Dominance and Sexual Dimorphism Pervade the *Salix purpurea* L. Transcriptome

Craig H. Carlson[1], Yongwook Choi[2], Agnes P. Chan[2], Michelle J. Serapiglia[1], Christopher D. Town[2], and Lawrence B. Smart[1],*

[1]Horticulture Section, School of Integrative Plant Science, Cornell University, New York State Agricultural Experiment Station, Geneva, New York 14456 USA

[2]J. Craig Venter Institute, Rockville, Maryland 20850 USA

*Corresponding author: E-mail: lbs33@cornell.edu.

Abstract

The heritability of gene expression is critical in understanding heterosis and is dependent on allele-specific regulation by local and remote factors in the genome. We used RNA-Seq to test whether variation in gene expression among F_1 and F_2 intraspecific *Salix purpurea* progeny is attributable to *cis*- and *trans*-regulatory divergence. We assessed the mode of inheritance based on gene expression levels and allele-specific expression for F_1 and F_2 intraspecific progeny in two distinct tissue types: shoot tip and stem internode. In addition, we explored sexually dimorphic patterns of inheritance and regulatory divergence among F_1 progeny individuals. We show that in *S. purpurea* intraspecific crosses, gene expression inheritance largely exhibits a maternal dominant pattern, regardless of tissue type or pedigree. A significantly greater number of *cis*- and *trans*-regulated genes coincided with upregulation of the maternal parent allele in the progeny, irrespective of the magnitude, whereas the paternal allele was higher expressed for genes showing *cis* × *trans* or compensatory regulation. Importantly, consistent with previous genetic mapping results for sex in shrub willow, we have delimited sex-biased gene expression to a 2 Mb pericentromeric region on *S. purpurea* chr15 and further refined the sex determination region. Altogether, our results offer insight into the inheritance of gene expression in *S. purpurea* as well as evidence of sexually dimorphic expression which may have contributed to the evolution of dioecy in *Salix*.

Key words: differential expression, dioecy, regulatory divergence, sex determination, shrub willow, ZW system.

Introduction

Allele-specific expression (ASE) reflects the regulatory status of each parent allele inherited in an individual and has become an informative phenotype for biologists in understanding nonadditive phenotypic expression (Stupar and Springer 2006). Without further knowledge of parent pedigree, ASE can only be considered for sites that differ between the parents of an F_1 cross, whereby a single copy of each homozygous parent allele exists in a heterozygous state in the F_1 hybrid. For any biallelic site, the normalized expression ratio of the female parent (P1) allele and the male parent (P2) allele is contrasted to the same ASE ratio (P1H/P2H) in the hybrid. Statistically significant deviations of ASE in the F_1 from the expected contribution of each parent ($P = 0.5$) are based on binomial exact tests, which lay the groundwork for the estimation of *cis*-regulatory divergent gene expression. Despite the fact that the *extent* of regulatory divergence is largely dependent on predetermined global significance thresholds (i.e., False Discovery Rate, FDR), the overall *patterns* of divergent expression do not drastically change (Suvorov et al. 2013). These patterns are broadly subject to sequence variation observed in the domains of local *cis*-regulatory elements or remote *trans*-acting factors.

There is evidence that nonadditive gene expression can confer novel transgressive phenotypes in hybrids (Springer and Stupar 2007) and is alleged to be a major driver of hybrid speciation (Rieseberg et al. 2003). For instance, in interspecific crosses of *Drosophilia melanogaster* and *D. simulans*, *cis*-effects were shown to account for a majority of regulatory divergent

expression (Wittkopp et al. 2008a), whereas *trans*-effects accounted for a higher proportion of expression variation between parents of the same species (Wittkopp et al. 2004, 2008b). Hybridization can introduce substantial divergence in offspring gene expression when compared with that of the parents (McManus et al. 2010). Such a merger provides new allelic variation within the regulatory domains of genes (e.g., promoters) as well as new targets of *trans*-acting factors (e.g., transcription factors). Although mutations in *cis*-regulatory elements have been shown to account for evolutionarily significant phenotypic change, *trans*-regulatory evolution can also affect adaptive morphological change (Wittkopp and Kalay 2011).

Depending on the effective population size, the effects of *cis*-mutations on gene expression are generally considered to be less deleterious as they only affect a single gene and are more likely to become fixed, whereas *trans*-effects can alter the expression of a number of genes (pleiotropy) and are more likely to be subject to purifying selection (McManus et al. 2010). Consequently, the conservation of gene expression (P1 $\approx$ H $\approx$ P2) should be less pronounced in wide hybrids, but more common with inbreeding or sib-mating, because the transcriptional activity of ASE is simply a function of the two *cis*-regulatory parent alleles in a common *trans*-regulated background. Nevertheless, studies of ASE in intraspecific progeny derived from closely related parents have attributed parental expression divergence to both *cis*- and *trans*- regulatory components (Bell et al. 2013; Suvorov et al. 2013; Combes et al. 2015).

A bulk of ASE work in plants have used hybrids derived from inbred parents to study the effects of hybridization on gene expression (Stupar and Springer 2006; Song et al. 2013). Although there are notable exceptions (Bell et al. 2013), there is a general lack of understanding on the evolution of gene expression with regards to the hybridization of heterozygous parents from natural, obligate outcrossing populations. Previous expression studies in dioecious shrub willow (*Salix* spp., Salicaceae) have predominantly focused on correlating functional variation of candidate gene family members to lignocellulosic composition traits in contrasting pedigrees (Puckett et al. 2012; Serapiglia et al. 2012). In this study, we examined the variation in transcriptome-wide expression within and among full-sib F$_1$ and F$_2$ intraspecific families generated from heterozygous parents collected from naturalized *S. purpurea* L. populations.

Shrub willow has been bred as a dedicated energy crop since the early 1970's with the goal of producing fast-growing bioenergy feedstock cultivars that are high-yielding, genetically diverse, pest and disease resistant, and able to grow on marginal land without competing with food crops (Stoof et al. 2015). The heterogeneity and adaptive plasticity of *Salix* spp. provides an abundant germplasm pool for trait improvement and phylogenetic characterization (Hanley and Karp 2014). Hybridization is a key component in the development of shrub willow bioenergy crops, as hybrids often

display heterosis for yield (Fabio et al. 2017a, 2017b). While significant improvements in biomass yield has been realized in interspecific crosses of *Salix* (Kopp et al. 2001; Cameron et al. 2008), heterosis is more pronounced in triploid progeny derived from the hybridization of diploid and tetraploid parents (Smart and Cameron 2008; Serapiglia et al. 2014a; Carlson and Smart 2016). These high-yielding triploid shrub willow outperform foundation commercial cultivars and show promise for the future of the biomass production industry (Serapiglia et al. 2014b; Fabio et al. 2017a, 2017b).

With the public release of the *S. purpurea* genome reference assembly (phytozome.jgi.doe.gov; last accessed September 10, 2017), *Salix* has become a powerful model to study the genomic basis of heterosis in dioecious species. *Salix purpurea* has a relatively compact genome ($\sim$400 Mb) with $\sim$37,500 primary gene models and $\sim$65,000 alternatively spliced isoforms (Smart et al., in prep). Although the genome of *S. purpurea* is remarkably collinear to that of *Populus trichocarpa* (Berlin et al. 2010), major differences in the overall arrangement and abundance of coding and noncoding DNA (Hou et al. 2016) has radically affected the ecology, habit, and reproductive paths since the Salicoid duplication and divergence of the genera (Rodgers-Melnick et al. 2012). For instance, the sex determining region (SDR) of *P. trichocarpa* resides within a peritelomeric region on *Populus* chr19 (Yin et al. 2008; McKown et al. 2017), whereas the SDR of *Salix* spp. has been mapped to a pericentromeric region on *Salix* chr15 (Pucholt et al. 2015; Zhou et al. 2017). In addition to contrasting genomic locations of the SDR, *S. purpurea* exhibits a ZW sex determining system with heterogametic females (Zhou et al. 2017) and *P. trichocarpa*, an XY system with heterogametic males (Tuskan et al. 2012). To date, the mechanism of sex determination in the Salicaceae has not been completely resolved.

The main objectives of this study were 1) to test for differential gene expression among the shoot tip and internode transcriptome of segregating F$_1$ and F$_2$ intraspecific *S. purpurea* family progeny, 2) to categorize gene expression by modes of inheritance, 3) to assess the magnitude and direction of regulatory divergent expression, and 4) to examine the regulatory components of sexually dimorphic gene expression that may have contributed to the evolution of dioecy in *Salix*.

Materials and Methods

Plant Material and Growing Conditions

The full-sib intraspecific F$_1$ *S. purpurea* family was generated from a cross between the female clone 94006 and the male clone 94001, collected from a population of *S. purpurea* in Central New York. Two F$_1$ siblings from this family were selected and crossed (9882-41 $\times$ 9882-34) to generate the F$_2$ population (supplementary fig. S1A, Supplementary Material online). Dormant first-year postcoppice vegetative shoots of

all family parents and progeny were taken from nursery beds in the winter of 2013. Cuttings of equal length (20 cm) and diameter were cut from shoots and planted into 2.5 l containers filled with Farfard PV-1 potting media and grown under environmentally controlled greenhouse conditions with supplemental lighting provided on a 14 h day:10 h night regimen with a max daytime temperature of 26 °C and a nighttime temperature of 18 °C. All plots were completely randomized over five replicate blocks. Liquid fertilizer (Peter's 15-16-17 Peat-Lite Special; Scotts Miracle-Gro Company, Marysville, Ohio, U.S.A.) was applied weekly at 100 ppm after the third week from planting cuttings, until the study was terminated. Herein, we refer to parents of the F_1 and F_2 families by their clone identifiers and discriminate the female and male parents as P1 and P2, respectively.

Determination of Ploidy Level

The relative DNA content (pg 2 C value^{-1}) of family parents and progeny was determined by flow cytometric analysis using young leaf material harvested from actively growing shoots in greenhouse conditions. Analysis of 50 mg of mature leaf tissue from parental genotypes and selected progeny was performed at the Flow Cytometry and Imaging Core Laboratory at Virginia Mason Research Center in Seattle, WA as was previously described (Serapiglia et al. 2014b). A minimum of four replicates of all samples were independently assessed using either the diploid *S. purpurea* female genome reference clone 94006 or the diploid *S. purpurea* male clone 94001 as an internal standard. Diploid parent clones from multiple runs were averaged and then divided by the value of the check for that run. This factor was then multiplied by each sample value within the same run as the check. When a clone was analyzed more than once, 2C values were averaged. All parents and progeny described in this study are diploid ($2n = 38$) according to flow cytometric and genetic marker analysis (Argus 1997; Serapiglia et al. 2014b).

RNA Sample Preparation and Sequencing

Total RNA was extracted from three biological replicates of ten random family progeny and their parents for the F_1 and F_2 families. Both shoot tip and internode tissues from each individual were collected and processed using the Spectrum™ Total Plant RNA Kit with DNase digestion (Sigma, St. Louis, MO), following the manufacturers procedures. Cold-ethanol precipitations were performed by addition of 10 μl actetic acid and 280 μl 100% cold ethanol to 100 μl eluate and placed in −80 °C for at least 3 h. Samples were centrifuged at 17,000 × *g* for 30 min at 4 °C, washed with 80% cold ethanol, then centrifuged at 17,000 × *g* for 20 min at 4 °C. After centrifugation, the ethanol supernatant was discarded and the pellet resuspended in ribonuclease-free 10 mM Tris–HCl (pH = 8). Quantification of sample quality and concentration was performed using the Experion RNA StdSens kit

(Bio-Rad Laboratories, Inc., Hercules, CA), following manufacturers' procedures.

Independent extractions were performed on three replicate plants of each of ten progeny individuals within each family and subsequently pooled in equal concentrations. For each tissue type, three RNA-Seq libraries were constructed representing the female parent, the male parent, and a pool of of ten progeny. In addition, for comparisons between progeny within the F_1 family, ten F_1 progeny were individually barcoded and sequenced. Libraries were constructed using the NEBNext Ultra Directional RNA Library Prep Kit and sequenced on the Illumina platform (1x100 bp) at the J. Craig Venter Institute (Rockville, MD).

Shoot tips were defined as the shoot axis that is the most distal part of a shoot system, comprised of a shoot apical meristem and the youngest leaf primordia. Stem internodes were defined as the cardinal organ part that is the part of a shoot axis between two nodes of the axis.

Read Filtering, Alignment, and Variant Discovery

Low-coverage paired-end genomic DNA sequencing of the parental lines for the F_1 and F_2 families was performed to validate variants from RNA-Seq data. Biallelic SNPs were used to quantify allele-specific expression and regulatory divergence within and among intraspecific family progeny. Parent DNA libraries were sequenced (Illumina HiSeq 2 × 101) and aligned to the *S. purpurea* v1 reference genome using BWA-MEM (-M -R) (Li and Durbin 2009). Subsequent SAM files were sorted, marked for duplicates, and indexed in Picard (broadinstitute.github.io/picard). Indel realignment and variant calling was performed using *HaplotypeCaller* (emit_conf 10, call_conf = 30) in the Genome Analysis Toolkit (GATK) (DePristo et al. 2011). For all samples, RNA-Seq reads were trimmed (min length = 50) and mapped (length = 0.8, similarity = 0.9) to the *S. purpurea* v1 reference genome using CLC Genomics Workbench (www.qiagenbioinformatics.com; last accessed September 10, 2017). The *S. purpurea* L. v1.0 genome reference assembly and annotation is available online via the Joint Genome Institute Comparative Plant Genomics Portal, Phytozome v12 (phytozome.jgi.doe.gov).

Gene Expression Inheritance Classifications

To determine the mode of inheritance for genes, the number of RNA-Seq reads mapped to individual genes was counted for each of the female (P1) and male (P2) parents and progeny (H). Expression levels were compared based on normalized read counts using the edgeR package (Robinson et al. 2010). Differentially expressed genes were determined using an exact test implemented in edgeR for negative-binomially distributed counts (disp = 0.1, FDR = 0.005). We used a custom R script to sort genes into the following six inheritance categories: 1) P1-dominant: $H \approx P1$ and $H \neq P2$, 2) P2-dominant: $H \approx P2$ and $H \neq P1$, 3) additive: $P1 < H < P2$ or

P2 < H < P1, 4) overdominant: H > P1 and H > P2, 5) under-dominant: H < P1 and H < P2, and 6) conserved: all others.

The absolute magnitude of dimorphic gene expression inheritance was determined as the Euclidean distance (L^2) between vectors m_{xy} and f_{xy}, such that: $L^2 = \langle p, q \rangle = \left[(|m_x| - |f_x|)^2 + (|m_y| - |f_y|)^2 \right]^{0.5}$, where m_x is the male coordinate derived from $log_2(m/p_1)$, m_y is the male coordinate derived from $log_2(m/p_2)$, f_x is the female coordinate derived from $log_2(f/p_1)$, and f_y is the female coordinate derived from $log_2(f/p_2)$, p is the squared absolute difference of the vectors $(|m_x|, |f_x|)$ and $(|m_x|, |m_y|)$, and q is the squared absolute difference of the vectors $(|m_x|, |f_y|)$ and $(|m_x|, |m_y|)$.

Regulatory Divergence Classifications

For regulatory divergence classification of genes, the sequence reads in progeny need to be assigned to their parental origins. For each gene, the expression levels of the two parental alleles were estimated based on the nucleotide allele counts across all SNP sites detected based on parent DNA libraries (described above), and where the nucleotide alleles present in each of the parents are distinct and therefore allow unambiguous assignment of parental origins. Categories of regulatory functions considered conserved, compensatory, ambiguous, *cis*, *trans*, *cis* + *trans*, and *cis* × *trans*, were sorted using R scripts, as described in Landry et al. (2005) and McManus et al. (2010). Regulatory divergence assignments were based on two sets of tests: 1) a binomial exact test between P1 and P2 in the parents and between P1H and P2H in the progeny, and 2) Fisher's Exact Test on P1, P2, P1H, and P2H (supplementary table S1, Supplementary Material online). The percent divergence due to *cis*- and *trans*-effects were calculated such that % $cis = [|cis|/(|cis| + |trans|)] \times 100$ and % $trans = [|trans|/(|cis| + |trans|)] \times 100$, where $cis = log_2(P1H/P2H)$ and $trans = log_2(P1H/P2H) - log_2(P1/P2)$.

Tests for Differential Expression

All statistical analyses were performed in the open-source statistical computing environment, R (Team 2015). Tests for differential expression were conducted in the package edgeR (Robinson et al. 2010). Normalization factors and dispersion estimates (robust = T) were computed prior to tests for differential expression. A general linear model was used to fit normalized count data using *glmFit* and *glmLRT* to conduct likelihood ratio tests for the model coefficients. Tests for paired comparisons were conducted to investigate the effect of tissue type over the individuals within the F_1 family, using an additive linear model with clone as a blocking factor ($y \sim 0 + clone + tissue$). In order to test for differential expression by sex ($y \sim pedigree + sex$) in the F_1 family, a conservative and

robust quasi-likelihood model using tag-wise dispersion estimates was used to fit the data with *glmQLFit* (robust = T). For tests of differential expression, genes were only considered to be significant at a False Discovery Rate of 0.05. To explore variation in gene content, we considered the total number of genes expressed per sample library (transcriptome-normalized expression), without inferring relative expression per gene copy or per cell, as is described in Coate and Doyle (2010). A summary of analyses and corresponding libraries used is in supplementary table S2, Supplementary Material online.

Gene Ontology

Gene ontology (GO) enrichment was performed in agriGO (Du et al. 2010) using the subset of the *S. purpurea* v1 transcriptome (reference set) that passed filtering prior to tests of differential expression. Only significant ontologies were reported from query lists. For *S. purpurea* gene annotations which encode for hypothetical proteins, gene models and associated GO-terms were inferred using the best-hit (BLASTP E-value ≤ 0.1) to *Populus trichocarpa* (Phytozome v10.3 annotation) and Arabidopsis (TAIR10 and Araport11 annotations) proteome.

Results

Transcriptome Analysis

In order to define the factors contributing to variation in global expression among the F_1 and F_2 parents and progeny, the RNA-Seq data were subjected to multidimensional scaling (MDS) analysis, considering only genes with a sum cpm-normalized expression >1.0 for ≥50% of the samples (supplementary fig. S1*B*, Supplementary Material online). As expected, the parent transcriptomes of the F_1 were the most distantly clustered, because they are the least related in this study. The first MDS dimension clustered samples based entirely on tissue type, and the second dimension split F_1 and F_2 parents by sex, then by pedigree. Although 9882-41 and 9882-34 are F_1 siblings, their gene expression levels clustered more towards the parent of the same sex. The pedigrees, ploidy levels, and read mapping statistics are summarized in supplementary table S3, Supplementary Material online.

Inheritance of Gene Expression

The mode of gene expression inheritance assessed in F_1 and F_2 intraspecific progeny was based on two distinct tissue types: shoot tip and stem internode. Differentially expressed genes were defined as having normalized expression levels significantly higher or lower for treatment comparisons at an FDR of 0.005. In general, gene expression in both the F_1 and F_2 family was largely conserved, yet expression differences between the F_1 parents was far more pronounced than among the
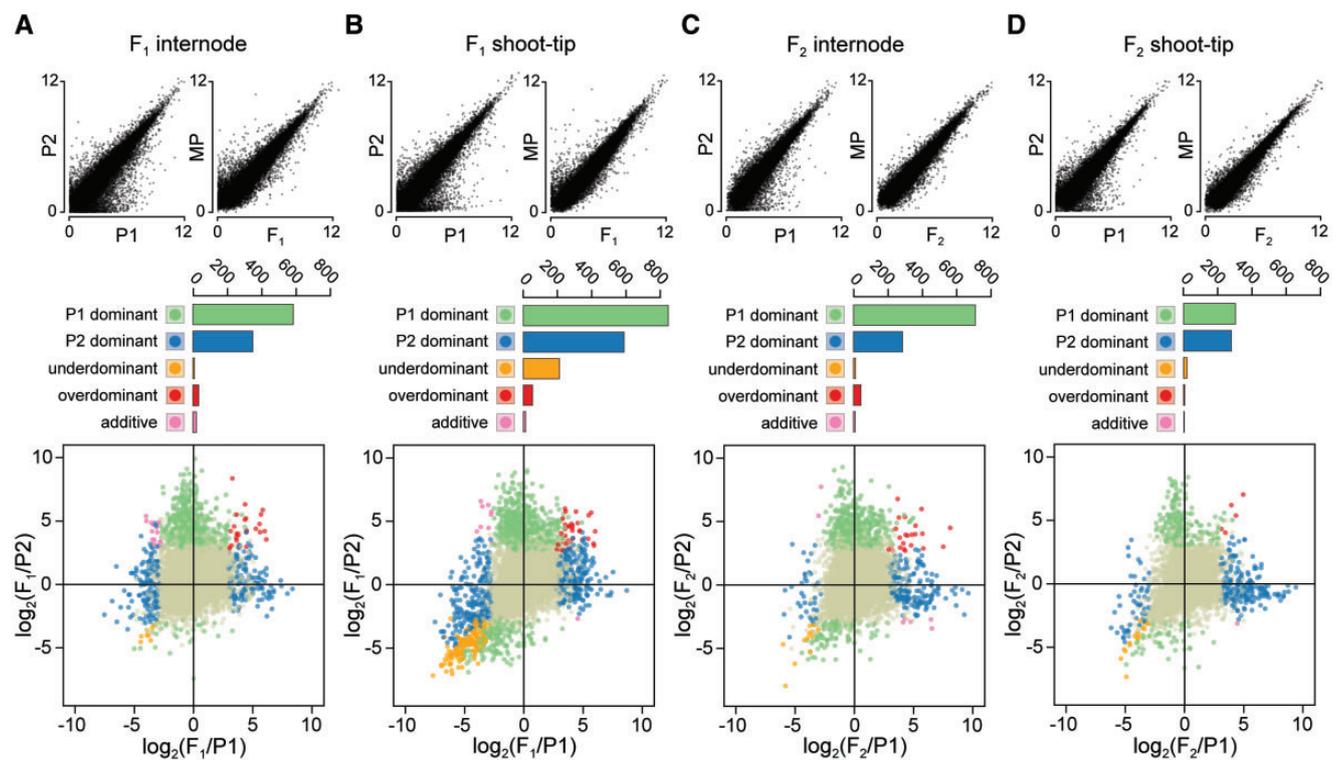
Fig. 1.—Inheritance of gene expression in intraspecific F_1 and F_2 *S. purpurea* families. Pairwise comparison of global gene expression between parents (P1 vs. P2) as well as the respective midparent to family progeny (MP vs. F_1 or F_2) for each tissue transcriptome. RNA-Seq count data were normalized using log_2 counts-per million mapped reads (cpm) in each library with a prior count of 1. Before normalization, rows with low expression (cpm $\leq$ 1.0) over 50% of the samples were removed from the analysis. Inheritance of gene expression is summarized in barplots for (*A*) F_1 internode, (*B*) F_1 shoot tip, (*C*) F_2 internode, and (*D*) F_2 shoot tip tissues, and color-coded with respect to classes. Conserved inheritance class (beige) are not shown (reported in table 1). Scatterplots below barplots for each tissue type within a family depict the ratio of log_2 normalized read counts of family progeny to the respective female (*x* axis) and male parent (*y* axis). Single points within each scatterplot represents unique genes colored according to inheritance classifications (same as boxplots).

F_1 progeny and the F_1 midparent. Likewise, the parents of the F_2 family had dramatically fewer differentially expressed genes than among the parents of the F_1, whereas gene expression in the F_2 was considerably more conserved than in the F_1. Regardless of the tissue type, <0.07% of all differentially expressed genes in the F_1 and F_2 could be classified as having an additive mode of expression inheritance. Among differentially expressed genes, the greatest proportion were classified as dominant (fig. 1); accounting for 95% and 84% of all nonadditive (i.e., dominant, overdominant, and underdominant) F_1 gene expression and 94% and 96% F_2 gene expression in the internode and shoot tip transcriptome, respectively (table 1).

Further, both the F_1 and F_2 families showed a significantly greater proportion of dominant expression biased in the direction of the maternal parent (P1 dominant). The most extreme case of maternal dominance was identified in the F_1 shoot tip transcriptome (845 P1-dominant genes) and F_2 internode transcriptome (711 P1-dominant genes). The total number of maternally dominant genes in common among the F_1 and F_2 shoot tip (99) or among the F_1 and F_2 internode (92) transcriptomes was relatively greater than paternally

dominant genes shared among the F_1 and F_2 shoot tip (27) or internode (13) transcriptomes. In addition to the extensive expression-level dominance present within these families, >10-fold the number of transgressive genes in the F_1 shoot tip transcriptome were classified as underdominant (209), compared with the F_2 shoot tip transcriptome (20). Likewise, there were fewer genes with overdominant expression in the F_2 shoot tip transcriptome (6) than in the F_1 shoot tip transcriptome (53).

There was a relatively large subset of genes differentially expressed between parents and progeny of the F_1 family (supplementary table S4, Supplementary Material online). Many of these genes were classified as transcriptional regulators and hypothetical proteins with domains of unknown function (DUF). Gene ontology (GO) term analysis for P1-dominant genes in internode tissues showed significant enrichment for response to stimulus, catalytic activity and binding, as well as the cellular components intracellular membrane-bound organelle and intracellular part. Significant GO enrichment for P1-dominant genes in the shoot tip was similar to those in the internode for biological processes and cellular components, but included unique molecular functions of

Table 1

Summary of Gene Expression Inheritance and Regulatory Divergence Classifications of *Salix purpurea* F$_1$ and F$_2$ Families

Class	F$_1$ Family 10X-082				F$_2$ Family 10X-317			
	Internode		Shoot Tip		Internode		Shoot Tip	
Expression Inheritance								
P1 dominant	**586**	**2.14%**	**845**	**3.10%**	**711**	**2.63%**	**303**	**1.13%**
P2 dominant	**350**	**1.28%**	**587**	**2.15%**	286	1.06%	279	1.04%
Overdominant	34	0.12%	53	0.19%	**43**	**0.16%**	**6**	**0.02%**
Underdominant	**10**	**0.04%**	**209**	**0.77%**	12	0.04%	20	0.07%
Additive	20	0.07%	13	0.05%	**10**	**0.04%**	**1**	**0.01%**
Conserved	26,441	96.4%	25,545	93.7%	25,953	96.1%	26,260	97.7%
Total	27,441		27,252		27,015		26,869	
Regulatory Divergence								
cis only	**145**	**2.80%**	**231**	**5.90%**	29	4.40%	41	6.30%
trans only	85	1.60%	187	4.80%	28	4.20%	15	2.30%
cis + *trans*	5	0.10%	13	0.30%	2	0.30%	1	0.20%
cis × *trans*	19	0.40%	140	3.60%	10	1.50%	6	0.90%
Compensatory	38	0.70%	512	13.0%	9	1.40%	18	2.80%
Ambiguous	833	16.1%	811	20.7%	129	19.4%	100	15.4%
Conserved	**4,063**	**78.3%**	**2,033**	**51.8%**	458	68.9%	467	72.1%
Total	5,188		3,927		665		648	

NOTE.—The total number and percentage of genes among those classified within the F$_1$ and F$_2$ shoot tip and internode transcriptome are partitioned by their inheritance and regulatory divergence classes (False Discovery Rate = 0.005). Numbers in boldface indicate significant ($P < 0.01$) deviations from a 1:1 ratio according to a χ^2 test.

catalytic and UDP-glycosyltransferase activity. Enrichment of P2-dominant genes in internode tissues included response to stimulus, transcription regulator activity, and cell wall components, whereas postembryonic development, programmed cell death, binding, and catalytic activity were overrepresented in shoot tip tissue.

ASE Analysis

In order to discern the overall proportion of *cis*- and *trans*-regulation of gene expression in *S. purpurea*, ASE tests were conducted using RNA-Seq expression data which was based on biallelic sites called from DNA-Seq and RNA-Seq of the parents. For both families, the regulation of gene expression was primarily conserved, regardless of the tissue type assayed (fig. 2). However, for those genes showing nonconserved regulatory classes, moderate proportions of pure *trans*-regulated gene expression were identified in the F$_1$ shoot tip (187, 11%) and internode (85, 14.5%) transcriptome, but substantially less in the F$_2$ shoot tip (15, 2.5%) and internode (28, 2.6%) transcriptome (table 1). The proportion of ASE in the F$_1$ shoot tip transcriptome was nearly twice that of the F$_1$ internode transcriptome, yet there was no difference among F$_2$ tissues. On average, the F$_1$ had a greater number of SNPs per gene than the F$_2$ (fig. 2; supplementary table S5, Supplementary Material online). As the number of SNPs per gene increased, the log$_2$ (P1/P2) expression ratio decreased. The only major discrepancy in the average number of SNPs per gene was between F$_1$ tissues for those genes showing *cis* + *trans* regulatory interactions.

A significantly greater number genes showing *cis*- and *trans*-regulation in the F$_1$ coincided with upregulation of the maternal P1 allele, irrespective of the magnitude, whereas the paternal P2 allele was higher expressed for genes showing *cis* × *trans* or compensatory regulation. Compensatory regulation accounted for 13% (512) of gene expression in the F$_1$ shoot tip transcriptome, but only 0.7% (38) in the internode transcriptome. Without considering conserved or ambiguous classes in the F$_1$ shoot tip transcriptome, 30% exhibited compensatory patterns and 8.3% with *cis* × *trans* regulation, compared with 13.5% and 11% of genes with pure *cis*- and *trans*-regulation, respectively. Further analysis of the F$_1$ revealed significantly greater levels of nonadditive and regulatory divergent expression in the shoot tip transcriptome compared with the internode transcriptome. Among the unpooled libraries of F$_1$ progeny individuals (supplementary fig. S2, Supplementary Material online), the degree of midparent differential expression corresponded linearly to the variation in regulatory divergent expression, whereby increased levels of compensatory expression coincided with increased levels of *cis* × *trans* regulation (supplementary fig. S3, Supplementary Material online). However, compensatory regulation was negligible in the F$_2$ family and did not show significant differences by tissue type.

Tissue-Biased Gene Expression

Individuals within the F$_1$ family were independently tested to investigate the average effect of tissue type. After library
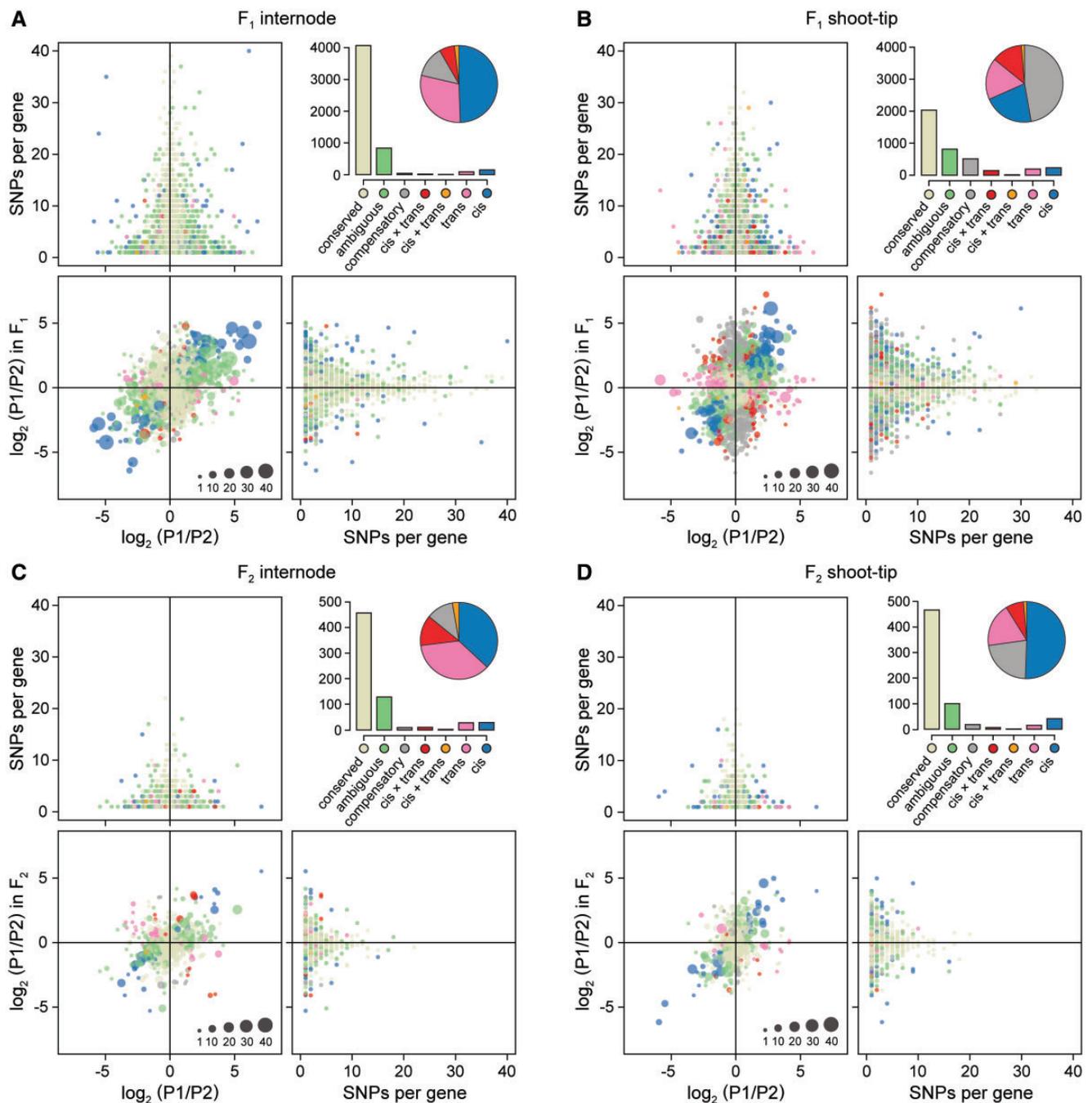
Fig. 2.—Allele-specific expression in intraspecific F_1 and F_2 *S. purpurea* families. Regulatory divergence classifications are summarized in barplots and proportions in pie charts for the (*A*) F_1 internode, (*B*) F_1 shoot tip, (*C*) F_2 internode, and (*D*) F_2 shoot tip tissues. Scatterplots in lower left panel for each tissue transcriptome within a family depict regulatory divergence as a ratio of allele-specific expression of the parents to that of F_1 and F_2 family progeny. For panels *A–D*, the number of SNPs is plotted against ASE ratios. Single points within each scatterplot represents unique genes colored according to respective regulatory classifications and scaled using the log number of SNPs per gene. Regulatory divergence assignments were based on binomial exact tests performed between the female parent (P1) and the male parent (P2) and Fisher's Exact test of the female and male parent alleles in the hybrid at an FDR global significance threshold of 0.005.

normalization and filtering for low expression, we identified a total of 262 genes as differentially expressed between the F_1 shoot tip and internode transcriptome. A subset of 46 genes were compiled (supplementary table S6, Supplementary

Material online) that were highly expressed in internode tissues with log_2 fold-differences ranging from 3.2 (SapurV1A.0130s0080) to 6.0 (SapurV1A.0216s0270). Genes encoding for fasiclin-like arabinogalactan (FLA)
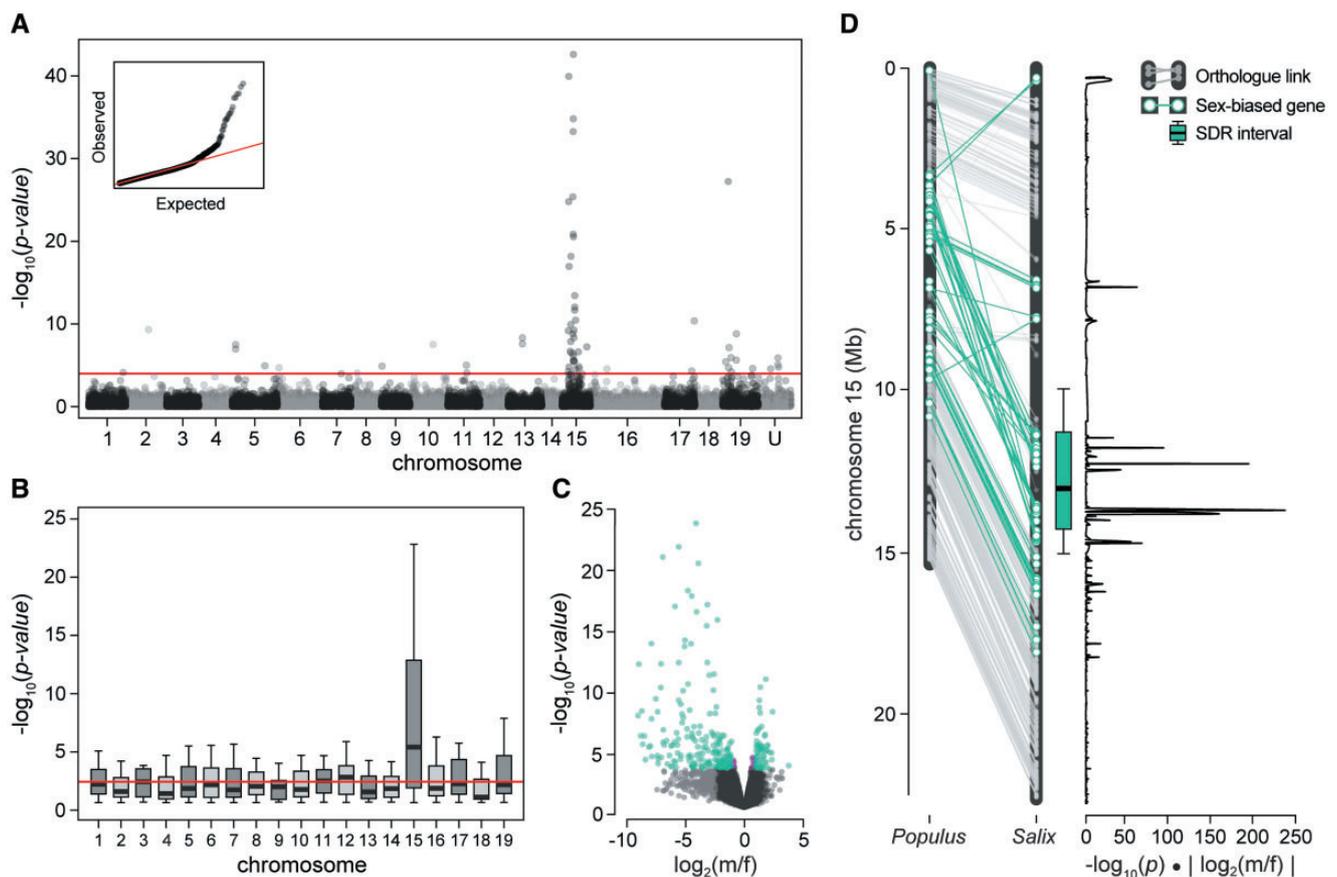
FIG. 3.—Sex-biased expression maps to *Salix purpurea* chr15. Significant sex-biased gene expression is depicted in the Manhattan plot (*A*), where each point represents the $-\log_{10}$ of the adjusted *P* value for each gene. Chromosomes are ordered from 1 to 19 (unplaced scaffolds not shown), the horizontal red line represents the global significance threshold, and QQ-plot showing model-fit in the upper left. Within each boxplot demonstrating the chromosomal distribution of differentially expressed genes (*B*), the solid black line represents the median $-\log_{10}(P)$ for each chromosome with whiskers extending an interquartile range of 1.5, and the median genome-wise $-\log_{10}(P)$ is represented by the red horizontal line. A volcano plot (*C*) depicts the magnitude of the $\log_2(m/f)$ change in gene expression (*x* axis) and the $-\log_{10}(P)$ significance (*y* axis) for sex, where negative values depict upregulated genes in females and positive values, upregulation in males. The magnitude of differential gene expression along *Salix purpurea* chr15 (*D*) is portrayed as the product of the $-\log_{10}(P)$ and absolute $\log_2(m/f)$ values. Lines connect orthologue pairs along *P. trichocarpa* (left) and *S. purpurea* (right) chr15 assemblies. For panels (*C*) and (*D*), teal points represent genes considered differentially expressed (False Discovery Rate < 0.05). The SDR interval boxplot was derived from mapping sex QTL within the F$_2$ *S. purpurea* family (n = 497).

proteins were most represented. A total of 20 FLAs from the 57 FLA gene family members were identified as differentially expressed (supplementary fig. S4, Supplementary Material online), all of which are annotated *FLA11* or *FLA12* with a single *FLA2* representative (SapurV1A.0054s0430). GO term analyses of internode predominant genes showed significant enrichment for cell wall and biosynthetic processes.

Sex-Biased Gene Expression

In order to test for differential expression between male and female *S. purpurea*, we utilized the shoot tip transcriptome of three male and three female F$_1$ individuals as well as their parents, such that each sex was represented by four related individuals. We identified a total of 315 genes in the F$_1$ shoot tip transcriptome as having significant sex-biased expression (fig. 3*A* and table 2). In stark contrast, there were no genes in the F$_1$ internode transcriptome that showed significant sex-biased expression. Of the 315 sex-biased genes, 62 map to *S. purpurea* chr15. In addition, 77 genes with best BLAST hits to *P. trichocarpa* v3 chr15 orthologues accounted for ~24% of the sex-biased genes identified (fig. 3*B*). From this list, 231 genes were more highly expressed in females and 84 were more highly expressed in males, indicating that a significantly higher proportion of the shoot tip transcriptome is female-biased than male-biased (fig. 3*C*). Nearly all female-biased genes reported in this study localize to *Salix* chr15 or unplaced scaffolds syntenic to *Populus* chr15. By contrast, male-biased genes were not predominantly on *S. purpurea* chr15, many mapped to *S. purpurea* chr19, especially for genes with higher

Table 2

Sex-Biased Gene Information

Salix purpurea v1 Gene Model	Chr or Scaffold	Functional Annotation	Homolog	log$_2$	−log$_{10}$(*P*)
Up in Females					
SapurV1A.0582s0010	582	NIPA, interacting partner of ALK	Potri.015G052400	6.7	42.6
SapurV1A.2504s0020	2504	AGL98, agamous 98-like	AT5G39810	7.7	39.9
SapurV1A.4040s0010	4040	Di-glucose binding, kinesin	Potri.001G436200	3.3	27.3
SapurV1A.0301s0070	15	REM1, reproductive meristem 1	Potri.009G103300	8.8	25.3
SapurV1A.2504s0010	2504	GPI-anchored protein	Potri.015G040900	6.0	24.6
SapurV1A.0301s0160	15	DR1/NF-YB, TBP-associated	Potri.015G052800	4.7	21.0
SapurV1A.0301s0170	15	BRCA, fragile-X-F-associated	Potri.015G050300	6.8	20.8
SapurV1A.0475s0170	15	peptidase M50B-like protein	Potri.015G045900	9.8	18.2
SapurV1A.1892s0010	15	MOS4/SPF27, modifier of SNC1	Potri.015G041800	5.1	17.0
SapurV1A.2524s0010	2524	PHYB, phytochrome protein B	Potri.008G105200	4.1	13.7
SapurV1A.2212s0030	2212	activating signal cointegrator 1, 3	Potri.015G056700	5.5	12.5
SapurV1A.4349s0010	4349	SCD1, cytokinesis-defective 1	Potri.015G049500	3.9	11.7
SapurV1A.1210s0090	1210	PME36, pectinesterase inhibitor 36	Potri.015G127700	6.9	11.6
SapurV1A.1386s0030	15	HOT101, heat shock protein 101	Potri.015G056900	3.1	10.3
SapurV1A.0530s0090	15	LRK10, serine/threonine kinase	Potri.015G044800	4.7	9.9
SapurV1A.0178s0110	15	18S pre-ribosomal, gar2-related	Potri.015G048400	5.9	9.2
SapurV1A.1146s0050	15	LP-1, thaumatin protein 1	Potri.015G039200	2.2	9.2
SapurV1A.0301s0080	15	CaS, extracellular Ca^{2+} receptor	Potri.015G052200	3.2	8.8
SapurV1A.0107s0110	3	UBC2/RAD6, Ub conjugating E2, 1	Potri.013G064400	4.0	8.8
SapurV1A.0582s0060	582	GUS2/HPSE1, heparanase 1-like	Potri.015G049100	5.0	8.3
SapurV1A.0530s0130	15	RTNLB9, reticulon-like protein	Potri.015G044300	7.9	7.8
SapurV1A.2212s0020	2212	activating signal cointegrator 1, 3	Potri.015G056500	5.0	7.6
SapurV1A.0107s0070	3	IMPA-2, importin alpha 2	Potri.005G020400	4.5	7.5
SapurV1A.0107s0060	3	GATA Znf protein	Potri.005G020500	2.8	7.0
SapurV1A.1538s0020	15	TCP-1/cpn60, delta chaperonin	Potri.015G042600	5.6	6.7
SapurV1A.1002s0030	15	WOX5, wuschel-related homeobox 5	Potri.015G065400	2.9	6.6
SapurV1A.0530s0070	15	delta-ADR, AP-3 complex delta-1	Potri.015G045600	2.5	6.5
SapurV1A.1254s0040	19	AMY-1, associate of c-MYC	Potri.019G014100	7.1	5.9
SapurV1A.2772s0010	15	RPL19e/EMB2386, ribosomal 19e	Potri.015G037100	4.7	5.9
SapurV1A.0582s0100	582	LAP4, less-adhesive pollen 4	Potri.015G048800	5.2	5.6
SapurV1A.0582s0040	582	meiotic endonuclease, putative	Potri.015G049800	6.7	5.5
SapurV1A.2535s0010	2535	suppressor of protein silencing	Potri.018G137400	6.6	5.4
SapurV1A.0178s0160	15	RING/U-box Znf protein	Potri.015G047900	2.6	5.4
SapurV1A.1596s0050	15	DYNLL1, dynein light chain 1-like	Potri.015G067800	4.6	4.8
SapurV1A.0307s0060	19	PIF1, phytochrome interacting 1	Potri.008G203700	7.3	4.6
Up in Males					
SapurV1A.0934s0010	15	RPS3, 40S ribosomal protein S3-1	Potri.015G071700	1.9	6.4
SapurV1A.0830s0010	830	NIPA, interacting partner of ALK	Potri.015G052400	1.2	6.1
SapurV1A.0934s0060	15	DR1/NF-YB, TBP-associated	Potri.015G052800	1.3	5.5
SapurV1A.3555s0010	3555	CLO1-2, caleosin 1, seed gene 1	Potri.010G066600	1.3	5.1
SapurV1A.1765s0050	1765	fertility restorer (Rf)	Potri.015G036400	1.4	4.8
SapurV1A.0530s0040	15	peptidase M50B	Potri.015G045900	1.3	3.9
SapurV1A.1246s0030	15	transmembrane protein	AT3G18215	1.3	3.9
SapurV1A.1510s0020	15	GPI-anchored protein	Potri.015G040900	1.1	3.8
SapurV1A.0391s0140	19	ankyrin repeat, SAM domain 1	Potri.019G106200	6.7	3.8
SapurV1A.0704s0100	15	TB2/DP1, HVA22 family protein	Potri.015G062800	1.8	3.8
SapurV1A.0391s0170	19	ankyrin repeat protein	Potri.019G107700	6.3	3.6
SapurV1A.1421s0010	15	NOF1/Utp25, nucleolar factor 1	Potri.003G010000	5.0	3.5
SapurV1A.1515s0010	15	CAAX amino terminal protease	Potri.019G101100	2.7	3.5

NOTE.—Rows within the table are ordered by the −log$_{10}$(*P* value) significance of each respective gene. *Salix* v1 homologs are reported as the best BLAST hit (*E*-value ≤ 0.01) to the *Populus trichocarpa* v3 or Arabidopsis TAIR v10 proteome.

expression in males. Over half of the sex-biased genes on *S. purpurea* chr19 which were highly expressed in males encode proteins related to signaling and response (e.g., ankyrin repeat, patatin phospholipase, and nucleotide-binding site leucine-rich repeat proteins), but do not appear to localize to any particular region of the chromosome.

There were two primary gene clusters on *S. purpurea* chr15 that were significantly enriched for sex-biased expression span chr15: 11.4–12.3 Mb and chr15: 13.6–14.6 Mb (fig. 3D); the latter with more sex-biased genes and with higher significance than the former. Differentially expressed genes located on unplaced *Salix* scaffolds 265, 582, 830, 1765, 2212, 2504, and 4349 align to *S. purpurea* chr15 gene clusters and corresponding regions on *Populus* chr15. For those genes located on unplaced *S. purpurea* scaffolds and chr15 which have high identity to *Populus* chr15 homologs, when ordered according to *Populus* chr15 positions, a substantial number of those genes appear to be duplicated within one or the other cluster.

Sex determination regions of dioecious species are often highly polymorphic, so we investigated the presence and absence of gene expression specific to F_1 females or males. Complete absence of expression in males was observed for *REM1* (SapurV1A.0301s0070), reticulon *RTNLB9* (SapurV1A.0530s0130), peptidase *M50B* (SapurV1A.0475s0170), chaperonin *TCP-1* (SapurV1A.1538s0020), *PMEI1* (SapurV1A.1210s0090), a gar2-related 18 S preribosomal assembly protein (SapurV1A.0178s0110), terpene synthase *TPS21* (SapurV1A.1522s0030), and *AGL98* (SapurV1A.2504s0020); all of which are located within or align to the pericentromeric SDR on *S. purpurea* chr15. Complete loss or very low-levels of gene expression in females was accompanied by correspondingly low-levels in males; nearly 10% of gene models were filtered from analyses for this reason.

Sex-biased genes highly expressed in females were enriched for GO terms in the biological processes of signaling, signal transmission and transduction, cation binding, and ion binding, as well as the molecular functions of copper ion binding, magnesium ion binding, signal transducer activity, and lyase activity. Genes showing higher expression in males were enriched for cell death, death, apoptosis, and programmed cell death, the molecular functions of ATP-binding, structural constituent of the ribosome, and structural molecule activity, and the cellular components of intracellular organelle, ribonucleoprotein complex, and ribosome.

Sexually Dimorphic Inheritance of Gene Expression

Although we were not able to identify sex-biased expression in the F_1 internode transcriptome, considerable variation within the shoot tip transcriptome offered us a unique opportunity to dissect the heritable components of sexually dimorphic patterns of expression. Genes considered to exhibit

sexually dimorphic inheritance were only reported for those with a significant nonadditive or transgressive inheritance class for at least one sex. Genes with expression inheritance classifications that did not show significant sex dimorphism were classified as having same-sex inheritance. Although a majority of the shoot tip transcriptome retained same-sex inheritance, 3.8% (1,055 genes) displayed sexually dimorphic patterns of inheritance (fig. 4, supplementary table S7, Supplementary Material online). While there were no significant differences in the median expression level for genes with same-sex inheritance, the expression levels of dimorphic genes were significantly greater in females than in males (fig. 4A). Broadly, sexually dimorphic inheritance in the F_1 shoot tip transcriptome was associated with a greater number of genes with conserved expression in males (65%) (fig. 4B) and nonadditive expression in females (75%) (fig. 4C). In addition, for those genes with sexually dimorphic inheritance, there was a marginally greater number of P2-dominant genes (174) in males compared with P1-dominant genes (148), whereas the opposite was found in females. Nearly five-times the number of dimorphic genes in females were classified as P1-dominant (499) compared with P2-dominant (119).

The most drastic instances of sex dimorphism were for genes with transgressive (overdominant and underdominant) expression inheritance. Of the 175 overdominant genes in females, 167 had conserved expression in males, with only eight showing same-sex inheritance (supplementary table S7, Supplementary Material online). Further, of the 33 genes with an underdominant mode of expression inheritance, all were restricted to males. Relative to the results of the pooled F_1 family inheritance classifications (table 1), additive inheritance plays a negligible role in both the female (7) and male (1) shoot tip transcriptome.

In order to quantify the extent of dimorphism for each gene, the magnitude of dimorphic expression was calculated as the Euclidean distance between male and female vectors on the same Cartesian plane of expression levels (supplementary fig. S5, Supplementary Material online). The direction and magnitude of dimorphic gene expression in the F_1 shoot tip transcriptome is illustrated in figure 4D. The contribution of nonadditive inheritance on dimorphic gene expression was examined directly, because gene expression inheritance in an individual is simply a function of the difference in progeny and parent expression ratios. This method demonstrates that there was a positive relationship between the magnitude of dimorphic inheritance and midparent differential expression. In F_1 males, sexually dimorphic genes with conserved inheritance exhibited significantly greater differences in the magnitude of differential expression than genes with nonadditive expression. However, dimorphic genes with conserved inheritance in females (and nonadditive inheritance patterns in males) exhibited a significantly lower magnitude of expression than genes with nonadditive expression (fig. 4E). Likewise, the absolute magnitude of genes showing conserved expression
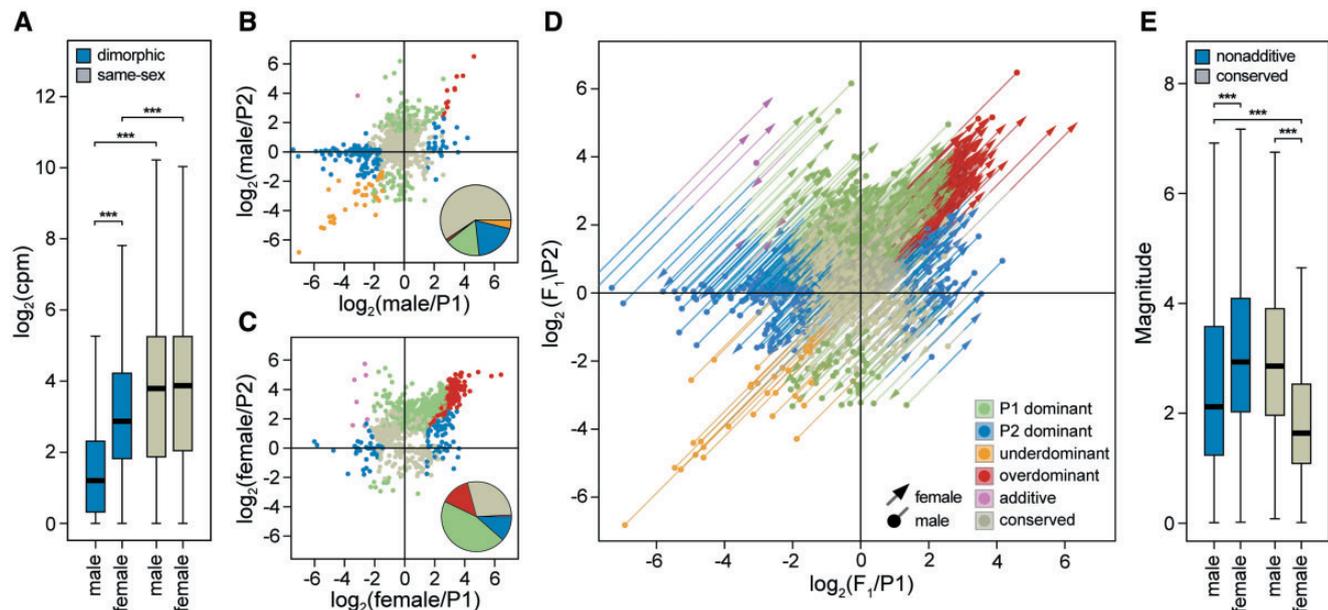
FIG. 4.—Sexually dimorphic inheritance in the F_1 shoot tip transcriptome. Boxplots (A) summarize the $\log_2$ normalized expression differences for genes with sexually dimorphic inheritance patterns (teal) and those with same-sex inheritance (beige), by sex. Asterisks above boxplots represent significant differences (Wilcoxon $P < 0.001$). Scatterplots compare $\log_2$ normalized expression of F_1 males (B) and females (C) to the maternal (P1, x axis) and paternal (P2, y axis) expression. Points represent only genes with dimorphic inheritance patterns (same-sex inheritance not shown). Pie charts within the scatterplots summarize patterns of gene expression inheritance for genes with dimorphic gene expression for each sex. The scatterplot (D) illustrates overlain coordinates of gene expression inheritance for males and females, where each gene is represented by two vectors, one male (m_{xy}, points) and one female (f_{xy}, arrows), connected by a single line segment. Each segment is equally divided by two colors which correspond to the male and female inheritance class for each gene. The magnitude of dimorphic gene expression inheritance was calculated for each gene as the absolute Euclidean distance between the vectors, m_{xy} and f_{xy}, on the same Cartesian plane. For those genes with dimorphic inheritance, boxplot distributions (E) of nonadditive (blue) and conserved (beige) inheritance patterns for males and females depict differences in their absolute magnitude.

in males was not significantly different from the magnitude of nonadditive expression in females, signifying that a higher frequency of sexually dimorphic genes with conserved inheritance in males coincided with nonadditive gene expression in females. Simply, as the distance between female and male vectors (magnitude) increased, the frequency of nonadditive and conserved patterns of inheritance increased, respectively.

Sexually Dimorphic ASE

To test the hypothesis that sex dimorphic gene expression was attributable differences in regulation, we compared the magnitude of *cis*- and *trans*-regulation of genes expressed predominantly in either females or males. In order to determine the extent to which patterns of regulatory divergent expression varies for the same gene, we contrasted the magnitude and direction of regulatory divergence and the proportion of their effects, based solely on sex in the F_1. Sexually dimorphic ASE was considered to be the total number of genes that significantly differ by regulatory class between three F_1 male individuals and three F_1 female individuals. Genes assigned regulatory classes which did not differ between males and females were characterized as having same-sex ASE. For each gene showing significant regulatory divergence in either

sex, these results illustrate the total number of genes that show contrasting or sexually dimorphic regulatory patterns. Only genes from the ASE analysis which were unique to either females or males were considered.

In the shoot tip transcriptome, a total of 297 genes displayed sexually dimorphic ASE (table 3 and fig. 5), where 143 had higher expression levels in F_1 males and 154 were expressed greater in F_1 females. For genes showing sexually dimorphic ASE, conserved or *cis*-regulation in females coincided with *cis* × *trans* or compensatory regulation for the same genes in males. For instance, there were significantly greater numbers of dimorphic genes expressed in the male internode transcriptome, with 54 *cis* × *trans* and 88 compensatory regulated genes, compared with only 4 *cis* × *trans* and 9 compensatory regulated genes uniquely expressed in females (table 3). By comparing ASE of female and male F_1 progeny individuals, genes with significant *cis*-regulation were significantly biased towards maternal parent (P1) in the shoot tip and internode transcriptome of females, as well as the shoot tip transcriptome of males, irrespective of the magnitude of ASE (table 4). On a per-site basis, antagonistic *cis* × *trans* and compensatory regulation in the F_1 shoot tip transcriptome was driven by upregulation of the paternal (P2) allele, whereas the maternal (P1) allele was upregulated for

Table 3

Sexually Dimorphic Patterns of Regulatory Divergence

	cis	*trans*	*cis + trans*	*cis × trans*	Compensatory	Ambiguous	Conserved
Internode							
dimorphic							
males	67	60	2	**54**	**88**	142	**39**
females	46	52	2	**4**	**9**	171	**168**
same-sex	50	62	2	17	22	698	4008
Shoot tip							
dimorphic							
males	23	35	2	17	14	**162**	45
females	46	42	3	16	12	**89**	89
same-sex	29	31	0	6	3	757	3857

Note.—Dimorphic ASE was considered as genes within regulatory divergence classifications (FDR $= 0.005$) that differ between males and females, and same-sex ASE were those with the same regulatory classifications. Numbers in boldface indicate significant ($P < 0.01$) deviations from a 1:1 ratio according to a χ^2 test.

either pure *cis*- or *trans*-regulation or *cis + trans* regulation (fig. 5*B*).

The absolute magnitude of *trans*-regulated genes was greater than that of *cis*-regulated genes for both males and females (fig. 5*C*). Although the magnitude of effect class sizes of genes that showed pure *cis*- or *trans*-regulation were not significantly different by sex, a significantly greater frequency of *cis*-effects were present at the tails (<0.5 and >2-fold) and a greater frequency of *trans*-effects between tails (>0.5 and <2-fold) (fig. 5*D*). By scaling *cis*- and *trans*-regulatory divergence by the total regulatory divergence, as in Coolon et al. (2013), we obtained a relative percent divergence due to *cis*- and *trans*-effects. With regard to females, the proportion of genes with evidence of *cis*-regulatory divergence was significantly greater than was observed in males. The proportion of genes showing significant *trans*-regulatory divergence was greater in females than in males, but with weaker significance than the *cis*-regulated genes (fig. 5*E*).

Discussion

Dominance and Regulatory Divergence in F$_1$ and F$_2$ *S. purpurea*

The majority of gene expression inheritance and ASE studies in both plants and animals have focused on F$_1$ hybrids generated from crossing stable, inbred parents. Unlike what has been described in model crop plants, like maize (Stupar and Springer 2006) and rice (Song et al. 2013), we show that in *S. purpurea*, the greatest proportion of differentially expressed genes did not exhibit a primarily additive mode of inheritance, but rather showed strong patterns of dominance. Preference of uniparental expression in progeny is thought to be orchestrated by epistatic interactions, which primarily function to silence one of the parental alleles in a parent-of-origin manner (Chen and Pikaard 1997; Stupar et al. 2007). Here, maternal dominance represented the greater proportion of the nonadditive gene expression in both tissue transcriptomes of the

F$_1$ and F$_2$ families. Other cases of expression-level dominance has been described in hybrids of intraspecific thistle (Bell et al. 2013), interspecific coffee (Combes et al. 2015), synthetic allotetraploid rice (Xu et al. 2014), as well as allotetraploid Arabidopsis (Shi et al. 2012). Reciprocal hybridization has become a useful technique to examine genomic imprinting by comparing common patterns of uniparental expression of alleles in reciprocal family progeny (Donoghue et al. 2014; Baldauf et al. 2016). In developing seeds of Arabidopsis, there is strong evidence that imprinting genes regulate early endosperm development and nutrient translocation and partitioning in the seed (McKeown et al. 2011). While epistasis may well contribute to the prodigious levels of dominant gene expression observed in F$_1$ and F$_2$ *S. purpurea*, since *Salix* spp. are dioecious, reciprocal crosses cannot be generated in the classical sense.

The inheritance and regulatory patterns described in this study were most similar to what was reported in hybrids derived from heterozygous parents that were collected from natural *C. arvense* populations (Bell et al. 2013). For instance, for both F$_1$ *S. purpurea* and F$_1$ *C. arvense*, more divergently expressed genes showed higher expression of the maternal allele than the paternal allele and that a significantly greater proportion of dominant cases had maternal expression patterns. Importantly, there were similar trends between *S. purpurea* and *C. arvense* in *cis*- and *trans*-divergence, such that *trans*-divergence was associated with higher expression of the paternal allele, whereas *cis*-divergence tended to increase expression of the maternal allele. However, there were differences in transgressive inheritance classes between the species. The frequency of overdominant gene expression in *S. purpurea* (shoot tip) was significantly less than underdominant expression, whereas the opposite was detected in *C. arvense*.

The transcription of two divergent parental alleles in the hybrid can be controlled by *cis*-regulatory elements as well as *trans*-acting factors, but parental expression differences may be reduced in the hybrid for genes subject to strong
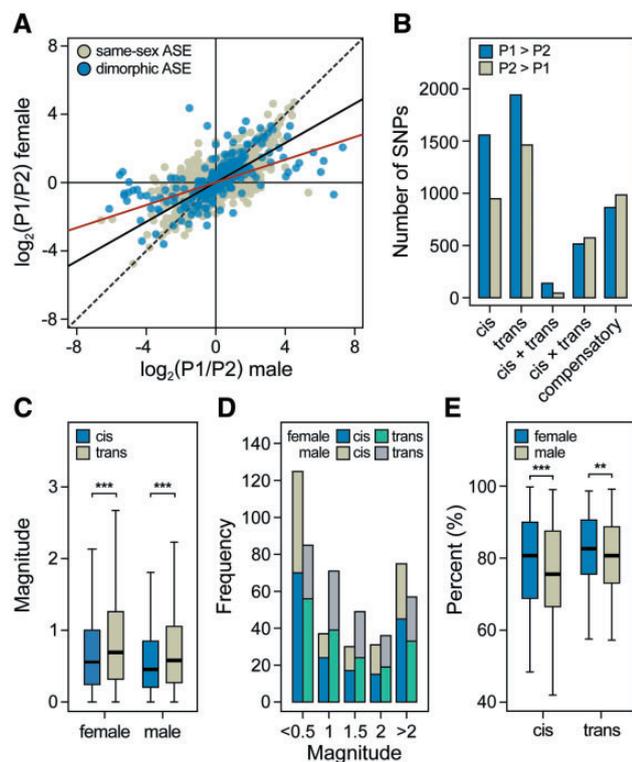
FIG. 5.—Sexually dimorphic regulatory divergence in the F_1 shoot tip transcriptome. The ratio of expression of parent alleles (A) is plotted for females (y axis) versus males (x axis) showing dimorphic ASE (blue points) and same-sex ASE (beige points) of the shoot tip transcriptome. The solid red line and solid black line represent the slopes of dimorphic ASE and same-sex ASE, respectively, where the black dashed line has an intercept = 0 and slope = 1. The barplots (B) explore the cases of parental dominance in hybrid regulatory patterns. Boxplots (C) show the distribution of the absolute differences between males and females for *cis* (blue) and *trans* (beige) for genes showing dimorphic ASE. Asterisks ***, ** above boxplots denote significant differences at a Wilcoxon $P < 0.001$ and $P < 0.01$, respectively. Barplot (D) distributions of genes with significant *cis*- and *trans*-regulation were binned according to the magnitude of their effect class size for females (blue and teal, respectively) and males (beige and grey, respectively). Boxplots (E) summarize the percent divergence due to *cis*- and *trans*-effects, where *cis* = $\log_2$(P1H/P2H) and *trans* = $\log_2$(P1H/P2H) − $\log_2$(P1/P2), and % *cis* = [|*cis*|/(|*cis*| + |*trans*|)] × 100 and % *trans* = [|*trans*|/(|*cis*| + |*trans*|)] × 100.

trans-regulation (Xu et al. 2014). Allele-specific tests in hybrids of inbred maize (Stupar and Springer 2006) established that *cis*-acting regulatory variation accounted for the majority of the observed parental expression divergence and that pure *cis*-regulation correlated with additive expression patterns in the F_1 hybrid. In synthetic allotetraploids of *indica* and *japonica* rice subspecies, parental expression differences were found to be intensified in the allotetraploid but showed a reduction of expression divergence in reciprocal F_1 hybrids, owning to the effects of a common *trans* environment on divergent *cis*-factors (Xu et al. 2014). Using allotetraploids of *Arabidopsis thaliana* as well as resynthesized allotetraploids

Table 4

Number of Pure *cis*- and *trans*-Regulated Genes in Male and Female Tissues

	Shoot Tip				Internode			
	cis		*Trans*		*cis*		*trans*	
	P1	P2	P1	P2	P1	P2	P1	P2
Males								
All genes	40	16	37	34	66	51	68	54
>1.5-fold	**31**	**10**	1	4	**48**	**36**	2	5
>2-fold	**21**	**7**	1	1	**34**	**29**	0	4
Females								
All genes	64	25	46	36	**79**	**37**	71	50
>1.5-fold	**47**	**10**	1	2	**59**	**23**	3	1
>2-fold	**38**	**5**	1	1	**49**	**18**	0	1

NOTE.—Numbers in boldface indicate significant ($P < 0.01$) deviations from a 1: 1 ratio according to a χ^2 test.

of *A. arenosa*, Shi et al. (2012) showed that the higher level of sequence divergence (and dominance) in *A. arenosa* promote flexibility of *trans*-factors for their binding to interacting factors and *cis*-elements of *A. thaliana* and *A. arenosa* alleles. In F_1 and F_2 *S. purpurea*, both *cis*- and *trans*-effects account for nonadditive or transgressive gene expression. These data also show significantly greater *cis* × *trans* and *cis*-*trans*-compensatory regulation in the F_1 shoot tip transcriptome compared with the internode transcriptome, fitting the model in which greater transgressive inheritance tends to show greater proportions of *cis* × *trans* regulatory divergence (McManus et al. 2010). Further, for genes showing both *cis*- and *trans*-effects, antagonistic *cis* × *trans* and compensatory interactions were driven by the up-regulation of the paternal allele. The compatibility of novel *trans*-factors with their target binding sites could explain the high proportion of expression-level dominance in *S. purpurea*, which suggests that compensatory interactions and stabilizing selection may play an important role in maintaining parental gene expression levels.

The Fasciclin Gene Family Is Highly Expressed in the F_1 Internode

Fasciclin-like AGPs were found to be over-represented and highly expressed in stem internode tissue, whereas genes encoding defense-related proteins were primarily upregulated in shoot tip tissue. Found in most angiosperms, AGPs are a class of Hyp-rich glycoproteins that are highly abundant in the cell wall and plasma membrane which are involved in various aspects of plant growth (Kitazawa et al. 2013) and primarily function in cell adhesion (Johnson et al. 2003). In *Populus* and *Eucalyptus*, FLAs were highly expressed during xylem differentiation (Hefer et al. 2015) and implicated in secondary cell wall thickening (Lafarguette et al. 2004). *Salix FLA11* and *FLA12* gene models are closely related to Arabidopsis *FLA11*

and *FLA12* homologs and have been shown to contribute to stem tensile strength and the modulus of stem elasticity (MacMillan et al. 2010, 2015), which is consistent with FLA roles in cellulose deposition during secondary xylogenesis. These data suggest that constitutive defense in the shoot tips of *S. purpurea* is coordinated with the rapid development of secondary xylem in stem internodes. Functional analysis of FLA homologs in *Salix* may prove to be useful in characterizing genes involved in the regulation of stem strength for the genetic improvement of shrub willow bioenergy crops.

Sexually Dimorphic Gene Expression

The effect of sex may explain a significant amount of the variation driving the evolution of gene expression in dioecious plants. While differential expression in the F_1 family shoot tip transcriptome was found to be almost entirely nonadditive, the primary contributors of nonadditive expression were only discernable by barcoding and sequencing the F_1 progeny individuals. This study provides evidence of sexually dimorphic expression in intraspecific F_1 *S. purpurea*, where both *cis*- and *trans*-effects accounted for the observed differences in magnitude among regulatory patterns; however, the effect of sex on gene expression in the F_1 was tissue-specific. Although allelic effects were comparable, Meiklejohn et al. (2014) described sexually dimorphic regulatory divergence among *Drosophilia simulans* and *D. mauritiana* introgression hybrids, proposing that pure-species genotypes carry modifier alleles that increase sexually dimorphic expression.

Although *cis*-effects accounted for more of the regulatory divergent expression in F_1 females, compensatory regulation was enriched in F_1 males. It may be the sex determining system itself can help explain sexually dimorphic ASE in *S. purpurea*. In females, the Z haplotype is paternally (ZZ) inherited and the W is maternally (ZW) inherited, but in males, both the maternal and paternal Z haplotypes are inherited equally. If there is substantial divergence between the Z and W, theoretically, *cis*-effects should outweigh *trans*-effects in females. In contrast, the two Z haplotypes present in males are considerably less polymorphic, such that parental expression differences on the Z would be explained by more antagonistic *trans*-regulatory interactions.

It is possible that a subset of the sexually dimorphic genes in F_1 *S. purpurea* actually reflect unresolved conflicts between females and males and that sex-biased gene expression on autosomes or in pseudo-autosomal regions could be due to sexual antagonism (Alström-Rapaport et al. 1997; Nagamitsu and Futamura 2014; Su et al. 2016; Zhai et al. 2016). The accumulation of sexually antagonistic loci are predicted to occur in newly formed SDRs, where differential selection via tight linkage is beneficial to one sex and harmful to the other (Rice 1992). Over time, accumulation of sexually antagonistic genes within a population may eventually lead to a considerable conflict between sexes such that adaptation by each sex

would be compromised. As female-benefit and male-detriment genes accumulated, Rice (1992) found that the sex ratios (m/f) of Drosophilia declined, suggesting that pseudo-autosomal regions (PARs) near the SDR can act as a hot spot for the accumulation of genes detrimental to the homogametic sex.

There are a number of studies which have reported consistent 2: 1 (f: m) sex ratio-biases in natural willow populations (Ueno et al. 2007; Che-Castaldo et al. 2015). Disproportionate sex-ratios in natural populations of *Salix* spp. could promote evolutionary biases by sustaining the regulatory roles of the predominant sex. The higher proportion of genes with female-specific expression in *S. purpurea* may be a consequence of cyclic asexual reproduction leading to a relaxation of purifying selection on male-biased genes, whereby conflicting modes of inheritance and regulatory divergence patterns could lead to unequal investments in reproduction and reproductive strategies (Parsch and Ellegren 2013). However, sex-biased gene expression in *S. purpurea* may reflect resolved conflicts in favor of females.

Sex-Biased Expression Localizes to the SDR of *S. purpurea* chr15

The genus *Salix* exhibits a ZW sex determination system (Alström-Rapaport et al. 1997; Gunter et al. 2003; Semerikov et al. 2003; Liu et al. 2013), where the female is heterogametic (ZW) and the male is homogametic (ZZ), in contrast to the male heterogametic XY system of *P. trichocarpa* (Yin et al. 2008; Geraldes et al. 2015). Although the genomes are rather collinear, is not entirely clear whether the SDR developed before or after the Salicoid duplication (Rodgers-Melnick et al. 2012) and divergence of *Salix* from *Populus* (Hou et al. 2016). We have delimited the SDR of *S. purpurea* to a centromeric region on chr15 (Zhou et al. 2017) using a full-sib F_2 *S. purpurea* mapping population (described here) and a diverse panel of *S. purpurea* naturalized North American genotypes. Zhou et al. (2017) show in *S. purpurea* that females are heterozygous (or hemizygous) and are males homozygous at polymorphic sites within the coding sequences of genes found in the SDR of *S. purpurea* chr15. As low recombination near the *Salix* SDR has prevented the identification of a single gene responsible for sex determination in experimental mapping populations (Pucholt et al. 2015; Chen et al. 2016; Pucholt et al. 2017), increasing the overall experimental scale is likely required to pin-point the causative sex-determining genes.

Of the previous gene expression studies on sex determination in *Salix*, none have yet identified the gene(s) responsible. Even though a relatively small proportion of genes in the F_1 shoot tip transcriptome of *S. purpurea* are sex-biased (<0.1%), there was no substantial evidence of sex-biased expression in the F_1 internode transcriptome. At the time shoot tips were collected, it was likely that floral buds were

developing near the shoot apical meristem and it is possible this amalgam of cells included a subset of that were expressing genes involved in the determination of male or female flowers. Genes differentially expressed among the F_1 male and female shoot tip transcriptome are not necessarily indicative of sex determination alone, but altogether represent a cascade of developmental gene expression involved in patterning, signaling, and organ suppression during the vegetative-to-reproductive transition leading to sexual dimorphism (Fairbairn and Roff 2006).

Tests for differential expression between mature catkins of female and male willows likely confound the search for sex determining genes because there are profound morphological and phenological differences between females and males at floral maturity. For instance, RNA-sequencing of male and female catkins of *S. suchowensis* (Liu et al. 2013) and *S. viminalis* (Pucholt et al. 2017) identified a plethora of differentially expressed genes, but failed to establish any biological links among those with significant associations. Yet, a link between sex determination and meristem fate is well-described in oil palm (Ho et al. 2016) and in maize via RNA-induced silencing of *TS1* and *AP2* by miRNA172 (Hartwig 2011). Genes related to DNA methylation, *MET1* and *DDM1*, as well as SAUR-like auxin responsive genes were implicated in sex determination in andromonecious *Populus tomentosa* (Song et al. 2013) and *P. trichocarpa* (McKown et al. 2017). In *Asparagus officinalis*, the MYB-like transcription factor, *MSE1*, is specifically expressed in males and has been shown to induce male sterility in knockouts of Arabidopsis (Murase et al. 2017). The identification of *MeGI*, an autosomal homeobox transcription factor in *Diospyros lotus*, has been shown to dominantly suppress male organ development can be suppressed by the small RNA *OGI* on the Y chromosome that targets *MeGI* for gene silencing (Akagi et al. 2014).

The most significant sex-biased gene highly expressed in males ($-\log_{10}(P) = 6.4$) was most similar to the 40S ribosomal subunit, *RPS3* (SapurV1A.0582s0010), a positive regulator of apoptosis during UV irradiation in Arabidopsis (Lee et al. 2010). Conversely, the most significant upregulated gene in females ($-\log_{10}(P) = 42.6$) was most similar to an inhibitor of apoptosis in Arabidopsis, a C3CH zinc finger homolog of *NUCLEAR INTERACTING PARTNER OF ALK* (*NIPA*) (SapurV1A.0934s0010). Developmental requirements controlling the vegetative-to-reproductive transition in meristematic tissues are likely to differ by sex. The Arabidopsis *ASSOCIATE OF C-MYC* (*AMY1*) homolog in *S. purpurea* (SapurV1A.1254s0040), was highly expressed in females and has been shown to interact with numerous chromatin modifiers and transcription factors in Arabidopsis (Taira et al. 1998) and implicated as a universal amplifier of gene expression, acting to increase output at all active promoters. Programmed cell death (PCD) in developing floral buds could lead to the specification of sex at the cost of structural reproductive components integral to the opposite sex. If PCD is a primary component in the determination of male or female flowers in *S. purpurea*, both the greater numbers and magnitude of differential gene expression in females could indicate that the role of PCD is more active in the development of female flowers.

REPRODUCTIVE MERISTEM 1 (*REM1*) (SapurV1A.0301s0070), *AGAMOUS-LIKE 98* (AGL98) (SapurV1A.2504s0020), and *DOWN-REGULATOR OF TRANSCRIPTION 1* (*DR1*) (SapurV1A.0301s0160) were highly expressed in F_1 female shoot tip tissues and among the top differentially expressed genes that encode for known flowering-time genes in Arabidopsis (Pagnussat et al. 2005). The *REM1* gene aligns to a region within the SDR of *S. purpurea* chr15 and highly expressed in females, but showed complete null expression in males. In maturing inflorescences of Arabidopsis, *REM1* expression localized to only a few vegetative cells in the shoot apical meristem, but during the vegetative-to-reproductive transition, *REM1* was progressively restricted to the gynoecium which gives rise to the stigma, style, and septum (Franco-Zorrilla et al. 2002). Although SapurV1A.1250s0040 is paralogous to SapurV1A.0301s0070, the former is highly expressed in males and is homologous to *A. thaliana* AGO9. Implicated in the vegetative-to-reproductive transition during male gametogenesis in *A. thaliana*, the primary role of *AGO9* is to silence transposable elements (TEs) in the female gametophyte; thereby establishing the transgenerational epigenetic information required to control gametophytic fate (Hernandez-Lagana et al. 2016). Thus, it is conceivable that there is sRNA-induced silencing of W-specific genes via AGO loading in the RdRM pathway.

We have identified *TCP-1* (SapurV1A.1538s0020) as one of the few sex-biased genes that exhibited complete null expression in male shoot tip. TCP transcription factors play pivotal roles in the control of shoot morphogenesis by negatively regulating the expression of boundary-specific genes (Koyama et al. 2007; Koyama et al. 2010; Li et al. 2012), such as suppression of secondary wall thickening of the anther endothecium in Arabidopsis (Wang et al. 2015). Geraldes et al. (2015) identified two *TCP-1* chaperonin family cpn-60 proteins associated with sex determination in the T52 *P. trichocarpa* association population, and this gene could play a major role in suppression of the anther endothecium in female *S. purpurea* catkins. Two sex-biased genes encoding the DExH-box ATP-dependent RNA helicase, *BRR2C*, are highly conserved components of the spliceosome and are required for efficient splicing of *FLC* introns, as well as regulation of *FT* and *SOC1* in Arabidopsis (Mahrez et al. 2016). Of the two sex-biased Dr1/NF-Y paralogs in *S. purpurea*, SapurV1A.0934s0060 was highly expressed in males, whereas SapurV1A.0301s0160 was highly expressed in females. Dr1 represses RNAP II transcription by binding to TBP to prevent the formation of an active transcription complex. Members of the heterotrimeric NF-Y transcription factor

family in Arabidopsis initiate photoperiod-dependent flowering and also required for activation of the *FT* promoter by initiating downstream events leading to floral transition (Siriwardana et al. 2016).

We observed a number of genes with null expression in males that are highly expressed in females, yet we found no evidence for the converse. Rather, low levels of gene expression in females were always accompanied by low expression levels in males. Nearly 65% of all sex-biased genes were more highly expressed in females than in males, indicating disproportionate W-specific genes or Z-specific pseudogenes in the SDR. Given our findings on sexually dimorphic expression in *S. purpurea*, a reasonable hypothesis is that the transcription of genes involved in early floral meristem identity is differentially regulated by the relative abundance of RNAP II core subunits, whose promoter site specificity, initially guided by TBP components, are likely sex-specific.

We conclude that chr15 contributes to sexual dimorphism in *S. purpurea*, as genes with sex-biased expression were vastly over-represented within or near the SDR, compared with other autosomes. Although there was no apparent localization of sex-biased genes along chr19 as was found for chr15, over 12% of the sex-biased genes were most similar to *P. trichocarpa* chr19 gene models, many of which were highly expressed in males. In addition to mapping sex QTL on all three linkage maps to chr15 in F_2 *S. purpurea* (family 317), Zhou et al. (2017) identified a secondary sex QTL on chr19; however, this QTL was only present in the male backcross map. It is not clear whether chr19 is epistatic to chr15, yet it seems likely that chr19 is the ancestral sex determining chromosome of *Salix* and *Populus*, and may very well continue to contribute to sex determination, sex ratio bias, and sex dimorphism in *Salix*.

Conclusions

This study provides the first detailed analysis of transcriptome-wide regulatory divergent expression in *Salix*. Expression-level dominance and sexual dimorphism are prevailing features of differential gene expression in *S. purpurea*. Expanding upon transcriptomic resources in *Salix* will not only contribute to our understanding of the evolution of dioecy in the Salicaceae, but also facilitate the functional characterization of genes underlying sex determination in dioecious species.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

Literature Cited

Akagi T, Henry IM, Tao R, Comai L. 2014. A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. Science 346(6209): 646–650.

Alström-Rapaport C, Lascoux M, Gullberg U. 1997. Sex determination and sex ratio in the dioecious shrub *Salix viminalis* L. Theor Appl Genet. 94(3–4): 493–497.

Argus GW. 1997. Infrageneric classification of *Salix* (Salicaceae) in the New World. Syst Bot Monogr. 52: 1–121.

Baldauf JA, Marcon C, Paschold A, Hochholdinger F. 2016. Nonsyntenic genes drive tissue-specific dynamics of differential, nonadditive, and allelic expression patterns in maize hybrids. Plant Physiol. 171(2): 1144–1155.

Bell GD, Kane NC, Rieseberg LH, Adams KL. 2013. RNA-seq analysis of allele-specific expression, hybrid effects, and regulatory divergence in hybrids compared with their parents from natural populations. Genome Biol Evol. 5(7): 1309–1323.

Berlin S, Lagercrantz U, von Arnold S, Öst T, Rönnberg-Wästljung A. 2010. High-density linkage mapping and evolution of paralogs and orthologs in *Salix* and *Populus*. BMC Genomics 11(1): 129.

Cameron KD, et al. 2008. Quantitative genetics of traits indicative of biomass production and heterosis in 34 full-sib F_1 *Salix eriocephala* families. BioEnergy Res. 1(1): 80–90.

Carlson CH, Smart LB. 2016. Electrical capacitance as a predictor of root dry weight in shrub willow (*Salix*; Salicaceae) parents and progeny. Appl Plant Sci. 4(8): 1600031.

Che-Castaldo C, Crisafulli CM, Bishop JG, Fagen WF. 2015. What causes female bias in the secondary sex ratios of the dioecious woody shrub *Salix sitchensis* colonizing a primary successional landscape? Am J Bot 102(8): 1309–1322.

Chen CZ, Pikaard CS. 1997. Transcriptional analysis of nucleolar dominance in polyploid plants: biased expression/silencing of progenitor rRNA genes is developmentally regulated in *Brassica*. Proc Natl Acad Sci U S A. 94(7): 3442–3447.

Chen Y, Wang T, Fang L, Li X, Yin T. 2016. Confirmation of single-locus sex determination and female heterogamety in willow based on linkage analysis. PLoS One 11(2): e0147671.

Coate JE, Doyle JJ. 2010. Quantifying whole transcriptome size, a prerequisite for understanding transcriptome evolution across species: an example from a plant allopolyploid. Genome Biol Evol. 2: 534–546.

Combes MC, et al. 2015. Regulatory divergence between parental alleles determines gene expression patterns in hybrids. Mol Biol Evol. 7(4): 1110–1121.

Coolon JD, Webb W, Wittkopp PJ. 2013. Sex-specific effects of *cis*-regulatory variants in *Drosophila melanogaster*. Genetics 195(4): 1419–1422.

DePristo MA, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 43(5): 491–498.

Donoghue MTA, et al. 2014. C(m)CGG methylation-independent parent-of-origin effects on genome-wide transcript levels in isogenic reciprocal F1 triploid plants. DNA Res. 21(2): 141–151.

Du Z, Zhou X, Ling Y, Zhang Z, Su Z. 2010. agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res. 38(Issue suppl2): W64–W70. https://doi.org/10.1093/nar/gkq310.

Fabio ES, Kemanian AR, Montes F, Miller RO, Smart LB. 2017a. A mixed model approach for evaluating yield improvements in interspecific hybrids of shrub willow, a dedicated bioenergy crop. Ind Crops Prod. 96: 57–70.

Fabio ES, et al. 2017b. Genotype × environment interactions analysis of North American shrub willow yield trials confirms superior performance of triploid hybrids. GCB Bioenergy 9(2): 445–459.

Fairbairn DJ, Roff DA. 2006. The quantitative genetics of sexual dimorphism: assessing the importance of sex-linkage. Heredity 97(5): 319–328.

Franco-Zorrilla JM, et al. 2002. *AtREM1*, a member of a new family of B3 domain-containing genes, is preferentially expressed in reproductive meristems. Plant Physiol. 128(2): 418–427.

Geraldes A, et al. 2015. Recent Y chromosome divergence despite ancient origin of dioecy in poplars (*Populus*). Mol Ecol. 24(13): 3243–3256.

Gunter LE, Roberts GT, Lee K, Larimer FW, Tuskan GA. 2003. The development of two flanking SCAR markers linked to a sex determination locus in *Salix viminalis* L. J Heredity 94(2): 185–189.

Hanley SJ, Karp A. 2014. Genetic strategies for dissecting complex traits in biomass willows (*Salix* spp.). Tree Physiol. 34(11): 1167–1180.

Hartwig T, et al. 2011. Brassinosteroid control of sex determination in maize. Proc Natl Acad Sci U S A. 108(49): 19814–19819.

Hefer CA, Mizrachi E, Myburg AA, Douglas CJ, Mansfield SD. 2015. Comparative interrogation of the developing xylem transcriptomes of two wood-forming species: *Populus trichocarpa* and *Eucalyptus grandis*. New Phytol. 206(4): 1391–1405.

Hernandez-Lagana E, Rodriguez-Leal D, Lua J, Vielle-Calzada JP. 2016. A multigenic network of ARGONAUTE4 clade members controls early megaspore formation in *Arabidopsis*. Genetics 204(3): 1045–1056.

Ho H, Low JZ, Gudimella R, Tammi MT, Harikrishna JA. 2016. Expression patterns of inflorescence- and sex-specific transcripts in male and female inflorescences of African oil palm (*Elaeis guineensis*). Ann Appl Biol. 168(2): 274–289.

Hou J, et al. 2016. Major chromosomal rearrangements distinguish willow and poplar after the ancestral "Salicoid" genome duplication. Genome Biol Evol. 8(6): 1868–1875.

Johnson KL, Jones BJ, Bacic A, Schultz CJ. 2003. The fasciclin-like arabinogalactan proteins of Arabidopsis. A multigene family of putative cell adhesion molecules. Plant Physiol. 133(4): 1911–1925.

Kitazawa K, et al. 2013. beta-galactosyl Yariv reagent binds to the beta-1, 3-galactan of arabinogalactan proteins. Plant Physiol. 161(3): 1117–1126.

Kopp RF, et al. 2001. The development of improved willow clones for eastern North America. For Chron. 77(2): 287–292.

Koyama T, Furutani M, Tasaka M, Ohme-Takagi M. 2007. TCP transcription factors control the morphology of shoot lateral organs via negative regulation of the expression of boundary-specific genes in *Arabidopsis*. Plant Cell 19(2): 473–484.

Koyama T, Mitsuda N, Seki M, Shinozaki K, Ohme-Takagi M. 2010. TCP transcription factors regulate the activities of *ASYMMETRIC LEAVES1* and miR164, as well as the auxin response, during differentiation of leaves in *Arabidopsis*. Plant Cell 22(11): 3574–3588.

Lafarguette F, et al. 2004. Poplar genes encoding fasciclin-like arabinogalactan proteins are highly expressed in tension wood. New Phytol. 164(1): 107–121.

Landry CR, et al. 2005. Compensatory *cis-trans* evolution and the dysregulation of gene expression in interspecific hybrids of Drosophila. Genetics 171: 1813–1822.

Lee SB, et al. 2010. Ribosomal protein S3, a new substrate of Akt, serves as a signal mediator between neuronal apoptosis and DNA repair. J Biol Chem. 285(38): 29457–29468.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14): 1754–1760.

Li Z, Li B, Shen WH, Huang H, Dong A. 2012. TCP transcription factors interact with AS2 in the repression of class-I *KNOX* genes in *Arabidopsis thaliana*. Plant J. 71(1): 99–107.

Liu J, et al. 2013. Transcriptome analysis of the differentially expressed genes in the male and female shrub willows (*Salix suchowensis*). PLoS One 8(4): e60181.

MacMillan CP, Mansfield SD, Stachurski ZH, Evans R, Southerton SG. 2010. Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in Arabidopsis and *Eucalyptus*. Plant J. 62(4): 689–703.

MacMillan CP, et al. 2015. The fasciclin-like arabinogalactan protein family of *Eucalyptus grandis* contains members that impact wood biology and biomechanics. New Phytol. 206(4): 1314–1327.

Mahrez W, et al. 2016. BRR2a Affects Flowering Time via FLC Splicing. PLoS Genet. 12(4): e1005924.

McKeown PC, et al. 2011. Identification of imprinted genes subject to parent-of-origin specific expression in *Arabidopsis thaliana* seeds. BMC Plant Biol. 11: 113.

McKown AD, et al. 2017. Sexual homomorphism in dioecious trees: extensive tests fail to detect sexual dimorphism in *Populus*. Sci Rep.7: 1831.

McManus CJ, et al. 2010. Regulatory divergence in *Drosophila* revealed by mRNA-seq. Genome Res. 20(6): 816–825.

Meiklejohn CD, Coolon JD, Hartl DL, Wittkopp PJ. 2014. The roles of *cis*- and *trans*-regulation in the evolution of regulatory incompatibilities and sexually dimorphic gene expression. Genome Res. 24(1): 84–95.

Murase K, et al. 2017. MYB transcription factor gene involved in sex determination in *Asparagus officinalis*. Genes Cells 22(1): 115–123.

Nagamitsu T, Futamura N. 2014. Sex expression and inbreeding depression in progeny derived from an extraordinary hermaphrodite of *Salix subfragilis*. Bot Stud. 55(1): 1–9.

Pagnussat GC, et al. 2005. Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. Development 132(3): 603–614.

Parsch J, Ellegren H. 2013. The evolutionary causes and consequences of sex-biased gene expression. Nat Rev Genet. 14(2): 83–87.

Pucholt P, Rönnberg-Wästljung A-C, Berlin S. 2015. Single locus sex determination and female heterogamety in the basket willow (*Salix viminalis* L.). Heredity 114(6): 575–583.

Pucholt P, Wright AE, Conze LL, Mank JE, Berlin S. 2017. Recent sex chromosome divergence despite ancient dioecy in the willow *Salix viminalis*. Mol Biol Evol. 34(8): 1991–2001.

Puckett EE, et al. 2012. Differential expression of genes encoding phosphate transporters contributes to arsenic tolerance and accumulation in shrub willow (*Salix* spp.). Environ Exp Bot. 75: 248–257.

R Core Team. 2015. R: A language and environment for statistical computing. Vienna, Australia: R Foundation for Statistical Computing. ISBN: 3-900051-07-0. Available online at http://www.R-project.org/, last accessed September 10, 2017

Rice WR. 1992. Sexually antagonistic genes: experimental evolution. Science 256(5062): 1436–1439.

Rieseberg LH, Widmer A, Arntz AM, Burke JM. 2003. The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. Philos Trans R Soc Lond B Biol Sci. 358(1434): 1141–1147.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1): 139–140.

Rodgers-Melnick E, et al. 2012. Contrasting patterns of evolution following whole genome versus tandem duplication events in *Populus*. Genome Res. 22(1): 95–105.

Semerikov V, et al. 2003. Genetic mapping of sex-linked markers in *Salix viminalis* L. Heredity 91(3): 293–299.

Serapiglia MJ, Cameron KD, Stipanovic AJ, Smart LB. 2012. Correlations of expression of cell wall biosynthesis genes with variation in biomass composition in shrub willow (*Salix* spp.) biomass crops. Tree Genet Genomes 8(4): 775–788.

Serapiglia MJ, et al. 2014a. Ploidy level affects important biomass traits of novel shrub willow (*Salix*) hybrids. BioEnergy Res. 8: 259–269.

Serapiglia MJ, Gouker FE, Smart LB. 2014b. Early selection of novel triploid hybrids of shrub willow with improved biomass yield relative to diploids. BMC Plant Biol. 14: 74.

Shi X, et al. 2012. *Cis*- and *trans*-regulatory divergence between progenitor species determines gene-expression novelty in *Arabidopsis* allopolyploids. Nat Commun. 3: 950.

Siriwardana CL, et al. 2016. NUCLEAR FACTOR Y, Subunit A (NF-YA) proteins positively regulate flowering and act through FLOWERING LOCUS T. PLoS Genet. 12: e1006496.

Smart LB, Cameron KD. 2008. Genetic improvement of willow (*Salix* spp.) as a dedicated bioenergy crop. In: Vermerris W, editor. Genetic improvement of bioenergy crops. New York: Springer. p. 377–396.

Song G, et al. 2013. Global RNA sequencing reveals that genotype-dependent allele-specific expression contributes to differential expression in rice F1 hybrids. BMC Plant Biol. 13(1): 221.

Song Y, et al. 2013. Sexual dimorphic floral development in dioecious plants revealed by transcriptome, phytohormone, and DNA methylation analysis in *Populus tomentosa*. Plant Mol Biol. 83(6): 559–576.

Springer NM, Stupar RM. 2007. Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. Plant Cell 19(8): 2391–2402.

Stoof C, et al. 2015. Untapped potential: opportunities and challenges for sustainable bioenergy production from marginal lands in the Northeast USA. BioEnergy Res. 8(2): 482–501.

Stupar RM, Hermanson PJ, Springer NM. 2007. Nonadditive expression and parent-of-origin effects identified by microarray and allele-specific expression profiling of maize endosperm. Plant Physiol. 145(2): 411–425.

Stupar RM, Springer NM. 2006. *Cis*-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid. Genetics 173: 2199–2210.

Su X, et al. 2016. How does long-term complete submergence influence sex ratio and resource allocation of a dioecious shrub, *Salix variegata* Franch.? Ecol Eng. 87: 218–223.

Suvorov A, et al. 2013. Intra-specific regulatory variation in *Drosophila pseudoobscura*. PLoS One 8: e83547.

Taira T, et al. 1998. AMY-1, a novel C-MYC binding protein that stimulates transcription activity of C-MYC. Genes Cells 3(8): 549–565.

Tuskan GA, et al. 2012. The obscure events contributing to the evolution of an incipient sex chromosome in *Populus*: a retrospective working hypothesis. Tree Genet Genomes 8(3): 559–571.

Ueno N, Suyama Y, Seiwa K. 2007. What makes the sex ratio female-biased in the dioecious tree *Salix sachalinensis*? J Ecol. 95: 951–959.

Wang H, Mao Y, Yang J, He Y. 2015. *TCP24* modulates secondary cell wall thickening and anther endothecium development. Front Plant Sci. 6: 436.

Wittkopp PJ, Haerum BK, Clark AG. 2004. Evolutionary changes in *cis* and *trans* gene regulation. Nature 430(6995): 85–88.

Wittkopp PJ, Haerum BK, Clark AG. 2008a. Regulatory changes underlying expression differences within and between *Drosophila* species. Nat Genet. 40: 346–350.

Wittkopp PJ, Haerum BK, Clark AG. 2008b. Independent effects of *cis*- and *trans*-regulatory variation on gene expression in *Drosophila melanogaster*. Genetics 178: 1831–1835.

Wittkopp PJ, Kalay G. 2011. *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nat Rev Genet. 13(1): 59–69.

Xu C, et al. 2014. Genome-wide disruption of gene expression in allopolyploids but not hybrids of rice subspecies. Mol Biol Evol. 31(5): 1066–1076.

Yin T, et al. 2008. Genome structure and emerging evidence of an incipient sex chromosome in *Populus*. Genome Res. 18(3): 422–430.

Zhai F, et al. 2016. Male and female subpopulations of *Salix viminalis* present high genetic diversity and high long-term migration rates between them. Front Plant Sci. 7: 330.

Zhou R, et al. 2017. Sex determination in *Salix purpurea* (in preparation).

Associate editor: Patricia Wittkopp