Data Article

Data on overlapping brain disorders and emerging drug targets in human Dopamine Receptors Interaction Network

Avijit Podder, N. Latha*

Bioinformatics Infrastructure Facility, Sri Venkateswara College (University of Delhi), Benito Juarez Road, Dhaula Kuan, 110021 New Delhi, India

ARTICLE INFO

Article history:
Received 3 January 2017
Received in revised form 13 February 2017
Accepted 4 April 2017
Available online 8 April 2017

Keywords:
Dopamine receptors
Interactions network
Brain disorders
Drug targets

ABSTRACT

Intercommunication of Dopamine Receptors (DRs) with their associate protein partners is crucial to maintain regular brain function in human. Majority of the brain disorders arise due to malfunctioning of such communication process. Hence, contributions of genetic factors, as well as phenotypic indications for various neurological and psychiatric disorders are often attributed as sharing in nature. In our earlier research article entitled “Human Dopamine Receptors Interaction Network (DRIN): a systems biology perspective on topology, stability and functionality of the network” (Podder et al., 2014) [1], we had depicted a holistic interaction map of human Dopamine Receptors. Given emphasis on the topological parameters, we had characterized the functionality along with the vulnerable properties of the network. In support of this, we hereby provide an additional data highlighting the genetic overlapping of various brain disorders in the network. The data indicates the sharing nature of disease genes for various neurological and psychiatric disorders in dopamine receptors connecting protein-protein interactions network. The data also indicates toward an alternative approach to prioritize proteins for overlapping brain disorders as valuable drug targets in the network.

© 2017 Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

* Corresponding author.
E-mail address: lata-bic@svc.ac.in (N. Latha).

http://dx.doi.org/10.1016/j.dib.2017.04.001
2352-3409/© 2017 Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Specifications Table

| Subject area                  | Computational Systems Biology |
|-------------------------------|------------------------------|
| More specific subject area    | Human Dopamine Receptors Network |
| Type of data                  | Text files, tables and figures |
| How data was acquired         | Systemic database curation and statistical analysis |
| Data format                   | Analyzed                     |
| Experimental factors          | Common genetic factors responsible for various brain disorders in the network |
| Experimental features         | Comprehensive mapping of disease genes and identification of overlapping disease modules in the network |
| Data source location          | n/a                          |
| Data accessibility            | Data are within this article  |

Value of the data

- The data hints an overlapping disease spectrum of neurological disorders in DRIN.
- The data pinpoint towards common contributing genetic factors of neurological disorders in DRIN.
- The data might encourage researchers to identify additional drug targets in DRIN for future drug discovery endeavors.

1. Data

The data highlighted 25848 associations between 431 proteins and 4312 diseases, disorders and clinical or abnormal human phenotypes in Dopamine Receptors Interactions Network (DRIN) [1]. The associations were refined through stringent statistical analysis to reduce the over representation of

![Gene-disease enrichment analysis in DRIN](image)

**Fig. 1.** Gene-disease enrichment analysis in DRIN. In the plot the X-axis represents the percentage of genes associated with each disease term in DRIN and Y-axis represents the enrichment statistics ($P$-value cut off $< 0.05$) of disease terms in DRIN. A total of 143 different disease terms were significantly identified (blue dots) over 4312 disease association in DRIN after bonferroni correction of the $P$-value (two-sided hypergeometric test).
disease information in DRIN (Table S1). Further, related disease terms were combined together and significant associations of 29 disease modules in the DRIN were obtained (Figs. 1–3). The disease modules were classified based on the distributions of overlapping genetic factors (Figs. 4–7). In addition, the common contribution factors for different brain disorders were characterized based on functional enrichment analysis to pinpoint additional drug targets in human DRIN (Figs. 8–10). Data were also tabulated to get more insights into drug availability for those targets (Fig. 11).
2. Experimental design, materials and methods

2.1. Disease enrichment data analysis

Genes were mapped to respective disease data using DisGeNet database [2]. The disease data indicates an association of diseases, disorders and clinical or abnormal human phenotypes with a given set of gene list. The data covers different biomedical aspects of each disease from several sources.

To identify significant enrichment of diseases in our data two-sided hypergeometric test [3] were carried out. The \( P \)-values were further corrected by the bonferroni correction (\( P < 0.05 \)) to determine significant enrichment of diseases in our data two-sided hypergeometric test [3] were carried out. The \( P \)-values were further corrected by the bonferroni correction (\( P < 0.05 \)) to determine
a list of significant disease association in our data.

\[ P(X = x) = f(x) = \frac{(m/x)(N - m/n - x)}{(N/n)} \]

where

- \( N \) = Total number of genes in DisGeNet;
- \( n \) = Total number of genes in DRIN;
- \( m \) = Each disease associated genes in DisGeNet;
- \( x \) = Each disease associated genes in DRIN

2.2. Disease overlapping data analysis

Overlapping of genes between two diseases was measured using Jaccard or Tanimoto coefficient metrics \([4]\) which compared the similarity and diversity of disease sets in our data. It uses the ratio of
Fig. 7. Connectivity map within disease modules in neurological disorder cluster. Disease modules are connected to each other based on the overlapping ratio cutoff $<0.35$ as derived from Jaccard or Tanimoto coefficient metrics.

Fig. 8. Common set of genes within disease modules in neurological disorder cluster. Panel [A] highlights six genes (BDNF, DRD2, GAD1, GRIN2A, MAOA and MTHFR) which are common contributing factors for all eight disease modules in neurological disorder cluster (Fig. 7). Panel [B] shows pathway enrichment and functional classification of genes in DRIN. Panel [C] depicts Transcription Factors (TFs) binding information highlighting regulatory classification of genes in the network.
the intersecting set to the union set as the measure of similarity. Thus it equals to zero if there are no intersecting elements and equals to one if all elements intersect.

\[ T = \frac{N_c}{N_a + N_b - N_c} \]

where \( N_a \) = number of elements in set A; \( N_b \) = number of elements in set B; \( N_c \) = number of elements in intersecting set

Based on the similarity matrix, the disease sets were further classified into different clusters using Neighbor Joining (NJ) algorithm which is a bottom-up or agglomerative clustering method to identify similar group of elements in the data [5]. The Phylip tool [6] was used to implement the NJ method in our data.

2.3. Functional classification of data

The pathways information and transcription factor (TF) regulatory data were collected from the Molecular Signatures Database (MSigDB) [7]. Regulatory data is represented as a bipartite graph between a set of genes and transcription factors. Data for expression profiling of TFs in different brain tissues were obtained from GTEx portal web server [8].

2.4. Drug data analysis

A comprehensive search was made in DrugBank v4.0 (http://www.drugbank.ca/) database [9] to identify the known drugs related to our data set. Data was also obtained from a recent study by
Fig. 10. Tissue specific expression profiling of genes regulated by TF-Sp1 in human brain. Expression profiling of four genes and TF-SP1 in 13 different human brain tissues under normal physiological condition. Gene expression data were implemented through GTEx portal web server [8] where expression values are shown in RPKM (Reads Per Kilobase of transcript per Million mapped reads). Box plots are shown as median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range.
Rask-Andersen [10], to arrange the genes based on approved, investigational, not approved and withdrawn drug information.

Acknowledgements

This work is supported by Department of Biotechnology (DBT) BT/BI/25/010/2012(BIF-SVCND), Ministry of Science & Technology, Government of India under the DBT-BTISNET programme.

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2017.04.001.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2017.04.001.

References

[1] A. Podder, N. Jatana, N. Latha, Human Dopamine Receptors Interaction Network (DRIN): a systems biology perspective on topology, stability and functionality of the network, J Theor. Biol. 357 (2014) 169–183.
[2] A. Bauer-Mehren, M. Bundschus, M. Rautschka, M.A. Mayer, F. Sanz, L.I. Furlong, Gene-disease network analysis reveals functional modules in mendelian, complex and environmental diseases, PLoS One 6 (6) (2011) e20284.
[3] I. Rivals, L. Personnaz, L. Taing, M.C. Potier, Enrichment or depletion of a GO category within a class of genes: which test? Bioinformatics 23 (2007) 401–407.
[4] Toby Segaran, Programming Collective Intelligence: Building Smart Web 2.0 Applications, O'Reilly Media, Sebastopol, CA, 2007.
[5] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 4 (4) (1987) 406–425.
[6] J. Felsenstein, PHYLIP (Phylogeny Inference Package) Version 3.6 Distributed by the author, Department of Genome Sciences, University of Washington, 2005 (Seattle).
[7] Tamayo Subramanian, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, PNAS 102 (2005) 15545–15550.

[8] L.J. Carithers, H.M. Moore, The Genotype-Tissue Expression (GTEx) Project, Biopreservation Biobanking 13 (5) (2015) 307–308.

[9] D.S. Wishart, C. Knox, A.C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, J. Woolsey, DrugBank: a comprehensive resource for in silico drug discovery and exploration, Nucl. Acids Res. (2006) 668–672 1;34(Database issue):D.

[10] M. Rask-Andersen, S. Masuram, H.B. Schüth, The druggable genome: evaluation of drug targets in clinical trials suggests major shifts in molecular class and indication, Annu. Rev. Pharmacol. Toxicol. 54 (2013) 9–26.