



Database tool

Circad: a comprehensive manually curated resource of circular RNA associated with diseases

Mercy Rophina^{1,†}, Disha Sharma^{1,2,†}, Mukta Poojary^{1,2} and Vinod Scaria^{1,2,*}

¹Genome Informatics Department, CSIR-Institute of Genomics and Integrative Biology, Mathura Road, District-South Delhi, New Delhi-110025, India and ²Academy of Scientific and Innovative Research (AcSIR), AcSIR headquarters, CSIR-HRDC campus, Postal Staff College Area, Sector 19, Kamla Nehru Nagar, Ghaziabad, Uttar Pradesh -201002, India

*Corresponding author: Genome Informatics Department, CSIR-Institute of Genomics and Integrative Biology, Mathura Road, District- South Delhi, New Delhi-110025. Tel: +91-1125895615; Fax: +91-1127667471; Email: vinods@igib.in

†Contributed equally and would like to be known as joint first authors.

Citation details: Rophina, M., Sharma, D., Poojary, M. *et al.* Circad: a comprehensive manually curated resource of circular RNA associated with diseases. *Database* (2020) Vol. 2020: article ID baaa019; doi:10.1093/database/baaa019

Received 24 July 2019; Revised 22 January 2020; Accepted 17 February 2020

Abstract

Circular RNAs (circRNAs) are unique transcript isoforms characterized by back splicing of exon ends to form a covalently closed loop or circular conformation. These transcript isoforms are now known to be expressed in a variety of organisms across the kingdoms of life. Recent studies have shown the role of circRNAs in a number of diseases and increasing evidence points to their potential application as biomarkers in these diseases. We have created a comprehensive manually curated database of circular RNAs associated with diseases. This database is available at URL <http://clingen.igib.res.in/circad/>. The Database lists more than 1300 circRNAs associated with 150 diseases and mapping to 113 International Statistical Classification of Diseases (ICD) codes with evidence of association linked to published literature. The database is unique in many ways. Firstly, it provides ready-to-use primers to work with, in order to use circRNAs as biomarkers or to perform functional studies. It additionally lists the assay and PCR primer details including experimentally validated ones as a ready reference to researchers along with fold change and statistical significance. It also provides standard disease nomenclature as per the ICD codes. To the best of our knowledge, circad is the most comprehensive and updated database of disease associated circular RNAs.

Availability: <http://clingen.igib.res.in/circad/>

Introduction

Circular RNAs are unique transcript isoforms characterized by a circular conformation and produced by the back-splicing of 5' donor end and 3' acceptor ends of pre-mRNA transcripts (1). While the existence of circular isoforms has been previously established for viruses (2), recent evidence suggests they are ubiquitous (3). Circular RNAs have now been reported in a number of organisms including Humans (3), mouse (3), flies (4) and many plants (5). While earlier considered as splicing artefacts (6), recent reports suggest that they are produced through a coordinated biological process and regulated by key genes (7). The mechanism of action has been putatively reported via miRNAs, mRNAs, and RNA-binding proteins (RBPs) (8). For example, CDR1as circRNA has been shown to interact with miR-7 with further effect on cerebral development (3). The role of circular RNA candidates in key biological processes has also been now explored in detail, as in the case of circ-Sry, shown to have a role in sex differentiation (9). Notwithstanding emerging evidence on the functional roles of candidate circular RNAs, the functional annotation of majority of the candidate circular RNAs is still unexplored or needs detailed experimental study to understand their functions.

A number of studies in recent years have analysed the expression patterns of circular RNAs in disease states, and many suggest its potential utility as a biomarker. The closed-loop conformation also provides them an additional advantage as a biomarker that they are resistant to exonuclease digestion (1). These include the studies on a variety of diseases including brain disorders (10), for example, Alzheimer's disease (11); Parkinson disease (12); different cancers including gastric cancer (12–15), lung cancer (16) and colorectal cancer (16, 17); esophageal carcinoma (18); and adenocarcinoma (19), just to name a few.

The emerging corpus of information, while available as disparate pieces of evidence, would significantly benefit from a unified resource systematically cataloguing disease associations of circular RNAs. This is more significant, given the fact that there are a number of distinct approaches that have been employed to validate and characterize candidate circRNAs including microarray, quantitative RT-PCR and RNA *in situ* hybridization and distinct nomenclature for diseases and traits, which preclude a systematic survey and analysis of existing corpus of evidence on disease associations of circular RNAs.

In the present report, we describe a comprehensive and standardised resource of circular RNAs and disease associations manually curated from literature and following the International Statistical Classification of Diseases (ICD) codes for standardised nomenclature of diseases.

Results

The manual curation process identified a total of 1388 circRNA transcripts originating from 720 genes and associations with 150 disease conditions mapping to a total of 113 unique disease codes as per ICD. The user interface available at <http://clingen.igib.res.in/circad/> allows users to search the collection in a user-friendly manner. The web interface allows the user to query the database in different formats as detailed in the example searches, detailed on the homepage. The user can query the database using names of circRNAs, circRNA aliases, Gene Names, Locus, Organism, Disease, ICD-10 and Pubmed identifiers. A list of circRNAs, matching the search criteria, along with the gene, organism and associated diseases, will be displayed. Detailed information about the circRNA can be obtained by clicking on the name of the circRNA.

The detailed information about each circRNA provided by the resource includes the aliases as well as linkouts to relevant databases. For candidates which have been experimentally validated, the PCR primer sequences are provided, and if not, a list of predicted primers is also provided to aid the researchers. Appropriate linkages are also provided to literature evidence supporting the disease association. A comprehensive user manual is available with details on formatting specific queries and explaining the features, navigation and options available.

Discussion

A number of complementary resources list disease and trait associations of circular RNAs. The major resources include Circ2Traits (20), circ2disease (21), circRNA disease (22) and CircR2Disease (23). Circ2trait encompasses computational predictions of circular RNAs and disease associations, by virtue of interactions modulated through miRNAs (20). circRNA disease lists 330 disease-associated circular RNAs (22) with 48 diseases, circ2diseases covers only 237 circular RNAs with 217 diseases while circR2Disease lists a total of 661 disease-associated circular RNAs associated with 100 diseases, respectively (23). None of the databases mentions comparison group, and it is important because the higher the sample size, the more the significant biomarker. Also, the abovementioned databases do not mention the fold-change difference and *P* value significance for the upregulated/downregulated expression. Circad not only reports the number of samples and comparison group but also shows the fold change and *P* value. The database has also suggested primers as well as primers used by the respective studies for analysis which is not the case in other databases. Circad lists additional information, including

method of validation, status of experimental validation, primer information and above all, a standardised nomenclature using ICD10 codes, thereby enabling standardised interpretation of associations. Circad is updated till December 2019.

Conclusion

The evidence on circular RNAs and disease associations are continuously emerging, and a process of constant updation forms the key to be relevant and comprehensive. Circad is updated on a regular basis to keep abreast and updated of the relevant evidence as they emerge. In the future, we hope to integrate the database closely with information on expression levels of the circular RNAs in different disease states, as well as information on experimentally validated molecular interactions of circular RNAs with other biomolecules in the cell. This would enable mechanistic understanding of the circular RNA functions in diseases. In short, we foresee and envisage circad as a one-stop and comprehensive resource for circular RNAs associated with diseases and a ready reference in the field of circular RNA biomarkers.

Materials and methods

Manual curation of circular RNA-disease associations

Literature databases and resources including PubMed and Google scholar were used to query and retrieve relevant annotations using a combination of keywords. PubMed was queried using standard Query (((circRNA OR circular RNA)) AND “Diseases Category”[Mesh] to retrieve all relevant publications indexed on circular RNAs and diseases. The publications were then systematically checked for information including disease for which the role of circRNA had been reported, along with its method of experimental validation (microarray, qRT-PCR or any other method) and primer information, wherever available. The information was systematically collected in a pre-formatted template. We additionally retrieved information on the fold changes observed among comparison groups, apart from the statistical significance of the observed change. Towards providing a standardised nomenclature for Diseases/Traits, we used the latest ICD-10 nomenclature, which is the 10th revision of the International Statistical Classification of Disease and related health problems by World Health Organization (WHO). Additional information on aliases and links to relevant databases for candidate circRNAs were added through manual comparison.

Database and web interface

The information from the systematic curation was ported to respective database tables in MySQL. A user-friendly web interface which enables query of the database was built using LAMP stack architecture and hosted on the Apache HTTP server. The user-interface was coded in Perl-CGI and HTML, CSS.

Author’s contributions

M.W. and D.S. curated the data for database. M.P. designed the database. D.S. designed the primers and wrote the manuscript. V.S. conceived and designed the project. All authors approved the final manuscript.

Funding

Council of Scientific and Industrial Research (CSIR, India) [Grant MLP1601].

Conflict of interest. None declared.

References

1. Jeck, W.R. and Sharpless, N.E. (2014) Detecting and characterizing circular RNAs. *Nat. Biotechnol.*, **32**, 453–461.
2. Kos, A., Dijkema, R., Arnberg, A.C. *et al.* (1986) The hepatitis delta (delta) virus possesses a circular RNA. *Nature*, **323**, 558–560.
3. Memczak, S., Jens, M., Elefsinioti, A. *et al.* (2013) Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*, **495**, 333–338.
4. Westholm, J.O., Miura, P., Olson, S. *et al.* (2014) Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep.*, **9**, 1966–1980.
5. Sablok, G., Zhao, H. and Sun, X. (2016) Plant circular RNAs (circRNAs): transcriptional regulation beyond miRNAs in plants. *Mol. Plant*, **9**, 192–194.
6. Cocquerelle, C., Mascrez, B., Héтуin, D. *et al.* (1993) Mis-splicing yields circular RNA molecules. *FASEB J.*, **7**, 155–160.
7. Greene, J., Baird, A.-M., Brady, L. *et al.* (2017) Circular RNAs: biogenesis, function and role in human diseases. *Front Mol Biosci*, **4**, 38.
8. Rong, D., Sun, H., Li, Z. *et al.* (2017) An emerging function of circRNA-miRNAs-mRNA axis in human diseases. *Oncotarget*, **8**, 73271–73281.
9. Capel, B., Swain, A., Nicolis, S. *et al.* (1993) Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell*, **73**, 1019–1030.
10. Shao, Y. and Chen, Y. (2016) Roles of circular RNAs in neurologic disease. *Front. Mol. Neurosci.*, **9**, 25.
11. Lukiw, W.J. (2013) Circular RNA (circRNA) in Alzheimer’s disease (AD). *Front. Genet.*, **4**, 307.

12. Cheng,K.H. and Hui,S.W. (1986) Correlation between bilayer destabilization and activity enhancement by diacylglycerols in reconstituted Ca-ATPase vesicles. *Arch. Biochem. Biophys.*, **244**, 382–386.
13. Dang,Y., Ouyang,X., Zhang,F. *et al.* (2017) Circular RNAs expression profiles in human gastric cancer. *Sci. Rep.*, **7**, 9060.
14. Fang,Y., Ma,M., Wang,J. *et al.* (2017) Circular RNAs play an important role in late-stage gastric cancer: circular RNA expression profiles and bioinformatics analyses. *Tumour Biol.*, **39**, 1010428317705850.
15. Sui,W., Shi,Z., Xue,W. *et al.* (2017) Circular RNA and gene expression profiles in gastric cancer based on microarray chip technology. *Oncol. Rep.*, **37**, 1804–1814.
16. Hu,W., Bi,Z.-Y., Chen,Z.-L. *et al.* (2018) Emerging landscape of circular RNAs in lung cancer. *Cancer Lett.*, **427**, 18–27.
17. Taborda,M.I., Ramírez,S. and Bernal,G. (2017) Circular RNAs in colorectal cancer: possible roles in regulation of cancer cells. *World J. Gastrointest. Oncol.*, **9**, 62–69.
18. Xia,W., Qiu,M., Chen,R. *et al.* (2016) Circular RNA has_circ_0067934 is upregulated in esophageal squamous cell carcinoma and promoted proliferation. *Sci. Rep.*, **6**, 35576.
19. Qiu,M., Xia,W., Chen,R. *et al.* (2018) The circular RNA circPRKCI promotes tumor growth in lung adenocarcinoma. *Cancer Res.*, **78**, 2839–2851.
20. Ghosal,S., Das,S., Sen,R. *et al.* (2013) Circ2Traits: a comprehensive database for circular RNA potentially associated with disease and traits. *Front. Genet.*, **4**, 283.
21. Yao,D., Zhang,L., Zheng,M. *et al.* (2018) Circ2Disease: a manually curated database of experimentally validated circRNAs in human disease. *Sci. Rep.*, **8**, 11018.
22. Zhao,Z., Wang,K., Wu,F. *et al.* (2018) circRNA disease: a manually curated database of experimentally supported circRNA-disease associations. *Cell Death Dis.*, **9**, 475.
23. Fan,C., Lei,X., Fang,Z. *et al.* (2018) CircR2Disease: a manually curated database for experimentally supported circular RNAs associated with various diseases. *Database*, **2018**.