HGG Advances

REPORT

Gene, cell type, and drug prioritization analysis suggest genetic basis for the utility of diuretics in treating Alzheimer disease

Daria Pinakhina,[1,2] Alexander Loboda,[1,3] Alexey Sergushichev,[1] and Mykyta Artomov[1,4,5,6,7,8,*]

Summary

We introduce a user-friendly tool for risk gene, cell type, and drug prioritization for complex traits: GCDPipe. It uses gene-level GWAS-derived data and gene expression data to train a model for the identification of disease risk genes and relevant cell types. Gene prioritization information is then coupled with known drug target data to search for applicable drug agents based on their estimated functional effects on the identified risk genes. We illustrate the utility of our approach in different settings: identification of the cell types, implicated in disease pathogenesis, was tested in inflammatory bowel disease (IBD) and Alzheimer disease (AD); gene target and drug prioritization was tested in IBD and schizophrenia. The analysis of phenotypes with known disease-affected cell types and/or existing drug candidates shows that GCDPipe is an effective tool to unify genetic risk factors with cellular context and known drug targets. Next, analysis of the AD data with GCDPipe suggested that gene targets of diuretics, as an Anatomical Therapeutic Chemical drug subgroup, are significantly enriched among the genes prioritized by GCDPipe, indicating their possible effect on the course of the disease.

There is a common recognition that genome-wide association studies (GWASs) could be a powerful tool for drug development, as well as for drug repurposing.[1,2] The latter enables an effective transition of existing therapies into new applications in clinical practice. Several approaches to leverage GWAS data for drug discovery and repurposing have already been proposed.[1–13]

GWAS data are used as a source of the candidate genes to serve as targets for drug repositioning.[1] However, its utility is limited by the small yield of drug candidates from top associations in GWAS, as only a modest fraction (22%) of protein-coding genes is currently druggable.[2] Tools, such as Gentrepid,[14] use diverse information on pathway, protein-protein interaction, and/or protein domain homology to extend candidate gene lists and overcome this limitation. The other solution to candidate gene list extension is provided by pipelines for gene prioritization; however, the utility of such methods for drug and drug-target prioritization remains understudied.[15,16]

In addition to gene identification, GWAS findings can inform development of therapies by providing evidence for identification of disease-relevant cell types and tissues. This information could guide the development of drugs with increased cell type specificity to minimize potential side effects.[17,18] Despite a rapid increase in the amount of single-cell RNA sequencing data, aligning GWAS results with cell-type-specific expression patterns is the area of active development.[11,19–22] Most of the existing approaches are based on the linkage of variant-level association signals from GWAS to cell-type-specific expression through eQTL properties.

The complexity of direct implication of disease-causal genes from GWAS variant data remains the main challenge in understanding disease biology through genetic studies, indicating that more flexible solutions without a tight linkage to variant-level statistics would be required.

Here, we introduce a user-friendly random forest-based tool, GCDPipe, for GWAS-derived gene-level results analysis, which extends the list of disease gene candidates through the estimation of probability to influence disease risks; identifies gene expression profiles across cell types and tissues with the highest importance for the putative disease genes identification; and prioritizes drugs based on their affinity to the putative disease genes using drug-gene interaction databases.

Two types of disease-specific input are required to run GCDPipe: a feature matrix with expression values across cell types/tissues that can be constructed from a wide range of publicly available gene expression atlases (Table S1, Figure 1) and a gene list for training and testing of the classifier. The latter could be either a set of genes that are marked as "likely causal" and "likely non-causal" for training the model, alternatively a list of GWAS loci with the "likely causal" genes could be provided and GCDPipe will construct the training gene set automatically (supplemental methods).

There are two key steps in gene classification within GCDPipe. First, a random forest classifier is trained to

[1]ITMO University, 197101 Saint Petersburg, Russia; [2]Bekhterev National Medical Research Center, 192019 Saint Petersburg, Russia; [3]Almazov National Medical Research Center, 191014 Saint Petersburg, Russia; [4]Broad Institute, Cambridge, MA 02142, USA; [5]Massachusetts General Hospital, Boston, MA 02114, USA; [6]The Steve and Cindy Rasmussen Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, OH 43205, USA; [7]Department of Pediatrics, The Ohio State University College of Medicine, Columbus, OH 43210, USA
[8]Lead contact
*Correspondence: mykyta.artomov@nationwidechildrens.org
https://doi.org/10.1016/j.xhgg.2023.100203.

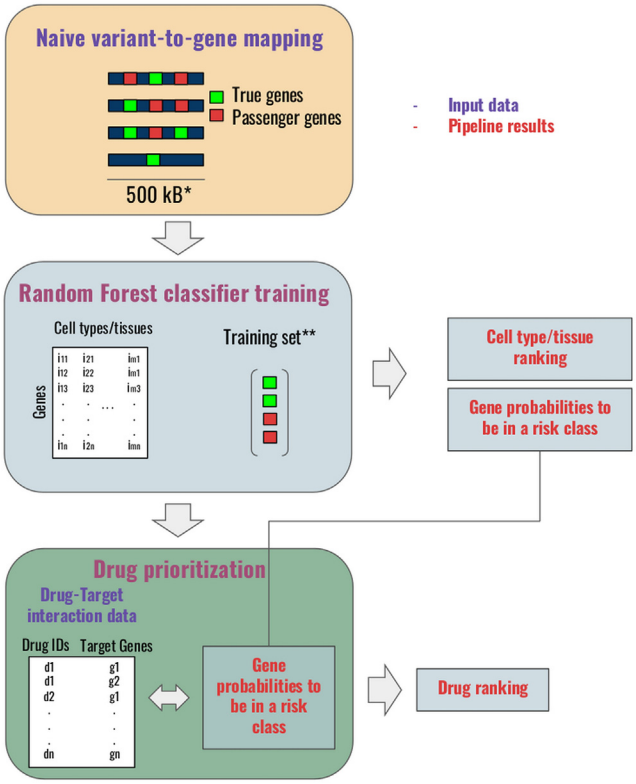

Figure 1. A general pipeline scheme for joint risk gene identification, cell/tissue type ranking, and drug prioritization for a disease based on GWAS variant-to-gene mapping results

*Adjustable parameter; **Is split into training and validation sets for cross-validation during hypermarameter optimization; a custom fraction of genes can also be allocated for the test set.

distinguish between risk and non-risk genes. The classifier training involves hyperparameter tuning with GridSearchCV using fivefold cross-validation (Figure S1A).

Second, the performance of this classifier is estimated to ensure the proper quality of training. An original set of input genes is initially split into two parts: a training and a testing set. The first one is used solely for the classifier training and the other is set aside for future performance estimation (Figures S1A and 1).

The identification of the disease-relevant expression profiles, representing cell types and tissues, is performed using a modification of the feature importance analysis with SHAP (SHapley Additive exPlanations), which assigns each feature an additive importance value (SHAP value) for a particular prediction.[23] The expression profile score is computed as follows:

$$Sf = \frac{cor(SHAP_{cls2}, f) \cdot \left(\underline{|SHAP_{cls2}|}\right)}{|cor(SHAP_{cls2}, f)|},$$

where Sf is a score for feature f, $SHAP_{cls2}$- SHAP values for the class 2 (the class of risk), and f- values of the feature f (supplemental methods: GCDPipe implementation and testing). The Sf score could be positive or negative depending on whether the putative risk genes are highly or lowly expressed in the selected expression profile. The greater the absolute value of Sf, the more important the expression profile is for the discrimination between disease and non-disease genes in the GCDPipe model based on the SHAP values.

Next, GCDPipe returns the drug prioritization: the drugs are arranged by mean probability of their gene targets to be assigned to the risk class. As an additional input for this step, GCDPipe requires a matrix with drug identifiers and their gene targets. Optionally, the pipeline will automatically estimate the quality of gene and drug ranking (supplemental methods).

GCDPipe returns: a range of quality metrics for the obtained classifier and their receiver operating characteristic (ROC) curves; a matrix with gene probabilities to influence disease risk and their risk class assignment using the probability threshold corresponding to the largest difference between true positive and false positive rates; and a matrix with the scores of gene expression profiles. If a drug analysis was performed, a matrix with drug scores is returned (supplemental methods).

We assembled the sets of genes for classifier training and testing based on variant-to-gene mapping from GWASs on inflammatory bowel disease (IBD), schizophrenia (SCHZ), and Alzheimer disease (AD) (supplemental methods). The feature matrices were obtained using publicly available expression datasets, and the information on drug targets was obtained from DrugCentral.[24] Only the drugs with DrugBank IDs were considered (Table S1).

The resultant risk gene classifier returned area under the curve-ROC of 0.91, 0.82, 0.82 for the test set in the studies of SCHZ, IBD, and AD, respectively (Figures S1B–S1D). The observed estimated quality of the classifiers suggests that current advancements in variant-to-gene mapping for GWAS provide sufficient gene-level data to achieve significant disease gene discrimination even in supervised training contexts.

To illustrate the practical utility of GCDPipe output, we compared it with the sets of genes that are differentially expressed across eight conditions. We performed an enrichment analysis of 500 top-scoring genes with top 200 differentially expressed genes for each phenotype (Table S1). As another source of independent information, a list of 200 genes with the lowest Q-values obtained from the SCHZ exome sequencing project (SCHEMA)[25] was included into this comparison (Table S1). Significant overlap between prioritized genes and independently assembled gene sets from differential expression analysis/exome sequencing was observed for all three case studies (SCHZ: SCHEMA - $P_{adj.} = 9.23 \times 10^{-9}$, GSE25673 - $P_{adj.} = 1.36 \times 10^{-3}$; IBD: GSE179285 - $P_{adj.} = 3.13 \times 10^{-8}$, GSE98820 - $P_{adj.} = 7.52 \times 10^{-8}$, GSE134025 - $P_{adj.} = 4.45 \times 10^{-2}$; AD: GSE97760–$5.32 \times 10^{-3}$, Figures S1E–S1G).

In addition, we performed Kyoto Encyclopedia of Genes and Genomes (KEGG)[26] pathway enrichment of 500 top-ranking risk genes and of the "true" genes from the classifier training and testing sets for SCHZ to assess the applicability of the GCDPipe for uncovering

disease-relevant molecular mechanisms. The top significantly enriched KEGG pathways among the 500 best-scoring risk genes, excluding those used for training and testing, were dopaminergic synapse ($P_{adj.} = 1.00 \times 10^{-11}$, hypergeometric), retrograde endocannabinoid signaling ($P_{adj.} = 6.59 \times 10^{-10}$, hypergeometric), nicotine addiction ($P_{adj.} = 2.38 \times 10^{-9}$, hypergeometric), and MAPK signaling pathway ($P_{adj.} = 4.22 \times 10^{-9}$, hypergeometric) (overall, there were 71 significantly enriched pathways), while the only enriched pathway for risk genes used for training and testing was GnRH secretion ($P_{adj.} = 4.54 \times 10^{-5}$, hypergeometric) (Figure S2). Alterations in dopaminergic neurotransmission are considered as one of the most robust pathophysiological observations in SCHZ patients.[27,28] Another contributor to SCHZ pathogenesis, proposed as a prospective target for its treatment and measurement, is the endocannabinoid system, and retrograde endocannabinoid signaling is the second most enriched pathway in the top-ranking genes.[29–33] Nicotine addiction is the third most enriched KEGG gene set, and it complies with the observation of an association between SCHZ and tobacco smoking.[34] These pathways are interconnected with other enriched gene sets, including glutamatergic synapse and calcium signaling. Further analysis of the resultant gene-pathway network could reveal how these mechanisms are interconnected at the molecular level. None of these pathways is enriched in the GWAS-derived risk gene set used for model development.

We also tested the pipeline performance with varying training risk gene count for SCHZ. It was run with 2, 5, 14, and 24 risk genes; enrichment of the top-ranking genes with the genes of the SCHEMA gene set and the enrichment of the gene ranking with SCHZ drug targets were estimated as performance metrics. In this assessment, enrichment with SCHEMA genes and with SCHZ drug targets is robust with respect to the size of the training gene set (Figure S3).

The scheme of the control experiments to assess GCDPipe utility for cell type, drug target, and drug prioritization is presented in Figure 2A. Two types of control data were used: the expression profile and gene scores obtained from training GCDPipe on the random training sets of the same size as original training data, and randomly generated gene ranking, implying no model training.

For IBD data, the ranking of the 490 expression profiles used for classifier training obtained with GCDPipe was significantly enriched with $CD4^+$ T cells (GSEA enrichment $p = 0.02$), and the leading prioritized expression profile belonged to this cell type. No enrichment with $CD4^+$ T cells was observed when GCDPipe was trained on a random training gene set analogous in size (GSEA enrichment $p = 0.55$) (Figure 2B).

For the AD, microglia expression profile ranked first among 137 features used for training (GSEA enrichment $p = 8 \times 10^{-3}$), while no enrichment was found in the control experiment (GSEA enrichment $p = 0.72$) (Figures 2C, 3A, and 3B). For SCHZ, the most relevant expression profile was obtained from the inhibitory L6 LHX6 GLP1R neurons, which belong to the PVALB cluster in neuronal cell taxonomy.[24,35]

Observed enrichments are concordant with the previously published results: $CD4^+$ cells are reported to play a key role in IBD pathogenesis,[36,37] whereas proliferation and activation of microglia in the brain, concentrated around amyloid plaques, is a prominent feature of AD.[38] No consensus exists regarding a specific cell type implicated in SCHZ: a complex cellular architecture may be involved in the condition, and different molecular subtypes of SCHZ with various underlying pathological cellular mechanisms may exist.[39] Deficits in the marker gene product of the leading prioritized cell type, Lhx6 - a transcription factor, regulating parvalbumin and somatostatin neuron development and migration, are associated with GABA-related disturbances in SCHZ.[40] The results of expression profile prioritization provided by GCDPipe correspond to existing evidence on cellular pathologies underlying IBD and AD and suggest the role of LHX6-expressing interneurons in SCHZ, which was previously mentioned in several reports.[41–43]

Next, to assess the utility of GCDPipe gene ranking for drug target search, we performed enrichment analyses of genes, arranged by the probabilities to affect disease risks, with targets of the drugs that are currently used for the four selected diseases (IBD and SCHZ as case studies; coronary heart disease [MIM: 607339] (CHD) and asthma [MIM: 600807] as a comparison, Figures 2D and 2E). The disease-specific drug sets were obtained from DrugBank.[44] We observed the most significant enrichment with the drugs for the corresponding conditions in both IBD and SCHZ cases compared with the other phenotypes (Figures 2D and 2E).

The prioritization of disease-specific drug targets in the GCDPipe output was assessed. The positions of gene drug targets for all disease-relevant drugs in the ranked list of genes provided by GCDPipe were averaged and compared with the same values obtained with training GCDPipe on the random set of genes or no training at all. Training GCDPipe on GWAS data allowed to obtain gene ranking with significantly higher mean drug-target ranks compared with the control experiments (Figures 2F and 2G).

An analogous comparison was performed for drugs that were ranked by the average of their gene targets' GCDPipe scores. A significant difference between GWAS-based training and both of the control cases was observed in this assessment (Figures 2H and 2I).

In the application study for the AD, we wanted to use GCDPipe to identify a drug ATC subgroup, relevant for the disease, based on target diversity. We ranked 82 drug ATC subgroups (second level) by enrichment of their target sets in the genes, arranged by probability to be assigned to the risk class provided by GCDPipe. We then prioritized the drugs within the leading ATC subgroup by enrichment of their targets in this gene profile, and selected the drugs with the highest NES scores among them as candidate

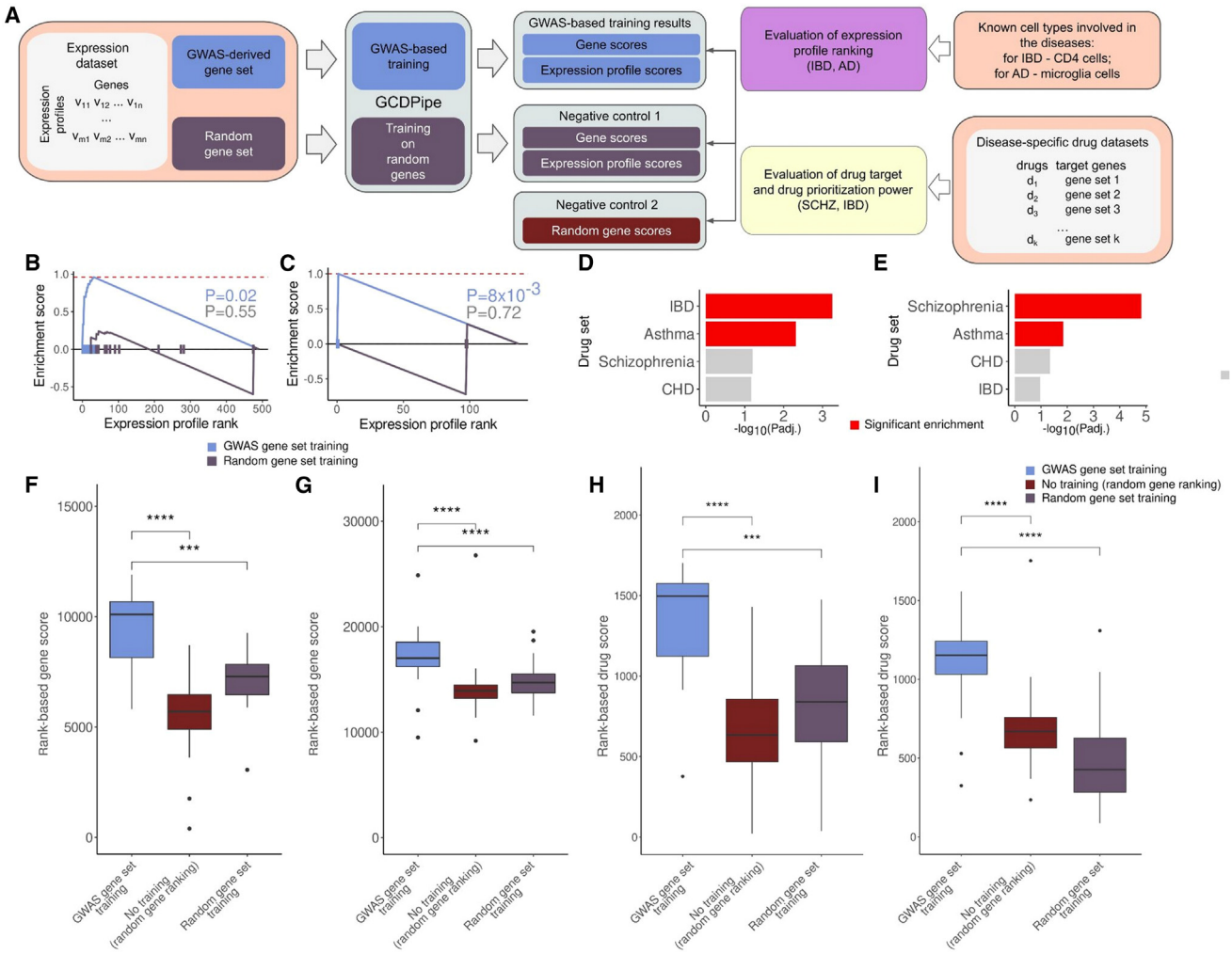

Figure 2. Assessment of GCDPipe performance in cell type/tissue, drug target, and drug prioritization

(A) A scheme of control experiments used to evaluate significance of GCDPipe usage for cell type/tissue, drug target, and drug prioritization.

(B) CD4+ T cell enrichment of expression profiles, arranged by GCDPipe scores, obtained in the IBD and control experiments.

(C) Microglia enrichment of expression profiles, arranged by GCDPipe scores, obtained in the IBD and control experiments.

(D) Comparison between significance of enrichment of the genes arranged by their probabilities to be assigned to the risk class with drug targets across a range of diseases in the IBD experiment.

(E) Comparison between significance of enrichment of the genes arranged by their probabilities to be assigned to the risk class with drug targets across a range of diseases in the schizophrenia experiment.

(F) Comparison between mean ranks of gene targets of IBD drugs in the IBD case study with control experiments.

(G) Comparison between mean ranks of gene targets of schizophrenia drugs in the schizophrenia case study with control experiments.

(H) Comparison of drug ranks of IBD drugs in the IBD case study with control experiments.

(I) Comparison of drug ranks of schizophrenia drugs in the schizophrenia case study with control experiments.

(F–I) Error bars represent standard deviation.

disease-relevant drugs (blood-brain barrier penetration potential was not assessed) (Figure 3). Interestingly, the leading enriched drug target category for AD appeared to be diuretics. The use of diuretics has been shown to be associated with a reduced risk of AD, and diverse experimental and real-world evidence appeared supporting the repurposing of diuretics for the treatment of the disease.[45–47]

Indapamide and furosemide displayed the highest normalized enrichment scores in the gene prioritization among all the drugs in this category considered in the study. Interestingly, indapamide was reported to suppress amyloid-β production in cellular models of AD as well as to improve the clearance of Aβ.[48] A decrease in activity of β-site APP-cleaving enzyme 1 (BACE1) in response to indapamide - the rate limiting step in Aβ generation, was also reported.[49]

The highest-scoring indapamide target, according to GCDPipe ranking, was potassium voltage-gated channel subfamily Q member 1 KCNQ1 (Figure 3D). Interestingly, BACE1, along with another agent important in AD pathogenesis, presenilin/γ-secretase, is found to modulate proteolytic processing of KCNQ1.[50,51] KCNQ channels are regarded as central plasticity molecules, and are

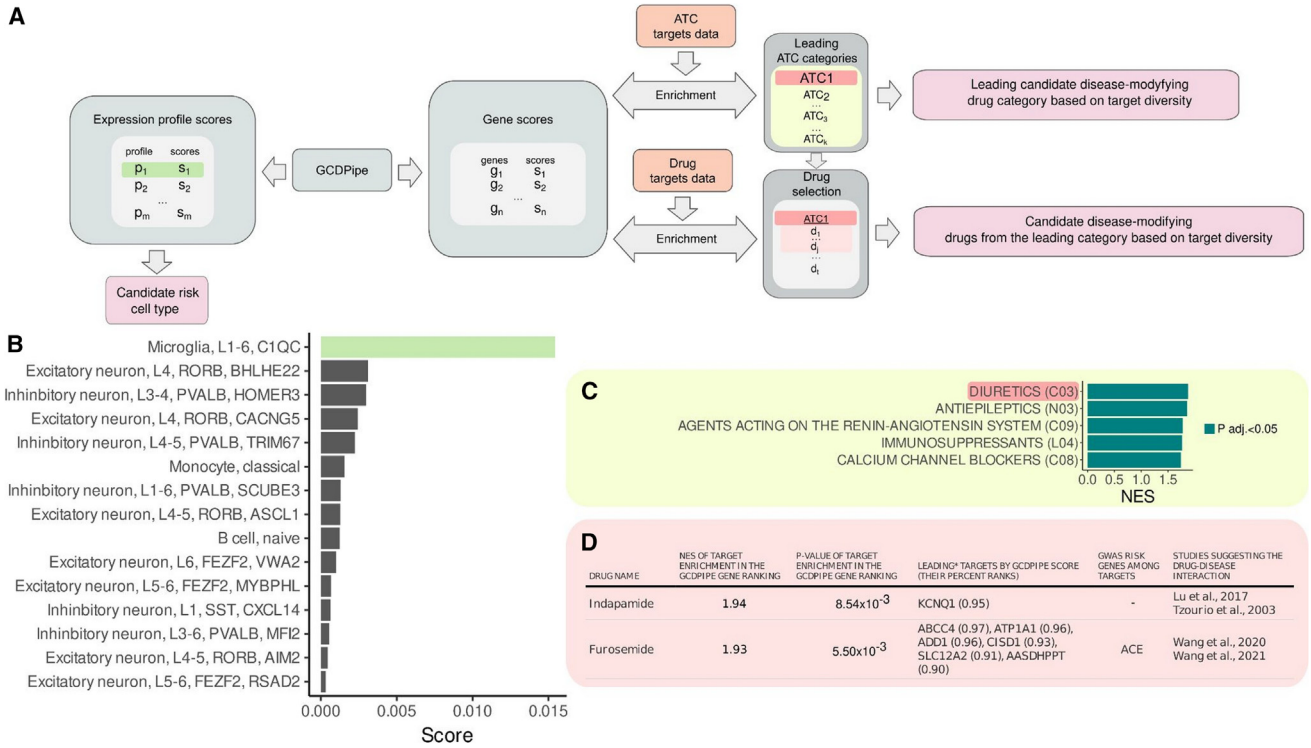

Figure 3. Application of GCDPipe to Alzheimer disease

(A) A scheme of experiments of application of the pipeline to Alzheimer disease in this study.

(B) Leading 15 expression profiles prioritized by GCDPipe for Alzheimer disease.

(C) Leading 5-second level drug ATC categories by NES scores of enrichment of their targets in the GCDPipe gene prioritization.

(D) Leading diuretic (C03) drugs with significant enrichment of their targets in GCDPipe gene prioritization demonstrating the highest NES scores within the subgroup, their gene targets with GCDPipe scores >0.90 driving the enrichment, targets present among Alzheimer disease GWAS-derived risk genes (if present), and studies considering their application to the disease.

reported to be involved in age-related memory impairment: these channels form an M-current that acts as a brake on neuronal excitability and their expression normally decreases in aging brains. In *Drosophila*, KCNQ overexpression in mushroom body neurons restores age-related memory impairment.[52] Other KCNQ-specific modulators have also been suggested to alleviate memory deficits associated with age-related memory diseases such as AD.[53] GCDPipe results also suggest KCNQ as a therapeutic target for AD and indicate that KCNQ-related mechanisms can be involved in the effect of some diuretics on cognitive health. Further studies of BACE1-KCNQ interaction mechanisms, including modulation of KCNQ with blood-brain barrier penetrating agents, might thus be of interest to target cognitive decline associated with AD.

The second most prioritized diuretic, furosemide, was already proposed as a probe molecule for the treatment of neuroinflammation in AD by Wang et al.[54] These authors conclude that furosemide has the capacity to downregulate the proinflammatory microglial M1 phenotype and upregulate the anti-inflammatory M2 phenotype, a potentially powerful and beneficial pharmacologic effect for inflammatory diseases, such as AD. In particular, they have shown that this drug inhibits cellular stress and proinflammatory cytokine production from LPS-stimulated SIM-A9 microglial cells and promotes their phagocytic activity, rescuing neuronal cells during Aβ-induced neuroinflammation and reducing LPS-induced endoplasmic reticulum (ER) stress.

The proposed molecular mechanism of action involves reduction of LPS-induced upregulation of ER stress marker genes, including Grp78, Atf4, Chop, tXbp1, and sXbp1. The gene targets with the highest GCDPipe scores (and which drive prioritization of this drug among diuretics), include ABCC4, ATP1A1, ADD1, CISD1, SLC12A2, and AASDHPPT, and it also happens to interact with the ACE gene, which is known to be implicated in AD based on GWASs (Figure 3D). Interestingly, ABCC4, the leading gene, is also reported to be involved in ER stress reactions.[55] GCDPipe thus allows us to link the furosemide activity observed by Wang et al. with its known drug targets.

The careful assessment of the overall quality of gene, cell type, or drug prioritization is impossible due to the incomplete knowledge of the disease-relevant instances. However, GCDPipe is intended as a user-friendly tool for specific disease applications to formulate testable hypotheses that would require further experimental validation. Thus, we provide a convenient and quick tool to rapidly integrate many types of data and obtain computational evidence to design future experiments, rather than a global solution to complex trait genetics challenges in the identification of disease genes and cell types.

The results of our analysis show that GCDPipe converges genetic risk factors for complex traits with findings on their pathophysiology at various hierarchical levels, and suggests the links between GWAS-derived disease genetic data with drug activity mechanisms.

Data and code availability

The code, installation and usage instructions, examples, and input data described in the case studies are available at https://github.com/ArtomovLab/GCDPipe.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.xhgg.2023.100203.

Acknowledgments

D.P. and A.S. were supported by the Ministry of Science and Higher Education of the Russian Federation (Priority 2030 Federal Academic Leadership Program).

A.L. was supported by the Ministry of Science and Higher Education of Russia (Award #075-15-2022-301).

M.A. was in part supported by the Nationwide Pediatric Innovation Fund.

Author contributions

D.P. and M.A. designed and conceived the study. D.P. and M.A. designed the computational algorithm. D.P., A.L., A.S., and M.A. analyzed the data. D.P. and M.A. wrote the manuscript. All authors reviewed and approved the manuscript.

Declaration of interests

The authors declare no competing interests.

Received: August 9, 2022
Accepted: April 25, 2023

References

1. Lau, A., and So, H.C. (2020). Turning genome-wide association study findings into opportunities for drug repositioning. Comput. Struct. Biotechnol. J. *18*, 1639–1650. https://doi.org/10.1016/j.csbj.2020.06.015.
2. Reay, W.R., and Cairns, M.J. (2021). Advancing the use of genome-wide association studies for drug repurposing. Nat. Rev. Genet. *22*, 658–671. https://doi.org/10.1038/s41576-021-00387-z.
3. Gamazon, E.R., Wheeler, H.E., Shah, K.P., Mozaffari, S.V., Aquino-Michaels, K., Carroll, R.J., Eyler, A.E., Denny, J.C., GTEx Consortium, Nicolae, D.L., et al. (2015). A gene-based association method for mapping traits using reference transcriptome data. Nat. Genet. *47*, 1091–1098. https://doi.org/10.1038/ng.3367.
4. Gusev, A., Ko, A., Shi, H., Bhatia, G., Chung, W., Penninx, B.W., Jansen, R., de Geus, E.J.C., Boomsma, D.I., Wright, F.A., et al. (2016). Integrative approaches for large-scale transcriptome-wide association studies. Nat. Genet. *48*, 245–252. https://doi.org/10.1038/ng.3506.
5. Fang, H., ULTRA-DD Consortium, De Wolf, H., Knezevic, B., Burnham, K.L., Osgood, J., Sanniti, A., Lledó Lara, A., Kasela, S., De Cesco, S., et al. (2019). A genetics-led approach defines the drug target landscape of 30 immune-related traits. Nat. Genet. *51*, 1082–1091. https://doi.org/10.1038/s41588-019-0456-1.
6. Sakaue, S., and Okada, Y. (2019). GREP: genome for REPositioning drugs. Bioinformatics *35*, 3821–3823. https://doi.org/10.1093/bioinformatics/btz166.
7. Ghoussaini, M., Mountjoy, E., Carmona, M., Peat, G., Schmidt, E.M., Hercules, A., Fumis, L., Miranda, A., Carvalho-Silva, D., Buniello, A., et al. (2021). Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. Nucleic Acids Res. *49*, D1311–D1320. https://doi.org/10.1093/nar/gkaa840.
8. Gaspar, H.A., Hübel, C., and Breen, G. (2019). Drug Targetor: a web interface to investigate the human druggome for over 500 phenotypes. Bioinforma Oxf Engl *35*, 2515–2517. https://doi.org/10.1093/bioinformatics/bty982.
9. Konuma, T., Ogawa, K., and Okada, Y. (2021). Integration of genetically regulated gene expression and pharmacological library provides therapeutic drug candidates. Hum. Mol. Genet. *30*, 294–304. https://doi.org/10.1093/hmg/ddab049.
10. Emon, M.A., Domingo-Fernández, D., Hoyt, C.T., and Hofmann-Apitius, M. (2020). PS4DR: a multimodal workflow for identification and prioritization of drugs based on pathway signatures. BMC Bioinf. *21*, 231. https://doi.org/10.1186/s12859-020-03568-5.
11. de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput. Biol. *11*, e1004219. https://doi.org/10.1371/journal.pcbi.1004219.
12. Hemani, G., Zheng, J., Elsworth, B., Wade, K.H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R., et al. (2018). The MR-Base platform supports systematic causal inference across the human phenome. Elife *7*, e34408. https://doi.org/10.7554/eLife.34408.
13. Reay, W.R., Atkins, J.R., Carr, V.J., Green, M.J., and Cairns, M.J. (2020). Pharmacological enrichment of polygenic risk for precision medicine in complex disorders. Sci. Rep. *10*, 879. https://doi.org/10.1038/s41598-020-57795-0.
14. Ballouz, S., Liu, J.Y., George, R.A., Bains, N., Liu, A., Oti, M., Gaeta, B., Fatkin, D., and Wouters, M.A. (2013). GentrepidV2.0: a web server for candidate disease gene prediction. BMC Bioinf. *14*, 249. https://doi.org/10.1186/1471-2105-14-249.
15. Kolosov, N., Daly, M.J., and Artomov, M. (2021). Prioritization of disease genes from GWAS using ensemble-based positive-unlabeled learning. Eur. J. Hum. Genet. *29*, 1527–1535. https://doi.org/10.1038/s41431-021-00930-w.
16. Isakov, O., Dotan, I., and Ben-Shachar, S. (2017). Machine learning-based gene prioritization identifies novel candidate risk genes for inflammatory bowel disease. Inflamm. Bowel Dis. *23*, 1516–1523. https://doi.org/10.1097/MIB.0000000000001222.
17. Ritprajak, P., Hayakawa, M., Sano, Y., Otsu, K., and Park, J.M. (2012). Cell type-specific targeting dissociates the therapeutic from the adverse effects of protein kinase inhibition in allergic skin disease. Proc. Natl. Acad. Sci. USA *109*, 9089–9094. https://doi.org/10.1073/pnas.1202984109.

18. Mathew, B., and Parambi, D.G.T. (2020). Principles of Neurochemistry (Springer). https://link.springer.com/book/10.1007/978-981-15-5167-3.
19. Watanabe, K., Umićević Mirkov, M., de Leeuw, C.A., van den Heuvel, M.P., and Posthuma, D. (2019). Genetic mapping of cell type specificity for complex traits. Nat. Commun. *10*, 3222. https://doi.org/10.1038/s41467-019-11181-1.
20. Finucane, H.K., Reshef, Y.A., Anttila, V., Slowikowski, K., Gusev, A., Byrnes, A., Gazal, S., Loh, P.R., Lareau, C., Shoresh, N., et al. (2018). Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nat. Genet. *50*, 621–629. https://doi.org/10.1038/s41588-018-0081-4.
21. Jagadeesh, K.A., Dey, K.K., Montoro, D.T., et al. (2021). Identifying disease-critical cell types and cellular processes across the human body by integration of single-cell profiles and human genetics. Preprint at bioRxiv. https://doi.org/10.1101/2021.03.19.436212.
22. Calderon, D., Bhaskar, A., Knowles, D.A., Golan, D., Raj, T., Fu, A.Q., and Pritchard, J.K. (2017). Inferring relevant cell types for complex traits by using single-cell gene expression. Am. J. Hum. Genet. *101*, 686–699. https://doi.org/10.1016/j.ajhg.2017.09.009.
23. Lundberg, S.M., and Lee, S.I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems, *Vol 30*, I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds. *(Curran Associates, Inc.), pp. 1–10. https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html*.
24. Transcriptomics Explorer. Allen brain atlas: cell types. https://celltypes.brain-map.org/rnaseq/human_ctx_smart-seq?selectedVisualization=Heatmap&colorByFeature=Cell+Type&colorByFeatureValue=GAD1.
25. Singh, S., Singh, T.G., Daly, M.J.; and on behalf of the Schizophrenia Exome Meta-Analysis (SCHEMA) Consortium (2020). Exome sequencing identifies rare coding variants in 10 genes which confer substantial risk for schizophrenia. Curr. Neuropharmacol. *18*, 918–935. https://doi.org/10.1101/2020.09.18.20192815.
26. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30. https://doi.org/10.1093/nar/28.1.27.
27. Matamales, M. (2021). How dopamine leads to hallucinations. Science *372*, 33–34. https://doi.org/10.1126/science.abh1310.
28. Conn, K.A., Burne, T.H.J., and Kesby, J.P. (2020). Subcortical dopamine and cognition in schizophrenia: looking beyond psychosis in preclinical models. Front. Neurosci. *14*, 542. https://www.frontiersin.org/article/10.3389/fnins.2020.00542.
29. Saito, A., Ballinger, M.D.L., Pletnikov, M.V., Wong, D.F., and Kamiya, A. (2013). Endocannabinoid system: potential novel targets for treatment of schizophrenia. Neurobiol. Dis. *53*, 10–17. https://doi.org/10.1016/j.nbd.2012.11.020.
30. Volk, D.W., and Lewis, D.A. (2016). The role of endocannabinoid signaling in cortical inhibitory neuron dysfunction in schizophrenia. Biol. Psychiatry *79*, 595–603. https://doi.org/10.1016/j.biopsych.2015.06.015.
31. Minichino, A., Senior, M., Brondino, N., Zhang, S.H., Godwlewska, B.R., Burnet, P.W.J., Cipriani, A., and Lennox, B.R. (2019). Measuring disturbance of the endocannabinoid system in psychosis: a systematic review and meta-analysis. JAMA Psychiatr. *76*, 914–923. https://doi.org/10.1001/jamapsychiatry.2019.0970.
32. Leweke, F.M., Piomelli, D., Pahlisch, F., Muhl, D., Gerth, C.W., Hoyer, C., Klosterkötter, J., Hellmich, M., and Koethe, D. (2012). Cannabidiol enhances anandamide signaling and alleviates psychotic symptoms of schizophrenia. Transl. Psychiatry *2*, e94. https://doi.org/10.1038/tp.2012.15.
33. Black, M.D., Stevens, R.J., Rogacki, N., Featherstone, R.E., Senyah, Y., Giardino, O., Borowsky, B., Stemmelin, J., Cohen, C., Pichat, P., et al. (2011). AVE1625, a cannabinoid CB1 receptor antagonist, as a co-treatment with antipsychotics for schizophrenia: improvement in cognitive function and reduction of antipsychotic-side effects in rodents. Psychopharmacology (Berl) *215*, 149–163. https://doi.org/10.1007/s00213-010-2124-0.
34. de Leon, J., and Diaz, F.J. (2005). A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. Schizophr. Res. *76*, 135–157. https://doi.org/10.1016/j.schres.2005.02.010.
35. Miller, J.A., Gouwens, N.W., Tasic, B., Collman, F., van Velthoven, C.T., Bakken, T.E., Hawrylycz, M.J., Zeng, H., Lein, E.S., and Bernard, A. (2020). Common cell type nomenclature for the mammalian brain. eLife *9*, e59928. https://doi.org/10.7554/eLife.59928.
36. Boden, E.K., and Lord, J.D. (2017). CD4 T cells in IBD: crossing the line? Dig. Dis. Sci. *62*, 2208–2210. https://doi.org/10.1007/s10620-017-4655-2.
37. Smids, C., Horjus Talabur Horje, C.S., Drylewicz, J., Roosenboom, B., Groenen, M.J.M., van Koolwijk, E., van Lochem, E.G., and Wahab, P.J. (2018). Intestinal T cell profiling in inflammatory bowel disease: linking T cell subsets to disease activity and disease course. J. Crohns Colitis *12*, 465–475. https://doi.org/10.1093/ecco-jcc/jjx160.
38. Hansen, D.V., Hanson, J.E., and Sheng, M. (2018). Microglia in Alzheimer's disease. J. Cell Biol. *217*, 459–472. https://doi.org/10.1083/jcb.201709069.
39. Price, A.J., Jaffe, A.E., and Weinberger, D.R. (2021). Cortical cellular diversity and development in schizophrenia. Mol. Psychiatry *26*, 203–217. https://doi.org/10.1038/s41380-020-0775-8.
40. Volk, D.W., Edelson, J.R., and Lewis, D.A. (2014). Cortical inhibitory neuron disturbances in schizophrenia: role of the ontogenetic transcription factor Lhx6. Schizophr. Bull. *40*, 1053–1061. https://doi.org/10.1093/schbul/sbu068.
41. Volk, D.W., Matsubara, T., Li, S., Sengupta, E.J., Georgiev, D., Minabe, Y., Sampson, A., Hashimoto, T., and Lewis, D.A. (2012). Deficits in transcriptional regulators of cortical parvalbumin neurons in schizophrenia. Am. J. Psychiatry *169*, 1082–1091. https://doi.org/10.1176/appi.ajp.2012.12030305.
42. Yuan, F., Fang, K.H., Hong, Y., Xu, S.B., Xu, M., Pan, Y., and Liu, Y. (2020). LHX6 is essential for the migration of human pluripotent stem cell-derived GABAergic interneurons. Protein Cell *11*, 286–291. https://doi.org/10.1007/s13238-019-00686-6.
43. Volk, D.W., Sampson, A.R., Zhang, Y., Edelson, J.R., and Lewis, D.A. (2016). Cortical GABA markers identify a molecular subtype of psychotic and bipolar disorders. Psychol. Med. *46*, 2501–2512. https://doi.org/10.1017/S0033291716001446.
44. Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., et al. (2018). DrugBank 5.0: a major update to the DrugBank

database for 2018. Nucleic Acids Res. *46*, D1074–D1082. https://doi.org/10.1093/nar/gkx1037.

45. Taubes, A., Nova, P., Zalocusky, K.A., Kosti, I., Bicak, M., Zilberter, M.Y., Hao, Y., Yoon, S.Y., Oskotsky, T., Pineda, S., et al. (2021). Experimental and real-world evidence supporting the computational repurposing of bumetanide for APOE4-related Alzheimer's disease. Nat. Aging *1*, 932–947. https://doi.org/10.1038/s43587-021-00122-7.
46. Chuang, Y.F., Breitner, J.C.S., Chiu, Y.L., Khachaturian, A., Hayden, K., Corcoran, C., Tschanz, J., Norton, M., Munger, R., Welsh-Bohmer, K., et al. (2014). Use of diuretics is associated with reduced risk of Alzheimer's disease: the Cache County Study. Neurobiol. Aging *35*, 2429–2435. https://doi.org/10.1016/j.neurobiolaging.2014.05.002.
47. DeLoach, T., and Beall, J. (2018). Diuretics: a possible keystone in upholding cognitive health. Ment. Health Clin. *8*, 33–40. https://doi.org/10.9740/mhc.2018.01.033.
48. Lu, M., Ma, L., Wang, X., Jiang, W., and Shan, P. (2017). Indapamide suppresses amyloid-β production in cellular models of Alzheimer's disease through regulating BACE1 activity. Int. J. Clin. Exp. Med. *10*, 5922–5930.
49. Lu, M., Ma, L., Wang, X., Jiang, W., and Shan, P. (2017). Indapamide suppresses amyloid-B production in cellular models of Alzheimer's disease through regulating BACE1 activity. Int. J. Clin. Exp. Med. *10*, 5922–5930.
50. Sachse, C.C., Kim, Y.H., Agsten, M., Huth, T., Alzheimer, C., Kovacs, D.M., and Kim, D.Y. (2013). BACE1 and presenilin/γ-secretase regulate proteolytic processing of KCNE1 and 2, auxiliary subunits of voltage-gated potassium channels. FASEB J. *27*, 2458–2467. https://doi.org/10.1096/fj.12-214056.
51. Agsten, M., Hessler, S., Lehnert, S., Volk, T., Rittger, A., Hartmann, S., Raab, C., Kim, D.Y., Groemer, T.W., Schwake, M., et al. (2015). BACE1 modulates gating of KCNQ1 (Kv7.1) and cardiac delayed rectifier KCNQ1/KCNE1 (IKs). J. Mol. Cell. Cardiol. *89*, 335–348. https://doi.org/10.1016/j.yjmcc.2015.10.006.
52. Cavaliere, S., Malik, B.R., and Hodge, J.J.L. (2013). KCNQ channels regulate age-related memory impairment. PLoS One *8*, e62445. https://doi.org/10.1371/journal.pone.0062445.
53. Barrese, V., Stott, J.B., and Greenwood, I.A. (2018). KCNQ-encoded potassium channels as therapeutic targets. Annu. Rev. Pharmacol. Toxicol. *58*, 625–648. https://doi.org/10.1146/annurev-pharmtox-010617-052912.
54. Wang, Z., Vilekar, P., Huang, J., and Weaver, D.F. (2020). Furosemide as a probe molecule for the treatment dof neuroinflammation in Alzheimer's disease. ACS Chem. Neurosci. *11*, 4152–4168. https://doi.org/10.1021/acschemneuro.0c00445.
55. Usuki, F., Fujimura, M., and Yamashita, A. (2017). Endoplasmic reticulum stress preconditioning modifies intracellular mercury content by upregulating membrane transporters. Sci. Rep. *7*, 12390. https://doi.org/10.1038/s41598-017-09435-3.