nature neuroscience

Analysis of genome-wide association data highlights candidates for drug repositioning in psychiatry

Hon-Cheong So[1,2], Carlos Kwan-Long Chau[1], Wan-To Chiu[3], Kin-Sang Ho[3], Cho-Pong Lo[3], Stephanie Ho-Yue Yim[4] & Pak-Chung Sham[5–8]

Knowledge of psychiatric disease genetics has advanced rapidly during the past decade with the advent of genome-wide association studies (GWAS). However, less progress has been made in harnessing these data to reveal new therapies. Here we propose a framework for drug repositioning by comparing transcriptomes imputed from GWAS data with drug-induced gene expression profiles from the Connectivity Map database and apply this approach to seven psychiatric disorders. We found a number of repositioning candidates, many supported by preclinical or clinical evidence. Repositioning candidates for a number of disorders were also significantly enriched for known psychiatric medications or therapies considered in clinical trials. For example, candidates for schizophrenia were enriched for antipsychotics, while those for bipolar disorder were enriched for both antipsychotics and antidepressants. These findings provide support for the usefulness of GWAS data in guiding drug discovery.

The last decade has witnessed a rapid growth of genotyping technologies, and GWAS have played a major role in unraveling the genetic bases of complex diseases. According to the latest GWAS catalog (http://www.ebi.ac.uk/gwas/), 3,020 GWAS studies have been performed as of 6 July 2017. It is clearly of great clinical and public interest to translate these findings into treatments for disease. Nevertheless, despite the large volume of GWAS literature, relatively few studies have been performed to discover novel drug candidates using GWAS data.

Psychiatric disorders inflict a significant burden on health globally, and the current treatment strategies are far from perfect. Despite the heavy health burden and increased awareness of mental health, drug discovery in psychiatry has largely been stagnant[1]. As argued by Hyman[1], the basic mechanisms of most antidepressants and antipsychotics, the most widely used drugs used in psychiatry, are similar to those of their prototypes, imipramine and chlorpromazine, which were discovered in the 1950s. On the other hand, in recent years GWAS studies have greatly advanced our knowledge of the genetic bases of many psychiatric disorders. A recent review[2] nicely summarizes the technical and analytic issues in translating GWAS findings into new therapeutics, which will be an important aim of the phase 3 study being conducted by the Psychiatric Genomics Consortium (PGC). Taking advantage of the rapid developments in genomics, we propose a new framework for identifying drug candidates based on GWAS results, and we have applied the method to a variety of psychiatric disorders.

Here we focus on drug repositioning, that is, finding new indications for existing drugs. As conventional drug development is an expensive and lengthy process, repositioning serves as a useful strategy to hasten the development cycle[3]. It is worth noting that, while we use the term 'drug repositioning' throughout this paper, the method is applicable to any chemical with measured expression profiles.

A few previous studies have investigated the use of GWAS data in drug repositioning. The most intuitive approach is to study whether the top genes identified in GWAS can serve as drug targets, for example in Sanseau *et al.*[4] and Lencz and Malhotra[5]. In another study, Cao *et al.*[6] considered interacting partners of GWAS hits to discover new drug targets. It has also been reported that the proportion of drugs with support from GWAS increases along the development pipeline[7].

While the approach of finding overlap between top GWAS hits and known drug target genes is useful, it has a number of limitations. First, many of the top GWAS genes may not be easily targeted by a drug, as shown by Cao and Moult[6]. In addition, many of the GWAS top single-nucleotide polymorphisms are within noncoding regions and do not encode a drug-targeted protein. The above approach might also miss 'multitarget' drugs. It has been argued that, as complex diseases involve the interplay of multiple genetic and environmental factors, they may be more easily managed by modulating multiple instead of single targets[8]. Lastly, as previous studies have mainly focused on the most significant hits, they have ignored the contribution of genetic variants with smaller effect sizes. As shown in polygenic score analyses

[1]School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong, China. [2]KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research of Common Diseases, Kunming Zoology Institute of Zoology and The Chinese University of Hong Kong, China. [3]Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong, China. [4]Univeristy of Exeter Medical School, Exeter, UK. [5]Department of Psychiatry, University of Hong Kong, Pokfulam, Hong Kong, China. [6]Centre for Genomic Sciences, University of Hong Kong, Pokfulam, Hong Kong, China. [7]State Key Laboratory for Cognitive and Brain Sciences, University of Hong Kong, Pokfulam, Hong Kong, China. [8]Centre for Reproduction, Development and Growth, University of Hong Kong, Pokfulam, Hong Kong, China. Correspondence should be addressed to H.-C.S. (hcso@cuhk.edu.hk).

Received 11 January; accepted 6 July; published online 14 August 2017; doi:10.1038/nn.4618

of many complex traits, variants with lower significance levels also often contribute to disease risks.

With the aforementioned limitations in mind, we developed a new strategy for drug repositioning by imputing gene expression profiles from GWAS summary statistics and comparing them with drug-induced expression changes. Analysis of the transcriptome of drugs versus diseases is an established approach for drug repositioning and has previously been successfully applied for complex diseases[9,10]. For example, by examining drugs in the Connectivity Map (Cmap)[11] that showed opposite patterns of expression to diseases, Sirota *et al.*[10] identified potential drugs for repositioning and experimentally validated a prediction about cimetidine for the treatment of lung adenocarcinoma. Using a similar method, Dudley *et al.*[9] identified topiramate as a novel treatment for inflammatory bowel disease and validated it in an animal model. Other studies (e.g., ref. 12) have also shown the potential of this approach in repositioning.

Building on this repositioning strategy, we proposed a new approach using imputed transcriptomes from GWAS instead of expression data from microarray or RNA-sequencing studies. This approach has several advantages. First, patients from expression studies are often medicated. This is particularly relevant for studies in psychiatry, as brain tissues can only come from postmortem patient samples. A history of psychiatric medications might confound the results, as our aim was to compare the expression patterns of diseases and those of drugs. Imputed transcriptomes, on the other hand, are not altered by medications or other environmental confounders. Second, current GWAS samples are usually orders of magnitude larger than expression studies (often 10^4 or more), and GWAS summary statistics are widely available. Also, for many diseases including as psychiatric disorders, the tissues of interest are not easily accessible. On the other hand, expression profiles can be readily imputed for a large number of tissues from GWAS data using appropriate statistical models.

RESULTS

We employed MetaXcan[13] to impute expression profiles for 10 brain regions from GWAS variants. The program applied a statistical model learned from a reference transcriptome data set to predict expressions in new data. We downloaded eight sets of GWAS summary statistics corresponding to seven psychiatric disorders, including schizophrenia (SCZ), major depressive disorder (MDD), bipolar disorder (BD), Alzheimer's disease (AD), anxiety disorders (ANX), autistic spectrum disorders (ASD) and attention deficit hyperactivity disorder (ADHD; Online Methods). Drug-induced transcriptomes were derived from the Cmap database version 2. We compared the expression profiles of drugs to those of diseases and looked for reversed expression patterns using various statistical methods. We manually curated the top 15 drugs (representing the top ~0.45% of all instances) identified for each disorder and brain region. In addition to manually inspecting the top candidates, to validate our approach, we also tested for enrichments of drugs that are (i) indicated for each disorder as documented in the Anatomical Therapeutic Chemical (ATC) classification system or the MEDication Indication high-precision subset (MEDI-HPS) database or (ii) included in clinical trials (as listed on https://clinicaltrials.gov/). The drug-set enrichment results were combined across brain regions by meta-analysis.

The sample sizes of the GWAS data sets we used are listed in **Supplementary Table 1**. The top 15 drug repositioned for each disorder and brain region (with manual curation of drug descriptions and potential therapeutic relationship to the disorder) are presented in full in **Supplementary Tables 2–9**. The approach for curation is described in Online Methods and the **Supplementary Note**. It should, however, be noted that owing to the very large number of drug–disorder pairs, we did not intend to perform a systematic review for each pair. Selected drugs within the top lists are presented and discussed below.

Repositioning hits for different psychiatric disorders

Schizophrenia. Table 1 shows selected repositioning hits for SCZ and BD. For SCZ, note that we identified a number of known antipsychotics such as thioproperazine, droperidol, trifluprom azine, thiethylperazine, spiperone and pimozide as top hits. It is worth noting that our repositioning method is blind to any knowledge about existing psychiatric drugs or known drug targets. The results provide further evidence for the role of the dopaminergic system in SCZ treatment. Some drugs listed among the top 15, such as idazoxan, have been shown to improve negative symptoms in clinical studies[14]. The antidepressant paroxetine was also tested in a double-blind clinical trial and shown to be efficacious in ameliorating negative symptoms in SCZ[15]. There are other candidates with preliminary support by preclinical or clinical studies with diverse mechanisms, such as the serotonin and dopamine antagonist metitepine[16] and the Na-K-Cl cotransporter-1 inhibitor bumetanide[17]. It is also noteworthy that a relatively large proportion of drugs with stronger support in literature were derived from comparison with gene expressions in the frontal cortex, a brain region strongly implicated in SCZ.

Bipolar disorder. We found a number of antipsychotics among the top list for BD (**Table 2**). Antipsychotics are well known to be effective for BD overall and for the associated psychotic symptoms. Our analyses also revealed other drugs with known or potential antidepressant effects, such as imipramine, metyrapone[18] and ketoconazole[19]. The latter two drugs are believed to exert antidepressant effects by reducing cortisol levels. Antidepressants are often used in the treatment of BD patients. While there are controversies regarding their use, antidepressants are included as valid treatment options in current guidelines, especially in bipolar II patients or when used with a mood stabilizer[20]. Notably, a few nonsteroidal anti-inflammatory drugs (NSAIDs) were also on the top 15 list, such as aspirin and a cyclooxygenase-2 (COX-2) inhibitor, SC-58125, which is supported by the neuroinflammatory hypothesis of BD[21]. In line with the observation of raised cardiovascular risks in BD patients and the possible shared pathophysiology between these disorders[22], simvastatin and metformin were also among the top 15. Notably, the relationship between statins and mood disorders has attracted much research attention recently. Some studies have reported a protective effect for depression[23,24]. While drugs like statins are widely prescribed, it is possible that their potential therapeutic effects remain unnoticed for a variety of reasons. For example, the effects may only be obvious at certain dosages, when combined with other drugs or in certain disease subtypes; also, many patients taking statins may not have any active psychiatric disorders. Further investigations are warranted to delineate the exact effects of common medications like statins on psychiatric traits.

Major depressive disorder. We used two sets of GWAS data for MDD in our analyses (**Table 3**), one from PGC (MDD-PGC) and one from the CONVERGE Consortium (MDD-CONVERGE; Online Methods), with the latter focusing on recurrent melancholic depression. We observed that in results based on MDD-PGC, fluoxetine, a widely used selective serotonin reuptake inhibitor (SSRI), was among the top results. We also observed quite a few antipsychotics on the list, such as sulpiride, promazine, perphenazine and loxapine.

Table 1 Selected repositioning hits for SCZ (ordered by *q* values)

Drug	Cell line	Brain region	Rank in brain region	*P* value	*q* value	Brief description
Tetrahydroalstonine	PC3	COR	1	7.19×10^{-5}	0.155	Indole alkaloid; antipsychotic properties in rodents; clinically used in Nigeria
Hesperidin	MCF7	COR	2	8.91×10^{-5}	0.155	Flavanones; antioxidant
Droperidol	HL60	CER	4	5.12×10^{-4}	0.445	D2 antagonist; antipsychotic properties
Triflupromazine	PC3	COR	5	8.40×10^{-4}	0.487	First-generation antipsychotic
Bumetanide	MCF7	COR	6	8.57×10^{-4}	0.487	Na-K-Cl cotransporter-1 inhibitor; improved hallucinations in an RCT
Meclofenamic acid	MCF7	FCOR	9	2.24×10^{-3}	0.759	NSAID; under clinical trial for cognitive symptoms
Spiradoline	PC3	FCOR	10	2.24×10^{-3}	0.759	Selective κ-opioid agonist; effective in animal studies
Metitepine	PC3	PUT	1	5.75×10^{-6}	0.787	5-HT- and dopamine-receptor antagonist with possible antipsychotic properties
Spiperone	PC3	NUC	8	3.37×10^{-3}	0.854	First-generation antipsychotic
Pimozide	PC3	NUC	13	4.14×10^{-3}	0.854	First-generation antipsychotic
Bromopride	MCF7	HYP	4	1.46×10^{-3}	0.860	Dopamine antagonist
Scopolamine	HL60	CAU	3	1.43×10^{-3}	0.883	Muscarinic antagonist; antidepressant properties
Ranitidine	MCF7	CAU	12	3.36×10^{-3}	0.883	H2 antagonist; shown to reduce negative symptoms in clinical study
Paroxetine	PC3	CAU	14	4.59×10^{-3}	0.883	SSRI; shown to improve negative symptoms in clinical study
Tremorine	PC3	AC	7	3.19×10^{-3}	0.893	Affect amine levels in brain; a closely related drug, oxotremorine, was effective for SCZ in animal studies
Thioproperazine	PC3	CEH	7	2.56×10^{-3}	0.899	First-generation antipsychotic
Thiethylperazine	PC3	FCOR	13	3.40×10^{-3}	0.904	First-generation antipsychotic
Idazoxan	PC3	FOCR	14	3.64×10^{-3}	0.904	α-2 antagonist; clinical studies showed potential in improving negative symptoms

'Rank' refers to the rank within each brain region for the studied disorder; see **Supplementary Table 2** for references. AC, anterior cingulate cortex (Brodmann area 24); CAU, caudate (of basal ganglia); CER, cerebellum; CEH, cerebellar hemisphere; COR, cortex; FCOR, frontal cortex (Brodmann area 9); HIP, hippocampus; HYP, hypothalamus; NUC, nucleus accumbens (of basal ganglia); PUT, putamen (of basal ganglia); RCT, randomized controlled trial; PC3, HL60 and MCF7 are the three cell lines used in Cmap; PC3, human prostate cancer cell line; HL60, human promyelocytic leukemia cell line; MCF7, human breast adenocarcinoma cell line.

Other repositioning hits were more diverse in their mechanisms, such as phosphodiesterase inhibitors (papaverine[25]), muscarinic antagonists (scopolamine[26]) and cortisol-lowering agents (ketoconazole[19]).

The repositioning results from the MDD-CONVERGE study included a known SSRI (zimelidine), a norepinephrine–dopamine reuptake inhibitor (nomifensine) and a monoamine oxidase (MAO) inhibitor (isocarboxazid) ranked among the top 15, although the former two drugs have been withdrawn due to other, unrelated, adverse effects. Two antipsychotics, sulpiride and risperidone, were also on the list[27]. Notably, we found three drugs with actions on glutamatergic transmission among the top hits. Arcaine sulfate and ifenprodil are NMDA antagonists, and cycloserine is a partial agonist at the glycine site of the NMDA glutamate receptor. Ifenprodil showed antidepressant-like effects in a mouse model[28], while cycloserine was effective as an add-on therapy in a randomized controlled trial (RCT)[29].

Anxiety disorders. For ANX, we found the SSRI paroxetine and the tricyclic antidepressant protriptyline among the top 15. Another drug active on the serotonergic system, pirenperone, acts as a 5-HT2 antagonist and was shown to have anxiolytic actions in a small clinical study[30]. Other drugs with preliminary support from animal or clinical studies include bumetanide, a loop diuretic that also affects $GABA_A$ signaling[31]; ivermectin, an antiparasitic and GABA agonist[32]; and kawain, a kavalactone with anxiolytic properties shown in a number of clinical trials[33].

Alzheimer disease. AD is a relatively intense research area and we found numerous drugs with at least some support by preclinical or clinical studies. Some of the drugs ranked among the top 15 have been tested in clinical trials. Naftidrofuryl is a vasodilatory agent reported to be effective for functional outcomes, mood and cognitive function in a meta-analysis[34]. Vinpocetine is an alkaloid with possible benefits according to a meta-analysis of three RCTs[35]. Notably, quite a number of NSAIDs appeared on our top list of drugs, such as meclofenamic acid, ketorolac, celecoxib, naproxen and acemetacin. NSAIDs have been tested in clinical studies for possible prevention or treatment of AD, although the results were inconsistent and further investigations are required[36].

ADHD. A few drugs on the top list have been tested in clinical trials. The anticonvulsant carbamazepine was analyzed in a meta-analysis, which reported preliminary evidence that it can be used to treat ADHD[37]. The alkaloid lobeline may improve working memory in adult ADHD[38]. Tranylcypromine (a monoamine oxidase inhibitor) was effective in clinical trials, although its side effects need to be considered[39].

ASD. A few drugs at the top of the list are worth mentioning. Risperidone is one of the two FDA-approved medications for treating irritability in ASD. Two drugs, pentoxifylline (a phosphodiesterase inhibitor) and amantadine (an NMDA antagonist), have been tested in clinical trials of ASD as a combination treatment with risperidone. Both showed effectiveness in ameliorating behavioral problems[40,41]. Ribavirin is another interesting candidate for potential repurposing. ASD was recently shown to be associated with eIF4E overexpression, which in turn leads to excessive translation of neuroligins[42]. Ribavirin was found to be an inhibitor of eIF4E[43] and hence may serve as a potential treatment. Ribavirin was also listed in a patent for ASD treatment (http://www.google.ch/patents/US5008251).

Drug-set enrichment analyses

Enrichment for drugs listed in the ATC classification system. First we considered the enrichment test results from drugs listed in the ATC classification system. As shown in **Table 4**, antipsychotics were enriched in the repositioning results for SCZ (lowest *P* value across four tests = 4.69×10^{-6}) and BD (lowest $P = 2.26 \times 10^{-7}$). Notably, antipsychotics were also enriched among the repositioning hits (albeit less strongly) for MDD (lowest $P = 0.0285$), AD (lowest $P = 0.0256$) and ANX (lowest $P = 0.0054$). We also observed antidepressants and anxiolytics to be enriched in drugs repositioned for BD (lowest $P = 1.17 \times 10^{-5}$). In addition we found a trend toward significance for AD (lowest $P = 0.0507$). When all psychiatric medications were combined as a drug-set, evidence of enrichment was found for SCZ, BD, AD and ANX.

Enrichment for drugs listed in MEDI-HPS. The enrichment test results using MEDI-HPS drugs were in general consistent with those

Table 2 Selected repositioning hits for BD (ordered by *q* values)

Drug	Cell line	Brain region	Rank in brain region	*P* value	*q* value	Brief description
Metitepine	PC3	PUT	1	5.75×10^{-6}	0.019	5-HT- and dopamine-receptor antagonist with possible antipsychotic properties
Norcyclobenzaprine	PC3	CAU	1	3.31×10^{-4}	0.423	Structurally very similar to the antidepressant amitriptyline
Acetylsalicylsalicylic acid	PC3	CAU	2	3.65×10^{-4}	0.423	Closely related to aspirin, which may be useful for BD
Metyrapone	HL60	FCOR	1	1.58×10^{-4}	0.485	Cortisol synthesis inhibitor; RCT showed effects in treatment-resistant MDD
Thioridazine	HL60	CAU	6	1.03×10^{-3}	0.598	First generation antipsychotic
Acetylsalicylic acid	HL60	CEH	9	2.00×10^{-3}	0.712	Aspirin; clinical studies showed potential in BD treatment
Simvastatin	MCF7	COR	6	5.59×10^{-3}	0.756	Augmentation with lithium may be effective for BD
Prestwick-689 (androsterone)	MCF7	COR	13	2.53×10^{-3}	0.756	DHEA (precursor of this drug) may have antidepressant effects
Mesoridazine	PC3	NUC	6	1.99×10^{-3}	0.759	First generation antipsychotic
Alpha yohimbine	MCF7	HIP	1	3.82×10^{-4}	0.797	α-2 antagonist; an RCT showed yohimbine hastened antidepressant response if combined with fluoxetine
Trifluoperazine	MCF7	HIP	5	1.22×10^{-3}	0.797	Antipsychotic
Molindone	MCF7	HIP	12	3.72×10^{-3}	0.797	Antipsychotic
Metformin	MCF7	HIP	13	3.90×10^{-3}	0.797	Antidiabetic; being tested in clinical trial for refractory BD
Mexiletine	PC3	CER	6	2.25×10^{-3}	0.931	Class IB antiarrhythmic; may be useful for treatment-resistant BD
Dextromethorphan	HL60	CER	11	4.23×10^{-3}	0.931	Morphinan class; may be effective for bipolar depression
Imipramine	HL60	CER	14	5.56×10^{-3}	0.931	Tricyclic antidepressant
SC-58125	MCF7	HYP	4	1.12×10^{-3}	0.946	COX-2 inhibitor; a closely related drug, celecoxib, showed antidepressive actions in BD
Ketoconazole	PC3	HYP	7	2.27×10^{-3}	0.968	Case reports and open studies suggested efficacy for bipolar depression
Thioproperazine	MCF7	HYP	10	3.54×10^{-3}	0.968	First generation antipsychotic

'Rank' refers to the rank within each brain region for the studied disorder; see **Supplementary Table 3** for corresponding references. See **Table 1** for abbreviations. DHEA, dehydroepiandrosterone.

derived from the ATC drugs (**Table 5**). Again we observed that antipsychotics were enriched in the repositioning hits for SCZ (lowest $P = 2.24 \times 10^{-9}$). For the rest of the analyses with the antipsychotics drug-set, we found enrichment for BD and to a lesser extent for MDD; a nominally significant result was also observed for AD (**Table 5**).

When the drug-set was limited to medications for depression or anxiety, a significant enrichment was observed for SCZ (lowest $P = 1.52 \times 10^{-3}$). We also observed nominally significant results for BD (lowest $P = 0.0143$) and AD (lowest $P = 0.0458$), although they did not pass the false-discovery rate threshold of 0.05. When all psychiatric drugs were included, enrichment was found for SCZ, BD and AD.

Enrichment for drugs listed at ClinicalTrials.gov. We tested for enrichment for drugs listed on ClinicalTrials.gov for each of the corresponding disorders (**Table 6**). (We did not pursue tests of drugs across diagnoses as drugs listed in on this website are less certain and less well-defined as belonging to a 'drug class' compared to the previous two sources.) Nominally significant enrichments were observed for SCZ, BD, MDD and ANX (**Table 6**), though they did not pass the primary false-discovery rate threshold at 0.05.

DISCUSSION

In this study we developed a framework for drug repositioning by linking two apparently disparate sources of information, GWAS-imputed transcriptome profiles and drug-induced changes in gene expression (Cmap). We applied the methodology to seven psychiatric disorders and identified a number of interesting candidates for repositioning, many of which are supported by animal or clinical studies. The drug-set enrichment analyses further lend support to the validity of our approach.

There are a number of advantages of our repositioning framework. First, this approach is largely 'hypothesis-free' in that it does not assume any knowledge about known drug targets or drug–disease relationships. As a result the method may be able to find drugs of different mechanisms from the known treatments. This method may be particularly useful for diseases with few known treatments or for which existing treatments are highly similar in their mechanisms. A related advantage is that it can be applied to any chemicals as long as the expression profile is available, such as traditional Chinese medicine (TCM) products with no known drug targets or drugs shelved after previous unsuccessful trials. As described above, another advantage when compared to conventional connectivity mapping is that GWAS-imputed transcriptomes are relatively immune to confounding by medication effects. In addition, only GWAS summary statistics are required as input for imputation, which obviates the difficulties in obtaining raw genotype data and makes the approach easily applicable to a wide variety of traits. The current method is also intuitive and computationally simple to implement. Moreover, we have considered all genetic variants instead of just the most significant hits in our repurposing pipeline. One alternative for inclusion of subthreshold associations is to employ gene-set analysis to look for over-representation of genes acting as drug targets[44], but here we propose a different approach, which does not rely on knowledge of known drug targets and takes into account the direction of genetic associations.

It is encouraging to observe that our repositioning results are, in general, supported by the drug-set enrichment analyses. In particular, antipsychotics, which are known to treat SCZ and BD, are also enriched in the repositioning results of these two disorders. In a recent study, Gasper *et al.*[45] employed a pathway analysis approach and revealed that SCZ GWAS results are increasingly enriched for target genes of antipsychotics and other psychiatric medications as sample size increases. In a similar vein, Ruderfer *et al.*[46] reported that SCZ risk genes from GWAS and exome sequencing are associated with targets of antipsychotics. While we have used a different approach here that integrated GWAS with drug expression data, the findings of all these studies are highly concordant, providing converging evidence for SCZ GWAS data in guiding drug development. Similarly, antidepressants are an indicated treatment option for BD[20] and were also significantly enriched for this disorder. For MDD (using the PGC data) and ANX, while we did not observe enrichment in ATC or MEDI-HPS drug-sets, there was some evidence that our approach preferentially picked up drugs included in clinical trials. These results suggest that GWAS results contain useful information for drug discovery or repositioning and are in line with the conclusion of a previous study, which found that drugs supported by human genomic data increase along the development pipeline[7].

Table 3 Selected repositioning hits for MDD (ordered by *q* values)

Drug	Cell line	Brain region	Rank	*P* value	*q* value	Brief description
MDD-PGC						
Perphenazine	MCF7	HYP	1	2.59×10^{-5}	0.090	Typical antipsychotic with possible antidepressant properties
Papaverine	HL60	AC	3	2.93×10^{-4}	0.340	Cyclic nucleotide phosphodiesterase inhibitor; shown to be useful for depression in a case report
Fluoxetine	MCF7	CAU	3	6.87×10^{-4}	0.535	SSRI antidepressant
Sanguinarine	HL60	CER	1	3.59×10^{-4}	0.568	A selective mitogen-activated protein kinase phosphatase-1 (Mkp-1) inhibitor; showed antidepressant effects in rats
Vitexin	PC3	COR	7	2.00×10^{-3}	0.596	Flavone glucoside; antidepressant properties shown in mice
Thioperamide	MCF7	CER	4	8.80×10^{-4}	0.625	Histamine H3- and H4-receptor antagonist; antidepressant effects in rats
Palmatine	HL60	HIP	10	2.06×10^{-3}	0.633	An alkaloid in plants, shown to have antidepressive properties in mice
Pyridoxine	PC3	NUC	5	1.07×10^{-3}	0.646	Vitamin B6; may have antidepressive effects in premenopausal women
Ketoconazole	PC3	NUC	14	2.75×10^{-3}	0.651	Antidepressant properties suggested in case series, possibly via cortisol-lowering effects
Pregnenolone	PC3	CEH	4	1.09×10^{-3}	0.722	Endogenous steroid; shown to be useful in bipolar depression in RCT
Loxapine	HL60	HYP	7	1.94×10^{-3}	0.728	Typical antipsychotic with possible antidepressant properties
Piroxicam	HL60	PUT	10	2.46×10^{-3}	0.787	NSAID with possible antidepressant actions shown in mice
Scopolamine N-oxide	HL60	CEH	13	4.44×10^{-3}	0.887	Antimuscarinic agent that showed antidepressive effects in at least two RCTs
Pirlindole	MCF7	FCOR	9	2.85×10^{-3}	0.898	A reversible inhibitor of monoamine oxidase A
Sulpiride	PC3	CAU	12	4.46×10^{-3}	0.934	Antipsychotic with potential antidepressant properties
Promazine	MCF7	CAU	15	4.89×10^{-3}	0.934	Antipsychotic with potential antidepressant properties
MDD-CONVERGE						
Bromocriptine	PC3	FCOR	1	3.16×10^{-5}	0.110	Dopamine agonist; antidepressant effects in a small open study
Isocarboxazid	HL60	FOCR	5	1.00×10^{-3}	0.581	Monoamine oxidase inhibitor
Arcaine	HL60	COR	2	3.91×10^{-4}	0.640	NMDA antagonist; this drug class was shown to produce antidepressive effects in animal studies and RCT
Pioglitazone	MCF7	FOCR	13	2.95×10^{-3}	0.645	Antidepressant properties when combined with citalopram in an RCT
Idazoxan	PC3	CAU	5	1.10×10^{-3}	0.679	α–2 antagonist; clinical studies showed effects in bipolar depression
Risperidone	HL60	HYP	7	2.25×10^{-3}	0.799	Atypical antipsychotic useful for augmentation as shown in RCT
Sulpiride	HL60	HIP	9	2.34×10^{-3}	0.815	Antipsychotic with potential antidepressant properties
Kawain	PC3	AC	2	6.90×10^{-4}	0.897	Effective for treatment of anxiety shown in RCTs
Doxycycline	PC3	AC	6	1.79×10^{-3}	0.897	Antidepressant properties shown in mice
Nomifensine	HL60	CER	4	1.56×10^{-3}	0.912	A norepinephrine–dopamine reuptake inhibitor; antidepressant
Zimelidine	PC3	COR	6	1.94×10^{-3}	0.918	SSRI antidepressant
Cycloserine	HL60	PUT	3	9.86×10^{-4}	0.956	Acts on NMDA receptor; clinical trial showed benefits in MDD
Ifenprodil	HL60	PUT	4	1.18×10^{-3}	0.956	NMDA antagonist; found to potentiate the effects of SSRI and TCA in mice

'Rank' refers to the rank within each brain region for the studied disorder; see **Supplementary Tables 4** and **5** for references. See **Table 1** for abbreviations of brain regions. TCA, tricyclic antidepressant.

Notably, our analyses also revealed possible enrichment of antipsychotics for MDD, ANX and AD. Antipsychotics have long been used for treatment of depression, especially in more severe cases with psychotic symptoms. A recent meta-analysis also demonstrated the efficacy of a number of atypical antipsychotics as adjunctive treatment in depression[27]. Given the high comorbidity of depression and anxiety and their possibly shared pathophysiology, it is not unexpected that antipsychotics may be useful for ANX as well. Antipsychotics are not infrequently prescribed for ANX, although their safety and efficacy warrant further investigation[47]. In a similar vein, we observed that medications for depression and anxiety were over-represented in the top drug lists of SCZ and AD (with nominal significance for AD; **Table 5**). Psychotic and depressive symptoms are very commonly seen in AD patients, hence the enrichment for antipsychotics and antidepressants are expected. Antidepressants have been tested in clinical trials for SCZ, especially for negative symptoms, although further studies are required in this area[48]. To our knowledge, this is the first study to demonstrate a genetic basis for the potential effectiveness of psychiatric drugs across diagnoses based on GWAS data.

We observed that for some psychiatric disorders, either there was no significant enrichment of known drug indications or all significant enrichments came from other classes of drugs. There are a few possible explanations. First, the sample size may not have been large enough to detect enrichment of known drugs. Among the eight data sets, the SCZ GWAS was the largest, including almost 80,000 samples. For the other traits, the sample sizes range from 5,000 to 20,000. Limited sample sizes imply that some true associations may not be detected and the imputed transcriptome may be less accurate, which will affect the ability to find drugs with matched (reversed) expression profiles. For any high-throughput studies with limited sample sizes, it is possible for some signals (here, known drug indications) to be 'missed', although other associations may be detected. Analogously, earlier SCZ GWAS with smaller sample sizes did not detect the *DRD2* locus; with accumulation of samples, the latest GWAS meta-analysis did confirm *DRD2* as a susceptibility gene[49]. Hence limited sample sizes may explain, for instance, the enrichment of antipsychotics but not antidepressants for MDD.

Second, many psychiatric disorders are known to be heterogeneous. MDD, ANX and AD are relatively prevalent and may display a wider range of heterogeneity in clinical manifestations and underlying pathophysiologies than rarer disorders. The heterogeneity impairs study power and implies that a specific drug or drug class may only be effective for a certain group of patients. Third, there are limited available treatments for some disorders, especially AD, ASD and ADHD, which makes the detection of these known drug indications difficult.

In addition to the above, we observed that enrichment analyses were mostly negative for the MDD-CONVERGE dataset. One possible explanation is that the Genotype–Tissue Expression (GTEx) dataset, on which the transcriptome imputation is based, is not

Table 4 Drug enrichment analyses *P* values with drug-sets defined in the ATC classification system

Disorder	Self (Fisher)	Compet (Fisher)	Self (min*P*)	Compet (min*P*)
Enrichment for ATC antipsychotics				
SCZ	**6.78×10^{-6}**	**4.69×10^{-6}**	**0.0011**	**0.0010**
BD	**4.14×10^{-7}**	**2.26×10^{-7}**	**0.0031**	**0.0025**
MDD-PGC	0.1634	0.1532	*0.0322*	**0.0285**
MDD-CONVERGE	0.4484	0.4231	0.6674	0.6546
AD	0.6849	0.6684	*0.0294*	**0.0256**
ANX	*0.0249*	*0.0241*	**0.0058**	**0.0054**
ADHD	0.9215	0.9209	0.7887	0.7939
ASD	0.9956	0.9949	0.3816	0.3523
Enrichment for ATC antidepressants or anxiolytics				
SCZ	0.1141	0.1075	0.1301	0.1264
BD	**1.64×10^{-5}**	**1.17×10^{-5}**	**0.0055**	**0.0050**
MDD-PGC	0.9911	0.9908	0.9448	0.9420
MDD-CONVERGE	0.7199	0.7016	0.3904	0.3693
AD	0.3475	0.3323	0.0564	0.0507
ANX	0.1941	0.1953	0.2248	0.2173
ADHD	0.9904	0.9904	0.9654	0.9641
ASD	0.8343	0.8400	0.7034	0.6884
Enrichment for all ATC psychiatric drugs				
SCZ	**8.20×10^{-5}**	**4.22×10^{-5}**	**0.0160**	**0.0130**
BD	**5.70×10^{-9}**	**1.20×10^{-9}**	**0.0033**	**0.0022**
MDD-PGC	0.7234	0.6935	0.1738	0.1482
MDD-CONVERGE	0.6548	0.6130	0.6718	0.6329
AD	0.1695	0.1360	**0.0022**	**0.0013**
ANX	**0.0061**	**0.0048**	**0.0129**	**0.0110**
ADHD	0.9981	0.9982	0.8358	0.8367
ASD	0.9988	0.9988	0.9542	0.9406

Self (Fisher): self-contained test (one-sample *t* test) combined across brain regions by Fisher's method; Compet (Fisher), competitive test (two-sample one-sided *t* test) combined across brain regions by Fisher's method; Self (min*P*), self-contained test (one-sample one-sided *t* test) combined across brain regions by Tippett's minimum *P* method; Compet (min*P*), competitive test (two-sample *t* test) combined across brain regions by Tippett's minimum *P* method. Test results with $P < 0.05$ and $q < 0.05$ are in bold; results with $P < 0.05$ and $q < 0.2$ are in italics. Full lists of *q* values are presented in **Supplementary Tables 10–12**.

Table 5 Drug enrichment analyses *P* values with drug-sets defined by MEDI-HPS

Disorder	Self (Fisher)	Compet (Fisher)	Self (min*P*)	Compet (min*P*)
Enrichment for indicated drugs of each disorder				
SCZ	**4.46×10^{-9}**	**2.24×10^{-9}**	*0.0197*	*0.0170*
BD	0.9962	0.9960	0.9982	0.9982
MDD-PGC	0.4724	0.4545	0.4525	0.4210
MDD-CONVERGE	0.5120	0.4787	0.7166	0.6945
AD	0.5686	0.5646	0.1148	0.1140
ANX	0.7188	0.7224	0.6363	0.6195
ADHD	0.6374	0.6355	0.4404	0.4359
ASD	NA	NA	NA	NA
Enrichment for drugs for schizophrenia				
SCZ	**4.46×10^{-9}**	**2.24×10^{-9}**	*0.0197*	*0.0170*
BD	**5.82×10^{-3}**	**4.46×10^{-3}**	0.0950	0.0890
MDD-PGC	0.0850	0.0795	*0.0293*	*0.0290*
MDD-CONVERGE	0.4266	0.3933	0.2645	0.2321
AD	0.3318	0.3096	0.0546	*0.0479*
ANX	0.2963	0.2949	0.2099	0.2061
ADHD	0.6671	0.6578	0.2597	0.2419
ASD	0.9311	0.9235	0.1326	0.1149
Enrichment for drugs for depression or anxiety				
SCZ	**1.92×10^{-3}**	**1.52×10^{-3}**	*0.0316*	*0.0282*
BD	*0.0173*	*0.0143*	*0.0214*	*0.0189*
MDD-PGC	0.4724	0.4545	0.4525	0.4210
MDD-CONVERGE	0.5120	0.4787	0.7166	0.6945
AD	0.3293	0.3104	0.0530	*0.0458*
ANX	0.7188	0.7224	0.6363	0.6195
ADHD	0.9975	0.9975	0.7589	0.7435
ASD	0.9880	0.9879	0.8248	0.8112
Enrichment for all psychiatric drugs				
SCZ	**1.21×10^{-7}**	**4.41×10^{-8}**	**1.02×10^{-3}**	**7.44×10^{-4}**
BD	**7.67×10^{-4}**	**4.13×10^{-4}**	**6.59×10^{-3}**	**5.03×10^{-3}**
MDD-PGC	0.2691	0.2471	0.2859	0.2878
MDD-CONVERGE	0.5842	0.5316	0.2983	0.2422
AD	0.2307	0.1921	**0.0117**	**8.14×10^{-3}**
ANX	0.5701	0.5650	0.4802	0.4488
ADHD	0.9482	0.9455	0.6337	0.5999
ASD	0.9778	0.9738	0.5533	0.4971

See **Table 4** for abbreviations. Test results with $P < 0.05$ and $q < 0.05$ are in bold; results with $P < 0.05$ and $q < 0.2$ are in italics. Results are not presented for ASD as there were only two matched instances.

well-matched to the MDD-CONVERGE sample. The latter sample is composed of Chinese women, while the GTEx project is only 1.0% Asians and 34.4% female as of 22 December 2016 (https://gtexportal.org/home/tissueSummaryPage). While we expect overlap in the genetics of expression regulation, the imputation quality may be affected. Notwithstanding some negative results in the drug enrichment analyses, many of the top repositioning hits are suggested in preclinical or clinical studies and may be still worthy of further investigations.

Concerning the repositioning results for MDD, we also noted that despite inclusion of two GWAS data sets (MDD-PGC and MDD-CONVERGE), we do not observe much overlap in their repositioning hits. This may be due to a number of reasons. First, while both studies (MDD-PGC and MDD-CONVERGE) focused on depression, there were a number of differences between these two samples. The CONVERGE sample includes only Chinese women, while the PGC sample is composed of primarily Caucasian subjects and includes both males and females. In addition, the MDD-CONVERGE study recruited patients with severe recurrent depression with melancholia. On the other hand, the MDD-PGC samples contained subjects with mixed severities and subtypes of depression. With substantial differences in subject ascertainment, ethnicity and gender, differences in GWAS results and consequently repositioning hits may not be totally unexpected. Indeed, GWAS of the PGC sample did not replicate the top signals in MDD-CONVERGE, although polygenic scores from PGC were weakly predictive of case–control status in the MDD-CONVERGE cohort.

It is worth noting that current diagnosis of MDD relies on clinical symptoms, and no biomarkers are available; as such it is unlikely that this diagnosis reflects a single biological entity with a uniform genetic basis. Substantial heterogeneity among depressive patients has long been recognized[50], and given the relatively moderate sample sizes of both studies, differences in the drug repositioning results may be understandable. Another possible reason, as mentioned above, is that the GTEx sample mainly consists of Caucasians with a higher proportion of males, and this may impair the accuracy of imputation when using GTEx as reference.

Viewed from another angle, our proposed method is useful in the sense that one could focus the repositioning on a specific subtype (or particular phenotype) of a disease. The researcher can make use of GWAS data on particular disease subtypes or phenotypes and employ our methodology for repositioning. Notably, a recent paper also discussed how drug development for MDD might be affected by the heterogeneity and lack of reliability of the diagnosis[50].

There are a few general limitations to our approach. It is worth noting that the GWAS-imputed transcriptome captures the genetically regulated part of expression, and expression changes due to other factors (for example, environmental risk factors) cannot be modeled. The method of comparing expressions is largely hypothesis-free, as mentioned previously, but the assumptions of opposite

Table 6 Drug enrichment analyses *P* values with drug-sets defined by those listed on ClinicalTrials.gov

Disorder	Self (Fisher)	Compet (Fisher)	Self (min*P*)	Compet (min*P*)
SCZ	*0.0162*	*0.0116*	0.5155	0.4949
BD	*0.0167*	*0.0132*	*0.0178*	*0.0158*
MDD-PGC	*0.0448*	*0.0396*	0.1085	0.1068
MDD-CONVERGE	0.5465	0.4978	0.6540	0.5910
AD	0.4371	0.4159	0.4969	0.4642
ANX	0.0894	0.0783	*0.0100*	*0.0066*
ADHD	0.1765	0.1728	0.2834	0.2691
ASD	0.9502	0.9555	0.8904	0.9004

See **Table 4** for abbreviations. Test results with $P < 0.05$ and $q < 0.2$ are in italics.

expression patterns may not be completely true for every drug–disease pair. Our method may be improved by incorporating knowledge of drug targets or drug–disease relationships where such information is available. Drug expression profiles are not measured in brain tissues in Cmap, although the original publication on Cmap showed that drugs for neuropsychiatric diseases such as AD or SCZ can still be reasonably modeled[11]. There are also limitations to the curation process. Due to the very large number of drug–disorder pairs, we did not intend to perform a formal systematic review for each pair, and as such we have not assessed the quality of each study in detail. Inherent bias (for example, bias in sample selection, inadequate randomization or blinding) in at least some of these cited studies is inevitable, and publication bias is also probable. It is difficult to quantify the publication bias by statistical means, due to the small number of studies available for most drugs. Given these limitations, further critical and thorough reviews of relevant studies are recommended before follow-up of each repositioning hit. We also emphasize that the best approach for verifying the repositioning predictions should rest on further careful and adequately sized preclinical and clinical studies, and the current study does not provide confirmatory evidence for the repositioning hits.

Since a wide range of drugs and disorders are considered here, much work remains to be done to prioritize the most promising candidates for further investigations. As the most significant drug-set enrichment results were observed for SCZ and BD, repositioning candidates for these two disorders may hold greater promises at this stage. It may also be worthwhile to examine the ability of each drug in passing through the blood–brain barrier. In addition, combining our proposed approach with other (computational or experimental) drug repositioning methods may help further narrow down the candidates.

In conclusion, we have developed a framework for drug repositioning by linking up GWAS and drug expression profiling, and here we have applied the methodology to seven psychiatric disorders. This framework can potentially be applied to any complex diseases or traits. To our knowledge, this is also the first systematic analyses of drug repositioning covering all major psychiatric disorders. In addition, our analyses provide support for the hypothesis that psychiatric GWAS signals are enriched for known drug indications. We present a list of repositioning opportunities for each disorder, which we believe will serve as a useful resource for preclinical and clinical researchers pursuing further studies.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We would like to thank C.-W. Chan and C.-F. Wong for assistance in drug annotations. We are grateful to S.K.W. Tsui and S.S.Y. Lui for discussions and to the Hong Kong Bioinformatics Centre for computing support. This study was partially supported by the Lo-Kwee Seong Biomedical Research Fund and a Direct Grant from the Chinese University of Hong Kong to H.-C.S.

AUTHOR CONTRIBUTIONS

H.-C.S. conceived and designed the study. H.-C.S. and C.K.-L.C. performed data analyses. H.-C.S. interpreted the data. H.-C.S., C.K.-L.C., W.-T.C., K.-S.H., C.-P.L. and S.H.-Y.Y. performed drug annotations (here W.-T.C., K.-S.H., C.-P.L. and S.H.-Y.Y. are listed in alphabetical order). P.-C.S. provided advice on statistical and computational analyses. H.-C.S. wrote the manuscript and supervised the study. All authors approved the final version of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Hyman, S.E. Psychiatric drug development: diagnosing a crisis. *Cerebrum* **2013**, 5 (2013).
2. Breen, G. *et al.* Translating genome-wide association findings into new therapeutics for psychiatry. *Nat. Neurosci.* **19**, 1392–1396 (2016).
3. Dudley, J.T., Deshpande, T. & Butte, A.J. Exploiting drug-disease relationships for computational drug repositioning. *Brief. Bioinform.* **12**, 303–311 (2011).
4. Sanseau, P. *et al.* Use of genome-wide association studies for drug repositioning. *Nat. Biotechnol.* **30**, 317–320 (2012).
5. Lencz, T. & Malhotra, A.K. Targeting the schizophrenia genome: a fast track strategy from GWAS to clinic. *Mol. Psychiatry* **20**, 820–826 (2015).
6. Cao, C. & Moult, J. GWAS and drug targets. *BMC Genomics* **15** (Suppl. 4), S5 (2014).
7. Nelson, M.R. *et al.* The support of human genetic evidence for approved drug indications. *Nat. Genet.* **47**, 856–860 (2015).
8. Talevi, A. Multi-target pharmacology: possibilities and limitations of the "skeleton key approach" from a medicinal chemist perspective. *Front. Pharmacol.* **6**, 205 (2015).
9. Dudley, J.T. *et al.* Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. *Sci. Transl. Med.* **3**, 96ra76 (2011).
10. Sirota, M. *et al.* Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci. Transl. Med.* **3**, 96ra77 (2011).
11. Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
12. Raghavan, R. *et al.* Drug discovery using clinical outcome-based Connectivity Mapping: application to ovarian cancer. *BMC Genomics* **17**, 811 (2016).
13. Barbeira, A. *et al.* MetaXcan: summary statistics based gene-level association method infers accurate PrediXcan results. Preprint at https://doi.org/10.1101/045260 (2016).
14. Litman, R.E., Su, T.P., Potter, W.Z., Hong, W.W. & Pickar, D. Idazoxan and response to typical neuroleptics in treatment-resistant schizophrenia. Comparison with the atypical neuroleptic, clozapine. *Br. J. Psychiatry* **168**, 571–579 (1996).
15. Jockers-Scherübl, M.C. *et al.* Negative symptoms of schizophrenia are improved by the addition of paroxetine to neuroleptics: a double-blind placebo-controlled study. *Int. Clin. Psychopharmacol.* **20**, 27–31 (2005).
16. Foye, W.O., Lemke, T.L. & Williams, D.A. *Foye's Principles of Medicinal Chemistry* (Lippincott Williams & Wilkins, 2013).
17. Rahmanzadeh, R. *et al.* Effect of bumetanide, a selective NKCC1 inhibitor, on hallucinations of schizophrenic patients; a double-blind randomized clinical trial. *Schizophr. Res.* **184**, 145–146 (2017).
18. Sigalas, P.D., Garg, H., Watson, S., McAllister-Williams, R.H. & Ferrier, I.N. Metyrapone in treatment-resistant depression. *Ther. Adv. Psychopharmacol.* **2**, 139–149 (2012).
19. Brown, E.S., Bobadilla, L. & Rush, A.J. Ketoconazole in bipolar patients with depressive symptoms: a case series and literature review. *Bipolar Disord.* **3**, 23–29 (2001).
20. Pacchiarotti, I. *et al.* The International Society for Bipolar Disorders (ISBD) task force report on antidepressant use in bipolar disorders. *Am. J. Psychiatry* **170**, 1249–1262 (2013).
21. Muneer, A. Bipolar disorder: role of inflammation and the development of disease biomarkers. *Psychiatry Investig.* **13**, 18–33 (2016).
22. Henderson, D.C., Vincenzi, B., Andrea, N.V., Ulloa, M. & Copeland, P.M. Pathophysiological mechanisms of increased cardiometabolic risk in people with schizophrenia and other severe mental illnesses. *Lancet Psychiatry* **2**, 452–464 (2015).
23. Parsaik, A.K. *et al.* Statins use and risk of depression: a systematic review and meta-analysis. *J. Affect. Disord.* **160**, 62–67 (2014).

24. Redlich, C. *et al.* Statin use and risk of depression: a Swedish national cohort study. *BMC Psychiatry* **14**, 348 (2014).
25. Malison, R.T., Price, L.H., Nestler, E.J., Heninger, G.R. & Duman, R.S. Efficacy of papaverine addition for treatment-refractory major depression. *Am. J. Psychiatry* **154**, 579–580 (1997).
26. Furey, M.L. & Drevets, W.C. Antidepressant efficacy of the antimuscarinic drug scopolamine: a randomized, placebo-controlled clinical trial. *Arch. Gen. Psychiatry* **63**, 1121–1129 (2006).
27. Spielmans, G.I. *et al.* Adjunctive atypical antipsychotic treatment for major depressive disorder: a meta-analysis of depression, quality of life, and safety outcomes. *PLoS Med.* **10**, e1001403 (2013).
28. Poleszak, E. *et al.* Effects of ifenprodil on the antidepressant-like activity of NMDA ligands in the forced swim test in mice. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **46**, 29–35 (2013).
29. Heresco-Levy, U. *et al.* A randomized add-on trial of high-dose D-cycloserine for treatment-resistant depression. *Int. J. Neuropsychopharmacol.* **16**, 501–506 (2013).
30. Ansseau, M., Doumont, A., Thiry, D. & Gelders, Y. Pilot study of a specific serotonergic antagonist, pirenperone, in the treatment of anxiety disorders. *Acta Psychiatr. Belg.* **83**, 517–524 (1983).
31. Krystal, A.D., Sutherland, J. & Hochman, D.W. Loop diuretics have anxiolytic effects in rat models of conditioned anxiety. *PLoS One* **7**, e35417 (2012).
32. Spinosa, Hde.S., Stilck, S.R. & Bernardi, M.M. Possible anxiolytic effects of ivermectin in rats. *Vet. Res. Commun.* **26**, 309–321 (2002).
33. Pittler, M.H. & Ernst, E. Efficacy of kava extract for treating anxiety: systematic review and meta-analysis. *J. Clin. Psychopharmacol.* **20**, 84–89 (2000).
34. Lu, D., Song, H., Hao, Z., Wu, T. & McCleery, J. Naftidrofuryl for dementia. *Cochrane Database Syst. Rev.* (12): CD002955 (2011).
35. Szatmari, S.Z. & Whitehouse, P.J. Vinpocetine for cognitive impairment and dementia. *Cochrane Database Syst. Rev.* (1): CD003119 (2003).
36. Imbimbo, B.P., Solfrizzi, V. & Panza, F. Are NSAIDs useful to treat Alzheimer's disease or mild cognitive impairment? *Front. Aging Neurosci.* **2**, 19 (2010).
37. Silva, R.R., Munoz, D.M. & Alpert, M. Carbamazepine use in children and adolescents with features of attention-deficit hyperactivity disorder: a meta-analysis. *J. Am. Acad. Child Adolesc. Psychiatry* **35**, 352–358 (1996).
38. Martin, C.A. *et al.* Lobeline effects on cognitive performance in adult ADHD. *J. Atten. Disord.* https://dx.doi.org/10.1177/1087054713497791 (2013).
39. Zametkin, A., Rapoport, J.L., Murphy, D.L., Linnoila, M. & Ismond, D. Treatment of hyperactive children with monoamine oxidase inhibitors. I. Clinical efficacy. *Arch. Gen. Psychiatry* **42**, 962–966 (1985).
40. Akhondzadeh, S. *et al.* Double-blind placebo-controlled trial of pentoxifylline added to risperidone: effects on aberrant behavior in children with autism. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **34**, 32–36 (2010).
41. Mohammadi, M.R. *et al.* Double-blind, placebo-controlled trial of risperidone plus amantadine in children with autism: a 10-week randomized study. *Clin. Neuropharmacol.* **36**, 179–184 (2013).
42. Gkogkas, C.G. *et al.* Autism-related deficits via dysregulated eIF4E-dependent translational control. *Nature* **493**, 371–377 (2013).
43. Kentsis, A., Topisirovic, I., Culjkovic, B., Shao, L. & Borden, K.L. Ribavirin suppresses eIF4E-mediated oncogenic transformation by physical mimicry of the 7-methyl guanosine mRNA cap. *Proc. Natl. Acad. Sci. USA* **101**, 18105–18110 (2004).
44. de Jong, S., Vidler, L.R., Mokrab, Y., Collier, D.A. & Breen, G. Gene-set analysis based on the pharmacological profiles of drugs to identify repurposing opportunities in schizophrenia. *J. Psychopharmacol.* **30**, 826–830 (2016).
45. Gaspar, H.A. & Breen, G. Pathways analyses of schizophrenia GWAS focusing on known and novel drug targets. Preprint at https://doi.org/10.1101/091264 (2016).
46. Ruderfer, D.M. *et al.* Polygenic overlap between schizophrenia risk and antipsychotic response: a genomic medicine approach. *Lancet Psychiatry* **3**, 350–357 (2016).
47. Breier, A. Anxiety disorders and antipsychotic drugs: a pressing need for more research. *Am. J. Psychiatry* **168**, 1012–1014 (2011).
48. Singh, S.P., Singh, V., Kar, N. & Chan, K. Efficacy of antidepressants in treating the negative symptoms of chronic schizophrenia: meta-analysis. *Br. J. Psychiatry* **197**, 174–179 (2010).
49. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
50. Lieblich, S.M. *et al.* High heterogeneity and low reliability in the diagnosis of major depression will impair the development of new drugs. *BJPsych Open* **1**, e5–e7 (2015).

ONLINE METHODS

Imputation of expression profile from GWAS data. Recently, methods have been developed to impute expression from GWAS variants[13,51,52]. The main idea is to build a statistical model to predict expression levels from SNPs in a reference transcriptome data set, and this prediction model can then be applied to new genotype data. This approach estimates the component of gene expression that is contributed by (germline) genetic variations. The program PrediXcan[51] was developed for this purpose based on models built from the Genotype–Tissue Expression (GTEx) project[53] and the Depression Genes and Networks (DGN) cohort[54]. As most individual genotype data are not publicly available due to privacy concerns, we applied a recently developed algorithm called MetaXcan, which allows imputation of expression z-scores (i.e., z-statistics derived from association tests of expression changes with disease status) based on summary statistics alone. This method was shown to be in excellent concordance with predictions made from raw genotype data[13]. Assuming that a set of SNPs (SNP_1, SNP_2, …, SNP_k) contribute to the expression of gene g, the following formula can be used to compute the expression z-scores for the gene:

$$z_g \approx \sum_{i=1}^{k} w_{gi} \frac{\hat{\sigma}_i}{\hat{\sigma}_g} \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)}$$

where w_{gi} is the weight given to SNP_i for predicting the expression level of gene g, $\hat{\sigma}_i$ and $\hat{\sigma}_g$ denote the estimated variance of SNP_i and gene g respectively (estimated from a reference genotype data set), and $\hat{\beta}_i / SE(\hat{\beta}_i)$ denotes the summary z-statistic of SNP_i of the disease trait. The weights of SNPs w_{gi} were computed from reference datasets of expression quantitative trait loci (eQTL) studies. We downloaded precomputed weights derived from an elastic net prediction model with GTEx as the reference transcriptome data set, provided by the authors of MetaXcan. The above methodology is very similar to another transcriptome imputation algorithm, TWAS[52], but TWAS uses a different prediction method (based on a mixed model). We used MetaXcan in this study as the program readily allows transcriptome imputation for a much larger variety of tissues.

Numerous previous studies have shown that using GWAS to impute expression profiles and test for transcriptome–trait associations is both valid and provides important insight into the susceptibility genes underlying different diseases. For example, Gamazon *et al.*[51] showed that the prediction R^2 of expression using an elastic net model (the same model employed here) reached or exceeded the lower bound of heritability estimates in 94% of genes. Other studies reported that gene-based risk scores (GeRS) from GWAS-based transcriptome association tests predicted complex traits well[52]; for example, in a study on SCZ, GeRS outperformed SNP-based polygenic scores[52]. Susceptibility genes identified from this new approach have also been shown to replicate in independent cohorts and 'forecast' disease genes later discovered in larger GWAS[52]. In addition they demonstrate functional relevance to the diseases under study[13,52]. See the **Supplementary Note** for a more detailed explanation. In this study no statistical methods were used to predetermine sample sizes, but the GWAS samples from which transcriptome profiles are derived are large, mostly over 10^4.

GWAS-based transcriptome-wide association study of seven psychiatric disorders. GWAS summary statistics were obtained from the Psychiatric Genomics Consortium (PGC) website (http://www.med.unc.edu/pgc/results-and-downloads). We downloaded eight sets of GWAS summary statistics corresponding to seven psychiatric disorders, including schizophrenia (SCZ)[49], major depressive disorder (MDD)[55,56], bipolar disorder (BD)[57], Alzheimer's disease (AD)[58], anxiety disorders (ANX)[59], autistic spectrum disorders (ASD)[60] and attention deficit hyperactivity disorder (ADHD)[61]. We employed two different sets of summary statistics for MDD; the first set was from the PGC group[55], while the other was from the CONVERGE study, which recruited more severe MDD cases from Chinese women only[56]. Details of individual studies are described in the respective references. Transcriptome-wide association tests based on imputed expression profiles were conducted for ten brain regions included in GTEx, using default parameters in MetaXcan.

Drug-induced expression profiles. Drug-induced expression profiles were derived from the Cmap database, a resource of genome-wide expression profile of cultured cell lines treated with 1,309 different drugs or small molecules[11]. We downloaded raw expression data from Cmap and performed normalization with the MAS5 algorithm. Expression levels of genes represented on more than one probe set were averaged. Differential expression between treated cell lines and controls was tested using the limma package[62]. We performed analyses on each combination of drug and cell line, for a total of 3,478 instances. Statistical analyses were performed in R version 3.2.1, with the assistance of the R package 'longevityTools'.

Comparison of gene expression profiles of drugs versus diseases. Next, we compared the expression profiles (in z-scores) of drugs versus those of diseases. Reversed patterns of expression can be tested by Spearman or Pearson correlations. Yet another approach is to use only the K-most up- or downregulated genes in computing the correlations[63]. In this study we employed all five methods (i.e., Kolmogorov–Smirnov test, Spearman correlation with all or with the most differentially expressed genes and Pearson correlation with all or with the most differentially expressed genes) in our analyses and computed the average ranks. As there are no consensus methods to define K, we also set different values of K (50, 100, 250, 500) and averaged the results for each method. Drugs were then ranked in ascending order of their connectivity scores or correlation results (i.e., the most negative correlation or connectivity score ranked first).

Besides correlation measures, we also employed the Kolmogorov–Smirnov (KS) test to assess reversed expression profiles. The original study on Cmap employed this test[11] to compare expression patterns. In brief, the aim of the KS test is to evaluate whether a set of disease-related genes are ranked higher or lower than expected in a list of genes sorted by their drug-induced expression levels. The KS test was performed separately for upregulated and downregulated disease genes. For drug repositioning, we studied whether there was an enrichment of genes that are upregulated for disease but downregulated on drug treatment and vice versa. We adopted the formulae described in the original Cmap paper to calculate the 'connectivity scores'. To maintain consistency with the expression correlation tests of expression (as described above), we designated the same K-most up- or downregulated genes as disease-related genes and included the same range of values for K (50, 100, 250, 500). Indeed, the use of any thresholds (P value, q value or K thresholds) to determine 'true' associations is somewhat arbitrary, as we do not know exactly how many genes underlie each disorder and the corresponding genetic architecture. However, the arbitrariness of threshold selection is mitigated by aggregating a range of threshold values and combining various methodologies when determining the final ranks. The development of a more sophisticated, data-driven approach for threshold selection and correlation testing may further improve the current methodology. We have made our code open-source to facilitate further methodological development.

To assess the significance of the ranks, we performed permutation tests by shuffling the disease-expression z-scores and comparing them to drug transcriptomic profiles. We performed 100 permutations for each drug–disease pair and combined the distribution of ranks under the null across all drug–disease pairs (such that the null distribution was derived from 347,800 ranks under H_0).

We noted that the concentration of drugs given may be relatively high (10 μM in general; for details see Lamb *et al.*[11]), and thus it is possible that the expression levels may be higher than expected. However this is unlikely to affect our results, as Pearson correlations are not affected by linear transformations (for example, expressions multiplied by a constant), and Spearman correlations and KS tests are also rank-based methods invariant to monotone transformations. In addition, the final repositioning P values are also based on the ranks of the drugs (according to the degree of reversed expression compared to diseases).

Manual curation of the top repositioning hits. To assess the top results found in our drug repositioning algorithm, we manually curated the top 15 drugs (representing the top ~0.45% of all instances) identified for each disorder and brain region. A literature search was performed with predefined search strategies to look for evidence of therapeutic potential of the identified drugs. Details of the search methods and keywords employed are given in the **Supplementary Note**.

Tests for enrichment of known indicated drugs or drugs in clinical trial. In addition to manually inspecting the top repositioning hits, to validate our approach, we also tested for an enrichment of drugs that were (i) indicated for each disorder or (ii) included in clinical trials. The enrichment tests were similar in principle to gene-set analyses but with gene-sets replaced by drug-sets. We employed two types of tests for enrichment. In the first approach, we tested

doi:10.1038/nn.4618

whether a known drug-set, such as antipsychotics or antidepressants, was ranked significantly higher than it would be by chance; this approach is also known as a 'self-contained' test. In the second method, we compared medications in the drug-set against those outside the set and tested whether the former group was ranked significantly higher. This is also known as a 'competitive' test[64,65]. *P* values were first converted to *z*-scores via a probit function and one-tailed *t* tests were employed for drug-set analyses. Data distribution was assumed to be normal, but this was not formally tested. Details of the statistical methods are described in **Supplementary Note**.

We considered three sources of drug-sets in our analyses. The first set comes from the Anatomical Therapeutic Classification (ATC) drugs downloaded from KEGG (http://www.genome.jp/kegg/). The ATC is an established system for classifying medications. We extracted three groups of drugs: (i) all psychiatric drugs (coded 'N05' or 'N06'), (ii) antipsychotics (coded 'N05A') and (iii) antidepressants and anxiolytics (coded 'N05B' or 'N06A'). We grouped antidepressants and anxiolytics together in our analyses, as many antidepressants are indicated for ANX, and vice versa, anxiolytics are frequently prescribed to depressive patients clinically[66]. We did not specifically include drugs for dementia or psychostimulants as they are relatively few in number. Note that the ATC does not classify drugs specifically for some disorders, such as BD, ASD and ADHD.

The second source is Wei *et al.*[67], who compiled a MEDication Indication resource (MEDI) from four public medication resources, including RxNorm, Side Effect Resource 2 (SIDER2), Wikipedia and MedlinePlus. A random subset of the extracted indications was also reviewed by physicians. We used the MEDI high-precision subset (MEDI-HPS), which only includes drug indications found in RxNorm or in at least two of the other three sources, with an estimated precision of 92%.

As the aforementioned sources only include known drug indications, we also considered a wider set of drugs that are included for clinical trials on ClinicalTrials.gov. These drugs represent promising candidates that are often supported by preclinical or human studies. We downloaded a precompiled list of these drugs (created in May 2016) from https://doi.org/10.15363/thinklab.d212.

For each disorder, we compiled a list of drugs for repositioning using the imputed transcriptome profile of each brain region. We combined the drug-set analysis results across brain regions by meta-analyses of the respective *P* values. We employed two different algorithms, Fisher's method[68] and Tippett's minimum *P* approach[69] to combine the *P* values. Analyses were performed with the 'metap' R package.

Multiple testing correction was performed by the false discovery rate (FDR) approach[70]. The FDR approach controls the expected proportion of false positives and might serve as a more reasonable alternative to the control of family-wide error rate (for example, Bonferroni correction) when a relatively large numbers of tests are performed[71,72]. In our case, the Bonferroni approach was also likely to be overly conservative due to positive correlations among the tests. In this study we controlled the FDR for each combination of enrichment test method, drug class and reference database of drug indications, with a primary threshold at 0.05 (the overall FDR was thus also controlled at the same level[73]). To implement it, we employed the Benjamini-Hochberg method (as implemented by p.adjust in R), which essentially sets the proportion of null associations to 1. A **Life Sciences Reporting Summary** is available.

Data availability.

- GWAS summary statistics for psychiatric disorders: http://www.med.unc.edu/pgc/results-and-downloads;
- MEDI: https://www.vumc.org/cpm/center-precision-medicine-blog/medi-ensemble-medication-indication-resource
- List of drugs included in clinical trials (according to ClinicalTrials.gov) https://doi.org/10.15363/thinklab.d212
- Anatomical Therapeutic Chemical (ATC) Classification from KEGG: http://www.genome.jp/kegg-bin/get_htext?br08303.keg
- The Connectivity Map: https://portals.broadinstitute.org/cmap/
- MetaXcan: https://github.com/hakyimlab/MetaXcan

Code availability. The R codes for comparing expression profiles of drugs versus those of diseases and permutation testing to assess the significance of the drug ranks and for drug-set enrichment tests are available online at https://sites.google.com/site/honcheongso/software/gwascmap. The codes are for noncommercial use only.

51. Gamazon, E.R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nat. Genet.* **47**, 1091–1098 (2015).
52. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
53. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
54. Battle, A. *et al.* Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. *Genome Res.* **24**, 14–24 (2014).
55. Ripke, S. *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Mol. Psychiatry* **18**, 497–511 (2013).
56. CONVERGE consortium. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* **523**, 588–591 (2015).
57. Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat. Genet.* **43**, 977–983 (2011).
58. Lambert, J.C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).
59. Otowa, T. *et al.* Meta-analysis of genome-wide association studies of anxiety disorders. *Mol. Psychiatry* **21**, 1485 (2016).
60. Autism Spectrum Disorder Working Group of the Psychiatry Genomics Consortium. ASD GWAS - 2015 dataset. *Psychiatric Genomics Consortium* http://www.med.unc.edu/pgc/results-and-downloads (2015).
61. Neale, B.M. *et al.* Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **49**, 884–897 (2010).
62. Ritchie, M.E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
63. Cheng, J., Yang, L., Kumar, V. & Agarwal, P. Systematic evaluation of connectivity map for disease indications. *Genome Med.* **6**, 540 (2014).
64. de Leeuw, C.A., Mooij, J.M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLOS Comput. Biol.* **11**, e1004219 (2015).
65. de Leeuw, C.A., Neale, B.M., Heskes, T. & Posthuma, D. The statistical properties of gene-set analysis. *Nat. Rev. Genet.* **17**, 353–364 (2016).
66. Kanba, S. Although antidepressants and anxiolytics are frequently used together to treat depression in the acute phase, how effective is the concomitant use of these drugs? *J. Psychiatry Neurosci.* **29**, 485 (2004).
67. Wei, W.Q. *et al.* Development and evaluation of an ensemble resource linking medications to their indications. *J. Am. Med. Inform. Assoc.* **20**, 954–961 (2013).
68. Fisher, R.A. *Statistical Methods for Research Workers* (Oliver and Boyd, 1925).
69. Tippett, L.H.C. *The Methods of Statistics; an Introduction Mainly for Workers in the Biological Sciences* (Williams & Norgate Ltd., 1931).
70. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
71. Sham, P.C. & Purcell, S.M. Statistical power and significance testing in large-scale genetic studies. *Nat. Rev. Genet.* **15**, 335–346 (2014).
72. Glickman, M.E., Rao, S.R. & Schultz, M.R. False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *J. Clin. Epidemiol.* **67**, 850–857 (2014).
73. Efron, B. Simultaneous inference: when should hypothesis testing problems be combined? *Ann. Appl. Stat.* **2**, 197–223 (2008).

nature research

Corresponding author(s): Hon-Cheong SO

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

There is no specific calculation for sample size. However, the GWAS studies from which transcriptome profiles are derived are of large sample sizes, mostly over 10000.

2. Data exclusions

Describe any data exclusions.

No particular data exclusions.

3. Replication

Describe whether the experimental findings were reliably reproduced.

Not applicable.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Not applicable.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Not applicable.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☒	☐	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☒	☐	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☒	☐	A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☒	☐	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Nature Neuroscience: doi:10.1038/nn.4618

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Statistical analyses were performed in R 3.2.1. Imputation of transcriptome from GWAS data was performed with MetaXcan available at https://github.com/hakyimlab/MetaXcan. Differential gene expression was tested with the Bioconductor package "limma" 3.32.2. CMap data analyses was performed with the assistance of the R package "longevityTools" (https://github.com/tgirke/longevityTools). Meta-analysis of p-values was performed by the R package "metap". Custom codes for comparing GWAS-imputed transcriptome against drug-induced expression profiles, as well as drug-set enrichment tests, are available at https://sites.google.com/site/honcheongso/software/gwascmap.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	Not applicable.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	Not applicable.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	Not applicable.
b. Describe the method of cell line authentication used.	Not applicable.
c. Report whether the cell lines were tested for mycoplasma contamination.	Not applicable.
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	Not applicable.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	Not applicable.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	This study involves secondary analyses of GWAS summary data and is not primarily a clinical study involving patients and controls. Details of the human research participants of each of the eight GWAS datasets are given in the references.

Nature Neuroscience: doi:10.1038/nn.4618