific GWAS on widespread pain

Locus Rank	rsID	Chr	SNP position	Nearest Gene	Effect allele	Non effective allele	Frequency the effect allele	$p$ value	beta
1	rs137998089	10	78115898	<i>LRMDA</i>	A	G	0.994	$4.03 \times 10^{-8}$	-0.003

Chr: chromosome

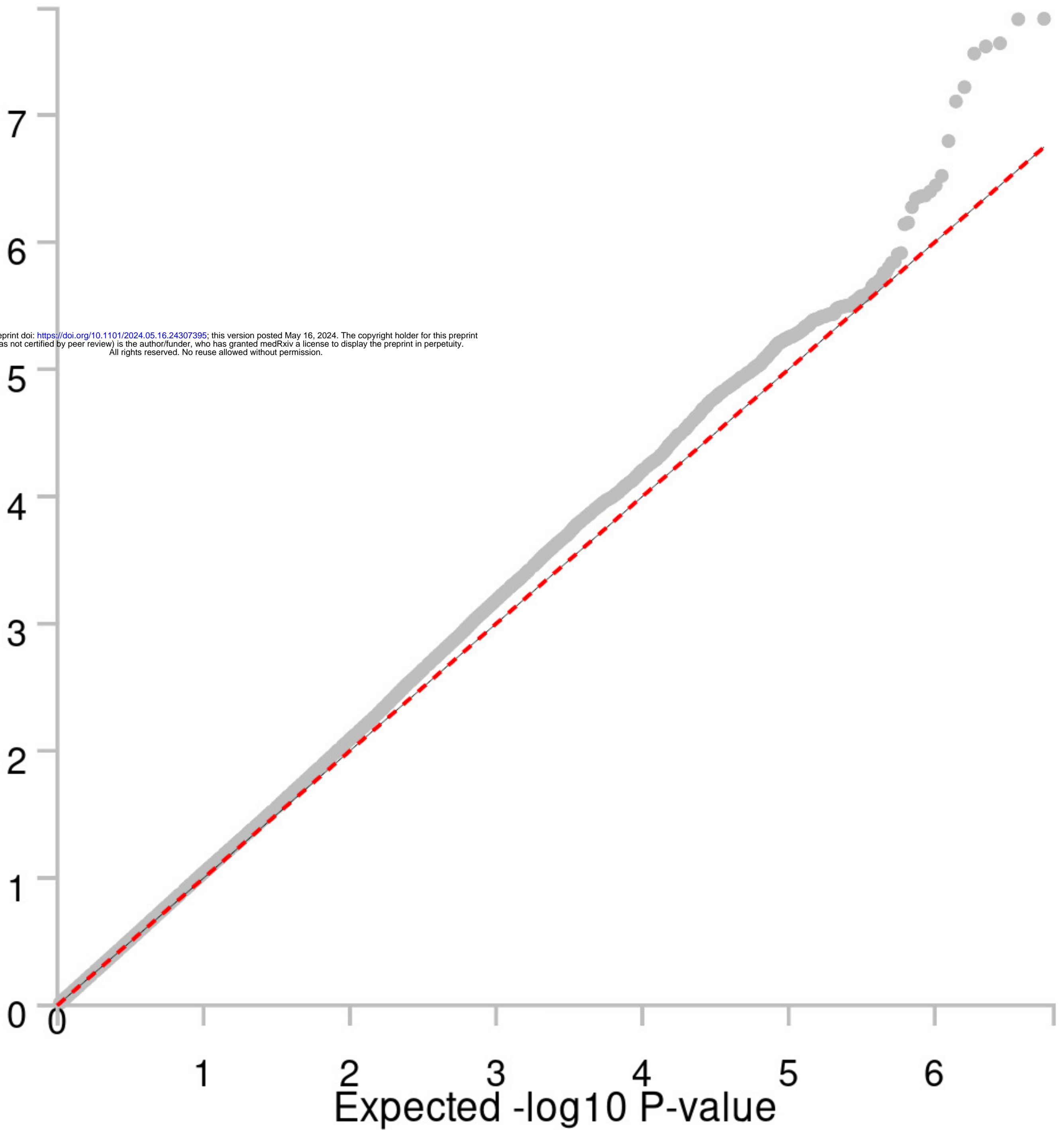
*ARHGEF28*



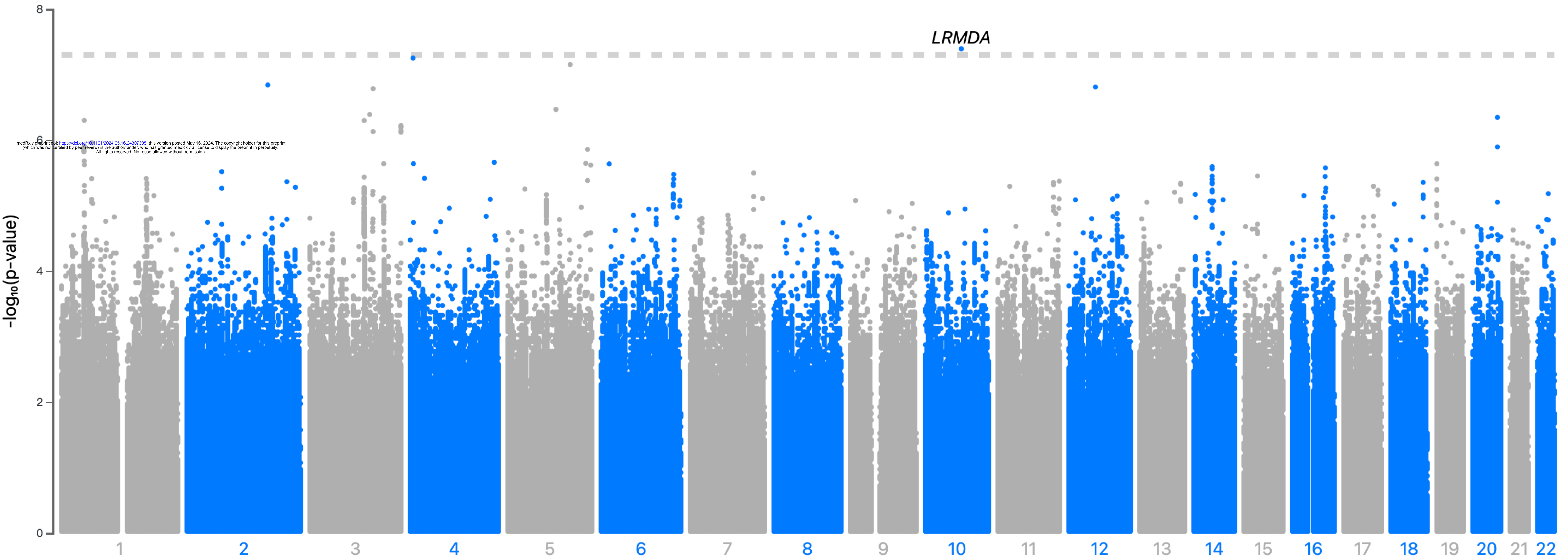


Observed  $-\log_{10}$  P-value

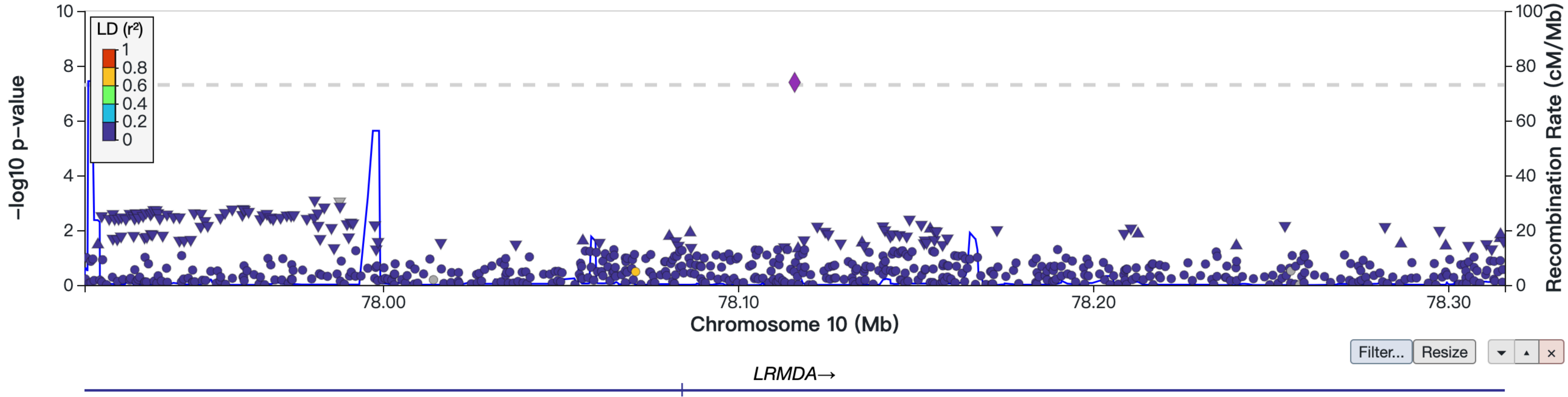
medRxiv preprint doi: <https://doi.org/10.1101/2024.05.16.24307395>; this version posted May 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.







medRxiv preprint doi: <https://doi.org/10.1101/2024.05.16.24307395>; this version posted May 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



Filter... Resize

LRMDA →

