5

Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases

Jie Zheng*[§1], Valeriia Haberland*[1], Denis Baird*[1], Venexia Walker*[1], Philip C. Haycock*[1], Mark R. Hurle[2], Alex Gutteridge[3], Pau Erola[1], Yi Liu[1], Shan Luo[1,4], Jamie Robinson[1], Tom G. Richardson[1], James R. Staley[1,5], Benjamin Elsworth[1], Stephen Burgess[5], Benjamin B. Sun[5], John Danesh[5,6,7,8,9,10], Heiko Runz[11], Joseph C. Maranville[12], Hannah M. Martin[13], James Yarmolinsky[1], Charles Laurin[1], Michael V. Holmes[1,14,15,16], Jimmy Z. Liu[11], Karol Estrada[11], Rita Santos[17], Linda McCarthy[3], Dawn Waterworth[2], Matthew R. Nelson[2], George Davey Smith*[1,18], Adam S. Butterworth*[5,6,7,8,9], Gibran Hemani*[1], Robert A. Scott*[§3], and Tom R. Gaunt*[§1,18]

[1]MRC Integrative Epidemiology Unit (IEU), Bristol Medical School, University of Bristol, Bristol, UK.
[2]Human Genetics, GlaxoSmithKline, Collegeville, PA, USA.
[3]Human Genetics, GlaxoSmithKline, Stevenage, Hertfordshire, UK.
[4]School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China.
[5]BHF Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
[6]BHF Centre of Research Excellence, School of Clinical Medicine, Addenbrooke's Hospital, Cambridge, UK.
[7]NIHR Blood and Transplant Research Unit in Donor Health and Genomics, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
[8]NIHR Cambridge Biomedical Research Centre, School of Clinical Medicine, Addenbrooke's Hospital, Cambridge, UK.
[9]Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Hinxton, UK.
[10]Department of Human Genetics, Wellcome Sanger Institute, Hinxton, UK.
[11]Translational Biology, Biogen, Cambridge, MA, USA.
[12]Informatics and Predictive Sciences, Celgene Corporation, Cambridge, MA, USA.
[13]School of Biological Sciences, University of Edinburgh, Edinburgh, UK.
[14]Medical Research Council Population Health Research Unit, University of Oxford, Oxford, UK.
[15]Clinical Trial Service Unit & Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK.
[16]National Institute for Health Research, Oxford Biomedical Research Centre, Oxford University Hospital, Oxford, UK.
[17]Functional Genomics, GlaxoSmithKline, Gunnels Wood Road, Stevenage, Hertfordshire, UK.
[18]NIHR Bristol Biomedical Research Centre, Bristol, UK.

35   *Proteome MR writing group
36   §e-mail: jie.zheng@bristol.ac.uk, robert.a.scott@gsk.com, tom.gaunt@bristol.ac.uk
37
38

39   **The human proteome is a major source of therapeutic targets. Recent genetic association**
40   **analyses of the plasma proteome enable systematic evaluation of the causal**
41   **consequences of variation in plasma protein levels. Here we estimated the effects of 1,002**
42   **proteins on 225 phenotypes using two-sample Mendelian randomization (MR) and**
43   **colocalization. Of 413 associations supported by evidence from MR, 130 (31.5%) were not**
44   **supported by results of colocalization analyses, suggesting that genetic confounding due**
45   **to linkage disequilibrium (LD) is widespread in naïve phenome-wide association studies of**
46   **proteins. Combining MR and colocalization evidence in cis-only analyses, we identified**
47   **111 putatively causal effects between 65 proteins and 52 disease-related phenotypes**
48   **(www.epigraphdb.org/pqtl/). Evaluation of data from historic drug development**
49   **programs showed that target-indication pairs with MR and colocalization support were**
50   **more likely to be approved, evidencing the value of this approach in identifying and**
51   **prioritizing potential therapeutic targets.**

52   Despite increasing investment in research and development (R&D) in the pharmaceutical
53   industry[1], the rate of success for novel drugs continues to fall[2]. Lower success rates make
54   new therapeutics more expensive, reducing availability of effective medicines and
55   increasing healthcare costs. Indeed, only one in ten targets taken into clinical trials reaches
56   approval[2], with many showing lack of efficacy (~50%) or adverse safety profiles (~25%) in
57   late stage clinical trials after many years of development[3,4]. For some diseases, such as
58   Alzheimer's disease, the failure rates are even higher[5].
59   Thus, early approaches to prioritize target-indication pairs that are more likely to be
60   successful are much needed. It has previously been shown that target-indication pairs for
61   which genetic associations link the target gene to related phenotypes are more likely to
62   reach approval[6]. Consequently, systematically evaluating the genetic evidence in support of
63   potential target-indication pairs is a potential strategy to prioritize development programs.
64   While systematic genetic studies have evaluated the putative causal role of both methylome
65   and transcriptome on diseases[7,8], studies of the direct relevance of the proteome are in
66   their infancy[9,10].
67   Plasma proteins play key roles in a range of biological processes and represent a
68   major source of druggable targets[11,12]. Recently published genome-wide association studies
69   (GWAS) of plasma proteins have identified 3,606 conditionally independent single
70   nucleotide polymorphisms (SNPs) associated with 2,656 proteins ('protein quantitative trait
71   loci', pQTL)[9,13,14,15,16]. These genetic associations offer the opportunity to systematically test
72   the causal effects of a large number of potential drug targets on the human disease
73   phenome through Mendelian randomization (MR)[17]. In essence, MR exploits the random
74   allocation of genetic variants at conception and their associations with disease risk factors
75   to uncover causal relationships between human phenotypes, and has been described in
76   detail previously[18,19].
77   For MR analyses of proteome, unlike more complex exposures, an intuitive way to
78   categorize protein-associated variants is into cis-acting pQTLs located in the vicinity of the
79   encoding gene (defined as ≤ 500 kb from the leading pQTL of the test protein in this study)
80   and trans-acting pQTLs located outside this window. The cis-acting pQTLs are considered to
81   have a higher biological prior and have been widely employed in relation to some phenome-
82   wide scans of drug targets such as CETP[20] and IL6R[21]. Trans-acting pQTLs may operate via
83   indirect mechanisms and are therefore more likely to be pleiotropic[22], although they may
84   support causal inference where they are likely to be non-pleiotropic.
85   Here we pool and cross-validate pQTLs from five recently published GWAS and use
86   them as instruments to systematically evaluate the causal role of 968 plasma proteins on
87   the human phenome, including 153 diseases and 72 risk factors available in the MR-Base
88   database[23]. Results of all analyses are available in an open online database
89   (www.epigraphdb.org/pqtl/), with a graphical interface to enable rapid and systematic
90   queries.
91
92

93   Results
94   Characterizing genetic instruments for proteins
95   **Figure 1** summarizes the genetic instrument selection and validation process. Briefly, we
96   curated 3,606 pQTLs associated with 2,656 proteins from five GWAS[9,13,14,15,16]. After
97   removing proteins and SNPs using criteria such as LD-pruning listed in **Online Methods**
98   (*Instrument selection*), we retained 2,113 pQTLs for 1,699 proteins as instruments for the
99   MR analysis (**Supplementary Table 1**). Among these instruments, we conducted further
100   validation by categorizing them into three tiers based on their likely utility for MR analysis
101   (**Online Methods**, *Instrument validation*): 1,064 instruments of 955 proteins with the
102   highest relative level of reliability (tier 1); 62 instruments that exhibited SNP effect
103   heterogeneity across studies (**Supplementary Figs. 1** and **2**), indicating uncertainty in the
104   reliability of one or all instruments for a given protein (tier 2; **Supplementary Tables 2** and
105   **3**); and 987 non-specific instruments that were associated with more than five proteins (tier
106   3). For the 263 tier 1 instruments associated with between two and five proteins, 68 of
107   them influenced multiple proteins in the sample biological pathway and thus are likely to
108   reflect vertical pleiotropy and remain valid instruments (**Supplementary Note**,
109   *Distinguishing vertical and horizontal pleiotropic instruments using biological pathway*
110   *data*)[22].
111   Among the 1,126 tier 1 and 2 instruments, 783 (69.5%) were cis-acting (within 500
112   kb of the leading pQTL) and 343 were trans-acting. Of 1,002 proteins with a valid instrument,
113   765 had only a single cis or trans instrument, 66 were influenced by both cis and trans SNPs
114   (**Supplementary Table 4**), and 153 had multiple conditionally distinct cis instruments (381
115   cis instruments shown in **Supplementary Table 5**).
116
117   Estimated effects of plasma proteins on human phenotypes
118   We undertook two-sample MR to systematically evaluate evidence for the causal effects of
119   1,002 plasma proteins (with tier 1 and tier 2 instruments) on 153 diseases and 72 disease-
120   related risk factors (**Supplementary Table 6** and **Online Methods**, *Phenotype selection*).
121   Overall, we observed 413 protein-trait associations with MR evidence ($P < 3.5 \times 10^{-7}$ at a
122   Bonferroni-corrected threshold) using either cis or trans instruments (or both for proteins
123   with multiple instruments).
124   Genetically filtering out predicted associations between proteins and phenotypes
125   may indicate four explanations: causality, reverse causality, confounding by LD between the
126   leading SNPs for proteins and phenotypes, or horizontal pleiotropy (**Supplementary Fig. 3**).
127   Given these alternative explanations, we conducted a set of sensitivity analyses to establish
128   whether the MR association reflects a causal effect of protein on phenotype: tests of
129   reverse causality using bi-directional MR[24] and MR Steiger filtering[25,26]; heterogeneity
130   analyses for proteins with multiple instruments[27], and colocalization analyses[28] to
131   investigate whether the genetic associations with both protein and phenotype shared the
132   same causal variant (**Fig. 1**). To avoid unreliable inference from colocalization analysis due to
133   the potential presence of multiple neighboring association signals, we also developed and
134   performed pairwise conditional and colocalization analysis (PWCoCo) of all conditionally
135   independent instruments against all conditionally independent association signals for the
136   outcome phenotypes (**Online Methods**, *Pairwise conditional and colocalization analysis*; **Fig.**
137   **2**). For this study, MR and colocalization were the two methods filtering reliable associations.
138   After the colocalization analysis, 283 of the 413 protein-phenotype associations had profiles
139   supportive of causality.

Estimating protein effects on human phenotypes using cis pQTLs

In the MR analyses using cis-pQTLs, we identified 111 putatively causal effects of 65 proteins on 52 phenotypes, with strong evidence of MR ($P < 3.5$ x 10^{-7}) and colocalization (posterior probability > 80%; after applying PWCoCo) between the protein- and phenotype-associated signals (**Fig. 3** and **Supplementary Table 7**). A further 69 potential associations had evidence from MR but did not have strong evidence of colocalization (posterior probability < 80%; **Supplementary Table 8**), highlighting the potential for confounding by LD and the importance of colocalization analyses in MR of proteins. Evidence of potentially causal effects supported by colocalization was identified across a range of disease categories, including anthropometric phenotypes and cardiovascular and autoimmune diseases (**Supplementary Note**, *Disease areas of protein-trait associations*), and our findings replicated some previous reported associations (**Supplementary Note**, *MR results replicated previous findings*).

Of 437 proteins with tier 1 or tier 2 cis instruments from Sun *et al.*[9] and Folkersen *et al.*[14], 153 (35%) had multiple conditionally independent SNPs in the cis region identified by GCTA-COJO[29] (**Supplementary Table 5**). We applied an MR model that takes into account the LD structure between conditionally independent SNPs in these cis regions[30]. In this analysis, we identified 10 additional associations that had not reached our Bonferroni corrected *P*-value threshold in the single-variant cis analysis. Generally, the MR estimates from the multi-cis MR analyses were consistent with the single-cis instrumented analyses (**Supplementary Table 9**).

In regions with multiple cis instruments, 16 of the 111 top cis MR associations only showed evidence of colocalization after conducting PWCoCo analysis for both the proteins and the human phenotypes, where none was observed between marginal results (**Supplementary Table 7**). For example, interleukin 23 receptor (IL23R) had two conditionally independent cis instruments: rs11581607 and rs3762318[9]. Conventional MR analysis combining both instruments showed a strong association of IL23R with Crohn's disease (OR = 3.22, 95% CI = 2.93 to 3.53, $P = 6.93$ x 10^{-131}; **Supplementary Table 9b**). There were four conditionally independent signals (conditional $P < 1$ x 10^{-7}) predicted for Crohn's disease in the same region (data from de Lange *et al.*[31]). In the marginal colocalization analyses, we observed no evidence of colocalization (**Fig. 4** and **Supplementary Fig. 4**, colocalization probability = 0). After performing PWCoCo with each distinct signal in an iterative fashion, we observed compelling evidence of colocalization between IL23R and one of the Crohn's disease signals for the top *IL23R* signal (rs11581607) (colocalization probability = 99.3%), but limited evidence for the second conditionally independent *IL23R* hit (rs7528804) (colocalization probability = 62.9%). Additionally, for haptoglobin, which showed MR evidence for LDL-cholesterol (LDL-C), there were two independent cis instruments. There was little evidence of colocalization between the two using marginal associations (colocalization probability = 0.0%). However, upon performing PWCoCo, we observed strong evidence of colocalization for both instruments (colocalization probabilities = 99%; **Supplementary Table 10** and **Supplementary Fig. 5**). Both examples demonstrate the complexity of the associations in regions with multiple independent signals and the importance of applying appropriate colocalization methods in these regions. Of the 413 associations with MR evidence (using cis and trans instruments), 283 (68.5%) also showed strong evidence of colocalization using either a traditional colocalization approach (260 associations) or after applying PWCoCo (23 associations), suggesting that one third of the

187 MR findings could be driven by genetic confounding by LD between pQTLs and other causal
188 SNPs.
189   Due to potential epitope-binding artefacts driven by protein-altering variants[32], we
190 also flag putatively causal links where the lead instrument is a protein-altering variant or is
191 in high LD ($r^2 > 0.8$) with one (**Supplementary Tables 7** and **8** filtered by column
192 "VEP_pQTL_Ldproxy" including missense, stop-lost/gained, start-lost/gained and splice-
193 altering variants).
194

Using trans-pQTLs as additional instrument sources
196 Trans pQTLs are more likely to influence targets though pleiotropic pathways. Among the
197 1,316 trans instruments we identified from five studies, 73.5% were associated with more
198 than five proteins, compared with 1.8 % of cis instruments (**Supplementary Table 1**).
199 However, in the context of MR, including non-pleiotropic trans-pQTLs may increase the
200 reliability of the protein-phenotype associations since (i) they will increase variance
201 explained of the tested protein and increase power of the MR analysis; (ii) the causal
202 estimate will not be reliant on a single locus, where multiple instruments exist; and (iii)
203 further sensitivity analyses, such as heterogeneity test of MR estimates across multiple
204 instruments, can be conducted. Therefore, we extended our MR analyses to include 343
205 non-pleiotropic trans instruments (**Supplementary Fig. 6**).
206   To utilize trans instruments, we first combined cis and trans instruments for 66
207 proteins that had both cis and trans instruments (noted as cis + trans analysis). However,
208 none reached our pre-defined Bonferroni-corrected threshold, and only two protein-
209 phenotype associations showed even suggestive evidence ($P < 1$ x 10^{-5}) (**Supplementary
210 Table 11**). Further, after including trans instruments, 17 of the cis-only signals were
211 attenuated. Secondly, we performed trans-only MR analyses of 293 proteins and identified
212 158 associations with 44 phenotypes that also had strong evidence (posterior probability >
213 0.8) of colocalization (**Supplementary Table 12**). A further 54 trans-only MR associations did
214 not have strong evidence of colocalization (**Supplementary Table 13**).
215   Some of the trans analyses with MR and colocalization evidence suggest causal
216 pathways that are confirmed by evidence from rare pathogenic variants or existing
217 therapies. For example, although we had no cis instrument for Protein C (Inactivator Of
218 Coagulation Factors Va And VIIIa) (PROC) (**Supplementary Fig. 7a**), we found evidence for a
219 causal association between PROC levels and deep venous thrombosis ($P = 1.27$ x 10^{-10};
220 colocalization probability > 0.9) using a trans pQTL, rs867186 (**Supplementary Fig. 7b**),
221 which is a missense variant in *PROCR*[33], the gene encoding the endothelial protein C
222 receptor (EPCR). Individuals with mutations in *PROC* have protein C deficiency, a condition
223 characterized by recurrent venous thrombosis for which replacement protein C is an
224 effective therapy.
225   From 47 proteins with multiple trans instruments, we identified four additional MR
226 associations, but none showed strong evidence of colocalization (**Supplementary Table 13**)
227 and little evidence of heterogeneity (**Supplementary Table 14**).
228

Estimating protein effects on human phenotypes using pQTLs with heterogeneous effects
230 *across studies*
231 Among the 2,113 selected instruments, we checked whether the 1,062 instruments with
232 association information in at least two studies showed consistent effect size across studies
233 (**Supplementary Table 15**). For these SNPs, we found that 62 showed evidence of difference

234 in effect size across studies (tier 2 instruments), for which we performed MR analyses using
235 the most significant SNP across studies and report the findings with caution. Some proteins
236 that are targets of approved drugs were found to have potential causal effects in this
237 analysis, such as interleukin-6 receptor (IL6R) on rheumatoid arthritis (RA)[34], and coronary
238 heart disease (CHD)[21] (**Supplementary Table 16**). Tocilizumab, a monoclonal antibody
239 against IL6R, is used to treat RA, while canakinumab, a monoclonal antibody against
240 interleukin-1 beta (an upstream inducer of interleukin-6), has been shown to reduce
241 cardiovascular events specifically among patients who showed reductions in interleukin-6[35].
242   As another test of heterogeneity across studies, where the same protein was
243 measured in two or more studies, we performed colocalization analysis of each pQTL (in one
244 study) against the same pQTL (in another study) for the two studies in which we had access
245 to full summary results (Sun *et al.*[9] and Folkersen *et al.*[14]). Of the 41 proteins measured in
246 both studies, 76 pQTLs could be tested using conventional colocalization and PWCoCo
247 (**Supplementary Table 15**). We found weak evidence of colocalization for 51 pQTLs
248 (posterior probability < 0.8), which suggested either two different signals were present
249 within the test region or the protein has a pQTL in one study but not in the other. In either
250 case, as one of the two distinct signals may be genuine, we performed MR analysis of these
251 25 pQTLs using instruments from each study separately. Eight associations had MR evidence,
252 but only one showed colocalization evidence (IL27 levels on human height; **Supplementary**
253 **Table 17**).
254
255 Sensitivity analyses to evaluate reverse causality
256 For potential associations between proteins and phenotypes identified in the previous
257 analyses, we undertook two sensitivity analyses to highlight results due to reverse causation:
258 bi-directional MR[24] and Steiger filtering[25] (**Online Methods**, *Distinguishing causal effects*
259 *from reverse causality*). In general, we found little evidence of reverse causality for genetic
260 predisposition to diseases on protein level changes (more details in **Supplementary Note**,
261 *Bi-directional MR and Steiger filtering results*; **Supplementary Data 1**).
262
263 Drug target prioritization and repositioning using phenome-wide MR
264 Given that human proteins represent the major source of therapeutic targets, we sought to
265 mine our results for targets of molecules already approved as treatments or in ongoing
266 clinical development. We first compared MR findings for 1,002 proteins against 225
267 phenotypes with historic data on progression of target-indication pairs in Citeline's
268 PharmaProjects (downloaded on 9th May 2018). Of 783 target-indication pairs with an
269 instrument for the protein and association results for a phenotype similar to the indication
270 for which the drug had been trialled, 9.2% (73 pairs) had successful (approved) drugs, 69.1%
271 had failed drugs (including 195 failed drugs in the clinical stage and 354 drugs that failed in
272 the preclinical stage) and 20.3% were for drugs still in development (161 pairs). The 268
273 pairs for successful (73) or failed (195) drugs were included in further analyses
274 (**Supplementary Table 18**). We observed eight target-indication pairs of successful drugs
275 with MR and colocalization evidence of a potentially causal relationship between protein
276 and disease (**Supplementary Table 19**). After removing duplicate genetic evidence for
277 related indications for the same therapy (**Online Methods**, *Drug target validation and*
278 *repositioning*), six successful drugs remained from 214 pairs (**Supplementary Table 20**). In
279 addition to the PROC and IL6R examples discussed earlier, we found Proprotein convertase
280 subtilisin/kexin type 9 (PCSK9) (target for evolocumab) for hypercholesterolemia and

281 hyperlipidaemia, Angiotensinogen (AGT) for hypertension, IL12B for psoriatic arthritis and
282 psoriasis, and TNF Receptor Superfamily Member 11a (TNFRSF11A) for osteoporosis. For
283 each of these examples, the direction of effect between circulating protein and disease risk
284 was consistent with the therapeutic mechanism, except IL6R and PROC at first sight.
285 However, for IL6R and PROC, the alleles associated with higher soluble protein levels have
286 been shown to also lead to lower intracellular pathway activation[36,37], indicating consistency
287 of direction with the therapeutic approach. These examples highlight the importance of
288 careful examination of the biological mechanisms underlying plasma pQTLs to enable
289 translation. Further removing associations potentially driven by protein-altering variants, as
290 well as drugs that were in large part motivated by genetic evidence (e.g. PCSK9 fits both
291 exclusion criteria), comparisons of the remaining 191 pairs indicated that protein-phenotype
292 associations with MR and colocalization evidence remained more likely to become
293 successful target-indication pairs (**Table 1**). Although we acknowledge the limited sample
294 size of the test set, this raises enthusiasm for the utility of pQTL MR analyses with
295 colocalization as a method for target prioritization.
296   Previous efforts have highlighted the opportunities and challenges of using genetics
297 for drug repositioning[38]. We identified three approved drugs for which we found pQTL MR
298 and colocalization evidence for five phenotypes other than the primary indication and 23
299 drug targets under development for 33 alternative phenotypes (**Supplementary Table 21**).
300 An example of urokinase-type plasminogen activator (PLAU) levels associated with lower
301 inflammatory bowel disease (IBD) risk is presented in the **Supplementary Note** (*Case study*
302 *for drug repurposing*) and **Supplementary Figure 8**.
303   We also evaluated drugs in current clinical trials and identified eight additional
304 protein-phenotype associations with MR and colocalization evidence (**Supplementary Table**
305 **22**), for which we observe MR evidence implicating an increased likelihood of success.
306   Finally, we compared the 1,002 instrumentable proteins (i.e. those that passed our
307 instrument selection procedure) against the druggable genome[39], and found that 682 of the
308 1,002 (68.1%) instrumentable proteins overlapped with the druggable genome
309 (**Supplementary Table 23** and **Online Methods**, *Enrichment of proteome-wide MR with the*
310 *druggable genome*). We conducted a further enrichment analysis to assess the overlap
311 between putative causal protein-phenotype associations and the druggable genome
312 (**Supplementary Table 24**). Of the 295 top findings (120 proteins on 70 phenotypes) with
313 both MR and colocalization evidence, 250 of them (87.7%) overlapped with the druggable
314 genome (**Fig. 5**). This enrichment analysis will become more valuable with the continuous
315 evolution of the druggable genome[38].
316

Discussion
318 MR analysis of molecular phenotypes against disease phenotypes provides a promising
319 opportunity to validate and prioritize novel or existing drug targets through prediction of
320 efficacy and potential on-target beneficial or adverse effects[40]. Our phenome-wide MR
321 study of the plasma proteome employed five pQTL studies to robustly identify and validate
322 genetic instruments for thousands of proteins. We used these instruments to evaluate the
323 potential effects of modifying protein levels on hundreds of complex phenotypes available
324 in MR-Base[23] in a hypothesis-free approach[17]. We confirmed that protein-phenotype
325 associations with both MR and colocalization evidence predicted a higher likelihood of a
326 particular target-indication pair being successful and highlight 283 potentially causal
327 associations. Collectively, we underline the important role of pQTL MR analyses as an
328 evidence source to support drug discovery and development and highlight a number of key
329 analytical approaches to support such inference.
330   In particular, we note the distinct opportunities and methodological requirements
331 for MR of molecular phenotypes, such as transcriptomics and proteomics, compared to
332 other complex exposures. For example, the number of instruments is often limited for
333 proteins, restricting the opportunity to apply recently developed pleiotropy robust
334 approaches[27,41]. New methods such as MR-robust adjusted profile scoring (MR-RAPS)[42]
335 allow inclusion of many weak instruments in the MR analysis and have been applied to a
336 recent proteome-wide MR study[10]. However, we note some examples where inclusion of
337 multiple weaker instruments can reduce power and yield different results to those based on
338 cis instruments alone[40,43], and we note very limited additional gain from inclusion of trans
339 instruments. A major advantage of proximal molecular exposures is the ability to include cis
340 instruments (or interpretable trans instruments) with high biological plausibility, limiting the
341 likelihood of horizontal pleiotropy[22,44]. Further, we note the limited gain from inclusion of
342 trans instruments in our analysis. However, undue focus on single SNP MR approaches
343 brings susceptibility to other pitfalls, such as the inability to examine heterogeneity of effect
344 and to evaluate and remove potential epitope artefacts.
345   To provide robust MR estimates for proteins, we note the important role of a
346 number of sensitivity analyses following the initial MR in order to distinguish causal effects
347 of proteins from those driven by horizontal pleiotropy, genetic confounding through LD[45]
348 and/or reverse causation[25]. Of note, only two-thirds of our putative causal associations had
349 strong evidence of colocalization, suggesting that a substantial proportion of the initial
350 findings were likely to be driven by genetic confounding through LD between pQTLs and
351 other disease-causal SNPs. To avoid misleading results, we suggest that for regions with
352 multiple molecular trait QTLs, it is important to consider methods such as PWCoCo, which
353 can avoid the assumptions of traditional colocalization approaches of just a single
354 association signal per region[46]. In the current study, application of PWCoCo identified
355 evidence of colocalization for 23 additional protein-phenotype associations hidden to
356 marginal colocalization[46]. We note that recent recommendations support the use of
357 colocalization as a follow up analysis to reduce false positives[47].
358   An important limitation of this work is that protein levels are known to differ
359 between cell types[48]. In this study, we have estimated the role of protein measured in
360 plasma on a range of complex human phenotypes but are unable to assess the relevance of
361 protein levels in other tissues. While eQTL studies highlight a large proportion of eQTLs
362 being shared across tissues[37], there are many which show cell type and state specificity[49],
363 highlighting the potential value of applying the current approach to data from proteomics

364 analyses in other cell types and tissues. We also hypothesize that, in instances with multiple
365 conditionally distinct pQTLs but where we observe colocalization of only certain
366 conditionally distinct pQTL-phenotype pairs, this may reflect underlying cell- and state-
367 specific heterogeneity in bulk plasma pQTLs, among which only certain cell-types or states
368 are causal[50]. Although pQTL studies have not yet been performed as systematically across
369 tissues or states as eQTL studies, it remains encouraging that our analyses using plasma
370 proteins identify associations across a range of disease categories, including for psychiatric
371 diseases for which we may expect key proteins to function primarily in the brain.
372   Evaluating the potential of MR to inform drug target prioritization, we demonstrated
373 that the presence of pQTL MR and colocalization evidence for a target-indication pair
374 predicts a higher likelihood of approval. One of the limitations of our approach is the lack of
375 comprehensive coverage of genetic data for all phenotypes for which drugs are in
376 development, as well as our inability to instrument the entire proteome through pQTLs. As
377 such, ongoing expansions in the scale, diversity and availability of GWAS will be important in
378 providing more precise estimates of the value of MR and colocalization in drug target
379 prioritization and in enabling its broader application.
380   Another potential limitation of our work is the presence of epitope-binding artefacts
381 driven by coding variants that may yield artefactual cis-pQTLs[32]. In particular, such instances
382 may lead to false negative conclusions where, in the presence of a silent missense variant
383 causing an artefactual pQTL but with no actual effect on protein function or levels, we do
384 not correctly instrument the target protein. In instances where the missense variant appears
385 to be driving the association with the phenotype, we suggest that causal inference may
386 remain valid but inference on direction of association is challenged. Finally, the limited
387 coverage of the proteome afforded by current technologies leaves the possibility of
388 undetected pleiotropy of instruments. While cis-pQTLs are less likely to be prone to
389 horizontal pleiotropy than trans-pQTLs, it is well known from studies of gene expression that
390 cis variants can influence levels of multiple neighboring genes and hence the same is likely
391 to be true for proteins. Future larger GWAS of the plasma proteome are likely to uncover
392 many more variant-protein associations, increasing the apparent pleiotropy of many pQTLs.
393   In conclusion, this study identified 283 putatively causal effects between the plasma
394 proteome and the human phenome using the principles of MR and colocalization. These
395 observations support, but do not prove, causality, as potential horizontal pleiotropy remains
396 an alternative explanation. Our study provides both an analytical framework and an open
397 resource to prioritize potential new targets and a valuable resource for evaluation of both
398 efficacy and repurposing opportunities by phenome-wide evaluation of on-target
399 associations.
400
401

482 Author contributions

Competing Interests Statement

513 References

514 1. Plenge, R. M., Scolnick, E. M. & Altshuler, D. Validating therapeutic targets through

515   human genetics. *Nat. Rev. Drug Discov.* **12**, 581–594 (2013).

516 2. Hay, M., Thomas, D. W., Craighead, J. L., Economides, C. & Rosenthal, J. Clinical

517   development success rates for investigational drugs. *Nat. Biotechnol.* **32**, 40–51 (2014).

518 3. Arrowsmith, J. & Miller, P. Phase II and Phase III attrition rates 2011–2012. *Nat. Rev.*

519   *Drug Discov.* **12**, 569 (2013).

520 4. Harrison, R. K. Phase II and phase III failures: 2013–2015. *Nat. Rev. Drug Discov.* **15**,

521   817 (2016).

522 5. Cummings, J. L., Morstorf, T. & Zhong, K. Alzheimer's disease drug-development

523   pipeline: few candidates, frequent failures. *Alzheimers Res. Ther.* **6**, 37 (2014).

524 6. Nelson, M. R. *et al.* The support of human genetic evidence for approved drug

525   indications. *Nat. Genet.* **47**, 856–860 (2015).

526 7. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts

527   complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).

528 8. Richardson, T. G. *et al.* Systematic Mendelian randomization framework elucidates

529   hundreds of CpG sites which may mediate the influence of genetic variants on disease.

530   *Hum. Mol. Genet.* **27**, 3293–3304 (2018).

531 9. Sun, B. B. *et al.* Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79 (2018).

532 10. Chong, M. *et al.* Novel drug targets for ischemic stroke identified through Mendelian

533   randomization analysis of the blood proteome. *Circulation* **140**, 819–830 (2019).

534 11. Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.*

535   **16**, 19–34 (2017).

536 12. Imming, P., Sinning, C. & Meyer, A. Drugs, their targets and the nature and number of

537   drug targets. *Nat. Rev. Drug Discov.* **5**, 821–834 (2006).

538   13. Suhre, K. *et al.* Connecting genetic risk to disease end points through the human blood

539   plasma proteome. *Nat. Commun.* **8**, 14357 (2017).

540   14. Folkersen, L. *et al.* Mapping of 79 loci for 83 plasma protein biomarkers in

541   cardiovascular disease. *PLoS Genet.* **13**, e1006706 (2017).

542   15. Yao, C. *et al.* Genome-wide association study of plasma proteins identifies putatively

543   causal genes, proteins, and pathways for cardiovascular disease. *Nat. Commun.* **9**, 3268

544   (2018).

545   16. Emilsson, V. *et al.* Co-regulatory networks of human serum proteins link genetics to

546   disease. *Science* **361**, 769–773 (2018).

547   17. Evans, D. M. & Davey Smith, G. Mendelian randomization: new applications in the

548   coming age of hypothesis-free causality. *Annu. Rev. Genomics Hum. Genet.* **16**, 327–350

549   (2015).

550   18. Davey Smith, G. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology

551   contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **32**,

552   1–22 (2003).

553   19. Zheng, J. *et al.* Recent developments in Mendelian randomization studies. *Curr.*

554   *Epidemiol. Rep.* **4**, 330–345 (2017).

555   20. Millwood, I. Y. *et al.* Association of *CETP* gene variants with risk for vascular and

556   nonvascular diseases among Chinese adults. *JAMA Cardiol.* **3**, 34–43 (2018).

557   21. Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6R MR) Consortium *et al.*

558   The interleukin-6 receptor as a target for prevention of coronary heart disease: a

559   mendelian randomisation analysis. *Lancet* **379**, 1214–1224 (2012).

560   22. Swerdlow, D. I. *et al.* Selecting instruments for Mendelian randomization in the wake of

561   genome-wide association studies. *Int. J. Epidemiol.* **45**, 1600–1616 (2016).

562   23. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the

563   human phenome. *Elife* **7**, e34408 (2018).

564   24. Timpson, N. J. *et al.* C-reactive protein levels and body mass index: elucidating direction

565   of causation through reciprocal Mendelian randomization. *Int. J. Obes.* **35**, 300–308

566   (2011).

567   25. Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationship between

568   imprecisely measured traits using GWAS summary data. *PLoS Genet.* **13**, e1007081

569   (2017).

570   26. Hemani, G. *et al.* Automating Mendelian randomization through machine learning to

571   construct a putative causal map of the human phenome. bioRxiv (2017)

572   doi:10.1101/173682.

573   27. Bowden, J. *et al.* A framework for the investigation of pleiotropy in two-sample summary

574   data Mendelian randomization. *Stat. Med.* **36**, 1783–1802 (2017).

575   28. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic

576   association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).

577   29. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics

578   identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).

579   30. Burgess, S., Dudbridge, F. & Thompson, S. G. Combining information on multiple

580   instrumental variables in Mendelian randomization: comparison of allele score and

581   summarized data methods. *Stat. Med.* **35**, 1880–1906 (2016).

582   31. de Lange, K. M. *et al.* Genome-wide association study implicates immune activation of

583   multiple integrin genes in inflammatory bowel disease. *Nat. Genet.* **49**, 256–261 (2017).

584   32. Solomon, T. *et al.* Identification of common and rare genetic variation associated with

585   plasma protein levels using whole-exome sequencing and mass spectrometry. *Circ.*

586   *Genom. Precis. Med.* **11**, e002170 (2018).

587   33. Taylor, F. B., Jr, Peer, G. T., Lockhart, M. S., Ferrell, G. & Esmon, C. T. Endothelial cell

588   protein C receptor plays an important role in protein C activation in vivo. *Blood* **97**,

589   1685–1688 (2001).

590   34. Hashizume, M. *et al.* Tocilizumab, a humanized anti-IL-6R antibody, as an emerging

591   therapeutic option for rheumatoid arthritis: molecular and cellular mechanistic insights.

592   *Int. Rev. Immunol.* **34**, 265–279 (2015).

593   35. Ridker, P. M. *et al.* Modulation of the interleukin-6 signalling pathway and incidence

594   rates of atherosclerotic events and all-cause mortality: analyses from the Canakinumab

595   Anti-Inflammatory Thrombosis Outcomes Study (CANTOS). *Eur. Heart J.* **39**, 3499–

596   3507 (2018).

597   36. Ferreira, R. C. *et al.* Functional IL6R 358Ala allele impairs classical IL-6 receptor

598   signaling and influences risk of diverse inflammatory diseases. *PLoS Genet.* **9**, e1003444

599   (2013).

600   37. Stacey, D. *et al.* Elucidating mechanisms of genetic cross-disease associations: an

601   integrative approach implicates protein C as a causal pathway in arterial and venous

602   diseases. *medRxiv* (2020) doi:10.1101/2020.03.16.20036822.

603   38. Sanseau, P. *et al.* Use of genome-wide association studies for drug repositioning. *Nat.*

604   *Biotechnol.* **30**, 317–320 (2012).

605   39. Finan, C. *et al.* The druggable genome and support for target identification and validation

606   in drug development. *Sci. Transl. Med.* **9**, eaag1166 (2017).

607   40. Holmes, M. V., Ala-Korpela, M. & Smith, G. D. Mendelian randomization in

608   cardiometabolic disease: challenges in evaluating causality. *Nat. Rev. Cardiol.* **14**, 577–

609   590 (2017).

610   41. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid

611   instruments: effect estimation and bias detection through Egger regression. *Int. J.*

612   *Epidemiol.* **44**, 512–525 (2015).

613   42. Zhao, Q., Wang, J., Hemani, G., Bowden, J. & Small, D. S. Statistical inference in two-

614   sample summary-data Mendelian randomization using robust adjusted profile score.

615   aRxiv (2018).

616   43. Evans, D. M. *et al.* Mining the human phenome using allelic scores that index biological

617   intermediates. *PLoS Genet.* **9**, e1003919 (2013).

618   44. Timpson, N. J. One size fits all: are there standard rules for the use of genetic instruments

619   in Mendelian randomization? *Int. J. Epidemiol.* **45**, 1617–1618 (2016).

620   45. Hemani, G., Bowden, J. & Davey Smith, G. Evaluating the potential role of pleiotropy in

621   Mendelian randomization studies. *Hum. Mol. Genet.* **27**, R195–R208 (2018).

622   46. Wu, Y. *et al.* Colocalization of GWAS and eQTL signals at loci with multiple signals

623   identifies additional candidate genes for body fat distribution. *Hum. Mol. Genet.* **28**,

624   4161–4172 (2019).

625   47. Wainberg, M. *et al.* Opportunities and challenges for transcriptome-wide association

626   studies. *Nat. Genet.* **51**, 592–599 (2019).

627   48. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).

628   49. GTEx Consortium *et al.* Genetic effects on gene expression across human tissues. *Nature*

629   **550**, 204–213 (2017).

630   50. Kim-Hellmuth, S. *et al.* Genetic regulatory effects modified by immune activation

631   contribute to autoimmune disease associations. *Nat. Commun.* **8**, 266 (2017).

632
633
634

Figure 1 | Study design of this phenome-wide MR study of the plasma proteome. The study included instrument selection and validation, outcome selection, four types of MR analyses, colocalization, sensitivity analyses, and drug target validation.

Figure 2 | A demonstration of pairwise conditional and colocalization (PWCoCo) analysis. Assume there are two conditional independent association pQTL signals (SNP 1 and SNP 2) and two conditional independent outcome signals (SNP 1 and SNP 3) in the tested region. A naïve colocalization analysis using marginal association statistics will return weak evidence of colocalization (showed in regional plots A and D). By conducting the analyses conditioning on SNP 2 (plot B) and 1 (plot C) for the pQTLs and conditioning on SNP 1 (plot E) and 3 (plot F) for the outcome phenotype, each of the nine pairwise combinations of pQTL and outcome association statistics (represented as lines with different colors in the middle of this figure) will be tested using colocalization. In this case, the combination of plot B and plot E shows evidence of colocalization but the remaining eight do not.

Figure 3 | Miami plot for the cis-only analysis, with circles representing the MR results for proteins on human phenotypes. The labels refer to top MR findings with colocalization evidence, with each protein represented by one label. The color refers to top MR findings with $P < 3.09 \times 10^{-7}$, where red refers to immune-mediated phenotypes, blue refers to cardiovascular phenotypes, green refers to lung-related phenotypes, purple refers to bone phenotypes, orange refers to cancers, yellow refers to glycemic phenotypes, brown refers to psychiatric phenotypes, pink refers to other phenotypes and grey refers to phenotypes that showed less evidence of colocalization. The x-axis is the chromosome and position of each MR finding in the cis region. The y-axis is the $-\log_{10} P$ value of the MR findings, MR findings with positive effects (increased level of proteins associated with increasing the phenotype level) are represented by filled circles on the top of the Miami plot, while MR findings with negative effects (decreased level of proteins associated with increasing the phenotype level) are on the bottom of the Miami plot.

Figure 4 | Regional association plots of IL23R plasma protein level and Crohn's disease in the *IL23R* region. a,b, Regional plots of IL23R protein level and Crohn's disease without conditional analysis. Plot in **b** lists the sets of conditionally independent signals for Crohn's disease in this region: rs7517847, rs7528924, rs183020189, rs7528804 (a proxy for the second *IL23R* hit rs3762318, $r^2 = 0.42$ in the 1000 Genome Europeans) and rs11209026 (a proxy for the top *IL23R* hit rs11581607, $r^2 = 1$ in the 1000 Genome Europeans), conditional P value $< 1 \times 10^{-7}$. **c,** Regional plot of IL23R with the joint SNP effects conditioned on the second hit (rs3762318) for *IL23R*. **d,** Regional plot of Crohn's disease with the joint SNP effects adjusted for other independent signals except the top *IL23R* signal rs11581607. **e,** Regional plot of IL23R with the joint SNP effects conditioned on the top hit (rs11581607) for *IL23R*. **f,** Regional plot of Crohn's disease with the joint SNP effects adjusted for other independent signals except the second *IL23R* signal rs3762318. The heatmap of the colocalization evidence for IL23R association on Crohn's disease (CD) in the *IL23R* region is presented in **Supplementary Figure 4**.

681 **Figure 5 | Enrichment of phenome-wide MR of the plasma proteome with the druggable**
682 **genome.** In this figure, we only show proteins with convincing MR and colocalization
683 evidence with at least one of the 70 phenotypes. The *x*-axis shows the categories of 70
684 human phenotypes, where the phenotypes have been grouped into 8 categories: 8
685 autoimmune diseases (red), 3 bone phenotypes (purple), 8 cancers (orange), 12
686 cardiovascular phenotypes (blue), 4 glycemic phenotypes (yellow), 2 lung phenotypes
687 (green), 4 psychiatric phenotypes (brown), and 29 other phenotypes (pink). The *y*-axis
688 presents the tiers of the druggable genome (as defined by Finan et al.[39]) of 120 proteins
689 under analysis, where the proteins have been classified into 4 groups based on their
690 druggability: tier 1 contains 23 proteins that are efficacy targets of approved small
691 molecules and biotherapeutic drugs, tier 2 contains 11 proteins closely related to approved
692 drug targets or with associated drug-like compounds, tier 3 contains 58 secreted or
693 extracellular proteins or proteins distantly related to approved drug targets, and 28 proteins
694 have unknown druggable status (Unclassified). The cells with colors are protein-phenotype
695 associations with strong MR and colocalization evidence. Cells in green are associations
696 overlapping with the tier 1 druggable genome, while cells in yellow, red or purple were
697 associations with tier 2, tier 3 or unclassified. More detailed information is shown in
698 **Supplementary Table 24**.

699 **Table 1 | Enrichment analysis comparing target-indication pairs with or without MR and colocalization evidence**

700

		Mendelian randomization and colocalization evidence	
		YES	NO
Target-indication pair approved after clinical trials	YES	4	40
	NO	0	147

701
702 The protein-phenotype association pairs were grouped into four categories: (i) pairs with both MR/colocalization indications of causality and
703 drug trial success; (ii) pairs with MR and colocalization evidence but no drug trial evidence; (iii) pairs with no strong MR or colocalization
704 evidence but with drug trial evidence; and (iv) pairs with no strong MR, colocalization or drug trial evidence. The cut-off for MR evidence was P
705 $< 3.5 \times 10^{-7}$; the cut off for colocalization evidence was posterior probability $> 80\%$. The drug trial evidence was obtained from PharmaProjects
706 database. The MR and colocalization analysis results involved in this analysis including both tier 1 and tier 2 instruments in both cis and trans
707 region. More results comparing MR and trial evidence for cis-only and tier 1 instruments can be found in **Supplementary Table 20**.
708

709   ## Methods

710   ### Instrument selection

711   pQTLs from five GWAS[9,13-16] were used for the instrument selection (**Fig. 1**). We first
712   mapped SNPs to genome build GRCh37.p13 coordinates and then used the following criteria
713   to select instruments:

714   - We selected SNPs that were associated with any protein (using a *P*-value threshold ≤
715   5×10^{-8}) in at least one of the five studies, including both cis and trans pQTLs.
716   - Due to the complex LD structure of SNPs within the human Major Histocompatibility
717   Complex (MHC) region, we removed SNPs and proteins coded for by genes within
718   the MHC region (chr6: from 26 Mb to 34 Mb).
719   - We then conducted linkage disequilibrium (LD) clumping for the instruments with
720   the TwoSampleMR R package[23] to identify independent pQTLs for each protein. We
721   used $r^2 < 0.001$ as the threshold to exclude dependent pQTLs in the cis (or trans)
722   gene region.

723   After instrument selection, 2,113 instruments were kept for further instrument validation
724   (**Supplementary Table 1**). The instrument selection process, and the number of instruments
725   for proteins at each step in the process, is illustrated in **Figure 1**.

726   We incorporated conditionally distinct signals from protein association data through
727   systematic conditional analysis. Of the five studies, Sun *et al.*[9] reported conditionally distinct
728   results for both cis and trans pQTLs, which have been used in our study. Folkersen *et al.*[14]
729   have shared summary statistics, with which we performed approximate conditional analyses
730   ourselves using GCTA-COJO[29], with genotype data from mothers in the Avon Longitudinal
731   Study of Parents and Children (ALSPAC) as the LD reference panel[51,52] (a description of the
732   ALSPAC cohort can be found in **Supplementary Note**, *Description of ALSPAC study*).
733   Conditionally independent signals in the cis region for Sun *et al.* and Folkersen *et al.* are
734   reported in **Supplementary Table 5**.

735

736   ### Instrument validation

737   For the 2,113 instruments, we further classified them into three groups (noted as tier 1, tier
738   2 and tier 3 instruments) using two major instrument-filtering steps: a specificity test and a
739   consistency test. More details of instrument validation, including harmonization of proteins
740   and instruments and statistical tests for consistency can be found in the **Supplementary**
741   **Note** (*The protocol of the instrument validation*).

742

743   *Test estimating instrument specificity*

744   Absence of horizontal pleiotropy is one of the core assumptions for MR. This assumes that
745   the genetic variant should only be related to the outcome of interest through the
746   instrumented exposure. We noted that some SNPs were associated with more than one
747   protein. For example, *APOE* SNP rs7412 is associated with a set of proteins such as ADAM11,
748   APBB2, and APOB. We plotted a histogram of the number of proteins each instrument was
749   associated with (**Supplementary Fig. 6**) and considered instruments associated with more
750   than 5 proteins as highly pleiotropic and assigned them as tier 3 instruments (which were
751   excluded from all analyses). For instruments associated with fewer than (or equal to) five
752   proteins, we reported the number of proteins each of them (and their proxies with LD $r^2 >$
753   0.5) was associated with to indicate the level of potential pleiotropy.

754   To further distinguish vertical and horizontal pleiotropy for these instruments, we
755   used biological pathway information from Reactome (https://reactome.org/) and protein-
756   protein interaction information from STRING DB (https://string-db.org/) implemented in
757   EpiGraphDB (www.epigraphdb.org; **Supplementary Note**, *Distinguishing vertical and*
758   *horizontal pleiotropic instruments using biological pathway data*). After this analysis, 68
759   instruments associated with multiple proteins were mapped to the same pathway (or same
760   PPI) and were considered as valid instruments. Given there are other pathways and PPIs
761   that may be not included in Reactome and STRING, we kept tier 1 and 2 instruments
762   associated with 1 to 5 proteins for the main MR analysis, but we recorded the number of
763   proteins and number of pathways these instruments are associated with as an indication of
764   potential pleiotropy.
765
766   *Consistency test estimating instrument heterogeneity across studies*
767   Among the 2,113 pQTLs selected as instruments, we looked up available protein GWAS
768   results (Sun *et al.*[9], Suhre *et al.*[13] and Folkersen *et al.*[14] with full GWAS summary statistics;
769   Yao *et al.*[15] and Emilsson *et al.*[16] with pQTLs only) and found 1,062 pQTLs (or proxies with $r^2 >$
770   0.8) with association information in at least two studies (**Supplementary Table 15**). We then
771   tested the beta-beta correlation using the Pearson correlation function in R. The results of
772   the beta-beta correlations of SNP effects for each pair of studies and the number of SNPs
773   included in each correlation analysis can be found in **Supplementary Table 2**.
774   We further performed two consistency tests on the instruments that were present
775   across studies: (i) pairwise Z test; (ii) colocalization analysis of proteins across studies
776   (details of the analyses in **Supplementary Note**, *The protocol of the instrument validation*).
777   Instruments showing evidence of high heterogeneity across studies using either the pair-
778   wise Z test (pairwise Z > 5) or colocalization analysis (PP < 80%), were flagged as tier 2
779   instruments. Recognizing that lack of replication and effect heterogeneity does not preclude
780   at least one of these effects being genuine, we used these instruments separately for the
781   follow-up genetic analyses (**Supplementary Table 3**) and reported the findings with caution.
782   We designated instruments passing both pleiotropy and consistency tests as tier 1
783   instruments and used them as primary instruments for the MR analysis.
784
785   *Identifying cis and trans instruments*
786   We further split tier 1 instruments into two groups: (i) *cis-acting pQTLs* within a 500-kb
787   window from each side of the leading pQTL of the protein were used for the initial MR
788   analysis (defined as the cis-only analysis)[45]; (ii) *trans-acting pQTLs* outside the 500-kb
789   window of the leading pQTL were designated as trans instruments. While trans instruments
790   may be more prone to pleiotropy, their inclusion could increase statistical power as well as
791   the scope of downstream sensitivity analyses (e.g. tests for heterogeneity between
792   instruments). Therefore, for the proteins with cis instruments, we also looked for additional
793   trans instruments, and if these were available, we conducted further MR analyses using
794   both sets of instruments (defined as the "cis + trans" analysis).
795   For cis instruments, we looked up their predicted consequence via Variant Effect
796   Predictor[53] hosted by Ensembl. We identified coding variants (including missense, stop-
797   lost/gained, start-lost/gained and splice-altering variants) since epitope-binding artefacts
798   driven by coding variants may yield artefactual cis pQTLs[32]. We then conducted a sensitivity
799   MR analysis that excluded cis instruments that are in the coding region to further avoid the
800   potential issue of epitope-binding artefacts driven by coding variants.

801

Phenotype selection

803 We obtained effect estimates for the association of the pQTLs with complex human
804 phenotypes using GWAS summary statistics that were included in the MR-Base database
805 (http://www.mrbase.org). We selected GWAS with the greatest excepted statistical power
806 when multiple GWAS records for the same phenotype were available in MR-Base. Diseases
807 were defined as primary outcomes. Risk factors were defined as secondary outcomes. After
808 selection, 153 diseases and 72 risk factors (such as lipids and glucose phenotypes) were
809 included as outcomes for the MR analyses (**Supplementary Table 6**).

810

Causal inference and sensitivity analyses

812 The following sections describe the two-sample MR analyses using single or small numbers
813 of instruments on 153 diseases and 72 risk factors. To identify possible violations of
814 assumptions of MR and to distinguish between the aforementioned scenarios in
815 **Supplementary Figure 3**, we therefore conducted the following sensitivity analyses:
816 colocalization analysis[28], tests for heterogeneity between instrumental SNPs[27], bi-directional
817 MR[24], and Steiger filtering[25,26] (**Fig. 1**).

818

Estimating the causal effects of proteins on human phenotypes using MR

820 In the initial MR analysis, proteins were treated as the exposures and 225 complex human
821 phenotypes as the outcomes (**Fig. 1**, Estimate putative causal relationship). Due to high
822 correlation among some of the tested phenotypes (e.g. coronary heart disease (CHD) and
823 myocardial infarction), we used the PhenoSpD method[54,55] to provide a more appropriate
824 estimate of the number of independent tests. We selected a *P*-value threshold of 0.05,
825 corrected for the number of independent tests, as our threshold for prioritizing MR results
826 for follow up analyses (number of tests = 142,857; $P < 3.5 \times 10^{-7}$).

827

MR analysis using single locus instruments

829 First, the strongest cis pQTL variants for each protein were used as the instrumental variable
830 (described as 'single cis' analysis). The Wald ratio[56] method was used to obtain MR effect
831 estimates. In this analysis, the MR effect estimates were sensitive to the particular choice of
832 pQTLs, since only the most strongly associated SNPs within each genomic region were used
833 as instruments. Burgess *et al.* recently suggested that more precise causal estimates can be
834 obtained using multiple genetic variants from a single gene region, even if the variants are
835 correlated[30,57]. We used multiple conditional independent cis SNPs (**Supplementary Table 5**)
836 against all 225 phenotypes to further evaluate the MR findings from our initial MR analysis
837 (described as 'multiple cis' analysis). A generalized inverse variance weighted (IVW) model
838 considering the LD pattern between the multiple cis SNPs was used to estimate the MR
839 effects, where the pairwise LD (r^2) were obtained from the 1000 Genomes European
840 ancestry reference samples.

841

MR analysis using multi-locus instruments

843 Among the measured proteins reported in Sun *et al.*[9], 34% had both cis and trans pQTLs and
844 30% had only trans pQTLs. We also conducted MR on proteins with both cis and trans pQTLs
845 (noted as the cis + trans MR analysis) and proteins with only trans pQTLs (noted as trans-
846 only analysis). In the cis + trans MR analysis, we tested the protein-phenotype associations
847 of 66 proteins with both cis and trans instruments. The IVW method was used to obtain MR

848   effect estimates. In the trans-only MR analysis, we used 351 trans instruments for 298
849   proteins. The IVW method was used when two or more trans instruments were included in
850   the analysis, whereas the Wald ratio method was used when only one trans instrument was
851   included in the analysis.
852
853   ***MR analysis software***
854   The majority of MR analyses (including Wald ratio, IVW, bi-directional MR, MR Steiger
855   filtering and heterogeneity test across multiple instruments) were conducted using the MR-
856   Base TwoSampleMR R package (github.com/MRCIEU/TwoSampleMR)[23]. The IVW analysis
857   considering LD pattern was conducted using the MendelianRandomization R package[58]. The
858   MR results were plotted as forest plots and Miami plots using code derived from the ggplot2
859   package in R.
860
861   *Distinguishing causal effects from genomic confounding due to linkage disequilibrium*
862   Results that survived the multiple testing threshold in the MR analysis were evaluated using
863   a stringent Bayesian model (colocalization analysis) to estimate the posterior probability (PP)
864   of each genomic locus containing a single variant affecting both the protein and the
865   phenotype[28]. For protein and phenotype GWAS lacking sufficient SNP coverage or missing
866   key information (e.g. allele frequency or effect size), we conducted the "LD check" analysis
867   (more details of the two methods in **Supplementary Note**, *Linkage disequilibrium check*).
868
869   *Pairwise conditional and colocalization analysis*
870   The presence of multiple conditionally distinct association signals within the same genomic
871   region will influence the performance of colocalization analysis. We therefore developed an
872   analysis pipeline to integrate conditional and colocalization approaches for regions with
873   multiple conditionally independent pQTLs. Where there was convincing MR evidence below
874   the *P*-value threshold of 3.5×10^{-7}, but no good evidence of colocalization using the marginal
875   SNP effects of the exposures and outcomes (in total 148 MR associations in both cis and
876   trans regions), we performed pairwise colocalization analyses of all conditionally distinct
877   pQTLs against all identified conditionally distinct association signals in the outcome data
878   (noted as pair-wise conditional and colocalization analysis: PWCoCo). The conditional
879   analysis for proteins and human phenotypes was conducted using the GCTA-COJO package[29],
880   with genotype data from mothers in the Avon Longitudinal Study of Parents and Children
881   (ALSPAC) as the LD reference panel[51,52] (a description of the ALSPAC cohort can be found in
882   **Supplementary Note**, *Description of ALSPAC study*). **Figure 2** demonstrates the nine possible
883   pair-wise combinations of various conditional signals for proteins and phenotypes at which
884   there are two independent signals in the region (**Supplementary Table 27**).
885   For protein-phenotype associations that only showed colocalization evidence after
886   we applied PWCoCo, we recorded the PWCoCo model that showed colocalization evidence
887   in a new column "PWCoCo_model", in **Supplementary Tables 7**, **8**, **11**, **12**, **13**, **16** and **17**.
888
889   *Heterogeneity test and directionality test of MR findings*
890   For MR analyses using two or more instruments, we conducted heterogeneity tests to
891   estimate the variability in the causal estimates obtained for each SNP (i.e. how consistent is
892   the causal estimate across all SNPs used as separate instruments) (**Fig. 1**, Consistency of the
893   causal estimate across all SNPs). Cochran's Q test statistic was calculated for the IVW
894   analyses, which is expected to be chi-squared distributed with number of SNPs minus one

895 degrees of freedom[27]. Lower heterogeneity suggests a lower chance of violations of
896 assumptions in MR estimates, such as the presence of confounding through horizontal
897 pleiotropy[59].
898   In order to mitigate the potential impact of reverse causality (i.e. the hypothesised
899 outcome actually has a causal effect on the hypothesised exposure and not vice versa), we
900 used two approaches to identify directions of causality: bi-directional MR and Steiger
901 filtering (more details in **Supplementary Note**, *Directionality test*).
902
903 *Drug target validation and repositioning*
904 Approved drug targets have previously been shown to be enriched for gene-phenotype
905 associations[6]. We therefore wished to assess whether approved drug targets were enriched
906 for protein-phenotype associations, as obtained in the present study using MR. We assessed
907 the support for approved drug targets among our MR findings using Fisher's exact test.
908 Target-indication pairs for successful and failed drugs were identified using a manually
909 annotated version of PharmaProjects database from Citeline
910 (https://pharmaintelligence.informa.com/). The phenotypes used in the MR analyses and
911 the indications listed in Citeline's PharmaProjects (downloaded on 9th May 2018) were then
912 manually mapped to MeSH headings as a common ontology. This allowed us to match the
913 protein-phenotype associations with corresponding target-indication pairs. To improve this
914 matching, we implemented a similarity matrix, derived from all MeSH headings in the
915 manual mapping, and retained matches with a relative similarity greater than 0.7 for our
916 analyses (the similarity matrix has been previously described in Nelson *et al.*[6]). We then
917 compared whether the target-indication pair represented a successful or failed drug against
918 whether there was a signal or not for the corresponding protein-phenotype pair among our
919 MR findings. For the purposes of this test, a signal was defined as an MR result with $P < 3.5$ x
920 10^{-7} (which is the Bonferroni *P*-value threshold of the MR analysis) with supporting evidence
921 from colocalization analysis. We further conducted a set of sensitivity analyses based on the
922 following criteria to increase the reliability of the enrichment analysis:
923   1. We checked the direction of effect of MR findings and drug trial results for the eight
924   approved drugs using therapeutic direction information from PharmaProjects.
925   2. For target-indication pairs linked to similar phenotypes (for example, the same
926   target associated with angina and myocardial infarction), we removed one of them
927   to avoid double counting the same association.
928   3. To avoid the influence of epitope-binding artefacts, we removed MR results
929   estimated using missense variants as an instrument.
930   4. We checked whether approved drugs had been motivated by genetics from Drug
931   Bank (https://www.drugbank.ca/), which may have inflated the OR estimate.
932 In total, we removed 75 target-indication pairs based on criteria 2 (45 pairs), 3 (23 pairs) and
933 4 (2 pairs; some pairs appeared in multiple situations) and conducted the comparison
934 between protein-phenotype associations using MR and target-indication pairs from
935 PharmaProjects, both on each criterion separately and on all criteria together
936 (**Supplementary Table 20**).
937   Phenome-wide MR has demonstrated the potential to validate, repurpose and
938 predict on-target side effects of drug targets. Of the protein-phenotype associations that
939 showed evidence of colocalization identified in the cis-only, cis+trans, trans-only or MR
940 analyses using pQTLs with heterogeneous effects across studies (noted as tier 2
941 instruments), we first looked up how many proteins with MR evidence were established

942   drug targets in the Informa PharmaProjects database. We then looked up how many of the
943   associations were established target-indication pairs in the PharmaProjects database. More
944   importantly, we predicted the potential adverse effects and repositioning opportunities of
945   all marketed drugs and drugs under development using phenome-wide MR.
946
947   *Enrichment of proteome-wide MR with the druggable genome*
948   Previously, Finan *et al.*[39] systematically identified 4479 genes as the newest druggable
949   genome compendium. This study stratified the druggable genome set into three tiers. Tier 1
950   (1,427 genes) included efficacy targets of approved small molecules and biotherapeutic
951   drugs, as well as targets modulated by clinical-phase drug candidates; tier 2 was composed
952   of 682 genes encoding proteins closely related to drug targets, or with associated drug-like
953   compounds; and tier 3 contained 2,370 genes encoding secreted or extracellular proteins,
954   distantly related proteins to approved drug targets, and members of key druggable gene
955   families not already included in tier 1 or tier 2. We assessed whether the 1,002 proteins we
956   selected for the MR analyses overlapped with the 4,479 genes from the druggable genome
957   (**Supplementary Table 23**). The proteins were mapped based on the HGNC name of the
958   encoding genes. We further assessed the overlap based on whether the protein had cis or
959   trans instruments and based on the druggable genome tiers.
960   In addition to the above comparison between instrumentable and druggable
961   genome, we also assessed the enrichment of top pQTL MR findings with the druggable
962   genome. 295 protein-phenotype associations (120 proteins on 70 phenotypes) with both
963   MR and colocalization evidence were selected for this analysis. We stratified the 120
964   proteins into 4 groups based on their druggability: tier 1 contained 23 proteins, tier 2
965   contained 11 proteins, tier 3 contained 58 proteins, and 28 proteins remained unclassified.
966   The 70 phenotypes were stratified into 8 groups: 8 autoimmune diseases, 3 bone
967   phenotypes, 8 cancer phenotypes, 12 cardiovascular phenotypes, 4 glycemic phenotypes, 2
968   lung phenotypes, 4 psychiatric phenotypes and 29 other phenotypes. The protein-
969   phenotype associations with MR and colocalization evidence were colored separately based
970   on their druggability tiers. More details of this enrichment analysis are shown in **Figure 5**
971   and **Supplementary Table 24**.
972
973   Data availability
974   The data (GWAS summary statistics) used in the analyses described here are freely
975   accessible in the MR-Base platform (www.mrbase.org). All our analysis results for 989
976   proteins against 225 human phenotypes are freely available to browse, query and download
977   in EpiGraphDB (http://www.epigraphdb.org/pqtl/). An application programming interface
978   (API) and R package documented on the website enable users to programmatically access
979   data from the database.
980
981   Code availability
982   The code used in the Mendelian randomization and colocalization analyses described here
983   are freely accessible via our GitHub repo (https://github.com/MRCIEU/epigraphdb-pqtl).
984   The MR analysis was conducted using TwoSampleMR R package
985   (https://github.com/MRCIEU/TwoSampleMR). We implemented the colocalization analysis
986   using the coloc R package (created by Chris Wallace *et al.*), which can be downloaded here
987   (https://cran.r-project.org/web/packages/coloc/index.html).

988   Methods-only references
989

990   51. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'—the index offspring of the Avon

991   Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).

992   52. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children:

993   ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).

994   53. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).

995   54. Nyholt, D. R. A simple correction for multiple testing for single-nucleotide

996   polymorphisms in linkage disequilibrium with each other. *Am. J. Hum. Genet.* **74**, 765–

997   769 (2004).

998   55. Cichonska, A. *et al.* metaCCA: summary statistics-based multivariate meta-analysis of

999   genome-wide association studies using canonical correlation analysis. *Bioinformatics* **32**,

1000   1981–1989 (2016).

1001   56. Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N. & Davey Smith, G.

1002   Mendelian randomization: using genes as instruments for making causal inferences in

1003   epidemiology. *Stat. Med.* **27**, 1133–1163 (2008).

1004   57. Burgess, S., Zuber, V., Valdes-Marquez, E., Sun, B. B. & Hopewell, J. C. Mendelian

1005   randomization with fine-mapped genetic data: Choosing from large numbers of

1006   correlated instrumental variables. *Genet. Epidemiol.* **41**, 714–725 (2017).

1007   58. Yavorska, O. O. & Burgess, S. MendelianRandomization: an R package for performing

1008   Mendelian randomization analyses using summarized data. *Int. J. Epidemiol.* **46**, 1734–

1009   1739 (2017).

1010   59. Haycock, P. C. *et al.* Best (but oft-forgotten) practices: the design, analysis, and

1011   interpretation of Mendelian randomization studies. *Am. J. Clin. Nutr.* **103**, 965–978

1012   (2016).

1013