

RESEARCH ARTICLE

Open Access

# Comparative transcriptome analysis of male and female flowers in *Spinacia oleracea* L



Ning Li, Ziwei Meng, Minjie Tao, Yueyuan Wang, Yulan Zhang, Shufen Li, Wujun Gao and Chuanliang Deng\*

## Abstract

**Background:** Dioecious spinach (*Spinacia oleracea* L.), a commercial and nutritional vegetable crop, serves as a model for studying the mechanisms of sex determination and differentiation in plants. However, this mechanism is still unclear. Herein, based on PacBio Iso-seq and Illumina RNA-seq data, comparative transcriptome analysis of male and female flowers were performed to explore the sex differentiation mechanism in spinach.

**Results:** Compared with published genome of spinach, 10,800 transcripts were newly annotated; alternative splicing, alternative polyadenylation and lncRNA were analyzed for the first time, increasing the diversity of spinach transcriptome. A total of 2965 differentially expressed genes were identified between female and male flowers at three early development stages. The differential expression of RNA splicing-related genes, polyadenylation-related genes and lncRNAs suggested the involvement of alternative splicing, alternative polyadenylation and lncRNA in sex differentiation. Moreover, 1946 male-biased genes and 961 female-biased genes were found and several candidate genes related to gender development were identified, providing new clues to reveal the mechanism of sex differentiation. In addition, weighted gene co-expression network analysis showed that auxin and gibberellin were the common crucial factors in regulating female or male flower development; however, the closely co-expressed genes of these two factors were different between male and female flower, which may result in spinach sex differentiation.

**Conclusions:** In this study, 10,800 transcripts were newly annotated, and the alternative splicing, alternative polyadenylation and long-noncoding RNA were comprehensively analyzed for the first time in spinach, providing valuable information for functional genome study. Moreover, candidate genes related to gender development were identified, shedding new insight on studying the mechanism of sex determination and differentiation in plant.

**Keywords:** Spinach, Sex determination and differentiation, Gibberellin, Auxin, Sugar, Full-length transcriptome

## Background

Most animals are dioecious, whereas only approximately 6% of angiosperm plants are dioecious [1]. Male and female sterile mutations in a pair of proto sex chromosomes result in dioecy production. To date, there are two viewpoints about sex determination mechanism in dioecious plants: one is two-gene mutant model and the other is single-factor mutant model. In two-gene mutant model, one mutant occurs in a stamen fertility-related

gene and results in male sterility; the other mutant occurs in pistil-related gene and suppresses female organ development [2–7]. In single-factor mutant model, the emergence of dioecious plants may be caused by two independent mutations in a single gene, one is functional inactivation mutation leading to male sterility, and the other is functional acquisition mutation leading to female sterility [8, 9]. The genes mutated in the two theoretical models are called sex-determining genes, and the autosomes that carry the sex-determining genes evolved into sex chromosomes, and finally XY or ZW sex-determination system were formed [9]. Sex-determining

\* Correspondence: [dengchuanliang@htu.edu.cn](mailto:dengchuanliang@htu.edu.cn)  
College of Life Sciences, Henan Normal University, Xinxiang 453007, China



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

gene is located in the non-recombining region of the Y/W chromosome, while, repetitive sequences are usually around the non-recombining region, which impedes the uncovering of sex determination mechanism in dioecious plants. Till now, only a few sex-determining genes have been reported in plants, such as *Diospyros lotus* [10, 11], *Asparagus officinalis* [12, 13], *Date palm* [14], *Actinidia chinensis* [15], *Vitis vinifera* [16, 17], *Ficus carica* [18], *Populus trichocarpa* [19] and *Fragaria octoploids* [20]. Although these identified genes are involved in or determine the emergence of unisexual flower, the exact mechanism of sex determination in plants remains unclear.

The differentiation of male and females is variable in dioecious plants. For example, the unisexual flower of some dioecious plants has both stamen primordia and pistil primordia and one is selectively aborted before maturity, such as garden asparagus. However, some other dioecious plants only form male or female primordia and then develop to male or female flower, such as spinach [21, 22]. Series of regulatory factors, such as transcription factors, noncoding RNAs, epigenetic marks, hormones and other metabolites are involved in sex differentiation [10, 21]. The sex-determining genes may cooperate with genetic networks and metabolic pathways involved in sex differentiation to control unisexual flower development. Hence, identification of the sex differentiation-related genes is helpful to elucidate the mechanism underlying the sex determination and differentiation.

Spinach (*Spinacia oleracea*,  $2n = 12$ ), an annual dioecious plant with short lifecycle, is a good model for studying sex determination of dioecious plants. Spinach is a member of *Spinacia* genus, which included another two species *S. turkestanica* Ilj. and *S. tetrandra* Stev.. The three species of the *Spinacia* genus are dioecious and occasionally monoecious, which is controlled by a pair of sex chromosomes XX/XY. The co-existence of homomorphic and heteromorphic sex chromosomes under genus *Spinacia* makes spinach as a good model plant for investigating the evolution of sex chromosomes. Researchers have exerted considerable effort to uncover the sex-determining region (SDR) in spinach. Many molecular markers were developed in spinach and serials of sex-linked molecular markers were identified, such as T11A, V20A, and SP\_0018 [21–28]. 45 s rDNA can also be used to discriminate X and Y chromosome by FISH [29]. On the basis of sex-linked markers, Onodera and colleagues found a non-recombining region around the male-determining locus on the Y chromosome [30, 31]. These molecular markers are helpful in locating the accurate SDR and further identifying sex-determining genes. Qian et al. found the SDR located at 66.98–69.72 and 75.48–92.96 cM regions of the sex

chromosome [32]. Spinach genome data was published in 2017 (<http://www.spinachbase.org/>) [33], and Chr4 is the sex chromosome where the sex-determining genes reside [27]. In 2020, an 18.4 Mb sex-linked region was released by Yu et al. [34]. These works provide valuable information for uncovering sex-determining genes. According to the classical ABC model for flower development, the B class floral identity genes *SpAPETALLA3* (*SpAP3*) and *SpPISTILLATA* (*SpPI*) have been identified as the masculinizing factors in spinach [35, 36]. Gibberellic acid (GA) induces sex conversion in spinach [37, 38]. Recently, a DELLA family gene, *GIBBERELLIC ACID INSENSITIVE* (*SpGAI*), has been characterized; it modulates GA level and further control the sex differentiation of spinach [39]. However, the mechanism of sex determination and differentiation in spinach remains unclear.

The length of a complete transcript is 1000–6000 bases in general. Due to the short reading length of the second generation sequencing technology, the resulting sequencing fragments need to be assembled, and the resulting transcripts may produce splicing errors and more chimeras, so that the complete transcripts cannot be obtained [40]. PacBio Iso-seq, the third-generation sequencing technology with long reading length (average 15 KB), can help us directly obtained complete full-length transcripts without splicing. It is beneficial to improve the genome sequence and study mRNA structure such as alternative splicing, fusion gene and allele expression and so on [41].

Hence, we utilized PacBio Iso-seq and Illumina RNA-seq to detect the candidate genes for sex determination and differentiation. Weighted gene co-expression network analysis (WGCNA) was further performed to construct the regulation network of the candidate genes. Our study show that sex differentiation may be regulated by the crosstalk among sugar, auxin and gibberellin signaling; different transcription factors between female and male may receive signals and target cell development-related genes to modulate floral whorl initiation.

## Results

### PacBio Iso-seq sequencing

The transcriptome of the spinach flower at the early developmental stage were sequenced using the PacBio Iso-seq platform (Additional file 1). A total of 373,155 circular consensus sequence reads were obtained. The reads length ranged from 50 bp to 15,000 bp, and the mean value was 2483 bp (Fig. 1a). We obtained 34,443 high-quality isoforms, of which 98.63% were mapped to the reference genome (Fig. 1b). A total of 15,267 isoforms encoded by 10,506 genes were mapped to the reference genome after clustering and removing redundancy. Among these mapped isoforms, 4467 known isoforms (mapped to 4445 annotated genes), 9786 new isoforms (mapped to the

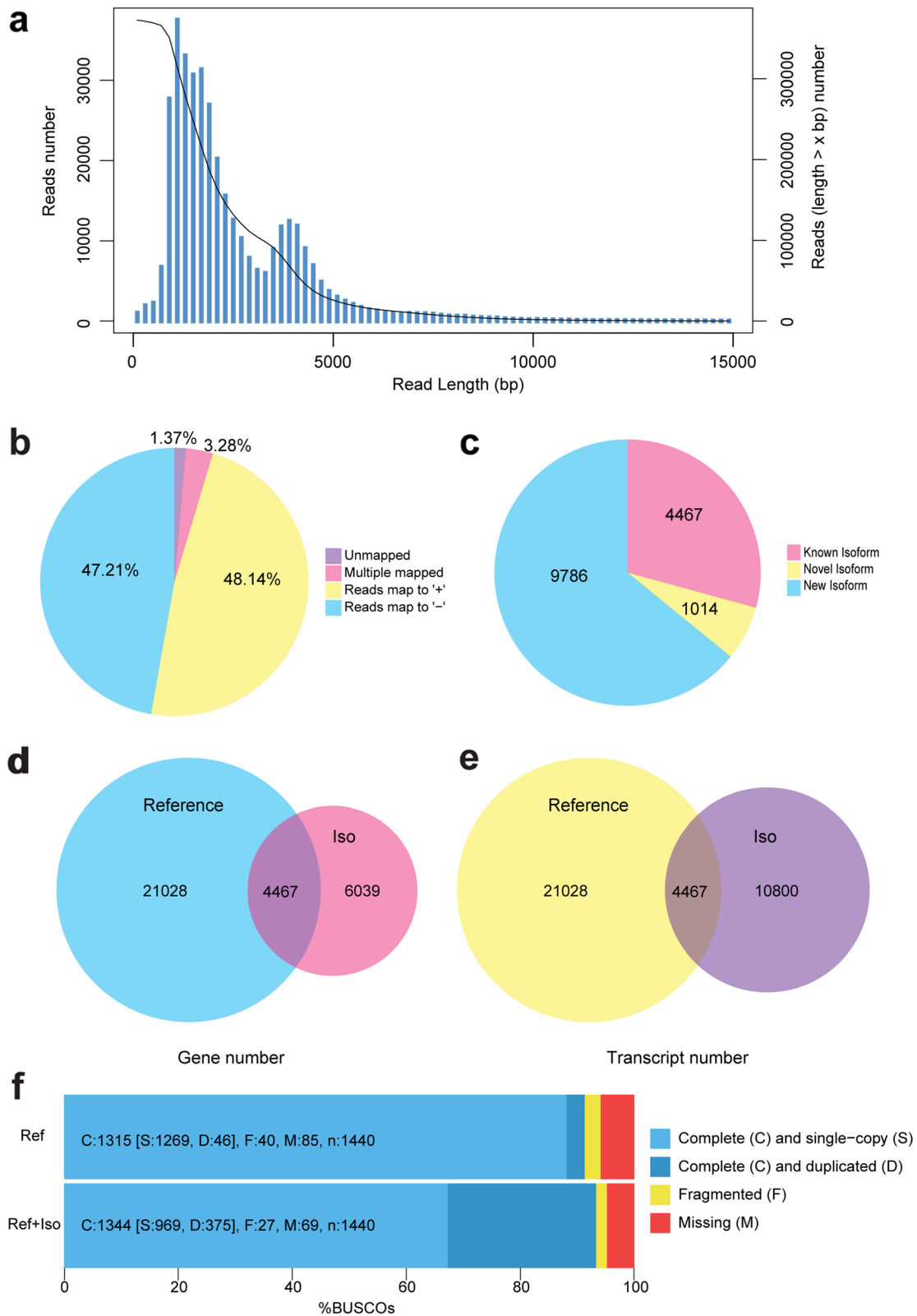


Fig. 1 (See legend on next page.)

(See figure on previous page.)

**Fig. 1** Summary of Iso-seq data. **a** CCS read length distribution, the X-axis represents the read length, the Y-axis on the left represents the column graph coordinate, indicating the number of reads with a certain length (X-axis), and the Y-axis on the right is the curve coordinate, indicating the number of reads whose length is bigger than a certain X-axis value. **b** The percentage of high quality isoforms mapped to the reference genome. **c** Classification and number of isoforms mapped to the reference genome. Known isoform was mapped to annotated genes in reference genome, new isoform was mapped to different exons of annotated gene in reference genome, and novel isoform was mapped to unannotated region in reference genome. **d** and **e** The number of genes and transcripts annotated in reference genome and Iso-seq data. **f** BUSCO assessment results of reference genome and optimized genome of spinach. Ref, reference genome; Ref + Iso, genome optimized via PacBio Iso-seq sequencing

different exons of annotated gene), and 1014 novel isoforms (mapped to unannotated region of the genome) were found (Fig. 1c). The isoforms were annotated with the NCBI non-redundant protein (Nr), Kyoto Encyclopedia of Genes and Genomes (KEGG), Swiss-Prot, and Gene Ontology (GO) databases. Conclusively, 6039 new genes can be added to the reference genome (Fig. 1d). Moreover, 10,800 transcripts encoded by 6061 genes were newly annotated, thereby increasing the diversity of the reference transcriptome (Fig. 1e). Moreover, we evaluated the BUSCO score of spinach genome optimized by our full length transcriptome, which is 93.33% and higher than published spinach genome (91.32%). Clearly, our work improved the spinach genome annotation (Fig. 1f).

#### The genes residing in sex chromosome

In the published spinach genome data, chr4 is the sex chromosome where the sex-determining genes reside [27]. According to Iso-seq data, 2190 isoforms encoded by 1489 genes reside in chr4. Blasting with spinach genome, we found 80 isoforms are single with no copy in other chromosomes or scaffolds, and the rest isoforms are duplicated. The expression of these isoforms residing in sex chromosome was further analyzed using RNA-seq data and 367 isoforms showed sex-biased expression (Fig. 2). The sex determining region (SDR) has not been fully uncovered in spinach. Till now, at least 12 genes in male-specific region were identified by Okazaki et al. (2019) [27] and a 18.4 Mb sex-linked region was revealed by Yu et al. (2020) [34]. Blasting with these genes or sequence, ten sex-biased genes fall within the 18.4 Mb sex-linked regions, which may be involved in sex determination or differentiation in spinach (Table 1).

SNP and InDel analysis were performed for RNA-seq data. We identified 131 SNPs in 114 sex-chromosomal genes, which were homozygous mutations (1|1) or no mutations (0|0) in all female samples and heterozygous mutations (0|1) in all male samples. Moreover, 11 of the 114 sex-chromosomal genes displayed sex-biased expression (Additional file 10). Of the 114 sex-chromosomal genes, two genes overlap with the Y-linked genes [27] and one gene falls within 18.4 Mb sex-linked region [34],

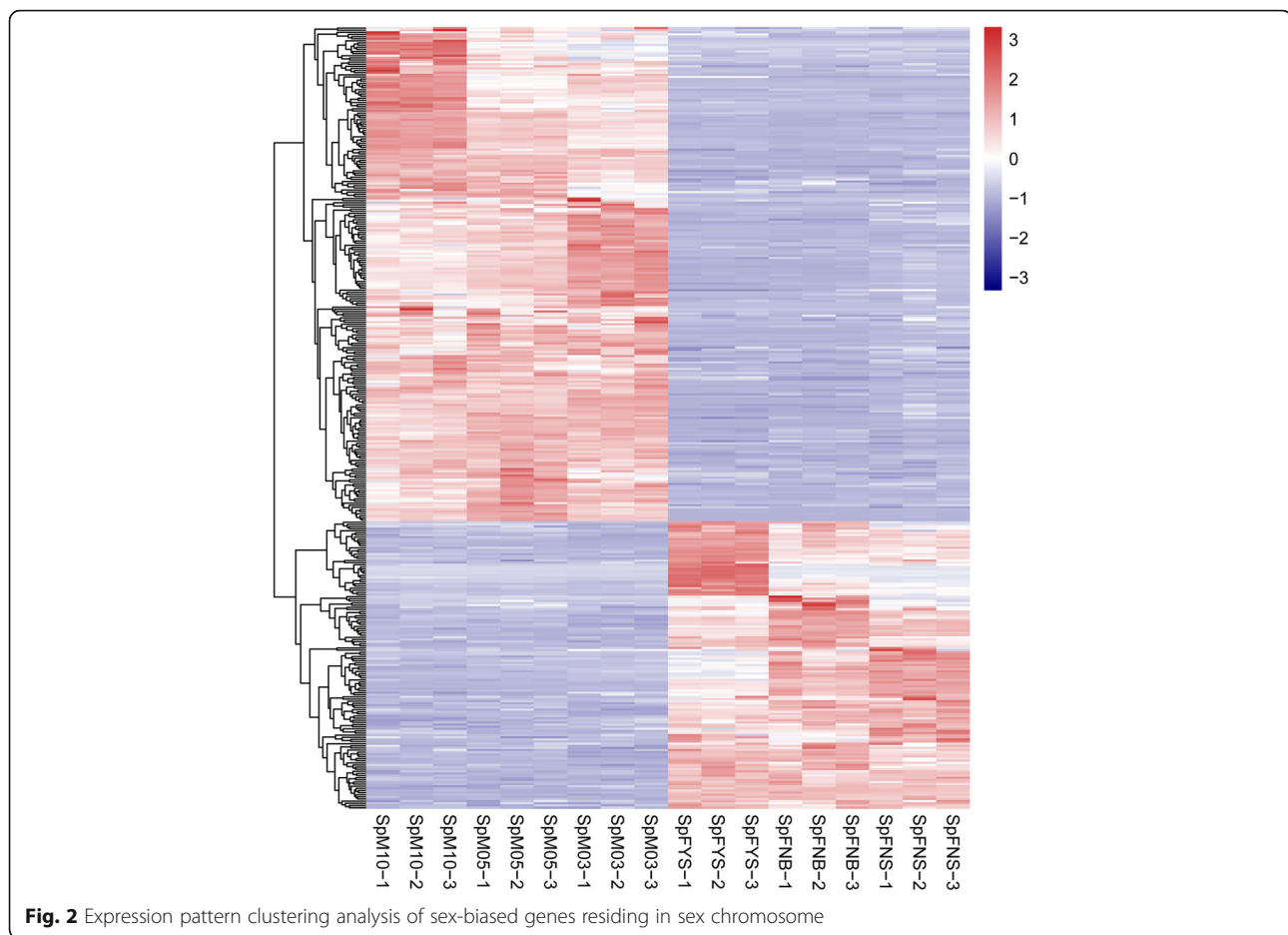
but their expression were not sex-biased (Additional file 10).

#### Alternative splicing analysis

Alternative splicing is an important mechanism of post-transcriptional regulation and influences protein coding to modulate plant development. Using the Iso-seq data, we investigated alternative splicing in spinach flower for the first time. There are 2391 alternative splicing events, including 252 skipping exon (SE), 405 alternative 5' splice sites (A5), 805 alternative 3' splice sites (A3), 10 mutually exclusive exons (MX), 773 retained intron (RI), 126 alternative first exons (AF), and 20 alternative last exons (AL) events (Fig. 3a). As shown in Fig. 3b, 7,337 genes (69.84%) generated only a single isoform, 3169 genes (30.16%) produced two to five transcripts, and 54 genes encoded 6–14 splice isoforms. Furthermore, 8442 annotated genes in the reference genome yielded splicing variants in our data. Five genes were randomly selected to validate the alternative splicing events by RT-PCR (Additional file 2). We analyzed the distribution of alternative splicing events in the six assembled chromosomes and found that the sex chromosome chr4 produced the most alternative splicing events (Fig. 3c). There are 1489 genes reside in chr4, 199 of which produced alternative splicing isoforms. Through alternative splicing, a gene may produce variant transcripts to regulate different biological process, such as induce of female or male floral primordia. According to RNA-seq results, 18 genes related to RNA splicing displayed differentially expressed and 15 of them showed sex-biased expression (described later) (Fig. 3f), indicating that alternative splicing may be involved in spinach sex differentiation.

#### Alternative polyadenylation analysis

Most eukaryotic genes contain more than one polyadenylation sites. Alternative polyadenylation influence the coding region, stability, and localization and translation efficiency of mRNA [42]. Alternative polyadenylation can increase the complexity of the transcriptome and regulate gene expression to control various plant developmental processes from cell cycle to flowering and floral organ identity [43, 44].



Herein, alternative polyadenylation analysis in spinach flower was first performed using Iso-seq. Of the 10,506 genes obtained by Iso-seq, 5032 have more than two alternative polyadenylation sites (Fig. 3d). Five genes were randomly selected to verify the alternative polyadenylation events by 3' rapid amplification of cDNA ends (3' RACE) (Additional file 2). The clear bias of nucleotides was determined by analyzing the nucleotide composition upstream ( $-50$  nts) and downstream ( $+50$  nts) of polyadenylation sites, as shown in Fig. 3e. The nucleotide bias is similar to 3' UTR profiles as previously reported [45]. According to RNA-seq results, three genes related to polyadenylation showed sex-biased expression (described later) (Fig. 3f). Hence, alternative polyadenylation occurring in flower development-related genes may be a crucial regulatory mechanism of sex differentiation.

#### LncRNA identification

Long non-coding RNAs (lncRNAs) can participate in eukaryotic gene regulation. However, lncRNAs are still largely unknown in spinach, and herein lncRNAs were analyzed in spinach for the first time. After assessing the protein-coding potential with CPC and CNCI software

and annotating protein with Swiss-Prot database, we identified 500 lncRNAs, which include 201 intergenic lncRNAs, 30 bidirectional lncRNAs, 61 intronic lncRNAs, 45 antisense lncRNAs, 149 sense lncRNAs, and 20 other lncRNAs (Fig. 4a, b). Five genes were randomly selected to verify the lncRNAs by PCR (Additional file 3). As shown in Fig. 4c, the expression profile of 42 lncRNAs (marked by asterisk) is distinct between male and female flowers, confirming the tissue-specific expression of lncRNA. The distinct expression of four lncRNAs (Gene004422, Gene002735, Gene004186 and Gene001186) was further validated by qPCR (Additional file 8). Hence, lncRNAs with sex-biased expression (described later) may be involved in the sex differentiation.

#### Sex-biased genes of spinach

Sex-biased genes, which exhibit significantly higher expression in one sex than in the other sex, always act downstream of sex-determining gene. Hence, identification of the sex-biased genes is helpful to uncover the sex-determination mechanism [46]. Herein, a total of 2965 DEGs (differentially expressed genes) were

**Table 1** List of sex-biased genes residing in 18.4 Mb of sex-linked region in spinach

Gene ID	Genome location	Type	Log <sub>2</sub> (Fold Change)			Description
			M03/FNS	M05/FNB	M10/FYS	
<i>Gene003616</i>	chr4: (+) 91,520 ... 105,982 bp	new	-1.9	-1.66	-1.07	cyclin-dependent kinase
<i>Gene004137</i>	chr4: (-) 50,043,916 ... 50,045,261 bp	novel	1.36	1.38	1.49	uncharacterized LOC104902451
<i>Gene004138</i>	chr4: (+) 50,048,233 ... 50,049,888 bp	novel	3.35	4.75	4.9	serine/threonine-protein kinase D6PKL1-like
<i>Gene004422</i>	chr4: (-) 90,648,657 ... 906,531,003 bp	novel	-1.67	-1.73	-2.54	uncharacterized LOC104906114
<i>Gene005061</i>	chr4: (+) 122,096,317 ... 122,097,168 bp	novel	7.12	F = 0	F = 0	-
<i>Spo16245</i>	chr4: (-) 57,224,763 ... 57,239,708 bp	known	F = 0	F = 0	F = 0	RNA-binding family protein
<i>Spo21056</i>	chr4: (+) 50,045,074 ... 50,047,854 bp	known	1.07	1.79	1.8	Microtubule associated family protein
<i>Spo26984</i>	chr4: (-) 41,263,089 ... 41,268,484 bp	known	3.34	3.96	3.1	ABC transporter B family member 11
<i>XLOC_021290</i>	chr4: (+) 57,293,175 ... 57,294,874 bp	new	-4.24	-3.1	-3.86	-
<i>XLOC_021752</i>	chr4: (+) 99,677,867 ... 99,680,635 bp	new	4.98	F = 0	F = 0	-

Note: "F = 0" means the gene expression is not detected in female sample

identified between female and male flowers at three early developmental stages (Fig. 5a). Of the DEGs, 1946 genes are male-biased genes and 961 genes are female-biased genes (Fig. 5b). Moreover, there are 124 male-specific genes and 25 female-specific genes (Fig. 5b). Through GO functional enrichment analysis of these DEGs, we found that the end term of Directed Acyclic Graph is sugar transmembrane transporter activity (GO:0051119) in molecular function and carbohydrate transport (GO:0008643) in biological processes (Additional file 4). Moreover, these genes were obviously enriched in 11 pathways via KEGG analysis. The top 5 pathways are glycosaminoglycan degradation (ko00531), protein processing in endoplasmic reticulum (ko04141), phenylpropanoid biosynthesis (ko00940), pentose and glucuronate interconversions (ko00040), and plant hormone signal transduction (ko04075) (Additional file 4).

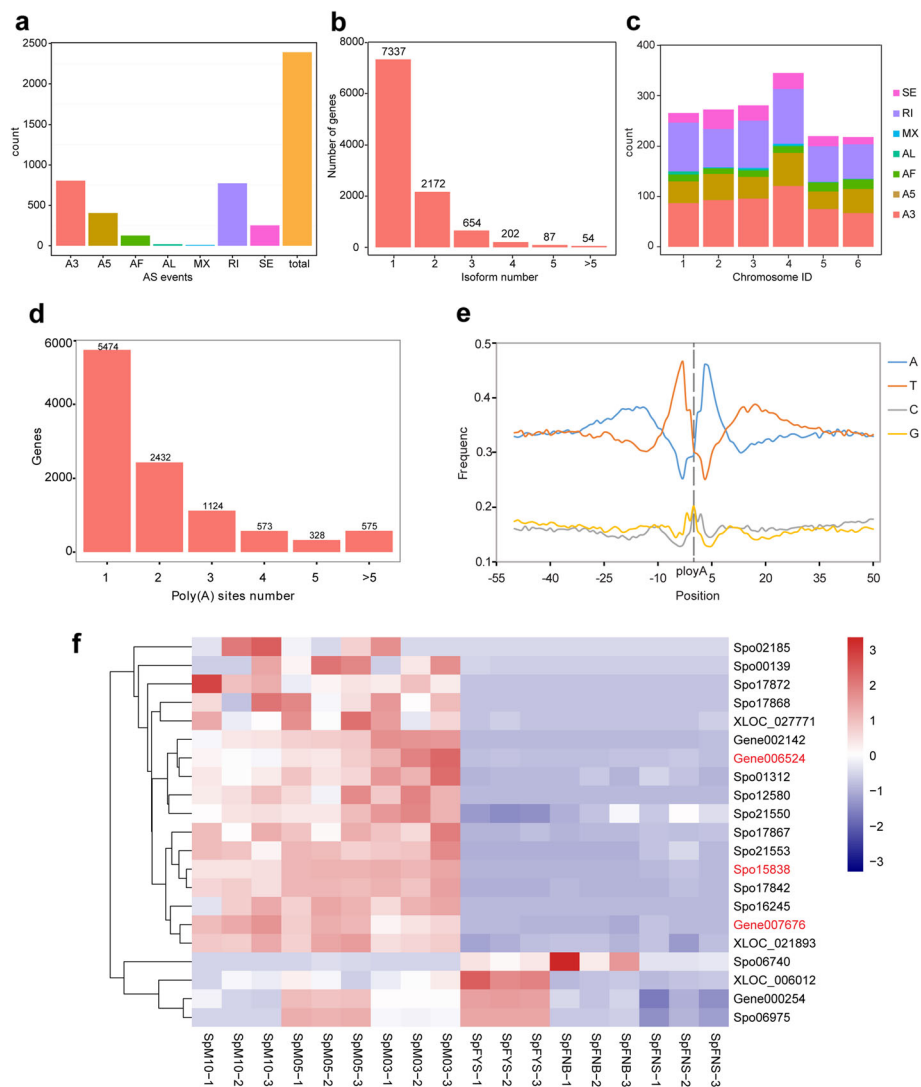
Among the DEGs, 206 transcription factors were differentially expressed between female and male flowers at the three early developmental stages. These genes are from 30 transcription factor families, of which MYB (including 26 DEGs), bHLH (including 23 DEGs), NAC (including 21 DEGs) and MADS-box (including 19 DEGs) family are the predominant transcription factor families (Fig. 6a). Of these differentially expressed transcription factors, 111 genes exhibited male-biased expression and 95 genes showed female-biased expression (Additional file 5). Moreover, six transcription factors, including two NAC genes (*Spo02680* and *Spo17876*), two

bHLH genes (*Spo26350* and *Spo26939*), one WRKY gene (*Spo09488*) and one C3H gene (*Spo24792*), show male-specific expression; and only one C2H2 gene (*Spo23846*) exhibits female-specific expression (Fig. 6c).

There were 67 DEGs involved in hormone pathway; of these DEGs, 13 were involved in JA pathway, 12 were involved in auxin pathway, 11 were involved in SA pathway, seven were involved in CK pathway, seven were involved in ethylene pathway, six were involved in gibberellin pathway, five were involved in BR pathway, and five were involved in ABA pathway (Fig. 6b). All these 67 hormone-related DEGs displayed sex-biased expression; 33 genes showed female-biased expression, while 34 genes exhibited male-biased expression. Moreover, we found that two auxin-related genes, *Spo12116* and *Spo22710*, displayed male-specific expression (Fig. 6c, Additional file 5).

#### Co-expression networks of female and male flower

We performed weighted gene co-expression network analysis (WGCNA) to construct a potential regulatory network of sex differentiation in spinach. As shown in Fig. 7a, 16 modules were established by WGCNA in 18 samples including male and female flowers at different stages. These genes are naturally clustered according to the sexuality and developmental stage. As shown in Fig. 7b, the connectivity between each two genes was calculated, and the genes in the same module displayed strong connectivity. The expression pattern of each

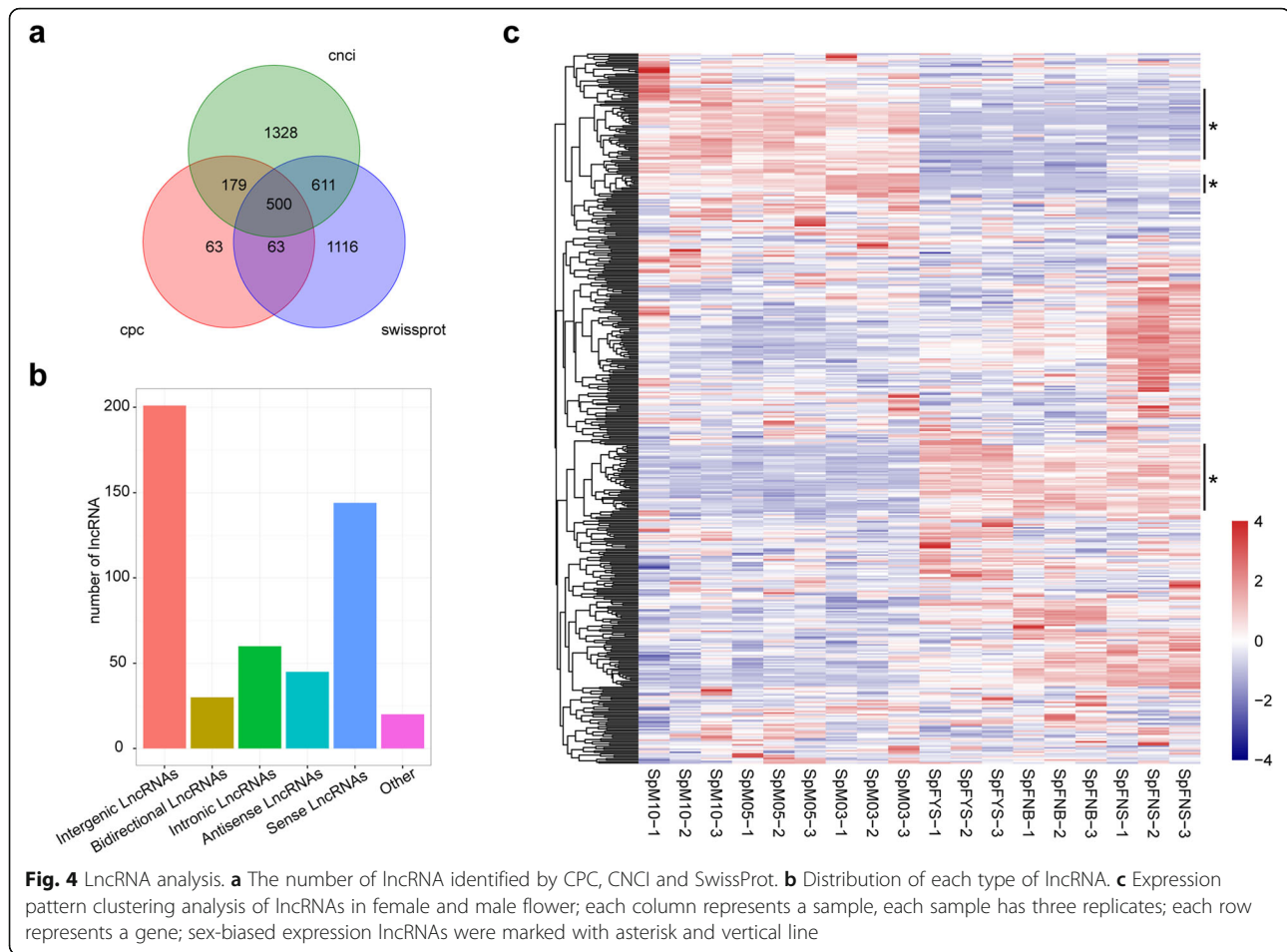


**Fig. 3** Summary of alternative splicing events and alternative polyadenylation events. **a** Alternative splicing events classification and number statistics. AS, alternative splicing; SE, skipping exon events; A5, 5' splice sites events; A3, 3' splice sites events; MX, mutually exclusive exons events; RI, retained intron events; AF, alternative first exons events; AL, alternative last exons events. **b** Distribution of genes that produce one or more splice isoforms. **c** Distribution of alternative splicing events in assembled chromosomes. **d** Distribution of the number of poly(A) sites per gene. **e** Relative frequency of each nucleotide around poly(A) cleavage sites. **f** The differentially expressed genes related to alternative splicing and alternative polyadenylation, the genes in red font are related to alternative polyadenylation, the genes in black font are related to alternative splicing

module was analyzed in all samples (Fig. 7c). Notably, the eigengene expression level of the blue module is female-biased, whereas that of the brown module is male-biased (Figs. 7c; 8); and these two modules were closely negatively correlated (Additional file 6,  $R = -0.90$ ,  $p = 5.17E-07$ ). We defined the blue module as female module and brown module as male module.

To construct the co-expression network of female module and male module, we screened the differentially expressed genes related to flower development and hormone and differentially expressed transcription factors from female and male module. As a result, 41 DEGs

were selected from female module, and 37 DEGs were selected from male module (Additional file 7). As shown in Fig. 9a, the hub gene of the female network is *Gene005140*, encoding a WRKY transcription factor. In the female network, the most closely correlated gene of the hub gene is *Spo24877* (weight value is 0.4899), encoding a Gibberellin-regulated protein and homologous to *AtGASA13* (*AT3G10185*); *Gene006949* (encoding cyclin D3 protein) is the most closely correlated gene of *Spo24877* (weight value is 0.4951) (Table 2; Fig. 9a). Moreover, eight genes related to gibberellin, auxin, abscisic acid, brassinosteroid, and salicylic pathway were



involved in female network; 32 transcription factors from 12 families were found in female network; additionally, there is one floral organ development-related gene, *Spo10951*, encoding an argonaute protein and homologous to *AtAGO5* (*AT2G27880*) (Fig. 9a).

As shown in Fig. 9b, the hub gene of the male network is *Gene000891*, encoding salicylic acid-binding protein 2-like and homologous to *AtMES3* (*AT2G23610*). In this male network, the most closely correlated gene of the hub gene is *Spo08103* (weight value is 0.5074), encoding a Gibberellin-regulated protein *GASA6* and homologous to *AtGASA6*; the most closely correlated gene of *Spo08103* is *Gene005514* (weight value is 0.5121), encoding a MADS-box protein and homologous to *AtSEP1* (Table 2; Fig. 9b). Moreover, nine genes related to gibberellin, auxin, abscisic acid, salicylic, jasmonate and cytokinin pathway were involved in male network; 24 transcription factors from 10 families were found in male network; additionally, there are three floral development-related genes, including two UDP-glucose 4-epimerase family proteins (*Spo00735* and *Spo00708*) and one Glyceraldehyde-3-phosphate dehydrogenase (*Spo21495*) (Fig. 9b).

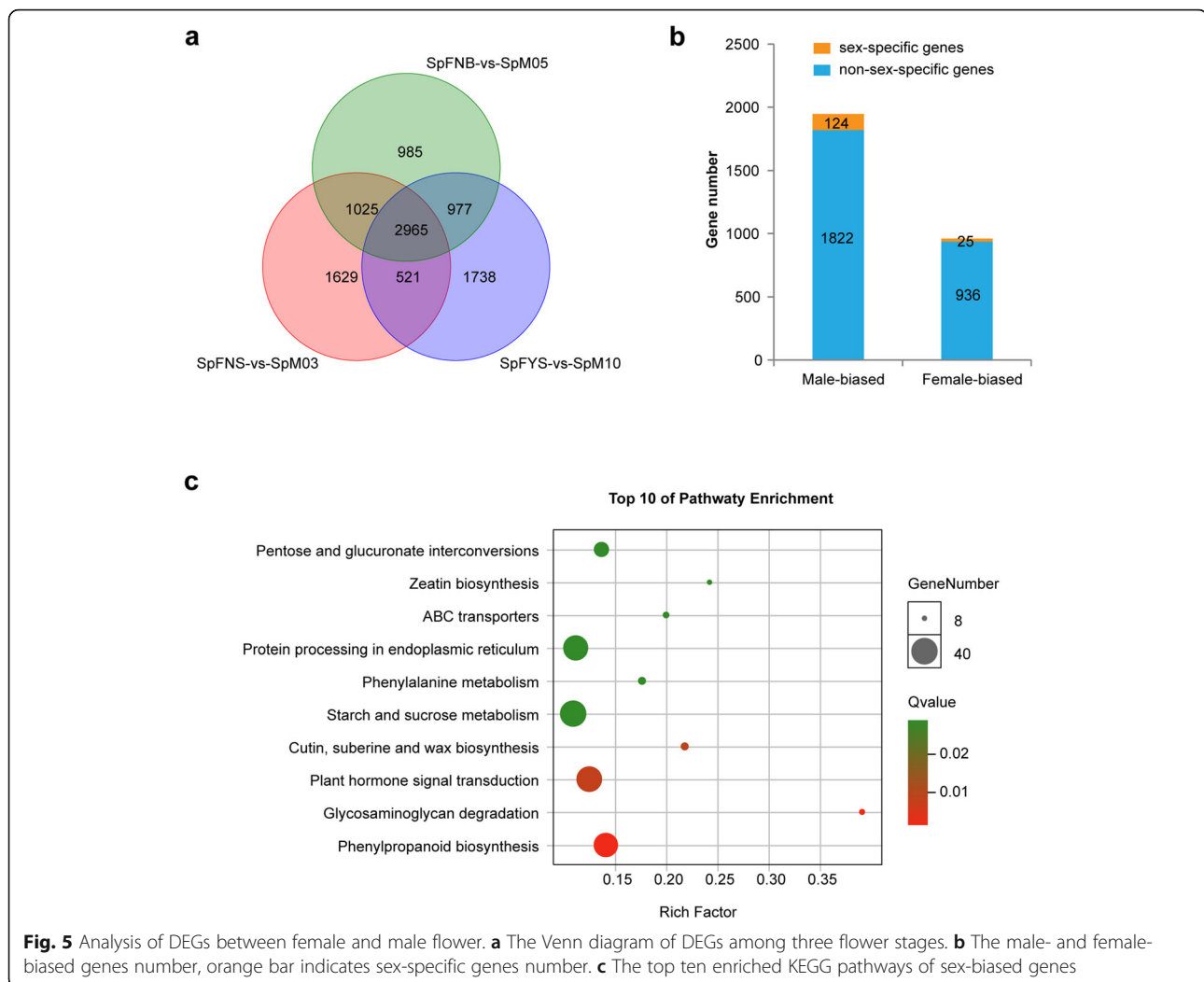
## Discussion

### Full-length transcriptome analysis

For the first time, we comprehensively analyzed the full-length transcriptome in spinach. In summary, Iso-seq data yielded a number of novel findings, including (i) identification of transcriptome-wide full-length isoforms in spinach at an unprecedented scale with 15,267 isoforms (including 10,800 newly annotated transcripts), (ii) detection of 2391 splicing events compared with the reference genome for the first time, (iii) discovery of extensive alternative polyadenylation of spinach transcripts (47.90% of expressed genes have multiple polyadenylation sites); and (iv) identification of 500 newly annotated lncRNAs in spinach for the first time. These newly annotated transcripts and analysis of alternative splice isoforms and lncRNA would provide valuable resource to improve spinach genome annotation and for future functional studies.

Alternative splicing is an important post-transcriptional regulation mechanism that affects protein coding to modulate plant development. Alternative splicing is sensitive to inner and outer cues, such as hormone and biotic





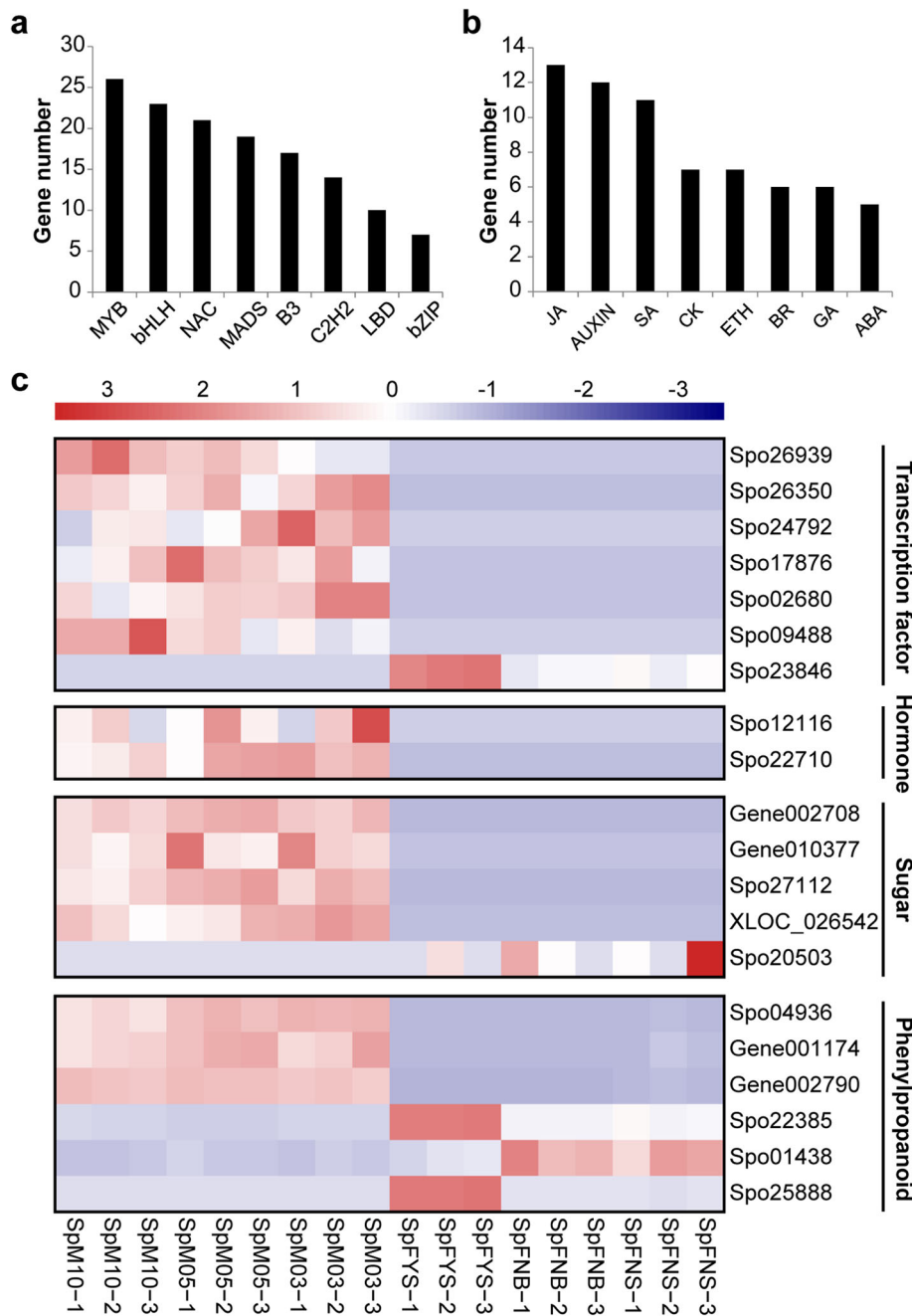
and abiotic factors, affecting plant growth [47–51]. In the current study, 30.26% genes had more than two alternative splicing isoforms. Alternative polyadenylation influences the coding region, stability, and localization and translation efficiency of mRNA [42]. Thus, it can regulate gene expression to control various plant developmental processes, such as floral organ identity [43, 44, 52–54]. Most of the DEGs associated with RNA splicing and polyadenylation showed sex-biased expression (Fig. 3). Hence, alternative splicing and alternative polyadenylation may play roles in sex differentiation. lncRNAs play roles in regulation of transcription, splicing, and nuclear structure and influences flowering time, root organogenesis, and reproduction [55–57]. As shown in Fig. 4c, 42 lncRNAs (marked by asterisks) showed distinct expression between male and female flowers, indicating their involvement in the regulation network of sex differentiation. The analysis of alternative splicing, alternative polyadenylation, and lncRNA revealed that various post-transcriptional regulations were involved in the sex differentiation and provided

more clues to explore the mechanism of sex determination and differentiation.

#### Candidate genes for spinach sex determination or differentiation

Sex-determining gene always locates in sex chromosome. In this study, 367 of 1489 sex-chromosomal genes showed sex-biased expression (Fig. 2). Moreover, ten of the 367 sex-biased genes fall within the 18.4 Mb sex-linked region [34], which can serve as candidates for sex determination or sex differentiation (Table 1). In addition, 114 sex-chromosomal genes identified by SNP analysis may be sex-linked genes, including 11 sex-biased genes; interestingly, two genes of them overlap with the reported Y-linked genes [27] and one gene falls within 18.4 Mb sex-linked region [34], but they were not sex-biased genes (Additional file 10). These candidate genes need to be validated in future work.

Identification of the sex-biased genes related to flower development is helpful for uncovering the sex determination

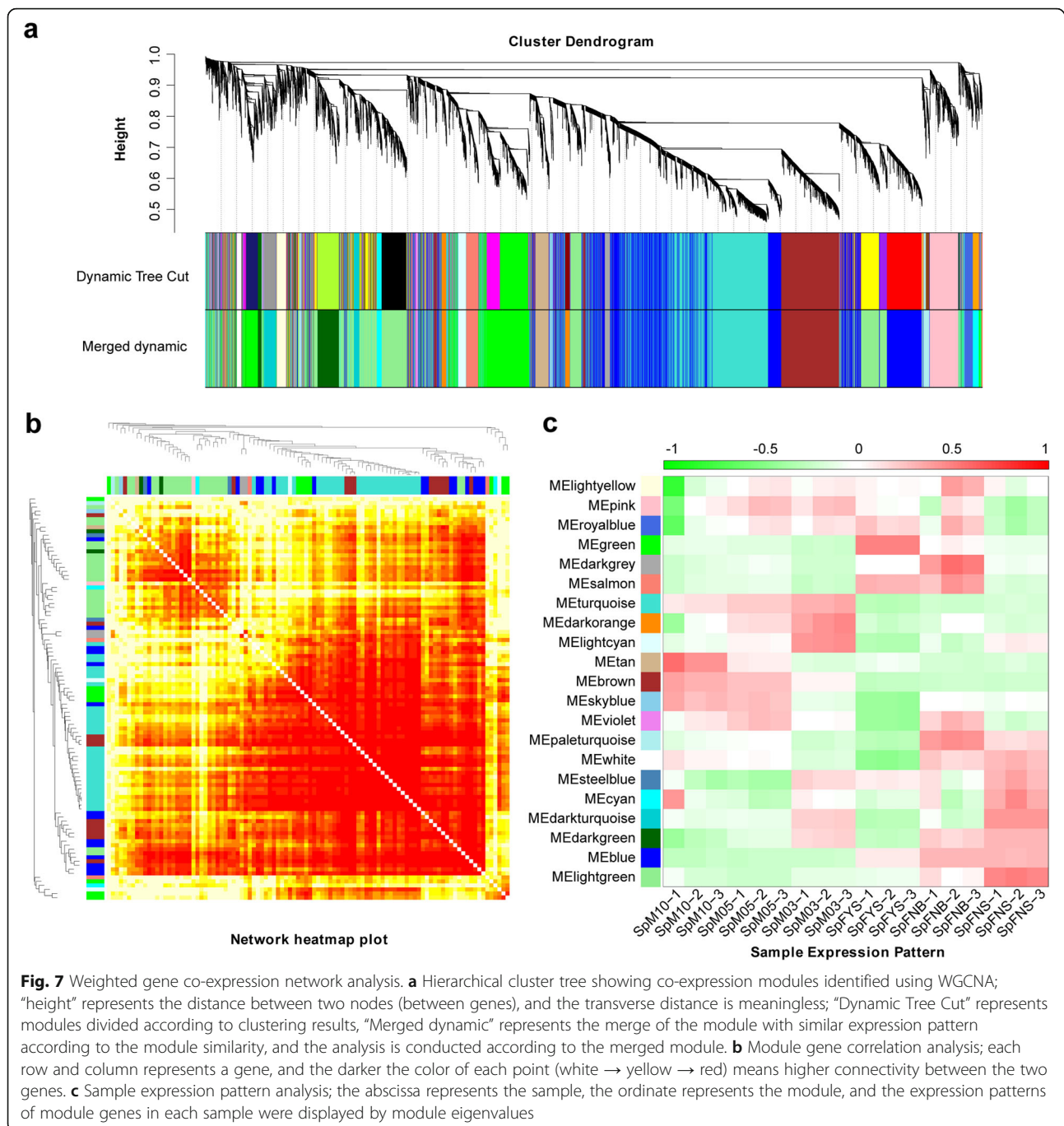


**Fig. 6** Analysis of sex-biased genes. **a** Statistics of sex-biased genes related to transcription factors (top eight families). **b** Statistics of sex-biased genes related to hormones. **c** Heat map of sex-specific DEGs related to transcription factors, hormones and sugar and top six sex-biased DEGs related to phenylpropanoid

and differentiation mechanism. Herein, 2907 sex-biased genes were found during the three early flower development stages, including 1946 male-biased genes and 961 female-biased genes; moreover, there are 124 male-specific genes and 25 female-specific genes (Fig. 5b). GO and KEGG analysis indicated that these sex-biased genes were predominantly functionally enriched in sugar metabolism,

phenylpropanoid biosynthesis, and plant hormone signal transduction (Fig. 5c).

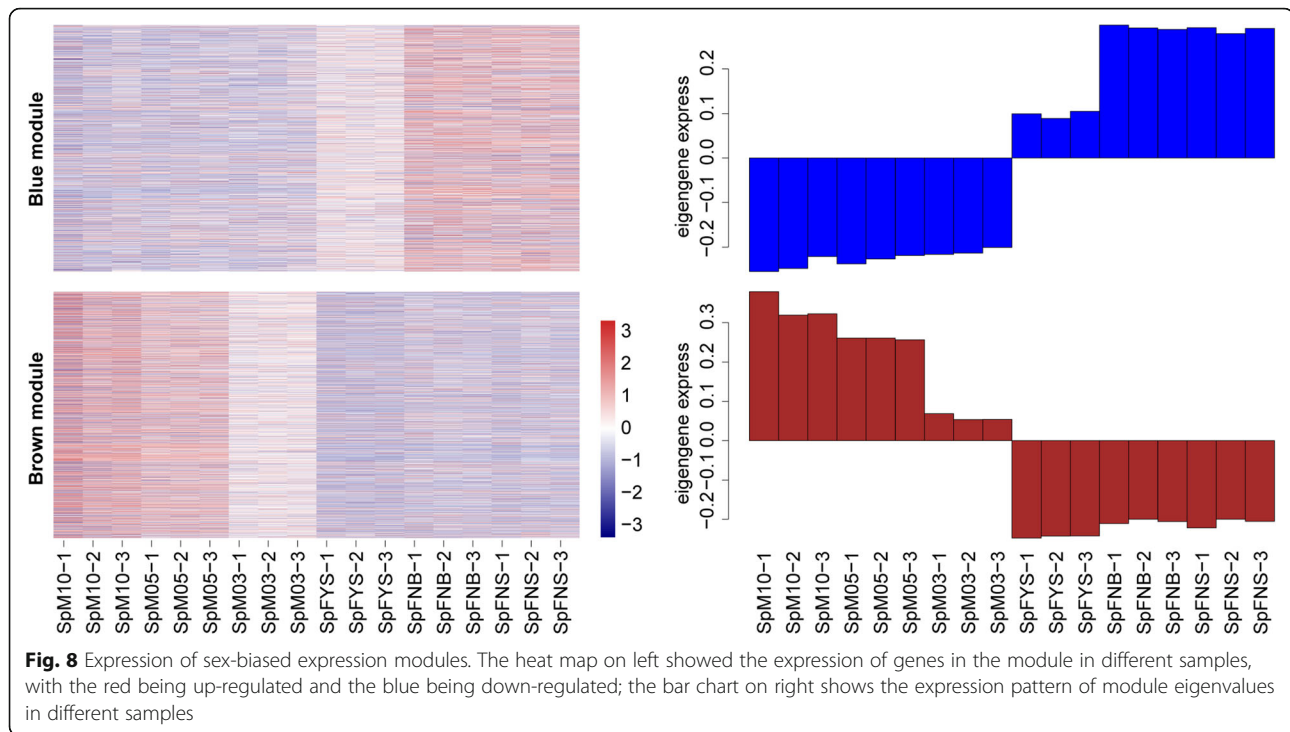
Sugars generated by photosynthesis and carbon metabolism in source and sink tissues can act as a signal to modulate various processes, such as photosynthesis, nutrient mobilization, and allocation, finally stimulating sink tissue growth, such as flower initiation and floral



organ development [58, 59]. Hexokinase (HXK) and trehalose pathways are well defined in relation to sugar signaling. For example, HXK-involved sugar signaling pathway affects monoecious cucumber flower development, that is, femaleness process [60]. In maize, a trehalose-6-phosphate phosphatase, RAMOSA3, controls inflorescence architecture by modifying trehalose-dependent sugar signaling [61]. In this study, we found 88 sex-biased genes related to sugar metabolism; of them, one gene exhibited female-specific expression (*Spo20503*) and four genes exhibited

male-specific expression (*Gene002708*, *Gene010377*, *Spo27112* and *XLOC\_026542*) (Fig. 6, Additional file 5); the roles of these genes in spinach sex differentiation will be further studied.

Phenylpropanoid and downstream flavonoid and anthocyanin pathway are involved in male fertility [62–67]. We found 43 phenylpropanoid-related genes showed sex-biased expression (Additional file 5), encoding 4-coumarate-CoA ligase (4CL), beta-glucosidase (BGLU), caffeic acid 3-O-methyltransferase (COMT), caffeoyl-CoA O-

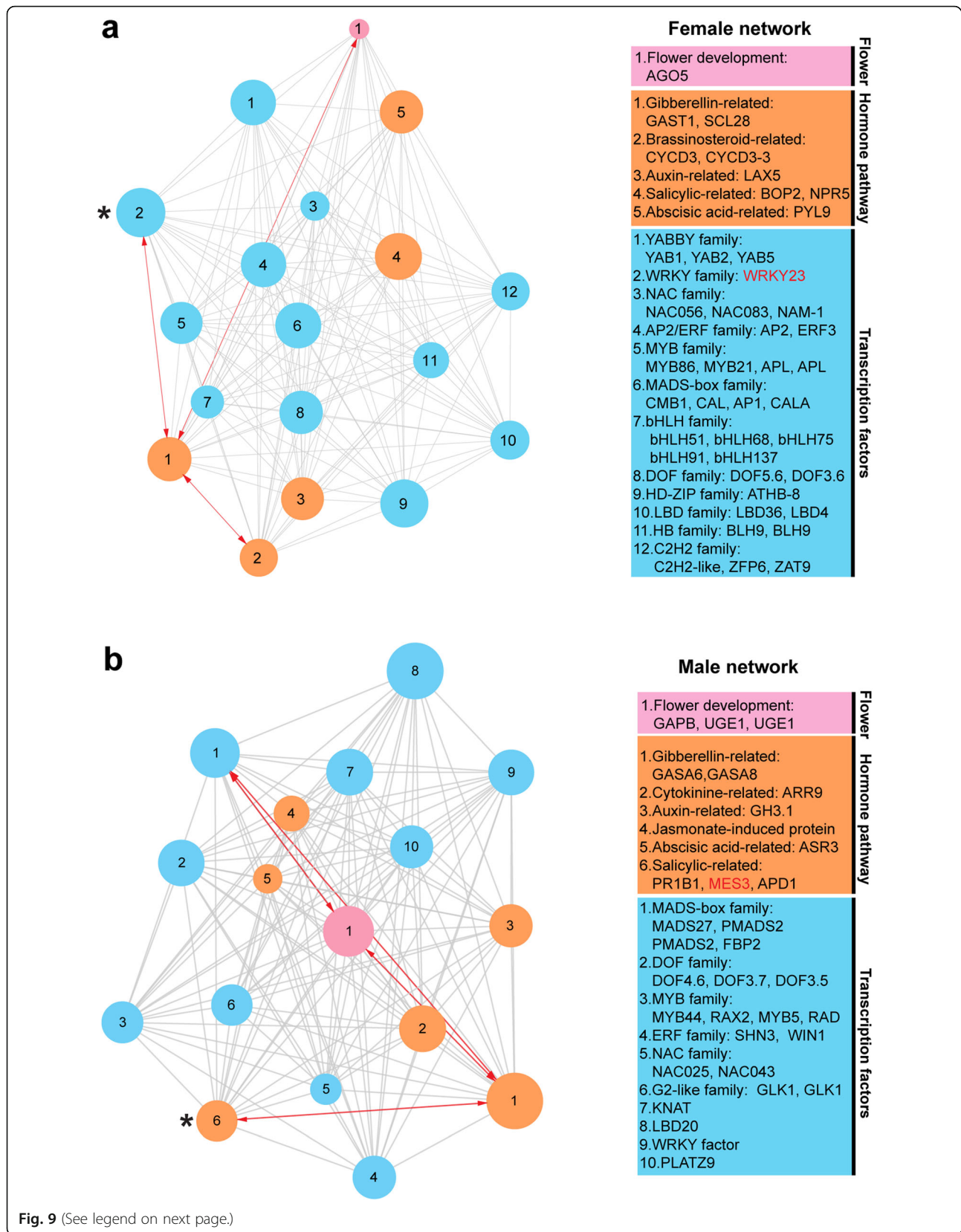


methyltransferase (CCoAOMT), chalcone synthase (CHS), cinnamyl-alcohol dehydrogenase (CAD), cytochrome P450 protein (CYP450), dihydroflavonol 4-reductase (DR), flavanone-3-hydroxylase (F3H) and peroxidase (PER). These genes are intriguing and need to be validated in future work.

Phytohormones are important regulators of flower development, such as auxin, ethylene, gibberellin (GA), cytokinin (CK), abscisic acid (ABA), brassinosteroid (BR) and jasmonate (JA) [68–70]. One hormone may crosstalk with other hormones to regulate floral organ development. For example, auxin can coordinate with JA and GA to regulate stamen development in *Arabidopsis* [68]. Ethylene promotes the femaleness of cucumber and a 1-aminocyclopropane-1-carboxylic acid synthase CmACS11 was identified as sex determinant [71], whilst GA has a negative effect on female flower formation [72, 73]. As reported in spinach, GA can promote maleness and a DELLA gene *GIBBERELLIC ACID INSENSITIVE (GAI)* was defined as feminizing factor [39]. In this paper, 67 DEGs were related to hormones, including auxin, ethylene, GA, CK, ABA, JA, BR and SA (Additional file 5). All of them displayed sex-biased expression and only two auxin-related genes, *Spo12116* and *Spo22710*, were found to display male-specific expression (Fig. 6, Additional file 5). Hence, these two auxin-related genes can be served as candidates for studying the mechanism of spinach sex differentiation.

Transcription factors are pivotal factors for transporting endogenous and exogenous signals to regulate plant

growth. Among the sex-biased genes, we found six male-specific transcription factors, including two NAC genes (*Spo02680* and *Spo17876*), two bHLH genes (*Spo26350* and *Spo26939*), one WRKY gene (*Spo09488*) and one C3H gene (*Spo24792*), and one female-specific transcription factor C2H2 gene (*Spo23846*) (Fig. 6). The male-specific NAC gene *Spo02680* located on sex chromosome chr04 (chr4: 88,744,812–88,747,497 bp) and its homologous gene in *Arabidopsis* is *SOMBRERO (AT1G79580)*, regulating the cell fate during root cap development [74]. The homologous gene in *Arabidopsis* of another male-specific NAC gene *Spo17876* is *TAPN AC (AT1G61110)*, a tapetal-specific NAC gene, affecting tapetum and pollen development and male fertility [75, 76]. The male-specific bHLH gene *Spo26350* locates on sex chromosome (chr4: 108,742,184–108,748,668 bp) and its homologous gene in *Arabidopsis* is *ABORTED MICRO-SPORES (AT2G16910)*, regulating tapetal cell development and affecting male fertility and pollen differentiation [77]. The male-specific C3H gene *Spo24792* is homologous to *Arabidopsis* gene *CALLOSE DEFECTIVE MICROSPORE 1 (AT1G68200)*, which influences the male fertility by regulating callose metabolism during microspores development [78]. The female-specific C2H2 gene *Spo23846* is homologous to *Arabidopsis* gene *TRANSPARENT TESTA 1 (AT1G34790)*, which was expressed during ovules and seeds development [79]. Clearly, these gender-specific transcription factors may be associated with spinach sex differentiation and we will verify their function in our future work.



(See figure on previous page.)

**Fig. 9** Visualization of female (a) and male (b) co-expression network. In the left network, the black stars marked hub gene, pink circles indicated flower-related genes, orange circles indicated hormone-regulated genes, blue circles indicated transcription factors, and the bigger the size of the circle means higher connectivity; the red line with arrow linked two most closely correlated genes. In the right box, genes involved in the network were listed and hub genes were in red font; gene in each circle can be found in the right box according to the color and Arabic numeral of the circle

### Female and male network

Uncovering the mechanism of sex differentiation is helpful to identify sex-determining gene. We constructed the regulation network of spinach sex differentiation via WGCNA. Interestingly, we found two modules with sex-biased expression, i.e. female module and male module (Figs. 7; 8; 9).

In female network (Fig. 9; Table 2; Additional file 7), the hub gene is *Gene005140* (*WRKY23*) and its homologue in *Arabidopsis* is *AtWRKY23*, an auxin-responsive transcription factor and playing a role in auxin canalization [80]; moreover, overexpression of tomato *WRKY23* (homologue of *AtWRKY23*) in *Arabidopsis* altered the flowering time [81]; hence, we speculate that auxin may be the key hormone involved in female flower development. *AGO5* (*Spo10951*) was the only floral-related gene except hormone and TFs in female network. As reported, *AtAGO5* can interact with miR156 to regulate flowering by suppressing SPL transcription factors [82]. As shown in Fig. 9 and Additional file 7, *Spo24877* (*GAST1*) is the most closely correlated gene both of hub gene *WRKY23* and floral-related gene *AGO5* (weight value is 0.4491); and *Gene006949* (cyclin D3 protein) is the most closely correlated gene of *GAST1*. Hence, auxin may interact with GA to regulate spinach female flower development via *WRKY23* and cell cycle-related gene cyclin D3.

In male network (Fig. 9; Table 2; Additional file 7), the hub gene is *Gene000891* (*MES3*) and its most correlated gene is *Spo08103* (*GASA6*); the homologue in *Arabidopsis* of *MES3* is *AtMES3*, which participates in auxin pathway and can hydrolyze MeIAA to active IAA [83]. Furthermore, of the DEGs-related to hormone, two auxin-related genes, *Spo12116* and *Spo22710*, displayed

male-specific expression (Fig. 6; Additional file 6). Hence, similar to female module, auxin may be the crucial hormone regulating male flower development. Except of TFs and the genes related to hormones, two UDP-glucose 4-epimerase family genes, both homologous to *Arabidopsis UGE1*, are involved in flower development. As reported, *AtUGE1* (AT1G12780) is involved in sugar flux and may be regulated by transcription factor HEME ACTIVATOR PROTEIN 3b to affect flowering [84, 85]. The closely correlated genes of spinach *UGE1* are *Spo08103* (*GASA6*) and *Gene005514* (*FBP2*) (Fig. 9; Additional file 7). Sugar signaling pathways can interact with various hormones to modulate plant growth. Daily fluctuations in GA sensitivity track the fluctuations in sugar levels and are regulated by the circadian clock [86, 87]. DELLA protein is an inhibitor of GA activity, whereas sucrose can stabilize DELLA proteins [87, 88], explaining the negative effect of GA on the sucrose-dependent induction of anthocyanin synthesis [89, 90] and indicating the involvement of DELLAs in the sucrose–GA interaction [91–93]. MADS-box family genes play important roles in flower development and form the ABCDE model to pattern floral whorls. The floral MADS-box proteins can target cell development-related genes to control flower development [94–97]. The cucumber *DELLA* gene *CsGAIP* may negatively regulate the B class floral homeotic genes *AP3* and *PI* to inhibit staminate development [98]. The B class floral identity genes, *AP3* and *PI*, act as masculinizing factors that participate in the establishment of sexual dimorphism in spinach [35, 36]. Hence, auxin, GA and sugar may cooperate together to regulate spinach male flower development through MADS-box transcription factor.

**Table 2** Hub gene in female and male network

Network	ID	Connectivity	Symbol	Arabidopsis homolog	Weight value
Female	<sup>a</sup> Gene005140	1217.92	WRKY23	AT2G47260	
	<sup>b</sup> Spo24877	1164.75	GAST1	AT3G10185	<sup>b</sup> 0.4899
	<sup>c</sup> Gene006949	1196.59	CYCD3–3	AT3G50070	<sup>c</sup> 0.4951
Male	<sup>d</sup> Gene000891	903.44	MES3	AT2G23610	
	<sup>e</sup> Spo08103	803.60	GASA6	AT1G74670	<sup>e</sup> 0.5074
	<sup>f</sup> Gene005514	723.19	FBP2	AT5G15800	<sup>f</sup> 0.5121

Note: “a” and “d” indicated the hub gene of female and male network, respectively; “b” and “e” indicated the most closely co-expressed gene of “a” and “d” gene, respectively, and the weight value between them; “c” and “f” indicated the most closely co-expressed gene of “b” and “e” gene, respectively, and the weight value between them

Comparing the co-expression network of female and male flower, we found auxin and GA were the common crucial factors in regulating female or male flower development; however, the co-expressed genes of these two factors were different, which may result in spinach sex differentiation. These hypotheses need to be verified in future work.

## Methods

### Plant materials and growth conditions

All plants used in this study were *Spinacia oleracea* L. cv DA JIAN YE BO CAI, a dioecious plant. Seeds were obtained from U.S. National Plant Germplasm System (<https://npgsweb.ars-grin.gov/gringlobal/search.aspx?>, accession number is PI 527332) and grown in an experimental field at Henan Normal University, Xinxiang, China (113.90°E, 35.32°N). Following the method of Sherry et al. [99], female flowers were collected when they were 0.3 mm in diameter (without stigma, FNS), 0.5 mm in diameter (without stigma, FNB), and mature with stigma (FYS). Male flowers were collected when they were 0.3 mm (M03), 0.5 mm (M05), and 1.0 mm (M10) in diameter. Flowers were sampled from five individuals at each stage and divided into three replicates. Female and male flowers collected at three stages were separately used for Illumina RNA-seq sequencing. The mixture of female and male flowers was used for PacBio Iso-seq sequencing.

### PacBio Iso-seq sequencing

Total RNA of the mixture of female and male flowers was prepared by TRIzol reagent (Life technologies). mRNA was enriched by Oligo (dT) magnetic beads. The enriched mRNA was reverse transcribed into cDNA using Clontech SMARTer PCR cDNA Synthesis Kit. PCR cycle optimization was performed to determine the optimal amplification cycle number for downstream large-scale PCR reactions. The optimized cycle number was then used to generate double-stranded cDNA. In addition, size selection of Cdna with > 4 kb was performed using the BluePippin™ Size-Selection System and mixed equally with no-size-selection cDNA. Large-scale PCR was performed for the next SMRTbell library construction. cDNAs were DNA damage repaired, end repaired, and ligated to sequencing adapters. The SMRTbell template was annealed to sequencing primer and bound to polymerase and sequenced on the PacBio Sequel platform using P6-C4 chemistry with 10 h movies by Gene Denovo Biotechnology Co. (Guangzhou, China).

### Annotation

The raw sequencing reads of cDNA libraries were classified and clustered into transcript consensus using the SMRT Link v5.0.1 pipeline [100] supported by Pacific Biosciences. The high-quality consensus sequences were then

mapped to the reference genome using GMAP (version 2018-05-30) with ( $\text{min\_fl\_count} = 1$ , other parameters were default value) [101]. Redundant transcripts were collapsed with minimum identity of 95% and minimum coverage of 99%. The obtained isoforms were compared with reference genome (<http://www.spinachbase.org/>, spinach genome sequence\_v1) [33] annotation and classified into three groups: known isoforms (mapped to annotated genes), new isoforms (mapped to different exons of annotated gene), and novel isoforms (mapped to unannotated region of reference genome).

To investigate the functions of new isoforms, we blasted new isoforms against the Nr database (<http://www.ncbi.nlm.nih.gov/>), the Swiss-Prot protein database (<http://www.expasy.ch/sprot/>), and the KEGG database (<http://www.genome.jp/kegg/>) with the BLASTx program (<http://www.ncbi.nlm.nih.gov/BLAST/>, version 2.6.0+) at an E-value threshold of  $1e-5$  to evaluate sequence similarity with genes of other species. GO annotation was analyzed by Blast2GO software (version 2.5.0) [102] with Nr annotation results of isoforms. Isoforms ranking the first 20 highest scores and no shorter than 33 high-scoring segment pair hits were selected to conduct Blast2GO analysis (default parameter). The functional classification of isoforms was performed using the WEGO software (version 2.0, default parameter) [103].

### Structural analyses

Alternative splicing events were identified using the SUPPA tool (version 2.2.1, default parameter) [104]. Seven major types of alternative splicing events, namely, ES, RI, A5, A3, MX, AF, and AL, were classified, counted, and compared between different samples.

Polyadenylation site was analyzed using all high-quality isoforms aligning to transcripts. The end position is the potential alternative polyadenylation site.

CNCI software (version 2) [105] and CPC software (<http://cpc.cbi.pku.edu.cn/>, version 0.9-r2) [106] were used to assess the protein-coding potential of novel and new isoforms by default parameters. Isoforms were mapped to the Swiss-Prot database to assess protein annotation. The intersection of both nonprotein-coding potential results and nonprotein annotation results were defined as lncRNAs. To further annotate lncRNAs at the evolutionary level, we used the software Infernal (<http://eddylab.org/infernal/>, version 1.1.1) [107] in sequence alignment. lncRNAs were classified by secondary structures and sequence conservation.

### Illumina RNA-seq sequencing and bioinformatics analyses

After the total RNA of male and female flowers was extracted by TRIzol reagent (Life technologies), mRNA was enriched by Oligo (dT) beads. The enriched mRNA was fragmented into short fragments using fragmentation

buffer and reverse transcribed into cDNA with random primers. Second-strand cDNA was synthesized by DNA polymerase I, RNase H, dNTP, and buffer. The cDNA fragments were then purified with QiaQuick PCR extraction kit, end repaired, poly(A) added, and ligated to Illumina sequencing adapters. The ligation products were size selected by agarose gel electrophoresis, PCR amplified, and sequenced using Illumina HiSeq™ 2500 by Gene Denovo Biotechnology Co. (Guangzhou, China). After filtering, high-quality clean reads were assembled and mapped to the reference genome (<http://www.spinachbase.org/>) using TopHat2 (version 2.1.1) and Cufflinks (version 2.2.1) [108, 109].

Gene abundances were quantified by RSEM software (version 1.2.19) [110]. Gene expression level was normalized using the method of fragments per kilobase of transcript per million mapped reads. edgeR package (<http://www.r-project.org/>, version 3.12.1) was used to identify differentially expressed genes across samples or groups. Genes with a fold change of  $\geq 2$  and a false discovery rate (FDR) of  $< 0.05$  in a comparison were designed as significant DEGs. DEGs were then subjected to enrichment analysis of GO functions and KEGG pathways.

The GATK (version 3.4–46) was used for calling variants of transcripts, and ANNOVAR (version 42, 401) was used for SNP/InDel annotation. The function, genome site and type of variation of SNPs were also analyzed.

#### Weighted gene co-expression network analysis

Co-expression networks were constructed using the WGCNA (v1.47) package in R (version 3.4.4) [111]. After filtering, gene expression values were imported into WGCNA to construct co-expression modules by using the automatic network construction function blockwiseModules with default settings, except that the power was 6, TOMType was unsigned, mergeCutHeight was 0.15, and minModuleSize was 50. Genes were clustered into 21 correlated modules.

To determine biologically significant modules, we used module eigengenes to calculate the correlation coefficient with samples. The intramodular connectivity (function softConnectivity) of each gene was calculated, and genes with high connectivity tended to be hub genes that may have important functions. The networks were visualized using Cytoscape\_3.3.0.

GO and KEGG pathway enrichment analyses were conducted to determine the biological functions of genes in each module. Significantly enriched GO terms and pathways in genes in a module compared with the background were defined by hypergeometric test and a threshold of FDR of  $< 0.05$ .

#### PCR and qPCR experiments

Total RNA was isolated using the TRIzol reagent (Invitrogen) according to the user's manual. Approximately 1  $\mu\text{g}$  of total RNA was used to synthesize the first-strand cDNA using PrimeScript™ RT reagent Kit with gDNA Eraser (TaKaRa). PCR validation of alternative splicing, novel isoform, and lncRNA was performed by *Trans-Start® FastPfu* Fly DNA Polymerase (TRANSGEN). For alternative polyadenylation validation, we used 3'-Full RACE Core Set with PrimeScript™ RTase for cDNA synthesis and PCR amplification (TaKaRa). PCR product was monitored on 1% agarose gel. To validate the DEGs expression, qPCR was carried out with FastSYBR Mixture (CWBIO) on LightCycler® 480 System (Roche) according to the manufacturer's instructions. UBQ was used as an internal control for normalization. The relative expression level of transcript was calculated with  $2^{-\Delta\Delta C_T}$  method and the result was shown in Additional file 8. All primers sequences are listed in Additional file 9.

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-020-07277-4>.

**Additional file 1.** Flower tissue samples used in transcriptome sequencing.

**Additional file 2.** Validation of alternative splicing and alternative polyadenylation. (a) PCR validation of alternative splicing events (left) and gene structure (right); exons are represented by boxes, introns by lines; "ref" means gene structure identified in reference genome, "Sp" means gene structure identified by Iso-seq; "F" means forward primer, "R" means reverse primer. (b) Validation of polyadenylation sites by 3' RACE PCR (left) and structure of the 3' end (right); "□" represents exon, "—" represents intron, "■" represents 3' UTR; "—A(n)—" represents poly(A) structure; "●" represents the adaptor used in 3' RACE; "F" means forward primer, "R" means reverse primer.

**Additional file 3.** PCR validation of novel genes and lncRNAs. "—" represents intron, "■" represents exon, "F" means forward primer, "R" means reverse primer.

**Additional file 4.** Summary of RNA-seq data. (a) PCA analysis of all samples. (b) The top 20 enriched KEGG pathway. (c) GO Directed Acyclic Graph of DEGs between female and male flower at three early development stages in biological process. (d) GO Directed Acyclic Graph of DEGs between female and male flower at three early development stages in molecular function.

**Additional file 5.** List of sex-biased genes related to transcription factors, hormones, sugar and phenylpropanoid.

**Additional file 6.** Correlation between WGCNA module memberships.

**Additional file 7.** List of genes in female and male network.

**Additional file 8.** Expression validation of the DEGs by qPCR.

**Additional file 9.** Primer list used in this study.

**Additional file 10.** Summary of sex-linked genes with SNPs.

#### Abbreviations

lncRNA: long noncoding RNA; WGCNA: Weighted gene co-expression network analysis; SDR: Sex-determining region; *GAI*: *GIBBERELLIC ACID INSENSITIVE*; *AP3*: *APETALLA3*; *PI*: *PISTILLATA*; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene Ontology; Nr database: NCBI non-redundant protein database; *HXK*: Hexokinase; SE: Skipping exon events; A5: 5' splice sites events; A3: 3' splice sites events; MX: Mutually exclusive exons events;



RI: Retained intron events; AF: Alternative first exons events; AL: Alternative last exons events; GA: Gibberellin; CK: Cytokinin; ABA: Abscisic acid; BR: Brassinosteroid; JA: Jasmonate

#### Acknowledgements

We would like to thank Prof. Ray Ming and Prof. Yongfang Li for their constructive suggestion and revision.

#### Authors' contributions

NL and CD designed experiments; NL, ZM, MT, YW and CD performed experiments and processed the data; NL, YZ, SL, WG and CD discussed the results and revised the manuscript; NL and CD wrote the manuscript with comments from the authors. The authors read and approved the final manuscript.

#### Funding

This work was supported by grants from the National Natural Science foundation of China (31770346, 31970240 and 32000243), and the key scientific research project plan for Henan provincial colleges and universities (20A180016). The funding bodies have no roles in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

#### Availability of data and materials

Sequencing data that support the findings of this study have been deposited in the NCBI SRA database [<https://www.ncbi.nlm.nih.gov/sra/PRJNA607896>]. All data generated or analyzed during this study are included in this published article.

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

Received: 11 August 2020 Accepted: 24 November 2020

Published online: 01 December 2020

#### References

- Renner SS. The relative and absolute frequencies of angiosperm sexual systems: dioecy, monoecy, gynodioecy, and an updated online database. *Am J Bot.* 2014;101(10):1588–96.
- Charlesworth B, Charlesworth D. A model for the evolution of Dioecy and Gynodioecy. *Am Nat.* 1978;112(988):975–97.
- Renner SS, Ricklefs RE. Dioecy and its correlates in the flowering plants. *Am J Bot.* 1995;82(5):596–606.
- Barrett SC. The evolution of plant sexual diversity. *Nat Rev Genet.* 2002;3(4):274–84.
- Charlesworth D. Plant contributions to our understanding of sex chromosome evolution. *New Phytol.* 2015;208(1):52–65.
- Charlesworth D. Evolution of recombination rates between sex chromosomes. *Philos Trans R Soc Lond Ser B Biol Sci.* 2017;372(1736):20160456.
- Charlesworth D. Does sexual dimorphism in plants promote sex chromosome evolution? *Environ Exp Bot.* 2018;146:5–12.
- Henry IM, Akagi T, Tao R, Comai L. One hundred ways to invent the sexes: theoretical and observed paths to Dioecy in plants. *Annu Rev Plant Biol.* 2018;69:553–75.
- Feng G, Sanderson BJ, Keefeovering K, Liu J, Ma T, Yin T, Smart LB, Difazio SP, Olson MS. Pathways to sex determination in plants: how many roads lead to Rome? *Curr Opin Plant Biol.* 2020;54:61–8.
- Akagi T, Henry IM, Tao R, Comai L. A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. *Science.* 2014;346(6209):646–50.
- Akagi T, Shirasawa K, Nagasaki H, Hirakawa H, Tao R, Comai L, Henry IM. The persimmon genome reveals clues to the evolution of a lineage-specific sex determination system in plants. *PLoS Genet.* 2020;16(2):e1008566.
- Harkess A, Zhou JS, Xu CY, Bowers JE, Van der Hulst R, Ayyampalayam S, Mercati F, Riccardi P, McKain MR, Kakrana A, et al. The asparagus genome sheds light on the origin and evolution of a young Y chromosome. *Nat Commun.* 2017;8(1):1279.
- Harkess A, Huang K, van der Hulst R, Tissen B, Caplan JL, Koppula A, Batish M, Meyers BC, Leebens-Mack J. Sex determination by two Y-linked genes in garden Asparagus. *Plant Cell.* 2020;32(6):1790–6.
- Torres MF, Mathew LS, Ahmed I, Al-Azwani IK, Krueger R, Rivera-Nunez D, Mohamoud YA, Clark AG, Suhre K, Malek JA. Genus-wide sequencing supports a two-locus model for sex-determination in Phoenix. *Nat Commun.* 2018;9(1):3969.
- Akagi T, Pilkington SM, Varkonyi-Gasic E, Henry IM, Sugano SS, Sonoda M, Firl A, McNeillage MA, Douglas MJ, Wang TC, et al. Two Y-chromosome-encoded genes determine sex in kiwifruit. *Nature Plants.* 2019;5(8):801–9.
- Picq S, Santoni S, Lacombe T, Latreille M, Weber A, Ardisson M, Ivorra S, Maghradze D, Arroyo-Garcia R, Chatelet P, et al. A small XY chromosomal region explains sex determination in wild dioecious *V. vinifera* and the reversal to hermaphroditism in domesticated grapevines. *BMC Plant Biol.* 2014;14:229.
- Fechter I, Hausmann L, Daum M, Sorensen TR, Viehover P, Weisshaar B, Topfer R. Candidate genes within a 143 kb region of the flower sex locus in *Vitis*. *Mol Genetics Genomics.* 2012;287(3):247–59.
- Mori K, Shirasawa K, Nogata H, Hirata C, Tashiro K, Habu T, Kim S, Himeno S, Kuhara S, Ikegami H. Identification of RANI orthologue associated with sex determination through whole genome sequencing analysis in fig (*Ficus carica* L.). *Sci Rep.* 2017;7:41124.
- Geraldes A, Hefer CA, Capron A, Kolosova N, Martinez-Nunez F, Soolanayakanahally RY, Stanton B, Guy RD, Mansfield SD, Douglas CJ, et al. Recent Y chromosome divergence despite ancient origin of dioecy in poplars (*Populus*). *Mol Ecol.* 2015;24(13):3243–56.
- Tennessen JA, Wei N, Straub S, Govindarajulu R, Liston A, Ashman TL. Repeated translocation of a gene cassette drives sex-chromosome turnover in strawberries. *PLoS Biol.* 2018;16(8):e2006062.
- Diggle PK, Di Stilio VS, Gschwend AR, Golenberg EM, Moore RC, Russell JR, Sinclair JP. Multiple developmental processes underlie sex differentiation in angiosperms. *Trends Genet.* 2011;27(9):368–76.
- Sobral R, Silva HG, Morais-Cecilio L, Costa MMR. The quest for molecular regulation underlying unisexual flower development. *Front Plant Sci.* 2016;7:160.
- Khattak JZK, Torp AM, Andersen SB. A genetic linkage map of *Spinacia oleracea* and localization of a sex determination locus. *Euphytica.* 2006;148(3):311–8.
- Onodera Y, Yonaha I, Masumo H, Tanaka A, Niikura S, Yamazaki S, Mikami T. Mapping of the genes for dioecism and monoecism in *Spinacia oleracea* L.: evidence that both genes are closely linked. *Plant Cell Rep.* 2011;30(6):965–71.
- Avşar B. Genetic diversity of Turkish spinach cultivars (*Spinacia oleracea* L.). Izmir: Izmir Institute of Technology; 2011.
- Deng C-I, R-y Q, Cao Y, Gao J, Li S-f, Gao W-j, Lu L-d. Microdissection and painting of the Y chromosome in spinach (*Spinacia oleracea*). *J Plant Res.* 2013;126(4):549–56.
- Okazaki Y, Takahata S, Hirakawa H, Suzuki Y, Onodera Y. Molecular evidence for recent divergence of X- and Y-linked gene pairs in *Spinacia oleracea* L. *PLoS One.* 2019;14(4):e0214949.
- Wadlington WH, Ming R. Development of an X-specific marker and identification of YY individuals in spinach. *Theor Appl Genet.* 2018;131(9):1987–94.
- Deng C, Qin R, Gao J, Cao Y, Li S, Gao W, Lu L. Identification of sex chromosome of spinach by physical mapping of 45s rDNAs by FISH. *Caryologia.* 2012;65(4):322–7.
- Kudoh T, Takahashi M, Osabe T, Toyoda A, Hirakawa H, Suzuki Y, Ohmido N, Onodera Y. Molecular insights into the non-recombining nature of the spinach male-determining region. *Mol Gen Genomics.* 2017;293(2):557–68.
- Takahata S, Yago T, Iwabuchi K, Hirakawa H, Suzuki Y, Onodera Y. Comparison of spinach sex chromosomes with sugar beet autosomes reveals extensive Synteny and low recombination at the male-determining locus. *J Hered.* 2016;107(7):679–85.
- Qian W, Fan GY, Liu DD, Zhang HL, Wang XW, Wu J, Xu ZS. Construction of a high-density genetic map and the X/Y sex-determining gene mapping in spinach based on large-scale markers developed by specific-locus amplified fragment sequencing (SLAF-seq). *BMC Genomics.* 2017;18(1):276.
- Xu C, Jiao C, Sun H, Cai X, Wang X, Ge C, Zheng Y, Liu W, Sun X, Xu Y, et al. Draft genome of spinach and transcriptome diversity of 120 *Spinacia* accessions. *Nat Commun.* 2017;8:15275.

34. Yu L, Ma XK, Deng B, Yue JJ, Ming R. Construction of high-density genetic maps defined sex determination region of the Y chromosome in spinach. *Mol Gen Genomics*. 2020. <https://doi.org/10.1007/s00438-020-01723-4>.
35. Pfent C, Pobursky KJ, Sather DN, Golenberg EM. Characterization of SpAPETALA3 and SpPSTILLATA, B class floral identity genes in *Spinacia oleracea*, and their relationship to sexual dimorphism. *Dev Genes Evol*. 2005; 215(3):132–42.
36. Sather DN, Jovanovic M, Golenberg EM. Functional analysis of B and C class floral organ genes in spinach demonstrates their role in sexual dimorphism. *BMC Plant Biol*. 2010;10:46.
37. Chailakhyan MK, Khryanin VN. The role of leaves in sex expression in hemp and spinach. *Planta*. 1979;144(2):205–7.
38. El-Gizawy A, El-Oksh I, Sharaf A, El-Habar M. Effect of gibberellic acid and alar on flowering and seed yield of spinach. *Egypt J Horticulture (Egypt)*. 1992;19(2):191–200.
39. West NW, Golenberg EM. Gender-specific expression of GIBBERELLIC ACID INSENSITIVE is critical for unisexual organ initiation in dioecious *Spinacia oleracea*. *New Phytol*. 2018;217(3):1322–34.
40. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y. A tale of three next generation sequencing platforms: comparison of ion torrent, Pacific biosciences and Illumina MiSeq sequencers. *BMC Genomics*. 2012;13(1):341.
41. Sharon D, Tilgner H, Grubert F, Snyder M. A single-molecule long-read survey of the human transcriptome. *Nat Biotechnol*. 2013;31(11):1009–14.
42. Elkon R, Ugalde AP, Agami R. Alternative cleavage and polyadenylation: extent, regulation and function. *Nat Rev Genet*. 2013;14:496.
43. Rodriguez-Cazorla E, Ripoll JJ, Andujar A, Bailey LJ, Martinez-Laborda A, Yanofsky MF, Vera A. K-homology nuclear ribonucleoproteins regulate floral organ identity and determinacy in *Arabidopsis*. *PLoS Genet*. 2015;11(2): e1004983.
44. Deng X, Cao X. Roles of pre-mRNA splicing and polyadenylation in plant development. *Curr Opin Plant Biol*. 2017;35:45–53.
45. Abdel-Ghany SE, Hamilton M, Jacobi JL, Ngam P, Devitt N, Schilkey F, Ben-Hur A, Reddy ASN. A survey of the sorghum transcriptome using single-molecule long reads. *Nat Commun*. 2016;7:11706.
46. Harkess A, Mercati F, Shan H-Y, Sunseri F, Falavigna A, Leebens-Mack J. Sex-biased gene expression in dioecious garden asparagus (*Asparagus officinalis*). *New Phytol*. 2015;207(3):883–92.
47. Kriechbaumer V, Wang P, Hawes C, Abell BM. Alternative splicing of the auxin biosynthesis gene YUCCA4 determines its subcellular compartmentation. *Plant J*. 2012;70(2):292–302.
48. Jiang J, Zhang C, Wang X. A recently evolved isoform of the transcription factor BES1 promotes brassinosteroid signaling and development in *Arabidopsis thaliana*. *Plant Cell*. 2015;27(2):361–74.
49. Wang Z, Ji H, Yuan B, Wang S, Su C, Yao B, Zhao H, Li X. ABA signalling is fine-tuned by antagonistic HAB1 variants. *Nat Commun*. 2015;6(1):8138.
50. Mandadi KK, Scholthof KB. Genome-wide analysis of alternative splicing landscapes modulated during plant-virus interactions in *Brachypodium distachyon*. *Plant Cell*. 2015;27(1):71–85.
51. Godoy Herz MA, Kubaczka MG, Brzyżek G, Servi L, Krzyszton M, Simpson C, Brown J, Swiezewski S, Petrillo E, Kornblihtt AR. Light Regulates Plant Alternative Splicing through the Control of Transcriptional Elongation. *Mol Cell*. 2019;73(5):1066–74 e1063.
52. Srivastava AK, Lu Y, Zinta G, Lang Z, Zhu J-K. UTR-dependent control of gene expression in plants. *Trends Plant Sci*. 2018;23(3):248–59.
53. Quesada V, Macknight R, Dean C, Simpson GG. Autoregulation of FCA pre-mRNA processing controls *Arabidopsis* flowering time. *EMBO J*. 2003;22(12): 3142–52.
54. Moreira A. Integrating transcription kinetics with alternative polyadenylation and cell cycle control. *Nucleus*. 2011;2(6):556–61.
55. Batista Pedro J, Chang Howard Y. Long noncoding RNAs: cellular address codes in development and disease. *Cell*. 2013;152(6):1298–307.
56. Bardou F, Ariel F, Simpson Craig G, Romero-Barrios N, Laporte P, Balzergue S, Brown John WS, Crespi M. Long noncoding RNA modulates alternative splicing regulators in *Arabidopsis*. *Dev Cell*. 2014;30(2):166–76.
57. Ding J, Lu Q, Ouyang Y, Mao H, Zhang P, Yao J, Xu C, Li X, Xiao J, Zhang Q. A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc Natl Acad Sci*. 2012;109(7):2654.
58. Rolland F, Baena-Gonzalez E, Sheen J. SUGAR SENSING AND SIGNALING IN PLANTS: conserved and novel mechanisms. *Annu Rev Plant Biol*. 2006;57(1): 675–709.
59. Sawicki M, Jacquens L, Baillieul F, Clément C, Vaillant-Gaveau N, Jacquard C. Distinct regulation in inflorescence carbohydrate metabolism according to grapevine cultivars during floral development. *Physiol Plant*. 2015;154(3): 447–67.
60. Miao M, Yang X, Han X, Wang K. Sugar signalling is involved in the sex expression response of monoecious cucumber to low temperature. *J Exp Bot*. 2010;62(2):797–804.
61. Satoh-Nagasawa N, Nagasawa N, Malcomber S, Sakai H, Jackson D. A trehalose metabolic enzyme controls inflorescence architecture in maize. *Nature*. 2006;441(7090):227–30.
62. van der Meer IM, Stam ME, van Tunen AJ, Mol JN, Stuitje AR. Antisense inhibition of flavonoid biosynthesis in petunia anthers results in male sterility. *Plant Cell*. 1992;4(3):253–62.
63. Matsuda N, Tsuchiya T, Kishitani S, Tanaka Y, Toriyama K. Partial male sterility in transgenic tobacco carrying antisense and sense PAL cDNA under the control of a tapetum-specific promoter. *Plant Cell Physiol*. 1996;37(2):215–22.
64. Dobritsa AA, Lei ZT, Nishikawa S, Urbanczyk-Wochniak E, Huhman DV, Preuss D, Sumner LW. LAP5 and LAP6 encode anther-specific proteins with similarity to Chalcone synthase essential for pollen Exine development in *Arabidopsis*. *Plant Physiol*. 2010;153(3):937–55.
65. Tang LK, Chu H, Yip WK, Yeung EC, Lo C. An anther-specific dihydroflavonol 4-reductase-like gene (DRL1) is essential for male fertility in *Arabidopsis*. *New Phytol*. 2009;181(3):576–87.
66. Grienenberger E, Kim SS, Lallemand B, Geoffroy P, Heintz D, Souza CD, Heitz T, Douglas CJ, Legrand M. Analysis of TETRAKETIDE alpha-PYRONE REDUCTASE function in *Arabidopsis thaliana* reveals a previously unknown, but conserved, biochemical pathway in Sporopollenin monomer biosynthesis. *Plant Cell*. 2010;22(12):4067–83.
67. Wang K, Guo ZL, Zhou WT, Zhang C, Zhang ZY, Lou Y, Xiong SX, Yao XZ, Fan JJ, Zhu J, et al. The regulation of Sporopollenin biosynthesis genes for rapid Pollen Wall formation. *Plant Physiol*. 2018;178(1):283–94.
68. Song S, Qi T, Huang H, Xie D. Regulation of stamen development by coordinated actions of jasmonate, auxin, and gibberellin in *Arabidopsis*. *Mol Plant*. 2013;6(4):1065–73.
69. Yuan Z, Zhang D. Roles of jasmonate signalling in plant inflorescence and flower development. *Curr Opin Plant Biol*. 2015;27:44–51.
70. Hartwig T, Chuck GS, Fujioka S, Klempien A, Weizbauer R, Potluri DP, Choe S, Johal GS, Schulz B. Brassinosteroid control of sex determination in maize. *Proc Natl Acad Sci U S A*. 2011;108(49):19814–9.
71. Boualem A, Troadec C, Camps C, Lemhemi A, Morin H, Sari M-A, Fraenkel-Zagouri R, Kovalski I, Dogimont C, Perl-Treves R, et al. A cucurbit androecy gene reveals how unisexual flowers develop and dioecy emerges. *Science*. 2015;350(6261):688.
72. Peterson CE, Anhder LD. Induction of staminate flowers on Gynoecious cucumbers with gibberellin A3. *Science*. 1960;131(3414):1673.
73. Atsmon D, Lang A, Light EN. Contents and recovery of gibberellins in Monoecious and Gynoecious cucumber plants. *Plant Physiol*. 1968;43(5):806.
74. Fendrych M, Van Hautegeem T, Van Durme M, Olvera-Carrillo Y, Huysmans M, Karimi M, Lippens S, Guérin Christopher J, Krebs M, Schumacher K, et al. Programmed cell death controlled by ANAC033/SOMBRERO determines root cap organ size in *Arabidopsis*. *Curr Biol*. 2014;24(9):931–40.
75. Alvarado VY, Tag A, Thomas TL. A cis regulatory element in the TAPNAC promoter directs tapetal gene expression. *Plant Mol Biol*. 2011;75(1):129–39.
76. Wellmer F, Riechmann JL, Alves-Ferreira M, Meyerowitz EM. Genome-wide analysis of spatial gene expression in *Arabidopsis* flowers. *Plant Cell*. 2004; 16(5):1314.
77. Sorensen A-M, Kröber S, Unte US, Huijser P, Dekker K, Saedler H. The *Arabidopsis* ABORTED MICROSPORES (AMS) gene encodes a MYC class transcription factor. *Plant J*. 2003;33(2):413–23.
78. Lu P, Chai M, Yang J, Ning G, Wang G, Ma H. The *Arabidopsis* CALLOSE DEFECTIVE MICROSPORO1 gene is required for male fertility through regulating Callose metabolism during Microsporogenesis. *Plant Physiol*. 2014;164(4):1893.
79. Sagasser M, Lu G-H, Hahlbrock K, Weisshaar B. A *thaliana* TRANSPARENT TESTA 1 is involved in seed coat development and defines the WIP subfamily of plant zinc finger proteins. *Genes Dev*. 2002;16(1):138–49.
80. Prát T, Hajný J, Grunewald W, Vasileva M, Molnár G, Tejos R, Schmid M, Sauer M, Friml J. WRKY23 is a component of the transcriptional network mediating auxin feedback on PIN polarity. *PLoS Genet*. 2018;14(1):e1007177.
81. Singh D, Debnath P, Roohi, Sane AP, Sane VA. Expression of the tomato WRKY gene, SIWRKY23, alters root sensitivity to ethylene, auxin and JA and

- affects aerial architecture in transgenic Arabidopsis. *Physiol Mol Biol Plants*. 2020;26(6):1187–99.
82. Roussin-Léveillé C, Silva-Martins G, Moffett P. ARGONAUTE5 represses age-dependent induction of flowering through physical and functional interaction with miR156 in Arabidopsis. *Plant Cell Physiol*. 2020;61(5):957–66.
  83. Yang Y, Xu R, Ma C-j, Vlot AC, Klessig DF, Pichersky E. Inactive methyl Indole-3-acetic acid Ester can be hydrolyzed and activated by several Esterases belonging to the AtMES esterase family of Arabidopsis. *Plant Physiol*. 2008;147(3):1034.
  84. Seifert GJ, Barber C, Wells B, Roberts K. Growth regulators and the control of nucleotide sugar flux. *Plant Cell*. 2004;16(3):723.
  85. Cai X, Ballif J, Endo S, Davis E, Liang M, Chen D, DeWald D, Kreps J, Zhu T, Wu Y. A putative CCAAT-binding transcription factor is a regulator of flowering timing in Arabidopsis. *Plant Physiol*. 2007;145(1):98.
  86. Covington MF, Harmer SL. The circadian clock regulates Auxin signaling and responses in Arabidopsis. *PLoS Biol*. 2007;5(8):e222.
  87. Arana MV, Marin-de la Rosa N, Maloof JN, Blazquez MA, Alabadi D. Circadian oscillation of gibberellin signaling in Arabidopsis. *Proc Natl Acad Sci U S A*. 2011;108(22):9292–7.
  88. Li Y, Van den Ende W, Rolland F. Sucrose induction of anthocyanin biosynthesis is mediated by DELLA. *Mol Plant*. 2014;7(3):570–2.
  89. Teng S, Keurentjes J, Bentsink L, Koornneef M, Smeekens S. Sucrose-specific induction of anthocyanin biosynthesis in Arabidopsis requires the MYB75/PAP1 gene. *Plant Physiol*. 2005;139(4):1840–52.
  90. Solfanelli C, Poggi A, Loreti E, Alpi A, Perata P. Sucrose-specific induction of the anthocyanin biosynthetic pathway in Arabidopsis. *Plant Physiol*. 2006;140(2):637–46.
  91. Loreti E, Povero G, Novi G, Solfanelli C, Alpi A, Perata P. Gibberellins, jasmonate and abscisic acid modulate the sucrose-induced expression of anthocyanin biosynthetic genes in Arabidopsis. *New Phytol*. 2008;179(4):1004–16.
  92. Ljung K, Nemhauser JL, Perata P. New mechanistic links between sugar and hormone signalling networks. *Curr Opin Plant Biol*. 2015;25:130–7.
  93. Li QF, Wang CM, Jiang L, Li S, Sun SSM, He JX. An interaction between BZR1 and DELLAs mediates direct signaling crosstalk between brassinosteroids and gibberellins in Arabidopsis. *Sci Signal*. 2012;5(244):ra72.
  94. Murmu J, Bush MJ, DeLong C, Li ST, Xu ML, Khan M, Malcolmsen C, Fobert PR, Zachgo S, Hepworth SR. Arabidopsis basic Leucine-zipper transcription factors TGA9 and TGA10 interact with floral Glutaredoxins ROXY1 and ROXY2 and are redundantly required for anther development. *Plant Physiol*. 2010;154(3):1492–504.
  95. Li S, Lauri A, Ziemann M, Busch A, Bhave M, Zachgo S. Nuclear activity of ROXY1, a Glutaredoxin interacting with TGA factors, is required for petal development in Arabidopsis thaliana. *Plant Cell*. 2009;21(2):429–41.
  96. O'Maoileidigh DS, Wuest SE, Rae L, Raganelli A, Ryan PT, Kwasniewska K, Das P, Lohan AJ, Loftus B, Graciet E, et al. Control of reproductive floral organ identity specification in Arabidopsis by the C function regulator AGAMOUS. *Plant Cell*. 2013;25(7):2482–503.
  97. Wuest SE, O'Maoileidigh DS, Rae L, Kwasniewska K, Raganelli A, Hanczaryk K, Lohan AJ, Loftus B, Graciet E, Wellmer F. Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA. *Proc Natl Acad Sci U S A*. 2012;109(33):13452–7.
  98. Zhang Y, Liu B, Yang S, An J, Chen C, Zhang X, Ren H. A cucumber DELLA homolog CsGAIP may inhibit staminate development through transcriptional repression of B class floral homeotic genes. *PLoS One*. 2014;9(3):e91804.
  99. Sherry RA, Eckard KJ, Lord EM. Flower development in Dioecious Spinacia oleracea (Chenopodiaceae). *Am J Bot*. 1993;80(3):283–91.
  100. Gordon SP, Tseng E, Salamov A, Zhang JW, Meng XD, Zhao ZY, Kang DW, Underwood J, Grigoriev IV, Figueroa M, et al. Widespread polycistronic transcripts in fungi revealed by single-molecule mRNA sequencing. *PLoS One*. 2015;10(7):e0132628.
  101. Wu TD, Watanabe CK. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics*. 2005;21(9):1859–75.
  102. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21(18):3674–6.
  103. Ye J, Fang L, Zheng HK, Zhang Y, Chen J, Zhang ZJ, Wang J, Li ST, Li RQ, Bolund L, et al. WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res*. 2006;34:W293–7.
  104. Alamancos GP, Pages A, Trincado JL, Bellora N, Eyras E. Leveraging transcript quantification for fast computation of alternative splicing profiles. *RNA*. 2015;21(9):1521–31.
  105. Sun L, Luo HT, Bu DC, Zhao GG, Yu KT, Zhang CH, Liu YN, Chen RS, Zhao Y. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res*. 2013;41(17):e166.
  106. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res*. 2007;35:W345–9.
  107. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*. 2013;29(22):2933–5.
  108. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.
  109. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks. *Nat Protoc*. 2012;7(3):562–78.
  110. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323.
  111. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

