BrainBase: a curated knowledgebase for brain diseases

Lin Liu [1,2,†], **Yang Zhang**[1,2,3,†], **Guangyi Niu**[1,2,3,†], **Qianpeng Li**[1,2,3], **Zhao Li**[1,2,3], **Tongtong Zhu**[1,2,3], **Changrui Feng** [1,2,3], **Xiaonan Liu**[1,2,3], **Yuansheng Zhang**[1,2,3], **Tianyi Xu**[1,2], **Ruru Chen**[1,2,3], **Xufei Teng**[1,2,3], **Rongqin Zhang**[1,2,3], **Dong Zou**[1,2], **Lina Ma** [1,2,3,*] **and Zhang Zhang** [1,2,3,*]

[1]National Genomics Data Center & CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China, [2]China National Center for Bioinformation, Beijing 100101, China and [3]University of Chinese Academy of Sciences, Beijing 100049, China

ABSTRACT

Brain is the central organ of the nervous system and any brain disease can seriously affect human health. Here we present BrainBase (https://ngdc.cncb.ac.cn/brainbase), a curated knowledgebase for brain diseases that aims to provide a whole picture of brain diseases and associated genes. Specifically, based on manual curation of 2768 published articles along with information retrieval from several public databases, BrainBase features comprehensive collection of 7175 disease–gene associations spanning a total of 123 brain diseases and linking with 5662 genes, 16 591 drug–target interactions covering 2118 drugs/chemicals and 623 genes, and five types of specific genes in light of expression specificity in brain tissue/regions/cerebrospinal fluid/cells. In addition, considering the severity of glioma among brain tumors, the current version of BrainBase incorporates 21 multi-omics datasets, presents molecular profiles across various samples/conditions and identifies four groups of glioma featured genes with potential clinical significance. Collectively, Brain-Base integrates not only valuable curated disease–gene associations and drug–target interactions but also molecular profiles through multi-omics data analysis, accordingly bearing great promise to serve as a valuable knowledgebase for brain diseases.

INTRODUCTION

Brain is the central organ of the nervous system, not only controlling thoughts, memory, speech, and movement, but also regulating the function of many organs (1). Any brain disease that alters brain function or structure, can seriously affect human health. For instance, glioma, one of the most common types of brain tumors, represents ∼ 80% of malignant brain tumors and exhibits low resection rate and high recurrence risk (2,3). Nowadays, increasing brain-related projects have been launched throughout the world, e.g. Human Brain Project in Europe (4), BRAIN Initiative in US (5), Brain/MINDS in Japan (6), China Brain Project (7), etc., with the aim to deepen our understanding of brain diseases, structure and function and to accelerate brain-derived applications in health, computing, and technology.

Particularly, powered by high-throughput sequencing technologies, great efforts have been devoted to deciphering complex associations between genes and brain diseases from multiple omics levels (2,8–11). However, these valuable disease–gene associations (as well as omics datasets) are scattered in massive scientific publications, thus making their retrieval, integration and visualization very arduous and time-consuming. Although several related resources have already been developed with different purposes, there still lacks a comprehensive resource dedicated for brain diseases that includes high-quality associations extracted from published literatures. Specifically, the Developmental Brain Disorders Database (DBDB, https://www.dbdb.urmc.rochester.edu) aims to provide a number of genes, phenotypes, and syndromes associated only with neurodevelopmental disorders (12), the Ivy Glioblastoma Atlas Project (Ivy GAP, http://glioblastoma.alleninstitute.org) focuses on the anatomic and genetic basis of glioblastoma at the cellular and molecular levels (13), the Chinese Glioma Genome Atlas (CGGA, http://www.cgga.org.cn) is devoted to providing open access to glioma datasets and offering interactive visualization of multi-omics profiles (14), Glio-Vis (http://gliovis.bioinfo.cnio.es) is a web application for

*To whom correspondence should be addressed. Tel: +86 10 84097261; Fax: +86 10 84097720; Email: zhangzhang@big.ac.cn
Correspondence may also be addressed to Lina Ma. Email: malina@big.ac.cn
†The authors wish it to be known that, in their opinion, first three authors should be regarded as Joint First Authors.

visualization and analysis of brain tumors' expression datasets (15) and STAB (http://stab.comp-sysbio.org) is a spatio-temporal cell atlas with the aim to understand both the etiology of neuropsychiatric disorders and the development of the normal human brain (16). Despite this, a comprehensive collection of high-quality disease–gene associations covering a diversity of brain diseases is highly desirable.

Here, we present BrainBase (https://ngdc.cncb.ac.cn/brainbase), a curated knowledgebase for brain diseases that is dedicated to the curation, integration and visualization of brain diseases-related knowledge. Unlike extant relevant databases, BrainBase features comprehensive collection of high-quality disease–gene associations, drug–target interactions and multi-omics datasets. Accordingly, BrainBase not only integrates a large number of valuable knowledge associations and interactions manually curated from published literatures but also includes a wide range of molecular profiles through multi-omics data analysis, thus bearing great promise to serve as a valuable knowledgebase for uncovering molecular mechanisms underlying the progression of brain diseases.

MATERIALS AND METHODS

Knowledge curation and integration

Disease-gene associations were extracted from 2768 published literatures, with literature curation on 510 glioma publications due to its severity among brain tumors and information retrieval from related databases, including DBDB (12), Brain Disease Knowledgebase (http://birg1.fbb.utm.my/bddb), HMDD (17), LncBook (18), EWAS Atlas (19), GWAS Catalog (20), EDK (21) and CR-Marker (22). Drug-target interactions were collected from ChEMBL (23), DrugBank (24), CMap (25), PharmGKB (26) and CTD (27). In the current version of BrainBase, it was found that drugs were designed for 8 common brain diseases, including Alzheimer's disease, Parkinson's disease, glioma, glioblastoma, epilepsy, autism spectrum disorder, medulloblastoma and cerebral palsy.

To unify disease name and definition, terms or identifiers of Disease Ontology (DO) were retrieved and mapped to the corresponding diseases (28). In order to provide consistent names for all collected genes in BrainBase, gene names were unified with the help of the gene symbol-alias conversion table from the HGNC database (2021.4.23 version) (29).

Data collection

Multi-omics datasets were obtained from several well-known public databases, including GEO (30) TCGA (31), GTEx (32), HPA (33), Ivy GAP (13) and CGGA (14). Totally, 21 open-access datasets were included in BrainBase and publicly available without restrictions at ftp://download.big.ac.cn/brainbase. Detailed information about these datasets can be accessible at https://ngdc.cncb.ac.cn/brainbase/faq. Among these datasets, GSE50161, GSE59612, GSE4290, GSE111260, GSE36278, GSE60274, GSE61160, GSE50923, CGGA_301, CGGA_325, CGGA_693, CGGA_methylation, TCGA_expression, TCGA_methylation and TCGA_CNV were used to identify glioma featured genes.

Computational identification of specific genes

We identified genes specifically expressed in brain tissue, regions and cells, as well as genes that are detected in the special 'tissue' cerebrospinal fluid (CSF): (i) brain-specific genes: a dataset from GTEx v7 (32) (15 January 2016) containing 11 688 samples across 53 tissue sites was used for identification of brain-specific genes. Since tissues may have multiple different sites, gene expression levels were averaged over sites that were from the same tissue. To reduce background noise, genes with maximum expression level smaller than 1 TPM (Transcripts Per Million) were removed. Finally, based on the expression profiles across 31 tissues, we calculated the tissue specificity index τ for each gene to identify tissue-specific genes (34,35). In this study, brain-specific genes were defined as those genes that are specifically expressed in brain with $\tau > 0.9$. As a consequence, a list of 639 brain-specific genes were identified, including 475 mRNAs, 2 miRNAs, 127 lncRNAs and 4 pseudogenes; (ii) brain-region-specific genes: based on the GTEx dataset, 546 brain-region-specific genes were identified with the tissue specificity index τ, expressed specifically in 12 brain regions, including cerebellum, cerebellar hemisphere, spinal cord (cervical c-1), hypothalamus, nucleus accumbens (basal ganglia), substantia nigra, frontal cortex (BA9), cortex, hippocampus, putamen (basal ganglia), amygdala and caudate (basal ganglia); (iii) brain-specific proteins: a list of 215 proteins specifically expressed in brain were obtained from the HPA database (33); (iv) cell markers: a total of 328 brain cell marker genes were collected from CellMarker (36) and Cell Taxonomy (https://ngdc.cncb.ac.cn/celltaxonomy); and (v) CSF proteins: 1126 proteins detected in CSF with their fluorescence intensity were obtained from GEO (GSE83710) (37,38).

Identification of glioma featured genes

Glioma, a severe brain tumor, represents ∼80% of malignant brain tumors, can be classified into low-grade glioma (LGG) and glioblastoma multiforme (GBM) (39). Powered by high-throughput sequencing technologies, a set of molecular biomarkers have been found to benefit glioma diagnosis and prognosis; among them, isocitrate dehydrogenase (*IDH*) mutation and 1p/19q co-deletion (codel) are two most important genetic events (35,40). Since it was reported that glioma is related to other brain diseases (41,42), identification of glioma featured genes by comparing the following four pairs, namely, Glioma vs Normal, GBM versus LGG, *IDH* wildtype vs *IDH* mutation, and 1p19q noncodel vs 1p19q codel, is of great significance for studying glioma and other brain diseases. Therefore, to identify glioma featured genes, student's t-test was performed at multi-omics levels and the *P*-values were adjusted by the false discovery rate (FDR) described by Benjamini and Hochberg (43), namely, four pairs (as mentioned above) at the expression level, 4 pairs (as mentioned above) at the DNA methylation level (promoter region) and 3 pairs (except Glioma vs Normal) at the CNV level, which were denoted as E4, M4 and C3, respectively. As a result, four

groups of featured genes were identified: (i) ubiquitously differential genes (UDGs): genes exhibit significant differences ubiquitously in E4, M4 and C3 (P-value < 0.001, FDR < 0.01); (ii) differentially expressed genes (DEGs): genes exhibit significant differences in E4 (P-value < 0.001, FDR < 0.01); (iii) differentially DNA-methylated genes (DMGs): genes exhibit significant differences in M4 (P-value < 0.001, FDR < 0.01) and (iv) differential CNV genes (DCGs): genes exhibit significant differences in C3 (P-value < 0.001, FDR < 0.01).

Database construction and web interface implementation

BrainBase was built based on Apache Tomcat Server (http://tomcat.apache.org) and MySQL (http://www.mysql.org). Web interfaces were developed by JSP, HTML5, CSS3, AJAX, JQuery and BootStrap (version 3.3.7). Omics profiles were visualized by HTML widgets and R packages including plumber (version 1.1.0), ggplot2 (version 3.3.4), ggpubr (version 0.4.0), survminer (version 0.4.9) and survival (version 3.2). ECharts (version 4.1.0) was also adopted to generate interactive charts.

DATABASE CONTENTS AND USAGE

BrainBase presents a comprehensive collection of high-quality knowledge entries including disease–gene associations and drug–target interactions, houses five types of specific genes in terms of expression specificity, incorporates multi-omics molecular profiles across various samples/conditions, and identifies four groups of glioma featured genes with potential significance in clinical application (Figure 1).

Disease-gene associations

Given great efforts that have been devoted to the identification of genes associated with brain diseases, it is critical to build a comprehensive view of disease–gene associations across diverse brain diseases (44,45). Hence, based on literature curation and integration with controlled vocabularies (see details in Materials and Methods), the current version of BrainBase houses a total of 7175 disease–gene associations spanning a total of 123 brain diseases and linking with 5662 genes. Based on this, BrainBase is capable to capture a whole picture of brain diseases and associated genes (Figure 2). Notably, the top five diseases by associations, according to the current accumulation of brain-related studies, are Alzheimer's disease, autism spectrum disorder, glioma, multiple sclerosis and Parkinson's disease (Figure 2A). In addition, BrainBase provides a landscape of hot genes (with different types, including mRNA, miRNA and lncRNA) that are closely associated with brain diseases. Strikingly, *hsa-mir-146* (miRNA), *GRIA2* (mRNA) and *MEG3* (lncRNA) are representative genes associated with 13, 7 and 9 brain diseases, respectively (Figure 2B).

Considering that glioma is highly malignant among brain diseases and associated with either aberrant activities of genes or abnormal upstream regulators/downstream targets (46,47), here we propose a curation model with particular focus on glioma to standardize the curation process and establish controlled vocabularies and descriptive terms (that are abstracted and categorized based on literature curation) to depict each item (for details see https://ngdc.cncb.ac.cn/brainbase/faq). Based on this model, a wide range of important items are curated from publications, including molecular role (e.g. gene type, omics information, tumor process, pathway, description and PMID), regulation axis (e.g., regulator type/name/effect and target type/name/effect) and experimental sample (e.g. subtype, grade, species and tissue/cell line). As a consequence, among all the 7175 associations, 656 glioma entries are further curated and obtained, which are associated with 363 genes; the top ten genes in terms of disease–gene associations are *IDH1*, *EGFR*, *MGMT*, *TP53*, *IDH2*, *PTEN*, *FGFR1*, *PDGFRA*, *EGFRVIII* and *H3-3A* (Figure 2C), consistent well with previous findings (11,48–49) that these genes are closely related with glioma. Additionally, Brain-Base obtains a whole picture of glioma pathways and processes; the top five pathways are AKT, Wnt/β-catenin, Purine Biosynthesis, PI3K/AKT/mTOR and Notch (Figure 2D) and the top five processes are proliferation, invasion, suppression, aggressiveness and growth (Figure 2E), potentially providing an integrated landscape of underlying etiology of glioma.

Drug–target interactions

Accumulating evidence have shown that drug targeted genes can be involved in multiple diseases (50,51). Therefore, BrainBase incorporates a large number of 16,591 drug–target interactions involving 2118 drugs/chemicals and 623 genes and focusing on 8 common diseases (see details in Materials and Methods). For example, *TNF*, a gene encoding a proinflammatory cytokine (52), is a targeted gene shared by glioma, Parkinson's disease and Alzheimer's disease, and *BDNF*, one of the neurotrophic factors (53), is a targeted gene shared by autism spectrum disorder, Alzheimer's disease, epilepsy and Parkinson's disease (Figure 2F), implying the potential of these core targeted genes as therapeutic opportunities. In addition, considering that the trend of drug design is from single target to multiple targets (54,55), we find that in BrainBase, the collected drugs/chemicals overall have six targets on average. Among them, noticeably, afimoxifene, associating with the seven diseases, interacts with three genes in autism spectrum disorder, 14 genes in Alzheimer's disease, seven genes in epilepsy, eight genes in Parkinson's disease, seveb genes in medulloblastoma, 12 genes in glioblastoma and 13 genes in glioma (https://ngdc.cncb.ac.cn/brainbase/drugs). Collectively, BrainBase features comprehensive integration of drug–target interactions, yielding a set of genes for drug target selection in brain diseases.

Potential significance of specific genes

Specifically expressed genes at multi-omics levels are likely associated with diseases and targeted by drugs with potential clinical application (56–58). Thus, through computational identification based on the GTEx v7 (15 January 2016) dataset, BrainBase provides five types of specific genes, namely, 639 brain-specific genes specifically expressed in brain tissue (RNA level), 546 brain-region-specific genes in 12 brain regions (RNA level), 328 cell
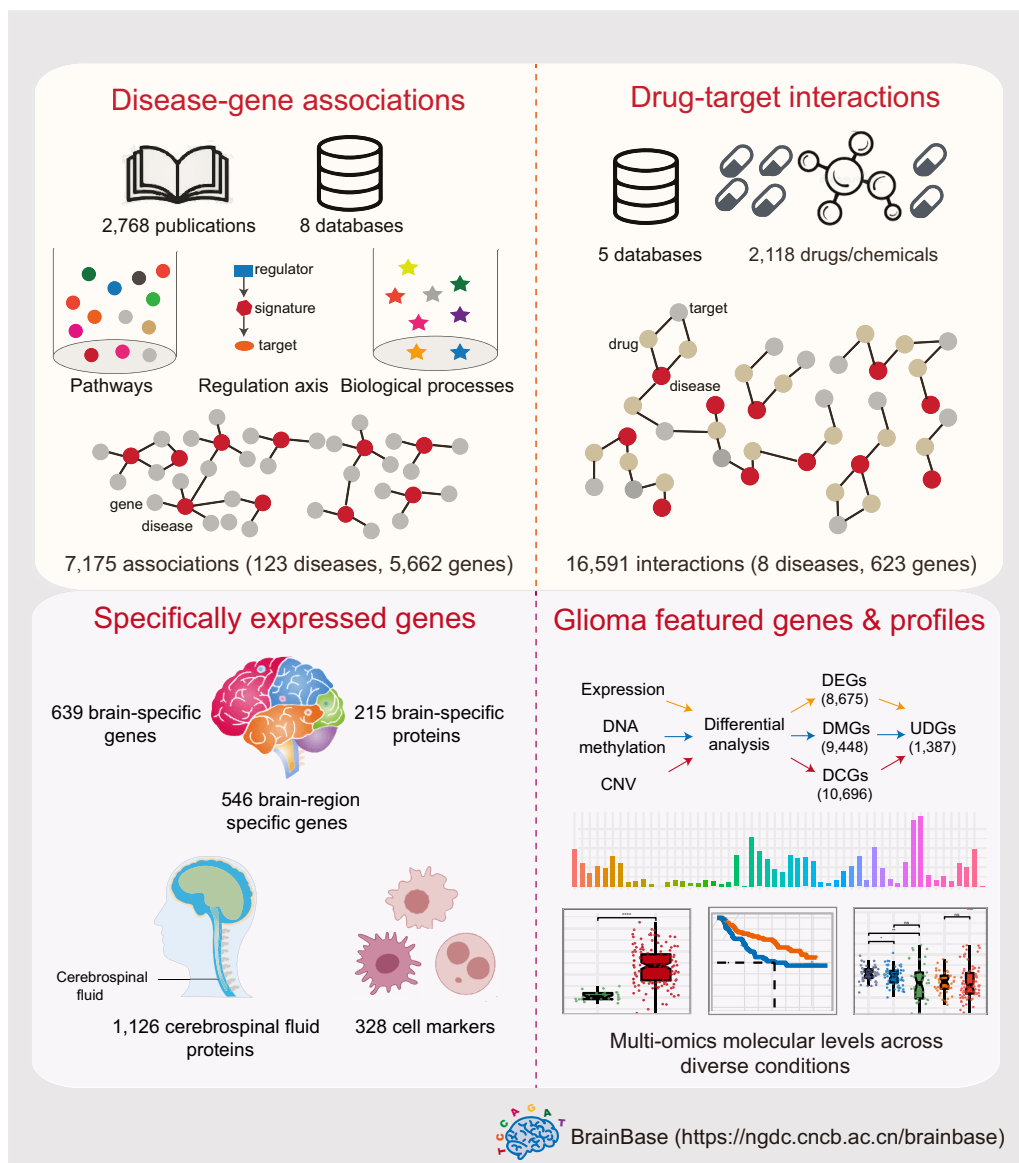
Figure 1. Database contents and features.

markers in 23 cell types (RNA level), 1126 CSF proteins (protein level) and 215 brain-specific proteins (protein level), respectively (see Materials and Methods). Intriguingly, we find among the disease-associated/drug–targeted genes, 192 are brain-specific, 91 are brain-region-specific, 134 are brain cell markers, 373 are CSF associated genes and 94 are brain-specific proteins (Supplementary Tables S1–S5). Clearly, these results indicate that the five types of specific genes are of great potential for clinical applications. Meanwhile, it should be noted that brain-region-specific genes may not reflect the unique features of different regions, since it has been reported that brain regions have close functional connectivity with other regions (59,60). To facilitate users to perform in-depth investigations on these specific genes, BrainBase provides open access to these genes as well as their details (e.g. cell type, brain region, expression level and gene type) at https://ngdc.cncb.ac.cn/brainbase/genes.

Glioma multi-omics profiles and featured gene groups

Considering the severity of glioma among brain tumors, the current version of BrainBase collects 21 datasets from worldwide public resources, and exhibits multi-omics molecular profiles (expression, DNA methylation and CNV) across various samples/conditions (Figure 3). For any given gene of interest, BrainBase provides a wealth of multi-omics molecular levels across diverse conditions, including glioma vs normal, different grades (G4 vs G3 vs G2), *IDH* status, 1p19q status, *IDH* & 1p19q status, *MGMT* status, cell subtype and survival (Figure 3A-C). In addition, BrainBase offers a series of charts for visualizing omics profiles, which can be also downloaded in PDF format.

In order to provide candidate genes for glioma studies, based on these collected multi-omics datasets, four groups of glioma featured genes are identified (Figure 3D), namely, UDGs, DEGs, DMGs and DCGs (see Materials
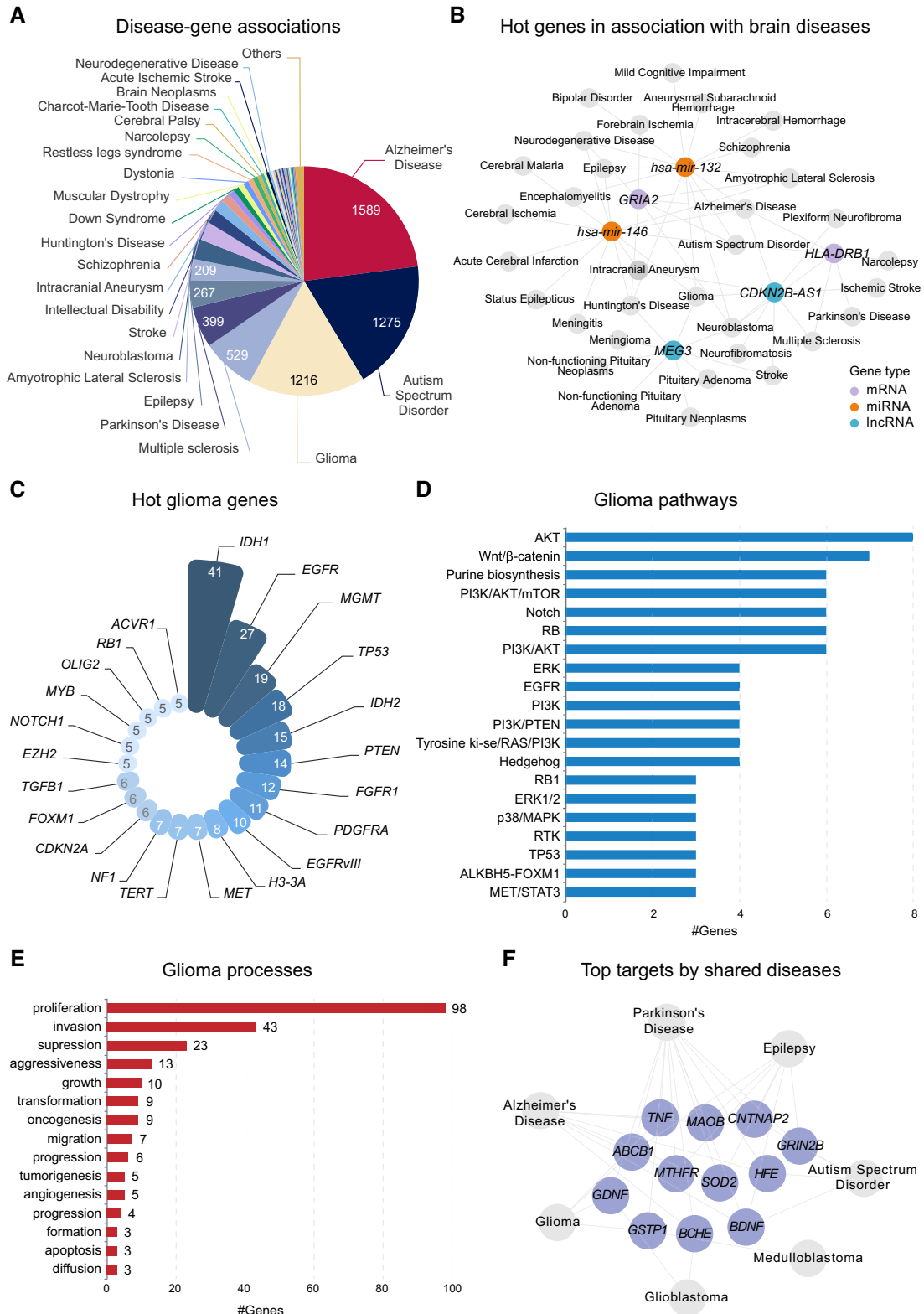
Figure 2. Statistics derived from curated disease–gene associations and drug–target interactions. (**A**) Brain diseases in terms of disease–gene associations, (**B**) Hot genes in association with brain diseases, (**C**) Hot glioma genes, (**D**) Glioma pathways, (**E**) Glioma processes and (**F**) Top targets by shared diseases.
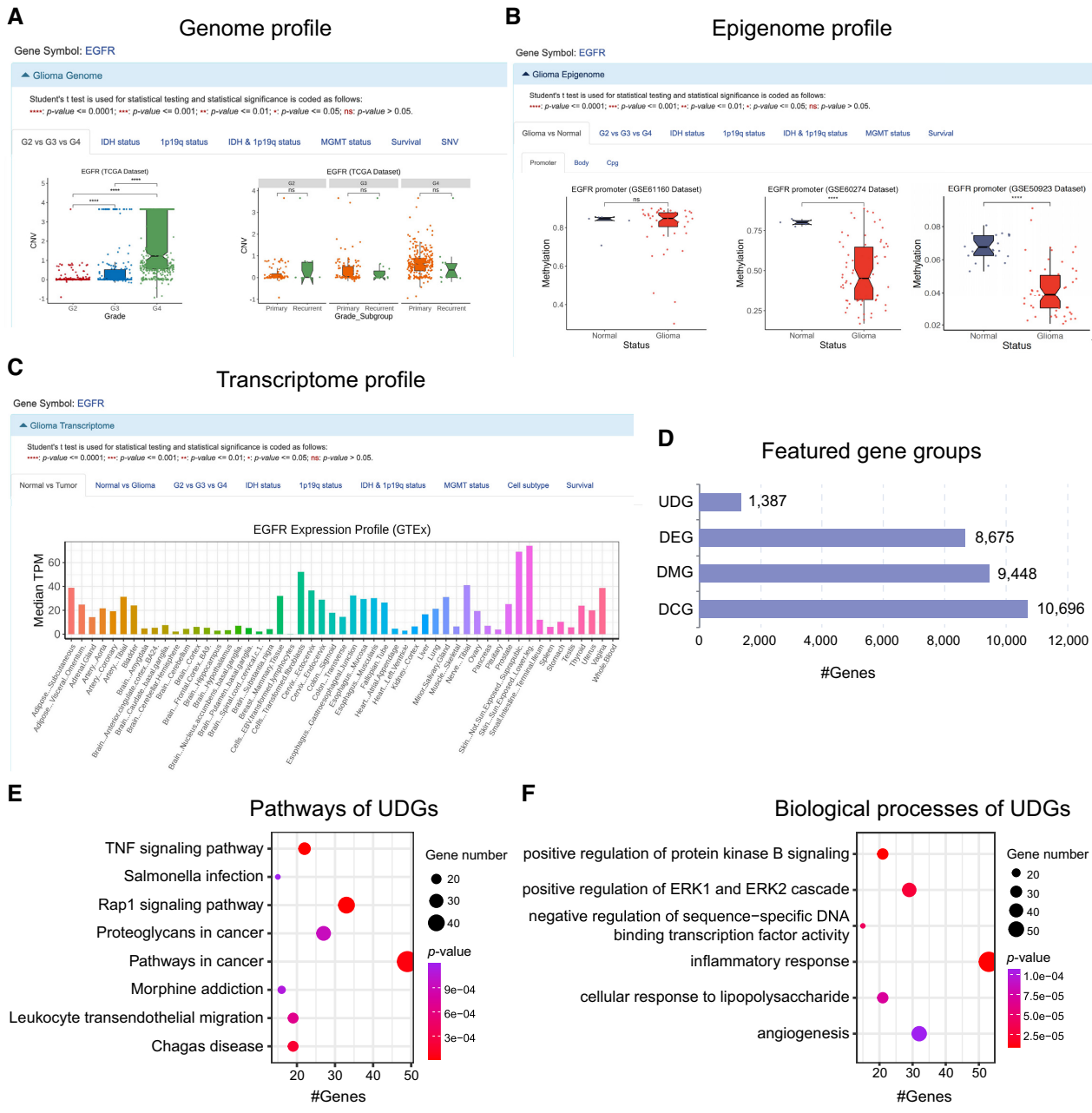
Figure 3. Glioma multi-omics profiles. (**A**) Screenshot of genome profile, (**B**) screenshot of epigenome profile, (**C**) screenshot of transcriptome profile, (**D**) four groups of featured genes, (**E**) pathways of UDGs, and (**F**) biological processes of UDGs.

and Methods). In this study, UDGs are defined as differential genes that exhibit significant differences ubiquitously at multi-omics levels, DEGs/DMGs/DCGs are defined as genes with significant differences at the expression/DNA methylation/CNV level, respectively. As a result, a total of 1387 UDGs, 8675 DEGs, 9448 DMGs and 10 696 DCGs are obtained, which are publicly available at https://ngdc.cncb.ac.cn/brainbase/featured_genes.

Remarkably, we find that these featured genes cover several well-known biomarkers, e.g. *MGMT* in UDGs, *IDH1* in DEGs, *ADAR* in DMGs, and *EGFR* in DCGs, respectively. Specifically, *MGMT*, a key gene for DNA repair, is

associated with glioma survival and clinical treatment (11). *ADAR*, a gene encoding the enzyme responsible for RNA editing (61), is associated with GBM cell proliferation (62). In our previous study, *PRKCG*, a member of protein kinase C (PKC) family (63), is in UDGs and can be used for glioma diagnosis, prognosis and treatment prediction (35). In addition, functional enrichment analysis of these 1,387 UDGs shows that UDGs are significantly associated with TNF signaling pathway, rap1 signaling pathway, proteoglycans in cancer, positive regulation of *ERK1* and *ERK2* cascade, inflammatory response and angiogenesis (Figure 3E and F). Importantly, among 5,976 brain-diseases' associated genes,

285 genes are UDGs, including 65 glioma-associated genes (Supplementary Table S6) and 236 genes associated with other brain diseases (Supplementary Table S7), clearly indicating that UDGs are of utmost importance with potential clinical application for brain diseases.

DISCUSSION AND FUTURE DEVELOPMENTS

In this study, we present BrainBase, a comprehensive knowledgebase that provides a high-quality collection of brain diseases, associated genes, drugs and omics profiles. Based on manual curation and integration from published literatures and related databases, the current version of BrainBase houses 7,175 disease–gene associations, 16 591 drug–target interactions, 21 multi-omics datasets, five types of specific genes and four groups of featured genes. As an important resource of the National Genomics Data Center (NGDC, https://ngdc.cncb.ac.cn) (64), BrainBase is devoted to serving as an open-access resource for studying brain diseases. Future directions include: (i) frequent curation and incorporation of disease–gene associations and drug–target interactions; (ii) integration and analysis of more multi-omics datasets for common brain diseases and (iii) improvement of web interfaces and development of tools in aid of multi-omics data mining and visualization. We also call for worldwide scientists to work together to build BrainBase into a valuable resource covering more comprehensive associations, interactions and omics datasets and further providing high-quality curated knowledge for brain research.

DATA AVAILABILITY

BrainBase is freely available online at https://ngdc.cncb.ac.cn/brainbase and does not require user to register.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank a number of users for reporting bugs and providing suggestions as well as Zhuojing Fan and Zheng Zhao for their valuable comments on this work.

REFERENCES

1. Wang,Y., Pan,Y. and Li,H. (2020) What is brain health and why is it important? *BMJ*, **371**, m3683.
2. Ceccarelli,M., Barthel,F.P., Malta,T.M., Sabedot,T.S., Salama,S.R., Murray,B.A., Morozova,O., Newton,Y., Radenbaugh,A., Pagnotta,S.M. *et al.* (2016) Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell*, **164**, 550–563.
3. Schwartzbaum,J.A., Fisher,J.L., Aldape,K.D. and Wrensch,M. (2006) Epidemiology and molecular pathology of glioma. *Nat. Clin. Pract. Neurol.*, **2**, 494–503.
4. Amunts,K., Ebell,C., Muller,J., Telefont,M., Knoll,A. and Lippert,T. (2016) The human brain project: creating a European research infrastructure to decode the human brain. *Neuron*, **92**, 574–581.
5. Mott,M.C., Gordon,J.A. and Koroshetz,W.J. (2018) The NIH BRAIN Initiative: Advancing neurotechnologies, integrating disciplines. *PLoS Biol.*, **16**, e3000066.
6. Okano,H., Miyawaki,A. and Kasai,K. (2015) Brain/MINDS: brain-mapping project in Japan. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **370**, 20140310.
7. Poo,M.M., Du,J.L., Ip,N.Y., Xiong,Z.Q., Xu,B. and Tan,T. (2016) China brain project: basic neuroscience, brain diseases, and brain-inspired computing. *Neuron*, **92**, 591–596.
8. Capper,D., Jones,D.T.W., Sill,M., Hovestadt,V., Schrimpf,D., Sturm,D., Koelsche,C., Sahm,F., Chavez,L., Reuss,D.E. *et al.* (2018) DNA methylation-based classification of central nervous system tumours. *Nature*, **555**, 469–474.
9. Jia,P., Manuel,A.M., Fernandes,B.S., Dai,Y. and Zhao,Z. (2021) Distinct effect of prenatal and postnatal brain expression across 20 brain disorders and anthropometric social traits: a systematic study of spatiotemporal modularity. *Brief Bioinform*, bbab214.
10. Verhaak,R.G., Hoadley,K.A., Purdom,E., Wang,V., Qi,Y., Wilkerson,M.D., Miller,C.R., Ding,L., Golub,T., Mesirov,J.P. *et al.* (2010) Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*, **17**, 98–110.
11. Hegi,M.E., Diserens,A.C., Gorlia,T., Hamou,M.F., de Tribolet,N., Weller,M., Kros,J.M., Hainfellner,J.A., Mason,W., Mariani,L. *et al.* (2005) MGMT gene silencing and benefit from temozolomide in glioblastoma. *N. Engl. J. Med.*, **352**, 997–1003.
12. Mirzaa,G.M., Millen,K.J., Barkovich,A.J., Dobyns,W.B. and Paciorkowski,A.R. (2014) The Developmental Brain Disorders Database (DBDB): a curated neurogenetics knowledge base with clinical and research applications. *Am. J. Med. Genet. A*, **164A**, 1503–1511.
13. Puchalski,R.B., Shah,N., Miller,J., Dalley,R., Nomura,S.R., Yoon,J.G., Smith,K.A., Lankerovich,M., Bertagnolli,D., Bickley,K. *et al.* (2018) An anatomic transcriptional atlas of human glioblastoma. *Science*, **360**, 660–663.
14. Zhao,Z., Zhang,K.N., Wang,Q., Li,G., Zeng,F., Zhang,Y., Wu,F., Chai,R., Wang,Z., Zhang,C. *et al.* (2021) Chinese Glioma Genome Atlas (CGGA): a comprehensive resource with functional genomic data from Chinese glioma patients. *Genomics Proteomics Bioinformatics*, **19**, 1–12.
15. Bowman,R.L., Wang,Q., Carro,A., Verhaak,R.G. and Squatrito,M. (2017) GlioVis data portal for visualization and analysis of brain tumor expression datasets. *Neuro. Oncol.*, **19**, 139–141.
16. Song,L., Pan,S., Zhang,Z., Jia,L., Chen,W.H. and Zhao,X.M. (2021) STAB: a spatio-temporal cell atlas of the human brain. *Nucleic Acids Res.*, **49**, D1029–D1037.
17. Huang,Z., Shi,J., Gao,Y., Cui,C., Zhang,S., Li,J., Zhou,Y. and Cui,Q. (2019) HMDD v3.0: a database for experimentally supported human microRNA-disease associations. *Nucleic Acids Res.*, **47**, D1013–D1017.
18. Ma,L., Cao,J., Liu,L., Du,Q., Li,Z., Zou,D., Bajic,V.B. and Zhang,Z. (2019) LncBook: a curated knowledgebase of human long non-coding RNAs. *Nucleic Acids Res.*, **47**, D128–D134.
19. Li,M., Zou,D., Li,Z., Gao,R., Sang,J., Zhang,Y., Li,R., Xia,L., Zhang,T., Niu,G. *et al.* (2019) EWAS Atlas: a curated knowledgebase of epigenome-wide association studies. *Nucleic Acids Res.*, **47**, D983–D988.
20. Buniello,A., MacArthur,J.A.L., Cerezo,M., Harris,L.W., Hayhurst,J., Malangone,C., McMahon,A., Morales,J., Mountjoy,E., Sollis,E. *et al.* (2019) The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.*, **47**, D1005–D1012.

21. Niu,G., Zou,D., Li,M., Zhang,Y., Sang,J., Xia,L., Li,M., Liu,L., Cao,J., Zhang,Y. *et al.* (2019) Editome Disease Knowledgebase (EDK): a curated knowledgebase of editome-disease associations in human. *Nucleic Acids Res.*, **47**, D78–D83.

22. Zhang,J., Yan,S., Li,R., Wang,G., Kang,S., Wang,Y., Hou,W., Wang,C. and Tian,W. (2021) CRMarker: a manually curated comprehensive resource of cancer RNA markers. *Int. J. Biol. Macromol.*, **174**, 263–269.

23. Gaulton,A., Bellis,L.J., Bento,A.P., Chambers,J., Davies,M., Hersey,A., Light,Y., McGlinchey,S., Michalovich,D., Al-Lazikani,B. *et al.* (2012) ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.*, **40**, D1100–D1107.

24. Wishart,D.S., Knox,C., Guo,A.C., Cheng,D., Shrivastava,S., Tzur,D., Gautam,B. and Hassanali,M. (2008) DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.*, **36**, D901–D906.

25. Subramanian,A., Narayan,R., Corsello,S.M., Peck,D.D., Natoli,T.E., Lu,X., Gould,J., Davis,J.F., Tubelli,A.A., Asiedu,J.K. *et al.* (2017) A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell*, **171**, 1437–1452.

26. Thorn,C.F., Klein,T.E. and Altman,R.B. (2013) PharmGKB: the Pharmacogenomics Knowledge Base. *Methods Mol. Biol.*, **1015**, 311–320.

27. Davis,A.P., Grondin,C.J., Johnson,R.J., Sciaky,D., Wiegers,J., Wiegers,T.C. and Mattingly,C.J. (2021) Comparative Toxicogenomics Database (CTD): update 2021. *Nucleic Acids Res.*, **49**, D1138–D1143.

28. Schriml,L.M., Arze,C., Nadendla,S., Chang,Y.W., Mazaitis,M., Felix,V., Feng,G. and Kibbe,W.A. (2012) Disease Ontology: a backbone for disease semantic integration. *Nucleic Acids Res.*, **40**, D940–D946.

29. Braschi,B., Denny,P., Gray,K., Jones,T., Seal,R., Tweedie,S., Yates,B. and Bruford,E. (2019) Genenames.org: the HGNC and VGNC resources in 2019. *Nucleic Acids Res.*, **47**, D786–D792.

30. Clough,E. and Barrett,T. (2016) The Gene Expression Omnibus Database. *Methods Mol. Biol.*, **1418**, 93–110.

31. Hutter,C. and Zenklusen,J.C. (2018) The Cancer Genome Atlas: creating lasting value beyond its data. *Cell*, **173**, 283–285.

32. Consortium,GTEx. (2013) The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.*, **45**, 580–585.

33. Thul,P.J. and Lindskog,C. (2018) The human protein atlas: a spatial map of the human proteome. *Protein Sci.*, **27**, 233–244.

34. Yanai,I., Benjamin,H., Shmoish,M., Chalifa-Caspi,V., Shklar,M., Ophir,R., Bar-Even,A., Horn-Saban,S., Safran,M., Domany,E. *et al.* (2005) Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*, **21**, 650–659.

35. Liu,L., Wang,G., Wang,L., Yu,C., Li,M., Song,S., Hao,L., Ma,L. and Zhang,Z. (2020) Computational identification and characterization of glioma candidate biomarkers through multi-omics integrative profiling. *Biol. Direct*, **15**, 10.

36. Zhang,X., Lan,Y., Xu,J., Quan,F., Zhao,E., Deng,C., Luo,T., Xu,L., Liao,G., Yan,M. *et al.* (2019) CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res.*, **47**, D721–D728.

37. Sasayama,D., Hattori,K., Ogawa,S., Yokota,Y., Matsumura,R., Teraishi,T., Hori,H., Ota,M., Yoshida,S. and Kunugi,H. (2017) Genome-wide quantitative trait loci mapping of the human cerebrospinal fluid proteome. *Hum. Mol. Genet.*, **26**, 44–51.

38. Sasayama,D., Hattori,K., Yokota,Y., Matsumura,R., Teraishi,T., Yoshida,S. and Kunugi,H. (2020) Increased apolipoprotein E and decreased TNF-alpha in the cerebrospinal fluid of nondemented APOE-epsilon4 carriers. *Neuropsychopharmacol. Rep.*, **40**, 201–205.

39. Louis,D.N., Ohgaki,H., Wiestler,O.D., Cavenee,W.K., Burger,P.C., Jouvet,A., Scheithauer,B.W. and Kleihues,P. (2007) The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol.*, **114**, 97–109.

40. Wiestler,B., Capper,D., Holland-Letz,T., Korshunov,A., von Deimling,A., Pfister,S.M., Platten,M., Weller,M. and Wick,W. (2013) ATRX loss refines the classification of anaplastic gliomas and identifies a subgroup of IDH mutant astrocytic tumors with better prognosis. *Acta Neuropathol.*, **126**, 443–451.

41. Liu,T., Ren,D., Zhu,X., Yin,Z., Jin,G., Zhao,Z., Robinson,D., Li,X., Wong,K., Cui,K. *et al.* (2013) Transcriptional signaling pathways inversely regulated in Alzheimer's disease and glioblastoma multiform. *Sci. Rep.*, **3**, 3467.

42. Ghosh,M.K., Chakraborty,D., Sarkar,S., Bhowmik,A. and Basu,M. (2019) The interrelationship between cerebral ischemic stroke and glioma: a comprehensive study of recent reports. *Signal Transduct. Target Ther.*, **4**, 42.

43. Benjamini,Y. and Hochberg,Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57**, 289–300.

44. Chen,W., Hu,Y. and Ju,D. (2020) Gene therapy for neurodegenerative disorders: advances, insights and prospects. *Acta Pharm. Sin. B*, **10**, 1347–1359.

45. Qi,M., Fan,S., Wang,Z., Yang,X., Xie,Z., Chen,K., Zhang,L., Lin,T., Liu,W., Lin,X. *et al.* (2019) Identifying common genes, cell types and brain regions between diseases of the nervous system. *Front Genet*, **10**, 1202.

46. Mehta,S. and Lo Cascio,C. (2018) Developmentally regulated signaling pathways in glioma invasion. *Cell. Mol. Life Sci.*, **75**, 385–402.

47. Baldwin,R.M., Parolin,D.A. and Lorimer,I.A. (2008) Regulation of glioblastoma cell invasion by PKC iota and RhoB. *Oncogene*, **27**, 3587–3595.

48. Brennan,C.W., Verhaak,R.G., McKenna,A., Campos,B., Noushmehr,H., Salama,S.R., Zheng,S., Chakravarty,D., Sanborn,J.Z., Berman,S.H. *et al.* (2013) The somatic genomic landscape of glioblastoma. *Cell*, **155**, 462–477.

49. Park,S.H., Won,J., Kim,S.I., Lee,Y., Park,C.K., Kim,S.K. and Choi,S.H. (2017) Molecular testing of brain tumor. *J Pathol Transl Med*, **51**, 205–223.

50. Sun,J., Zhu,K., Zheng,W. and Xu,H. (2015) A comparative study of disease genes and drug targets in the human protein interactome. *BMC Bioinformatics*, **16**(Suppl. 5), S1.

51. Suthram,S., Dudley,J.T., Chiang,A.P., Chen,R., Hastie,T.J. and Butte,A.J. (2010) Network-based elucidation of human disease similarities reveals common functional modules enriched for pluripotent drug targets. *PLoS Comput. Biol.*, **6**, e1000662.

52. Loetscher,H., Steinmetz,M. and Lesslauer,W. (1991) Tumor necrosis factor: receptors and inhibitors. *Cancer Cells*, **3**, 221–226.

53. Bathina,S. and Das,U.N. (2015) Brain-derived neurotrophic factor and its clinical implications. *Arch Med Sci*, **11**, 1164–1178.

54. Ramsay,R.R., Popovic-Nikolic,M.R., Nikolic,K., Uliassi,E. and Bolognesi,M.L. (2018) A perspective on multi-target drug discovery and design for complex diseases. *Clin. Transl. Med.*, **7**, 3.

55. Lu,J.J., Pan,W., Hu,Y.J. and Wang,Y.T. (2012) Multi-target drugs: the trend of drug research and development. *PLoS One*, **7**, e40262.

56. Ryaboshapkina,M. and Hammar,M. (2019) Tissue-specific genes as an underutilized resource in drug discovery. *Sci. Rep.*, **9**, 7233.

57. Vasmatzis,G., Klee,E.W., Kube,D.M., Therneau,T.M. and Kosari,F. (2007) Quantitating tissue specificity of human genes to facilitate biomarker discovery. *Bioinformatics*, **23**, 1348–1355.

58. McKenzie,A.T., Wang,M., Hauberg,M.E., Fullard,J.F., Kozlenkov,A., Keenan,A., Hurd,Y.L., Dracheva,S., Casaccia,P., Roussos,P. *et al.* (2018) Brain cell type specific gene expression and co-expression network architectures. *Sci. Rep.*, **8**, 8868.

59. Tanimizu,T., Kenney,J.W., Okano,E., Kadoma,K., Frankland,P.W. and Kida,S. (2017) Functional connectivity of multiple brain regions required for the consolidation of social recognition memory. *J. Neurosci.*, **37**, 4103–4116.

60. Greicius,M.D., Krasnow,B., Reiss,A.L. and Menon,V. (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 253–258.

61. Nishikura,K. (2010) Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.*, **79**, 321–349.

62. Lv,X., Gu,C. and Guo,S. (2020) Activation of BDNF-AS/ADAR/p53 positive feedback loop inhibits glioblastoma cell proliferation. *Neurochem. Res.*, **45**, 508–518.

63. do Carmo,A., Balca-Silva,J., Matias,D. and Lopes,M.C. (2013) PKC signaling in glioblastoma. *Cancer Biol. Ther.*, **14**, 287–294.

64. CNCB-NGDC Members and Partners. (2021) Database Resources of the National Genomics Data Center, China National Center for Bioinformation in 2021. *Nucleic Acids Res.*, **49**, D18–D28.