# Translating Genomewide Association Findings into New Therapeutics for Psychiatry

Gerome Breen[1,2]*, Qingqin Li[3], Bryan L Roth[4,5,6], Patricio O`Donnell[7], Michael Didriksen[8], Ricardo Dolmetsch[9], Paul O`Reilly[1], Helena Gaspar[1,2], Husseini Manji[3], Christopher Huebel[1,2], John R Kelsoe[10], Dheeraj Malhotra[11], Alessandro Bertolino[12], Danielle Posthuma[13,14], Pamela Sklar[15,16,17], Shitij Kapur[18], Patrick F Sullivan[19,20], David A Collier[1,2,21], Howard J Edenberg[22,23]

* Corresponding author – gerome.breen@kcl.ac.uk

1. MRC Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London,
2. UK National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health, South London and Maudsley Hospital, London, UK.
3. Neuroscience Therapeutic Area, Janssen Research & Development, LLC, 1125 Trenton-Harbourton Road, Titusville, NJ 08560
4. Department of Pharmacology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7365, USA
5. National Institute of Mental Health Psychoactive Drug Screening Program (NIMH PDSP), School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7365, USA
6. Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7360, USA.
7. Neuroscience Research Unit, Pfizer Inc, Cambridge, MA, USA
8. H. Lundbeck A/S, Synaptic Transmission, Neuroscience Research DK, Ottiliavej 9, Valby 2500, Denmark.
9. Department of Neuroscience, Novartis Institutes for BioMedical Research, Cambridge, MA, United States.
10. Department of Psychiatry, University of California San Diego, San Diego, La Jolla, California; Veterans Affairs San Diego Healthcare System, La Jolla, California.
11. Neuroscience Discovery and Translational Area, Pharma Research & Early Development, F. Hoffmann - La Roche, CH-4070 Basel, Switzerland
12. Institute of Psychiatry, Department of Basic Medical Science, Neuroscience and Sense Organs, University of Bari 'Aldo Moro', Italy.
13. Department of Complex Trait Genetics, Centre for Neurogenomics and Cognitive Research/VU University Amsterdam, Amsterdam 1081 HV, Netherlands.
14. Department of Clinical Genetics, VU University Medical Centre Amsterdam, Neuroscience Campus Amsterdam, Amsterdam 1007 MB, Netherlands.
15. Departments of Psychiatry and Neuroscience, Icahn School of Medicine at Mount Sinai, New York, USA.
16. Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA.
17. Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, USA.
18. Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, England.
19. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden;
20. Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina.
21. Discovery Neuroscience Research, Eli Lilly and Company Ltd, Windlesham, Surrey, UK.
22. Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, 46202, USA.
23. Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, 46202, USA.

---

This is the author's manuscript of the article published in final edited form as:
Breen, G., Li, Q., Roth, B. L., O'Donnell, P., Didriksen, M., Dolmetsch, R., … Edenberg, H. J. (2016). Translating genome-wide association findings into new therapeutics for psychiatry. Nature Neuroscience, 19(11), 1392–1396. https://doi.org/10.1038/nn.4411

**Abstract**

Genome-wide association studies (GWAS) in psychiatry, once they reach sufficient sample size and power, have been enormously successful. The Psychiatric Genomics Consortium (PGC) aims for mega-analyses with sample sizes that will grow to (cumulatively) >1 million individuals in the next 5 years. This should lead to hundreds of new findings for common genetic variants across nine psychiatric disorders studied by the PGC. The new targets discovered by GWAS have the potential to restart largely stalled psychiatric drug development pipelines, and the translation of GWAS findings into the clinic is a key aim of the recently funded phase 3 of the PGC. This is not without considerable technical challenges. These approaches complement the other main aim of GWAS studies on risk prediction approaches for improving detection, differential diagnosis, and clinical trial design. This paper outlines the motivations, technical and analytical issues, and the plans for translating PGC3 findings into new therapeutics.

**The state of drug discovery in psychiatry.**

In psychiatry, conventional drug discovery is at an impasse [1]. In 2015, three (cariprazine, aripiprazole lauroxil, and brexpiprazole) out of 45 new drugs approved by FDA were related to psychiatry. The mechanisms of action of these drugs are not novel as their pharmacology primarily targets dopamine and serotonin receptors. There still remain significant unmet medical needs and societal costs for psychiatric disorders that necessitate novel therapeutics. [2] In disorders where partially effective treatments already exist, drug development has a higher investment risk, because any new drug has to exceed the clinical efficacy of existing treatments, or show equivalent efficacy together with significant improvements in safety and tolerability, as well as competing for market share with established standards of care. This is particularly difficult where there is a lack of novel targets with adequate validation. This has resulted in relatively higher drug discovery and development costs and longer than average cycle time in both clinical trial execution and regulatory agency review. Some companies have paused or de-prioritised their drug discovery and clinical trial efforts in psychiatry [3]. However, there are many (183) clinical trials underway or registered, showing there is still considerable investment in the field. (Supp Table 1 provides details of current and recent trials in psychiatry, including the nine PGC3 disorders).

The challenges in developing novel therapeutics for psychiatric disorders result from the paucity of novel, valid targets. This results from etiological heterogeneity, the complex and polygenic nature of genetic risk and the definition of psychiatric disorders based on the range and duration of symptoms (that are subjective, self-reported or observational). In addition, the complexity of the human brain means that large gaps exist in our knowledge of how brain expressed biochemical pathways relate to identified brain circuits and neuronal networks. The few examples of aetiology relevant higher order human behavioural functional domains and behavioural quantitative trait dimensions [4] limit the potential targets and measurable readouts that can used in animal and human experimental medicine studies. .While target identification based on genetics and biology looks increasingly feasible, concerns about the validity of existing model systems, especially rodents, have hampered the assessment of the value of potential new drug targets (target

qualification) and have led to calls for proof of concept human studies as the ultimate approach in hypothesis testing for target validation. [5] However clinical proof-of concept validation studies are expensive and carry risk, and will always be limited in number. Other challenges arise from the lack of informative biomarkers to guide proof of concept clinical studies and clinical development (for example by patient stratification), subjective clinical endpoints, and high placebo response rates (particularly in major depression) {Shorter, 2011 #629}.

**What can genetic studies offer for drug discovery?**

Human genetic studies have made tremendous progress in identifying loci linked to human disorders. Outside of psychiatry, these include high-risk mutations in single genes that identify specific targets for manipulation[4]. These include *PCSK9*, where individuals with 'knockout' mutations have lower LDL cholesterol without obvious deleterious effects, that has led to promising results in clinical trials [6], loss of function mutations in *SLC30A8* [7] which reduce the risk of type 2 diabetes, and loss of function *LPA* mutations which reduce plasma lipoprotein(a) levels and cardiovascular disease risk. [8]

With the notable exception of autism with intellectual disability, however, rare mutations account for a relatively small proportion of cases in psychiatry, although this varies among disorders and their exact contribution is debated. Where they have been found, there is evidence that they converge on the same biological pathways as common variants: genes in schizophrenia GWAS associated regions overlap with those identified by sequencing studies focussed on de-novo damaging mutations in intellectual disability and autism [9] [10] [11].

It may be more straightforward to identify a new target via rare mutations, but it is often not clear whether manipulating these targets will be effective in the wider disease population. The common disease-associated polymorphisms identified by GWAS in psychiatry and other complex disorders also have the potential to identify novel drug targets as well as new aetiologies that can kindle the generation of new model systems for therapeutic development in the wider population. [12] Several examples indicate that although GWAS loci have small effect sizes, they nonetheless may help identify targets for novel therapeutics, as shown in GWAS

meta-analyses of lipid levels, [13] or existing drugs that can be repurposed for the treatment of diseases that they were not initially developed to treat, an approach known as drug repositioning [14,15]. Integration of genetic data can be used for target selection, matching targets to indications while allowing a reduction in clinical trial costs such as by allowing more accurate identification of high risk individuals. Targets with genetic support have been shown to have a higher chance of success [16].

**What genomics can offer**

The discovery of common genetic variants associated with risk for psychiatric illness has the capability of restarting hypothesis-led drug discovery. As for other complex genetic disorders, the application of human genetics to schizophrenia, led by the PGC (URLs), has identified multiple disease susceptibility loci with increasing sample sizes. In 2014, over 100 robustly associated loci were identified through case-control GWAS meta-analysis by the PGC [9]. Similar progress is underway in other psychiatric disorders, with new successful GWAS reports expected for ADHD, autism, major depressive disorder, anorexia nervosa, and bipolar disorder in the next year.

The discovery of GWAS loci for these disorders is likely to continue for many years to come with, ultimately, many hundreds or thousands of independent genetic associations expected for each disorder [17]. This does not mean the whole genome will eventually be implicated - rather we expect thousands of physically overlapping and independently associated loci to cluster onto hundreds of gene regions. The available evidence suggests these hits will converge onto both specific genes and biological pathways.

Insight into which genes (and which gene-products) are implicated and the direction of effect is needed to determine the most appropriate therapeutic strategy. A general understanding of the additional steps in the target identification and qualification process has developed: GWAS locus-to-gene mapping to determine which gene(s) give rise to the association, plus functional studies of how the disease-associated SNPs operate (modality), either via regulatory effects (e.g. affecting RNA splicing or levels) or through direct functional effects (affecting the nature and function of a

protein). In this way, therapeutics targeting single GWAS identified targets, such as *HMGCR* in the LDL cholesterol metabolism responsible for hypercholesterolemia [18], have been successfully developed. This process is beginning for schizophrenia [19], and the PGC aims to accelerate this for all psychiatric disorders.

One problem is that GWAS hits identify variants, usually SNPs, that mark regions of the genome, so-called 'loci', but in most cases do not directly identify the genes themselves nor their causal alleles. A GWAS locus often includes multiple genes within the region of statistical significance, and a hit within a gene does not guarantee that that is the gene involved; the functional effect of the variants is not usually obvious, and it may even have a regulatory effect on a gene outside the GWAS risk locus. Data from large scale genomic and systems biology experiments are being used to identify expression, protein and methylation quantitative trait loci (e, p and m-QTLs) to try to better map causal alleles [20 21]. This includes imputation of gene expression profiles [22 23]. A caveat is that linkage disequilibrium between markers often results in multiple genes in a region being implicated by expression imputation, recapitulating the initial problem. In addition, the lack of large samples of available brain tissues from both patients and healthy donors at appropriate stages of development as yet hampers the wide scale application of this approach, although the CommonMind (http://commonmind.org) and Brainseq [24] initiatives are taking strides in this direction (discussed below). It remains the case that each GWAS locus requires careful and bespoke examination (see Geschwind et al this issue [25]).).

The available data indicate that psychiatric disorders are highly 'polygenic' and we now expect hundreds or thousands of individual variants to be associated with each disorder. A promising strategy to deal with the small effect sizes and plethora of results is to adopt a pathway- and network-informed interpretation of GWAS hits. An analysis by Cao and Moult [26] found that while only a small fraction of known drug targets are in GWAS loci (12 of 353 drug targets for 81 diseases), known drug targets are enriched three-fold in the nearest neighbour interactors (proteins that physically interact with a given protein) of genes in GWAS loci and are also enriched in second order interactors. This is supported by GWAS results in type 2 diabetes [27] which found that pathways targeted by anti-diabetes drugs are enriched in genes from GWAS and their direct protein interactors. This pool of GWAS hits, their

interacting partners and networks provides a resource for the identification of novel drug targets and for drug repositioning.

**How can genetic and genomic data be used in the psychiatric drug development pipelines?**

A critical issue in the field is how to use genetics information to drive drug discovery. As reviewed above, it often is not clear what genes are driving the association for GWAS significant loci. A potentially paradigmic example has recently emerged. C4 copy number was recently confirmed as a schizophrenia risk locus potentially affecting synaptic pruning in neurodevelopment; this study used PGC2 schizophrenia GWAS data, expression data from 700 postmortem brains, and genetic engineering of mice to confirm a potential mechanism [19]. This is already encouraging the development of new therapeutics, because synaptic pruning occurs as the brain develops to full maturity in the late teens/early adulthood, providing time during which therapeutic interventions may be possible.

Relatively few GWAS hits have thus far been studied in such detail. However, much GWAS evidence converges on particular biological pathways which are in themselves more druggable than single genes [28]. The pharmaceutical industry has also embarked on efforts to understand gene associations and the biological pathways impacted [5]. We need to link risk loci information to our understanding of pathways to help identify relevant biological processes, cell-types and brain circuits and to hone in on new molecular hypotheses and possible novel targets [29]. This need has sparked several academic projects and industry-academia pre-competitive collaborations. There are currently a large number of open-source and/or publically available efforts. These include large databases, ranging from ChEMBL. DiGB, Drug Bank to KiDB from the Psychoactive Drug Screening Program (listed in Table c), which serve as portals for identifying known molecular targets of drugs and drug-like small molecules. PHAROS (https://pharos.nih.gov/idg/index; http://targetcentral.ws/) is a new resource enabled by the NIH Druggable Genome Initiative, which serves as a portal for a variety of useful information regarding druggable targets. Likewise the Open Targets (formerly the Centre for Therapeutic Target Validation) public-private initiative in the UK integrates a large number of data sources into one searchable platform for single targets (https://www.targetvalidation.org/).

In order to enable the integration of functional genomic data from post-mortem brain samples from cases and controls new technologies are needed that enable the accurate identification of cell type specific omics profiles and individual level neuronal circuitry. Key examples driving the generation of large relevant datasets are industry-academia partnerships including the BrainSeq [24], CommonMind (URLs), and psychENCODE (URLs) projects, which allow investigators to map genes identified in GWAS onto transcriptomics in postmortem tissue from controls and cases with schizophrenia or bipolar disorder (as well as iPSC neuronal cell lines from cases and controls [30]). A primary goal is to elucidate molecular mechanisms driven by risk variants with the additional benefit that using genetic data can allow causal anchoring of molecular changes and pathology thus avoiding incidental, downstream effects of the disorders themselves and their treatments [24].

In order to advance our ability to understand GWAS data, the field will need to undertake further large-scale efforts to generate sufficient functional characterization of changes in brain gene and protein expression in patients and during development, and to move beyond schizophrenia and bipolar disorder to address many other disorders. The exploration and availability of large patient data sets is valuable. There are a number of initiatives in large, deeply phenotyped longitudinal samples aimed at mapping psychiatric genetic discoveries onto imaging, neurophysiological, and behavioral traits, to establish aetiologically related intermediate phenotypes that could be useful in the development of novel therapeutics. These and many other efforts aimed at linking genetic variations associated with risk with circuitry and molecular targets are a needed next step.

| Name | Bioactivities | Link | Summary | Last updated |
|---|---|---|---|---|
| ChEMBL | Various bioactivities ($K_i$, EC50…) | https://www.ebi.ac.uk/chembl/ | ~1.6M compounds, 14M activities, 11K targets | 2016 |
| $K_i$ DB | $K_i$ | http://kidbdev.med.unc.edu/databases/kidb.php | ~10K compounds, 59K interactions, 738 targets | 2016 |
| BindingDB | Various bioactivities | https://www.bindingdb.org/bind/index.jsp | ~542K compounds, 1.2M activities, 5K targets | 2016 |
| PharmGKB | Drug response data | https://www.pharmgkb.org/ | - | 2016 |
| Guide to Pharmacology | Various bioactivities | http://www.guidetopharmacology.org/ | ~8K compounds, 14K bioactivities, 2.7K targets | 2016 |
| DrugBank | Drug/target interactions | http://www.drugbank.ca/ | ~8K drugs, 15K drug/target associations, 4K targets | 2016 |
| CTD | Chemical-gene interactions, gene-disease and chemical-disease associations | http://www.ctdbase.org/ | ~1.4M chemical-gene interactions, 20M gene-disease associations, 2M chemical-disease associations | 2016 |
| STITCH | Association scores | http://stitch.embl.de/<br>new beta: http://stitch-beta.embl.de/ | interactions between 300K small molecules and 2.6K proteins from 1133 organisms | 2016 |
| PubChem | Various bioactivities | https://pubchem.ncbi.nlm.nih.gov/ | ~2M compounds, 230M bioactivities, 10K targets | 2016 |
| PHAROS | Various bioactivities, target-disease score | https://pharos.nih.gov/ | ~134K compounds, 140K bioactivities, 1.8K targets, 2.6K diseases | 2016 |
| Open Targets | Target-disease and drug-target associations | https://www.targetvalidation.org/ | ~2.1M target-disease associations covering 7.9K diseases and 25K targets | 2016 |
| DGIdb | Drug/gene interactions | http://dgidb.genome.wustl.edu/ | Without PharmGKB: ~12K compounds, 26K structure/gene pairs, ~3.1K targets | 2016 |
| CARLSBAD | CARLSBAD activity | http://carlsbad.health.unm.edu/ | ~435K structures, 933K structure/target pairs, 3.7K targets | 2014 |
| ChemProt | ChemProt activity | http://potentia.cbs.dtu.dk/ChemProt/ | ~1.7M structures, 7.8M structure/target pairs, 19K targets | 2016 |

**Table 1 Large and commonly used chemoinformatics resources.**

## Precision medicine for psychiatry and polygenic risk scores.

The customization of diagnosis and treatment to individuals - is likely to have a role in clinical psychiatry. However, the extent to which this will be important and the proportions of individuals with a particular psychiatric disorder who might benefit from precision medicine is unclear and is now the subject of considerable research. Genomics is an important tool in the precision medicine toolbox. It is already important for several disorders and becoming common in clinical practice (e.g., in the evaluation of children with intellectual disability and pervasive developmental delay). However, these studies are mostly focused upon rare genetic variants of

uncommonly large effect. For most individuals with serious psychiatric disorders whose risk is mediated by the cumulative effect of large numbers of common genetic variant with or without important environmental impacts, it is not yet clear whether genomics will be an important part of precision medicine in psychiatry. We know that these genetic effects significantly impact risk [9,28] but the effects are not deterministic.

An key approach is to use polygenic risk scores (extensively reviewed and discussed elsewhere [31]). A polygenic risk score (PRS) [32] is an approximate measure of an individual's common variant genetic propensity for a given disorder and, at a population level shows some predictive power [33] for case-control status. PRS approaches provide several potential routes to drug development, including identification of genetically associated endophenotypes and biomarkers. PRS can also be exploited to improve clinical trial efficacy. *Super controls* can be chosen by selecting participants with very low PRS for the disease, or PRS for low risk of side-effects or where differential diagnosis is unclear. This may convey particular benefit in trials for diseases such as Alzheimer's (being investigated by a new workgroup in the PGC), where defining cases and controls is challenging. Furthermore, prevention trials could enlist high risk individuals from the top end of the PRS distribution [34], which, amongst other benefits, may be less expensive and confounded than the sibling design [35]. Current studies in psychiatry are attempting to improve prediction of diagnosis or treatment response, for example in first episode psychosis [36].

**PGC phase 3: Target identification in Psychiatric GWAS data.**

To fully exploit GWAS data for drug development, we need to complement the direct identification of single targets and their interactors and the use of polygenic risk scores with pathway-driven approaches, explicitly targeting sets of GWAS implicated regions/proteins together. In our view, this may be a powerful means to discover new drug indications/targets that gains power by exploiting the underlying polygenic nature of these disorders. This mirrors the observation that many successful psychiatric (and other) drugs have complex receptor pharmacology profiles binding multiple targets with different affinities. The PGC is planning to exploit pathway analysis methods[37] that show better control for type 1 error alongside chemoinformatically generated gene sets to identify drugs or molecules with sets of targets significantly enriched for association in GWAS data. Applying drug pathway

analyses to psychiatric GWAS results will allow us to derive hypotheses about drug mechanisms of action and rational drug repurposing [38]. Rare variants, discovered by large scale sequencing efforts, can also be included in these analyses, particularly the known recurrent Copy Number Variations in Autism and Schizophrenia [39]. These are complemented by ongoing large scale sequencing efforts in these disorders. Although rare mutations are only found in a small percentage of cases with most common disorder [40 41], integrative pathway analysis including common and rare variants might increase power to detect statistically significant enriched pathways.

Using these data sources, three broad strategies are possible (see Figure 1). First, pathway analysis using the genetic variants found to be associated with psychiatric disorders using gene-sets (pathways) annotated for their drug associations or corresponding to sets of ligands in publically available resources such as ChEMBL and KiDB to test whether these gene sets together harbour a significant association signal using the PGC pathway analysis pipeline [42]. Second, use relevant gene expression profiles identified from case-control transcriptome data and examine their similarity to induced gene expression changes in cell lines, as identified by the NIH LINCS project (URLs) or in studies of neuronal cells derived from iPSC, to identify potential pathways and molecules which impact the expression and/or function of identified targets [43]. This strategy of 'connectivity mapping' allows identification of compounds with a similar or opposite effect on gene expression as our findings and can point to possible new treatment targets. Finally, we can layer onto these approaches "traditional" pathway annotations and ontologies (particularly GO and REACTOME) and newer data sources that may be less biased and more complete[44] to allow us to develop a mechanistic understanding.

## Conclusions

These approaches require substantial and integrated efforts, involving consortia such as the PGC, other academic groups, and industry in pre-competitive framework to drive forward target identification and qualification to the point where confidence will be high enough to begin a clinical validation process; sharing of data and expertise will be essential. It will only be through collaborative work that the field will muster enough breadth of data and resources for this effort to fulfill its translational potential beyond polygenic risk score and prediction, to the identification of new

biology and eventually towards resolving the current blockages in psychiatric drug discovery.

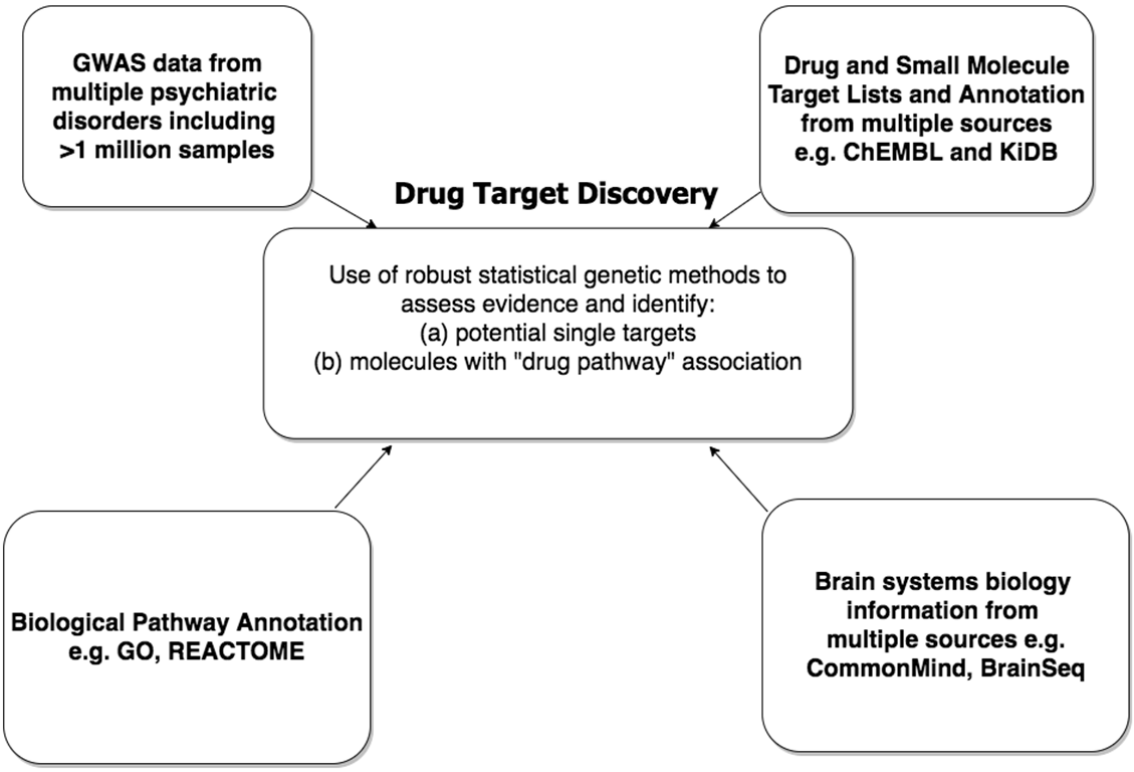


**Figure 1. PGC GWAS Drug Target Analysis Strategy: utilising diverse information sources for drug target discovery.**

## URLs

PGC (https://pgc.unc.edu), PHAROS (https://pharos.nih.gov), CHEMBL (https://www.ebi.ac.uk/chembl), Open Targets (https://www.targetvalidation.org), DGIdb (http://dgidb.genome.wustl.edu), CommonMind (http://commonmind.org), psychENCODE (http://psychencode.org), NIH LINCS (http://apps.lincscloud.org), GO project (http://geneontology.org/), REACTOME (http://www.reactome.org/), Pharos (https://pharos.nih.gov/), Open Targets (https://www.opentargets.org/).

1. Fibiger, H.C. Psychiatry, the pharmaceutical industry, and the road to better therapeutics. *Schizophr Bull* **38**, 649-50 (2012).
2. Rizzo, S.J., Edgerton, J.R., Hughes, Z.A. & Brandon, N.J. Future viable models of psychiatry drug discovery in pharma. *J Biomol Screen* **18**, 509-21 (2013).
3. Papassotiropoulos, A. & de Quervain, D.J. Failed drug discovery in psychiatry: time for human genome-guided solutions. *Trends Cogn Sci* **19**, 183-7 (2015).
4. Jones, H.J. *et al.* Phenotypic Manifestation of Genetic Risk for Schizophrenia During Adolescence in the General Population. *JAMA Psychiatry* **73**, 221-8 (2016).
5. O'Donnell, P. & Ehlers, M.D. Opportunities for New Drug Development in Psychiatry: A Glass Half-Full. *JAMA Psychiatry* **72**, 1067-8 (2015).
6. Fitzgerald, K. *et al.* Effect of an RNA interference drug on the synthesis of proprotein convertase subtilisin/kexin type 9 (PCSK9) and the concentration of serum LDL cholesterol in healthy volunteers: a randomised, single-blind, placebo-controlled, phase 1 trial. *Lancet* **383**, 60-8 (2014).
7. Pearson, E. Zinc transport and diabetes risk. *Nat Genet* **46**, 323-4 (2014).
8. Lu, W. *et al.* Evidence for several independent genetic variants affecting lipoprotein (a) cholesterol levels. *Hum Mol Genet* **24**, 2390-400 (2015).
9. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-7 (2014).
10. Purcell, S.M. *et al.* A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* **506**, 185-90 (2014).
11. Fromer, M. *et al.* De novo mutations in schizophrenia implicate synaptic networks. *Nature* **506**, 179-84 (2014).
12. Okada, Y. From the era of genome analysis to the era of genomic drug discovery: a pioneering example of rheumatoid arthritis. *Clin Genet* **86**, 432-40 (2014).
13. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707-13 (2010).
14. Manolio, T.A. *et al.* Implementing genomic medicine in the clinic: the future is here. *Genet Med* **15**, 258-67 (2013).
15. Sanseau, P. *et al.* Use of genome-wide association studies for drug repositioning. *Nat Biotechnol* **30**, 317-20 (2012).
16. Nelson, M.R. *et al.* The support of human genetic evidence for approved drug indications. *Nat Genet* **47**, 856-60 (2015).
17. Ripke, S. *et al.* Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet* **45**, 1150-9 (2013).
18. Mangravite, L.M. *et al.* Combined influence of LDLR and HMGCR sequence variation on lipid-lowering response to simvastatin. *Arterioscler Thromb Vasc Biol* **30**, 1485-92 (2010).
19. Sekar, A. *et al.* Schizophrenia risk from complex variation of complement component 4. *Nature* **530**, 177-83 (2016).
20. Psych, E.C. *et al.* The PsychENCODE project. *Nat Neurosci* **18**, 1707-12 (2015).
21. insert, S.R.t.
22. Gusev, A. *et al. Integrative approaches for large-scale transcriptome-wide association studies*, (2015).
23. Gamazon, E.R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet* **47**, 1091-8 (2015).

24. BrainSeq, A.H.B.G.C. BrainSeq: Neurogenomics to Drive Novel Target Discovery for Neuropsychiatric Disorders. *Neuron* **88**, 1078-83 (2015).
25. Geschwind, D.H.
26. Cao, C. & Moult, J. GWAS and drug targets. *BMC Genomics* **15 Suppl 4**, S5 (2014).
27. Segre, A.V. *et al.* Pathways targeted by antidiabetes drugs are enriched for multiple genes associated with type 2 diabetes risk. *Diabetes* **64**, 1470-83 (2015).
28. Cross-Disorder Group of the Psychiatric Genomics, C. *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* **45**, 984-94 (2013).
29. Schubert, C.R., Xi, H.S., Wendland, J.R. & O'Donnell, P. Translating human genetics into novel treatment targets for schizophrenia. *Neuron* **84**, 537-41 (2014).
30. Sun, Y. & Dolmetsch, R.E. How induced pluripotent stem cells are informing drug discovery in psychiatry. *Swiss Med Wkly* **146**, w14241 (2016).
31. Chatterjee, N., Shi, J. & Garcia-Closas, M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat Rev Genet* **17**, 392-406 (2016).
32. Wray, N.R., Goddard, M.E. & Visscher, P.M. Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Res* **17**, 1520-8 (2007).
33. Dima, D. & Breen, G. Polygenic risk scores in imaging genetics: Usefulness and applications. *J Psychopharmacol* **29**, 867-71 (2015).
34. Hu, Y. *et al.* The benefits of using genetic information to design prevention trials. *Am J Hum Genet* **92**, 547-57 (2013).
35. Agerbo, E. *et al.* Polygenic Risk Score, Parental Socioeconomic Status, Family History of Psychiatric Disorders, and the Risk for Schizophrenia: A Danish Population-Based Study and Meta-analysis. *JAMA Psychiatry* **72**, 635-41 (2015).
36. Vassos, E. *et al.* An Examination of Polygenic Score Risk Prediction in Individuals with First Episode Psychosis. *Biological Psychiatry* **AOP** (2016). DOI: http://dx.doi.org/10.1016/j.biopsych.2016.06.028
37. 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).
38. De Jong, S., Vidler, L., Mokrab, Y., Collier, D.A. & Breen, G.D. Gene-set analysis based on the pharmacological profiles of drugs to identify repurposing opportunities in schizophrenia. *Journal of Psychopharmacology* (2016).
39. Stefansson, H. *et al.* CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* **505**, 361-6 (2014).
40. Wellcome Trust Case Control, C. *et al.* Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature* **464**, 713-20 (2010).
41. Rucker, J.J. *et al.* Genome-wide association analysis of copy number variation in recurrent depressive disorder. *Mol Psychiatry* **18**, 183-9 (2013).
42. Network & Pathway Analysis Subgroup of Psychiatric Genomics, C. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci* **18**, 199-209 (2015).
43. Prathipati, P. & Mizuguchi, K. Systems Biology Approaches to a Rational Drug Discovery Paradigm. *Curr Top Med Chem* **16**, 1009-25 (2016).
44. Greene, C.S. *et al.* Understanding multicellular function and disease with human tissue-specific networks. *Nat Genet* **47**, 569-76 (2015).

| Name | Bioactivities | Link | Some Stats | Last update |
|---|---|---|---|---|
| ChEMBL | Various bioactivities ($K_i$, EC50…) | https://www.ebi.ac.uk/chembl/ | ~1.6M compounds, 14M activities, 11K targets | 2016 |
| $K_i$ DB | $K_i$ | http://kidbdev.med.unc.edu/databases/kidb.php | ~10K compounds, 59K interactions, 738 targets | 2016 |
| BindingDB | Various bioactivities | https://www.bindingdb.org/bind/index.jsp | ~542K compounds, 1.2M activities, 5K targets | 2016 |
| PharmGKB | Drug response data | https://www.pharmgkb.org/ | - | 2016 |
| Guide to Pharmacology | Various bioactivities | http://www.guidetopharmacology.org/ | ~8K compounds, 14K bioactivities, 2.7K targets | 2016 |
| DrugBank | Drug/target interactions | http://www.drugbank.ca/ | ~8K drugs, 15K drug/target associations, 4K targets | 2016 |
| CTD | Chemical-gene interactions, gene-disease and chemical-disease associations | http://www.ctdbase.org/ | ~1.4M chemical-gene interactions, 20M gene-disease associations, 2M chemical-disease associations | 2016 |
| STITCH | Association scores | http://stitch.embl.de/<br><br>new beta: http://stitch-beta.embl.de/ | interactions between 300K small molecules and 2.6K proteins from 1133 organisms | 2016 |
| PubChem | Various bioactivities | https://pubchem.ncbi.nlm.nih.gov/ | ~2M compounds, 230M bioactivities, 10K targets | 2016 |
| PHAROS | Various bioactivities, target-disease score | https://pharos.nih.gov/ | ~134K compounds, 140K bioactivities, 1.8K targets, 2.6K diseases | 2016 |
| Open Targets | Target-disease and drug-target associations | https://www.targetvalidation.org/ | ~2.1M target-disease associations covering 7.9K diseases and 25K targets | 2016 |
| DGIdb | Drug/gene interactions | http://dgidb.genome.wustl.edu/ | Without PharmGKB: ~12K compounds, 26K structure/gene pairs, ~3.1K targets | 2016 |

| | | | | |
|---|---|---|---|---|
| CARLSBAD | CARLSBAD activity | http://carlsbad.health.unm.edu/ | ~435K structures, 933K structure/target pairs, 3.7K targets | 2014 |
| ChemProt | ChemProt activity | http://potentia.cbs.dtu.dk/ChemProt/ | ~1.7M structures, 7.8M structure/target pairs, 19K targets | 2016 |