

1 **Directional anchor genes refine polygenic informed treatment selection in**
2 **schizophrenia and bipolar disorder**

3
4 William R. Reay^{1,2}, Michael P. Geaghan³, Joshua R. Atkins¹, Vaughan J. Carr^{4,5,6}, Melissa J.
5 Green^{4,5}, Murray J. Cairns^{1,2*}

6
7 ¹School of Biomedical Sciences and Pharmacy, The University of Newcastle, Callaghan,
8 NSW, Australia

9 ²Centre for Brain and Mental Health Research, Hunter Medical Research Institute,
10 Newcastle, NSW, Australia

11 ³Kinghorn Centre for Clinical Genomics, Garvan Medical Research Institute, Darlinghurst,
12 NSW, Australia

13 ⁴School of Psychiatry, University of New South Wales, Randwick, NSW, Australia

14 ⁵Neuroscience Research Australia, Sydney, NSW, Australia

15 ⁶Department of Psychiatry, Monash University, Melbourne, VIC, Australia

16
17 To whom correspondence should be addressed: Professor Murray Cairns

18 (Murray.Cairns@newcastle.edu.au)

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33 NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

34 **ABSTRACT**

35 Genetically informed drug development and repurposing is an attractive prospect for
36 improving outcomes in patients with psychiatric illness; however, the effectiveness of these
37 endeavours can be confounded by heterogeneity. In this study, we propose a novel approach
38 that links interventions implicated by disorder-associated genetic risk, at the population level,
39 to a framework that can target these compounds to individuals. Specifically, results from
40 genome-wide association studies are integrated with expression data to prioritise individual
41 “directional anchor” genes for which the predicted risk-increasing direction of expression
42 could be counteracted by an existing drug. While these compounds represent plausible
43 therapeutic candidates, they are not likely to be equally efficacious for all individuals. To
44 account for this heterogeneity, we constructed polygenic scores restricted to variants annotated
45 to the network of genes that interact with each directional anchor gene. These metrics, we call
46 pharmagenic enrichment scores (PES), identify individuals with a higher burden of genetic
47 risk, localised in biological processes related to the candidate drug target, to inform precision
48 drug repurposing. We used this approach to investigate schizophrenia and bipolar disorder and
49 reveal several compounds targeting specific directional anchor genes that could be plausibly
50 repurposed, including antioxidants, vitamins, antiarrhythmics, and lipid modifying agents.
51 These genetic risk scores, mapped to the networks associated with target genes, revealed
52 biological insights that cannot be observed in undifferentiated genome-wide polygenic risk
53 score (PRS). For example, an enrichment of these partitioned scores in schizophrenia cases
54 with otherwise low PRS and distinct correlations with measured biochemical traits. In
55 summary, genetic risk could be used more specifically to direct drug repurposing candidates
56 that target particular genes implicated in psychiatric and other complex disorders.

57

58

59

60

61

62

63

64

65

66

67

68 INTRODUCTION

69 Psychiatric disorders remain difficult to effectively manage in some patients, with treatment
70 resistance observed in a notable proportion of individuals prescribed conventional
71 pharmacotherapies [1–3]. Moreover, a key challenge in psychiatric practice is the selection of
72 a suitable course of treatment for newly diagnosed patients. Novel treatment opportunities for
73 these disorders would be of great clinical benefit, however, the drug development pipeline
74 remains arduous, expensive, and unproductive [4, 5]. Drug repurposing, whereby an approved
75 compound is redeployed for a new indication, is a promising avenue to more rapidly alter
76 psychiatric practice relative to the *de novo* drug development process [6, 7]. There has already
77 been utility in this approach demonstrated in psychiatry, such as atomoxetine that has been
78 repurposed for attention-hyperactive deficit disorder (ADHD) and the anti-convulsant
79 valproate for bipolar disorder [8].

80

81 We know that psychiatric illnesses arise from a multifaceted interplay between genetic and
82 environmental factors that contribute to its aetiologic complexity. In recent years, genome-
83 wide association studies (GWAS) have confirmed that psychiatric disorders are polygenic in
84 nature [9–13], with common frequency variants constituting a significant portion of trait
85 heritability. This means that individual loci that have small to modest impact, but contribute to
86 a much larger polygenic effect of many such variants throughout the genome [14, 15].
87 Biological insights from genetic studies could lead to repurposing opportunities in psychiatry
88 – for example, schizophrenia GWAS have previously suggested that the common variant signal
89 is enriched amongst the targets of antiepileptics, as well as in genes involved in retinoid
90 (vitamin A derivative) pathways [16, 17]. The polygenic nature of these disorders, however,
91 present a challenge drug targeting because the genetic architecture of each individual will be
92 highly heterogeneous. This means that any given patient will carry a unique combination of
93 risk and protective alleles, which likely translates to different underlying biological processes
94 being affected. As a result, genetically informed drug candidates may not be efficacious at a
95 population level. These phenomena necessitate the consideration of how pharmacotherapies
96 could be targeted more specifically to individuals based on their underlying genetic and clinical
97 risk factors.

98

99 Our group has previously sought to address these challenges through the development of the
100 *pharmagenic enrichment score* (PES), which is a framework that seeks to use polygenic risk
101 to direct precision drug repurposing opportunities [7, 18–20]. Specifically, the PES approach

102 derives partitioned polygenic scores from variants annotated to pathways or networks that are
103 targeted by approved drugs, with the underlying hypothesis that individuals with elevated
104 genetic risk (PES) amongst those genes may benefit from a compound which modulates that
105 pathway. Furthermore, in prior work we have established that PES profiles provide distinct
106 insights from a biologically undifferentiated genome-wide polygenic risk score (PRS) [18–20].
107 However, a limitation of the PES approach is that it is not innately informative as to which of
108 the suite of drugs targeting a pathway will be most useful, particularly in regards as to whether
109 an agonist or antagonist of target genes should be investigated. While we addressed this
110 previously, in respiratory medicine, by triangulating prioritised pathways through causal
111 inference of pharmacologically sensitive biochemical traits [19, 20], these relationships are
112 more difficult to find in psychiatric disorders [21]. In the current study, we propose a new
113 implementation of the PES that is informed by genetically proxied mRNA or protein
114 expression of drug target genes. This approach identifies candidate psychiatric drug
115 repurposing opportunities at the population level that can then be more appropriately integrated
116 with genetic risk scores relevant to these target genes, to identify individuals who may benefit
117 more readily from these compounds. These are termed *directional anchor genes* as they inform
118 on the clinically useful direction of modulation is for the biological networks containing these
119 genes which can be utilised to construct PES. In this study, we investigated this novel approach
120 in two highly heritable psychiatric disorders, schizophrenia and bipolar disorder, and identified
121 novel drug repurposing opportunities from candidate directional anchor genes and propose how
122 these genes could be used in concert with the PES to direct the candidate compounds for
123 repurposing.

124

125 **MATERIALS AND METHODS**

126

127 **Overview of the directional anchor gene *pharmagenic enrichment score* approach**

128 We summarise the fundamental principles of directional anchor genes and their integration
129 with the *pharmagenic enrichment score* in this section, followed by more expansive details in
130 the subsequent sections. Briefly, we define the concept of a directional anchor gene (DA-gene)
131 as a gene where, i) the direction of expression associated with increased odds of the disorder
132 can be predicted, and ii) this disorder associated direction of expression could be counteracted
133 by an approved compound, thus constituting a drug repurposing opportunity. For instance, if
134 upregulation of a hypothetical gene, gene *X*, was associated with greater odds of a disease
135 phenotype, then an antagonist of gene *X* may be clinically useful. If this gene *X* antagonist is

136 already approved for another indication, this may inform drug repurposing. However, there is
137 immense heterogeneity between individuals for any given complex trait or disease in its genetic
138 architecture, which often translates to highly variable clinical manifestation. We therefore
139 hypothesise that individuals with a greater burden of disorder-associated polygenic risk in the
140 directional anchor gene, and its network of genes that physically and biologically interact with
141 it, may benefit more specifically from a drug repurposing candidate targeting the DA-gene.
142 Polygenic risk mapped to biological networks encompassing the directional anchor genes is
143 likely to incorporate disorder-associated impacts on upstream processes that would modify the
144 effect of a compound targeting the candidate gene, as well as downstream processes triggered
145 by modulating the directional anchor. As discussed in the introduction, our group has
146 previously developed the *pharmagenic enrichment score* (PES) methodology to utilise
147 polygenic scoring to direct drug repurposing, whereby polygenic scores are constructed
148 specifically using variants mapped to biological pathways targeted by known drugs [18, 19].
149 A limitation of the PES approach is that it does not inherently predict the direction of effect
150 genes in the pathway that would need to be targeted such that repurposing a drug for individuals
151 with high polygenic risk in said pathway would be efficacious. Directional anchor genes,
152 therefore, help address this limitation when used in conjunction with PES constructed using
153 networks or biological pathways in which the gene participates. In other words, drug
154 repurposing opportunities are predicted at the population level based on the expression of target
155 genes, with these compounds potentially able to be more specifically directed to individuals
156 with elevated disorder associated genetic risk within pathways or networks that contain the
157 directional anchor gene, that is, an elevated PES (Figure 1). It should be noted that whilst we
158 apply this approach to binary disease phenotypes in this study, it can also be utilised for
159 clinically relevant continuous traits. In that case, candidate directional anchor genes would be
160 those genes for which the drug repurposing candidate is genetically inferred to modulate the
161 trait in a clinically useful fashion.

162

163

164

165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198

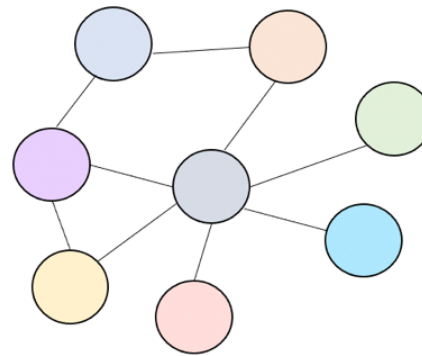
Candidate directional anchor gene



↑ Expression = ↑ Disease



Biological network with directional anchor as seed gene



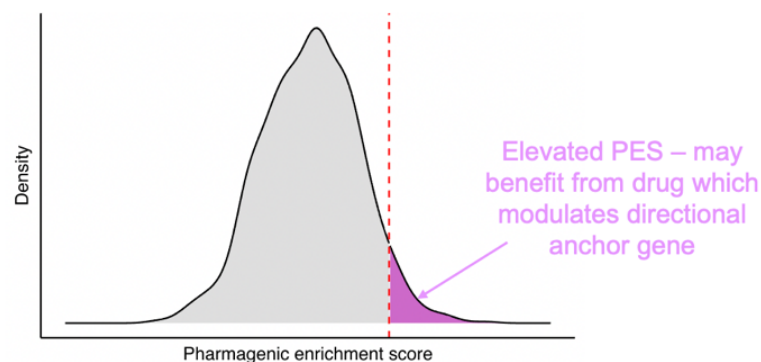
Effect size of SNPs in network on disorder

$$\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \dots, \hat{\beta}_k$$

Polygenic score for SNPs in network

Pharmagenic enrichment score [PES]

$$PES_i = \sum_{i=1} \hat{\beta}_j G_{ij}$$



199 **Figure 1. Overview schematic of the integration of candidate directional anchor**
200 **genes with *pharmagenic enrichment scores*.** Directional anchor genes are genes
201 targeted by an approved compound in what is genetically predicted through integration
202 with expression data to decrease the risk of the disorder or modulate the trait, if
203 continuous, in a clinically useful manner. For instance, if increased expression of a gene
204 is associated with a disorder through a transcriptome or proteome-wide association
205 study (TWAS/PWAS) or Mendelian randomisation (MR) using quantitative trait loci
206 as instrumental variables, then an antagonist of said gene may be a repurposing
207 opportunity. Directional anchor genes then act as seed genes to define a network of
208 other genes that interact with them. SNPs mapped to this network are then utilised to
209 construct a *pharmagenic enrichment score* (PES) for the network. In the case of a
210 binary disease phenotype, the interpretation of the PES would be that individuals with
211 an elevated score relative to an appropriate population reference may benefit from a
212 compound which modulates the directional anchor gene. The hypothesis underlying
213 this is that these individuals will have genetic risk that impacts upstream or downstream
214 processes relating to the directional anchor, as well as the anchor gene itself, which
215 could be addressed by the repurposing candidate in question.

216

217 **Schizophrenia and bipolar disorder genome-wide association studies**

218 We obtained GWAS summary statistics for schizophrenia (SZ) and bipolar disorder (BIP) from
219 the psychiatric genomics consortium [9, 10]. The SZ GWAS was a mega-analysis of mostly
220 European ancestry cohorts and comprised 67,390 cases and 94,015 controls, whilst the
221 European ancestry BIP GWAS mega-analysis had 20,352 cases and 31,358 controls. In
222 addition, we also utilised the same SZ GWAS with a constituent cohort removed (Australian
223 Schizophrenia Research Bank) when we profiled PES within that dataset, as described in a
224 proceeding section of the methods.

225

226 **Transcriptome and proteome-wide association studies**

227 A transcriptome-wide association study (TWAS) and a proteome-wide association study
228 (PWAS) was performed of SZ and BIP by leveraging genetically imputed models of mRNA
229 and protein expression, respectively. Specifically, we utilised the FUSION approach for
230 TWAS/PWAS, with full details outlined in the supplementary materials [22]. Expression
231 weights for the TWAS were derived from postmortem brain (GTEx v7, PsychENCODE) and
232 whole blood (GTEx v7), whilst protein expression weights were similarly from postmortem

233 brain (ROSMAP) and whole blood (ARIC) [22–25]. The FUSION methodology integrates
234 SNP-effects from the model of genetically predicted expression with the effects of the same
235 SNPs on SZ or BIP, after accounting for linkage disequilibrium, such that the TWAS Z score
236 can be a conceptualised measure of genetic covariance between mRNA or protein expression
237 of the gene and the GWAS trait of interest. We utilised a conservative method for multiple-
238 testing correction whereby the Bonferroni methodology was implemented to divide the alpha
239 level (0.05) by the total number of significantly *cis*-heritable models of genetically regulated
240 expression (GReX) tested from any brain tissue considered or whole blood (Supplementary
241 Text, Supplementary Tables 1-4). Several genes had GReX available in multiple-tissues, thus
242 rendering Bonferroni correction conservative, however, we implemented this approach to
243 capture only the most confidently associated genes that could constitute drug-repurposing
244 candidates. For candidate directional anchor genes derived from TWAS/PWAS, we
245 probabilistically finemapped those regions using the FOCUS methodology using the default
246 prior ($p = 1 \times 10^{-3}$) and prior variance ($n\sigma^2 = 40$) to approximate Bayes' factors, such that the
247 posterior inclusion probability (*PIP*) of each gene being a member of a credible set with 90%
248 probability of containing the causal gene could be derived [26]. We also investigated the impact
249 of using a more conservative prior as outlined in the supplementary text. Moreover, we tested
250 whether SNPs that constitute the GReX model and either SZ or BIP displayed statistical
251 colocalisation with the *coloc* package as implemented by FUSION [27].

252

253 **Mendelian randomisation**

254 In addition, we leveraged variants strongly correlated with mRNA (expression quantitative trait
255 loci – eQTL) and protein expression (protein expression quantitative trait loci – pQTL),
256 respectively, as instrumental variables (IVs) in a two-sample Mendelian randomisation (MR)
257 analysis [28]. Analogous to the TWAS/PWAS, eQTL/pQTL were derived from brain
258 (MetaBrain, ROSMAP) and blood (eQTLGen, Zhang *et al.*), with full details described in the
259 supplementary text [28–31]. Strict selection criteria were implemented to select suitable IVs,
260 including only retaining independent genome-wide significant ($P < 5 \times 10^{-8}$) SNPs that were
261 associated with three or less mRNA/proteins in each relevant tissue/study (Supplementary
262 Text). Moreover, we utilised a more stringent LD clumping procedure for eQTLs given the
263 greater power and sample sizes for these datasets also results in immense pleiotropy amongst
264 the SNP effects on mRNA. This was achieved by only selecting the most significant
265 independent SNPs using one megabase clumps, with LD estimated using the 1000 genomes
266 phase 3 panel. The effect of mRNA or protein expression for any given gene on SZ or BIP was

267 estimated using the Wald ratio (single IV) or an inverse-variance weighted estimator (multiple
268 IVs, with fixed effects due to the small number of IVs). As in the TWAS/PWAS, we utilised
269 Bonferroni correction across all tissues in the mRNA and protein analyses respectively and
270 then sought to identify candidate directional anchor genes from these signals. For any candidate
271 directional anchor genes, where an approved drug was predicted to reverse the odds increasing
272 direction of expression, we performed a series of sensitivity analyses (Supplementary Text).
273 Briefly, these involved: assessing the genomic locus of the IV SNP, for genes it may be
274 associated with, by testing for evidence of a shared causal variant through colocalisation
275 (default priors) and conducting a phenome-wide Mendelian randomisation analysis (MR-
276 pheWAS) using SNP effects from each trait in the IEUGWAS database. The above MR and
277 sensitivity analyses were performed using the R packages TwoSampleMR v0.5.5, ieugwas
278 v0.1.5, and coloc v4.0.4 [27, 32].

279

280 **Identifying drug repurposing candidates**

281 We searched genes prioritised from the TWAS/PWAS or MR analyses in the Drug-gene
282 interaction database (DGIdb v4.2.0 – accessed April 2021) to identify approved compounds
283 that could reverse the odds increasing direction of expression for SZ or BIP [33]. DGIdb
284 combines data from databases such as DrugBank, as well as curated literature sources. We
285 defined high confidence drug-gene interactions as those reported in DrugBank as well as at
286 least one other database or literature source.

287

288 **Identification of genes interacting with candidate directional anchor genes**

289 Protein-protein interaction data was downloaded from the STRING database v11 [34]. We
290 utilised each of the six candidate directional anchor genes as a seed gene, separately, and
291 constructed a network of genes predicted to interact with the seed gene by retaining high
292 confidence edges (confidence score > 0.7) derived from experimental evidence or curated
293 protein-complex and pathway databases, as this is generally considered the most rigorous
294 evidence from STRING. We then tested which gene-sets curated by the g:Profiler (version
295 e104_eg51_p15_3922dba) resource (GO, KEGG, Reactome, WikiPathways, TRANSFAC,
296 miRTarBase, Human Protein Atlas, CORUM, and Human phenotype ontology) were
297 overrepresented amongst the genes in each network, using the g:SCS (set counts and sizes)
298 multiple-testing correction method implemented by g:Profiler that has been shown to better
299 account for the complex, overlapping nature of these data [35]. We considered a corrected *P*
300 value < 0.05 as significant.

301 We then tested the association of the genes in each of these networks, with and without the
302 gene removed, with the common variant signal in the SZ and BIP GWAS using MAGMA
303 v1.09 [36]. Briefly, SNP-wise P values were aggregated at gene-level, with SNPs annotated to
304 genes using two different sets of genic boundary extensions to capture potential regulatory
305 variation, conservative (5 kilobases (kb) upstream, 1.5 kb downstream), and liberal (35 kb
306 upstream, 10 kb downstream). Gene-set association is implemented by MAGMA using linear
307 regression, whereby the probit transformed genic P values (Z scores) are the outcome with a
308 binary explanatory variable indicating whether a gene is in the set to be tested (β_S), covaried
309 for other confounders like gene size, as described previously. The test statistic of interest is a
310 one-sided test of whether $\beta_S > 0$, and thus, quantifies if the genes in the set are more associated
311 than all other genes. We also investigated the association of the approximately 34,000 gene-
312 sets collated by g:Profiler, such that we could demonstrate whether gene-sets overrepresented
313 in each network were also associated with SZ or BIP.

314

315 **Constructing directional anchor gene network PES**

316 We sought to utilise variants annotated to the genes within the network of each candidate
317 directional anchor genes to develop *pharmacogenic enrichment scores* for SZ and BIP,
318 respectively. As described previously, a PES is analogous to a genome-wide PRS in its
319 derivation, with the key defence that it only utilises variants mapped to the gene-set of interest
320 (equation 1) [18]. Specifically, a PES profile in individual i comprises the sum of the effect
321 size of j variants from the GWAS ($\hat{\beta}_j$) annotated to at least one gene in set M , multiplied by
322 the allelic dosage under an additive model ($G_{ij} \in G = 0, 1, 2$).

$$323 \quad PES_i = \sum_{j=1}^M \hat{\beta}_j G_{ij} \quad [1]$$

324

325 The genome wide PRS for SZ and BIP are essentially the same model but M incorporates the
326 entire genome. In accordance with the MAGMA analyses, we tested two genic boundary
327 configurations for evaluating the best performing PES for each directional anchor gene network
328 – conservative and liberal. Our previous PES related approaches utilised the LD clumping and
329 thresholding (C+T) approach, whereby SNPs are ‘clumped’ such that the retained SNPs are
330 largely independent and ‘thresholded’ based on their association P -value in the GWAS. In each
331 case the threshold was set at the optimum for the druggable gene-set association at the
332 population level. However, given we selected the gene-sets in this study based on interactions

333 with the candidate directional anchor gene, we tested four different P value thresholds
334 ($P_T \in T = \{0.005, 0.05, 0.5, 1\}$), which represent a model with all SNPs, nominally significant
335 SNPs, and a threshold an order of magnitude above or below the nominal threshold. These
336 choices of T have been discussed extensively elsewhere [18–20]. We utilised PRSice-2 v2.3.3
337 (linux) for the C+T models [37]. In addition, we utilised a penalised regression framework to
338 shrink SNP effect sizes to optimise the model for each PES, as implemented by the standalone
339 version of lassosum v0.4.5 [38]. The implementation for this method has been outlined
340 extensively elsewhere, with the optimal tuning parameter (λ) based on the score that displays
341 the highest correlation with the phenotype and the best performing constraint parameter (s)
342 chosen from a range of *a priori* specified values to decrease computational burden (0.2, 0.5,
343 0.9, and 1).

344

345 **Training and validation of directional anchor gene network PES**

346 We utilised the prospective UK Biobank (UKBB) cohort to define the best performing PES for
347 each directional anchor gene network. Our group has previously processed the UKBB genotype
348 data such that unrelated individuals of white British ancestry were retained, along with other
349 sample and variant level quality control considerations applied [20]. As a result, the
350 composition of the full UKBB cohort in this study was 336,896 participants for which up to
351 13,568,914 autosomal variants were available (imputation INFO > 0.8). SZ and BIP cases were
352 defined in the UKBB using a combination of self-report data both from the general assessment
353 visit and the mental health questionnaire (MHQ), along with hospital inpatient data (primary
354 or secondary ICD-10 codes), with full details in the supplementary text. In total, there were
355 631 UKBB participants from the study cohort defined as having SZ, with 1657 BIP cases
356 identified. The controls were double the number of the respective case cohorts randomly, and
357 independently for SZ and BIP, derived from 75,201 individuals with genotype data that
358 completed the MHQ and did not self-report any mental illness. The full complement of SZ
359 cases with the aforementioned controls ($N = 1262$) was utilised as the training set for the SZ
360 scores given the relatively small number of cases. As a result, we utilised the Australian
361 Schizophrenia Research Bank (ASRB) cohort as a validation set to attempt to replicate the
362 associations observed with the scores, which has been described elsewhere [18, 39]. The ASRB
363 was a component of the PGC3 SZ GWAS, and thus, we retrained all the best performing PES
364 scores using summary statistics with the ASRB cohort removed before testing them in that
365 dataset. The BIP analyses employed a 70/30 split for the training and validation cohort in the

366 UKBB, with double the number of independent MHQ derived healthy controls utilised for each
367 case-set. Further information regarding the demographic composition of these cohorts is
368 provided in the supplementary text.

369

370 The PES and PRS constructed using the C+T configurations and penalised regression were
371 scaled to have a mean of zero and unit variance before evaluating their association with SZ or
372 BIP, for the respective scores in the UKBB training cohorts, using binomial logistic regression
373 covaried for sex, age, genotyping batch, and five SNP derived principal components. The
374 optimal PES for each network was selected for each disorder separately by calculating the
375 variance explained on the liability scale (Nagelkerke's R^2), assuming a 0.7% and 1%
376 prevalence for SZ and BIP, respectively [40]. These PES/PRS that explained the most
377 phenotypic variance were then profiled and tested in the validation sets. For PES that were
378 significantly associated with either disorder, we conservatively constructed another model that
379 also included genome wide PRS, with a χ^2 test of residual deviance performed to ascertain
380 whether adding the PES in addition to the PRS significantly improved model fit. Correlations
381 (Pearson's) amongst scaled PES and PRS were visualised using the corrplot package v0.84.
382 Individuals with at least one elevated PES in the training cohorts (highest decile) were
383 identified, with this binary variable tested for association with SZ or BIP using another logistic
384 regression model. Finally, we also considered *residualised* PES, whereby the residuals were
385 extracted and scaled from a linear model that regressed genome wide PRS against principal
386 components and genotyping batch on the score in question. All analyses described in this
387 paragraph were performed utilising R v3.6.0.

388

389 **Biochemical and mental health phenome-wide association studies**

390 We then investigated the correlations between the best performing PES for each network and
391 i) blood or urine biochemical traits, and ii) self-reported mental health disorders besides SZ or
392 BIP. The biochemical analyses were performed in up to 70,625 individuals who did not self-
393 report any mental health disorders in the MHQ and were also not included in the SZ or BIP
394 training/validation sets as controls. There were 33 biochemical traits tested (raw values -
395 Supplementary Table 5) in a linear model with each PES or PRS as an explanatory variable
396 covaried for sex, age, sex x age, age², 10 principal components, and genotyping batch. We also
397 performed sex-stratified analyses, with oestradiol additionally considered in females. A
398 number of sensitivity analyses were performed for PES-biochemical trait pairs that were
399 significantly correlated after FDR correction— i) adjustment for genome-wide PRS, ii) natural

400 log transformation of the biochemical outcome variable, iii) inverse-rank normal transformed
401 residuals as the outcome variable from a model that regressed sex, sex x age, and age², and iv)
402 adjustment for statin use (given the number of lipid related signals uncovered). These
403 correlations are observational in nature, and thus, there are several other potential confounders
404 that could be considered – however, given the potential biases induced by adjusting for
405 heritable covariates, we utilised the above strategies as a baseline suite of sensitivity analyses
406 [41]. A specific test of sexual dimorphisms between the regression results in males and females
407 was also performed based on the sex-specific regression estimates and standard errors, as
408 outlined elsewhere [42]. Moreover, we then evaluated the association between each score and
409 14 non-SZ or BIP mental health disorders which individuals who completed the MHQ were
410 asked to self-report (Supplementary Table 6). In all instances, we used the 70,625 individuals
411 who did not self-report any mental disorders as the controls in binomial logistic regression
412 models covaried for the same terms as in the biochemical analyses.

413

414 **RESULTS**

415

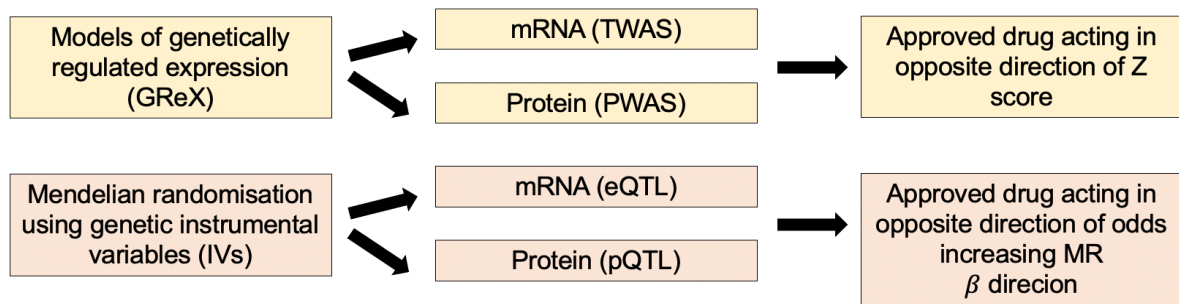
416 **Candidate directional anchor genes reveal novel drug repurposing opportunities in** 417 **psychiatry**

418 We sought to identify candidate directional anchor genes for SZ or BIP by integrating GWAS
419 summary statistics for these traits with transcriptomic and proteomic data collected from either
420 blood or post-mortem brain (Figure 2a). Specifically, we utilised genetically imputed models
421 of mRNA or protein expression to conduct a TWAS and PWAS, respectively (Supplementary
422 Tables 1-4). Genome-wide significant eQTLs and pQTLs were also leveraged as instrumental
423 variables in a more conservative two-sample Mendelian randomisation analysis to explicitly
424 test for any causal effects of mRNA or protein expression (Supplementary Tables 7-10). After
425 implementing Bonferroni correction within each analysis set (TWAS, PWAS, eQTL-MR,
426 pQTL-MR), we found several genes for which their expression was associated with at least
427 one of the psychiatric phenotypes at the mRNA or protein level that was also putatively
428 modulated by an approved compound in a risk decreasing direction. There were 13 druggable
429 genes from TWAS for which the direction of genetically predicted mRNA expression
430 correlated with SZ could be pharmacologically counteracted, whilst there were two such genes
431 for BIP, some examples of which are visualised in Figure 2b. For instance, imputed mRNA
432 expression of the calcium voltage-gated channel subunit gene *CACNA1C* was negatively
433 correlated with SZ ($P = 3.65 \times 10^{-15}$), and thus, an activator of this gene like the antiarrhythmic

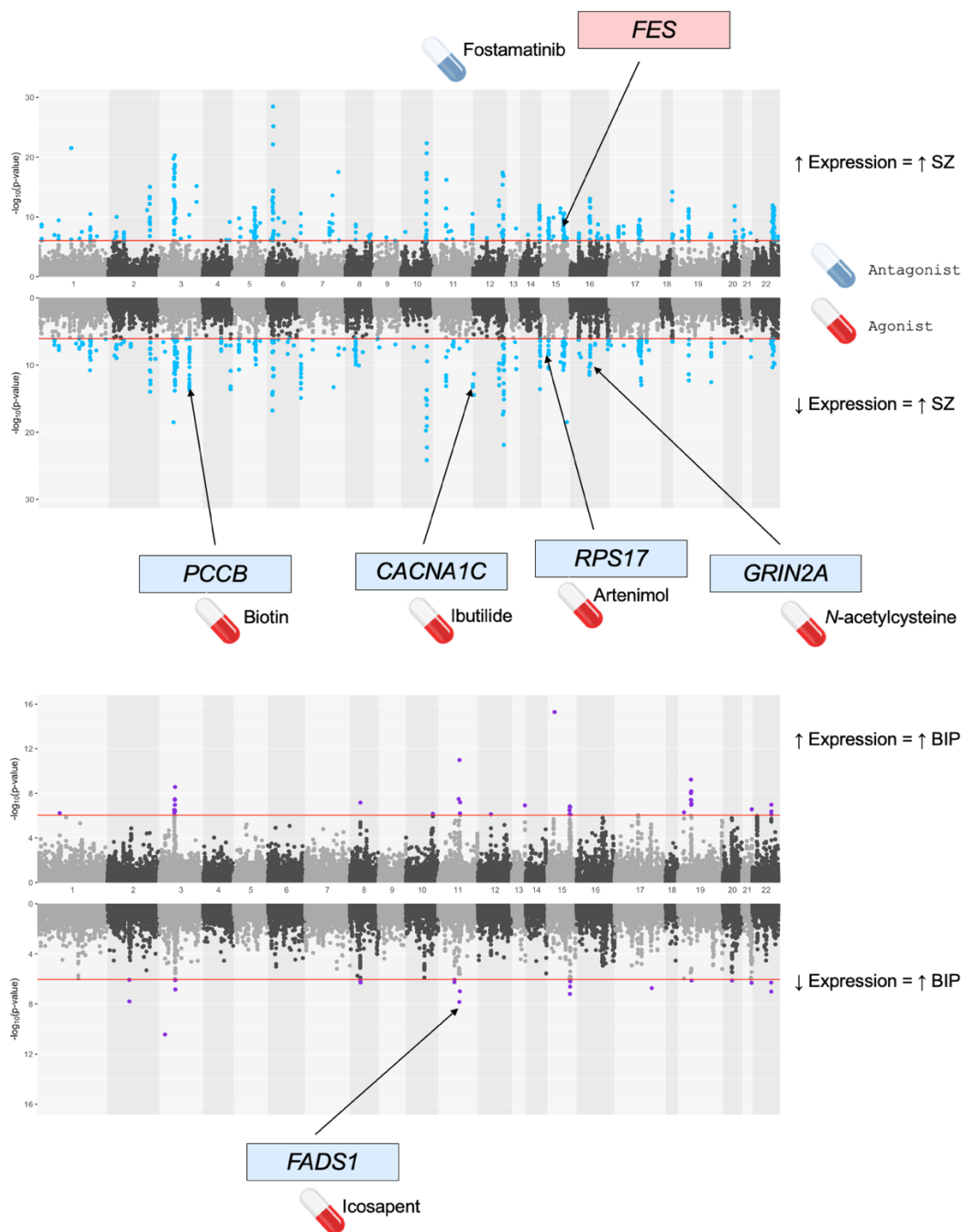
434 agent Ibutilide may be a repurposing candidate for SZ, with this gene also trending towards
435 association with BIP in the same direction ($P = 3.18 \times 10^{-5}$). We compared the TWAS results
436 to that of a PWAS using data from blood or brain tissue, although the number of proteins
437 assayed in these studies was considerably smaller than that of the number of mRNA available,
438 and thus, most of the candidate genes derived using TWAS did not have protein measurements
439 available for a direct comparison of the effect of protein expression relative to mRNA.
440 However, there were two Bonferroni significant TWAS genes that represented a plausible
441 repurposing candidate with protein expression data available (*NEK4* and *CTSS*), with both
442 genes showing a similar strength of association in the PWAS. For any given genetically
443 predicted mRNA or protein expression, the gene is not necessarily causal due to LD complexity
444 and other phenomena such as co-regulation [43]. This is critical for using these approaches to
445 direct drug repurposing as the target the genes need to correspond to the genetic association for
446 the disorder. As a result, we implemented a Bayesian finemapping procedure for each TWAS
447 candidate gene to identify plausible causal genes in each locus (Supplementary Text). We
448 found four repurposing candidate genes for SZ with strong evidence for membership of a
449 credible set with 90% probability of containing the causal gene ($PIP > 0.8$ – *PCCB*, *GRIN2A*,
450 *FES*, and *CACNAID*). However, *CACNAID* was excluded from further analyses due to the
451 poor performance of its imputed model and complexity of its locus on chromosome three, as
452 outlined more extensively in the supplementary text. We then considered a more lenient
453 posterior inclusion probability of 0.4, which identified two more genes for SZ (*CACNAIC* and
454 *RPS17*), and a BIP gene (*FADSI*). Colocalisation analyses were also performed to test a
455 related, but distinct hypothesis that the GWAS signal and SNP weights in the expression model
456 share an underlying single causal variant. Interestingly, for the genes selected using the lower
457 confidence $PIP > 0.4$ threshold, we found strong evidence for a shared causal variant ($PP_{H4} >$
458 0.9), supporting their inclusion as putative drug repurposing targets. We did not consider the
459 two genes shared with the PWAS any further as they did not display strong finemapping
460 support in the TWAS, which is a more accurate representation of any given locus due to the
461 more expansive number of genes with RNAseq available. In summary, using a genetically
462 imputed expression approach we identified five candidate directional anchor genes for SZ and
463 one for BIP (Table 1). For example, imputed *GRIN2A* mRNA expression was negatively
464 correlated with SZ ($P = 1.44 \times 10^{-9}$), with a trend also observed for BIP ($P = 5.07 \times 10^{-3}$), with
465 compounds of interest in psychiatry, such as *N*-acetylcysteine, known to agonise this subunit
466 [44, 45].
467

468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501

a



b



502

503 **Figure 2. Identification of candidate directional anchor genes for schizophrenia**
504 **and bipolar through integration of transcriptomic and proteomic data. (a)**

505 Schematic for the prioritisation of candidate directional anchor genes through models
506 of genetically regulated expression (GReX, yellow) and Mendelian randomisation
507 (orange). In both instances, approved compounds are derived for implicated genes that
508 reverse the odds increasing direction of mRNA or protein expression. TWAS =
509 transcriptome-wide association study, PWAS = proteome-wide association study (b)
510 Results of the multi-tissue (brain and blood) TWAS for schizophrenia (SZ, top) and
511 bipolar disorder (BIP, bottom). The Miami plot visualises the $-\log_{10}$ transformed P
512 value of association with genes exhibiting a negative genetic covariance between
513 expression and the trait, that is, TWAS $Z < 0$, plotted in the downward direction. The
514 red line denotes the Bonferroni threshold. The candidate directional anchor genes from
515 the TWAS approach are highlighted on the plot along with their putative repurposing
516 candidate that corrects the odds-increasing direction of expression. For example,
517 predicted *PCCB* expression is negatively correlated with SZ, and thus, a *PCCB* agonist
518 like biotin may be clinically useful.

519

520 To prioritise candidate directional anchor genes, we utilised eQTL and pQTL as IVs to estimate
521 the causal effect of mRNA or protein expression on either disorder outcome in a Mendelian
522 randomisation analysis, given more onerous assumptions are met (Supplementary Materials,
523 Supplementary Tables 7-10). This is critical as the use of molecular QTLs related to variables
524 like mRNA expression as IVs is challenging due to LD complexity and the potential effect of
525 QTLs on multiple genes [28, 46]. As a result, we sought to complement the above discovery
526 orientated TWAS/PWAS with more conservative selection criteria for an eQTL or pQTL to be
527 an IV, particularly in the case of eQTLs where sample sizes for some tissues are now very
528 large. Independent SNPs ($LD\ r^2 < 0.001$) acting as eQTLs or pQTLs at a threshold of genome-
529 wide significance ($P < 5 \times 10^{-8}$) were selected from post-mortem brain or blood, as outlined in
530 the methods and supplementary materials. Due to the conservative nature of these analyses,
531 many of the genes considered in the TWAS/PWAS did not have a suitable IV available.
532 Conversely, a small number of genes that did not display adequate multivariate *cis*-heritability
533 in the TWAS/PWAS weights could now be included. The mRNA models after Bonferroni
534 correction uncovered four genes for which expression exerted a potential causal effect on SZ
535 or BIP with a suitable compound approved to reverse the odds-increasing direction of effect.

536 There were three for SZ (*PCCB*, *NEK1*, and *PTK2B*), as well as *FADS1* for bipolar.
537 Interestingly, *PCCB* and *FADS1* overlapped with the TWAS results – as an example, each
538 standard deviation increase in cortical *FADS1* expression was associated with approximately
539 15.23% [95% CI: 8.69%, 21.77%] decrease in the odds of BIP, which could be accentuated by
540 a *FADS1* agonist like the omega-3 fatty acid supplement icosapent (Ethyl eicosapentaenoic
541 acid). We then performed a series of sensitivity analyses to assess IV validity and for evidence
542 of confounding pleiotropy (Supplementary Materials). These analyses supported *PCCB* and
543 *FADS1* as candidate directional anchor genes. The index IV-SNPs mapped to *PCCB* and
544 *FADS1* expression, respectively, was then utilised to perform a phenome-wide scan spanning
545 over 10,000 GWAS of the effect of expression of these two genes using SNP effect sizes from
546 the IEUGWAS database (Supplementary Tables 11-13). Firstly, we found that increased
547 cortical expression of *PCCB*, which was associated with decreased odds of SZ from a previous
548 GWAS, was also linked to a reduction in other psychiatric phenotypes from self-reported UK
549 Biobank GWAS such as worry, neuroticism, nervousness, and tenseness, supporting the utility
550 of a *PCCB* agonist, like biotin, as a repurposing candidate. Secondly, the phenome-wide data
551 for increased cortical *FADS1* expression demonstrated, as expected, a strong effect on lipids,
552 including increased HDL and decreased triglycerides. While another candidate for BIP
553 (*MAP2K2*) was suggested as a *trans*-pQTL by this approach, it was excluded to retain the most
554 biologically confident associations. Although the MR approach did not add any additional
555 candidate directional anchor genes (after exclusion of *MAP2K2*), it provided more support to
556 *PCCB* and *FADS1*. A less conservative MR paradigm in terms of IV selection would likely
557 yield more genes but as our TWAS/PWAS analyses were already discovery focused, we
558 believe this would not be appropriate given the underlying assumptions of MR. We summarise
559 the candidate directional anchor genes in table 1.

560
561
562
563
564
565
566
567
568
569

570 **Table 1. Candidate directional anchor genes for schizophrenia and bipolar along with**
571 **their associated drug repurposing candidates.**

572

Gene	Disorder	Protective direction	Repurposing candidates	Sensitivity analyses
<i>PCCB</i>	SZ	Increased expression	Biotin	High confidence causal gene from TWAS ($PIP > 0.8$). Further supported by MR
<i>FADSI</i>	BIP	Increased expression	Icosapent Linolenic acid	Moderate confidence causal gene from TWAS ($PIP > 0.4$, additionally supported by colocalisation – $PP_{H4} > 0.8$). Further supported by MR
<i>GRIN2A</i>	SZ	Increased expression	<i>N</i> -acetylcysteine Glycine	High confidence causal gene from TWAS ($PIP > 0.8$).
<i>CACNA1C</i>	SZ	Increased expression	Ibutilide	Moderate confidence causal gene from TWAS ($PIP > 0.4$, additionally supported by colocalisation – $PP_{H4} > 0.8$).
<i>RPS17</i>	SZ	Increased expression	Artemimol	Moderate confidence causal gene from TWAS ($PIP > 0.4$, additionally supported by colocalisation – $PP_{H4} > 0.8$).
<i>FES</i>	SZ	Decreased expression	Fostamatinib Lorlatinib	High confidence causal gene from TWAS ($PIP > 0.8$).

573

574

575 **Interaction networks related to directional anchor genes capture biology associated with**
576 **schizophrenia and bipolar disorder**

577 We sought to define a network of genes that display high confidence interactions with each
578 candidate directional anchor gene using data from the STRING database, such that we can then
579 construct a *pharmagenic enrichment score* using variants annotated to these genes. The number
580 of direct interactions identified for each of the six candidate genes, excluding the gene itself,
581 were as follows (Supplementary Tables 14): *CACNA1C* network (83 genes), *FADSI* network
582 (16 genes), *FES* network (37 genes), *GRIN2A* network (54 genes), *PCCB* network (26 genes),
583 and *RPS17* network (254 genes). All of these networks displayed significantly more
584 interactions than what would be expected by chance alone ($P < 1 \times 10^{-16}$, Supplementary Table
585 15), with an example of two of these networks (*CACNA1C* and *FADSI*) visualised in Figure

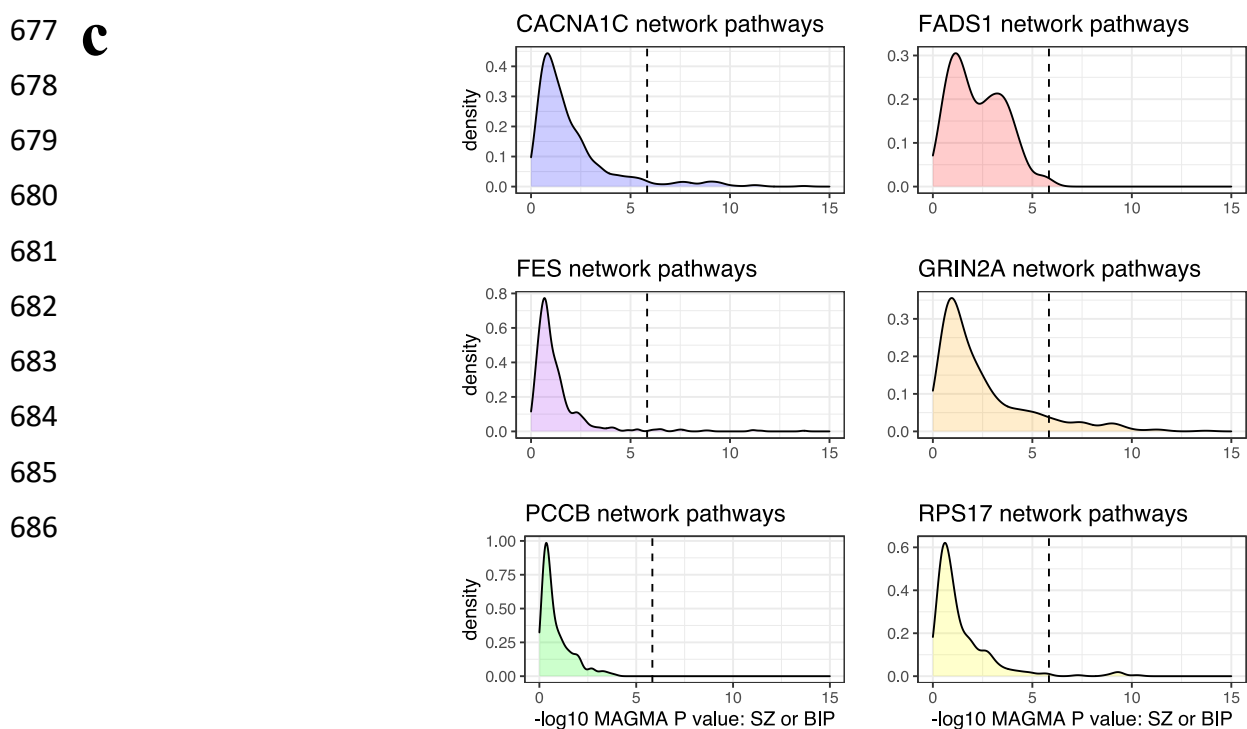
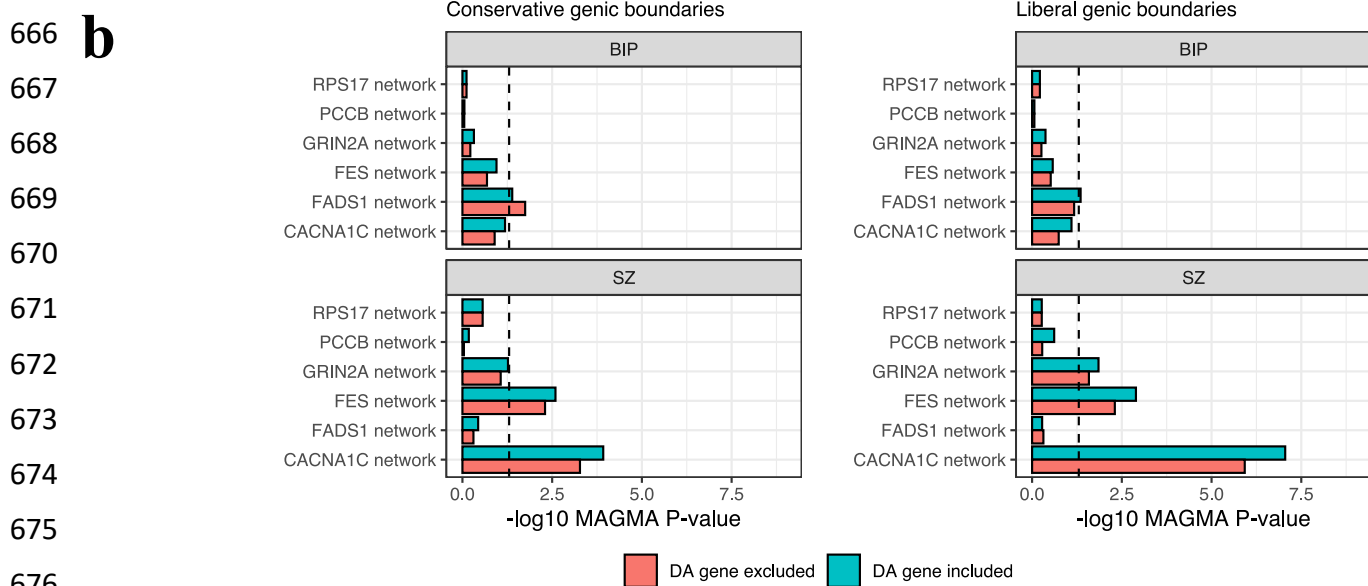
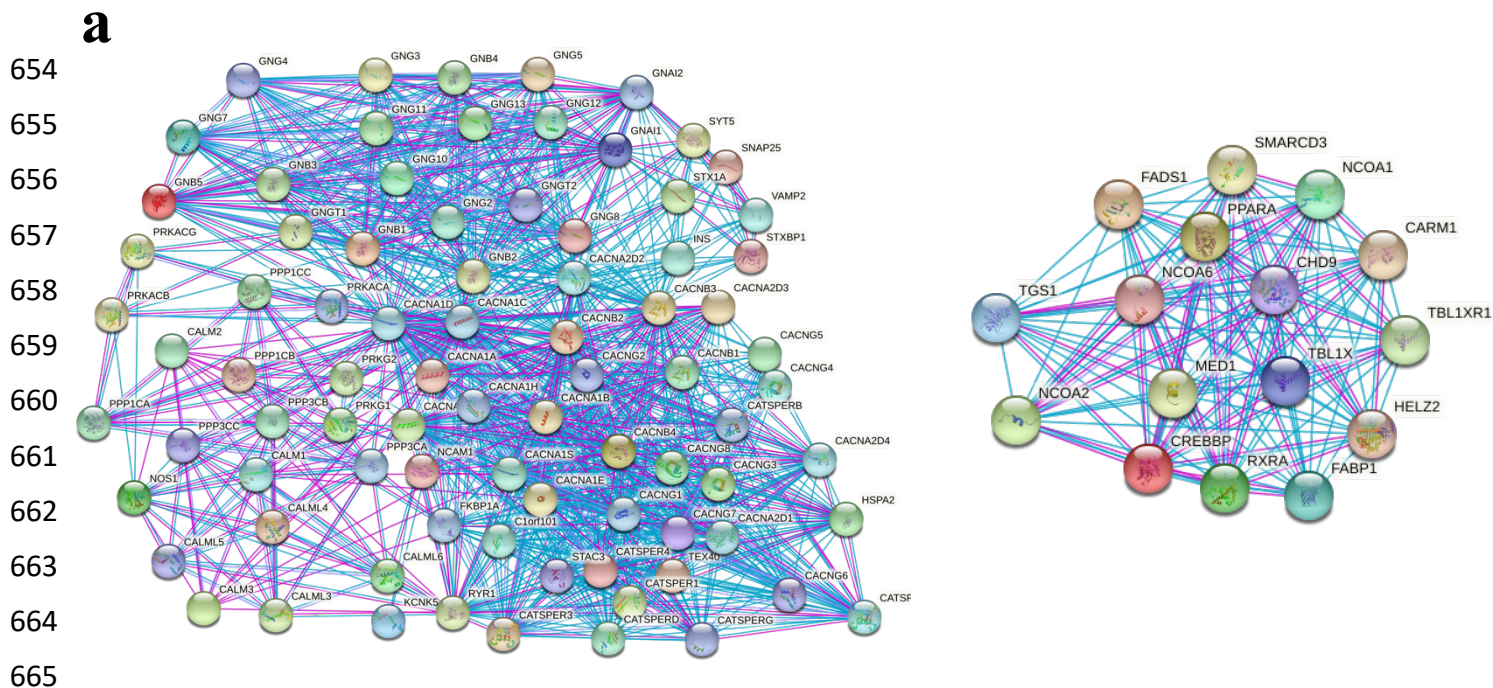
586 3a. These networks likely represent heterogenous biological processes in which the directional
587 anchor gene may participate, and thus, we sought to better understand the biology of these
588 interacting genes by testing their overrepresentation within biological pathways and other
589 ontological gene-sets. The six directional anchor gene networks each displayed
590 overrepresentation in pathways related to the known biology of the candidate gene
591 (Supplementary Tables 16-21). For instance, the *CACNA1C* network genes were enriched
592 within several hundred gene-sets, many of which related to neuronal calcium channel biology
593 along with systemic processes known to involve calcium signalling such as pancreatic insulin
594 secretion. Furthermore, the *FADS1* network genes displayed an overrepresentation in several
595 lipid and other metabolic related pathways, whilst *GRIN2A* network genes demonstrated a
596 strong link to neuronal biology.

597

598 We then tested whether there was enrichment of the common variant signal for SZ or BIP in
599 any of these networks, with and without the directional anchor gene included, using MAGMA
600 (Figure 3b, Supplementary Table 22). The *CACNA1C* network was strongly associated with
601 SZ ($P = 8.87 \times 10^{-8}$), even after removing *CACNA1C* itself ($P = 1.19 \times 10^{-6}$). The *FES* and
602 *GRIN2A* networks demonstrated a nominal enrichment of the SZ common variant signal
603 relative to all other genes, $P = 1.28 \times 10^{-3}$ and $P = 0.014$, respectively, remaining significant
604 upon removing the relevant directional anchor genes. None of the other networks were
605 associated with SZ when considering all genes, with only the *FADS1* network demonstrating a
606 nominal association with BIP ($P = 0.04$). Given that these networks represent several different
607 biological processes, we further hypothesised that specific gene-sets for which they were
608 overrepresented may specifically display a stronger association with SZ or BIP (Supplementary
609 Tables 23-28). Indeed, we show that all of the networks had at least one overrepresented
610 pathway that was associated with SZ or BIP using Bonferroni ($\text{FWER} < 0.05$) and Benjamini-
611 Hochberg ($\text{FDR} < 0.05$) correction, with the exception of the sets enriched in the *PCCB*
612 network that only survived correction using FDR. Kernel density estimation plots of the
613 MAGMA gene-set association P values are visualised in figure 3c, which show pathway-
614 associations reaching these thresholds. We briefly describe the results for the *CACNA1C* and
615 *GRIN2A* networks below for illustration. Pathways overrepresented in the *CACNA1C* network
616 related to calcium channel activity displayed strong association with SZ, for instance, *voltage*
617 *gated calcium channel process* ($P = 2.80 \times 10^{-10}$, $q = 6.83 \times 10^{-7}$), whilst the *regulation insulin*
618 *secretion* pathway that also was enriched in the network was associated with SZ and trended
619 towards surviving multiple testing correction for BIP. *GRIN2A* network members also

620 displayed an enrichment amongst several neuronal pathways strongly associated with SZ, such
621 as *synaptic signalling* ($P = 3.88 \times 10^{-8}$, $q = 2.82 \times 10^{-5}$). Taken together, these results suggest
622 that pathways in which genes in each network participate are associated with psychiatric illness
623 and reinforces the biological salience of these networks.

624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653



687 **Figure 3. Biological networks interacting with candidate directional anchor genes.**
688 (a) Visualisation of two networks of genes that putatively interact with *CACNA1C* (left)
689 and *FADS1* (right) based on experimental and curated database evidence. Blue edges
690 represent evidence from curated databases, whilst purple edges denote experimentally
691 determined evidence. (b) Gene-set association (MAGMA) of the entire network for
692 each candidate directional anchor (DA) gene, with and without the DA gene included
693 from the model. Dotted line represents nominal significance ($P < 0.05$). The MAGMA
694 P value is derived from a model which tests whether the common variant signal within
695 genes in the network is greater than what is observed amongst all remaining genes. Two
696 genic boundaries were utilised to annotate SNPs to genes from the GWAS: conservative
697 (5kb upstream, 1.5 kb downstream, left panel) and liberal (35 kb upstream, 10 kb
698 downstream, right panel). (c) Kernel density estimation plots of the MAGMA gene-set
699 association P value for each gene-set tested using either schizophrenia or bipolar
700 results, whichever was more significant, which had a significant overrepresentation of
701 genes within that network. The dotted line represents the Bonferroni significance level
702 for approximately 34,000 gene-sets considered in the full analysis of all gene-sets that
703 were tested for overrepresentation.

704
705

706 **Directional anchor gene network *pharmagenic enrichment scores* display significant trait**
707 **associations after adjustment for genome-wide polygenic risk score**

708 *Pharmagenic enrichment scores* (PES) were then constructed for the genes in each directional
709 anchor gene network using SNP weights for SZ and BIP, respectively. SZ and BIP PES were
710 considered for all six networks given the high genetic correlation between SZ and BIP, as well
711 as extensive phenotypic overlap. We defined a training set of SZ ($N = 631$) and BIP cases (N
712 $= 1161$) in the UK Biobank, with double the number of controls randomly, and independently,
713 selected from individuals with no self-reported mental health conditions for each training set.
714 Two methods were utilised to find the most parsimonious PES profile for each network, along
715 with a genome-wide PRS for SZ and BIP - clumping and thresholding (C+T), and penalised
716 regression (Table 2).

717
718
719

720 **Table 2. Characteristics of the best performing schizophrenia and bipolar genome wide**
721 **PRS along with a *pharmacogenic enrichment score* for each directional anchor gene network**

Score	N _{SNPs} ¹	Beta (SE) ²	P	R ²	Type ³
<i>Bipolar disorder</i>					
PRS	1096096	0.62 (0.04)	2.91x10 ⁻⁵¹	4.96%	Penalised regression
<i>CACNA1C</i> network	396	0.11 (0.04)	2.97x10 ⁻³	0.16%	Penalised regression
<i>FADS1</i> network	196	0.06 (0.04)	0.1	0.05%	C+T
<i>FES</i> network	5	0.10 (0.04)	4.87x10 ⁻³	0.14%	Penalised regression
<i>GRIN2A</i> network	143	0.17 (0.04)	3.81x10 ⁻⁶	0.39%	Penalised regression
<i>PCCB</i> network	520	0.14 (0.04)	2.22x10 ⁻⁴	0.25%	Penalised regression
<i>RPS17</i> network	13724	0.15 (0.04)	1.07x10 ⁻⁴	0.28%	Penalised regression
<i>Schizophrenia</i>					
PRS	1861450	1.01 (0.07)	2.65x10 ⁻⁵¹	9.55%	Penalised regression
<i>CACNA1C</i> network	128	0.09 (0.05)	0.1	0.09%	C+T
<i>FADS1</i> network	36	0.04 (0.05)	0.4	0.02%	C+T
<i>FES</i> network	141	0.17 (0.05)	1.41x10 ⁻³	0.32%	C+T
<i>GRIN2A</i> network	5037	0.18 (0.05)	9.23x10 ⁻⁴	0.35%	Penalised regression
<i>PCCB</i> network	2393	0.07 (0.05)	0.16	0.06%	Penalised regression
<i>RPS17</i> network	77	0.15 (0.05)	3.53x10 ⁻³	0.27%	Penalised regression

722 ¹SNPs with a non-zero coefficient after reweighting in the penalised regression model or independent
723 SNPs after linkage disequilibrium clumping and thresholding (C+T).
724 ²Bipolar disorder or schizophrenia log odds per standard deviation increase in the score (standard error)
725 ³The two models evaluated were clumping and thresholding (C+T) or penalised regression (as
726 implemented by the lassosum package)
727

728 *Schizophrenia*

729 In the SZ cohort, there were three network SZ PES which were significantly associated with
730 increased odds of SZ after multiple testing correction including networks for *FES*, *GRIN2A*,
731 and *RPS17* (Table 2, Supplementary Figure 1a). In the *GRIN2A* network PES featuring 5037
732 variants constructed using penalised regression, explained approximately 0.35% of phenotypic
733 variance on the liability scale (OR per SD in score = 1.19 [95% CI: 1.09, 1.29], $P = 9.23 \times 10^{-4}$).
734 We then conservatively adjusted for the best performing genome-wide SZ PRS and found
735 that the *GRIN2A* network PES remained significantly associated with SZ. In the *FES* and
736 *RPS17* networks, their PES were just below the threshold for significance after PRS adjustment
737 (Supplementary Table 29). It is notable that the SZ network PES profiles were only marginally
738 correlated with genome-wide SZ PRS (all $r < 0.11$, Supplementary Figure 2a), which suggests

739 these scores may capture biologically aggregated risk which is distinct from the
740 undifferentiated genome-wide signal. Interestingly, the majority of individuals with SZ
741 (53.72%) had at least one elevated PES (for the disorder) in the highest decile, which was
742 significant even after covariation for genome-wide PRS – OR = 1.45 [95% CI: 1.22, 1.67], P
743 = 1.57×10^{-3} . Interestingly, amongst individuals in this cohort with relatively low SZ PRS
744 (lowest decile), 12 out of the 19 SZ cases had an elevated PES (63%), with a nominally
745 significant association remaining between elevated PES and SZ amongst those with low
746 genome wide PRS ($P = 0.027$). Upon considering only SZ cases in terms of low PRS, we found
747 that 46.88% had at least one elevated PES. Taken together, these data suggest that some
748 individuals with otherwise low SZ PRS have localised genetic risk within these biological
749 networks. Given the relatively small number of SZ cases in the UKBB, we sought to replicate
750 our results using an independent case-control cohort from the ASRB ($N_{\text{Cases}} = 425$, $N_{\text{Controls}} =$
751 251) rather than splitting the UKBB cohort into a training and validation set. The PES and PRS
752 models were retrained in the UKBB from the same GWAS with the ASRB cohort removed.
753 We were able to nominally replicate the association of the *FES* network PES with SZ in the
754 ASRB (OR per SD = 1.21 [95% CI: 1.04, 1.38], $P = 0.024$), whilst the observed association
755 between the *GRIN2A* and *RPS17* network PES and SZ and in the UKBB was not replicated
756 (Supplementary Table 30).

757

758 *Bipolar disorder*

759 BIP PES within these networks was then profiled in the UKBB training set (Table 2,
760 Supplementary Figure 1b). Interestingly, there were more of the directional anchor gene
761 network PES associated with BIP than SZ, which may have been a reflection the greater
762 statistical power afforded by the larger number of BIP cases in the UKBB. Specifically, all of
763 the network BIP PES were significantly higher in cases, with the exception of the *FADSI*
764 network PES for which there was only a trend towards significance. Analogous to the SZ
765 cohort, the *GRIN2A* network PES explained the most phenotypic variance on the liability scale
766 (0.39%), with each SD in the score associated with a 19% [95% CI: 12%, 26%] increase in the
767 odds of BIP. Moreover, adjustment for BIP genome wide PRS did not ablate the significance
768 of the *GRIN2A* network, *RPS17* network PES, and *FES* network PES, whilst the *PCCB* network
769 PES trended towards significance ($P = 0.1$) after PRS covariation (Supplementary Table 31).
770 The correlations between each PES and BIP PRS were also small (Supplementary Figure 2b),
771 however, the *RPS17* network PES ($r = 0.13$), *CACNA1C* network PES ($r = 0.14$), and *PCCB*
772 network PES ($r = 0.13$) were slightly larger in terms of their PRS correlation than what was

773 observed for the SZ scores. We then investigated the characteristics of individuals with
774 elevated BIP PES and found like SZ that almost half of the BIP cases (49.1%) had at least one
775 PES greater than or equal to the 90th percentile. There was also enrichment of BIP cases
776 amongst participants with an elevated PES after adjusting for BIP PRS (OR = 1.19 [95% CI:
777 1.04, 1.34], $P = 0.027$). Considering BIP cases in the lowest decile of the BIP PRS distribution,
778 36% of them had at least one top decile PES despite their low genome-wide burden, although
779 unlike SZ the association between elevated PES and case-status in this subcohort was not
780 statistically significant. An independent BIP case-control cohort from the UKBB was utilised
781 to attempt to replicate these associations (Supplementary Table 32), and we found that the
782 network *RPS17* PES was significantly enriched in BIP cases in this validation cohort: whilst
783 there was a trend for the *GRIN2A* network PES ($P = 0.052$).

784

785 *Sensitivity analyses for the GRIN2A network PES*

786 The *GRIN2A* network PES explained the most phenotypic variance for SZ and BIP, and
787 survived covariation for genome-wide PRS – thus, we wanted to test whether constructing a
788 PES for this network with the *GRIN2A* gene excluded would still be associated. In other words,
789 we investigated whether there was an effect from variants mapped to the network without the
790 directional anchor gene itself. For example, the *GRIN2A* network PES with the *GRIN2A* gene
791 removed, was still significantly enriched in both SZ and BIP ($P_{SZ} = 9.23 \times 10^{-4}$ and $P_{BIP} = 3.81$
792 $\times 10^{-6}$). The relationship between genome-wide PRS and this PES was also examined in further
793 detail by constructing a ‘residualised PES’ whereby we obtained the normalised residuals from
794 a model that regressed SNP derived principal components, genotyping batch, and genome-
795 wide PRS for BIP and SZ, respectively, on the *GRIN2A* network PES for either disorder. We
796 posit that the individuals with an elevated residualised PES are more likely to represent true
797 enrichment in that network given that the effect of the genome wide PRS, along with variables
798 related to technical artefacts and population stratification, have been adjusted for.
799 Encouragingly, we find that the correlation between the raw *GRIN2A* network PES for either
800 disorder and their respective residualised PES were highly concordant, with the majority of
801 individuals with an elevated *GRIN2A* PES ($\geq 90^{\text{th}}$ percentile) also in that same quantile for the
802 residualised PES (Figure 4a-b).

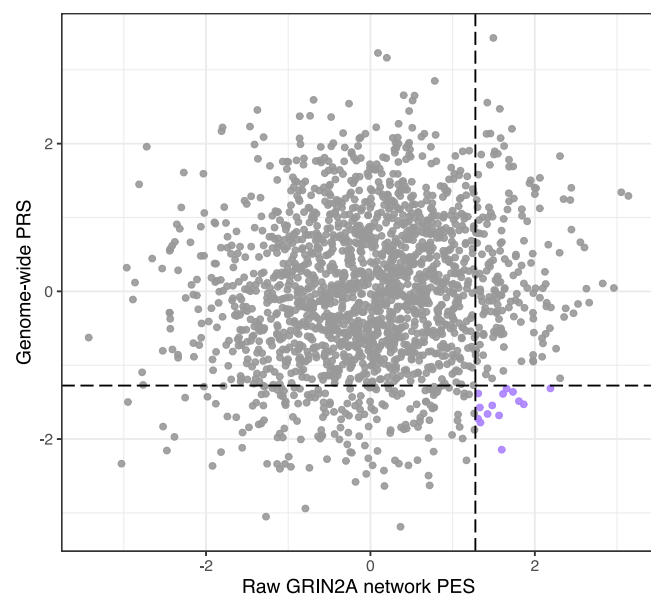
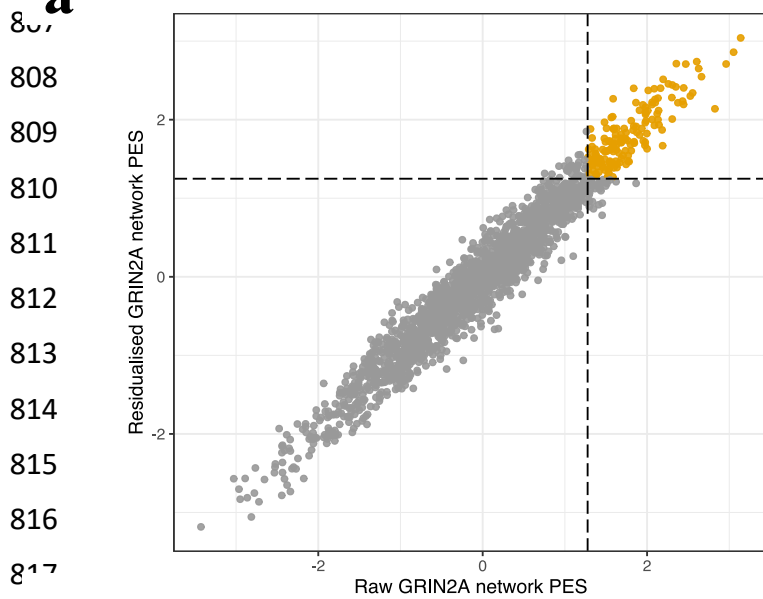
803

804

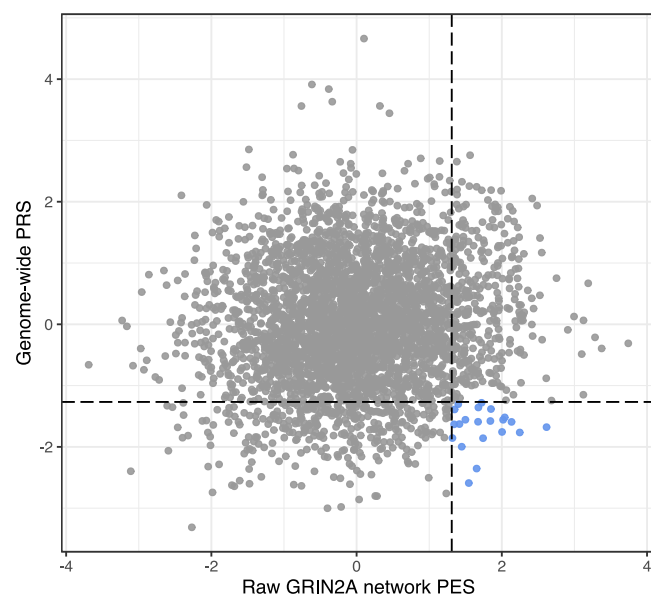
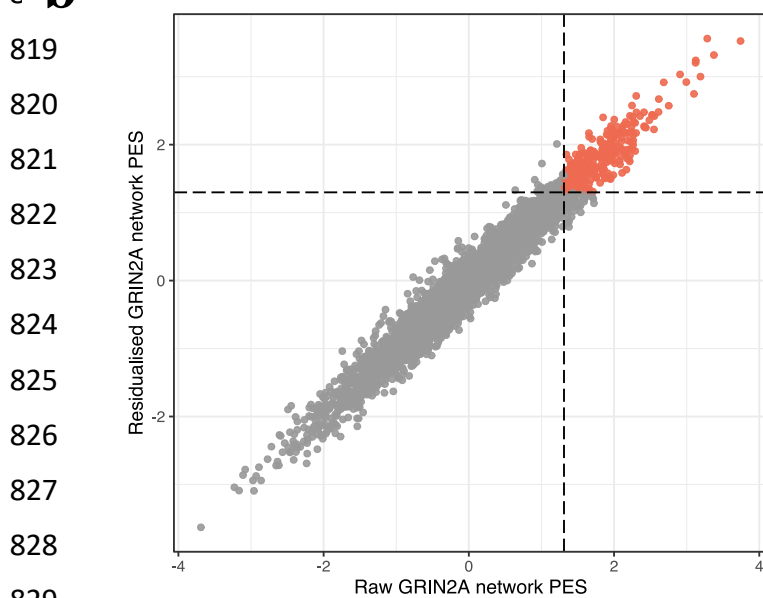
805

806

a



b



830

831

832

833

834

835

836

837

838

839

840

841 **Figure 4. Schizophrenia and bipolar disorder *GRIN2A* directional anchor gene**
842 **network *pharmagenic enrichment scores* and their relationship with PRS.** The
843 scatter plots denote the concordance between the scaled unadjusted (raw) *GRIN2A*
844 network PES for SZ (a) and BIP (b) and both a residualised score and genome wide
845 PRS. Specifically, the left-most scatterplots visualise the relationship between the raw
846 *GRIN2A* network PES and the residuals from a model which regressed genotyping
847 batch, ten SNP derived principal components, and genome wide PRS for the disorder
848 in question (Residualised *GRIN2A* PES). The dotted lines represent the 90th percentile
849 of the raw PES and residualised PES, respectively. The points coloured orange (SZ)
850 and red (BIP) indicate individuals with a PES in the 90th percentile or above for both
851 the raw and residualised scores. The right scatterplots plot the relationship between
852 genome wide PRS for SZ or BIP and the *GRIN2A* network PES. In these instances, the
853 dotted vertical line denotes the 90th percentile of the *GRIN2A* PES, whilst the horizontal
854 dotted line denotes the 10th percentile of genome wide PRS. As a result, the points
855 coloured purple and blue in the SZ and BIP plots, respectively, are individuals with low
856 relative genome wide PRS (lowest decile) but high *GRIN2A* PES (highest decile).

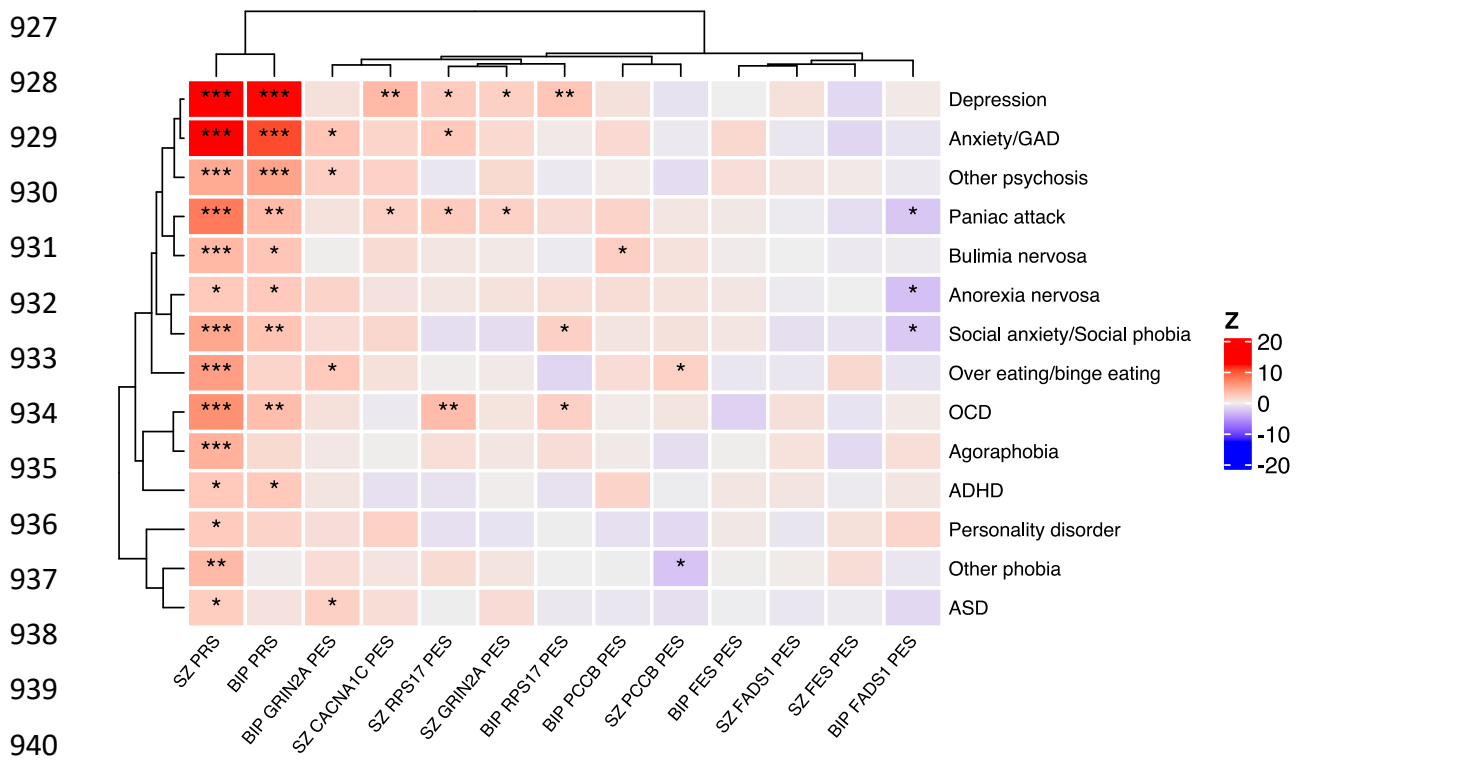
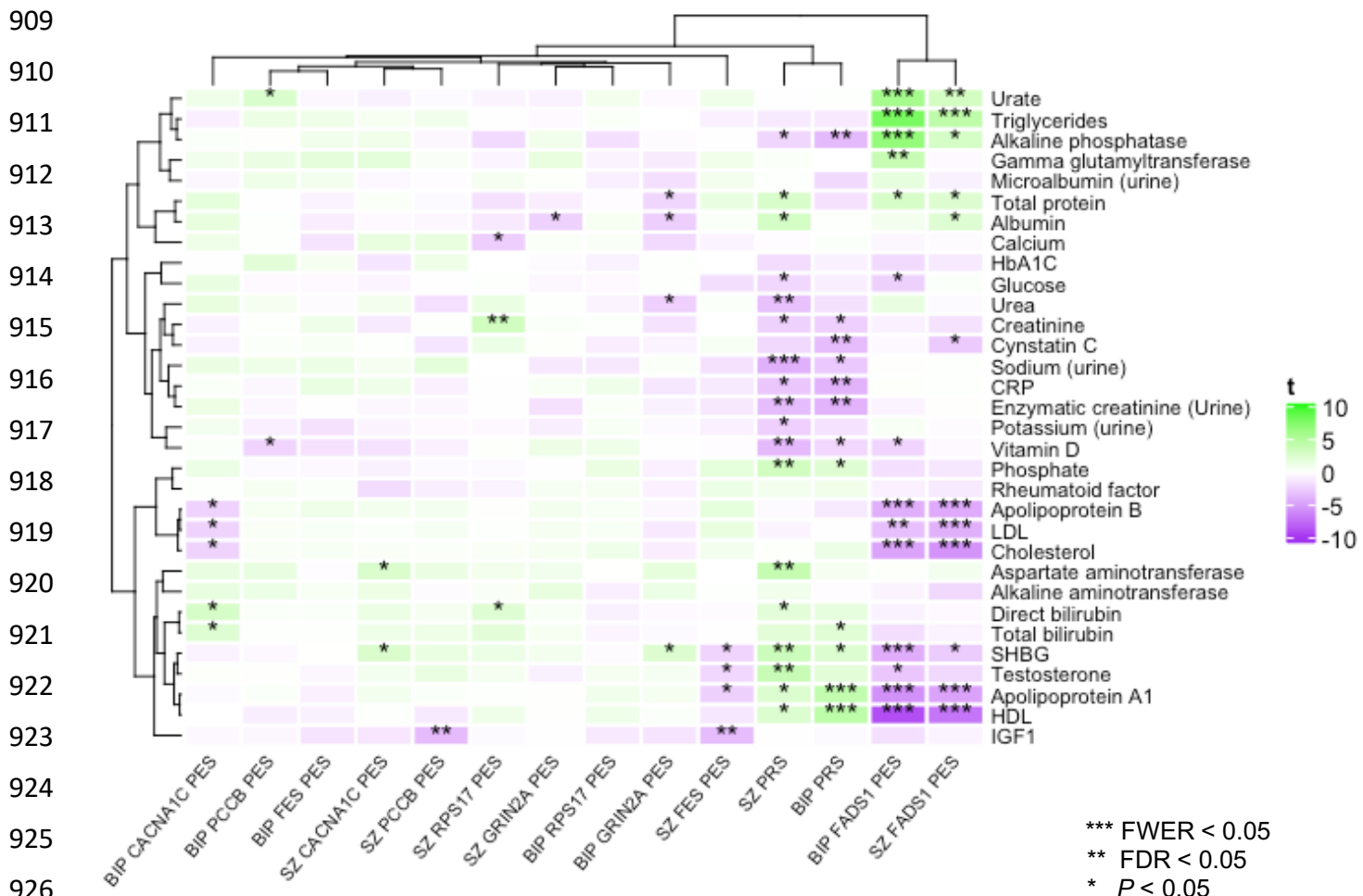
857

858 **Phenotypic correlations of directional anchor gene network scores support their** 859 **biological relevance**

860 We then investigated the association of the directional anchor gene PES in an independent
861 subset of the UKBB with other mental health phenotypes and systemic biochemical measures
862 (Supplementary Tables 33-35 Figure 5). The correlation profile of each PES relative to these
863 phenotypes may also support its clinical utility, whilst it also provides an opportunity to further
864 establish what distinct properties these scores have from a genome wide PRS. Firstly, all SZ
865 and BIP network PES, along with their respective PRS, were regressed against 33 blood and
866 urine measures in up to 70,625 individuals, whilst oestradiol in females was also additionally
867 considered in a sex stratified analysis (Figure 5a). In both sexes, we found that the *FADS1*
868 network PES for SZ and BIP was significantly correlated with lipid related traits after
869 conservative Bonferroni correction for all PES/PES-trait pairs tested ($P < 1.11 \times 10^{-4}$). For
870 instance, these *FADS1* network PES were negatively correlated with HDL cholesterol and
871 apolipoprotein A1 levels, whereas increase in the same PES was associated with higher
872 measured triglycerides. The *FADS1* network PES were also significantly associated other non-
873 lipid biochemical traits including alkaline phosphatase, sex-hormone binding globulin
874 (SHBG), and urate. Notably, after adjusting PES for a genome-wide PRS for SZ or BIP, did

875 not ablate its association with the disorder, suggesting that these signals are not a product of
876 genome-wide polygenic inflation (Supplementary Tables 36-39). To the contrary, there was
877 evidence that PRS was correlated in the opposite direction with lipids to that of the
878 corresponding PES. Given the strong lipid related signals, we also adjusted for statin use in an
879 additional sensitivity analysis, but this similarly did not markedly impact the findings
880 (Supplementary Table 39). Using less stringent FDR correction ($FDR < 0.05$), revealed more
881 PES association with biochemical measures. This included a negative correlation between both
882 the *PCCB* and *FES* network PES for SZ and insulin-like growth factor 1 (IGF1); as well as
883 positive correlation between *RPS17* network PES for SZ and creatinine. However, there was
884 no direct effect of SZ or BIP PRS on IGF-1 or creatinine, with *FES*-related tyrosine kinase
885 activity previously shown in the literature to be associated with IGF-1 biology [47]. Sex
886 stratified analyses identified even more PES associated with a biochemical trait
887 (Supplementary Table 34) – for example, in males the SZ *PCCB* network PES was positively
888 correlated with SHBG, which interestingly is in the opposite direction to the correlation of
889 SHBG observed with the *FADS1* network PES, further highlighting biological heterogeneity
890 amongst different networks. The BIP *CACNA1C* network PES in males was also positively
891 correlated with direct bilirubin using an FDR cut-off, whilst the BIP *GRIN2A* network PES
892 was negatively correlated with measured total protein. Finally, we formally tested for evidence
893 of sexual dimorphic effects of PES/PRS on each biochemical measure and revealed nominal
894 evidence of heterogeneity between sexes in these associations for some traits such as the effect
895 of the *CACNA1C* network PES on direct bilirubin (Supplementary Table 40).

896
897
898
899
900
901
902
903
904
905
906
907
908



943 **Figure 5. Phenome-wide association studies (pheWAS) of each network PES or**
944 **PRS related to serum or urine biochemical measures and mental health disorders.**

945 Heatmap of the association between each network PES and PRS with each trait tested
946 for the biochemical measures (top) and self-reported mental health disorders (bottom).
947 Traits ordering derived from clustering by Pearson's distance. The variable visualised
948 in the heatmaps for the continuous biochemical traits was the regression t value
949 (beta/SE), whilst for the binary mental health phenotypes it was the corresponding Z
950 value from the logistic regression, whereby $Z > 0$ equates to an odds ratio for the
951 disorder > 0 . Asterisks were utilised to denote the significance of the association - * =
952 $P < 0.05$, ** = false discovery rate (FDR) > 0.05 , and *** = family-wise error rate
953 (FWER) < 0.05 .

954
955 We also performed a phenome-wide association study of each score with 14 self-reported
956 mental health disorders in the UKBB cohort, excluding SZ and BIP (Figure 5, Supplementary
957 Table 35). The number of cases ranged from 66 for attention deficit/hyperactivity disorder
958 (ADHD) to 22,974 for depression, with the same cohort of 70,625 individuals without a self-
959 reported mental health condition not featured in the SZ or BIP training/validation sets
960 leveraged as controls. Unsurprisingly, we found that SZ and BIP PRS were strongly associated
961 with increased odds of several mental health disorders after Bonferroni correction, but we also
962 found network PES associated with some of these phenotypes using FDR < 0.05 as the
963 multiple-testing correction threshold. Specifically, there was an association between the SZ
964 *CACNA1C* PES and increased odds of depression, whilst the SZ *RPS17* network PES was
965 associated with increased odds of self-reported OCD. These disorders were also associated
966 with elevated SZ PRS, however, both PES remained significantly higher in those with the
967 respective self-reported phenotypes even after covariation for the effect of the SZ PRS. There
968 were also several other nominal associations ($P < 0.05$), including one of particular interest in
969 the case of the BIP *FADS1* network PES, for which a higher score displayed some evidence of
970 a protective effect on self-reported anorexia nervosa. Whilst this association does not survive
971 multiple testing correction, and thus should be interpreted cautiously, it is notable as the *FADS1*
972 network PES was associated with lipid profiles in an analogous direction to what has previously
973 shown to be genetically correlated with anorexia nervosa GWAS via LD score regression [11,
974 21]. In summary, these data coupled with the biochemical associations support the distinct
975 nature of network PES from PRS and emphasise the unique insights that can be afforded by
976 these partitioned scores.

977 DISCUSSION

978 In this study, we leverage genetics to identify novel drug repurposing candidates for psychotic
979 disorders and show how they may be directed specifically to patients. This precision medicine
980 approach is critical given that phenotypic and genetic heterogeneity confounds traditional
981 interventions designed to target the entire disease population. We believe that a key advance
982 in this study is that it provides a direct link between compounds with evidence for efficacy at
983 the population level, a putative expression related mechanism, and genetic risk scores in the
984 network of genes that interact with the prioritised drug target.

985

986 Transcriptomic and proteomic data integrated with GWAS through GReX (TWAS/PWAS) and
987 causal inference (MR) revealed six interesting target genes for SZ or BIP that could be
988 modulated in a risk decreasing direction by an approved drug. Compounds which target
989 *GRIN2A* in the risk decreasing direction, one of the genes prioritised for SZ, have previously
990 been subjected to randomised control trials as an adjuvant to antipsychotic treatment.
991 Specifically, *N*-acetylcysteine and glycine intervention studies suggested that these compounds
992 could be effective in improving multiple symptom domains including negative and cognitive
993 related dimensions of the disorder [50, 51]. Omega-3 fatty acids, which are related to the
994 biology of the BIP candidate gene *FADS1*, have also been of interest in that disorder, although
995 trials have had contradictory results in terms of benefit [52]. This heterogeneity is unsurprising
996 given the complexity of the BIP phenotype and demonstrates the utility of an approach such as
997 ours which could more specifically target these interventions. The remaining genes and
998 respective repurposing candidates all had plausible biological links to neuronal biology or
999 psychiatric illness [53–56], and thus, warrant further investigation of their utility even without
1000 genetic stratification.

1001

1002 While these six candidates were supported by genetically regulated mRNA expression this was
1003 not confirmed by genetically regulated protein expression, however, this was probably due to
1004 reduced power to detect proteome related associations as this data is still relatively limited [48,
1005 49] is still notably. There are also some limitations to using both the GReX and MR frameworks
1006 for target identification [7, 43, 46], although these are mitigated by a suite of sensitivity
1007 analyses we performed, including probabilistic finemapping and colocalization, which
1008 strengthen our confidence in these six genes. Ideally, future statistical and molecular study of
1009 these association signals should be undertaken to refine our understanding of the role of these
1010 genes in the pathophysiology of SZ and BIP.

1011 We outlined a mechanism whereby genetic risk for the disorder could be profiled amongst the
1012 network of genes which were prioritised as repurposing opportunities and act as candidate
1013 directional anchor genes based on their direction of effect (network PES). Crucially, we
1014 observed a notable portion of cases with low overall PRS but at least one elevated PES, further
1015 supporting the biologically unique insights that may be gained from PES relative to an
1016 undifferentiated genome wide approach. Most of the scores also considered demonstrated at
1017 least nominal significance for association with either SZ or BIP, with PES like the *GRIN2A*
1018 network PES significantly enriched in both SZ and BIP even after conservatively correcting
1019 for the effect of a genome wide PRS. However, it should be noted that the training sets we used
1020 in this study were modestly powered, and larger training sets would be beneficial given that
1021 partitioned scores like PES will have smaller overall effect sizes than PRS. The relatively small
1022 trait-effect sizes of PES in terms of their phenotypic variance explained does not also
1023 necessarily preclude individual level relevance, particularly because PES like the *FADS1*
1024 network PES displayed strong correlations with traits relevant to that network such as measured
1025 lipids. Penalised regression was also applied to the construction of these network PES rather
1026 than just clumping and thresholding as was undertaken in previous PES studies [18–20]. These
1027 penalised regression PES did explain more variance in most of the networks considering the
1028 training set compared to clumping and thresholding, although selecting the appropriate
1029 constraint and tuning parameters when many combinations perform similarly is an ongoing
1030 challenge for such approaches. For example, the best performing *FES* network PES for BIP
1031 was derived through penalised regression, however, it only included five variants, which is not
1032 representative of the overall network. These issues highlight the need for future study in the
1033 construction of PES, particularly as popular genome wide approaches for PRS like LDpred and
1034 SBayesR would need to be methodologically adapted for a local gene-set or network
1035 implementation [15]. The proportion of phenotypic variance explained by PES derived using
1036 these networks could also be boosted in future by incorporating rare and structural variation,
1037 as well as re-weighting effect sizes informed by functional annotation.

1038

1039 While there is still work needed to confirm that the PES can effectively triage an individual's
1040 suitability for drug repurposing candidates in psychiatry and other disorders, this study
1041 represents a key methodological advance as it predicts the desired direction of effect needed to
1042 modulate a given target within the context of a PES. Previously, it was not clear whether
1043 agonism or antagonism of genes in the set would be clinically useful. Randomised placebo
1044 control trials of the PES approach would be a further step to demonstrate its utility and could

1045 involve repurposed drugs stratified by the relevant PES or more complex study designs, such
1046 as the multi-crossover “N-of-one” approach [57, 58]. In summary, we present a novel
1047 framework to inform precision drug repurposing in psychiatry that could account for individual
1048 level heterogeneity in genetic risk factors, and therefore, improve patient outcomes.

1049

1050 **DATA AND CODE AVAILABILITY**

1051 The analysis code for this study can be found in the following GitHub repository:
1052 https://github.com/Williamreay/Directional_anchor_gene_psychiatric_PES. Researchers can
1053 access the full UK Biobank data upon approval ([https://www.ukbiobank.ac.uk/enable-your-](https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access)
1054 [research/apply-for-access](https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access)). The ASRB cohort is available upon reasonable request and ethics
1055 approval (<https://www.neura.edu.au/discovery-portal/asrb/>). The remaining data are all
1056 publicly available, as outlined in the studies cited at relevant positions throughout the main
1057 text.

1058

1059 **ACKNOWLEDGEMENTS**

1060 This work was supported by a National Health and Medical Research Council (NHMRC) grant
1061 (1188493). M.J.C. is supported by an NHMRC Senior Research Fellowship (1121474), URL:
1062 <https://www.nhmrc.gov.au/>. The funders had no role in study design, data collection and
1063 analysis, decision to publish, or preparation of the paper. W.R.R was supported by an
1064 Australian Government Research Training Program Stipend. This research was conducted
1065 using the UK Biobank resource under the application 58432.

1066

1067

1068 **DISCLOSURES**

1069 W.R.R and M.J.C have filed a patent related to the pharmagenic enrichment score approach –
1070 WIPO Patent Application: WO/2020/237314. The remaining authors declare no competing
1071 financial interests.

1072

1073 **AUTHOR CONTRIBUTIONS**

1074 W.R.R. conceived and designed the study with input from M.J.C. W.R.R. performed the
1075 primary analyses. M.P.G assisted with the quality control and processing of the UK Biobank
1076 cohort. J.R.A performed the SNP imputation in the ASRB cohort. M.J.C, V.J.C, and M.J.G
1077 contributed the ASRB cohort, as well as to critical interpretation of the results and drafting of
1078 the manuscript.

1079

1080 REFERENCES

1081

- 1082 1. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, et al.
1083 Acute and longer-term outcomes in depressed outpatients requiring one or several
1084 treatment steps: a STAR*D report. *Am J Psychiatry*. 2006;163:1905–1917.
- 1085 2. Wimberley T, Støvring H, Sørensen HJ, Horsdal HT, MacCabe JH, Gasse C. Predictors
1086 of treatment resistance in patients with schizophrenia: a population-based cohort study.
1087 *Lancet Psychiatry*. 2016;3:358–366.
- 1088 3. Souery D, Oswald P, Massat I, Bailer U, Bollen J, Demyttenaere K, et al. Clinical
1089 factors associated with treatment resistance in major depressive disorder: results from a
1090 European multicenter study. *J Clin Psychiatry*. 2007;68:1062–1070.
- 1091 4. Hyman SE. Revolution stalled. *Sci Transl Med*. 2012;4:155cm11.
- 1092 5. Psychiatric drug discovery on the couch. *Nat Rev Drug Discov*. 2007;6:171–171.
- 1093 6. Pushpakom S, Iorio F, Eyers PA, Escott KJ, Hopper S, Wells A, et al. Drug
1094 repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov*.
1095 2019;18:41–58.
- 1096 7. Reay WR, Cairns MJ. Advancing the use of genome-wide association studies for drug
1097 repurposing. *Nat Rev Genet*. 2021. 23 July 2021. [https://doi.org/10.1038/s41576-021-](https://doi.org/10.1038/s41576-021-00387-z)
1098 [00387-z](https://doi.org/10.1038/s41576-021-00387-z).
- 1099 8. Fava M. The promise and challenges of drug repurposing in psychiatry. *World*
1100 *Psychiatry*. 2018;17:28–29.
- 1101 9. Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke S,
1102 Walters JT, O'Donovan MC. Mapping genomic loci prioritises genes and implicates
1103 synaptic biology in schizophrenia. *Genetic and Genomic Medicine*; 2020.
- 1104 10. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-
1105 wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet*.
1106 2019;51:793–803.
- 1107 11. Watson HJ, Yilmaz Z, Thornton LM, Hübel C, Coleman JRI, Gaspar HA, et al.
1108 Genome-wide association study identifies eight risk loci and implicates metabo-
1109 psychiatric origins for anorexia nervosa. *Nat Genet*. 2019;51:1207–1214.
- 1110 12. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, et al. Discovery
1111 of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder.
1112 *Nat Genet*. 2019;51:63–75.
- 1113 13. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al.
1114 Genome-wide association analyses identify 44 risk variants and refine the genetic
1115 architecture of major depression. *Nat Genet*. 2018;50:668–681.
- 1116 14. International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM,
1117 O'Donovan MC, et al. Common polygenic variation contributes to risk of schizophrenia
1118 and bipolar disorder. *Nature*. 2009;460:748–752.
- 1119 15. Ni G, Zeng J, Revez JA, Wang Y, Zheng Z, Ge T, et al. A Comparison of Ten
1120 Polygenic Score Methods for Psychiatric Disorders Applied Across Multiple Cohorts.
1121 *Biological Psychiatry*. 2021:S0006322321012634.
- 1122 16. Gaspar HA, Breen G. Drug enrichment and discovery from schizophrenia genome-wide
1123 association results: an analysis and visualisation approach. *Sci Rep*. 2017;7:12460.
- 1124 17. Reay WR, Cairns MJ. The role of the retinoids in schizophrenia: genomic and clinical
1125 perspectives. *Mol Psychiatry*. 2020;25:706–718.
- 1126 18. Reay WR, Atkins JR, Carr VJ, Green MJ, Cairns MJ. Pharmacological enrichment of
1127 polygenic risk for precision medicine in complex disorders. *Sci Rep*. 2020;10:879.

- 1128 19. Reay WR, El Shair SI, Geaghan MP, Riveros C, Holliday EG, McEvoy MA, et al.
1129 Genetic association and causal inference converge on hyperglycaemia as a modifiable
1130 factor to improve lung function. *ELife*. 2021;10:e63115.
- 1131 20. Reay WR, Geaghan MP, 23andMe Research Team, Cairns MJ. Genome-wide meta-
1132 analysis of pneumonia suggests a role for mucin biology and provides novel drug
1133 repurposing opportunities. *MedRxiv*. 2021.01.24.21250424 DOI
1134 10.1101/2021.01.24.21250424
- 1135 21. Reay WR, Kiltchewskij DJ, Geaghan MP, Atkins JR, Carr VJ, Green MJ, Cairns MJ.
1136 Genetic estimates of correlation and causality between blood-based biomarkers and
1137 psychiatric disorders. *Psychiatry and Clinical Psychology*; 2021.
- 1138 22. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, et al. Integrative
1139 approaches for large-scale transcriptome-wide association studies. *Nat Genet*.
1140 2016;48:245–252.
- 1141 23. Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, et al. Transcriptome-
1142 wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science*.
1143 2018;362.
- 1144 24. Wingo AP, Liu Y, Gerasimov ES, Gockley J, Logsdon BA, Duong DM, et al.
1145 Integrating human brain proteomes with genome-wide association data implicates new
1146 proteins in Alzheimer’s disease pathogenesis. *Nat Genet*. 2021;53:143–146.
- 1147 25. Zhang J, Dutta D, Köttgen A, Tin A, Schlosser P, Grams ME, et al. Large Bi-Ethnic
1148 Study of Plasma Proteome Leads to Comprehensive Mapping of *cis*-pQTL and Models
1149 for Proteome-wide Association Studies. *Genetics*; 2021.
- 1150 26. Mancuso N, Freund MK, Johnson R, Shi H, Kichaev G, Gusev A, et al. Probabilistic
1151 fine-mapping of transcriptome-wide association studies. *Nat Genet*. 2019;51:675–682.
- 1152 27. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al.
1153 Bayesian test for colocalisation between pairs of genetic association studies using
1154 summary statistics. *PLoS Genet*. 2014;10:e1004383.
- 1155 28. Zheng J, Haberland V, Baird D, Walker V, Haycock PC, Hurler MR, et al. Phenome-
1156 wide Mendelian randomization mapping the influence of the plasma proteome on
1157 complex diseases. *Nat Genet*. 2020. 7 September 2020. [https://doi.org/10.1038/s41588-](https://doi.org/10.1038/s41588-020-0682-6)
1158 [020-0682-6](https://doi.org/10.1038/s41588-020-0682-6).
- 1159 29. de Klein N, Tsai EA, Vochteloo M, Baird D, Huang Y, Chen C-Y, et al. Brain
1160 expression quantitative trait locus and network analysis reveals downstream effects and
1161 putative drivers for brain-related diseases. *Genomics*; 2021.
- 1162 30. Robins C, Liu Y, Fan W, Duong DM, Meigs J, Harerimana NV, et al. Genetic control of
1163 the human brain proteome. *Am J Hum Genet*. 2021;108:400–410.
- 1164 31. Vösa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, et al. Unraveling
1165 the polygenic architecture of complex traits using blood eQTL metaanalysis. *Genomics*;
1166 2018.
- 1167 32. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base
1168 platform supports systematic causal inference across the human phenome. *ELife*.
1169 2018;7:e34408.
- 1170 33. Freshour S, Kiwala S, Cotto KC, Coffman AC, McMichael JF, Song J, et al. Integration
1171 of the Drug-Gene Interaction Database (DGIdb) with open crowdsourcing efforts.
1172 *Bioinformatics*; 2020.
- 1173 34. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING
1174 v11: protein-protein association networks with increased coverage, supporting
1175 functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*.
1176 2019;47:D607–D613.

- 1177 35. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web
1178 server for functional enrichment analysis and conversions of gene lists (2019 update).
1179 Nucleic Acids Research. 2019;47:W191–W198.
- 1180 36. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set
1181 analysis of GWAS data. PLoS Comput Biol. 2015;11:e1004219.
- 1182 37. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale
1183 data. GigaScience. 2019;8:giz082.
- 1184 38. Mak TSH, Porsch RM, Choi SW, Zhou X, Sham PC. Polygenic scores via penalized
1185 regression on summary statistics. Genet Epidemiol. 2017;41:469–480.
- 1186 39. Loughland C, Draganic D, McCabe K, Richards J, Nasir A, Allen J, et al. Australian
1187 Schizophrenia Research Bank: a database of comprehensive clinical, endophenotypic
1188 and genetic data for aetiological studies of schizophrenia. Aust N Z J Psychiatry.
1189 2010;44:1029–1035.
- 1190 40. Lee SH, Goddard ME, Wray NR, Visscher PM. A better coefficient of determination for
1191 genetic profile analysis. Genet Epidemiol. 2012;36:214–224.
- 1192 41. Aschard H, Vilhjálmsson BJ, Joshi AD, Price AL, Kraft P. Adjusting for heritable
1193 covariates can bias effect estimates in genome-wide association studies. Am J Hum
1194 Genet. 2015;96:329–339.
- 1195 42. Martin J, Khramtsova EA, Goleva SB, Blokland GAM, Traglia M, Walters RK, et al.
1196 Examining Sex-Differentiated Genetic Effects Across Neuropsychiatric and Behavioral
1197 Traits. Biological Psychiatry. 2021;89:1127–1137.
- 1198 43. Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D,
1199 et al. Opportunities and challenges for transcriptome-wide association studies. Nat
1200 Genet. 2019;51:592–599.
- 1201 44. Lipton SA, Choi Y-B, Takahashi H, Zhang D, Li W, Godzik A, et al. Cysteine
1202 regulation of protein function--as exemplified by NMDA-receptor modulation. Trends
1203 Neurosci. 2002;25:474–480.
- 1204 45. Yawalkar R, Changotra H, Gupta GL. Protective influences of N-acetylcysteine against
1205 alcohol abstinence-induced depression by regulating biochemical and GRIN2A,
1206 GRIN2B gene expression of NMDA receptor signaling pathway in rats. Neurochem Int.
1207 2018;118:73–81.
- 1208 46. Neumeyer S, Hemani G, Zeggini E. Strengthening Causal Inference for Complex
1209 Disease Using Molecular Quantitative Trait Loci. Trends in Molecular Medicine.
1210 2020;26:232–241.
- 1211 47. Stanicka J, Rieger L, O'Shea S, Cox O, Coleman M, O'Flanagan C, et al. FES-related
1212 tyrosine kinase activates the insulin-like growth factor-1 receptor at sites of cell
1213 adhesion. Oncogene. 2018;37:3131–3150.
- 1214 48. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, et al. Aptamer-based
1215 multiplexed proteomic technology for biomarker discovery. PLoS One. 2010;5:e15004.
- 1216 49. Rohloff JC, Gelinis AD, Jarvis TC, Ochsner UA, Schneider DJ, Gold L, et al. Nucleic
1217 Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as
1218 Diagnostic and Therapeutic Agents. Mol Ther Nucleic Acids. 2014;3:e201.
- 1219 50. Javitt DC, Silipo G, Cienfuegos A, Shelley AM, Bark N, Park M, et al. Adjunctive high-
1220 dose glycine in the treatment of schizophrenia. Int J Neuropsychopharmacol.
1221 2001;4:385–391.
- 1222 51. Yolland CO, Hanratty D, Neill E, Rossell SL, Berk M, Dean OM, et al. Meta-analysis of
1223 randomised controlled trials with *N*-acetylcysteine in the treatment of schizophrenia.
1224 Aust N Z J Psychiatry. 2020;54:453–466.

- 1225 52. Nasir M, Bloch MH. Trim the fat: the role of omega-3 fatty acids in
1226 psychopharmacology. *Therapeutic Advances in Psychopharmacology*.
1227 2019;9:204512531986979.
- 1228 53. Moon AL, Haan N, Wilkinson LS, Thomas KL, Hall J. CACNA1C: Association With
1229 Psychiatric Disorders, Behavior, and Neurogenesis. *Schizophrenia Bulletin*.
1230 2018;44:958–965.
- 1231 54. Okorji UP, Velagapudi R, El-Bakoush A, Fiebich BL, Olajide OA. Antimalarial Drug
1232 Artemether Inhibits Neuroinflammation in BV2 Microglia Through Nrf2-Dependent
1233 Mechanisms. *Mol Neurobiol*. 2016;53:6426–6443.
- 1234 55. Chen W, Jin D, Shi Y, Zhang Y, Zhou H, Li G. The underlying mechanisms of
1235 lorlatinib penetration across the blood-brain barrier and the distribution characteristics
1236 of lorlatinib in the brain. *Cancer Med*. 2020;9:4350–4359.
- 1237 56. Hjelm BE, Rollins B, Mamdani F, Lauterborn JC, Kirov G, Lynch G, et al. Evidence of
1238 Mitochondrial Dysfunction within the Complex Genetic Etiology of Schizophrenia. *Mol*
1239 *Neuropsychiatry*. 2015;1:201–219.
- 1240 57. Schork NJ. Personalized medicine: Time for one-person trials. *Nature*. 2015;520:609–
1241 611.
- 1242 58. Percha B, Baskerville EB, Johnson M, Dudley JT, Zimmerman N. Designing Robust N-
1243 of-1 Studies for Precision Medicine: Simulation Study and Design Recommendations. *J*
1244 *Med Internet Res*. 2019;21:e12641.
- 1245