Title

A Global Analysis of Alternative Splicing of Dichocarpum Medicinal Plants, Ranunculales.

Journal

Current Genomics, 23(3)

ISSN

1389-2029

Authors

Hao, Da-Cheng

Chen, Hao

Xiao, Pei-Gen

et al.

Publication Date

2022-07-05

DOI

10.2174/1389202923666220527112929

Peer reviewed

Current Genomics, 2022, *23*, 207-216

RESEARCH ARTICLE

A Global Analysis of Alternative Splicing of *Dichocarpum* Medicinal Plants, Ranunculales

Da-Cheng Hao[1,2,*,#], Hao Chen[3,4,#], Pei-Gen Xiao[5] and Tao Jiang[3,6,*]

[1]*Biotechnology Institute, School of Environment and Chemical Engineering, Dalian Jiaotong University, Dalian 116028, China;* [2]*Institute of Molecular Plant Sciences, University of Edinburgh, Edinburgh EH9 3BF, UK;* [3]*Department of Computer Science and Engineering, University of California, Riverside, CA, USA;* [4]*Department of Computational Biology, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA;* [5]*Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Beijing 100193, China;* [6]*Bioinformatics Division, BNRIST/Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China*

ARTICLE HISTORY

Received: March 28, 2022
Revised: April 19, 2022
Accepted: April 26, 2022

DOI:
10.2174/1389202923666220527112929

Abstract: ***Background*:** The multiple isoforms are often generated from a single gene *via* Alternative Splicing (AS) in plants, and the functional diversity of the plant genome is significantly increased. Despite well-studied gene functions, the specific functions of isoforms are little known, therefore, the accurate prediction of isoform functions is exceedingly wanted.

***Methods*:** Here we perform the first global analysis of AS of *Dichocarpum*, a medicinal genus of Ranunculales, by utilizing full-length transcriptome datasets of five Chinese endemic *Dichocarpum* taxa. Multiple software were used to identify AS events, the gene function was annotated based on seven databases, and the protein-coding sequence of each AS isoform was translated into an amino acid sequence. The self-developed software DIFFUSE was used to predict the functions of AS isoforms.

***Results*:** Among 8,485 genes with AS events, the genes with two isoforms were the most (6,038), followed by those with three isoforms and four isoforms. Retained intron (RI, 551) was predominant among 1,037 AS events, and alternative 3' splice sites and alternative 5' splice sites were second. The software DIFFUSE was effective in predicting functions of *Dichocarpum* isoforms, which have not been unearthed. When compared with the sequence alignment-based database annotations, DIFFUSE performed better in differentiating isoform functions. The DIFFUSE predictions on the terms GO:0003677 (DNA binding) and GO: 0010333 (terpene synthase activity) agreed with the biological features of transcript isoforms.

***Conclusion*:** Numerous AS events were for the first time identified from full-length transcriptome datasets of five *Dichocarpum* taxa, and functions of AS isoforms were successfully predicted by the self-developed software DIFFUSE. The global analysis of *Dichocarpum* AS events and predicting isoform functions can help understand the metabolic regulations of medicinal taxa and their pharmaceutical explorations.

Keywords: Alternative splicing, *Dichocarpum*, isoform function, DNA sequence, gene expression profile, deep learning.

1. INTRODUCTION

In the process of alternative splicing (AS; Fig. **1**), there are various combinations of splicing sites in a precursor of mRNA (pre-mRNA), and variably spliced mRNAs are generated, which could encode structurally/functionally variable proteins from a sole gene [1, 2]. The transcriptome diversity can be enhanced *via* AS in all eukaryotic organisms, which is indeed indispensable in botanical tissue identity and adapting to stress conditions [3, 4]. As critical machinery of transcription regulation [5, 6], AS might contribute to the dynamic reprogramming of the plant transcriptome during both beneficial and detrimental interactions [7]. One of the deterministic consequences is the adaptive shift of specialized metabolite biosynthesis. For example, 20,015 AS events of 6,324 *Panax notoginseng* genes were identified in PacBio transcriptome sequencing [8] (Table **S1**), many of which were associated with the ginsenoside biosynthesis. In the full-length transcriptome of *Artemisia annua* [9], 39 isoforms of 15 genes of the artemisinin biosynthesis pathway and upstream MVA and MEP pathways were identified. In *Cannabis sativa*, AS occurred in the cannabinoid biosynthetic genes LOX, AAE, PT, DXR, MCT and HDR [10]. The terpene synthase gene CsLIS/NES has two splicing forms, which influence the biosynthesis of nerolidol/linalool

*Address correspondence to these authors at the School of Environment and Chemical Engineering, Dalian Jiaotong University, Dalian 116028, China; Tel: 0086-411-84572552; E-mail: hao@djtu.edu.cn; and Department of Computer Science and Engineering, University of California, Riverside, CA, USA; Tel/Fax: 001-951-827-2991; E-mail: jiang@cs.ucr.edu

 # These authors contributed equally to this work.

1875-5488/22 $65.00+.00

in *Camellia sinensis* [11]. The AS events of TCS (caffeine synthase), CHS (chalcone synthase) and LAR (leucoanthocyanidin reductase) genes were unearthed *via* full-length sequencing of the transcriptome [12]. The AS transcripts of C4H (cinnamate-4-hydroxylase) 1, FLS (flavonol synthase) 1, PAL (phenylalanine ammonia-lyase) 2, CCR (cinnamoyl-CoA reductase) 2, UGT (UDP-glycosyltransferase) 75L12 and MYB113-1 might be crucial in controlling anthocyanin production of *C. sinensis* var. assamica [13]. The swift shifts of gene expression might be partially due to altered AS patterns [14], especially upon the environmental stimuli. However, little is known about the AS events and their putative association with biosynthesis in the majority of medicinal species-rich families, such as Ranunculaceae and other phylogenetically related Ranunculales families [15].

The second-generation high-throughput sequencing and assembly approaches are weak in the identification of AS events [16]. In third-generation sequencing, full-length transcripts are obtained with no short reads assembly, ensuring the accurate detection of AS isoforms [9]. In species with known whole genome sequences, *e.g*. *Brassica napus* [16], the number of novel isoforms identified *via* third-generation transcriptome sequencing, including novel isoforms from both novel and known genes, accounts for > 60% of all isoforms, illustrating the high effectiveness of full-length transcriptome sequencing.

The basal eudicot order Ranunculales are rich in terpenoids, alkaloids, flavonoids and other bioactive chemicals [17]; numerous Ranunculales species are traditionally used in ethnomedicine and folk medicine [18]. *Dichocarpum* is a Ranunculaceae genus named by Pei-Gen Xiao and Wen-Cai Wang [19]. Like many other Ranunculaceae genera, *Dichocarpum* might undergo multiple rounds of genome duplication and is abundant in polyploid species [20]. Numerous phytometabolites are identified in *Dichocarpum* taxa, many of which could be used clinically [21, 22]. The whole-genome sequencing has not been performed for most Ranunculaceae species, making the AS detection an awesome task. Recently, the full-length transcriptomes of five *Dichocarpum* taxa endemic in China, *i.e*., *Dichocarpum basilare*, *D. auriculatum*, *D. fargesii*, *D. malipoense* and *D. lobatipetalum*, were sequenced on the PacBio sequencing platform [22], generating 71,598 non-redundant full-length transcripts; many transcripts participate in immunity, defense response and stress adaptation, and those involved in the production of secondary metabolites are particularly intriguing. We extracted 27 orthologs extracted from transcriptome datasets, which were jointly used to construct the evolutionary tree; the tree topology was validated by the Pearson correlation between gene expression modes of *Dichocarpum* taxa and metabolome-based clustering analysis. In the transcriptome wide analysis of metabolite biosynthetic genes, the inferred evolutionary courses could partially explain the chemodiversity of specialized metabolites identified in *Dichocarpum* and phylogenetically close taxonomic groups; the biosynthetic pathways can be suggested based on the combined omics data. We suggest that the gene duplications are not rare in Ranunculales, and there might be a genus-specific Whole Genome Duplication (WGD) in *Dichocarpum*. There are widespread correlations between gene expressions and phytometabolite levels, and genes essential for the biosynthesis of secondary metabolites might undergo co-evolution. Under such a scenario, we hypothesize that the global analysis of *Dichocarpum* AS could be achieved *via* full-length transcriptome datasets and state-of-the-art software, which facilitate deeper studies of omics and utility of Ranunculales species. The functions of genes/ isoforms

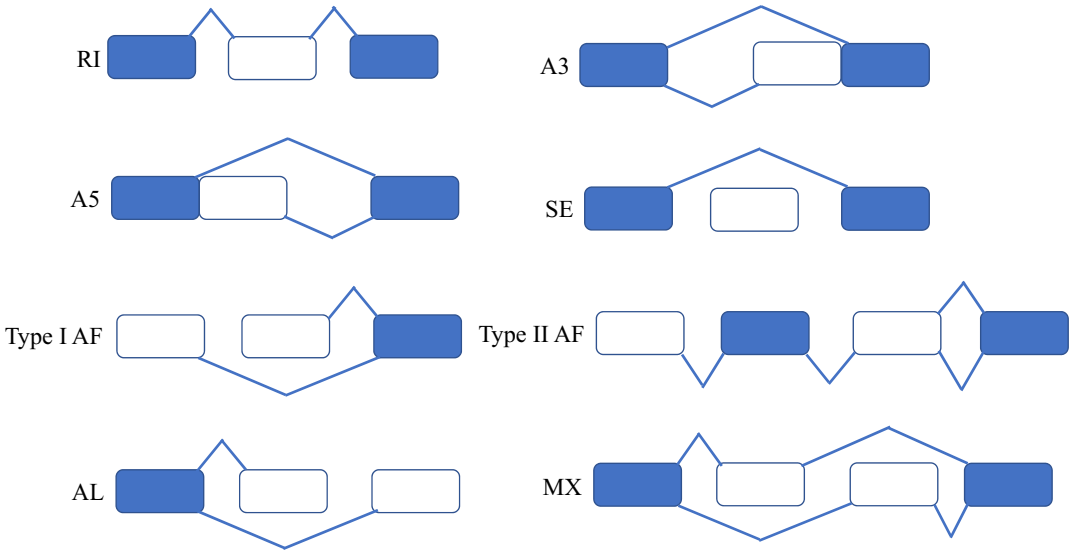

Fig. (1). Commonly found AS types in medicinal plants. Blue box, exon; white box, alternative region. In intron retention (RI), the intron is retained in the mature mRNA, instead of being spliced out as usual. In alternative 3' splice site (A3), the alternative 3' acceptor site (splice junction) is utilized, altering the 5' boundary of downstream exon. In alternative 5' splice site (A5), the alternative 5' donor site is used, changing the 3' boundary of upstream exon. In exon skipping (SE), the exon is spliced out of the primary transcript, which is the most common mode in mammalian (but not in plant) pre-mRNAs. AL, alternative last exon. Type I AF (alternative first exon): Alternative 1st exons mutually exclusive in different gene structures; Type II AF: The 1st exon is (part of) a downstream exon of other transcripts. In mutually exclusive exons (MX), one of two exons is retained in the mRNA after splicing, but not both. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

of *Dichocarpum* have not been fully probed so far. Although the gene functions could be analyzed by sequence alignment and based on annotations of homologs, it is challenging to resolve the isoform functions by this approach, given that the majority of their sequences are identical. Therefore, the self-developed software DIFFUSE [23] is used to predict isoform functions of *Dichocarpum*. DIFFUSE is a deep learning-based methods, which is able to predict isoform functions from gene-level functional annotations and isoform amino acid sequences and expression profiles. The predicted isoform functions help understand the metabolic control in *Dichocarpum*, as well as subsequent studies of naturally occurring botanical treasures.

2. MATERIALS AND METHODS

The GMAP (Genome Mapping and Alignment Program)-2017-06-20 was used to map error-free non-redundant *Dichocarpum* transcript sequences to UniTransModels (unique transcript model(s) reconstructed from each transcript family/cluster assembled based on PacBio full-length sequencing of five *Dichocarpum* taxa) [22]. We examined splicing junctions of transcripts mapped to the same UniTransModels, to collapse transcripts with identical splicing junctions, then transcripts with dissimilar splicing junctions were regarded as transcript isoforms of UniTransModels. SUPPA [24] (https://github.com/comprna/SUPPA) was used to identify AS events under default settings.

The gene function was annotated [22, 25] based on NT [26] (NCBI non-redundant nucleotide sequences), NR [26] (NCBI non-redundant protein sequences), KOG/COG [27] (Clusters of Orthologous Groups of proteins), Pfam [28] (Protein family), KO [29] (KEGG Ortholog database), Swiss-Prot [30] (manually annotated and reviewed protein sequence database) and GO [31] (Gene Ontology) databases. In NT database analysis, the software BLAST [32] and e-value '1e-10' were used; in NR, KOG, Swiss-Prot and KEGG analyses, the software Diamond BLASTX [33] and e-value '1e-10' were used; in Pfam analysis, the software Hmmscan [34] was used.

The ANGEL pipeline (https://github.com/PacificBiosciences/ANGEL), a long-read implementation of ANGLE [35], was used to determine protein-coding sequences (CDSs) of *Dichocarpum*. The assured protein sequences of one taxon or its closely related taxon were used in ANGEL training, and CDSs of given sequences were predicted by ANGEL. The software bowtie2 [36] was used to align sequencing reads to the reference transcriptome, and RSEM [37] was used to quantify expression levels of transcripts *via* the expected number of Fragments Per Kb of transcript sequence per Millions bp sequenced (FPKM).

The software DIFFUSE was used to analyze isoform functions of *Dichocarpum* taxa. Here we adopted the functional categories defined by GO as done in the original paper of DIFFUSE [23]. To predict GO terms associated with each isoform, DIFFUSE requires the input data of gene-level functional annotations, the amino acid sequences and conserved domains of each isoform, as well as networks of isoforms measuring their expression similarities. The GO annotations of each gene were obtained as mentioned above through sequence alignment. Using the same sequence alignment approach, we also collected the GO annotations for each isoform in order to make comparisons with the predictions of DIFFUSE. Generally, the AS isoforms from the same gene got identical annotations by sequence alignment. After aggregating all the GO terms annotated at the gene level, 1,900 GO terms were considered in the predictions of DIFFUSE, in which GO terms from MF (molecular function), BP (biological process) and CC (cellular component) branches were 821, 807 and 272 respectively.

The CDS of each AS isoform was translated into an amino acid sequence, and the NCBI Conserved Domain Database [38] was used to get conserved domains by searching. The isoform expression profiles were used to construct expression similarity networks. Due to the lack of biological replication in each species, correlations between expression profiles cannot be calculated, leading to the inability of constructing the co-expression network of isoforms. Therefore, the isoform networks were built from the Euclidean distance between the logarithm transformed expression levels of AS isoforms: $w(e_i, e_j) = \max(10 - |\log(1 + e_i) - \log(1 + e_j)|, 0)$, where e_i and e_j represent the expression values of isoforms i and j, and $w(e_i, e_j)$ represents the edge weight between two isoforms in the expression similarity network. Since each species had only one PacBio sequencing sample and biological replicates were not available, the correlation analysis was not feasible. The workflow of AS analysis and functional prediction is graphically represented in Fig. (**2**).

3. RESULTS

The AS was identified from 21,163 sequence isoforms (Table **S1**, Dataset **S1**), corresponding to 8,485 assembled UniTransModels (*i.e*. reconstructed genes) of five *Dichocarpum* taxa. All isoform sequences of *Dichocarpum* species used to identify AS events are shown in Dataset S1. The genes with two isoforms were the most (6,038), followed by those with three isoforms (1,566) and four isoforms (487) (Table **S2**); PB.2409_0_path1 had the most isoform (17). The length of genes with isoforms varied between 200 and 10,000 bp (Table **S3**); genes of 1,900 bp were the most (392), followed by 2,000 bp (387) and 2,300 bp (380). Totally 1,037 AS events were identified (Table **S4**), and retained intron (RI, 551; Fig. **1**) was the most; alternative 3' splice sites (A3, 253) was more than alternative 5' splice sites (A5, 160) and skipping exon (SE, 59). In other plants (Table **S1**), RI was also the most commonly found AS type. The number of alternative first exons (AF) and alternative last exons (AL) was 10 and four respectively, and there were no mutually exclusive exons (MX).

3.1. Isoform Functions Predicted by DIFFUSE

Totally 1,900 GO terms, including 821, 807 and 272 from MF, BP and CC branches, respectively, were considered in the predictions of DIFFUSE. We obtained 959,254 function predictions on 70,199 AS isoforms, 350,717 of which were *de novo* ones, meaning that we predicted the functions of isoforms for those genes that are unknown to have these putative functions.

3.2. Divergence of Isoform Functions

Given the expected functional divergence of AS isoforms [39], it might be enlightening to predict the func-

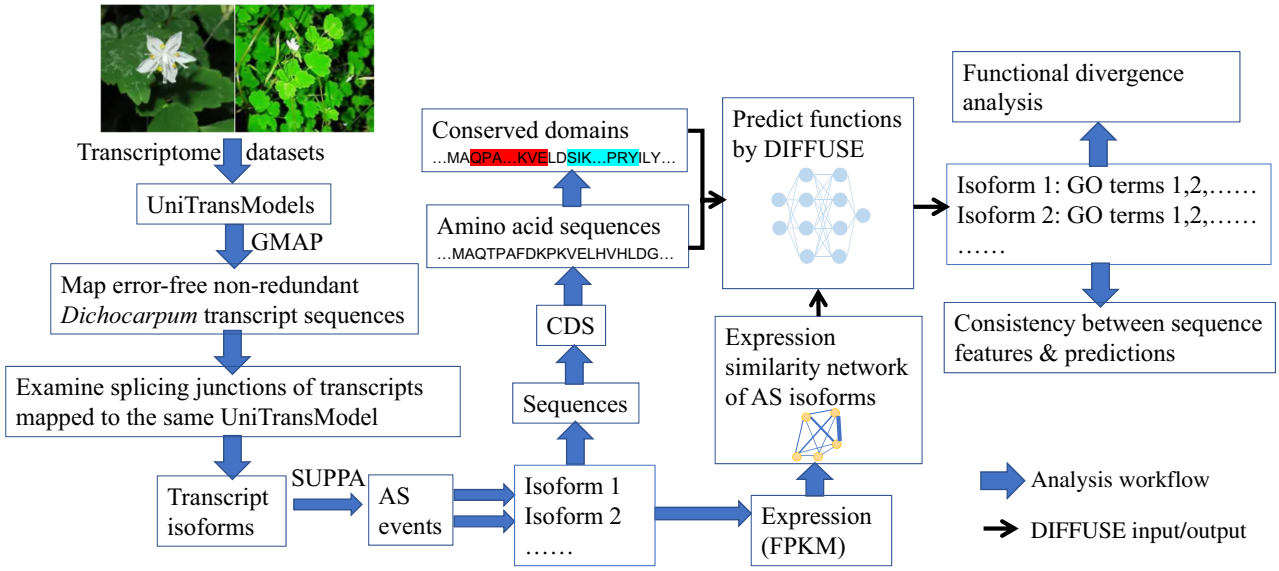

Fig. (2). The work flow of AS analysis and functional prediction of AS isoforms. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

tional divergence of isoforms [23, 40]. For each pair of isoforms of each multiple-isoform gene, we estimated their functional similarity according to the semantic similarity score obtained *via* GOssTo [41], and three GO branches were considered discretely. We defined isoform pairs with similarity scores 1.0 as "identical pairs", those with similarity scores 0.0 as "distinct pairs", and the others as "intermediate pairs". The analysis was performed for both the annotations obtained from sequence alignment and the functions predicted by DIFFUSE. The distributions of GO similarity scores of all considered isoform pairs (Fig. **3**) were obtained based on their sequence alignment annotations. Only 19.5%, 22.1% and 24.8% of isoform pairs had differences (intermediate pairs + distinct pairs) in their functional explanations in MF, BP and CC branches respectively. In contrast, in DIFFUSE predictions, 54.1%, 38.7% and 50.3% of isoform pairs had the functional differences in MF, BP and CC respectively (Fig. **4**). These suggest the better performance of DIFFUSE in discriminating isoform functions over the database annotations. The degree of functional divergence predicted by DIFFUSE might be more realistic when prior findings in the literature were taken into account [39].

3.3. Validation of Predicted Isoform Functions

A validation was conducted to examine the constancy between the presence of vital biological features in isoforms and DIFFUSE predictions. The functions concerning DNA binding and terpene synthase activity were analyzed here. Transcription factors (TFs) are essential in regulating specialized plant metabolism [25], which could be directly involved in producing medicinal compounds in plants. Therefore, we checked our predictions on the DNA binding function (GO:0003677) through examining domains of a set of isoforms from plant TF genes. Specifically, 207 isoforms were selected from multiple-isoform genes with two criteria: (*i*) The gene is annotated with the function of DNA binding (GO:0003677); (*ii*) The gene contains at least one DNA binding domain of plant TFs. We then checked the consistency between predictions on GO:0003677 of selected isoforms and the presence of DNA binding domains in their sequences. Isoforms annotated with DNA binding domains could be more possible to bind DNA. The Jaccard index between the presence of DNA binding domains and our predictions is 0.872 (Table **1**), suggesting high consistency. Specific predictions on GO:0003677 and domains of isoforms are shown in Table **S5**. Besides, we did the predictions on terpene synthase activity (GO: 0010333) and terpene synthase domain, and 27 AS isoforms were selected with the same criteria. The Jaccard index between having the terpene synthase activity or not and the existence of terpene synthase domain is 0.750 (Table **2**), which is lower than that of DNA binding, but still plausible; the term GO: 0010333 is more specific than GO:0003677, thus it can be harder to predict with fewer training samples. Details of functional predictions and domains for this case are shown in Table **S6**. In addition, predictions of DIFFUSE on five representative specialized metabolism-related GO terms are listed in Table **S7**.

4. DISCUSSION

In the present study, for the first-time numerous AS events are identified from full-length transcripts of *Dichocarpum*, a representative Ranunculales medicinal genus, which significantly upsurges the diversity of Ranunculaceae proteome and metabolome [21, 22], sometimes possibly leading to completely different proteins/metabolites. In *Arabidopsis*, around 61% of multi-exon genes are alternatively spliced under normal growth conditions [5] (Table **S1**); in cucumber, 58% of multi-exon genes undergo AS [42]; in polyploid cotton, divergent AS isoforms were generated from 51.4% of homoeologous genes in each subgenome [43]. However, it was challenging to identify most AS events based on the second generation high-throughput sequencing, *e.g.* Illumina-based one, when the whole genome sequences were not available [25, 44]; in the Illumina sequencing, full-length transcripts have to be assembled for

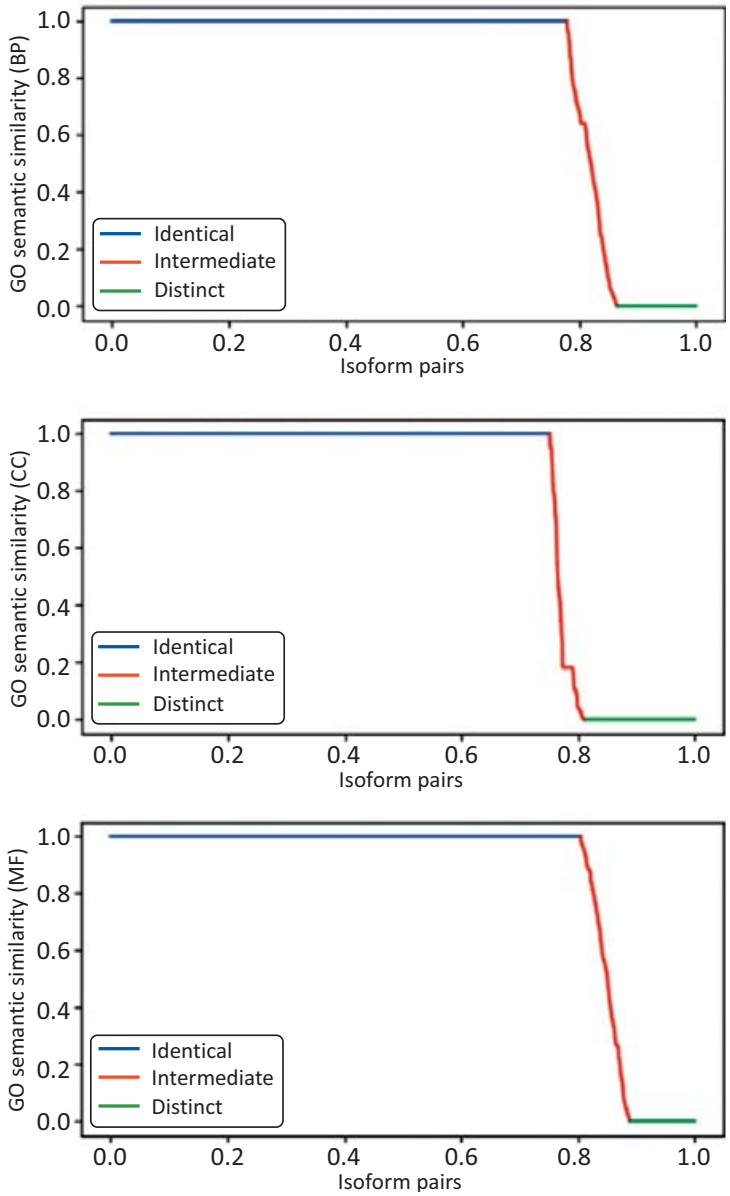

Fig. (3). Spreading of GO similarity scores of isoform pairs inferred from database annotations on MF, BP and CC. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

the AS identification, as only short sequencing reads are generated, and the PCR amplification bias cannot be eliminated during sequencing library construction. The Pacific Biosciences (PacBio) full-length sequencing enables the global mining of isoforms independent of whole-genome sequences [22]. When compared with Illumina short read sequencing, the PacBio isoform sequencing (Iso-Seq) is more efficient in precisely finding AS isoforms in polyploid species, where the sequence similarity between coexisting subgenomes is prevalent [16, 43]. In the integrated framework of response to Jasmonic Acid (JA), AS represents an essential regulatory layer between TFs and protein abundance/modification changes [45], among others. AS can be elicited by environmental cues rapidly, and then participates in transcriptome reprogramming, thus precisely modifying functional responses [46]. For example, a negative feedback loop was created by a repressor JAZ10 isoform to make cells insensitive to the JA cue [47], which could influence the biosynthesis of numerous specialized metabolites (*e.g.* flavonoids and terpenes) [25].

In polyploid species, total mRNA levels may not alter, but splicing isoform ratios of some genes may change substantially under the biological crosstalk [7, 16]. The AS type ratios vary in different plants. In *Dichocarpum* and *Arabidopsis*, the shares of A3 and A5 were more than that of SE [5] (Table **S1**), analogous to other dicots such as soybean [48], *Panax notoginseng* [8] and *Camellia sinensis var. assamica* purple cultivar [13], among others; the same is true for monocots rice [4] and moso bamboo [49]. SE was more common than A3 and A5 in *Panax ginseng* [50] and *Camellia sinensis* [51] (Table **S1**). AF and AL were identified unambiguously only in the minority species such as *C. sinensis var. assamica* purple cultivar and *Dichocarpum*, which may have unique evolutionary trajectory. MX was identified in *C. sinensis*, *P. ginseng* and rice, rather than in *Dichocarpum*. The differences between proportions of each AS type in

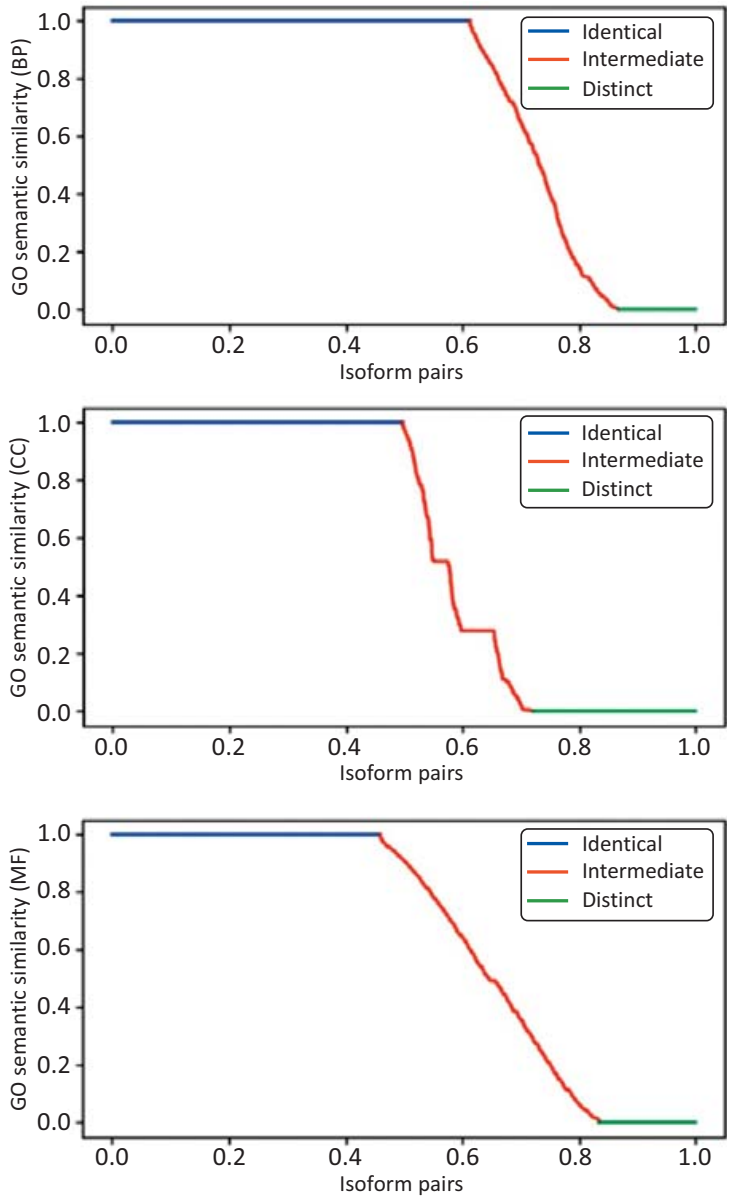

Fig. (4). Spreading of GO similarity scores of paired isoforms inferred from DIFFUSE predictions on MF, BP and CC. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

Table 1. Constancy between the existence/lack of DNA binding domains and functional predictions of GO:0003677 (DNA binding) of 207 selected isoforms.

-	With DNA Binding Domains	No. DNA Binding Domains
With the function	163	10
Without the function	14	20
Jaccard index	0.872	-

different taxa show that the approaches of intron processing and/or AS could have diverged extensively under various selection pressure.

As the whole genome sequencing of *Dichocarpum* has not been performed, much less AS events were identified in *Dichocarpum* when compared with other species. In addition, the AS frequency arguably decreases after gene /genome duplication [1]; the AS disparity of duplicated genes was influenced by age [52]: A positive correlation was identified in anciently duplicated genes, but in recently du-

Table 2. Constancy between the existence/lack of terpene synthase domain and functional predictions of GO: 0010333 (terpene synthase activity) of 27 selected isoforms.

-	With Terpene Synthase Domains	No. Terpene Synthase Domains
With the function	15	3
Without the function	2	7
Jaccard index	0.750	-

plicated ones it was a negative correlation. There were multiple rounds of WGD during the evolution of Ranunculales [53]; in Miocene, the *Dichocarpum* genome might underwent WGD and/or segmental duplication [22], meanwhile the evergreen broadleaved forests rapidly rose in subtropical regions of East Asia. Interestingly, the Ranunculales families Berberidaceae/Lardizabalaceae had overlapping ecological distributions with *Dichocarpum*, and they all diversified biogeographically. The gene families involved in specialized metabolite biosynthesis and transport had rapid and recent expansion in *Dichocarpum* and possibly phylogenetically close genera. Some gene duplications necessary for subsequent neofunctionalization/subfunctionalization of essential genes involved in the production of flavonoid and anthocyanin occurred exclusively in the Ranunculaceae lineage, which, along with AS events, makes this basal eudicot family a bonanza of ecological/pharmaceutical resources.

In a full-length transcriptome of *C. sinensis* [51], 28,980 AS events were identified; RI is the most common form of AS, which is similar to *Dichocarpum*, polyploid cotton [43], cucumber [42], *Arabidopsis* [5] and maize [3]. However, AF and AL were not included in this AS analysis of *C. sinensis*. RI events of chalcone isomerase and dihydroflavonol 4-reductase genes were found from the genomic and transcriptome data of *C. sinensis* [54]. However, it is more difficult to identify the AS events in *Dichocarpum* without available whole genome sequences. DIFFUSE is effective in predicting isoform functions of *Dichocarpum* taxa; 207 isoforms were identified from genes with DNA binding domain (Table **S5**), which belong to TF families such as WRKY, bZIP, Myb, PHD, SBP, tify, zf-C2H2 and B3. Interestingly, MYB113-1 could be a crucial AS transcript for regulating anthocyanin biosynthesis in *C. sinensis var. assamica* [13]. Seventeen *Salvia miltiorrhiza* bZIP genes underwent AS events [55], which could regulate the flux through MVA- and MEP-dependent isoprenoid biosynthetic pathways [56]. Therefore, it worths studying whether the AS isoform with the above DNA binding domain is important in the biosynthesis of any *Dichocarpum* metabolites, *e.g*. anthocyanin and triterpenoid [22].

CONCLUSION

DIFFUSE performs better than the annotations based on sequence alignment in distinguishing isoform functions. In case studies, the predictions on terms GO:0003677 (DNA binding) and GO: 0010333 (terpene synthase activity) agree with the biological characteristics of isoforms. As no proper biological features can be selected for some GO terms such as terpenoid metabolic process GO:0006721, secondary metabolic process GO:0019748 (Table **S7**), their DIFFUSE predictions have to be validated by alternative evidence, *e.g*., characterization of recombinant enzymes. The investigation of posttranscriptional regulation is necessary to gain more information about the complexity of various metabolic/signaling pathways. Predicting AS isoform functions can facilitate understanding the metabolic control of medicinal taxa and help the related pharmaceutical studies. We provide novel visions of the complexity of AS and isoform function and will boost our understanding of AS in polyploid species. Our methods of Iso-Seq data scrutiny are a useful reference for dissecting AS of other species. The alterations in AS numbers and transcript expression could contribute to swift alterations in gene expression and metabolite content under biotic/abiotic stress [6, 57-61], indicating that AS events play a vital regulatory role against environmental cues in plants. However, the exact link between AS isoforms and medicinal phytometabolites has not been established, which warrants further explorations.

LIST OF ABBREVIATIONS

A3	=	Alternative 3' splice site
A5	=	Alternative 5' splice site
AF	=	Alternative First exon
AL	=	Alternative Last exon
AS	=	Alternative Splicing
BIA	=	Benzylisoquinoline Alkaloid
DIFFUSE	=	Deep learning- based prediction of Iso-Form Functions from Sequences and Expression
JA	=	Jasmonic Acid
MX	=	Mutually Exclusive Exon
RI	=	Retained Intron
SE	=	Skipping Exon
WGD	=	Whole Genome Duplication

AUTHOR CONTRIBUTIONS

D.C.H. and P.G.X. designed the research; D.C.H. and H.C. performed the experiments and informatics analysis; P.G.X. and T.J. supervised the project; D.C.H., P.G.X. and T.J. provided resources and financial support; D.C.H. and H.C. wrote the manuscript. All authors read and approved of its content.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The data that support the findings of this study are available within the article.

FUNDING

This work is supported by the Scientific Research Funds Project of the Liaoning Education Department (JDL 2019012), China Scholarship Council (202108210156), and the US National Institute of Health (grant 1R01NS125018) and the National Key Research and Development Program of China (grant 2018YFC0910404).

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's website along with the published article.

REFERENCES

[1] Zhang, P.G.; Huang, S.Z.; Pin, A.L.; Adams, K.L. Extensive divergence in alternative splicing patterns after gene and genome duplication during the evolutionary history of Arabidopsis. *Mol. Biol. Evol.,* **2010**, *27*(7), 1686-1697. http://dx.doi.org/10.1093/molbev/msq054 PMID: 20185454

[2] Wei, X.P.; Li, H.J.; Che, P.; Guo, H.J.; Zhang, B.G.; Liu, H.T.; Qi, Y.D. Comparing chloroplast genomes of traditional Chinese herbs Schisandra sphenanthera and S. chinensis. *Chin. Herb. Med.,* **2020**, *12*(3), 247-256. http://dx.doi.org/10.1016/j.chmed.2019.09.009

[3] Thatcher, S.R.; Zhou, W.; Leonard, A.; Wang, B.B.; Beatty, M.; Zastrow-Hayes, G.; Zhao, X.; Baumgarten, A.; Li, B. Genome-wide analysis of alternative splicing in Zea mays: Landscape and genetic regulation. *Plant Cell,* **2014**, *26*(9), 3472-3487. http://dx.doi.org/10.1105/tpc.114.130773 PMID: 25248552

[4] Dong, C.; He, F.; Berkowitz, O.; Liu, J.; Cao, P.; Tang, M.; Shi, H.; Wang, W.; Li, Q.; Shen, Z.; Whelan, J.; Zheng, L. Alternative splicing plays a critical role in maintaining mineral nutrient homeostasis in rice (*Oryza sativa*). *Plant Cell,* **2018**, *30*(10), 2267-2285. http://dx.doi.org/10.1105/tpc.18.00051 PMID: 30254029

[5] Marquez, Y.; Brown, J.W.; Simpson, C.; Barta, A.; Kalyna, M. Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. *Genome Res.,* **2012**, *22*(6), 1184-1195. http://dx.doi.org/10.1101/gr.134106.111 PMID: 22391557

[6] Zheng, Y.; Luo, L.; Chen, Q.; Yang, D.; Gong, Y.; Yang, Y.; Qin, X.; Wang, Y.; Kong, X.; Yang, Y. Cold response transcriptome analysis of the alternative splicing events induced by the cold stress in D. catenatum. *Int. J. Mol. Sci.,* **2022**, *23*(2), 981. http://dx.doi.org/10.3390/ijms23020981 PMID: 35055168

[7] Rigo, R.; Bazin, J.R.M.; Crespi, M.; Charon, C.L. Alternative splicing in the regulation of plant-microbe interactions. *Plant Cell Physiol.,* **2019**, *60*(9), 1906-1916. http://dx.doi.org/10.1093/pcp/pcz086 PMID: 31106828

[8] Zhang, D.; Li, W.; Chen, Z.J.; Wei, F.G.; Liu, Y.L.; Gao, L.Z. SMRT- and Illumina-based RNA-seq analyses unveil the ginsinoside biosynthesis and transcriptomic complexity in Panax notoginseng. *Sci. Rep.,* **2020**, *10*(1), 15310. http://dx.doi.org/10.1038/s41598-020-72291-1 PMID: 32943706

[9] Ma, T.; Gao, H.; Zhang, D.; Sun, W.; Yin, Q.; Wu, L.; Zhang, T.; Xu, Z.; Wei, J.; Su, Y.; Shi, Y.; Ding, D.; Yuan, L.; Dong, G.; Leng, L.; Xiang, L.; Chen, S. Genome-wide analysis of light-regulated alternative splicing in *Artemisia annua* L. *Front. Plant Sci.,* **2021**, *12*, 733505. http://dx.doi.org/10.3389/fpls.2021.733505 PMID: 34659300

[10] Wu, B.; Li, Y.; Li, J.; Xie, Z.; Luan, M.; Gao, C.; Shi, Y.; Chen, S. Genome-wide analysis of alternative splicing and non-coding RNAs reveal complicated transcriptional regulation in Cannabis sativa L. *Int. J. Mol. Sci.,* **2021**, *22*(21), 11989. http://dx.doi.org/10.3390/ijms222111989 PMID: 34769433

[11] Liu, G.F.; Liu, J.J.; He, Z.R.; Wang, F.M.; Yang, H.; Yan, Y.F.; Gao, M.J.; Gruber, M.Y.; Wan, X.C.; Wei, S. Implementation of CsLIS/NES in linalool biosynthesis involves transcript splicing regulation in *Camellia sinensis*. *Plant Cell Environ.,* **2018**, *41*(1), 176-186. http://dx.doi.org/10.1111/pce.13080 PMID: 28963730

[12] Xu, Q.; Zhu, J.; Zhao, S.; Hou, Y.; Li, F.; Tai, Y.; Wan, X.; Wei, C. Transcriptome profling using single-molecule direct RNA sequencing approach for in-depth understanding of genes in secondary metabolism pathways of *Camellia sinensis*. *Front. Plant Sci.,* **2017**, *8*, 1205. http://dx.doi.org/10.3389/fpls.2017.01205 PMID: 28744294

[13] Chen, L.; Shi, X.; Nian, B.; Duan, S.; Jiang, B.; Wang, X.; Lv, C.; Zhang, G.; Ma, Y.; Zhao, M. Alternative splicing regulation of anthocyanin biosynthesis in *Camellia sinensis* var. assamica unveiled by PacBio Iso-Seq. *G3 (Bethesda),* **2020**, *10*(8), 2713-2723. http://dx.doi.org/10.1534/g3.120.401451 PMID: 32518082

[14] Li, Y.; Dai, C.; Hu, C.; Liu, Z.; Kang, C. Global identification of alternative splicing *via* comparative analysis of SMRT- and Illumina-based RNA-seq in strawberry. *Plant J.,* **2017**, *90*(1), 164-176. http://dx.doi.org/10.1111/tpj.13462 PMID: 27997733

[15] Hao, C.; Yang, L. Drug metabolism and disposition diversity of Ranunculales phytometabolites: A systems perspective. *Expert Opin. Drug Metab. Toxicol.,* **2016**, *12*(9), 1047-1065. http://dx.doi.org/10.1080/17425255.2016.1201068 PMID: 27295138

[16] Li, M.; Hu, M.; Xiao, Y.; Wu, X.; Wang, J. The activation of gene expression and alternative splicing in the formation and evolution of allopolyploid *Brassica napus*. *Hortic. Res.,* **2022**, [Epub ahead of print]. http://dx.doi.org/10.1093/hr/uhab075

[17] Hao, D.C.; He, C.N.; Shen, J.; Xiao, P.G. Anticancer chemodiversity of Ranunculaceae medicinal plants: Molecular mechanisms and functions. *Curr. Genomics,* **2017**, *18*(1), 39-59. http://dx.doi.org/10.2174/1389202917666160803151752 PMID: 28503089

[18] Hao, D.C.; Zhang, Y.; He, C.N.; Xiao, P.G. Distribution of therapeutic efficacy of Ranunculales plants used by ethnic minorities on the phylogenetic tree of Chinese species. *Evid. Based Complement. Alternat. Med.,* **2022**, *2022*, 9027727. http://dx.doi.org/10.1155/2022/9027727 PMID: 35069772

[19] Xiao, P.G.; Wang, W.C. A new genus of Ranunculaceae—Dichocarpum, W. T. Wang et Hsiao. *Acta Phytotaxon. Sin.,* **1964**, *9*, 315-333.

[20] Xiang, K.L.; Zhao, L.; Erst, A.S.; Yu, S.X.; Jabbour, F.; Wang, W. A molecular phylogeny of Dichocarpum (Ranunculaceae): Implications for eastern Asian biogeography. *Mol. Phylogenet. Evol.,* **2017**, *107*, 594-604. http://dx.doi.org/10.1016/j.ympev.2016.12.026 PMID: 28025001

[21] Li, P.; Shen, J.; Li, Y.; Yao, H.; Yu, M.; He, C.; Xiao, P. Metabolite profiling based on UPLC-Q-TOF-MS/MS and the biological evaluation of medicinal plants of Chinese Dichocarpum (Ranunculaceae). *Chem. Biodivers.,* **2021**, *18*(10), e2100432. http://dx.doi.org/10.1002/cbdv.202100432 PMID: 34351062

[22] Hao, D.C.; Li, P.; Xiao, P.G.; He, C.N. Dissection of full-length transcriptome and metabolome of *Dichocarpum* (Ranunculaceae): Implications in evolution of specialized metabolism of Ranunculales medicinal plants. *PeerJ,* **2021**, *9*, e12428.

http://dx.doi.org/10.7717/peerj.12428 PMID: 34760397

[23] Chen, H.; Shaw, D.; Zeng, J.; Bu, D.; Jiang, T. DIFFUSE: Predicting isoform functions from sequences and expression profiles *via* deep learning. *Bioinformatics,* **2019**, *35*(14), i284-i294. http://dx.doi.org/10.1093/bioinformatics/btz367 PMID: 31510699

[24] Trincado, J.L.; Entizne, J.C.; Hysenaj, G.; Singh, B.; Skalic, M.; Elliott, D.J.; Eyras, E. SUPPA2: Fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome Biol.,* **2018**, *19*(1), 40. http://dx.doi.org/10.1186/s13059-018-1417-1 PMID: 29571299

[25] Hao, C.; Chen, S.L.; Osbourn, A.; Kontogianni, V.G.; Liu, L.W.; Jordán, M.J. Temporal transcriptome changes induced by methyl jasmonate in *Salvia sclarea*. *Gene,* **2015**, *558*(1), 41-53. http://dx.doi.org/10.1016/j.gene.2014.12.043 PMID: 25536164

[26] O'Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufo, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; Astashyn, A.; Badretdin, A.; Bao, Y.; Blinkova, O.; Brover, V.; Chetvernin, V.; Choi, J.; Cox, E.; Ermolaeva, O.; Farrell, C.M.; Goldfarb, T.; Gupta, T.; Haft, D.; Hatcher, E.; Hlavina, W.; Joardar, V.S.; Kodali, V.K.; Li, W.; Maglott, D.; Masterson, P.; McGarvey, K.M.; Murphy, M.R.; O'Neill, K.; Pujar, S.; Rangwala, S.H.; Rausch, D.; Riddick, L.D.; Schoch, C.; Shkeda, A.; Storz, S.S.; Sun, H.; Thibaud-Nissen, F.; Tolstoy, I.; Tully, R.E.; Vatsan, A.R.; Wallin, C.; Webb, D.; Wu, W.; Landrum, M.J.; Kimchi, A.; Tatusova, T.; DiCuccio, M.; Kitts, P.; Murphy, T.D.; Pruitt, K.D. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.,* **2016**, *44*(D1), D733-D745. http://dx.doi.org/10.1093/nar/gkv1189 PMID: 26553804

[27] Tatusov, R.L.; Natale, D.A.; Garkavtsev, I.V.; Tatusova, T.A.; Shankavaram, U.T.; Rao, B.S.; Kiryutin, B.; Galperin, M.Y.; Fedorova, N.D.; Koonin, E.V. The COG database: New developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.,* **2001**, *29*(1), 22-28. http://dx.doi.org/10.1093/nar/29.1.22 PMID: 11125040

[28] El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S.R.; Luciani, A.; Potter, S.C.; Qureshi, M.; Richardson, L.J.; Salazar, G.A.; Smart, A.; Sonnhammer, E.L.L.; Hirsh, L.; Paladin, L.; Piovesan, D.; Tosatto, S.C.E.; Finn, R.D. The Pfam protein families database in 2019. *Nucleic Acids Res.,* **2019**, *47*(D1), D427-D432. http://dx.doi.org/10.1093/nar/gky995 PMID: 30357350

[29] Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.,* **2016**, *44*(D1), D457-D462. http://dx.doi.org/10.1093/nar/gkv1070 PMID: 26476454

[30] Boutet, E.; Lieberherr, D.; Tognolli, M.; Schneider, M.; Bairoch, A. UniProtKB/Swiss-Prot. *Methods Mol. Biol.,* **2007**, *406*, 89-112. PMID: 18287689

[31] Gene Ontology Consortium: going forward. *Nucleic Acids Res.,* **2015**, *43*, D1049-D1056. PMID: 25428369

[32] Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic Acids Res.,* **2008**, *36*, W5-9. http://dx.doi.org/10.1093/nar/gkn201

[33] Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods,* **2015**, *12*(1), 59-60. http://dx.doi.org/10.1038/nmeth.3176 PMID: 25402007

[34] Eddy, S.R. Accelerated profile HMM searches. *PLOS Comput. Biol.,* **2011**, *7*(10), e1002195. http://dx.doi.org/10.1371/journal.pcbi.1002195 PMID: 22039361

[35] Shimizu, K.; Adachi, J.; Muraoka, Y. ANGLE: A sequencing errors resistant program for predicting protein coding regions in unfinished cDNA. *J. Bioinform. Comput. Biol.,* **2006**, *4*(3), 649-664. http://dx.doi.org/10.1142/S0219720006002260 PMID: 16960968

[36] Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods,* **2012**, *9*(4), 357-359. http://dx.doi.org/10.1038/nmeth.1923 PMID: 22388286

[37] Li, B.; Dewey, C.N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.,* **2011**, *12*(1), 323. http://dx.doi.org/10.1186/1471-2105-12-323 PMID: 21816040

[38] Marchler-Bauer, A.; Derbyshire, M.K.; Gonzales, N.R.; Lu, S.; Chitsaz, F.; Geer, L.Y.; Geer, R.C.; He, J.; Gwadz, M.; Hurwitz, D.I.; Lanczycki, C.J.; Lu, F.; Marchler, G.H.; Song, J.S.; Thanki, N.; Wang, Z.; Yamashita, R.A.; Zhang, D.; Zheng, C.; Bryant, S.H. CDD: NCBI's conserved domain database. *Nucleic Acids Res.,* **2015**, *43*, D222-D226. http://dx.doi.org/10.1093/nar/gku1221 PMID: 25414356

[39] Li, W.; Kang, S.; Liu, C.C.; Zhang, S.; Shi, Y.; Liu, Y.; Zhou, X.J. High-resolution functional annotation of human transcriptome: Predicting isoform functions by a novel multiple instance-based label propagation method. *Nucleic Acids Res.,* **2014**, *42*(6), e39. http://dx.doi.org/10.1093/nar/gkt1362 PMID: 24369432

[40] Shaw, D.; Chen, H.; Jiang, T. DeepIsoFun: A deep domain adaptation approach to predict isoform functions. *Bioinformatics,* **2019**, *35*(15), 2535-2544. http://dx.doi.org/10.1093/bioinformatics/bty1017 PMID: 30535380

[41] Caniza, H.; Romero, A.E.; Heron, S.; Yang, H.; Devoto, A.; Frasca, M.; Mesiti, M.; Valentini, G.; Paccanaro, A. GOssTo: A stand-alone application and a web tool for calculating semantic similarities on the Gene Ontology. *Bioinformatics,* **2014**, *30*(15), 2235-2236.

[42] Sun, Y.; Hou, H.; Song, H.; Lin, K.; Zhang, Z.; Hu, J.; Pang, E. The comparison of alternative splicing among the multiple tissues in cucumber. *BMC Plant Biol.,* **2018**, *18*, 5.

[43] Wang, M.; Wang, P.; Liang, F.; Ye, Z.; Li, J.; Shen, C.; Pei, L.; Wang, F.; Hu, J.; Tu, L.; Lindsey, K.; He, D.; Zhang, X. A global survey of alternative splicing in allopolyploid cotton: Landscape, complexity and regulation. *New Phytol.,* **2018**, *217*(1), 163-178. http://dx.doi.org/10.1111/nph.14762 PMID: 28892169

[44] Hao, C.; Ge, G.; Xiao, P.; Zhang, Y.; Yang, L. The first insight into the tissue specific taxus transcriptome *via* Illumina second generation sequencing. *PLoS One,* **2011**, *6*(6), e21220. http://dx.doi.org/10.1371/journal.pone.0021220 PMID: 21731678

[45] Zander, M.; Lewsey, M.G.; Clark, N.M.; Yin, L.; Bartlett, A.; Saldierna Guzmán, J.P.; Hann, E.; Langford, A.E.; Jow, B.; Wise, A.; Nery, J.R.; Chen, H.; Bar-Joseph, Z.; Walley, J.W.; Solano, R.; Ecker, J.R. Integrated multi-omics framework of the plant response to jasmonic acid. *Nat. Plants,* **2020**, *6*(3), 290-302. http://dx.doi.org/10.1038/s41477-020-0605-7 PMID: 32170290

[46] Hartmann, L.; Drewe-Boß, P.; Wießner, T.; Wagner, G.; Geue, S.; Lee, H.C.; Obermüller, D.M.; Kahles, A.; Behr, J.; Sinz, F.H.; Rätsch, G.; Wachter, A. Alternative splicing substantially diversifies the transcriptome during early photomorphogenesis and correlates with the energy availability in Arabidopsis. *Plant Cell,* **2016**, *28*(11), 2715-2734. http://dx.doi.org/10.1105/tpc.16.00508 PMID: 27803310

[47] Moreno, J.E.; Shyu, C.; Campos, M.L.; Patel, L.C.; Chung, H.S.; Yao, J.; He, S.Y.; Howe, G.A. Negative feedback control of jasmonate signaling by an alternative splice variant of JAZ10. *Plant Physiol.,* **2013**, *162*(2), 1006-1017. http://dx.doi.org/10.1104/pp.113.218164 PMID: 23632853

[48] Shen, Y.; Zhou, Z.; Wang, Z.; Li, W.; Fang, C.; Wu, M.; Ma, Y.; Liu, T.; Kong, L.A.; Peng, D.L.; Tian, Z. Global dissection of alternative splicing in paleopolyploid soybean. *Plant Cell,* **2014**, *26*(3), 996-1008. http://dx.doi.org/10.1105/tpc.114.122739 PMID: 24681622

[49] Wang, T.; Wang, H.; Cai, D.; Gao, Y.; Zhang, H.; Wang, Y.; Lin, C.; Ma, L.; Gu, L. Comprehensive profiling of rhizome-associated alternative splicing and alternative polyadenylation in moso bamboo (*Phyllostachys edulis*). *Plant J.,* **2017**, *91*(4), 684-699. http://dx.doi.org/10.1111/tpj.13597 PMID: 28493303

[50] Han, Y.; Zhu, L.; Li, L.; Wang, Y.; Zhao, M.; Wang, K.; Sun, C.; Chen, J.; Liu, L.; Chen, P.; Lei, J.; Wang, Y.; Zhang, M. Characteristics of RNA alternative splicing and its potential roles in ginsenoside biosynthesis in a single plant of ginseng, Panax ginseng C.A. Meyer. *Mol. Genet. Genomics,* **2021**, *296*(4), 971-983. http://dx.doi.org/10.1007/s00438-021-01792-z PMID: 34008042

[51] Qiao, D.; Yang, C.; Chen, J.; Guo, Y.; Li, Y.; Niu, S.; Cao, K.; Chen, Z. Comprehensive identification of the full-length transcripts and alternative splicing related to the secondary metabolism pathways in the tea plant (*Camellia sinensis*). *Sci. Rep.,* **2019**, *9*(1), 2709. http://dx.doi.org/10.1038/s41598-019-39286-z PMID: 30804390

[52] Chen, T.W.; Wu, T.H.; Ng, W.V.; Lin, W.C. Interrogation of alternative splicing events in duplicated genes during evolution. *BMC Genom.,* **2011**, *12*(Suppl. 3), S16. http://dx.doi.org/10.1186/1471-2164-12-S3-S16 PMID: 22369477

[53] Hao, D.C. *Ranunculales medicinal plants: Biodiversity, chemodiversity and pharmacotherapy*; Elsevier/Academic Press: London, **2018**.

[54] Wang, W.; Zhou, Y.; Wu, Y.; Dai, X.; Liu, Y.; Qian, Y.; Li, M.; Jiang, X.; Wang, Y.; Gao, L.; Xia, T. Insight into catechins metabolic pathways of *Camellia sinensis* based on genome and transcriptome analysis. *J. Agric. Food Chem.,* **2018**, *66*(16), 4281-4293.
http://dx.doi.org/10.1021/acs.jafc.8b00946 PMID: 29606002

[55] Zhang, Y.; Xu, Z.; Ji, A.; Luo, H.; Song, J. Genomic survey of bZIP transcription factor genes related to tanshinone biosynthesis in *Salvia miltiorrhiza. Acta Pharm. Sin. B,* **2018**, *8*(2), 295-305.
http://dx.doi.org/10.1016/j.apsb.2017.09.002 PMID: 29719790

[56] Xu, Z.; Peters, R.J.; Weirather, J.; Luo, H.; Liao, B.; Zhang, X.; Zhu, Y.; Ji, A.; Zhang, B.; Hu, S.; Au, K.F.; Song, J.; Chen, S. Full-length transcriptome sequences and splice variants obtained by a combination of sequencing platforms applied to different root tissues of Salvia miltiorrhiza and tanshinone biosynthesis. *Plant J.,* **2015**, *82*(6), 951-961.
http://dx.doi.org/10.1111/tpj.12865 PMID: 25912611

[57] Li, Y.; Mi, X.; Zhao, S.; Zhu, J.; Guo, R.; Xia, X.; Liu, L.; Liu, S.; Wei, C. Comprehensive profiling of alternative splicing landscape during cold acclimation in tea plant. *BMC Genomics,* **2020**, *21*(1), 65.
http://dx.doi.org/10.1186/s12864-020-6491-6 PMID: 31959105

[58] Lin, A.; Ma, J.; Xu, F.; Xu, W.; Jiang, H.; Zhang, H.; Qu, C.; Wei, L.; Li, J. Differences in alternative splicing between yellow and black-seeded rapeseed. *Plants,* **2020**, *9*(8), 977.
http://dx.doi.org/10.3390/plants9080977 PMID: 32752101

[59] Chao, Q.; Gao, Z.F.; Zhang, D.; Zhao, B.G.; Dong, F.Q.; Fu, C.X.; Liu, L.J.; Wang, B.C. The developmental dynamics of the Populus stem transcriptome. *Plant Biotechnol. J.,* **2019**, *17*(1), 206-219.
http://dx.doi.org/10.1111/pbi.12958 PMID: 29851301

[60] Wang, L.; Jiang, X.; Wang, L.; Wang, W.; Fu, C.; Yan, X.; Geng, X. A survey of transcriptome complexity using PacBio single-molecule real-time analysis combined with Illumina RNA sequencing for a better understanding of ricinoleic acid biosynthesis in *Ricinus communis*. *BMC Genomics,* **2019**, *20*(1), 456.
http://dx.doi.org/10.1186/s12864-019-5832-9 PMID: 31170917

[61] Gao, T.; Xu, Z.; Song, X.; Huang, K.; Li, Y.; Wei, J.; Zhu, X.; Ren, H.; Sun, C. Hybrid sequencing of full-length cDNA transcripts of the medicinal plant *Scutellaria baicalensis*. *Int. J. Mol. Sci.,* **2019**, *20*(18), 4426.
http://dx.doi.org/10.3390/ijms20184426 PMID: 31505762