BioOne COMPLETE

Transcriptomic Resources for an Endemic Neotropical Plant Lineage (Gesneriaceae)

Authors: Serrano-Serrano, Martha L., Marcionetti, Anna, Perret, Mathieu, and Salamin, Nicolas

Source: Applications in Plant Sciences, 5(4)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1600135

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Downloaded From: https://complete.bioone.org/journals/Applications-in-Plant-Sciences on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Applications in Plant Sciences

Applications in Plant Sciences 2017 5(4): 1600135

Genomic Resources Article

Transcriptomic resources for an endemic Neotropical plant lineage (Gesneriaceae)[1]

Martha L. Serrano-Serrano[2,3,4,6], Anna Marcionetti[2,3,4], Mathieu Perret[5,7], and Nicolas Salamin[2,3,4,7]

[2]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland; [3]Department of Computational Biology, University of Lausanne, 1015 Lausanne, Switzerland; [4]Swiss Institute of Bioinformatics, Quartier Sorge, 1015 Lausanne, Switzerland; and [5]Conservatoire et Jardin botaniques de la Ville de Genève and Laboratory of Plant Systematics and Biodiversity, University of Geneva, Chemin de l'Impératrice 1, 1292 Chambésy, Geneva, Switzerland

- *Premise of the study:* Despite the extensive phenotypic variation that characterizes the Gesneriaceae family, there is a lack of genomic resources to investigate the molecular basis of their diversity. We developed and compared the transcriptomes for two species of the Neotropical lineage of the Gesneriaceae.
- *Methods and Results:* Illumina sequencing and de novo assembly of floral and leaf samples were used to generate multigene sequence data for *Sinningia eumorpha* and *S. magnifica*, two species endemic to the Brazilian Atlantic Forest. A total of 300 million reads were used to assemble the transcriptomes, with an average of 92,038 transcripts and 43,506 genes per species. The transcriptomes showed good quality metrics, with the presence of all eukaryotic core genes, and an equal representation of clusters of orthologous groups (COG) classifications between species. The orthologous search produced 8602 groups, with 15–20% of them annotated using BLAST tools.
- *Discussion:* This study provides the first step toward a comprehensive multispecies transcriptome characterization of the Gesneriaceae family. These resources are the basis for comparative analyses in this species-rich Neotropical plant group; they will also allow the investigation of the evolutionary importance of multiple metabolic pathways and phenotypic diversity, as well as developmental programs in these nonmodel species.

Key words: Brazilian Atlantic Forest; de novo transcriptome assembly; floral/flower evolution; gene expression; Gesneriaceae; habit; *Sinningia*.

The understanding of the genetic basis of phenotypic and ecological divergences between species is a central question in evolutionary biology (Coyne and Orr, 2004). The investigation of these questions in nonmodel plant species is motivated by their patterns of ecological and phenotypic differentiation during lineage diversification (Elmer and Meyer, 2011), but represents a challenge due to the lack of genomic resources. This challenge can now be overcome with the advent of next-generation sequencing (NGS) methods that allow a better integration of the patterns of molecular, chromosomal, and epigenetic evolution into plant speciation and diversification (Rieseberg and Wendel, 2004; Lexer and Widmer, 2008). Recent studies highlighted the value of genomic resources for the phylogenomics of highly diverse lineages (Yang et al., 2015) and the advances and future research of NGS into broad-scale patterns of diversification (Lexer et al., 2013).

One of the strategies that has revolutionized NGS projects is whole-transcriptome sequencing, or RNA-Seq (Wang et al., 2009). This technique directly accesses the expressed protein-coding genes in a sample, allowing the investigation of differences in gene expression and sequence between conditions, populations, or species, without any previous knowledge of the biological system. The increased adoption of this technique to address ecological and evolutionary questions relies partially on its applicability to nonmodel organisms and its growing toolkit and bioinformatics support (Orsini et al., 2013; Wolf, 2013).

The Ligeriinae subtribe, previously named Sinningieae (Weber et al., 2013), is a clade of Gesneriaceae endemic to the Brazilian Atlantic Forest, which includes three genera (*Sinningia* Nees, *Vanhouttea* Lem., and *Paliavana* Vand.) and 91 species. Ligeriinae shows a large variation in habit, inflorescence form, and corolla shape that evolved during the ca. 30 Myr history of this clade (Perret et al., 2003, 2013). Previous studies on the mechanisms of speciation in Ligeriinae suggested that geographical isolation and repeated adaptations to different functional

[1]Manuscript received 7 November 2016; revision accepted 6 March 2017.

The authors thank Alain Chautems for sharing his expertise about the species and the Vital-IT facilities of the Swiss Institute of Bioinformatics for the computational support. The authors thank the staff of the Jardin botaniques de la Ville de Genève for the maintenance of the Gesneriaceae collection. Collection permits were granted by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) of Brazil (EX-15/89). This work was funded by the University of Lausanne and the Swiss National Science Foundation (CRSIII3-147630) to N.S. M.L.S.S. was funded by a fellowship from the University of Lausanne.

[6]Author for correspondence: martha.serrano@unil.ch

[7]Both authors share the last authorship.

doi:10.3732/apps.1600135

Applications in Plant Sciences 2017 5(4): 1600135; http://www.bioone.org/loi/apps © 2017 Serrano-Serrano et al. Published by the Botanical Society of America. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC-BY-NC-SA 4.0), which permits unrestricted noncommercial use and redistribution provided that the original author and source are credited and the new work is distributed under the same license as the original.

Downloaded From: https://complete.bioone.org/journals/Applications-in-Plant-Sciences on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

groups of pollinators and environments have promoted their diversification and geographic expansion (Perret et al., 2006, 2007). Although these results and the recent advances in the knowledge of the evolutionary history in the Gesneriaceae (Serrano-Serrano et al., 2015; Roalson and Roberts, 2016) provided new insights on the diversification of tropical plants, the scarce genomic resources for this group still limit our understanding of the molecular mechanisms underlying their evolution. Recent efforts for generating genomic or transcriptomic resources in the Gesneriaceae have been focused on the Old World members of the family. The leaf transcriptome and the genome of *Boea hygrometrica* (Bunge) R. Br. have provided evidence for the regulation of physiological and cellular functions during plant dehydration (Zhu et al., 2015; Xiao et al., 2015). Similarly, transcriptomic resources from seedlings of *Streptocarpus rexii* (Bowie ex Hook.) Lindl. have contributed to the available genomic data for this family (Chiara et al., 2013). In contrast, data for New World members are restricted to two still unpublished National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) projects (i.e., *Achimenes* Pers. species [accession no. PRJNA340450] and *Sinningia speciosa* (Lodd.) Hiern [accession no. PRJNA282582]) that will soon complement the efforts for the exploration of the genomics of Neotropical Gesneriaceae.

Here, we contribute to these resources by developing novel NGS resources in the subtribe Ligeriinae. This tribe is particularly suited to address questions about the mechanisms generating morphological and ecological differentiation, such as changes in growth habits or the evolution of floral traits associated with pollinator types (Perret et al., 2003, 2007). This article describes the generation of de novo reference transcriptomes for two perennial tuber species within the *Dircaea* clade: *Sinningia magnifica* (Otto & A. Dietr.) Wiehler and *S. eumorpha* H. E. Moore, with around 9 Myr divergence (Serrano-Serrano et al., 2017b). The data produced include nucleotide and translated protein sequences for the assembled genes, annotation tables that will pave the way for future investigations including population genomic studies, development of new phylogenetic markers, and comparative transcriptomics in these species and their related lineages. We discuss the main differences in the transcriptomic data between the two species, providing information on functional categories that may be the groundwork for the study of the genomic regions associated with the morphological evolution of this clade of Gesneriaceae.

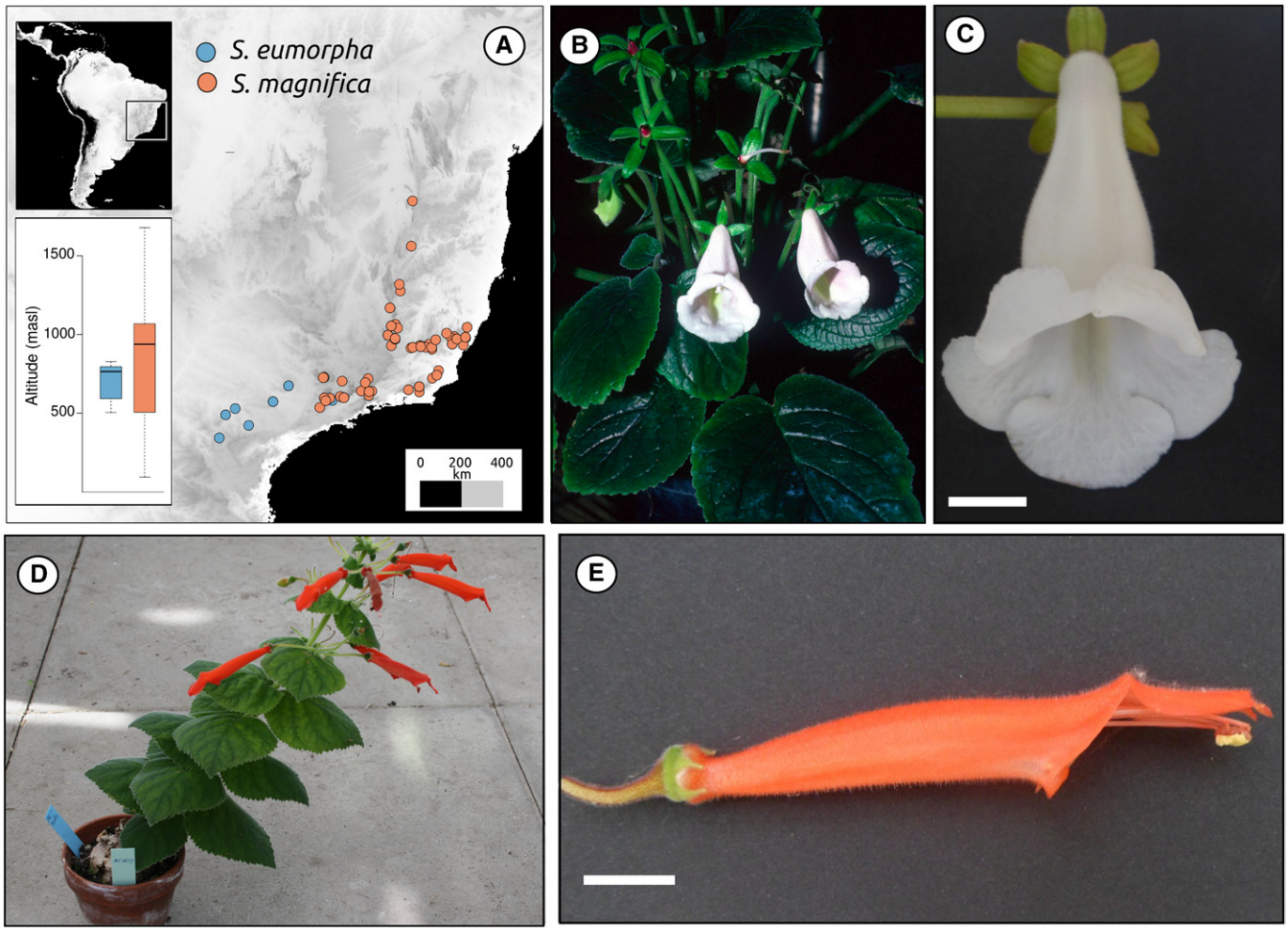


Fig. 1. (A) Distribution of *Sinningia eumorpha* and *S. magnifica* in southeastern Brazil. *Sinningia eumorpha* habit (B) and flower (C). *Sinningia magnifica* habit (D) and flower (E). Scale bars: 1 cm. Photos by A. Chautems (B) and M. L. Serrano-Serrano (C–E).

Downloaded From: https://complete.bioone.org/journals/Applications-in-Plant-Sciences on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

METHODS AND RESULTS

***Plant material, RNA extraction, and library construction*—**All plants were cultivated in the greenhouses of the Botanical Garden of Geneva in Switzerland (Conservatoire et Jardin botaniques de la Ville de Genève [CJB]; Fig. 1). Tissue material was collected from young leaves and flowers; flowers were sampled at three developmental time points based on the percentage of the total flower size measured from the receptacle to the end of the corolla tube: bud1 = 0–33%, bud2 = 33–75%, and flower = 75–100%. We collected each species at every floral stage and leaf material from two biological replicates represented by different accessions (*S. eumorpha* accession no.: 200807305 and 20090391J0; *S. magnifica* accession no.: AC3615 and AC23105). Each sample contained multiple floral tissues, such as sepals, the limb portion of the flower, flower tube, anthers including the filament, and the stigma including the style, which were all combined. All samples were collected from May 2013 to November 2014, and were immediately frozen in liquid nitrogen and stored at −80°C. RNA was extracted with the QIAGEN RNeasy Plant Kit (catalog no. 74904; QIAGEN, Hilden, Germany) and treated with DNase I (QIAGEN) according to the manufacturer's instructions. Illumina TruSeq (San Diego, California, USA) stranded paired-end mRNA libraries were performed using 2 µg of total RNA, following the library prep kit instructions (protocol version 15031047, Revision D, September 2012) for 300-bp fragments. Library concentration, integrity, and size were determined with the Agilent Fragment Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California, USA) and Qubit fluorometric quantitation (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Eight libraries were constructed for each species (three floral and one leaf, in two replicates each). Illumina sequencing was performed with 100 cycles of paired-end reads in a HiSeq 2500 at the Lausanne Genomic Technologies Facility (Lausanne, Switzerland).

***Transcriptome assembly and annotation*—**We generated Illumina raw reads from each tissue and replicate that were pooled. Sequencing was even between the two species, with 316 and 332 million reads for *S. eumorpha* and *S. magnifica*, respectively. Raw reads were trimmed and filtered with a minimum length of 80 nucleotides and an average quality score higher than 20, using the FASTX-Toolkit (version 0.0.13.2, http://hannonlab.cshl.edu/fastx_toolkit/). Reads have been deposited in the NCBI SRA under BioProject ID PRJNA349129 (Biosamples SAMN05930205: Illumina reads for *S. eumorpha* [TaxId: 189015], and SAMN05930295: Illumina reads for *S. magnifica* [TaxId: 189031]). De novo transcriptomes for each species were assembled using the Trinity pipeline (version 2.0.3; Grabherr et al., 2011) using a minimum contig length of 200. The large amount of raw reads from all libraries per species was concatenated and in silico normalized to a maximum coverage of 50 with the Trinity settings *–normalize_max_read_cov*. We filtered the lowest 5% of the distribution of transcript lengths and a minimum of 1 FPKM (fragments per kilobase of transcript per million reads mapped) to obtain the final transcript set per species (Table 1). The transcriptome sequences are available from the Dryad Digital Repository as a FASTA file (http://dx.doi.org/10.5061/dryad.4r5p1; Serrano-Serrano et al., 2017a). Open reading frames (ORFs) were predicted with TransDecoder (Trinity plug-ins version 2.0.3; Haas et al., 2013). All genes and ORFs (predicted proteins) were annotated using BLASTX and BLASTP against the SwissProt database, producing matches with high-quality annotations. Sequence contaminants were screened and removed using the BLAST results that matched bacterial, fungal, or any other nonplant genetic material. BLAST annotations were filtered to avoid spurious hits using a threshold for the *E*-value and identity higher than 1×10^{-6} and 55%, respectively. Trinotate (version 2014.07.07; http://trinotate.github.io) was used to integrate the functional annotation, selecting one unique top BLAST hit and gene ontology (GO) annotation. Gene ontologies were plotted and compared between species using the WEb Gene Ontology annotation plot (WEGO) webtool (http://wego.genomics.org.cn/cgi-bin/wego/index.pl; Ye et al., 2006). The assembly statistics are presented in Table 1.

The number of Trinity transcript clusters (hereafter called "genes"), transcripts, and ORFs, and the distribution of mean and total sizes of transcripts were similar between species, although these values were overall higher for *S. eumorpha* when compared with *S. magnifica* (Table 1; Fig. 2A, 2B). The comparison of libraries showed that a large proportion of genes are shared by all stages and tissues (18,636 and 18,386 for *S. eumorpha* and *S. magnifica*, respectively; see Fig. 3). Genes found uniquely in leaves, flower buds, and anthetic flowers ranged between 1418 to 6810 and 1234 to 4345 for *S. magnifica* and *S. eumorpha*, respectively (Fig. 3). We performed GO enrichment (http://geneontology.org/) for these unique genes (Table 2 shows the results with $P < 0.05$), which represent the specific GO functional categories enriched at a specific developmental time (flowers vs. early and late buds) and in particular tissues (leaves vs. all flower stages). For instance, the genes uniquely expressed in early stages of flower development (B1) included categories related to reproductive developmental processes and regulation of transcription for both species. For the middle flower stage (B2), a unique GO category (xylan biosynthetic pathway) was enriched in both species. This pathway produces xylan, which is an important polymer found in plant cell walls that is likely related to growth processes. The GO categories at the anthetic flower stage showed strong differences between species. The genes uniquely expressed in flowers of *S. eumorpha* were enriched for categories such as translation, as well as metabolism of heterocycle macromolecules. In contrast, *S. magnifica* showed fewer GO categories enriched (response to hormones and inorganic anion transport). The GO categories found between the leaf libraries showed similarities in functions related to hormone responses.

Table 1. Summary statistics for the two de novo–assembled transcriptomes.

Statistics	*Sinningia eumorpha*	*Sinningia magnifica*
No. of genes	46,974	40,039
No. of transcripts	87,053	97,023
Median contig length (bp)	1179	1089
Average contig length (bp)	1687	1545
N50 (bp)	2597	2794
Percentage of annotated genes	13.48%	17.37%
No. of ORFs	171,889	131,000
Total no. of SSRs	3428	2966
No. of contigs with SSRs	2236	1968
No. of simple SSRs	3213	2747
No. of compound SSRs	215	219
No. of contigs with 1 SSR	1615	1398
No. of contigs with 2 SSRs	341	356
No. of contigs with 3 SSRs	151	117
No. of contigs with 4 SSRs	61	44
No. of loci with dinucleotide repeats	600	524
No. of loci with trinucleotide repeats	1402	1284
No. of loci with tetranucleotide repeats	24	17

Note: ORF = open reading frame; SSR = simple sequence repeat.

The annotation of the two transcriptomes showed that only a low proportion of the genes have significant BLAST matches with existing curated protein and nucleotide sequences (maximum of 13% and 20% for BLASTX and BLASTP, respectively; see the BLAST annotation file in the Dryad Digital Repository [http://dx.doi.org/10.5061/dryad.4r5p1; Serrano-Serrano et al., 2017a]). This pattern is common for nonmodel species, and BLAST tools are known to fail for annotation of around 75% of the genes for de novo assemblies (Chiara et al., 2013; DeBiasse and Kelly, 2016). Gene ontologies indicated that the proportion of assembled genes associated with each functional category is similar between species, and that the de novo assemblies are comparable (Fig. 4). High and low abundance terms are mostly shared between species, except for three categories at the cellular component level (envelope, membrane-enclosed lumen, and symplast), one category at the molecular function level (electron carrier), and six categories at the biological process level (biological regulation, immune system process, locomotion, multiorganism process, pigmentation, and rhythmic process) that have significantly different gene counts ($P < 0.001$). These categories most likely contain pathways and genes with a different pattern of expression between the two species.

***Transcriptome quality checking*—**The two assembled transcriptomes were examined for quality and completeness using two measures: the ortholog hit ratio (OHR; O'Neil et al., 2010) and the Core Eukaryotic Genes Mapping Approach (CEGMA) analysis (http://korflab.ucdavis.edu/datasets/cegma/; Parra et al., 2007). The OHR is computed as the percentage of a gene in the transcriptome that matches a putative ortholog in a related species (*Solanum lycopersicum* L. [tomato] with ca. 79 Myr divergence; http://www.timetree.org/). It is calculated by dividing the length of the putative coding region by the total length of the orthologous gene. We performed a BLASTX of our genes against the set of transcripts from tomato (International Tomato Annotation Group [ITAG] version 2.4 predicted protein database: 34,725 sequences [December 2014]) and

Downloaded From: https://complete.bioone.org/journals/Applications-in-Plant-Sciences on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

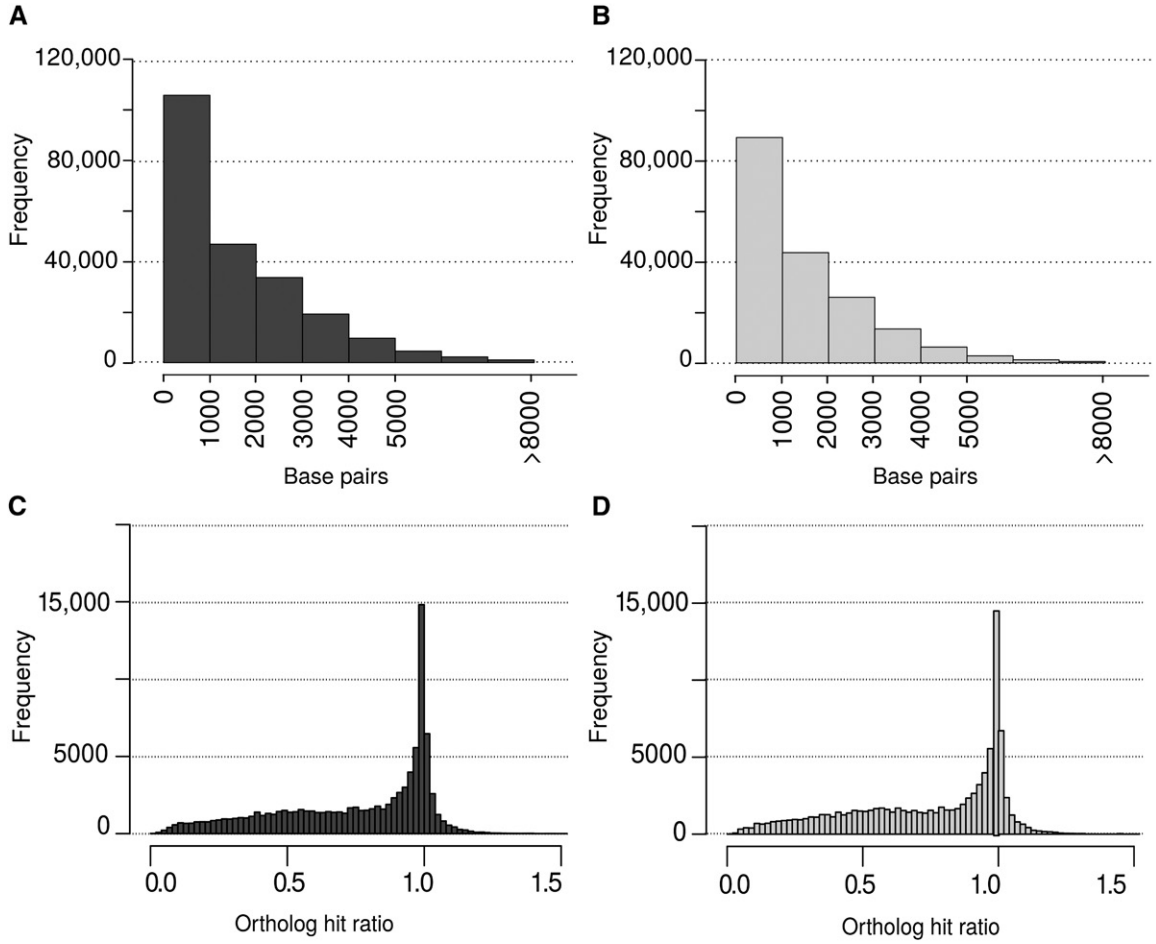


Fig. 2. Distribution of transcripts on size classes (0–1000, to >8000 bp) for *Sinningia eumorpha* (A) and *S. magnifica* (B). Distribution of ortholog hit ratios between tomato and *S. eumorpha* (C) and between tomato and *S. magnifica* (D).

considered the best hit with an *E*-value $< 1 \times 10^{-6}$ to be the putative orthologs. For the CEGMA analysis, we used clusters of orthologous groups (COGs) for eukaryotes to search the 458 highly conserved core proteins that matched our predicted ORFs. The evaluation of the transcriptome completeness performed using the CEGMA analysis showed that 99.6% of the core eukaryotic genes mapped to the generated genes on each species. Additionally, the evaluation of the OHR indicated that the assembled genes covered a large proportion of the putative reference orthologs in tomato (Fig. 2C, 2D). For *S. eumorpha* and *S. magnifica*, 51.4% and 51.3% of genes have an overlap with the reference orthologous genes higher than 0.8.

***Orthologous search and SSR markers*—**The analysis with OrthoMCL (version 2.0.9; Li et al., 2003) identified 8602 orthologous groups between the two species, with 2274 putative single-copy genes (data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.4r5p1; Serrano-Serrano et al., 2017a). This number is comparable with other surveys between closely related species (Zhang et al., 2013). These orthologous groups provide the information for future molecular evaluations and gene-specific analyses. Simple sequence repeat (SSR) markers were identified using the MicroSAtellite Identification Tool (MISA; Thiel et al., 2003), with a minimum number of repeats of six, five, and five for dinucleotides, trinucleotides, and tetranucleotides, respectively. A total of 3428 and 2966 SSRs were identified for *S. eumorpha* and *S. magnifica*, respectively. These SSRs are located in 1968 and 2236 genes for *S. eumorpha* and *S. magnifica*, respectively, which indicates that approximately 2% of the genes contain at least one SSR marker (1.98% and 2.61% for *S. eumorpha* and *S. magnifica*, respectively). Most of these genes contain a single SSR, and the most common SSRs are trinucleotides for both species (Table 1). Although we are unable to currently test these markers at the population level because of the scarcity of samples, we provide the SSRs as a valuable resource for marker development (data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.4r5p1; Serrano-Serrano et al., 2017a).

DISCUSSION

Our results provide large-scale sequence data for two related species (only six nuclear genes were available for the Ligeriinae previous to this work), and expression and population-level genomic analyses can now be developed. The resources we developed will facilitate the investigation of ecological and evolutionary questions within the *Sinningia* genus and the Gesneriaceae family. The transcriptomes described in this study generated at least 40,000 Trinity "genes" per species, and 8600 potentially orthologous groups were identified between them. The quality and quantity of data are comparable between the two species and to other plant studies (Zhang et al., 2013; Chapman, 2015), providing valuable transcriptomic data for the Gesneriaceae family.

The number of genes uniquely present in one of the developmental stages differed between species. We found a larger number of genes at the anthetic flower stage compared with other stages (6810 genes in Fig. 3A) in *S. eumorpha.* In contrast, a

Downloaded From: https://complete.bioone.org/journals/Applications-in-Plant-Sciences on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

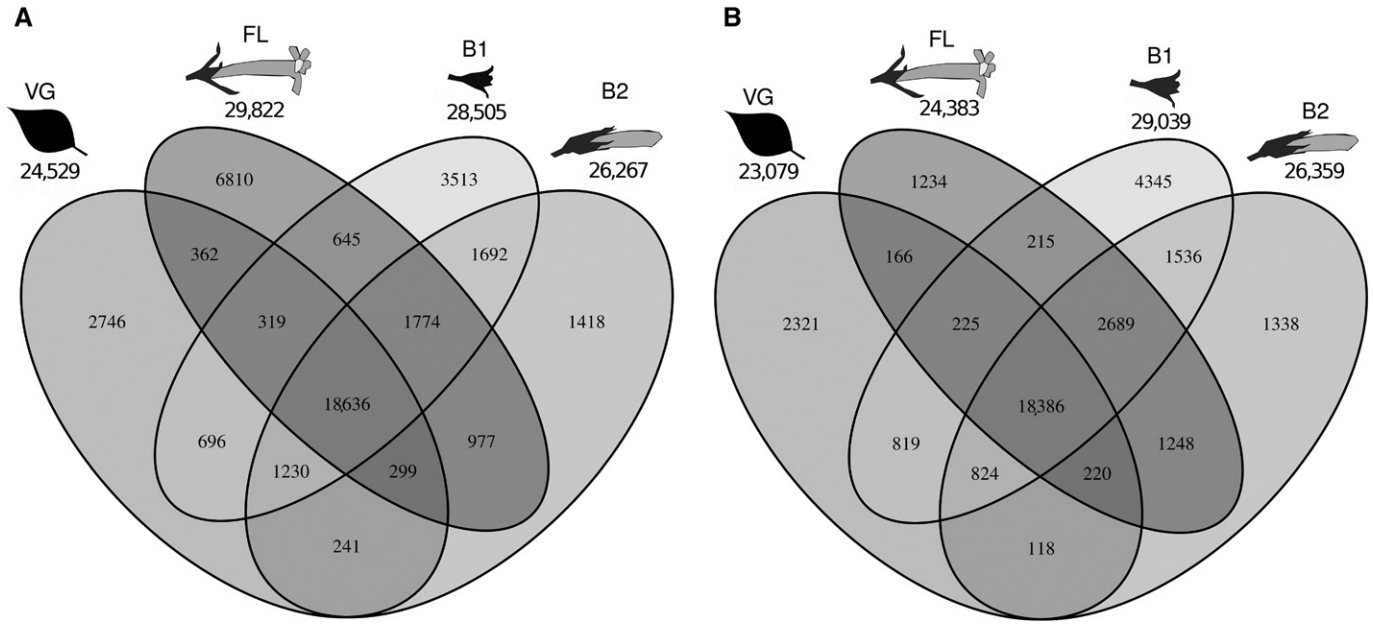


Fig. 3. Venn diagram for the shared and unique number of genes between libraries for *Sinningia eumorpha* (A) and *S. magnifica* (B). Pictographs show the stages: B1 = small bud, B2 = medium buds, FL = anthetic flower, VG = leaf. Boldfaced numbers outside the diagram correspond to the total number of genes in a library.

higher number of unique genes in *S. magnifica* was found in middle buds (4345 in Fig. 3B). These differences may indicate that most of the stage-specific expression occurs later in the floral development of *S. eumorpha* compared with *S. magnifica.* We should be careful, however, because the tissue identity and growth phases have not been characterized in detail for the species in question. The functional categories associated with the stage-specific genes also differed between species mostly at the anthetic flower stage (Table 2), pointing again to the fact that the late stages of flower development could carry relevant genes for the morphological differentiation between species. Figure 4 showed additional categories where the whole assembled transcriptomes differed between species, such as pigmentation, driving our attention toward flower color, one of the most distinctive traits between the species (Fig. 1C, 1E).

The sequences reported here constitute the raw material for designing probes and primers for phylogenetic and population studies. Additional steps are essential to complement any genomic or experimental survey for this plant family. For instance, the identification of differences in gene expression between taxa and the evaluation of evolutionary constraints or divergent use of genes will help our understanding of the evolution of reproductive isolation and habitat adaptation in plants (Lexer and Widmer, 2008). Furthermore, the identification of candidate genes for flower evolution and the investigation of floral developmental genetics will contribute to our understanding of flower

Table 2. Gene ontology enrichment for the genes uniquely expressed in a developmental stage and tissue.

Species	Stage[a]	Gene ontologies
S. eumorpha	B1	Gamete generation, transmembrane receptor protein tyrosine kinase, leaf development, reproductive shoot system development, protein phosphorylation, transcription and DNA-templated, regulation of transcription and DNA-templated
S. eumorpha	B2	Xylan biosynthetic process, plant-type secondary cell wall biogenesis, anther dehiscence
S. eumorpha	FL	Translation, ribosome biogenesis, response to cadmion ion, macromolecule complex subunit organization, oxoacid metabolic process, macromolecule modification, heterocycle metabolic process
S. eumorpha	VG	Hormone-mediated signaling pathway, regulation of developmental process, response to acid chemical, response to oxygen-containing compound, response to abiotic stimulus
S. magnifica	B1	Photomorphogenesis, cell surface receptor signaling pathway, protein phosphorylation, translation, reproductive structure development, single organism reproductive process, response to hormone, regulation of transcription and DNA-templated, cellular component organization or biogenesis
S. magnifica	B2	Xylan biosynthetic process, cellular response to jasmonic acid stimulus, regulation of response to stimulus, aromatic compound biosynthetic process, single-organism developmental process
S. magnifica	FL	Inorganic anion transport, response to hormone
S. magnifica	VG	Ribosome assembly, translation, transcription, regulation of cellular macromolecule biosynthesis, regulation of nitrogen compound metabolic process, response to chemical, regulation of gene expression, regulation of primary metabolic process

[a] Stages: B1 = small bud, B2 = medium buds, FL = anthetic flower, VG = leaf. See details in Methods.

Downloaded From: https://complete.bioone.org/journals/Applications-in-Plant-Sciences on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

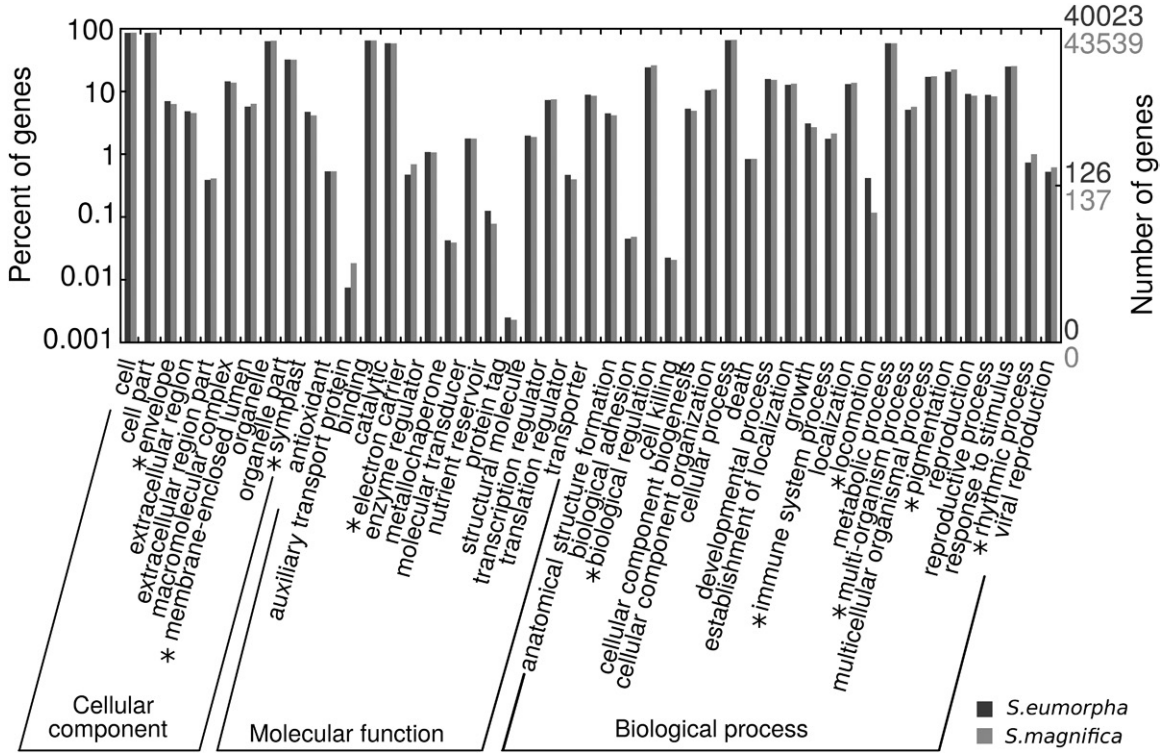


Fig. 4. Histogram of gene ontology classifications for the two species. The WEGO plot shows the three main ontology categories: cellular component, molecular function, and biological process. Asterisks show the categories for which the number of genes between the two species is significantly different ($P < 0.001$).

diversity (Chanderbali et al., 2016). The two Gesneriaceae species that we sampled will enable these genetic and comparative analyses, and will provide unprecedented research opportunities for the study of Neotropical plant diversity.

LITERATURE CITED

Chanderbali, A. S., B. A. Berger, D. G. Howarth, P. S. Soltis, and D. E. Soltis. 2016. Evolving ideas on the origin and evolution of flowers: New perspectives in the genomic era. *Genetics* 202: 1255–1265.

Chapman, M. A. 2015. Transcriptome sequencing and marker development for four underutilized legumes. *Applications in Plant Sciences* 3: 1400111.

Chiara, M., D. S. Horner, and A. Spada. 2013. De novo assembly of the transcriptome of the non-model plant *Streptocarpus rexii* employing a novel heuristic to recover locus-specific transcript clusters. *PLoS ONE* 8: e80961.

Coyne, J., and H. Orr. 2004. Speciation. Sinauer, Sunderland, Massachusetts, USA.

DeBiasse, M. B., and M. W. Kelly. 2016. Plastic and evolved responses to global change: What can we learn from comparative transcriptomics? *Journal of Heredity* 107: 71–81.

Elmer, K. R., and A. Meyer. 2011. Adaptation in the age of ecological genomics: Insights from parallelism and convergence. *Trends in Ecology & Evolution* 26: 298–306.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, et al. 2011. Trinity: Reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology* 29: 644–652.

Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* 8: 1494–1512.

Lexer, C., and A. Widmer. 2008. Review. The genic view of plant speciation: Recent progress and emerging questions. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 363: 3023–3036.

Lexer, C., S. Mangili, E. Bossolini, F. Forest, K. N. Stölting, P. B. Pearman, N. E. Zimmermann, and N. Salamin. 2013. 'Next generation' biogeography: Towards understanding the drivers of species diversification and persistence. *Journal of Biogeography* 40: 1013–1022.

Li, L., C. J. Stoeckert, and D. S. Roos. 2003. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research* 13: 2178–2189.

O'Neil, S., J. D. Dzurisin, R. D. Carmichael, N. F. Lobo, S. J. Emrich, and J. J. Hellmann. 2010. Population-level transcriptome sequencing of nonmodel organisms *Erynnis propertius* and *Papilio zelicaon*. *BMC Genomics* 11: 310.

Orsini, L., R. Andrew, and C. Eizaguirre. 2013. Evolutionary ecological genomics. *Molecular Ecology* 22: 527–531.

Parra, G., K. Bradnam, and I. Korf. 2007. CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics (Oxford, England)* 23: 1061–1067.

Perret, M., A. Chautems, R. Spichiger, G. Kite, and V. Savolainen. 2003. Systematics and evolution of tribe Sinningieae (Gesneriaceae): Evidence from phylogenetic analyses of six plastid DNA regions and nuclear ncpGS. *American Journal of Botany* 90: 445–460.

Perret, M., A. Chautems, and R. Spichiger. 2006. Dipersal-vicariance analyses in the tribe Sinningieae (Gesneriaceae): A clue to understanding biogeographical history of the Brazilian Atlantic forest. *Annals of the Missouri Botanical Garden* 93: 340–358.

Perret, M., A. Chautems, R. Spichiger, T. G. Barraclough, and V. Savolainen. 2007. The geographical pattern of speciation and floral diversification in the neotropics: The tribe Sinningieae (Gesneriaceae) as a case study. *Evolution* 61: 1641–1660.

Downloaded From: https://complete.bioone.org/journals/Applications-in-Plant-Sciences on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Perret, M., A. Chautems, A. O. DeAraujo, and N. Salamin. 2013. Temporal and spatial origin of Gesneriaceae in the New World inferred from plastid DNA sequences. *Botanical Journal of the Linnean Society* 171: 61–79.

Rieseberg, L. H., and J. F. Wendel. 2004. Plant speciation: Rise of the poor cousins. *New Phytologist* 161: 1–21.

Roalson, E. H., and W. R. Roberts. 2016. Distinct processes drive diversification in different clades of Gesneriaceae. *Systematic Biology* 65: 662–684.

Serrano-Serrano, M. L., M. Perret, M. Guignard, A. Chautems, D. Silvestro, and N. Salamin. 2015. Decoupled evolution of floral traits and climatic preferences in a clade of Neotropical Gesneriaceae. *BMC Evolutionary Biology* 15: 247.

Serrano-Serrano, M. L., A. Marcionetti, M. Perret, and N. Salamin. 2017a. Data from: Transcriptomic resources for an endemic Neotropical plant lineage (Gesneriaceae). Dryad Digital Repository. http://dx.doi.org/10.5061/dryad.4r5p1

Serrano-Serrano, M. L., J. Rolland, J. L. Clark, N. Salamin, and M. Perret. 2017b. Hummingbird pollination and the diversification of angiosperms: An old and successful association in Gesneriaceae. *Proceedings of the Royal Society B* 284: 20162816.

Thiel, T., W. Michalek, R. K. Varshney, and A. Graner. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 10: 411–422.

Wang, Z., M. Gerstein, and M. Snyder. 2009. RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews. Genetics* 10: 57–63.

Weber, A., J. L. Clark, and M. Möller. 2013. A new formal classification of Gesneriaceae. *Selbyana* 31: 68–94.

Wolf, J. B. 2013. Principles of transcriptome analysis and gene expression quantification: An RNA-seq tutorial. *Molecular Ecology Resources* 13: 559–572.

Xiao, L., G. Yang, L. Zhang, X. Yang, S. Zhao, Z. Ji, Q. Zhou, et al. 2015. The resurrection genome of *Boea hygrometrica*: A blueprint for survival of dehydration. *Proceedings of the National Academy of Sciences, USA* 112: 5833–5837.

Yang, Y., M. J. Moore, S. F. Brockington, D. E. Soltis, G. K.-S. Wong, E. J. Carpenter, Y. Zhang, et al. 2015. Dissecting molecular evolution in the highly diverse plant clade Caryophyllales using transcriptome sequencing. *Molecular Biology and Evolution* 32: 2001–2014.

Ye, J., L. Fang, H. Zheng, Y. Zhang, J. Chen, Z. Zhang, J. Wang, et al. 2006. WEGO: A web tool for plotting GO annotations. *Nucleic Acids Research* 34: W293–W297.

Zhang, L., H.-F. Yan, W. Wu, H. Yu, and X.-J. Ge. 2013. Comparative transcriptome analysis and marker development of two closely related Primrose species (*Primula poissonii* and *Primula wilsonii*). *BMC Genomics* 14: 329.

Zhu, Y., B. Wang, J. Phillips, Z. Zhang, H. Du, T. Xu, L. Huang, et al. 2015. Global transcriptome analysis reveals acclimation-primed processes involved in the acquisition of desiccation tolerance in *Boea hygrometrica*. *Plant and Cell Physiology* 56: 1429–1441.

Downloaded From: https://complete.bioone.org/journals/Applications-in-Plant-Sciences on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use