Circan: A Database of Circular RNAs Exploring Chromosomal Linkages in Human Cancers

Gaurav Kumar Bhagat[(✉)]

Department of Biochemistry and Biophysics, University of Kalyani, Nadia, West Bengal 741235, India
gauravbcbp22@klyuniv.ac.in, gbnitdgp@yahoo.co.in

Abstract. Circular RNAs (circRNAs) represent a milieu of non-coding RNAs that play a major role in gene regulation, development, and carcinogenesis. Circan is a database of cancer associated circular RNAs. As the interest in circRNA increases, Circan provides a systematic annotation of circRNA dysregulation in cancer and a helpful asset for studying their prominent role in cancer. Circan acts as a functional resource to efficiently browse and search various circRNAs and their annotations. This work includes circRNA name associated with each cancer subtype, gene, location of the circRNA on the chromosome, genomic length, spliced length, the PubMed reference, nucleotide sequence of each circRNA, sample description, experimental method, expression pattern, genomic context and miRNAs involved in cancer progression. On data analysis, it was seen that lesser number of cancers causing circRNAs were present on chromosomes 13, 21, 22, X, Y whereas more number were found on chromosomes 1, 2, 3, 6, 11 which might indicate the favored location of cancer causing circRNAs on the genome. As the interest in circRNAs keeps on increasing, the database will improve our knowledge of circRNA dysregulation. The URL is http://bhagatjeeesolutions.com/Cancer_circRNA.aspx

Keywords: Cancer · Chromosome · CircRNA · Database

1 Introduction

Circular RNAs (circRNAs) are distinctive transcript isoforms (regulatory non-coding RNA) forming a covalently closed continuous loop characterized by a circular conformation. They are mostly discovered by high throuput sequencing and found to be evolutionary conserved, stable and abundant among eukaryotes [1].

The first circular RNA was found four decades ago. Circular RNAs are found in circular form in the cytoplasm of eukaryotic cells. Latest research shows that they are regulated by key genes and produced via a coordinated biological process having characteristic biological functions [2]. Emerging evidence on the functions of circRNAs depicts that the functional annotation of most of the circRNAs is still not known properly and requires detailed experiments [3]. The potential of circRNAs to act as a biomarker is suggested by many research studies after analyzing their expression patterns in cancer patients. These studies include different types of cancers like colorectal cancer [4], gastric cancer [5], lung cancer [6, 7], cervical cancer [8], hepatocellular cancer [9, 10] and
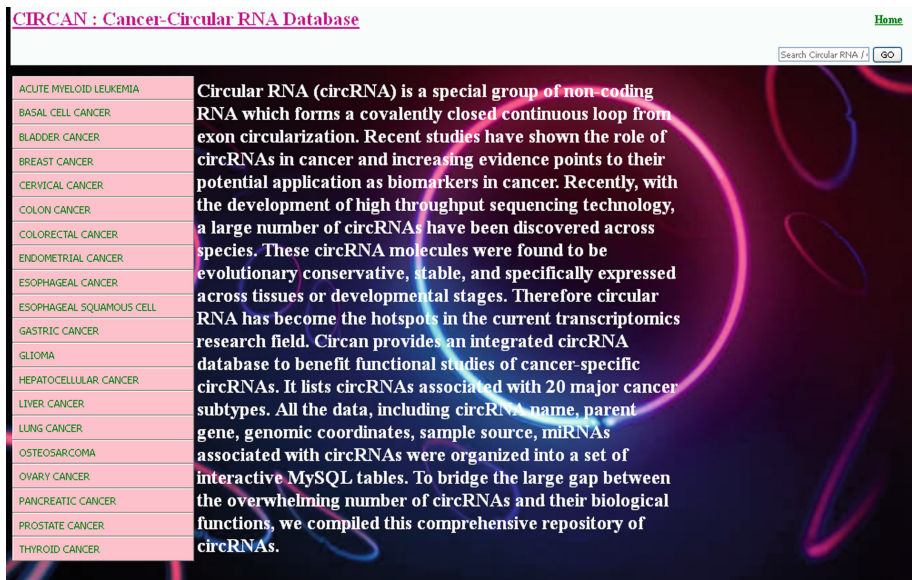
Fig. 1. The home page of circan database.

esophageal cancer [11]. Circular RNA expression can be cell-specific [12] or tissue-specific [13]. A fusion-circRNA obtained from fusion genes interacts with the fusion protein and plays a major role in cancer development [14]. Alternative splicing has a major impact on the biogenesis of circRNA [15]. Till today, the scientific community has found it difficult to characterize the functional features of most circRNAs. CircHIPK3 is a lethal circular RNA which has been found to be involved in all types of cancer. CircRNAs are found in large quantities in exosomes, saliva and blood samples making them excellent diagnostic biomarkers for cancer. They work like miRNA sponges and form RNA-protein complexes by binding to RNA-associated proteins that regulate gene transcription. Circular RNAs are found to be more stable than linear RNAs due to the absence of 3' termini [16].

The present knowledge on circRNA function remains inadequate. Circan database helps in integrating all the relevant information found from various studies to facilitate further research on their role in cancer. We developed this database so that the user can get useful knowledge regarding the circRNAs and their relations with various cancer types. This database will help in analyzing several characteristics of circRNAs in different chromosomes in various cancer types and identifying the pre-cursors of carcinogenesis.

2 Materials and Methods

2.1 Data Extraction from Literature

The circular RNA details were collected from research publications. Scientific publications related to circRNAs interacting with cancer were manually curated. The keywords

'circRNA' and 'circular RNA' were used to query and obtain relevant information from literature databases like PubMed. The disease was limited to 'cancer' mentioned in publications and species to human. Around 5000 research articles were found using these keywords**.** The publications were then examined for the type of cancer with which the circRNA had been reportedly associated with. The location of circRNA associated with the gene on the chromosome was noted along with the strand (either positive or negative) on which it was found. Other relevant information like gene symbol, genomic length, spliced length, PubMed-Id, expression pattern, sample description, method, genomic context and miRNAs involved in cancer progression were also collected from the publications.

2.2 Retrieval of Sequence Information of the CircRNAs

The sequences of various circRNAs were collected from UCSC genome browser [17] using the data collected from the publications [18]. Since the genome locus is known it is easy to extract the sequence information from UCSC genome browser. For example, if the location of the circular RNA is chr11:33307958–33309057, then the starting and end positions are known and this is entered in UCSC genome browser to obtain the sequence information.

2.3 Methodology of Development

All the data related to circRNAs was collected in Excel sheets. This data from the manual curation was exported to MySQL tables. The website has been developed to execute the SQL queries dynamically. The user-interface was coded using Dot Net framework. The website home page contains links (on the left side, Fig. 1) to 20 types of cancers arranged in alphabetical order for easy access to researchers. On clicking any type of cancer, a new webpage appears which displays all the circRNA information involved with that type of cancer.

3   Results and Discussions

3.1 Categorization of CircRNAs

The various circRNAs were inscribed as per the type of cancer. This led to a proper arrangement of various circRNAs involved in cancer proliferation. The study was conducted for 20 major cancer subtypes and a total of 1161 circRNA transcripts were found to be involved.The website is available at http://bhagatjeeesolutions.com/Cancer_cir cRNA.aspx facilitating users to browse the database. A user can search the database on the homepage using the name of circRNA or gene. A collection of circRNAs corresponding to the criteria along with all other relevant information will be listed. The nucleotide stretch of the corresponding circRNA can be viewed by clicking on the 'Sequence' button in the relevant column.
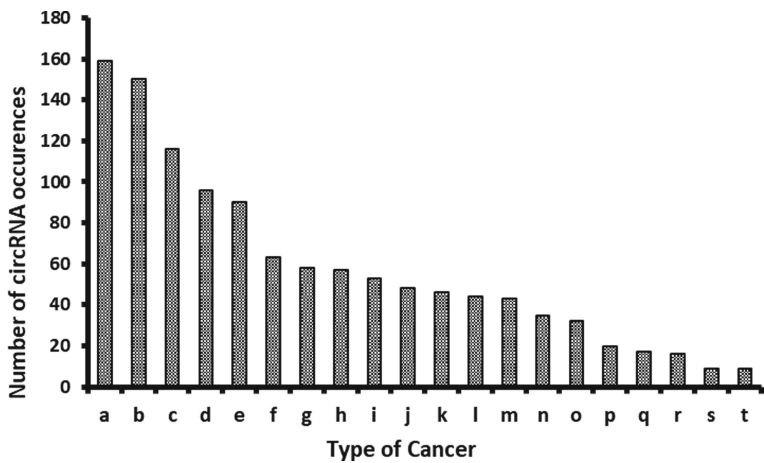
Fig. 2. Statistical analysis of the database datasets (from left to right based on the alphabetic order: gastric, breast, lung, glioma, colorectal, bladder, hepatocellular, liver, esophageal squamous cell, cervical, osteosarcoma, thyroid, pancreatic, AML, ovary, colon, basal cell, prostate, esophageal, endometrial).

3.2   Analysis of CircRNAs

An analysis was performed on the entire database dataset (Fig. 2) and it was found that gastric cancer had the highest number of circRNA associations: around 159. The least number of circRNAs (6) were found in endometrial cancer subtype.

The total number of circRNAs responsible for cancer associated with each chromosome was calculated from the database dataset. It was then compared to the total size of the chromosome in Fig. 3. Thus, a comparative analysis was performed on the cancer causing circRNAs with the total size of each chromosome in Fig. 3. Similarly, the total number of circRNAs responsible for cancer associated with each chromosome was compared to the count of protein coding genes on each chromosome in Fig. 4. It was done to evaluate the correlation between cancer causing circRNAs and their preferred location on the chromosome. From the figures it was seen that lesser number of cancers causing circRNAs were present on chromosomes 13, 21, 22, X, Y whereas more number were found on chromosomes 1, 2, 3, 6, 11 which might indicate the favored location of cancer causing circRNAs on the genome. This can be fortified by the discovery of more circRNAs which would elucidate intricate details of the circular RNA functions.

Circan can act as a ready reference and comprehensive resource for cancer associated circular RNAs. Its contribution could prove significant to research involved with regulation of cancer-associated circRNAs. We believe that Circan will help in understanding the various relationships between circRNAs and carcinogenesis. Circan will be updated regularly in the future when more sequencing data related to circRNA-cancer association is obtained.
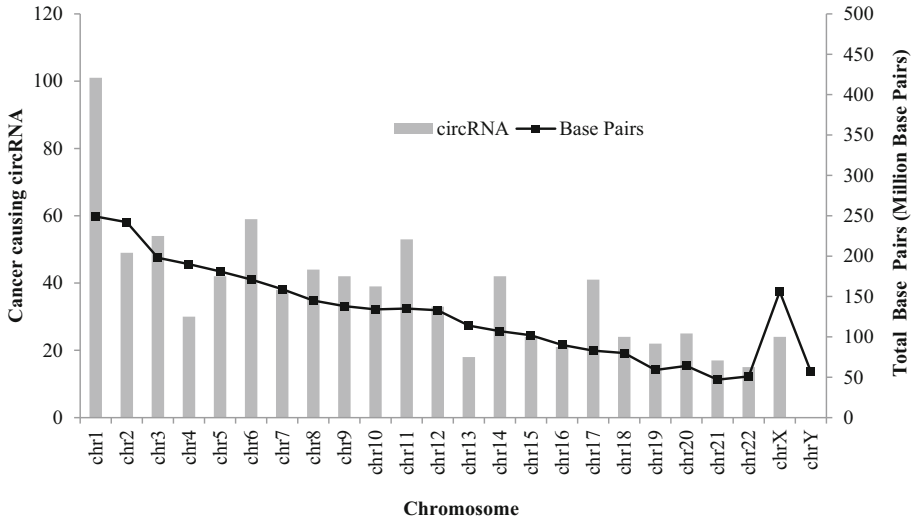
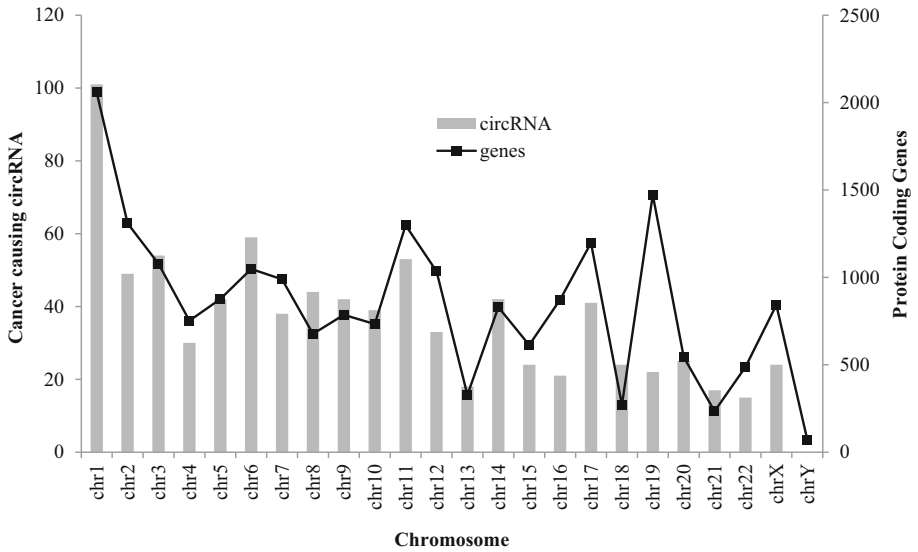**Fig. 3.** Comparative analysis of chromosome-wise data.

Fig. 4. Comparative analysis for each chromosome.

References

1. Jeck,W.R. and Sharpless,N.E. Detecting and characterizing circular RNAs. Nat. Biotechnol., 2014, 32:453–461. doi: https://doi.org/10.1038/nbt.2890.

2. Greene,J., Baird,A.M., Brady,L. *et al.* Circular RNAs: biogenesis, function and role in human diseases. Front Mol Biosci., 2017, 4:38. doi: https://doi.org/10.3389/fmolb.2017.00038.

3. Rong,D., Sun,H., Li,Z. *et al.* An emerging function of circRNA-miRNAs-mRNA axis in human diseases. Oncotarget, 2017, 8(42):73271–73281. doi: https://doi.org/10.18632/oncotarget.19154.

4. Taborda,M.I., Ramírez,S. and Bernal,G. Circular RNAs in colorectal cancer: possible roles in regulation of cancer cells. World J. Gastrointest. Oncol, 2017, 9(2):62–69. doi: https://doi.org/10.4251/wjgo.v9.i2.62.

5. Dang,Y., Ouyang,X., Zhang,F. *et al.* Circular RNAs expression profiles in human gastric cancer. Sci. Rep., 2017, 7(1):9060. doi: https://doi.org/10.1038/s41598-017-09076-6.

6. Hu,W., Bi,Z.-Y., Chen,Z.-L. *et al.* Emerging landscape of circular RNAs in lung cancer. Cancer Lett., 2018, 427:18–27. doi: https://doi.org/10.1016/j.canlet.2018.04.006.

7. Peng H, Zhang W, Dong H, Yuan J, Li Y, Li F, Yu D, Guan Y, Zhang F. CircFAT1 Promotes Lung Adenocarcinoma Progression by Sequestering miR-7 from Repressing IRS2-ERK-mediated CCND1 Expression. Int J Biol Sci., 2022;18(10):3944-3960. doi: https://doi.org/10.7150/ijbs.70889.

8. Liu Y, Li L, Yang Z, Wen D, Hu Z. Circular RNA circACAP2 Suppresses Ferroptosis of Cervical Cancer during Malignant Progression by miR-193a-5p/GPX4. J Oncol., 2022, 2022:5228874. doi: https://doi.org/10.1155/2022/5228874.

9. Fei Z, Wang Y, Gu Y, Xie R, Hao Q, Jiang Y. CircKIF5B Promotes Hepatocellular Carcinoma Progression by Regulating the miR-192 Family/XIAP Axis. Front Oncol., 2022;12:916246.doi: https://doi.org/10.3389/fonc.2022.916246.

10. Huang J, Zhou H, Diao Y, Yang Z. Hsa_circ_0000285 knockdown inhibits the progression of hepatocellular carcinoma by sponging miR-582-3p to regulate CCNB2 expression. Hum Exp Toxicol., 2022;41:9603271221115400. doi: https://doi.org/10.1177/09603271221115400.

11. Xi Y, Shen Y, Wu D, Zhang J, Lin C, Wang L, Yu C, Yu B, Shen W. CircBCAR3 accelerates esophageal cancer tumorigenesis and metastasis via sponging miR-27a-3p. Mol Cancer. 2022 ;21(1):145. doi: https://doi.org/10.1186/s12943-022-01615-8.

12. Salzman,J., Chen,R.E., Olsen,M.N., Wang,P.L. and Brown,P.O. Cell-type specific features of circular RNA expression. PLoS Genet., 2013, 9(9):e1003777. doi:https://doi.org/10.1371/journal.pgen.1003777.

13. Xia,S., Feng,J., Lei,L., Hu,J., Xia,L., Wang,J., Xiang,Y., Liu,L., Zhong,S., Han,L. *et al.* Comprehensive characterization of tissue-specific circular RNAs in the human and mouse genomes. Brief. Bioinformatics, 2017, 18(6):984-992. doi: https://doi.org/10.1093/bib/bbw081.

14. Guarnerio,J., Bezzi,M., Jeong,J.C., Paffenholz,S.V., Berry,K., Naldini,M.M., Lo-Coco,F., Tay,Y., Beck,A.H. and Pandolfi,P.P. Oncogenic role of fusion-circRNAs derived from cancer-associated chromosomal translocations. Cell, 2016, 165(2):289–302. doi: https://doi.org/10.1016/j.cell.2016.03.020.

15. Qu,S., Yang,X., Li,X., Wang,J., Gao,Y., Shang,R., Sun,W., Dou,K. and Li,H. Circular RNA: a new star of noncoding RNAs. Cancer Lett., 2015, 365(2):141–148. doi: https://doi.org/10.1016/j.canlet.2015.06.003.

16. Zheng,Q., Bao,C., Guo,W., Li,S., Chen,J., Chen,B., Luo,Y., Lyu,D., Li,Y., Shi,G. *et al.* Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. Nat. Commun., 2016, 7:11215. doi: https://doi.org/10.1038/ncomms11215.

17. Karolchik,D., Baertsch,R., Diekhans,M., Furey,T.S., Hinrichs,A., Lu,Y.T., Roskin,K.M., Schwartz,M., Sugnet,C.W., Thomas,D.J. *et al.* The UCSC Genome Browser Database. *Nucleic Acids Research*, 2003, 31(1):51–54. doi: https://doi.org/10.1093/nar/gkg129.

18. Fang S, Zhang L, Guo J, Niu Y, Wu Y, Li H, Zhao L, Li X, Teng X, Sun X, Sun L, Zhang MQ, Chen R, Zhao Y. NONCODEV5: a comprehensive annotation database for long non-coding RNAs. *Nucleic Acids Research*; 2018, 46(D1):D308-D314. doi: https://doi.org/10.1093/nar/gkx1107.