

Sex Determination in *Ceratopteris richardii* Is Accompanied by Transcriptome Changes That Drive Epigenetic Reprogramming of the Young Gametophyte

Nadia M. Atallah,* Olga Vitek,[†] Federico Gaiti,[‡] Milos Tanurdzic,^{§,1} and Jo Ann Banks**,¹

*Purdue University Center for Cancer Research and **Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907, [†]College of Computer and Information Science, Northeastern University, Boston, MA 02115,

[‡]Department of Medicine, Weill Cornell Medicine and New York Genome Center, New York, NY 10013, and [§]School of Biological Sciences, The University of Queensland, St. Lucia, Australia

ORCID IDs: 0000-0002-1819-7070 (N.M.A.); 0000-0002-7762-7261 (J.A.B.)

ABSTRACT The fern *Ceratopteris richardii* is an important model for studies of sex determination and gamete differentiation in homosporous plants. Here we use RNA-seq to *de novo* assemble a transcriptome and identify genes differentially expressed in young gametophytes as their sex is determined by the presence or absence of the male-inducing pheromone called antheridiogen. Of the 1,163 consensus differentially expressed genes identified, the vast majority (1,030) are up-regulated in gametophytes treated with antheridiogen. GO term enrichment analyses of these DEGs reveals that a large number of genes involved in epigenetic reprogramming of the gametophyte genome are up-regulated by the pheromone. Additional hormone response and development genes are also up-regulated by the pheromone. This *C. richardii* gametophyte transcriptome and gene expression dataset will prove useful for studies focusing on sex determination and differentiation in plants.

KEYWORDS

sex
determination
RNA-seq
Ceratopteris
gametophyte
epigenetics
gibberellin
antheridiogen
transcriptome
Genetics of Sex

Ceratopteris richardii is a homosporous fern that produces a single type of haploid spore, with each spore having the potential to develop as a free-living male or hermaphroditic gametophyte. In this and many other fern species, the sex of the gametophyte is determined by a male-inducing pheromone called antheridiogen (Warne and Hickok 1991; Banks 1999). In the absence of A_{CE} (for Antheridiogen

Ceratopteris), a spore develops as a hermaphrodite, which begins to secrete biologically detectable amounts of A_{CE} after it loses the competence to respond to its male-inducing effects. In the presence of A_{CE} , a spore develops as a male gametophyte. Thus, in a population, spores that germinate first in the absence of A_{CE} develop as hermaphrodites that secrete A_{CE} , while spores that germinate later and in the presence of A_{CE} develop as males (Banks *et al.* 1993; Warne and Hickok 1991). Given that the self-fertilization of a hermaphroditic gametophyte results in a completely homozygous sporophyte (similar to a double haploid), this mechanism of sex determination is presumed to promote outcrossing by increasing the proportion of males in a population of gametophytes (Hauffer 2002).

Although small (<3mm), male and hermaphroditic gametophytes are dimorphic and easy to distinguish by size and shape. Each hermaphrodite forms a multicellular, lateral meristem that contributes to its heart-shaped appearance, with multiple archegonia developing after the lateral meristem forms (Figure 1g). The development of this lateral meristem coincides with the loss of competence to respond to A_{CE} as well as the production of A_{CE} in the hermaphrodite (Banks *et al.* 1993).

Copyright © 2018 Atallah *et al.*

doi: <https://doi.org/10.1534/g3.118.200292>

Manuscript received February 14, 2018; accepted for publication April 18, 2018; published Early Online May 2, 2018.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material available at Figshare: <https://doi.org/10.25387/g3.6100139>.

¹Corresponding authors: Department of Botany and Plant Pathology, 915 W. State St., Purdue University, West Lafayette, IN 47906. E-mail: banksj@purdue.edu; School of Biological Sciences, The University of Queensland, St Lucia, Australia. E-mail: m.tanurdzic@uq.edu.au

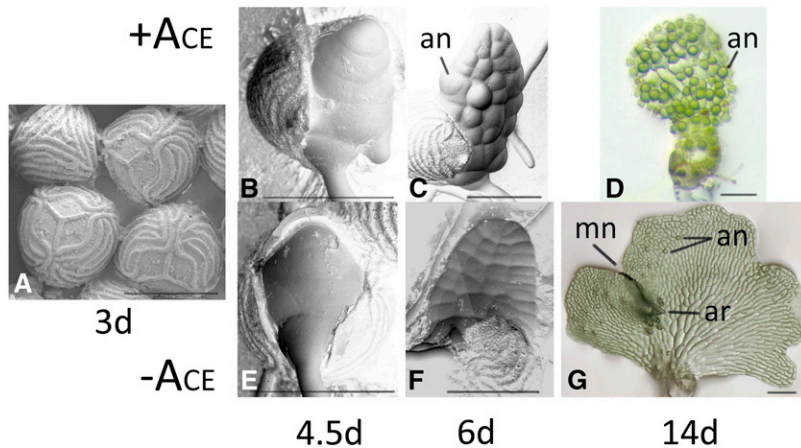


Figure 1 *Ceratopteris* gametophyte development. (a) SEM of spores three days after inoculation showing trilete markings. (b-d) SEMs of 4.5d, 6d and 14d gametophytes grown in the presence of A_{CE} . (e-g) SEMs of 4.5d, 6d and 14d gametophytes grown in the absence of A_{CE} . The mature hermaphrodite (g) has a meristem notch (mn), archegonia (ar) and antheridia (an) while the mature male (d) has only antheridia (an). Bars = 0.15mm.

Male gametophytes never develop a lateral meristem and are much smaller than hermaphrodites (Figure 1d), with nearly all cells of the male gametophyte terminally differentiating as antheridia. If removed from media containing A_{CE} , undifferentiated cells of a male can form a new hermaphroditic prothallus. Based upon these observations, A_{CE} has many functions in gametophyte development: it prevents the establishment of the lateral meristem; it promotes the rapid differentiation of antheridia; it prevents its own synthesis or secretion in the male; and is necessary to maintain the male program of expression.

To date, all fern antheridiogens characterized are gibberellins (GAs) (Yamane *et al.* 1987; Yamane 1998; Takeno *et al.* 1989; Furber *et al.* 1989). Although the structure of A_{CE} is unknown, the GA biosynthetic inhibitors ancymidol, AMO-1618, and uniconazole-P reduce the proportion of males in a population of *Ceratopteris* gametophytes suggesting that A_{CE} and GA share a common biosynthetic pathway (Warne and Hickok 1989). That ABA completely blocks the A_{CE} response in *Ceratopteris* is also consistent with A_{CE} being a GA (Hickok 1983; McAdam *et al.* 2016).

To understand how A_{CE} determines the sex of the *Ceratopteris* gametophyte, mutations affecting sexual phenotype have been characterized and used to develop a genetic model of the sex-determining pathway (Strain *et al.* 2001; Banks 1994, 1997; Eberle and Banks 1996). Cloning these genes is challenging because of the large genome size of *C. richardii*, ca. 11Gb (Li *et al.* 2015), and the lack of a reference genome sequence for this or any other homosporous fern species. An alternative approach to identifying sex-determining genes involves *de novo* transcriptome assembly using RNA-seq, which provides a means to perform sensitive gene expression studies in organisms that do not have a reference genome (Grabherr *et al.* 2011; Robertson *et al.* 2010; Schulz *et al.* 2012). Here we describe the *de novo* assembly of the transcriptome of young *Ceratopteris* gametophytes and identify genes whose expression differs as their sex is being determined by the absence or presence of A_{CE} , thus providing a snapshot of the transcriptional changes that occur as the sex of the spore becomes determined and prior to the differentiation of male or female traits in the developing gametophyte.

MATERIALS AND METHODS

Plants and Growth Conditions

The origins of Hn-n, an isogenic, wild-type strain of *Ceratopteris richardii* used in this study, is described in (Hickok *et al.* 1987). The conditions for spore sterilization and gametophyte culture are as previously described (Banks 1994). Medium used to culture gametophytes

in the absence of exogenous A_{CE} is as described in (Banks *et al.* 1993) and is referred to as fern medium, or FM. A_{CE} was obtained as a crude aqueous filtrate from media previously supporting gametophyte growth in FM as described in (Banks *et al.* 1993) and is referred to as conditioned FM (CFM). Scanning electron micrographs (SEMs) were performed on a FEI NOVA nanoSEM on samples prepared as previously described (Banks 1994).

For both RNA-seq and qRT-PCR, spores were grown aseptically in liquid FM at 28° in a growth chamber, shaken at 100 rpm, and at a density of 1g spores/L. Three days after spore inoculation, gametophytes were filtered from media; 1/6 of the spores were added to each of three flasks containing 200 mL sterile FM, which is the $-A_{CE}$ treatment, and 1/6 were added to each of three flasks containing 200 mL sterile CFM, which is the $+A_{CE}$ treatment. After 36 hr, gametophytes were vacuum filtered from media and frozen in $N_2(l)$. Tissue was subsequently stored at -80°. All samples were randomized throughout incubators and during sample preparation and harvesting protocols.

Library Preparation and Sequencing

Frozen tissue was ground under liquid nitrogen for 20 min and total RNA extracted using the RNeasy Plant Mini Kit (Qiagen, CA). The TruSeq kit (Illumina, CA) was used to select poly-adenylated mRNA and prepare libraries for sequencing. Libraries were sequenced on an Illumina HiSeq2000 platform using paired-end technology.

Transcriptome Assembly and Annotation

DeconSeq v.0.4.1 was run on each of the FASTQ read files to remove reads aligning to bacterial, viral, rRNA, mitochondrial RNA, and chloroplast DNA (Schmieder *et al.* 2011; Schmieder and Edwards 2011). After removing adapter sequences and trimming reads based on quality score with Trimmomatic v.0.22 (Lohse *et al.* 2012), reads were assembled using the *de novo* transcriptome assembler Trinity (Grabherr *et al.* 2011), with a minimum contig length cutoff of 150 and a fixed k-mer size of 25. An assembly with unique genes was generated by selecting the longest component from each Trinity de Bruijn graph. These were used in subsequent differential expression analyses in order to avoid biasing analyses toward genes that were more difficult to assemble and thus had many more contigs (subcomponents). The program Assembly Stats in the iPlant Discovery environment was utilized to obtain basic assembly statistics (Goff *et al.* 2011; Earl *et al.* 2011). Protein-encoding, differentially expressed genes were annotated using the Trinotate workflow (Ashburner *et al.* 2000;

Finn *et al.* 2011; Grabherr *et al.* 2011; Kanehisa *et al.* 2011) using a 50 amino acid minimum cutoff.

Differential Expression Analysis

Paired reads were aligned to the assembled transcriptome using RSEM v.1.0.1 (Li and Dewey 2011; Grabherr *et al.* 2011; Li *et al.* 2015). Only the transcripts with at least one read aligned in at least one of six samples were used. edgeR v.3.0.8 (Robinson *et al.* 2010), DESeq v.1.10.1 (Anders and Huber 2010), and EBseq v.1.1.4 (Leng *et al.* 2013) were used to identify differentially expressed genes at a Benjamini-Hochberg (Benjamini *et al.* 2001) corrected FDR of $q = 0.01$. In edgeR, dispersion was estimated as tagwise dispersion. To retain as much rigor in our methods as possible, genes that were identified as statistically significantly differentially expressed in all three packages and displayed at least a twofold expression difference between conditions were identified as “consensus DEGs” and used in all downstream analyses.

GO Enrichment and Assembly Validation

Because there is no reference genome sequence for *C. richardii*, GO enrichment was performed by annotating the *C. richardii* transcriptome and the differentially expressed genes against the *Arabidopsis* proteome (Araport11) (Hanlon *et al.* 2015), the non-redundant database, and the *Selaginella moellendorffii* proteome v. 1.0 (Banks *et al.* 2011), using BLASTx and an e-value threshold of 10^{-8} . Gene Ontology (GO) terms were then assigned from the *Arabidopsis* accession identifier from the best hit associated with the differentially expressed transcripts and the reference *C. richardii* transcriptome. ClueGO (version 2.3.4), a Cytoscape (version 3.5.1) plug-in (Cline *et al.* 2007; Smoot *et al.* 2010; Saito *et al.* 2012; Shannon *et al.* 2003), and GO Term Fusion were used to distill and visualize the GO term enrichments within the biological processes category using default parameters, with the following exceptions: the minimum number of genes/cluster was set to 5, the Benjamini-Hochberg method was used to correct the p-values for multiple testing, with a significance threshold of $P < 0.05$ and a custom background model supplied. The GO terms mapping to the entire non-redundant *Ceratopteris richardii* transcriptome was used as the background when assessing the enrichment of GO terms. To assess the quality of the *C. richardii* Trinity assembly, the 5133 *C. richardii* Sanger-generated ESTs available in GenBank were used to blast the entire *Ceratopteris* transcriptome assembly using BLASTn and a BUSCO (Benchmarking Universal Single-Copy Orthologs) analysis was performed using BUSCO v.2.0 to assess completeness of the assembled transcriptome using the ‘eukaryotic’ dataset, which consists of 303 highly conserved genes (Simão *et al.* 2015; Waterhouse *et al.* 2018).

Expression Analysis Validation

Total RNA was isolated from six gametophyte populations cultured and harvested in the same manner as that used to generate the RNA-seq data. RNA was reverse transcribed using the Tetro cDNA Synthesis Kit (Bioline, MA); qRT-PCR was performed using the SYBR green PCR Master Mix (Applied Biosystems), 3ng cDNA template and the StepOne Real-Time PCR System (Applied Biosystems, NY). PCR conditions were: one cycle of 20 min at 95°, 40 cycles of 3 sec at 95° and 30 sec at 60°. Melt curves were analyzed and only those reactions producing a single T_m peak were used. Three technical replicates were performed for each sample. Measurements were normalized to the amount of *CrEF1α* (GenBank accession number BE642078) transcript in the samples. Reactions without template added served as the negative control. The ΔC_t method was used in calculating relative fold changes (Livak and Schmittgen 2001). The primer sequences used are listed in Table S1.

Data Availability

Strains are available upon request. Table S1 contains primers used in qRT-PCR and supplemental figures. Table S3 contains a list of all 1,163 differentially expressed genes found by all three statistics packages with annotation and statistical support included. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GBGN000000000. The version with 82,870 genes used in the differential expression analysis is the second version, GBGN020000000. RSEM results and statistical support for all Trinity predicted transcripts are available upon request. Supplemental material available at Figshare: <https://doi.org/10.25387/g3.6100139>.

RESULTS AND DISCUSSION

Gametophyte Morphology and Selection of Tissue Samples

The early development of *Ceratopteris* gametophytes can be divided into distinct stages (Banks *et al.* 1993). During the first stage (0–3d after spore inoculation), the spore swells but remains intact (Figure 1a). During stage 2 (3–4d), the spore coat cracks along its trilete markings. The first rhizoid emerges from the spore during stage 3 and the two-dimensional protonema (Figure 1b and e) emerges during stage 4 (4–6d). The male and hermaphrodite gametophytes become morphologically distinct at stage 5 (6–7d; Figure 1c and f) at which time hermaphrodites begin to secrete *A_{CE}*. For a gametophyte to develop as a male, it must continuously be exposed to *A_{CE}* during stages 2 and 3 (Banks *et al.* 1993). Because we are interested in identifying genes that are differentially expressed by *A_{CE}* treatment during the period of time that the sex of the gametophyte is determined, three populations of gametophytes were grown without *A_{CE}* for three days; at day three (end of stage 1), each population was divided into two and either media without *A_{CE}* or media with *A_{CE}* was added to the split samples. All gametophytes were harvested and processed 36hr later (stage 3; Figure 1b and e)) when the sex of the gametophyte was determined but male and hermaphrodites were morphologically indistinguishable.

Transcriptome Assembly and Annotation

The *Ceratopteris* transcriptome was assembled from 188 million Illumina paired end reads generated from the six gametophyte samples (see Table S2 for a summary of run metrics, analysis and assembly of the transcriptome). Three biological replicate samples were sequenced and analyzed for each treatment condition. A Trinity (Grabherr *et al.* 2011) *de novo* assembly resulted in 82,820 genes with read support of which 24% could be annotated with the *Arabidopsis* proteome, and 23% could be annotated by the *Selaginella* proteome. A large number of genes (1,064) could be annotated using the *Selaginella* proteome but did not have hits in the *Arabidopsis* proteome. Most of the top hits of these sequences are from the *Selaginella moellendorffii* genome (302), however many are also from *Physcomitrella patens* (175), and *Marchantia polymorpha* (131). Of these, the majority (738 sequences) had hit descriptions of predicted/hypothetical, uncharacterized, or unknown proteins. That most of these sequences do not have known gene functions is not surprising given that *Arabidopsis* is generally used in annotation of plant datasets. Of the remaining sequences which were annotated, many are sperm-related. Motile sperm are a characteristic of early divergent land plants such as *Selaginella* and *Ceratopteris* (reviewed in (Hodges *et al.* 2012)), and thus it is not surprising that such proteins would be present in these assemblies but notably absent from *Arabidopsis*. A total of 44 sequences have blast hits to dynein related proteins and 18 have hits to flagellar associated proteins. Additional sequences are also present which are sperm-related,

including radial spoke protein 9 and sporangia induced deflagellation-inducible protein. Of the annotated sequences, 43 are annotated with the cellular component GO term cilium, and 13 are annotated with cilium or flagellum-dependent cell motility; these are likely sperm-related proteins as flagellum are solely found in sperm cells in seedless vascular land plants (Raven *et al.* 2005).

Following the assembly and annotation of the *Ceratopteris* gametophyte transcriptome, the quality of the assembly was assessed. First, the quality of the Trinity assembly was assessed by comparing 5,133 *Ceratopteris* Sanger EST sequences available in GenBank to transcript sequences generated by Trinity using BLASTn. 87% of the Sanger ESTs, generated either from *C. richardii* sporophyte and gametophyte tissues were identical or almost identical (E-value of 0.0) to transcripts in the transcriptome assembly, indicating that Trinity accurately assembled transcript sequences from the short Illumina reads. The expression of the Sanger ESTs not represented in the transcriptome assembly may be age or tissue specific and thus not captured in the transcriptome assembly described here. A BUSCO analysis (Waterhouse *et al.* 2018; Simão *et al.* 2015) was also performed to assess the completeness of the transcriptome. BUSCO identifies highly conserved genes as complete, complete and single-copy, fragmented, or missing in the transcriptome. Of the 303 total BUSCO groups searched, 290 were complete (95.7%), 181 were complete and single-copy (59.7%), 10 were fragmented (3.3%), and only 3 were missing (1%). This suggests that the assembled *Ceratopteris* gametophyte transcriptome is quite complete.

Identification and Validation of Differentially Expressed Genes by Antheridiogen Treatment

Three programs, edgeR (Robinson *et al.* 2010), DESeq (Anders and Huber 2010), and EBSeq (Leng *et al.* 2013), were used to identify genes that differ in their expression by A_{CE} treatment (See Table S2 for number of differentially expressed genes found by each package). A scatterplot (Figure 2) that assesses the overall expression pattern across all transcripts shows that the expression of most transcripts is similar regardless of treatment, as expected. The majority (88%) of differentially expressed genes were more highly expressed in A_{CE} treated gametophytes (Figure 2). Of the 1,183 DEGS identified using DESeq, 1,163 were also identified by EBSeq and edgeR; these 1,163 DEGS were used in subsequent analyses. A list of the 1,163 DEGS, their annotation and supporting statistics is provided in Table S3. Of the 133 DEGS more abundant in the non- A_{CE} -treated gametophytes, 55% were annotated as protein-encoding genes, while 71% of the 1,030 DEGS more abundant in the A_{CE} treated samples could be annotated.

To test the validity of the DEG analysis, the expression of 10 genes, including genes more abundant in A_{CE} -treated samples, genes more abundant in the non- A_{CE} -treated samples and genes showing a less than twofold difference in abundance between treatments were assessed by qRT-PCR. As shown in Figure S1, the qRT-PCR expression data are consistent with the RNA-seq expression data in the direction of the fold change.

GO-Enrichment of Differentially Expressed Genes

The enrichment of Gene Ontology (GO) Biological Process terms associated with the genes that are up-regulated by A_{CE} (Figures 3 and Figure S2) reveals four major networks of enriched GO terms. One cluster includes genes related to various aspects of development, including meristem, shoot and tissue development. Another cluster includes genes involved in hormone (ABA, auxin, ethylene and GA) signaling or responses. A third cluster includes genes that affect chromatin

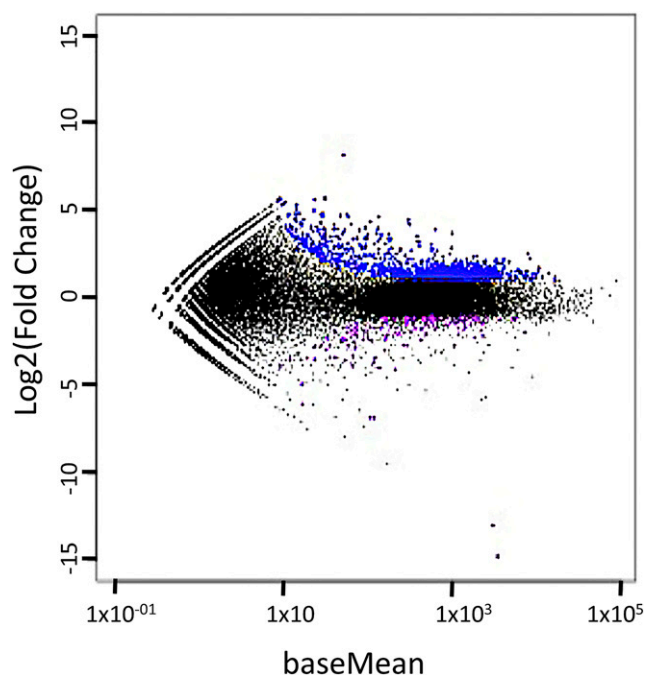


Figure 2 MA plot showing the \log_2 Fold change vs. the baseMean (normalized average expression), as calculated by DESeq (Anders and Huber 2010). Genes which are more highly expressed in $+A_{CE}$ treatment are shown in blue whereas those more highly expressed in $-A_{CE}$ treatment are shown in purple. The majority (88%) of differentially expressed genes were more highly expressed in A_{CE} treated gametophytes.

structure and epigenetic regulation of gene expression. The fourth cluster includes genes broadly involved in regulating gene expression; genes within this cluster are included in the “chromatin” cluster. Only a single GO term (response to light stimulus) was enriched for genes that are up-regulated in the non- A_{CE} treated samples.

Hormone and Development Genes Responsive to A_{CE}

Given that all characterized fern antheridiogens are gibberellins (Yamane 1998), genes involved in GA hormone biosynthesis, signaling and responses are likely to be involved in sex determination in *Ceratopteris*. Of the differentially expressed genes, *COPALYL DIPHOSPHATE SYNTHASE/KAURENE SYNTHASE (CPS/KS)*, which encodes a key enzyme in GA biosynthesis (Sun and Kamiya 1994; Hedden and Thomas 2012), is more abundant in gametophytes that will become A_{CE} secreting hermaphrodites (Table 1). No other known GA biosynthetic genes, including *kaurene oxidase* and *GA20 oxidase* are differentially expressed in *C. richardii*, indicating that sex-specific A_{CE} biosynthesis may be regulated or limited by the expression of the *CPS/KS* gene in *Ceratopteris*, and that its expression is down-regulated by A_{CE} (males do not secrete A_{CE}). All major ABA and GA signaling genes (Yamauchi *et al.* 2004; Chan 2012; Sun 2008) are present in the *Ceratopteris* transcriptome and are listed in Table S4. *Ceratopteris* seems to have all the components seen in *Arabidopsis*, though instead of the 7 DELLA proteins, responsible for repressing GA responses, *Ceratopteris* has only 2 and two F-box protein encoding genes (*SNE* and *SLY1*) involved in GA response in *Arabidopsis* are not present in the assembled *Ceratopteris* gametophyte transcriptome. Similarities in the sex determining pathway and the GA (Atallah and Banks 2015) and ABA (McAdam *et al.* 2016; Sussmilch *et al.* 2017) signaling pathways

■ Table 1 Differentially Expressed Genes Discussed in Text

Ceratopteris Gene	Annotation	Arabidopsis Accession	Blast E-value	AdjPval	FoldChange
Genes more abundant in $-A_{CE}$ treated gametophytes					
GA					
comp112296	copalyl diphosphate synthase	AT4G02780.1	3.00E-159	0.001459703	2.2
ABA					
comp103387	ABA 8'-hydroxylase	AT4G19230.1	0	0.003233766	2.4
comp112296	copalyl diphosphate synthase	AT4G02780.1	3.00E-159	0.001459703	2.2
comp112296	copalyl diphosphate synthase	AT4G02780.1	3.00E-159	0.001459703	2.2
comp112296	copalyl diphosphate synthase	AT4G02780.1	3.00E-159	0.001459703	2.2
CYtokinin					
comp80125	ARR9	AT2G41310.1	3.00E-42	1.18E-08	5.3
comp82535	ARR9	AT2G41310.1	2.00E-48	1.30E-08	4.2
comp119738	KAR-UP F-box 1	AT1G31350.1	4.00E-32	9.17E-05	2.3
Genes more abundant in $+A_{CE}$ treated gametophytes					
GA					
comp116986	SCARECROW-like (SCL)	AT5G66770.1	1.00E-87	0.006147273	2.4
comp82755	GRAS family transcription factor	AT1G63100.1	1.00E-92	0.000454384	2.7
comp103126	LOST MERISTEMS (LOM)	AT3G60630.1	5.00E-49	2.53E-05	2.9
comp81241	Lateral root primordium (LRP)	AT3G51060.1	2.00E-30	4.22E-06	3.6
comp42166	MOTHER of FT and TF 1 (MFT)	AT1G18100.1	5.00E-62	0.000371991	2.5
ABA					
comp82182	ARM repeat protein	AT5G19330.1	0	0.005713374	2.2
comp100365	ABA-insensitive 3 (ABI3)	AT3G24650.1	2.00E-40	0.000180668	2.6
comp103619	Protein phosphatase 2C	AT1G72770.3	1.00E-38	2.88E-05	3.2
comp114719	KEEP ON GOING (KEG)	AT5G13530.1	0	1.01E-07	3.7
Ethylene					
comp106297	ETHYLENE-INSENSITIVE2 (EIN2)	AT5G03280.1	1.00E-64	0.001387265	2.5
Auxin					
comp101920	NO VEIN (NOV)	AT4G13750.1	0	0.000253886	2.7
comp106375	PIN-FORMED 4 (PIN4)	AT2G01420.1	4.00E-166	2.68E-08	4.6
comp105872	PIN-FORMED 3 (PIN3)	AT1G70940.1	5.00E-156	0.009233133	2.2
comp98976	BIG auxin transport protein	AT3G02260.1	0	7.20E-09	4.2
comp109704	ABC transporter	AT3G28860.1	0	7.48E-12	4.7
comp97116	SART-1 family protein DOT2	AT5G16780.1	3.00E-132	0.008328934	2.5
comp114948	SAR1	AT1G33410.2	0	4.83E-05	3
comp105798	auxin response factor (ARF)	AT1G19220.1	5.00E-53	0.000175818	6.5
Cytokinin					
comp111805	AHK4; cytokinin receptor CRE1a	AT2G01830.1	0	0.000132766	3.2
comp100079	CKI1	AT2G47430.1	4.00E-108	0.000725748	2.6
DNA methylation/demethylation					
comp115365	cytosine methyltransferase (MET1)	AT5G49160.1	0	1.45E-06	3.3
comp82159	chromomethylase (CMT3)	AT1G69770.1	1.00E-155	0.006306721	2.3
comp112176	DEMETER-like protein 1 (ROS1)	AT2G36490.1	8.00E-83	0.001457903	2.7
comp101924	NERD	AT2G16485.1	7.00E-96	4.01E-05	3
Chromatin remodeling					
comp109662	CHR11; chromatin-remodeling 11	AT3G06400.2	0	0.007217854	2.2
comp83245	CHR5; chromatin remodeling 5	AT2G13370.1	0	2.28E-08	3.9
comp103550	CHR4; chromatin remodeling 4	AT5G44800.1	0.00E+00	6.59E-09	4.1
comp40502	PICKLE (PKL)	AT2G25170.1	0.00E+00	0.00059255	2.6
comp103233	PICKLE (PKL)	AT2G25170.1	5.00E-124	6.59E-05	2.8
comp39118	BRAHMA (BRM)	AT2G46020.2	0	5.18E-12	5
comp43532	CHR21/INO80	1.75E-05	3		
Histone modification					
comp81987	MBD09; methyl-CpG-binding domain	AT3G01460.1	5.00E-103	4.26E-09	4.1
comp99654	SUVH4/KYP	AT5G13960.1	0	0.001195543	2.5

(continued)

■ Table 1, continued

Ceratopteris Gene	Annotation	Arabidopsis Accession	Blast E-value	AdjPval	FoldChange
comp83034	CURLYLEAF (CLF)	AT2G23380.1	0	0.000158512	2.8
comp102724	ATX2	AT1G05830.2	0	0.000231694	2.8
comp83655	ATXR3	AT4G15180.1	2.00E-180	8.34E-08	3.8
comp98691	HAC12 histone acetyltransferase	AT1G16710.1	0	0.000576165	2.6
comp62161	HAC1 histone acetyltransferase	AT1G79000.1	0	0.001018811	2.6
comp108638	HAC1 histone acetyltransferase	AT1G79000.1	0	0.00334741	2.5
comp98650	subunit of Elongator	AT5G13680.1	0	0.009770466	2.2
comp106634	ASHH2 histone-lysine N-methyltransferase	AT1G77300.2	2.00E-94	5.83E-06	3.1
comp110316	IDM1 histone H3 acetyltransferase	AT3G14980.1	1.00E-111	7.37E-05	3
comp111521	histone deacetylase HDA14	AT4G33470.1	0	0.007411904	2.3
comp109495	SUVH6	AT2G22740.1	2.00E-142	0.00720228	2.5
RNA-mediated gene silencing pathways					
comp108491	ARGONAUTE1 (AGO1)	AT1G48410.1	0	0.000891723	2.5
comp82278	ARGONAUTE1 (AGO1)	AT1G48410.1	0	0.000345422	2.6
comp112142	DICER-LIKE 1 (DCL1)	AT1G01040.1	0	0.001158223	2.5
comp110523	DICER-LIKE 1 (DCL1)	AT1G01040.1	0	0.000162621	2.9
comp37939	DICER-LIKE 4 (DCL4)	AT5G20320.1	2.00E-179	0.00411352	2.4
comp81990	THO complex subunit 2	AT1G24706.1	0	2.46E-05	3
comp82821	SOU	AT3G48050.2	3.00E-91	2.98E-08	3.9
comp81850	NRPD2a	AT3G23780.1	0.00E+00	8.85E-03	2.2
comp111720	NRPD2b	AT3G18090.1	0.00E+00	1.06E-03	2.5
comp115970	XRN4	AT5G57610.1	8.00E-148	2.28E-07	2.9

treatment (Table 1). While ABA is known to affect sex determination by blocking the A_{CE} response, these results point to roles for additional hormones in the sex-determining process. Studies of the effects of exogenous auxin (Gregorich and Fisher 2006; Hickok and Kiriluk 1984), ethylene (Kazmierczak 2010) and cytokinin (Menéndez *et al.* 2009) on fern gametophyte development have shown that these hormones can affect the overall size and organization of the gametophyte as well as the number of sex organs in a gametophyte. However, neither auxin, ethylene or cytokinin substitute for or completely block the male-inducing effects of antheridiogen, indicating that A_{CE} may influence these hormones, or the crosstalk among these hormones, in modulating cell division and expansion in young gametophytes that will become important as they differentiate.

This DEG analysis suggests that A_{CE} affects the sex of the gametophyte by not only activating genes associated with development, but also by epigenetically reprogramming the nucleus that will divide and ultimately give rise to a male gametophyte. The relatively few genes that are up-regulated in gametophytes not treated with A_{CE} likely represent genes that are normally expressed in the gametophyte destined to become hermaphrodite but are repressed by A_{CE} .

An Epigenetic Response to A_{CE}

A striking number of DEGs up-regulated by A_{CE} encode factors involved in epigenetic regulation of gene expression or epigenetic reprogramming of the genome. These genes were sorted into five groups (Table 1) following the classification of Pikaard and Scheid (Pikaard and Scheid 2014): DNA modification, histone modification, Polycomb-group proteins and interacting components, chromatin formation/chromatin remodeling and RNA silencing.

The first group includes DNA modification genes that affect cytosine methylation. The DEGs assigned to this group encode *DNA METHYLTRANSFERASE 1 (MET1)*, which maintains CpG methylation (Saze *et al.* 2003; Jullien *et al.* 2012), *CHROMOMETHYLASE 3 (CMT3)*, which maintains CpHpG methylation (Law and Jacobsen 2010) and *REPRESSOR OF SILENCING 1 (ROS1)*, a cytosine demethylase (Gong *et al.* 2002). Differences in global DNA methylation

patterns between gametes and adjacent cells of both male and female gametophytes of *Arabidopsis* have been observed (Pillot *et al.* 2010; Calarco *et al.* 2012; Ibarra *et al.* 2012; Jullien *et al.* 2012) and are thought to silence transposable elements and reset silenced imprinted genes in sperm cells (Kawashima and Berger 2014; Martínez *et al.* 2016). While sex determination in a homosporous fern, which occurs during the gametophyte generation, differs from sex determination in the heterosporous angiosperms, which occurs during the sporophyte generation (Tanurdzic and Banks 2004), the up-regulation of these genes during sex determination in *Ceratopteris* adds another stage of plant development where DNA methylation may play an important role in stabilizing or destabilizing transposable elements and contributes to epigenetic reprogramming of the male gametophyte. Whether the observed differential expression of these DNA methylation genes alters DNA methylation patterns in the genomes of young *Ceratopteris* gametophytes, and whether additional changes in DNA methylation occur as their gametes differentiate, remain to be tested.

A number of A_{CE} -up-regulated DEGs encode proteins belonging to the second group, histone-modifying enzymes known to affect gene expression (Table 1). Among them are the histone acetyltransferases HAC1, HAC12 and ROS4, a histone deacetylase (HDA14), the histone methyltransferases TRITHORAX-LIKE PROTEIN 2 and 3 (ATX2 and 3), the SU(VAR)3-9 related proteins SUVH4/KYP and SUVH6, and EARLY FLOWERING IN SHORT DAYS (EFS/SDG8). These proteins are involved in either maintaining transcriptionally active states or transcriptionally inactive states (reviewed in (Liu *et al.* 2010; Bannister and Kouzarides 2011; Grossniklaus and Paro 2014; Pikaard and Scheid 2014; Xiao *et al.* 2016) and can contribute to the maintenance of DNA methylation at silenced loci. ATXR3 is notable in that it is essential for male and female gametophyte development (Berr *et al.* 2010) in angiosperms. Only one DEG, *CURLYLEAF (CLF)*, was classified as encoding proteins from the third group of chromatin modifiers: Polycomb proteins. Polycomb proteins and interacting partners are often involved in determining cell proliferation and identity through methylation and chromatin compaction (Grossniklaus and Paro 2014; Kingston and Tamkun 2014). The fourth group of genes, those involved in

chromatic formation/remodeling, are also represented among the genes up-regulated in response to A_{CE} . *PICKLE* (*PKL*), the gene encoding for a chromatin remodeling factor which is necessary for gibberellin modulated development in *Arabidopsis*, (Park *et al.* 2017) and INOSITOL-REQUIRING 80 (*INO80*), are both members of remodeling complexes and are required for normal development (Zhang *et al.* 2015).

The fifth group of genes involving epigenetic regulation is those relating to RNA-mediated gene silencing pathways (Table 1). Argonaute (*AGO*) 1, a core member of the RNA-induced silencing complex (*RISC*) which is involved post transcriptional gene silencing (*PTGS*) through cleavage or transcriptional inhibition (reviewed in (Czech and Hannon 2011; Martienssen and Moazed 2015)) is significantly up-regulated by A_{CE} . Also up regulated are genes encoding two Dicer endonucleases: *DCL1* which generates miRNAs of mostly 21nt and *DCL4*, which generates siRNAs that are 21nt (Pouch-Péllissier *et al.* 2008). Additional genes involved in RNA-mediated *PTGS* are *XRN4*, which encodes a nuclease involved in small RNA processing (Cao *et al.* 2014), and *SUO*, which encodes a component of the miRNA pathway (Yang *et al.* 2012). *NRPD2*, encoding the catalytic subunit of RNA polymerase IV and V in plants (Ream *et al.* 2009) is also up-regulated by A_{CE} . Pol IV and V are both required for intercellular RNA interference and are involved in *PTGS* maintenance (Onodera *et al.* 2005; Pontier *et al.* 2005; Kanno *et al.* 2005). Also modulated by A_{CE} is a component of the *THO/TREX* complex, which has a putative role in siRNA biosynthesis (Furumizu *et al.* 2010). Interestingly, the *THO* complex represses female germline specification in *Arabidopsis* (Su *et al.* 2017). Together, these results show that small RNA-mediated *PTGS* is involved in the suppression of female characteristics in *C. richardii* gametophytes.

All of the epigenetic mechanisms known to occur in plants are represented among the genes up regulated by A_{CE} . The importance of epigenetic regulation for sex determination in *C. richardii* should perhaps not come as a surprise. Gametophytes which are removed from A_{CE} containing media will over time develop into hermaphrodite gametophytes, thus the promotion of male/suppression of female traits must be reversible. Epigenetic regulation of sex determination would allow for such plasticity in development.

Conclusions

This work reports the first transcriptome of *Ceratopteris richardii*, along with a survey of significant differential gene expression changes between male and hermaphrodite gametophytes as sex is being determined. A high-quality reference gametophyte transcriptome was assembled and used in the identification of genes which may be involved in sex determination. The majority of differentially expressed genes were more highly expressed in the male gametophyte. Many of these up regulated genes are known to be involved in development and in response to hormones. A significant number of differentially expressed genes are involved in chromatin remodeling and epigenetic regulation. Outcomes of this research shed light on the molecular mechanisms involved in sex determination of *C. richardii* as well as provide a resource for other plant science researchers. Future work will probe and functionally classify these differentially expressed genes and as well as survey how these changes persist as the gametophyte moves from sex determination to differentiation.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (NSF 1258091 to JB) and the University of Queensland. We thank the Purdue University Rosen Center for Advanced Computing (RCAC)

for access to computational resources throughout the project, the Texas Advanced Computing Center (TACC) and Matthew Vaughn for computational resources and support. We are grateful to Michael Gribskov, Steve Kelly, Mikhail Atallah, Phillip SanMiguel, Joseph Ogas, Qiong Wu, and Rick Westerman for technical and computational support.

Author Contributions: NA performed tissue preparation and treatment as well as harvesting, performed bioinformatics analysis, performed qRT-PCR, interpreted data, and wrote part of the manuscript. FG provided bioinformatics guidance and support. OV participated in designing the experiment as well as provided statistical guidance and support. JB and MT designed the experiment, conceived of experimental design and setup, interpreted data, and wrote manuscript. All authors reviewed and approved the final manuscript.

LITERATURE CITED

- Anders, S., and W. Huber, 2010 Differential expression analysis for sequence count data. *Genome Biol.* 11: R106. <https://doi.org/10.1186/gb-2010-11-10-r106>
- Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler *et al.*, 2000 Gene ontology: tool for the unification of biology. *Nat. Genet.* 25: 25–29. <https://doi.org/10.1038/75556>
- Atallah, N. M., and J. A. Banks, 2015 Reproduction and the pheromonal regulation of sex type in fern gametophytes. *Front. Plant Sci.* 6: 100. <https://doi.org/10.3389/fpls.2015.00100>
- Aya, K., Y. Hiwatashi, M. Kojima, H. Sakakibara, M. Ueguchi-Tanaka *et al.*, 2011 The gibberellin perception system evolved to regulate a pre-existing gamyb-mediated system during land plant evolution. *Nat. Commun.* 2: 544. <https://doi.org/10.1038/ncomms1552>
- Banks, J. A., 1994 Sex-determining genes in the homosporous fern *Ceratopteris*. *Development* 120: 1949–1958.
- Banks, J. A., 1997 The transformer genes of the fern *Ceratopteris* simultaneously promote meristem and archegonia development and repress antheridia development in the developing gametophyte. *Genetics* 147: 1885–1897.
- Banks, J. A., 1999 Gametophyte development in ferns. *Annu. Rev. Plant Biol.* 50: 163–186. <https://doi.org/10.1146/annurev.arplant.50.1.163>
- Banks, J. A., L. Hickok, and M. A. Webb, 1993 The programming of sexual phenotype in the homosporous fern *Ceratopteris richardii*. *Int. J. Plant Sci.* 154: 522–534. <https://doi.org/10.1086/297135>
- Banks, J. A., T. Nishiyama, M. Hasebe, J. L. Bowman, M. Gribskov *et al.*, 2011 The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* 332: 960–963. <https://doi.org/10.1126/science.1203810>
- Bannister, A. J., and T. Kouzarides, 2011 Regulation of chromatin by histone modifications. *Cell Res.* 21: 381–395. <https://doi.org/10.1038/cr.2011.22>
- Benjamini, Y., D. Drai, G. Elmer, N. Kafkafi, and I. Golani, 2001 Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* 125: 279–284. [https://doi.org/10.1016/S0166-4328\(01\)00297-2](https://doi.org/10.1016/S0166-4328(01)00297-2)
- Berr, A., E. J. McCallum, R. Ménard, D. Meyer, J. Fuchs *et al.*, 2010 *Arabidopsis* SET DOMAIN GROUP2 is required for h3k4 trimethylation and is crucial for both sporophyte and gametophyte development. *Plant Cell* 22: 3232–3248. <https://doi.org/10.1105/tpc.110.079962>
- Calarco, J. P., F. Borges, M. T. Donoghue, F. Van Ex, P. E. Jullien *et al.*, 2012 Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell* 151: 194–205. <https://doi.org/10.1016/j.cell.2012.09.001>
- Cao, M., P. Du, X. Wang, Y.-Q. Yu, Y.-H. Qiu *et al.*, 2014 Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* 111: 14613–14618. <https://doi.org/10.1073/pnas.1407131111>

- Chan, Z., 2012 Expression profiling of ABA pathway transcripts indicates crosstalk between abiotic and biotic stress responses in Arabidopsis. *Genomics* 100: 110–115. <https://doi.org/10.1016/j.ygeno.2012.06.004>
- Cline, M. S., M. Smoot, E. Cerami, A. Kuchinsky, N. Landys *et al.*, 2007 Integration of biological networks and gene expression data using cytoscape. *Nat. Protoc.* 2: 2366–2382. <https://doi.org/10.1038/nprot.2007.324>
- Czech, B., and G. J. Hannon, 2011 Small RNA sorting: matchmaking for Argonautes. *Nat. Rev. Genet.* 12: 19–31. <https://doi.org/10.1038/nrg2916>
- Earl, D., K. Bradnam, J. St John, A. Darling, D. Lin *et al.*, 2011 Assemblathon 1: a competitive assessment of de novo short read assembly methods. *Genome Res.* 21: 2224–2241. <https://doi.org/10.1101/gr.126599.111>
- Eberle, J. R., and J. A. Banks, 1996 Genetic interactions among sex-determining genes in the fern *Ceratopteris richardii*. *Genetics* 142: 973–985.
- Finn, R. D., J. Clements, and S. R. Eddy, 2011 Hmmer web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39: W29–W37. <https://doi.org/10.1093/nar/gkr367>
- Furber, M., L. N. Mander, J. E. Nester, N. Takahashi, and H. Yamane, 1989 Structure of an antheridiogen from the fern *Anemia mexicana*. *Phytochemistry* 28: 63–66. [https://doi.org/10.1016/0031-9422\(89\)85008-3](https://doi.org/10.1016/0031-9422(89)85008-3)
- Furumizu, C., H. Tsukaya, and Y. Komeda, 2010 Characterization of EMU, the *arabidopsis* homolog of the yeast THO complex member HPR1. *RNA* 16: 1809–1817. <https://doi.org/10.1261/rna.2265710>
- Goff, S. A., M. Vaughn, S. McKay, E. Lyons, A. E. Stapleton *et al.*, 2011 The iPlant collaborative: cyberinfrastructure for plant biology. *Front. Plant Sci.* 2: 34. <https://doi.org/10.3389/fpls.2011.00034>
- Gong, Z., T. Morales-Ruiz, R. R. Ariza, T. Roldán-Arjona, L. David *et al.*, 2002 ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. *Cell* 111: 803–814.
- Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson *et al.*, 2011 Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* 29: 644–652. <https://doi.org/10.1038/nbt.1883>
- Gregorich, M., and R. Fisher, 2006 Auxin regulates lateral meristem activation in developing gametophytes of *Ceratopteris richardii*. *Botany* 84: 1520–1530.
- Grossniklaus, U., and R. Paro, 2014 Transcriptional silencing by polycomb-group proteins. *Cold Spring Harb. Perspect. Biol.* 6: a019331. <https://doi.org/10.1101/cshperspect.a019331>
- Gubler, F., P. M. Chandler, R. G. White, D. J. Llewellyn, and J. V. Jacobsen, 2002 Gibberellin signaling in barley aleurone cells. control of SLN1 and GAMYB expression. *Plant Physiol.* 129: 191–200. <https://doi.org/10.1104/pp.010918>
- Hanlon, M. R., M. Vaughn, S. Mock, R. Dooley, W. Moreira *et al.*, 2015 Araport: an application platform for data discovery. *Concurr. Comput.* 27: 4412–4422. <https://doi.org/10.1002/cpe.3542>
- Haufler, C. H., 2002 Homospory 2002: An odyssey of progress in pteridophyte genetics and evolutionary biology: Ferns and other homosporous vascular plants have highly polyploid chromosome numbers, but they express traits following diploid models and, although capable of extreme inbreeding, are predominantly outcrossing. *A.I.B.S. Bull.* 52: 1081–1093.
- Hedden, P., and S. G. Thomas, 2012 Gibberellin biosynthesis and its regulation. *Biochem. J.* 444: 11–25. <https://doi.org/10.1042/BJ20120245>
- Hickok, L. G., 1983 Absciscic acid blocks antheridiogen-induced antheridium formation in gametophytes of the fern *Ceratopteris*. *Can. J. Bot.* 61: 888–892. <https://doi.org/10.1139/b83-098>
- Hickok, L. G., and R. M. Kiriluk, 1984 Effects of auxins on gametophyte development and sexual differentiation in the fern *Ceratopteris thalictroides* (L.) brongn. *Bot. Gaz.* 145: 37–42. <https://doi.org/10.1086/337423>
- Hickok, L. G., T. R. Warne, and M. K. Slocum, 1987 *Ceratopteris richardii*: applications for experimental plant biology. *Am. J. Bot.* 74: 1304–1316. <https://doi.org/10.1002/j.1537-2197.1987.tb08743.x>
- Hodges, M. E., B. Wickstead, K. Gull, and J. A. Langdale, 2012 The evolution of land plant cilia. *New Phytol.* 195: 526–540. <https://doi.org/10.1111/j.1469-8137.2012.04197.x>
- Ibarra, C. A., X. Feng, V. K. Schoft, T.-F. Hsieh, R. Uzawa *et al.*, 2012 Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* 337: 1360–1364. <https://doi.org/10.1126/science.1224839>
- Jullien, P. E., D. Susaki, R. Yelagandula, T. Higashiyama, and F. Berger, 2012 DNA methylation dynamics during sexual reproduction in *Arabidopsis thaliana*. *Curr. Biol.* 22: 1825–1830. <https://doi.org/10.1016/j.cub.2012.07.061>
- Kanehisa, M., S. Goto, Y. Sato, M. Furumichi, and M. Tanabe, 2011 Kegg for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 40: D109–D114. <https://doi.org/10.1093/nar/gkr988>
- Kanno, T., B. Huettel, M. F. Mette, W. Aufsatz, E. Jaligot *et al.*, 2005 Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat. Genet.* 37: 761–765. <https://doi.org/10.1038/ng1580>
- Kawashima, T., and F. Berger, 2014 Epigenetic reprogramming in plant sexual reproduction. *Nat. Rev. Genet.* 15: 613–624. <https://doi.org/10.1038/nrg3685>
- Kaźmierczak, A., 2010 Gibberellic acid and ethylene control male sex determination and development of *Anemia phyllitidis* gametophytes, pp. 49–65 in *Working with Ferns*, Springer, New York. https://doi.org/10.1007/978-1-4419-7162-3_5
- Kingston, R. E., and J. W. Tamkun, 2014 Transcriptional regulation by trithorax-group proteins. *Cold Spring Harb. Perspect. Biol.* 6: a019349. <https://doi.org/10.1101/cshperspect.a019349>
- Law, J. A., and S. E. Jacobsen, 2010 Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11: 204–220. <https://doi.org/10.1038/nrg2719>
- Leng, N., J. A. Dawson, J. A. Thomson, V. Ruotti, A. I. Rissman *et al.*, 2013 EBSeq: an empirical bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* 29: 1035–1043. <https://doi.org/10.1093/bioinformatics/btt087>
- Li, B., and C. N. Dewey, 2011 RSEM: accurate transcript quantification from rna-seq data with or without a reference genome. *BMC Bioinformatics* 12: 323. <https://doi.org/10.1186/1471-2105-12-323>
- Li, Z., A. E. Baniaga, E. B. Sessa, M. Scascitelli, S. W. Graham *et al.*, 2015 Early genome duplications in conifers and other seed plants. *Sci. Adv.* 1: e1501084. <https://doi.org/10.1126/sciadv.1501084>
- Liu, C., F. Lu, X. Cui, and X. Cao, 2010 Histone methylation in higher plants. *Annu. Rev. Plant Biol.* 61: 395–420. <https://doi.org/10.1146/annurev.arplant.043008.091939>
- Livak, K. J., and T. D. Schmittgen, 2001 Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\delta\delta$ ct method. *Methods* 25: 402–408. <https://doi.org/10.1006/meth.2001.1262>
- Lohse, M., A. M. Bolger, A. Nagel, A. R. Fernie, J. E. Lunn *et al.*, 2012 RobNA: a user-friendly, integrated software solution for RNA-seq-based transcriptomics. *Nucleic Acids Res.* 40: W622–W627. <https://doi.org/10.1093/nar/gks540>
- Martienssen, R., and D. Moazed, 2015 RNAi and heterochromatin assembly. *Cold Spring Harb. Perspect. Biol.* 7: a019323. <https://doi.org/10.1101/cshperspect.a019323>
- Martínez, G., K. Panda, C. Köhler, and R. K. Slotkin, 2016 Silencing in sperm cells is directed by RNA movement from the surrounding nurse cell. *Nat. Plants* 2: 16030. <https://doi.org/10.1038/nplants.2016.30>
- McAdam, S. A., T. J. Brodribb, J. A. Banks, R. Hedrich, N. M. Atallah *et al.*, 2016 Absciscic acid controlled sex before transpiration in vascular plants. *Proc. Natl. Acad. Sci. USA* 113: 12862–12867. <https://doi.org/10.1073/pnas.1606614113>
- Menéndez, V., M. Revilla, M. Fal, and H. Fernández, 2009 The effect of cytokinins on growth and sexual organ development in the gametophyte of *Blechnum spicant* L. *Plant Cell Tissue Organ Cult.* 96: 245–250 (PCTOC). <https://doi.org/10.1007/s11240-008-9481-y>

- Onodera, Y., J. R. Haag, T. Ream, P. C. Nunes, O. Pontes *et al.*, 2005 Plant nuclear RNA polymerase iv mediates sirna and dna methylation-dependent heterochromatin formation. *Cell* 120: 613–622. <https://doi.org/10.1016/j.cell.2005.02.007>
- Park, J., D.-H. Oh, M. Dassanayake, K. T. Nguyen, J. Ogas *et al.*, 2017 Gibberellin signaling requires chromatin remodeler pickle to promote vegetative growth and phase transitions. *Plant Physiol.* 173: 1463–1474. <https://doi.org/10.1104/pp.16.01471>
- Pikaard, C. S., and O. M. Scheid, 2014 Epigenetic regulation in plants. *Cold Spring Harb. Perspect. Biol.* 6: a019315. <https://doi.org/10.1101/cshperspect.a019315>
- Pillot, M., C. Baroux, M. A. Vazquez, D. Autran, O. Leblanc *et al.*, 2010 Embryo and endosperm inherit distinct chromatin and transcriptional states from the female gametes in Arabidopsis. *Plant Cell* 22: 307–320. <https://doi.org/10.1105/tpc.109.071647>
- Pontier, D., G. Yahubyan, D. Vega, A. Bulski, J. Saez-Vasquez *et al.*, 2005 Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases iv in Arabidopsis. *Genes Dev.* 19: 2030–2040. <https://doi.org/10.1101/gad.348405>
- Pouch-Pélissier, M.-N., T. Pélissier, T. Elmayan, H. Vaucheret, D. Boko *et al.*, 2008 SINE RNA induces severe developmental defects in Arabidopsis thaliana and interacts with HYL1 (DRB1), a key member of the DCL1 complex. *PLoS Genet.* 4: e1000096. <https://doi.org/10.1371/journal.pgen.1000096>
- Raven, P. H., R. F. Evert, and S. E. Eichhorn, 2005 *Biology of Plants*. New York, W.H. Freeman, 7th edition.
- Ream, T. S., J. R. Haag, A. T. Wierzbicki, C. D. Nicora, A. D. Norbeck *et al.*, 2009 Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol. Cell* 33: 192–203. <https://doi.org/10.1016/j.molcel.2008.12.015>
- Robertson, G., J. Schein, R. Chiu, R. Corbett, M. Field *et al.*, 2010 De novo assembly and analysis of RNA-seq data. *Nat. Methods* 7: 909–912. <https://doi.org/10.1038/nmeth.1517>
- Robinson, M. D., D. J. McCarthy, and G. K. Smyth, 2010 edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–140. <https://doi.org/10.1093/bioinformatics/btp616>
- Saito, R., M. E. Smoot, K. Ono, J. Ruscheinski, P.-L. Wang *et al.*, 2012 A travel guide to cytoscape plugins. *Nat. Methods* 9: 1069–1076. <https://doi.org/10.1038/nmeth.2212>
- Saze, H., O. M. Scheid, and J. Paszkowski, 2003 Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nat. Genet.* 34: 65–69. <https://doi.org/10.1038/ng1138>
- Schmieder, R., and R. Edwards, 2011 Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PLoS One* 6: e17288. <https://doi.org/10.1371/journal.pone.0017288>
- Schmieder, R., Y. W. Lim, and R. Edwards, 2011 Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics* 28: 433–435. <https://doi.org/10.1093/bioinformatics/btr669>
- Schulz, M. H., D. R. Zerbino, M. Vingron, and E. Birney, 2012 Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28: 1086–1092. <https://doi.org/10.1093/bioinformatics/bts094>
- Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang *et al.*, 2003 Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13: 2498–2504. <https://doi.org/10.1101/gr.1239303>
- Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov, 2015 BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31: 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Smoot, M. E., K. Ono, J. Ruscheinski, P.-L. Wang, and T. Ideker, 2010 Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27: 431–432. <https://doi.org/10.1093/bioinformatics/btq675>
- Strain, E., B. Hass, and J. A. Banks, 2001 Characterization of mutations that feminize gametophytes of the fern *Ceratopteris*. *Genetics* 159: 1271–1281.
- Su, Z., L. Zhao, Y. Zhao, S. Li, S. Won *et al.*, 2017 The tho complex non-cell-autonomously represses female germline specification through the tas3-arf3 module. *Curr. Biol.* 27: 1597–1609.e2. <https://doi.org/10.1016/j.cub.2017.05.021>
- Sun, T., and Y. Kamiya, 1994 The Arabidopsis GA1 locus encodes the cyclase ent-kaurene synthetase a of gibberellin biosynthesis. *Plant Cell* 6: 1509–1518. <https://doi.org/10.1105/tpc.6.10.1509>
- Sun, T.-P., 2008 Gibberellin metabolism, perception and signaling pathways in Arabidopsis. *The Arabidopsis Book* 6: e0103. <https://doi.org/10.1199/tab.0103>
- Sussmilch, F. C., N. M. Atallah, T. J. Brodribb, J. A. Banks, and S. A. McAdam, 2017 Abscissic acid (aba) and key proteins in its perception and signaling pathways are ancient, but their roles have changed through time. *Plant Signal. Behav.* 12: e1365210. <https://doi.org/10.1080/15592324.2017.1365210>
- Takeno, K., H. Yamane, T. Yamauchi, N. Takahashi, M. Furber *et al.*, 1989 Biological activities of the methyl ester of gibberellin a73, a novel and principal antheridiogen in *Lygodium japonicum*. *Plant Cell Physiol.* 30: 201–205. <https://doi.org/10.1093/oxfordjournals.pcp.a077730>
- Tanaka, J., K. Yano, K. Aya, K. Hirano, S. Takehara *et al.*, 2014 Antheridiogen determines sex in ferns via a spatiotemporally split gibberellin synthesis pathway. *Science* 346: 469–473. <https://doi.org/10.1126/science.1259923>
- Tanurdzic, M., and J. A. Banks, 2004 Sex-determining mechanisms in land plants. *Plant Cell* 16: S61–S71. <https://doi.org/10.1105/tpc.016667>
- Warne, T. R., and L. G. Hickok, 1989 Evidence for a gibberellin biosynthetic origin of *Ceratopteris* antheridiogen. *Plant Physiol.* 89: 535–538. <https://doi.org/10.1104/pp.89.2.535>
- Warne, T. R., and L. G. Hickok, 1991 Control of sexual development in gametophytes of *Ceratopteris richardii*: Antheridiogen and abscissic acid. *Bot. Gaz.* 152: 148–153. <https://doi.org/10.1086/337874>
- Waterhouse, R. M., M. Seppey, F. A. Simão, M. Manni, P. Ioannidis *et al.*, 2018 BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* 35: 543–548. <https://doi.org/10.1093/molbev/msx319>
- Xi, W., C. Liu, X. Hou, and H. Yu, 2010 Mother of ft and tfl1 regulates seed germination through a negative feedback loop modulating aba signaling in arabidopsis. *Plant Cell* 22: 1733–1748. <https://doi.org/10.1105/tpc.109.073072>
- Xiao, J., U.-S. Lee, and D. Wagner, 2016 Tug of war: adding and removing histone lysine methylation in Arabidopsis. *Curr. Opin. Plant Biol.* 34: 41–53. <https://doi.org/10.1016/j.pbi.2016.08.002>
- Yamane, H., 1998 Fern antheridiogens. *Int. Rev. Cytol.* 184: 1–32. [https://doi.org/10.1016/S0074-7696\(08\)62177-4](https://doi.org/10.1016/S0074-7696(08)62177-4)
- Yamane, H., K. Nohara, N. Takahashi, and H. Schraudolph, 1987 Identification of antheridic acid as an antheridiogen in *Anemia rotundifolia* and *Anemia flexuosa*. *Plant Cell Physiol.* 28: 1203–1207.
- Yamauchi, Y., M. Ogawa, A. Kuwahara, A. Hanada, Y. Kamiya *et al.*, 2004 Activation of gibberellin biosynthesis and response pathways by low temperature during imbibition of arabidopsis thaliana seeds. *Plant Cell* 16: 367–378. <https://doi.org/10.1105/tpc.018143>
- Yang, L., G. Wu, and R. S. Poethig, 2012 Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA-mediated translational repression in Arabidopsis. *Proc. Natl. Acad. Sci. USA* 109: 315–320. <https://doi.org/10.1073/pnas.1114673109>
- Zhang, C., L. Cao, L. Rong, Z. An, W. Zhou *et al.*, 2015 The chromatin-remodeling factor AtINO80 plays crucial roles in genome stability maintenance and in plant development. *Plant J.* 82: 655–668. <https://doi.org/10.1111/tbj.12840>

Communicating editor: C. Pikaard