



Screening and identification of novel candidate biomarkers of focal cortical dysplasia type II via bioinformatics analysis

Jiang-ya Wang^{1,2} · Yang Li³ · Yuan-yuan Lv⁴ · Lian Jiang³

Received: 20 July 2021 / Accepted: 14 January 2022 / Published online: 2 February 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Purpose Focal cortical dysplasia (FCD) is the most common developmental malformation that causes refractory epilepsy. FCD II is a common neuropathological finding in tissues resected therapeutically from patients with drug-resistant epilepsy. However, its molecular genetic etiology remains unclear. This study aimed to identify potential molecular markers of FCD II using bioinformatics analysis.

Methods We downloaded two datasets for FCD II from the Gene Expression Omnibus data repository. Differentially expressed genes (DEGs) between FCD II and normal brain tissues were identified, and functional enrichment analysis was performed. A protein–protein interaction network was constructed, and hub genes were identified from the DEGs. The hub gene expression was validated using WB *in vitro*. IHC staining was performed to verify the feasibility of the target molecular markers identified in the bioinformatics analysis.

Results One hundred sixty-seven common DEGs were identified between the datasets. The GO and KEGG analyses showed that variations were prominently enriched in some functions associated with gene expression. Five hub genes (i.e., FANCI, FANCA, BRCA2, RAD18, and KEAP1) were identified. Western blotting confirmed that all hub gene expressions were higher in the FCD II tissue than in the normal brain tissue. IHC staining showed that the FANCI expression significantly increased in the FCD II tissue.

Conclusion There are DEGs between FCD II and normal brain tissues, which may be considered biomarkers for FCD II, along with FANCI. The DEGs and hub genes identified in the bioinformatics analysis could serve as candidate targets for diagnosing and treating FCD II.

Keywords Focal cortical dysplasia type II · Bioinformatics analysis · Differentially expressed genes · Hub genes

Introduction

Focal cortical dysplasia (FCD) is the most frequent developmental central nervous system malformation, which often causes drug-resistant epilepsy in children [1–3]. It is difficult

to control with currently available anti-epileptic drugs and accounts for up to 50% of cases requiring epilepsy surgery [4]. The International League Against Epilepsy Diagnostic Methods Commission has proposed a classification system for FCD entities based on their neuropathological features [5]. FCD I refers to isolated lesions that present radial and/or tangential dyslamination of the cortex; FCD II refers to isolated lesions characterized by cortical dyslamination and dysmorphic neurons without (FCD IIa) or with balloon cells (FCD IIb); and FCD III is associated with other brain lesions. FCD II is one of the most common neuropathological findings in tissues resected therapeutically from patients with drug-resistant epilepsy [6, 7]; 29–39% of all patients with FCD who undergo epilepsy surgery were diagnosed with FCD II. The histopathological features of FCD II resemble those of hemimegalencephaly, except for the lesion extending to unilateral hemispheric enlargement [3]. Most

✉ Lian Jiang
hlkycdx899@163.com

¹ Department of Pediatrics, Hebei Medical University, Shijiazhuang, China

² Department of Pediatrics, Hebei General Hospital, Shijiazhuang, China

³ Department of Pediatrics, The Fourth Hospital of Hebei Medical University, Chang'an district, Shijiazhuang 050000, China

⁴ Department of Pediatrics, Baoding First Central Hospital, Baoding, China

FCD cases are sporadic. Owing to the lack of understanding of the molecular genetic etiology of FCD, new and more effective FCD therapies remain elusive [8]. Therefore, it is necessary to explore FCD at the molecular level.

With the rapid development of high-throughput sequencing technology, bioinformatics analysis has become one of the most important core fields of molecular targeted therapy, which is a scientific approach that uses computers to analyze the sheer volume of data obtained from high-throughput sequencing. Based on the nucleic acid and protein sequences, analysis of the biological information of the differential expression in the sequences has been widely used in several research fields, such as tumors, and is a reliable method for identifying diagnostic and therapeutic targets. In recent years, the use of bioinformatics analysis in non-oncology research has also increased. For example, Nashiry identified the influence of coronavirus disease on cardiovascular and hypertensive comorbidities using bioinformatics analysis [9].

In this study, bioinformatics analysis was used to identify potential molecular markers of FCD II. A protein–protein interaction (PPI) network was constructed to determine the relationship between the differentially expressed genes (DEGs) of FCD II and normal brain tissues from the Gene Expression Omnibus (GEO) dataset. Gene function annotation was performed via Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses. Finally, we identified five hub genes using the MCC, MNC, BottleNeck, and Degree algorithms from the PPI network. Subsequently, experiments were conducted to test the expression of related genes in the *ex vivo* brain tissue. Bioinformatics analysis showed that the expression of the hub genes was significantly higher in the FCD II tissue than in the normal brain tissue. The five hub genes may be used as new diagnostic molecular markers for FCD II. Herein, the role of FANCI in FCD II was verified.

Methods

Gene expression profile from a public dataset

The GEO is a public repository for genetic data, including various diseased tissues and normal tissues (available online: <http://www.ncbi.nlm.nih.gov/geo>) [10]. We downloaded the FCD II RNA expression profiling datasets GSE128300 (GPL10999 platform) and GSE62019 (GPL10558 platform). The GSE128300 dataset included 15 surgical FCD II specimens, matched with five epilepsy and eight non-epilepsy brain tissue controls. The GSE62019 dataset included five cases with FCD II and three cases with tuberous sclerosis (TS) compared with three normal brain tissue controls. This study analyzed DEGs between FCD II tissues and normal

brain tissues. Thus, we selected 15 FCD II tissues and 8 normal brain tissues from GSE128300 and 5 FCD II tissues and 3 normal brain tissues from GSE62019 for the bioinformatics analysis.

DEGs identified

To identify the DEGs between the FCD II and normal brain tissues, we used the limma package in R to analyze the datasets downloaded from the GEO. The criteria for selecting the DEGs were set as follows: adjusted *p*-value of < 0.05 and \log_2 fold change of ≥ 1 . A volcano map was drawn to show the criteria for selecting the DEGs using the volcano plotting tool (<https://shengxin.ren>) [11]. A Venn plot was constructed to overlap the common DEGs between the two datasets using the FunRich software [12, 13].

Functional enrichment analysis

Metascape (<http://metascape.org/gp/index.html#/main/step1>), an online analysis tool suite with the function of integrated discovery and annotation, mainly provides typical batch annotation and GO term enrichment analysis to highlight the most relevant GO terms associated with a given gene [14]. It was used to perform the GO and KEGG pathway enrichment analyses of the DEGs. A bubble chart was created using the DAVID tool, which is a comprehensive set of functional annotation tools for investigators to understand the biological meaning behind a large list of genes (<https://david.ncifcrf.gov/>) [15]. Statistical significance was set at *p*-values of < 0.05.

PPI network construction and hub gene selection

The Search Tool for the Retrieval of Interacting Gene (STRING) database (<https://string-db.org>) was used to explore the comprehensive interaction of the DEGs. The PPI network based on the STRING database was generated using the Cytoscape software (www.cytoscape.org) [16]. The hub genes were evaluated using the Cytoscape algorithm.

Western blotting

We extracted proteins from the normal, TS, and FCD II tissues using RIPA buffer. The proteins were quantified via SDS-PAGE electrophoresis. The proteins extracted using cell lysis buffer were transferred to nitrocellulose membranes and incubated with primary 1:1000 rabbit antihuman antibodies specific for FANCI, FANCA, BRCA2, RAD18, and KEAP1 at 4 °C overnight. Thereafter, the membranes were washed with a buffer solution and incubated with secondary antibodies. Finally, we detected signals with image

acquisition using the Immobilon ECL substrate and X-ray Film Processor.

Histology and immunohistochemistry

Tissue biopsy samples were fixed in 4% paraformaldehyde overnight, dehydrated in 70% ethanol, and embedded in paraffin. The tissue Sects. (5- μ m-thick) were cut, mounted, deparaffinized, and pretreated. The samples were stained and incubated for 60 min with FANCI mouse monoclonal antibody and FANCI rabbit polyclonal antibody at a dilution ratio of 1:100.

Statistical analysis

Statistical significance was set at p -values of <0.05 , unless otherwise stated. All statistical analyses were conducted

using SPSS (version 26.0; Chicago, IL, USA) and R3.6.1 (<https://www.r-project.org/>).

Results

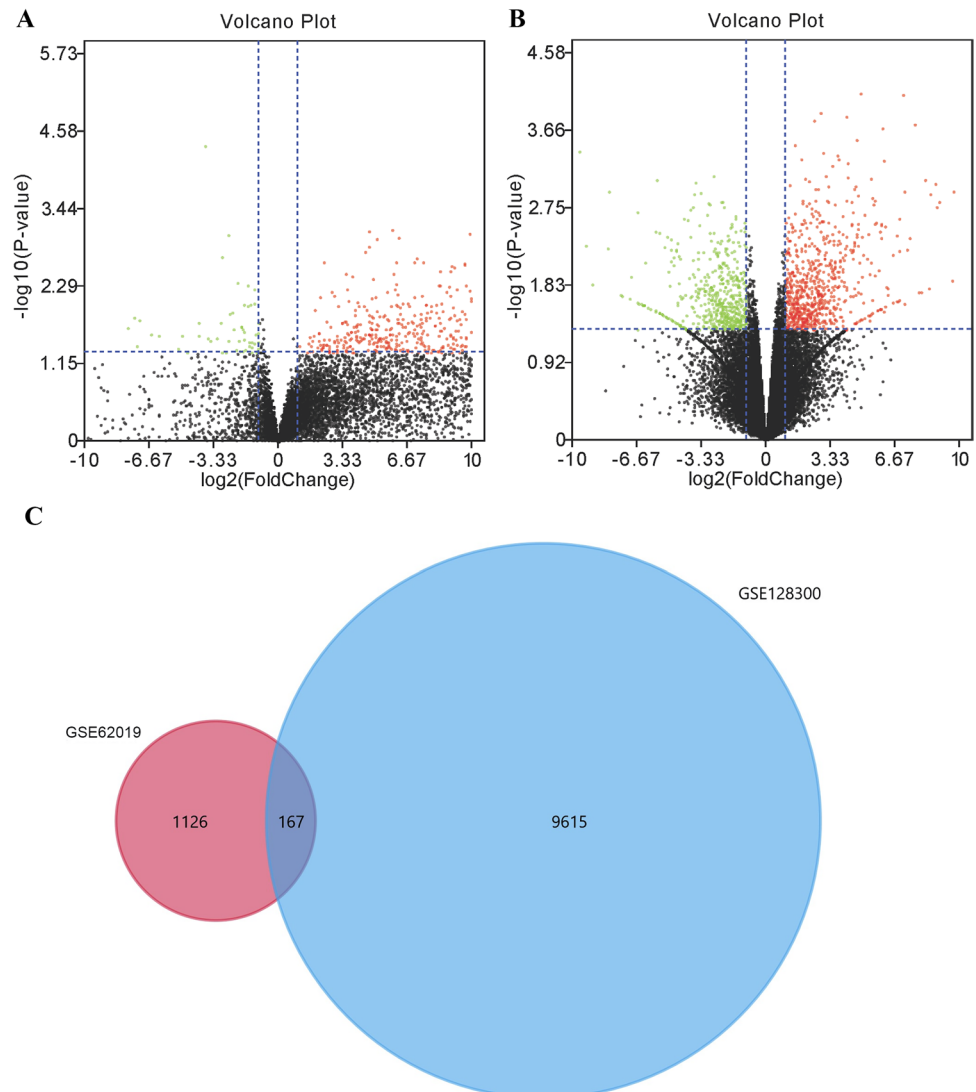
DEGs identified

The volcano map and heat map analysis results are shown in Fig. 1a and b. A total of 9782 DEGs were obtained in the GSE128300 dataset and 1293 DEGs in the GSE62019 dataset. A total of 167 common DEGs were obtained by overlapping the two datasets (Fig. 1c).

Functional enrichment analysis

The GO and KEGG analyses of the 167 DEGs were conducted using Metascape and the DAVID tool, which may

Fig. 1 Screening for the differentially expressed genes. **a** Volcano plot of GSE128300. **b** Volcano plot of GSE62019. **c** Venn graph of the two datasets



improve the biological understanding of these genes (Fig. 2). The enrichment analysis performed using Metascape showed that the DEGs were mainly enriched in the Fanconi anemia pathway, actin filament-based process, negative regulation of cell population proliferation, microtubule cytoskeleton organization, regulation of cardiac muscle contraction (signaling by Rho GTPases, Miro GTPases, and RHOBTB3),

CEN complex, neural tube closure, organelle fusion, prostate gland development, Hippo signaling pathway, antigen processing and presentation of exogenous peptide antigen via MHC class II, regulation of the extrinsic apoptotic signaling pathway via death receptors, regulation of embryonic development, myelination, protein deubiquitination, PID P75 NTR pathway, and PID CDC42 pathway (Fig. 2a–c). The

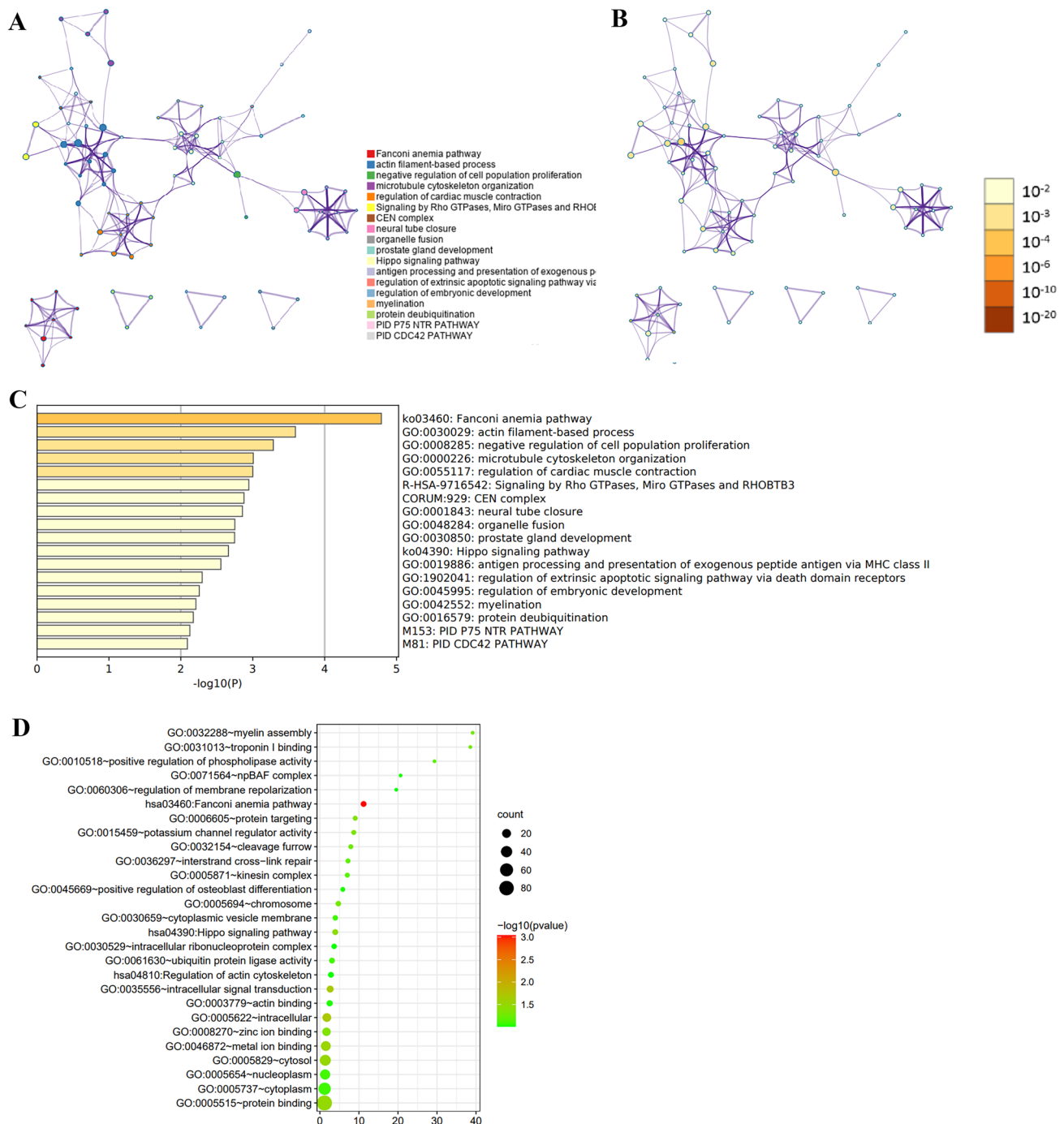


Fig. 2 Functional enrichment analysis of the differentially expressed genes. **a–c** Enrichment analysis results from Metascape. **d** Bubble chart of the functional enrichment analysis using the DAVID tool

enrichment analysis performed using the DAVID tool indicated that the functions of the DEGs were mainly enriched in myelin assembly, troponin I binding, positive regulation of phospholipase activity, npBAF complex, regulation of membrane repolarization, Fanconi anemia pathway, protein targeting, potassium channel regulator activity, cleavage furrow, interstrand cross-link complex, kinesin complex, positive regulation of osteoblast differentiation, chromosome, cytoplasmic vesicle membrane, Hippo signaling pathway, intracellular ribonucleoprotein complex, ubiquitin protein ligase activity, and regulation of actin cytoskeleton (Fig. 2d).

PPI network construction and hub gene selection

A total of 167 common DEGs were included in the STRING database, and the resulting file was further analyzed using Cytoscape. A PPI network was generated (Fig. 3a). We used

four algorithms (MCC, MNC, Degree, and BottleNeck) to calculate the top 10 hub genes separately and used a Venn plot to overlap the common hub genes (Fig. 3b). Finally, five hub genes were identified (FANCI, FANCA, BRCA2, RAD18, and KEAP1; Fig. 3c).

Hub gene expression level validation

To verify the results of the bioinformatics analysis, we used western blotting to detect the expression of the hub genes in the FCD II tissue and normal brain tissue. Consistent with the results of the bioinformatics analysis, western blotting showed that the FCD II tissue had significantly higher expression levels of the hub genes than the normal brain and TS tissues (Fig. 4a).

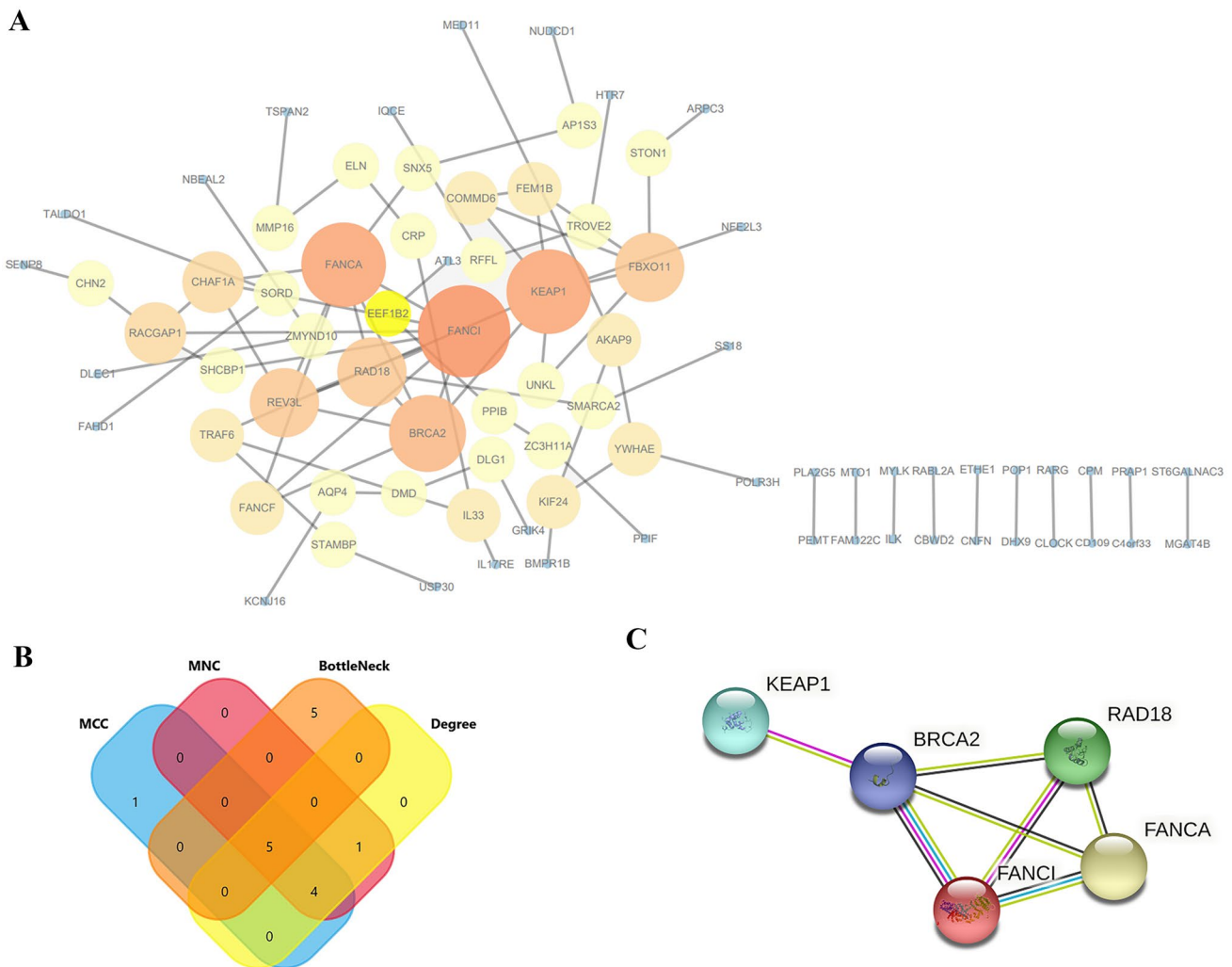
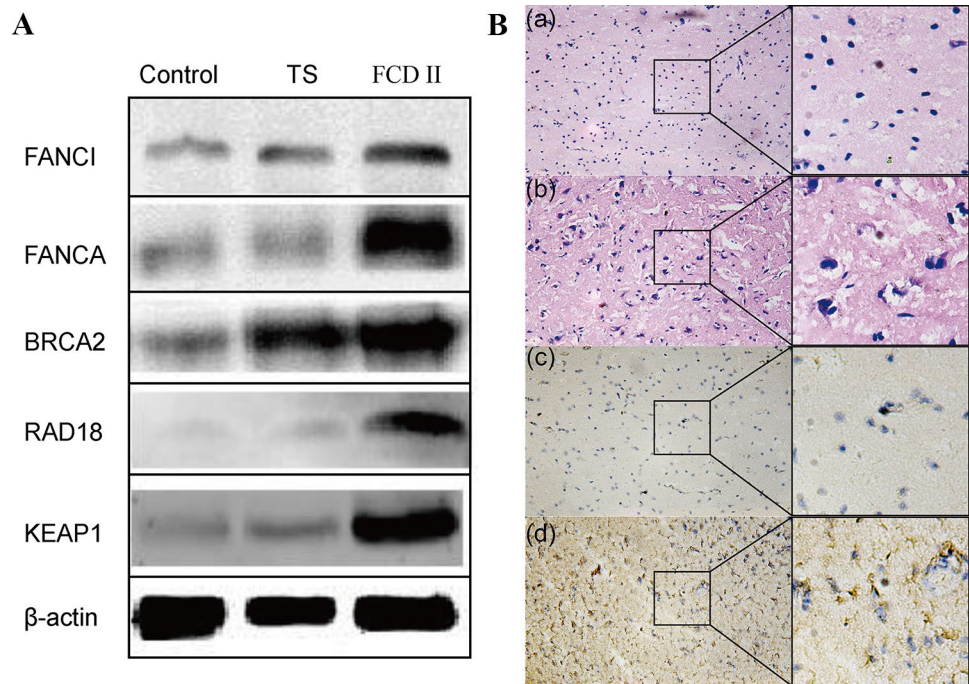


Fig. 3 Protein–protein network and hub gene selection. **a** Protein–protein network of the differentially expressed genes. **b** Venn graph of the four different algorithms (MCC, MNC, Degree, and BottleNeck).

c Five hub genes (FANCI, FANCA, BRCA2, RAD18, and KEAP1) visualized using the STRING database

Fig. 4 Validation of the expression levels of all hub genes and expression of FANCI in the FCD II tissue. **a** Western blotting of the five hub genes in the FCD II tissues and the other two tissues. **b** a HE stained-image of the histology of the normal brain tissues. **b** HE stained-image of the FCD II tissues. **c** IHC stained-image of the normal brain tissues. **d** Typical IHC stained-image of the FCD II tissues. TS, tuberous sclerosis; FCD, focal cortical dysplasia; HE, hematoxylin and eosin; IHC, immunohistochemistry



FANCI expression in the FCD II tissue

Morphological changes were determined using hematoxylin and eosin staining in the FCD II and normal brain tissues. Compared with the normal brain tissue, we observed numerous heteromorphic neurons in the FCD II tissue (Fig. 4b). Immunohistochemistry was used to detect the expression of FANCI and the matched normal brain tissues. The analysis showed that the expression of FANCI in the FCD II tissue was higher than that in the normal brain and TS tissues (Fig. 4b).

Discussion

FCD II is a malformation of cortical development associated with disrupted cortical lamination and specific cytologic abnormalities [17]. It is further divided into FCD IIa and FCD IIb. FCD is the most common histopathology found in lesions surgically resected from children with epilepsy, and FCD II is the most dominant type. Over the past 10 years, studies have identified somatic mutations in 10–63% of FCD II samples after examining them with deep next-generation sequencing and comparing data between surgically resected tissues and peripheral blood samples or salivary cells from the same patients [1].

In recent years, a large number of studies on the molecular mechanism of FCD II have focused on the mTOR pathway, including mTOR, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha, tuberous

sclerosis complex subunit 1, tuberous sclerosis complex subunit 2, AKT serine/threonine kinase 3, Ras homolog enriched in brain, and DEP domain containing 5 [18]. Their mutations lead to the activation of the mTOR signaling cascade, which plays a critical role in regulating cell proliferation during brain development. There are also some studies devoted to identifying molecular markers that do not belong to the mTOR pathway [1, 3].

Few studies have applied bioinformatics analysis to the study of FCD II; thus, our study attempted to use bioinformatics analysis to identify new candidate genes involved in FCD II. We obtained 9782 DEGs in the GSE128300 dataset, 1293 DEGs in the GSE62019 dataset, and 167 common DEGs in the two datasets. Functional annotation of the DEGs indicated that the functions of the DEGs mainly focused on myelin assembly, protein targeting, troponin I binding, positive regulation of phospholipase activity, actin filament-based process, negative regulation of cell population proliferation, and microtubule cytoskeleton organization. The occurrence of FCD II may also be affected by several pathways, such as the Fanconi anemia pathway, Hippo signaling pathway, and Rho pathway. We may find new molecular markers of FCD II by exploring the activation or inhibition of these pathways.

As previously known, nerve cells lack or inappropriately express cytoskeletal proteins necessary for migration and establishment of neuronal polarity, which may cause FCD II [19]. Furthermore, functional annotations show that DEGs influence the microtubule cytoskeleton

organization. In future studies, we will screen and verify the DEGs related to cytoskeletal protein formation.

Using the algorithms from Cytoscape, we obtained five genes (FANCI, FANCA, BRCA2, RAD18, and KEAP1), which may play an important role in the formation of FCD II. These genes are hub nodes in the PPI network and interact with a considerable number of proteins. The influence of changes in the expression of these genes is worthy of attention. FANCA, FANCI, and BRCA2 belong to the Fanconi anemia complementation group.

As previously reported, FANCI is broadly expressed in the brain, testis, and lymph node [20]. It has been proven that FANCI is related to Fanconi anemia complementation groups I and A [21]. Nuclear localization of the FANCA protein is vital for normal biological functions. BRCA2 is a nuclear protein involved in cytokinesis and homologous recombination. Accumulating evidence suggests [22] that the Fanconi anemia pathway forms part of a broader DNA damage response network that includes the BRCA genes known as the FA-BRCA pathway. DNA damage can induce many diseases, such as cancer and gene diseases. RAD18 plays an important role in DNA damage bypass and post-replication repair in vertebrates [23]. KEAP1 is associated with multiple sclerosis, which may cause medial temporal sclerosis [24].

The DEGs and hub genes identified in bioinformatics analyses could help us understand the molecular mechanisms underlying the occurrence and progression of FCD II. Studies have confirmed that cells that strongly express FNAC genes in the central nervous system also express proliferating cell nuclear antigen, which is associated with the mTOR pathway [20]. Aberrant activation of mTOR is associated with developmental malformations and epileptogenesis [25], which can increase the expression of FANCD2 [26]. Previous research [27, 28] has shown that the FANCD2/FANCI complex plays an important role in DNA repair, such as interstrand crosslinks and abnormal transcription. Therefore, we used immunohistochemistry to qualitatively measure the FANCI expression. Based on the overexpression of FANCI in the FCD II tissue and the relevant information, we believe that the occurrence and progression of FCD II are associated with mutations in important genes.

In summary, bioinformatics analysis could be a useful approach to explore the mechanism of the occurrence and development of FCD II by analyzing datasets from the GEO. The identified hub genes and DEGs could serve as a level of research target for the diagnosis and treatment of FCD II.

Author contribution Lian Jiang designed and analyzed the research study. Jiang-ya Wang wrote the manuscript and did the bioinformatics analysis. Yang Li and Yuan-yuan Lv wrote the original draft and performed the experiment in tissue.

Availability of data and material The datasets generated and analyzed during the current study are available in the [NCBI] repository.

Code availability The software application and custom code are available.

Declarations

Ethics approval The research was authorized by the Human Ethics and Research Ethics Committees of the Hebei General Hospital. The informed consents were obtained from all participants.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

- Zhang Z, Gao K, Liu Q et al (2020) Somatic variants in new candidate genes identified in focal cortical dysplasia type II. *Epilepsia* 61(4):667–678
- Tavakol S, Royer J, Lowe AJ et al (2019) Neuroimaging and connectomics of drug-resistant epilepsy at multiple scales: from focal lesions to macroscale networks. *Epilepsia* 60(4):593–604
- Majolo F, Marinowic DR, Machado DC, Da Costa JC (2018) MTOR pathway in focal cortical dysplasia type 2: What do we know. *Epilepsy Behav* 85:157–163
- Crino PB (2015) Focal cortical dysplasia. *Semin Neurol* 35(3):201–208
- Blümcke I, Thom M, Aronica E et al (2011) The clinicopathologic spectrum of focal cortical dysplasias: a consensus classification proposed by an ad hoc Task Force of the ILAE Diagnostic Methods Commission. *Epilepsia* 52(1):158–174
- Chimelli L, Melo A, Avvad-Portari E et al (2017) The spectrum of neuropathological changes associated with congenital Zika virus infection. *Acta Neuropathol* 133(6):983–999
- Jayalakshmi S, Nanda SK, Vooturi S et al (2019) Focal cortical dysplasia and refractory epilepsy: role of multimodality imaging and outcome of surgery. *AJNR Am J Neuroradiol* 40(5):892–898
- Jesus-Ribeiro J, Pires LM, Melo JD et al (2020) Genomic and epigenetic advances in focal cortical dysplasia types I and II: a scoping review. *Front Neurosci* 14:580357
- Nashiry A, Sarmin Sumi S, Islam S, Quinn J, Moni MA (2021) Bioinformatics and system biology approach to identify the influences of COVID-19 on cardiovascular and hypertensive comorbidities. *Brief Bioinform* 22(2):1387–1401
- Edgar R, Domrachev M, Lash AE (2002) Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30(1):207–210
- Wang D, Wang Y, Zou X et al (2020) FOXO1 inhibition prevents renal ischemia-reperfusion injury via cAMP-response element binding protein/PPAR- γ coactivator-1 α -mediated mitochondrial biogenesis. *Br J Pharmacol* 177(2):432–448
- Benito-Martin A, Peinado H (2015) FunRich proteomics software analysis, let the fun begin. *Proteomics* 15(15):2555–2556

13. Pathan M, Keerthikumar S, Ang CS et al (2015) FunRich: an open access standalone functional enrichment and interaction network analysis tool. *Proteomics* 15(15):2597–2601
14. Zhou Y, Zhou B, Pache L et al (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 10(1):1523
15. da Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4(1):44–57
16. Szklarczyk D, Gable AL, Lyon D et al (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47(D1):D607–D613
17. Spreafico R, Blümcke I (2010) Focal cortical dysplasias: clinical implication of neuropathological classification systems. *Acta Neuropathol* 120:359–367. <https://doi.org/10.1007/s00401-010-0714-x>
18. Marsan E, Baulac S (2018) Review: Mechanistic target of rapamycin (mTOR) pathway, focal cortical dysplasia and epilepsy. *Neuropathol Appl Neurobiol* 44(1):6–17
19. Crino PB, Trojanowski JQ, Eberwine J (1997) Internexin, MAP1B, and nestin in cortical dysplasia as markers of developmental maturity. *Acta Neuropathol* 93(6):619–627
20. Titus TA, Yan YL, Wilson C et al (2009) The Fanconi anemia/BRCA gene network in zebrafish: embryonic expression and comparative genomics. *Mutat Res* 668(1–2):117–132
21. Yuan F, Qian L, Zhao X et al (2012) Fanconi anemia complementation group A (FANCA) protein has intrinsic affinity for nucleic acids with preference for single-stranded forms. *J Biol Chem* 287(7):4800–4807
22. García MJ, Fernández V, Osorio A et al (2009) Mutational analysis of FANCL, FANCM and the recently identified FANCI suggests that among the 13 known Fanconi Anemia genes, only FANCD1/BRCA2 plays a major role in high-risk breast cancer predisposition. *Carcinogenesis* 30(11):1898–1902
23. Li P, He C, Gao A et al (2020) RAD18 promotes colorectal cancer metastasis by activating the epithelial-mesenchymal transition pathway. *Oncol Rep* 44(1):213–223
24. Taguchi K, Yamamoto M (2020) The KEAP1-NRF2 system as a molecular target of cancer treatment. *Cancers (Basel)* 13(1)
25. Mühlebner A, Bongaarts A, Sarnat HB, Scholl T, Aronica E (2019) New insights into a spectrum of developmental malformations related to mTOR dysregulations: challenges and perspectives. *J Anat* 235(3):521–542
26. Shen C, Oswald D, Phelps D et al (2013) Regulation of FANCD2 by the mTOR pathway contributes to the resistance of cancer cells to DNA double-strand breaks. *Cancer Res* 73(11):3393–3401
27. Alcón P, Shakeel S, Chen ZA, Rappsilber J, Patel KJ, Passmore LA (2020) FANCD2-FANCI is a clamp stabilized on DNA by monoubiquitination of FANCD2 during DNA repair. *Nat Struct Mol Biol* 27(3):240–248
28. Lopez-Martinez D, Kupculak M, Yang D et al (2019) Phosphorylation of FANCD2 inhibits the FANCD2/FANCI complex and suppresses the fanconi anemia pathway in the absence of DNA damage. *Cell Rep* 27(10):2990–3005.e5

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.