Database, 2023, 1–7
DOI: https://doi.org/10.1093/database/baad082
Original article

PlantIntronDB: a database for plant introns that host functional elements

Weiping Wang[‡], Jiming Hu[‡], Han Li, Jun Yan and Xiaoyong Sun[*]

Agricultural Big Data Research Center, College of Information Science and Engineering, Shandong Agricultural University, Taian, Shandong 271018, China

[*]Corresponding author: Tel: (86)538-8249879; Fax: (86)538-8241878; Email: sunx1@sdau.edu.cn
[‡]These authors contributed equally to this work.

Citation details: Wang, W., Hu, J., Li, H. *et al.* PlantIntronDB: a database for plant introns that host functional elements. *Database* (2023) Vol. 2023: article ID baad082; DOI: https://doi.org/10.1093/database/baad082

Abstract

Although more and more attention has been focused on introns and the important role of plant introns in plant growth and development has been discovered, there is still a lack of an open and comprehensive database on plant introns with functional elements in current research. In order to make full use of large-scale sequencing data and help researchers in related fields to achieve high-throughput functional verification of identified plant introns with functional elements, we designed a database containing five plant species, PlantIntronDB and systematically analyzed 358, 59, 185, 210 and 141 RNA-seq samples from *Arabidopsis thaliana (Arabidopsis), Gossypium raimondii* (cotton), *Zea mays* (maize), *Brassica napus* (oilseed rape) and *Oryza sativa Japonica Group* (rice). In total, we found 100 126 introns that host functional elements in these five species. Specifically, we found that among all species, the number of introns with functional elements on the positive and negative strands is similar, with a length mostly smaller than 1500 bp, and the Adenine/Thymine (A/T) content is much higher than that of Guanine/Cytosine (G/C). In addition, the distribution of functional elements in introns varies among different species. All the above data can be downloaded for free in this database. This database provides a concise, comprehensive and user-friendly web interface, allowing users to easily retrieve target data based on their needs, using relevant organizational options. The database operation is simple and convenient, aiming to provide strong data support for researchers in related fields to study plant introns that host functional elements, including circular RNAs, lncRNAs, etc.

Database URL: http://deepbiology.cn/PlantIntronDB/

Introduction

Introns are non-coding nucleotide sequences that do not appear in mature mRNAs (1). Central dogma states that DNA encodes information about genetic material that is first transcribed into RNA and then translated into proteins (2). Transcription and translation constitute the entire process of gene expression, and RNA splicing is a crucial step in this process (3). During splicing, introns in pre-mRNA are removed and exons are joined in an orderly manner, which are finally processed into mature mRNA for translation, and then, introns are rapidly degraded (4, 5). Therefore, introns have always been considered as dispensable by-products and have been neglected for a long time (6).

However, in recent years, several studies have revealed biological significance of introns and found that introns play an indispensable regulatory role (7–9). For example, two studies in 2019 reported that in yeast cells, one or more strains lacking the removed introns had more difficulty surviving under nutrient-poor conditions than wild-type yeast, demonstrating the importance of introns under nutrient-poor conditions and suggesting that introns are necessary for the survival of cells in nutrient deprivation (10, 11). In addition, introns with functional elements have been found to play an important role in regulating biological processes in tropical *Xenopus oocytes* (12–14), human cell lines (15–19), Drosophila (20–23), *Arabidopsis* (24, 25) and other organisms (26–28). What's more, in our recent study, we designed Intron-capture RNA-seq to study introns that host functional elements in *Arabidopsis* (29). Despite the increasing attention focused on introns, it seems that a comprehensive database on introns that host functional elements has not yet emerged as far as the current studies are concerned. Plant Intron-Splicing Efficiency Database focuses only splicing of introns, which may not reveal the detailed landscape of functional elements in introns (30).

In this study, we developed a user-friendly database of plant introns containing functional elements: PlantIntronDB. This database currently contains a total of 100 126 introns with functional elements, including 19 363 in *Arabidopsis*, 2334 in cotton, 13 079 in maize, 31 370 in oilseed rape and 33 980 in rice, which can be downloaded for free. We found that the introns with functional elements in all species have similar numbers of positive and negative strands, with lengths mostly less than 1500 bp, and the content of A/T is much higher than that of G/C. In addition, we also found that the distribution of intron functional elements in different species

Received 26 May 2023; Revised 3 October 2023; Accepted 27 October 2023

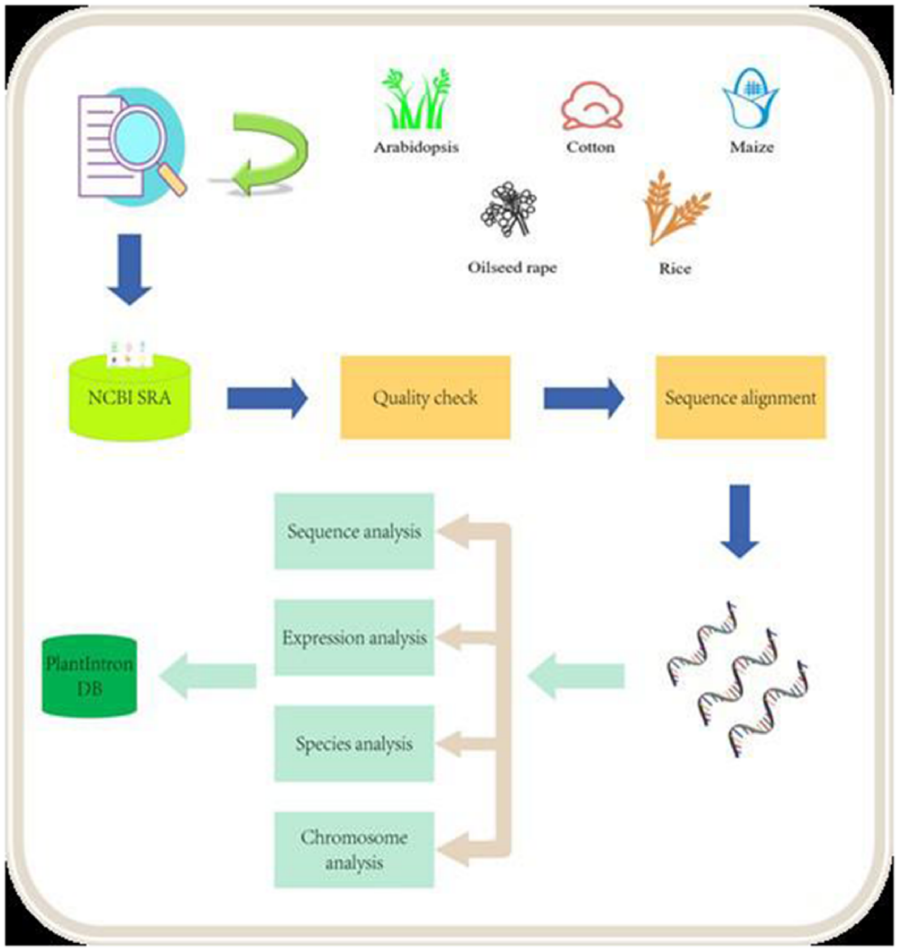

Figure 1. Data analysis pipeline. The raw data from NCB SRA were downloaded and processed after quality check. The alignment data were then analyzed for identifying functional introns. The final results were deposited into the PlantIntronDB.

has different characteristics. To the best of our knowledge, this is the first database currently available for the study of introns with functional components in plants. The database provides a clean and comprehensive, easy-to-use web interface so that users can easily search targets, using gene ID, chromosome, intron number, width, start, end, strand and species. The database is easy to operate and detailed information on all plant introns can be displayed. Placing these data on a unified browsing platform should allow high-throughput functional validation of identified plant introns and facilitate community studies on the regulatory roles of plant introns hosting functional elements.

Materials and methods

All plant data analysis pipeline

We selected five plant species, including *Arabidopsis*, cotton, maize, oilseed rape and rice, and searched the NCBI SRA database (https://www.ncbi.nlm.nih.gov/sra) for related RNA-seq samples on 8 July 2022. We used advanced search terms, including 'RNA-seq' and 'Illumina', and downloaded RNA-seq data for each species. Then, we used the SRAToolkit v3.0.2, Hisat2 v2.2.1 (31), and other software packages (details in 2.2) to process and analyze the data. We deleted files that did not generate valid sequence data during processing and then selected specific element features (geneID, intronNo, chr, strand, start, end, width, sampleNumber) to be entered into our database. Due to the fact that using more samples to process data can achieve better data purity, subsequent research will expand the sample size of the database to contain more species, which is also a suitable method to improve the reliability and comprehensiveness of the database. At the same time, the plant data we processed will be publicly released on our website. (http://deepbiology.cn/PlantIntronDB/).

Table 1. The number of functional introns in five species

Species	Introns hosting functional elements
Arabidopsis thaliana (Arabidopsis)	19 363
Gossypium raimondii (Cotton)	2334
Zea mays (Maize)	13 079
Brassica napus (Oilseed rape)	31 370
Oryza sativa Japonica Group (Rice)	33 980

Detection of introns hosting functional elements

In our study, we processed RNA-seq data and aligned the clean reads to the reference genomes of *Arabidopsis thaliana (Arabidopsis)* (TAIR10), *Gossypium raimondii*

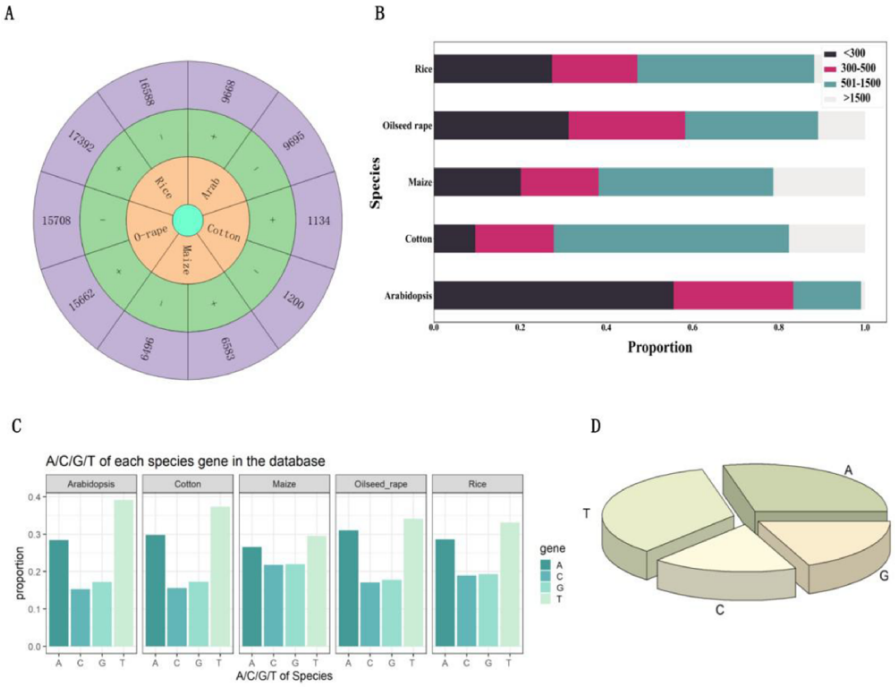

Figure 2. Analysis of introns that hosting functional elements. (A) Number of introns carrying functional elements on the positive and negative strands of *Arabidopsis*, cotton, maize, oilseed rape and rice. (B) Length of introns with functional elements in *Arabidopsis*, cotton, maize, oilseed rape and rice. (C) The proportion of each base in introns with functional elements of *Arabidopsis*, cotton, maize, oilseed rape and rice. (D) The proportion of bases in all introns with functional elements.

(cotton) (Graimondii2_0_v6), *Zea mays* (maize) (Zm-B73-REFERENCE-NAM-5.0), *Brassica napus* (oilseed rape) (AST_PRJEB5043_V1) and *Oryza sativa Japonica Group* (rice) (IRGSP-1.0) by using Hisat2. After alignment, we used Samtools v1.9 (32) to sort and index the bam files for subsequent analysis. Then, we used the Bioconductor packages [GenomicRanges v1.32.6 (33), GenomicAlignments v1.16.0 (33) and Biostrings v2.48.0 (34)] to analyze the bam files. Specifically, we compared the sequencing reads to the genome annotation by using the 'countOverlap', 'findOverlap' and 'subsetByOverlap' functions from the GenomicRanges and GenomicAlignments packages. We also used 'translateGTF' function from SplicingTypesAnno v1.0.2 (35) to input gene annotation files in GFF/GTF format to extract detailed features from introns, such as gene ID, start and end position, strand and intron number. We used the 'DNAStringSet' and 'substring' from the Biostrings package to extract nucleotide sequences. Finally, we used R packages [ggplot2 v3.3.5 (36) and lattice v0.20–38 (37)] for visualization.

To ensure the quality of the introns hosting functional elements, we followed our previous work (29) to process the reads. Specifically, when selecting reads, we followed the following criteria: (i) reads are inside introns, (ii) reads are strand-specific as introns and (iii) reads have no junctions. Besides, we removed all reads that overlapped both exons and introns. Then, all the reads within the introns were merged using the 'reduce' function. When identifying introns hosting functional elements, we selected candidates with the following criteria: (i) > 90% read coverage in the intron and (ii) found in at least seven mRNA-seq samples. We used the IGV (Integrative Genomics Viewer) v2.11.9 (38) to ensure the reliability of the selected data.

Database development

We developed a database to store plant intron data information, using PHP and MySQL. The data in this database not only include gene sequences but also essential information, such as species, gene ID, intron number, chromosome, strand, start and end position, width and sample number. Our data are limited to hundreds of RNA-seq samples from NCBI SRA for five species. However, as the amount of data increases, more than a single-layer MySQL database may be required to handle the data storage and retrieval requirements. Therefore, our database technology will shift towards distributed databases to better handle the challenges of storing massive amounts of data.

Results and discussion

Introns hosting functional elements for five plant species

In this study, we downloaded the data from NCBI SRA database and analyzed 358 samples of *Arabidopsis*, 59 samples of cotton, 185 samples of maize, 210 samples of oilseed rape and 141 samples of rice, respectively (details are available on the website). We used SRAToolkit, Hisat2 and other software packages to analyze the above RNA-seq samples. We removed files that did not generate valid sequence data during the analysis pipeline through data cleaning and only retained files that met the standards for analysis (Figure 1). In total, 100 126 introns that host functional elements were found in *Arabidopsis*, cotton, maize, oilseed rape, and rice, including 19 363 in *Arabidopsis*, 2334 in cotton, 13 079 in maize, 31 370 in oilseed rape and 33 980 in rice (Table 1).

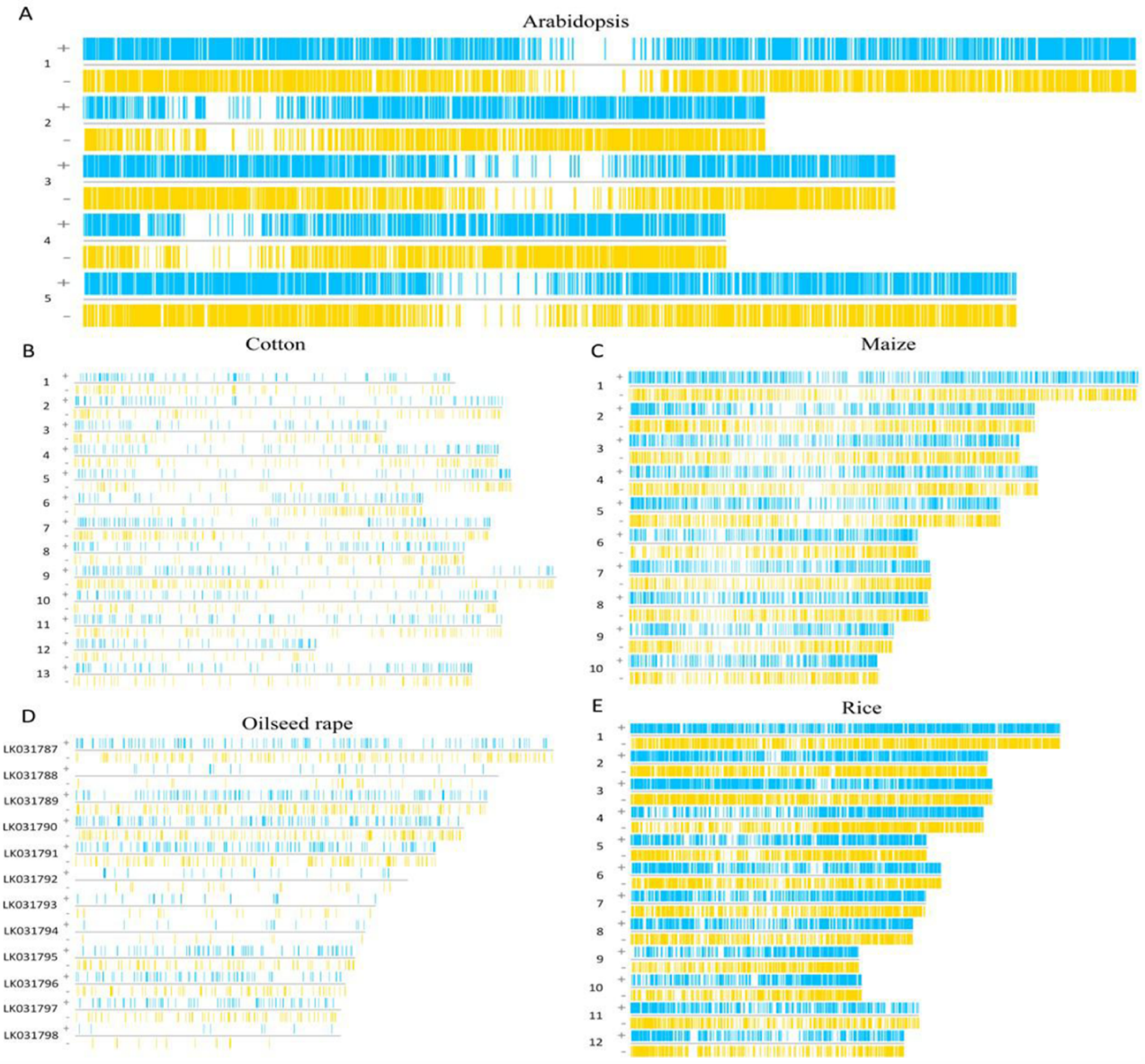

Figure 3. The region of functional elements in introns. (A) The functional elements of *Arabidopsis* introns are distributed on its chromosomes. (B) The functional elements of cotton introns are distributed on its chromosomes. (C) The functional elements of maize introns are distributed on its chromosomes. (D) The functional elements of oilseed rape introns are distributed on its chromosomes. (E) The functional elements of rice introns are distributed on its chromosomes.

Figure 2A provides the number of introns hosting functional elements on the positive and negative strands of each species, where '+' and '–' represent positive and negative strands, respectively. From the figure, we found that the number of introns hosting functional elements on both the positive and negative strands of five species is similar. In addition, this study also analyzed the length of introns with functional elements. The different sequence lengths in the five species are shown in Figure 2B. It can be seen that in all species, introns with functional elements >1500 bp in length are far less than those with lengths <1500 bp, indicating that most introns with functional elements have sequence lengths within 1500 bp. Figure 2C shows the content of each base in introns with functional elements. It can be seen from the figure that the content of A and T in introns with functional elements of various species is significantly higher than that of G and C. While integrating five species, it was discovered that the proportion of A and T is much greater than that of G and C (Figure 2D). These results will help future researchers to gain a deeper understanding of the biological characteristics and functions of introns hosting functional elements.

The functional elements in the introns of five species have roughly the same distribution on the positive and negative strands of their chromosomes. Specifically, functional elements in *Arabidopsis* introns are the most widely and densely distributed (Figure 3A) and shows similar representation in maize and rice (Figure 3C and E). In cotton and oilseed rape, the distribution of functional elements in introns is quite sparse (Figure 3B and D).

A

HOME USER GUIDE BROWSE SEARCH DOWNLOAD PUBLICATION TEAM

Introduction PlantIntronDB

B

Options	Example	Description
species	Arabidopsis	plant species
geneID	AT5G13140	the gene that hosting introns with functional elements
intronNO	intron_1	intron position
chr	5	chromosome number
strand	-	strand
start	4171302	intron start position
end	4171554	intron end position
width	252	intron width
sampleNumber	198	the number of sequencing samples from NCBI SRA

C

BROWSERS ALL

1 2 3 4 5 6 7 8 9 10 11 >> 1614

species	geneID	intronNO	chr	strand	start	end	width	sampleNumber
Arabidopsis	AT4G25820	intron_1	4	+	13130631	13130755	124	7
Arabidopsis	AT1G58060	intron_9	1	-	21491825	21492066	241	50
Arabidopsis	AT1G79720	intron_1	1	-	29998536	29998773	237	31
Arabidopsis	AT3G60370	intron_1	3	-	22315123	22315381	258	65
Arabidopsis	AT2G41710	intron_7	2	-	17402337	17402423	86	11
Arabidopsis	AT5G23390	intron_5	5	-	7872471	7872619	148	49
Arabidopsis	AT2G38330	intron_11	2	+	16066962	16067047	85	9
Arabidopsis	AT3G19800	intron_2	3	+	6876585	6876686	101	24
Arabidopsis	AT5G44790	intron_1	5	-	18076323	18076502	179	12
Arabidopsis	AT1G30490	intron_17	1	-	10800013	10800523	510	7
Arabidopsis	AT1G47900	intron_1	1	-	17647736	17648016	280	15
Arabidopsis	AT2G29530	intron_5	2	-	12641957	12642204	247	118

1 2 3 4 5 6 7 8 9 10 11 >> 1614

Arabidopsis Cotton Maize Oilseed rape Rice

D

SEARCHING CRITERIA

geneID
intronNo
chr
width
start end
species All
strand All
SEARCH

E

SEARCHING RESULT

species	geneID	intronNo	chr	strand	start	end	width	sampleNumber
Arabidopsis	AT5G13140	intron_1	5	-	4171302	4171554	252	198

F

1. Introns hosting functional elements in Arabidopsis [download]
   ID of the sequencing samples downloaded from NCBI SRA in Arabidopsis.[download]
2. Introns hosting functional elements in Cotton[download]
   ID of the sequencing samples downloaded from NCBI SRA in Cotton.[download]
3. Introns hosting functional elements in Maize.[download]
   ID of the sequencing samples downloaded from NCBI SRA in Maize.[download]
4. Introns hosting functional elements in Oilseed rape.[download]
   ID of the sequencing samples downloaded from NCBI SRA in Oilseed rape.[download]
5. Introns hosting functional elements in Rice.[download]
   ID of the sequencing samples downloaded from NCBI SRA in Rice.[download]

Figure 4. PlantIntronDB database. (A) PlantIntronDB navigation bar. (B) PlantIntronDB user guide page. (C) PlantIntronDB browsing page. (D) PlantIntronDB search page. (E) Search result of PlantIntronDB. (F) PlantIntronDB download page.

Database features

Currently, PlantIntronDB provides the following information: (i) gene ID. This ID integrates the higher-level ID to which the feature is subordinate, which helps to associate introns hosting functional elements with genes, and it enables users to compare and search different species of introns; (ii) intron-specific information for each species, including intron number, which refers to the position of the intron, chromosome, start position, end position, positive and negative strand, sequence length and sample number. The sample number is how many samples have been analyzed from all samples to contain introns with functional elements. At present, this database mainly supports the following functions: (i) intron information browsing: users can browse the specific attributes of all data in the database to understand its basic information, providing convenience for subsequent research; (ii) intron search: users can quickly locate an intron of interest by entering information such as gene ID, intron number, chromosome, start, end, width, strand and species; (iii) intron download: users can obtain complete information about all introns through the download function.

The web interface of PlantIntronDB consists of seven pages: Home, User Guide, Browse, Search, Download, Publication and Team (Figure 4A). Home introduces introns hosting functional elements and the database information. The User Guide page gives a guide to use the database and examples of searched introns hosting functional elements (Figure 4B). The Browse page provides detailed information about all data, i.e. species, gene ID, intron number, chromosome number, strand, start position, end position, sequence length and sample number. On the right side of the page, logos for five species are provided. Clicking on the logos will display detailed information on all introns hosting functional elements for the corresponding species (Figure 4C). The search portal is the main function of the website, providing eight attributes, and users can search for the target intron by entering and selecting the corresponding attribute of the search target (Figure 4D and E). The download module provides intron data and RNA-seq samples ID of five plants; users can download all data and sample IDs for free by clicking the corresponding option (Figure 4F). The Publication page shows our team's published papers related to this research, and the Team page gives a brief introduction of our team.

Conclusion

PlantIntronDB provides a complete and comprehensive platform for plant introns, including *Arabidopsis*, cotton, maize, oilseed rape and rice. It has detailed information about introns hosting functional elements, including gene ID, chromosome number, intron number, strand, start and end position and sample number. In addition, the platform also provides a search function to facilitate users to find the search target quickly. Finally, it also provides a free download function, which allows users to download by clicking the links of corresponding species. By using this database, users can not only grasp the rich details of plant introns hosting functional elements but also track these introns' resources and reveal their novel functions and regulatory roles. We believe that this database will help the community to have a deeper understanding of plant introns with functional elements.

Contribution statement

X.S. designed and supervised the project; X.S., W.W. and J.H. analyzed the data; W.W. and J.H. developed the database; H.L. and W.W. wrote the manuscript; X.S., W.W., J.H., J.Y. revised the manuscript.

Data availability

All related data are available at http://www.deepbiology.cn/PlantIntronDB/.

Funding

This work was supported by the National Natural Science Foundation of China (grant number 32070684, 31571306 to X.S.) and a Project of Shandong Province Higher Educational Program for Introduction and Cultivation of Young Innovative Talents in 2021. We thank Supercomputing Center in Shandong Agricultural University for technical support.

Conflict of interest

None declared.

References

1. Hesselberth,J. (2013) Lives that introns lead after splicing. *Wiley Interdiscip Rev. RNA*, **4**, 677–691.
2. Chang,H. and Qi,L. (2023) Reversing the central dogma: RNA-guided control of DNA in epigenetics and genome editing. *Mol. Cell*, **83**, 442–451.
3. Lee,Y. and Rio,D. (2015) Mechanisms and regulation of alternative pre-mRNA splicing. *Annu. Rev. Biochem.*, **84**, 291–323.
4. Neil,C. and Fairbrother,W. (2019) Intronic RNA: ad 'junk' mediator of post-transcriptional gene regulation. *Biochim. Biophys. Acta - Gene Regul. Mech.*, **1862**, 194439.
5. Nik,S. and Bowman,T. (2019) Splicing and neurodegeneration: insights and mechanisms. *Wiley Interdiscip Rev. RNA*, **10**, e1532.
6. Chan,S. and Pek,J. (2019) Stable intronic sequence RNAs (sisRNAs): an expanding universe. *Trends Biochem. Sci.*, **44**, 258–272.
7. Armakola,M., Higgins,M., Figley,M. *et al.* (2012) Inhibition of RNA lariat debranching enzyme suppresses TDP-43 toxicity in ALS disease models. *Nat. Genet.*, **44**, 1302–1309.
8. Laurent,G., Shtokalo,D., Tackett,M. *et al.* (2012) Intronic RNAs constitute the major fraction of the non-coding RNA in mammalian cells. *BMC Genom.*, **13**, 504.
9. Osman,I., Tay,M. and Pek,J. (2016) Stable intronic sequence RNAs (sisRNAs): a new layer of gene regulation. *Cell. Mol. Life Sci.*, **73**, 3507–3519.
10. Morgan,J., Fink,G. and Bartel,D. (2019) Excised linear introns regulate growth in yeast. *Nature*, **565**, 606–611.
11. Parenteau,J., Maignon,L., Berthoumieux,M. *et al.* (2019) Introns are mediators of cell response to starvation. *Nature*, **565**, 612–617.
12. Gardner,E., Nizami,Z., Talbot,C. *et al.* (2012) Stable intronic sequence RNA (sisRNA), a new class of noncoding RNA from the oocyte nucleus of Xenopus tropicalis. *Genes Dev.*, **26**, 2550–2559.
13. Talhouarne,G. and Gall,J. (2014) Lariat intronic RNAs in the cytoplasm of Xenopus tropicalis oocytes. *RNA*, **20**, 1476–1487.
14. Jin,J., He,X. and Silva,E. (2020) Stable intronic sequence RNAs (sisRNAs) are selected regions in introns with distinct properties. *BMC Genom.*, **21**, 287.
15. Guil,S., Soler,M., Portela,A. *et al.* (2012) Intronic RNAs mediate EZH2 regulation of epigenetic targets. *Nat. Struct. Mol. Biol.*, **19**, 664–670.

16. Talhouarne,G. and Gall,J. (2018) Lariat intronic RNAs in the cytoplasm of vertebrate cells. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, E7970–E7977.
17. Postnikova,O., Poliakov,E., Golestaneh,N. *et al.* (2019) Stable intronic sequences and exon skipping events in the human RPE65 gene: analysis of expression in retinal pigment epithelium cells and cell culture models. *Front. Genet.*, **10**, 634.
18. Pek,J. (2018) Stable intronic sequence RNAs engage in feedback loops. *Trends Genet.*, **34**, 330–332.
19. Zhang,Y., Zhang,X., Chen,T. *et al.* (2013) Circular intronic long noncoding RNAs. *Mol. Cell*, **51**, 792–806.
20. Pek,J., Osman,I., Tay,M. *et al.* (2015) Stable intronic sequence RNAs have possible regulatory roles in Drosophila melanogaster. *J. Cell Biol.*, **211**, 243–251.
21. Jia,N., Zheng,R., Osman,I. *et al.* (2018) Generation of Drosophila sisRNAs by independent transcription from cognate introns. *iScience*, **4**, 68–75.
22. Chan,S. and Pek,J. (2023) Distinct biogenesis pathways may have led to functional divergence of the human and Drosophila Arglu1 sisRNA. *EMBO Rep.*, **24**, e54350.
23. Osman,I. and Pek,J. (2018) A sisRNA/miRNA axis prevents loss of germline stem cells during starvation in Drosophila. *Stem Cell Rep.*, **11**, 4–12.
24. Li,Z., Wang,S., Cheng,J. *et al.* (2016) Intron lariat RNA inhibits microRNA biogenesis by sequestering the dicing complex in Arabidopsis. *PLoS Genet.*, **12**, e1006422.
25. Wu,H., Deng,S., Xu,H. *et al.* (2018) A noncoding RNA transcribed from the AGAMOUS (AG) second intron binds to CURLY LEAF and represses AG expression in leaves. *New Phytol.*, **219**, 1480–1491.
26. Moss,W. and Steitz,J. (2013) Genome-wide analyses of Epstein-Barr virus reveal conserved RNA structures and a novel stable intronic sequence RNA. *BMC Genom.*, **14**, 543.
27. Jiang,M., Zhang,S., Yang,Z. *et al.* (2018) Self-recognition of an inducible host lncRNA by RIG-I feedback restricts innate immune response. *Cell*, **173**, 906–919.
28. Tompkins,V., Valverde,D. and Moss,W. (2018) Human regulatory proteins associate with non-coding RNAs from the EBV IR1 region. *BMC Res. Notes*, **11**, 139.
29. Li,H., Zhang,Y., Bing,J. *et al.* (2023) Intron-capture RNA-seq reveals the landscape of intronic RNAs in Arabidopsis. *Plant Physiol. Biochem.*, **196**, 75–88.
30. Zhang,H., Jia,J. and Zhai,J. (2023) Plant Intron-Splicing Efficiency Database (PISE): exploring splicing of ∼1,650,000 introns in Arabidopsis, maize, rice, and soybean from ∼57,000 public RNA-seq libraries. *Sci. China Life Sci.*, **66**, 602–611.
31. Kim,D., Langmead,B. and Salzberg,S. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods*, **12**, 357–360.
32. Li,H., Handsaker,B., Wysoker,A. *et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
33. Lawrence,M., Huber,W., Pagès,H. *et al.* (2013) Software for computing and annotating genomic ranges. *PLoS Comput. Biol.*, **9**, e1003118.
34. Pagès,H., Aboyoun,P., Gentleman,R. *et al.* (2023) *Biostrings: Efficient Manipulation of Biological Strings*. R package version 2.48.0. https://bioconductor.org/packages/Biostrings (18 April 2023, date last accessed).
35. Sun,X., Zuo,F., Ru,Y. *et al.* (2015) SplicingTypesAnno: annotating and quantifying alternative splicing events for RNA-Seq data. *Comput. Methods Programs Biomed.*, **119**, 53–62.
36. Wickham,H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer, New York, ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org (18 April 2023, date last accessed).
37. Sarkar,D. (2008) *Lattice: Multivariate Data Visualization with R*. Springer, New York, ISBN 978-0-387-75968-5. http://lmdvr.r-forge.r-project.org (18 April 2023, date last accessed).
38. Thorvaldsdóttir,H., Robinson,J. and Mesirov,J. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinformatics*, **14**, 178–192.