Title

CircFunBase: a database for functional circular RNAs.

Permalink

Authors

Meng, Xianwen
Hu, Dahui
Zhang, Peijing
et al.

Publication Date

2019

DOI

10.1093/database/baz003

Copyright Information

Peer reviewed

Database tool

CircFunBase: a database for functional circular RNAs

Xianwen Meng[1,2,†], Dahui Hu[1,†], Peijing Zhang[1], Qi Chen[1] and Ming Chen[1,*]

[1]Department of Bioinformatics, the State Key Laboratory of Plant Physiology and Biochemistry, Institute of Plant Science, College of Life Sciences, Zhejiang University, Hangzhou 310058, China and [2]The State Key Laboratory of Crop Biology, Shandong Key Laboratory of Crop Biology, College of Agronomy, Shandong Agricultural University, Tai'an 271000, China

*Corresponding author: Tel: +86(0)571-88206612; Fax: +86(0)571-88206612; Email: mchen@zju.edu.cn
†These authors contributed equally to this work.

Abstract

Increasing evidence reveals that circular RNAs (circRNAs) are widespread in eukaryotes and play important roles in diverse biological processes. However, a comprehensive functionally annotated circRNA database is still lacking. CircFunBase is a web-accessible database that aims to provide a high-quality functional circRNA resource including experimentally validated and computationally predicted functions. The current version of CircFunBase documents more than 7000 manually curated functional circRNA entries, mainly including *Homo sapiens*, *Mus musculus* etc. CircFunBase provides visualized circRNA-miRNA interaction networks. In addition, a genome browser is provided to visualize the genome context of circRNAs. As a biological information platform for circRNAs, CircFunBase will contribute for circRNA studies and bridge the gap between circRNAs and their functions.

Database URL: http://bis.zju.edu.cn/CircFunBase

Introduction

Circular RNAs are a special group of endogenous RNAs characterized by covalent closed loops (1). CircRNAs are identified to be widespread in both animals and plants through high-throughput sequencing technology coupled with bioinformatics analysis (2–5). Although their biological functions remain largely unknown, increasing evidence suggests that they play important roles in the regulation of multiple biological processes, especially in human diseases (6, 7). One surprising finding is that a novel circular RNA, ciRS-7, could function as a designated miR-7 sponge (8, 9), and, by this way, ciRS-7 might be correlated with human cancers (10–12).

Table 1. Web-based resource for circRNAs

Name	Sample	Description	Address	Reference
circBase	Animals	A unified circRNA data set from previous publications	http://circbase.org	(14)
CIRCpedia	Animals	CircRNA annotations from over 180 RNA-seq data sets across six different species	http://www.picb.ac.cn/ rnomics/circpedia	(15)
CircR2Disease	Animals	Experimentally supported associations between circRNAs and diseases	http://bioinfo.snnu.edu. cn/CircR2Disease	(17)
Circ2Traits	Human	CircRNAs potentially associated with disease and traits	http://gyanxet-beta.com/ circdb	(18)
CircInteractome	Human	Interacting miRNAs and RNA-binding proteins of circRNAs	https://circinteractome. nia.nih.gov	(19)
CircNet	Human	CircRNA-miRNA-mRNA interaction networks	http://syslab5.nchu.edu. tw/CircNet	(20)
circRNADb	Human	CircRNAs with protein-coding potential	http://reprod.njmu.edu. cn/circrnadb	(21)
CSCD	Human	Cancer-specific circRNAs	http://gb.whu.edu.cn/ CSCD	(22)
TSCD	Human and mouse	Tissue-specific circRNAs	http://gb.whu.edu.cn/ TSCD	(23)
PlantcircBase	Plants	Plant circRNAs	http://ibi.zju.edu.cn/ plantcircbase	(24)
PlantCircNet	Plants	Providing plant circRNA-miRNA-gene regulatory networks, as well as circRNA information and circRNA expression profiles	http://bis.zju.edu.cn/ plantcircnet	(16)

New circRNA detection tools are constantly being developed (13). Together with high-throughput sequencing technologies, thousands of circRNAs in both animals and plants have been reported and integrated into the circRNA databases, such as circBase (14), CIRCpedia (15) and PlantCircNet (16) (Table 1). Although these circRNA databases provide the tissue information, the functions of circRNAs in a specific tissue are not clear. Circ2Traits associates circRNAs with human diseases based on the interactions of circRNAs with disease-associated miRNAs and disease-associated single nucleotide polymorphisms mapped on circRNA loci. However, these potential associations of circRNAs with diseases in human remain to be validated experimentally. Although CircR2Disease is a manually database, it is limited to the circRNAs in the context of diseases.

The number of publications about circRNA research keeps increasing rapidly in recent years (25), and the findings from these publications are essential for further studying circRNA functions. To bridge the gap between circRNAs and their functions, we collect current findings about circRNA functions from literature and form a unique functional circRNA resource (CircFunBase, http://bis.zju. edu.cn/CircFunBase). The current version of CircFunBase contains more than 7000 manually curated functional circRNAs, involving 15 organisms (such as *Homo sapiens* and *Mus musculus*). Hence, CircFunBase serves as a more specific functional circRNA resource to efficiently investigate, browse a particular circRNA and provide insights into its function. These functional circRNAs can be easily queried and downloaded through the webpage. In addition, CircFunBase allows researchers to submit novel functional circRNAs.

Materials and methods

In order to collect all functional circRNAs, we screened all of the literature in the PubMed database with the following keywords: 'circular RNA', 'circRNA' or 'RNA circularization'. The relevant hits were downloaded and further inspected manually. We extracted information on functional circRNAs, their related diseases and biological regulations. The GO annotations were downloaded from the Gene Ontology Consortium (26). Besides, the RNA-binding proteins (RBPs) matching to human functional circRNAs were from CircInteractome (19), which provides
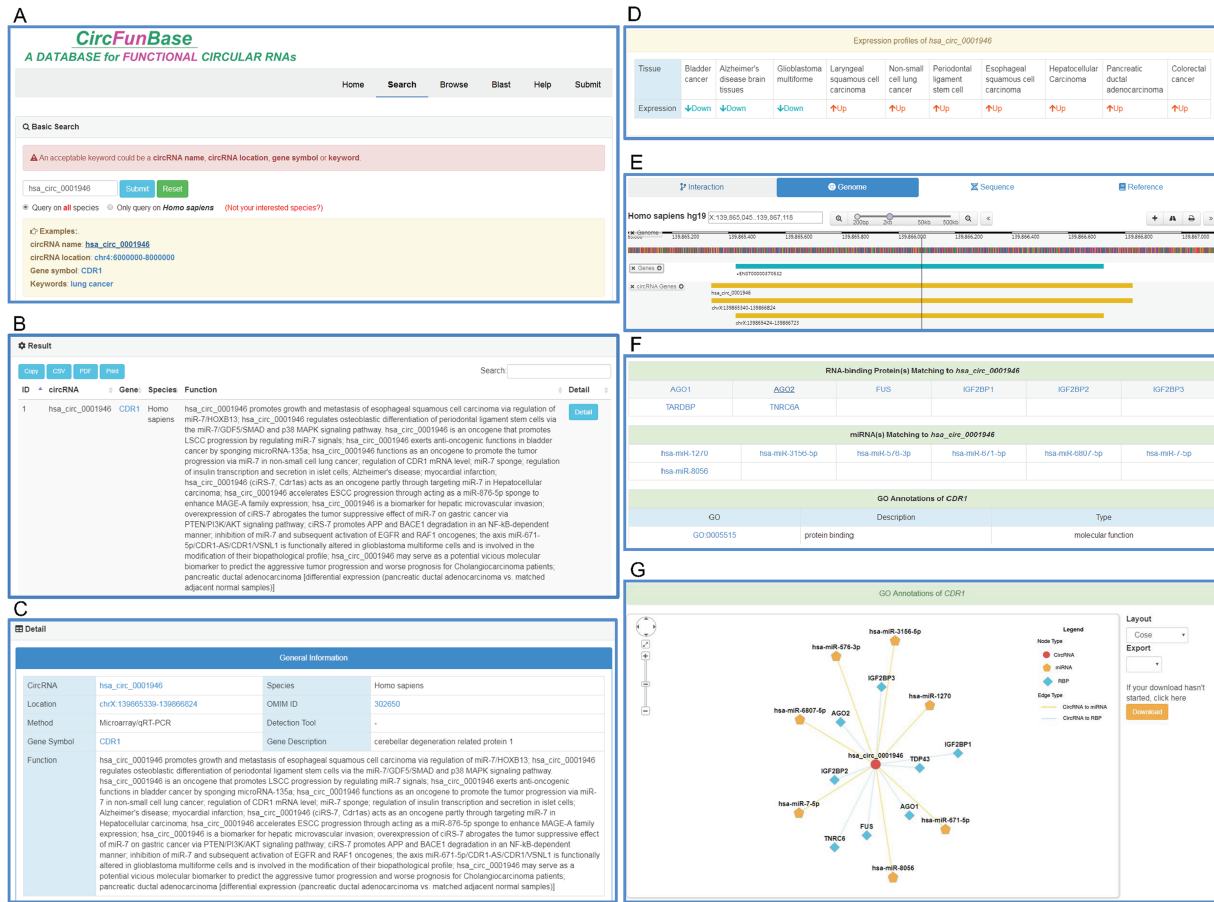
Figure 1. An overview of the user interface of CircFunBase. (A) The search page of CircFunBase. (B) Hsa_circ_0001946 as an example of input and the search result. (C) Detailed information of hsa_circ_0001946. (D) The expression pattern of hsa_circ_0001946. (E) Visualization of the circRNA genome location. (F) CircRNA-associated interactions and GO annotations of parent genes. (G) Visualization of circRNA-associated interactions.

a comprehensive binding map of RBPs to circRNA using cross-linking immunoprecipitation data. Mature miRNA sequences were acquired from miRBase (27), while animal and plant miRNA-circRNA interactions were predicted by miRanda (28) and TargetFinder (29), respectively.

The CircFunBase database is implemented using HTML and PHP languages with MySQL. The interface component consists of web pages designed and implemented in HTML/CSS. Cytoscape.js (30) was used to visualize the circRNA-associated networks, while Dalliance (31) was used to view the genome. The BLAST module was implemented using SequenceServer (32).

Results

Database contents

Functional circRNAs were manually obtained from articles published in the PubMed database before 1 May 2018. In the current version, CircFunBase documents 7059 functional circRNA entries from 15 organisms, including 7

plants and 8 animals. Each entry contains circRNA name, position, tissue, expression pattern, detection tool, function, gene symbol, gene description, PubMed ID, GO annotations and circRNA-associated miRNAs. Particularly, for human, the OMIM ID of parent gene and circRNA-associated RBPs are provided. In addition, a network viewer is used to visualize miRNA-circRNA and RBP-circRNA interactions, and a genome browser is used to view the circRNA genome. Figure 1 illustrates an overview of the user interface of CircFunBase database.

CircFunBase provides the 'Submit' page, inviting researchers to upload novel functional circRNAs. In the 'Help' page, instructions for using CircFunBase are available and functional circRNA list from each species could be downloaded.

Data querying, searching and browsing

CircFunBase provides a user-friendly interface for retrieval of functional circRNAs in the 'Search' page. Users can retrieve data by circRNA name, circRNA location, gene symbol or keywords (such as 'lung cancer' for human). Cir-
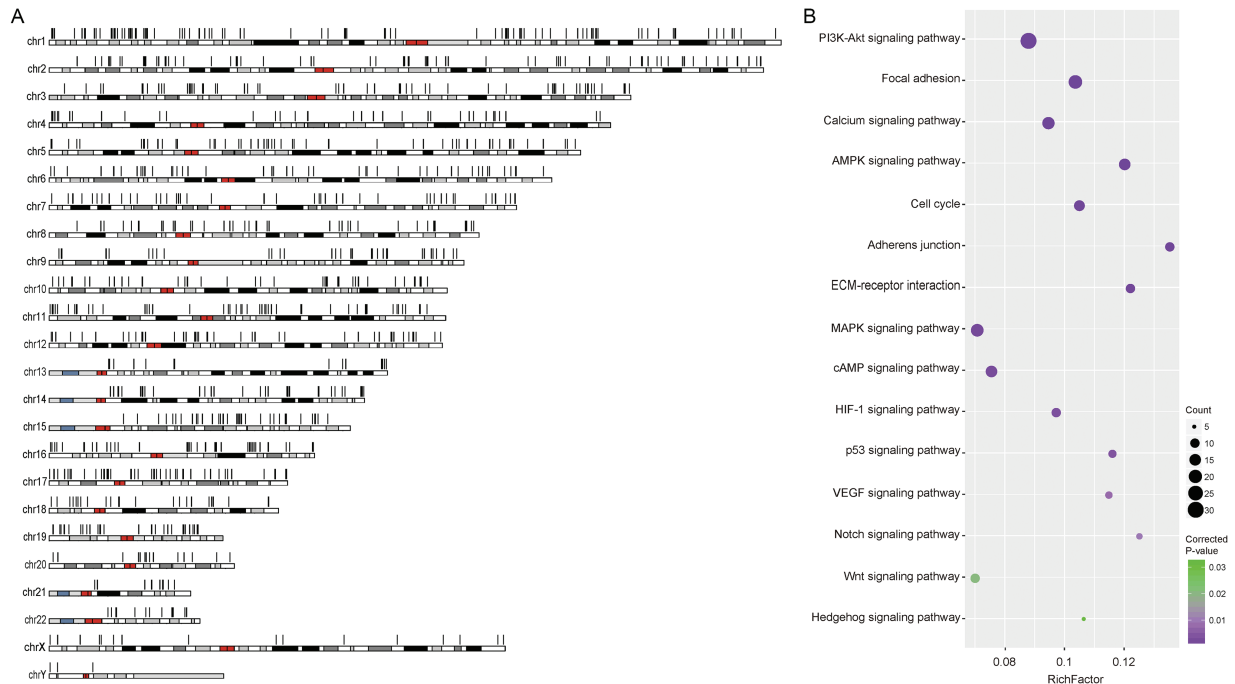
Figure 2. Human cancer circRNAs. (A) Genomic distribution of human cancer circRNAs. (B) Enriched KEGG pathways of parent genes of cancer circRNAs.

cFunBase provides brief description of search results in the 'Search Result' page. To gain more detail information of a specific circRNA, users can click the 'Details' button. Additional information such as PubMed ID, GO annotations and circRNA-associated miRNAs are displayed in the circRNA single-record page. We also provide many useful hyperlinks: (1) human and mouse circRNAs are linked to circBase and could be viewed in UCSC Genome Browser; (2) gene symbol is linked to the NCBI Gene (33) [for *Solanum lycopersicum*, gene symbol is linked to Ensembl (34)]; (3) clicking the GO term links to the GO Consortium website; and (4) miRNA name is linked to mirBase. A BLAST module was provided to query circRNAs in CircFunBase using RNA sequences. In 'Browse' page, users can browse all the functional circRNAs in a species by clicking the icon of corresponding species or using the lineage tree. In addition, CircFunBase provides a series of APIs to return detailed information about circR-NAs in JSON format, for example, circRNA information or miRNA interaction information.

The parent genes of cancer circRNAs are highly correlated with cancer pathways

To highlight the significance of CircFunBase, we presented a biological application based on this data resource. During data collection, we found that a large amount of human circRNAs were correlated to cancers. Previous studies have revealed that most human circRNAs are generated from exon regions of protein-coding genes. Whether the parent genes share similar functions with cancer circRNAs remains to be elucidated. Therefore, we explored this issue using the circRNAs annotated in CircFunBase.

First, we retrieved human cancer circRNAs in Circ-FunBase using the keyword 'cancer', and we got 1712 cancer circRNAs—of which, 92.87% circRNAs are exonic circRNAs (Figure 2A). Particularly, most cancer circRNAs (96.14%) were identified according to their differential expression in the context of cancer and the functional relationship between circRNAs and their parent genes remains to be investigated. Then, we performed functional enrichment analysis on the parent genes of these cancer circRNAs using KOBAS (35). The results showed that the parent genes of cancer circRNAs were highly correlated with cancer since most of them were significantly enriched in cancer-related pathways (Figure 2B). For example, both circITCH and ITCH were implicated in human cancers by regulating the Wnt signalling pathway (36, 37). Thus, the circRNAs share similar functions with their parent genes in the context of human cancer, and the circRNA functions in cancers can be predicted according to the functions of their parent genes. Actually, circRNA study provides new insights into the mechanisms of human diseases.

Discussion

As a comprehensive functional circRNA database, Cir-cFunBase is designed to provide a rich data resource for circRNA study. CircFunBase has collected recently

discovered functional circRNAs from relevant literature. Currently, more than 7000 functional circRNAs from 15 model species are deposited in CircFunBase. About 75.83% human circRNAs annotated in CircFunBase overlap with the current circRNA reference database, circBase. As circBase does not provide functional information of circRNAs, which is available in CircFunBase, our database will definitely help researchers to better understand circRNA biology. To gain insight into circRNA regulations, CircFunBase provides the circRNA associated RBPs and miRNAs. These interactions will benefit further studies on circRNA functions. Take hsa_circ_0001946 for example, except for experimentally validated miR-7-hsa_circ_0001946 and miR-671-hsa_circ_0001946 interactions, other five miRNAs (miR-576, miR-1270, miR-3156, miR-6807 and miR-8056) targeting hsa_circ_0001946 were also identified in CircFunBase, the biological functions of these novel interactions in specific contexts remain to be explored in the future. Another biological application based on the resource provided by CircFunbase revealed that the parent genes of cancer circRNAs are highly correlated with cancer pathways. The functional consistency between parent genes and circRNAs in human cancers is of great significance for exploring the mechanisms of cancers.

To our knowledge, CircFunBase is the first database focusing on circRNA functions in diverse species. It will bridge the gap in circRNAs and functional research and further facilitate biologists in unveiling the roles of circRNAs in diverse biological processes. We will update CircFunBase regularly with newly published data. In addition, direct data submission by the researchers is supported. With the improvement of CircFunBase, it is expected to become a valuable data resource and serve as a foundation for future circRNA study.

References

1. Wilusz,J.E. (2018) A 360 degrees view of circular RNAs: from biogenesis to functions. *Wiley Interdiscip. Rev. RNA*, **9**, 1–17.
2. Salzman,J., Chen,R.E., Olsen,M.N. *et al.* (2013) Cell-type specific features of circular RNA expression. *PLoS Genet.*, **9**, e1003777.
3. Wang,P.L., Bao,Y., Yee,M.C. *et al.* (2014) Circular RNA is expressed across the eukaryotic tree of life. *PLoS One*, **9**, e90859.
4. Ye,C.Y., Chen,L., Liu,C. *et al.* (2015) Widespread noncoding circular RNAs in plants. *New Phytol.*, **208**, 88–95.
5. Salzman,J., Gawad,C., Wang,P.L. *et al.* (2012) Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One*, **7**, e30733.
6. Wang,Y., Mo,Y., Gong,Z. *et al.* (2017) Circular RNAs in human cancer. *Mol. Cancer*, **16**, 25.
7. Meng,X., Li,X., Zhang,P. *et al.* (2017) Circular RNA: an emerging key player in RNA world. *Brief. Bioinform.*, **18**, 547–557.
8. 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.
9. Hansen,T.B., Jensen,T.I., Clausen,B.H. *et al.* (2013) Natural RNA circles function as efficient microRNA sponges. *Nature*, **495**, 384–388.
10. Peng,L., Yuan,X.Q. and Li,G.C. (2015) The emerging landscape of circular RNA ciRS-7 in cancer (Review). *Oncol. Rep.*, **33**, 2669–2674.
11. Yu,L., Gong,X., Sun,L. *et al.* (2016) The circular RNA Cdr1as act as an oncogene in hepatocellular carcinoma through targeting miR-7 expression. *PLoS One*, **11**, e0158347.
12. Weng,W., Wei,Q., Toden,S. *et al.* (2017) Circular RNA ciRS-7-A promising prognostic biomarker and a potential therapeutic target in colorectal cancer. *Clin. Cancer Res.*, **23**, 3918–3928.
13. Gao,Y. and Zhao,F. (2018) Computational strategies for exploring circular RNAs. *Trends Genet.*, **34**, 389–400.
14. Glazar,P., Papavasileiou,P. and Rajewsky,N. (2014) circBase: a database for circular RNAs. *RNA*, **20**, 1666–1670.
15. Zhang,X.O., Dong,R., Zhang,Y. *et al.* (2016) Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res.*, **26**, 1277–1287.
16. Zhang,P., Meng,X., Chen,H. *et al.* (2017) PlantCircNet: a database for plant circRNA-miRNA-mRNA regulatory networks. *Database*, **2017**, 1–8.
17. Yao,D., Zhang,L., Zheng,M. *et al.* (2018) CircR2Disease: a manually curated database for experimentally supported circular RNAs associated with various diseases. *Database*, **2018**, 1–6.
18. 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.
19. Dudekula,D.B., Panda,A.C., Grammatikakis,I. *et al.* (2016) CircInteractome: a web tool for exploring circular RNAs and their interacting proteins and microRNAs. *RNA Biol.*, **13**, 34–42.
20. Liu,Y.C., Li,J.R., Sun,C.H. *et al.* (2016) CircNet: a database of circular RNAs derived from transcriptome sequencing data. *Nucleic Acids Res.*, **44**, D209–D215.

21. Chen,X., Han,P., Zhou,T. *et al.* (2016) circRNADb: a comprehensive database for human circular RNAs with protein-coding annotations. *Sci. Rep.*, **6**, 34985.

22. Xia,S., Feng,J., Chen,K. *et al.* (2018) CSCD: a database for cancer-specific circular RNAs. *Nucleic Acids Res.*, **46**, D925–D929.

23. Xia,S., Feng,J., Lei,L. *et al.* (2017) Comprehensive characterization of tissue-specific circular RNAs in the human and mouse genomes. *Brief. Bioinform.*, **18**, 984–992.

24. Chu,Q., Zhang,X., Zhu,X. *et al.* (2017) PlantcircBase: a database for plant circular RNAs. *Mol. Plant*, **10**, 1126–1128.

25. Feng,J., Xiang,Y., Xia,S. *et al.* (2017) CircView: a visualization and exploration tool for circular RNAs. *Brief. Bioinform.*, **19**, 1310–1316.

26. The Gene Ontology Consortium (2015) Gene Ontology Consortium: going forward. *Nucleic Acids Res.*, **43**, D1049–D1056.

27. Kozomara,A. and Griffiths-Jones,S. (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.*, **42**, D68–D73.

28. Enright,A.J., John,B., Gaul,U. *et al.* (2003) MicroRNA targets in *Drosophila*. *Genome Biol.*, **5**, R1.

29. Allen,E., Xie,Z., Gustafson,A.M. *et al.* (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, **121**, 207–221.

30. Franz,M., Lopes,C.T., Huck,G. *et al.* (2016) Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics*, **32**, 309–311.

31. Down,T.A., Piipari,M. and Hubbard,T.J. (2011) Dalliance: interactive genome viewing on the web. *Bioinformatics*, **27**, 889–890.

32. Priyam,A., Woodcroft,B.J., Rai,V., *et al.* (2015) Sequenceserver: a modern graphical user interface for custom BLAST databases. Biorxiv, doi: 10.1101/033142.

33. Brown,G.R., Hem,V., Katz,K.S. *et al.* (2015) Gene: a gene-centered information resource at NCBI. *Nucleic Acids Res.*, **43**, D36–D42.

34. Kersey,P.J., Allen,J.E., Armean,I. *et al.* (2016) Ensembl Genomes 2016: more genomes, more complexity. *Nucleic Acids Res.*, **44**, D574–D580.

35. Xie,C., Mao,X., Huang,J. *et al.* (2011) KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.*, **39**, W316–W322.

36. Li,F., Zhang,L., Li,W. *et al.* (2015) Circular RNA ITCH has inhibitory effect on ESCC by suppressing the Wnt/beta-catenin pathway. *Oncotarget*, **6**, 6001–6013.

37. Wei,W., Li,M., Wang,J. *et al.* (2012) The E3 ubiquitin ligase ITCH negatively regulates canonical Wnt signaling by targeting dishevelled protein. *Mol. Cell. Biol.*, **32**, 3903–3912.