Pezzi, Pedro H., Leonardo T. Gonçalves, Alice Backes, Loreta B. Freitas, Stacy Pirro, and Maribeth Latvis. 2026. "Draft Genome Assemblies for 36 Species of False Foxglove (Agalinis, Orobanchaceae)." *Biodiversity Genomes*, November 10. https://doi.org/10.56179/001c.173537.

GENOME SEQUENCING

Draft Genome Assemblies for 36 Species of False Foxglove (Agalinis, Orobanchaceae)

Pedro H. Pezzi[1,2], Leonardo T. Gonçalves[3], Alice Backes[2], Loreta B. Freitas[2], Stacy Pirro[4], Maribeth Latvis[1,5a]

[1] Department of Biological Sciences, University of Arkansas, Fayetteville, AR, USA, [2] Department of Genetics, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil, [3] School of Biological Sciences, University of Nebraska–Lincoln, Lincoln, NE, USA, [4] Iridian Genomes, Silver Spring, MD, USA, [5] University of Arkansas Herbarium, University of Arkansas, Fayetteville, AR, USA

https://doi.org/10.56179/001c.173537

Biodiversity Genomes

Agalinis, commonly known as the false foxgloves, is a genus of hemiparasitic plants in Orobanchaceae, the largest family of parasitic plants. The genus comprises approximately 70 species distributed throughout the Americas, with about half of its diversity in North America and half in South America. Here we present draft genome assemblies for 36 Agalinis species and one Aureolaria species, generated from both herbarium specimens and silica-dried leaves. These assemblies are highly complete and provide valuable genomic resources for investigating the systematics and genome evolution of the group.

Introduction

Agalinis Raf. (Orobanchaceae), the false foxgloves, comprises about 70 species distributed across the Americas, from Canada to Argentina and Chile. About half of its diversity occurs in North America, mainly in the eastern United States, and the other half in South America, where the genus has centers of diversity in the northern Andes and in the highlands of southeastern Brazil. A densely sampled phylogeny based on six plastid and four nuclear regions placed *Esterhazya* within the South American clade of *Agalinis* (Latvis et al. 2024), and the genus now includes five additional species (Souza et al. 2025). Biogeographic reconstruction suggests that *Agalinis* is about 20 million years old and likely originated in eastern North America, followed by migration into South America, with subsequent species radiations in the Andes and Brazil (Pezzi et al. 2026). Migration scenarios through the Andes and/or South American grasslands remain unclear. In addition to recent 'dispersification', the parallel rapid radiations in South America are associated with changes in pollinators and life-history strategies (Pezzi et al. 2026). In North America, all species are bee-pollinated, and most are annuals, whereas in South America, *Agalinis* includes both bee- and hummingbird-pollinated species (Fig. 1), and most are perennials. South American *Agalinis* therefore offers a unique opportunity to study the genetic basis of rapid trait evolution during an ongoing species radiation, particularly the roles of ancestral polymorphism sorting and hybridization in the origin of hummingbird pollination. Moreover, many of the ~70 described

a Corresponding author. Email: mlatvis@uark.edu

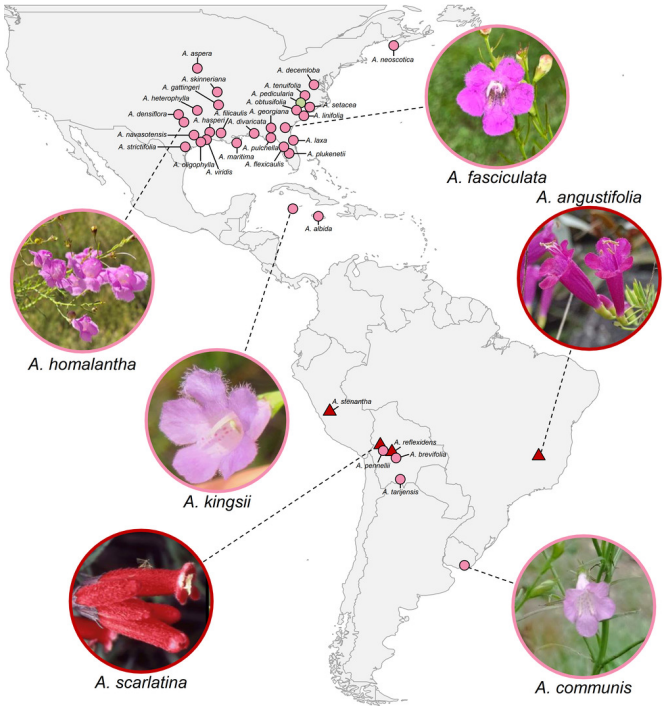


Figure 1. Collection localities of the 37 specimens sequenced in this study. Pink circles indicate bee-pollinated species and red triangles indicate hummingbird-pollinated species of *Agalinis*. A draft genome was also assembled for *Aureolaria pedicularia* (green circle), a close relative of *Agalinis*. Photos of *A. angustifolia*, *A. communis*, *A. fasciculata*, and *A. homalantha* by Pedro Pezzi; *A. kingsii* by Maribeth Latvis; and *A. scarlatina* by Judith Canne-Hilliker.

species have highly restricted distributions in both North and South America, and in the United States several are considered imperiled in at least one state. Here, we sequenced and assembled the genomes of 36 *Agalinis* species and *Aureolaria pedicularia* using mostly tissue from herbarium specimens. These data will improve our understanding of species relationships and serve as an important genomic resource for the group, supporting species conservation and, together with the chromosome-level reference genome of *Agalinis fasciculata* (Pezzi and Latvis 2026), studies of genome and trait evolution.

Material and Methods

We sampled tissue from herbarium specimens of the Missouri Botanical Garden Herbarium (MO), the Philecology Herbarium (BRIT), the Shirley C. Tucker Herbarium at Louisiana State University (LSU), the University of Florida Herbarium (FLAS), and the University of North Carolina at Chapel Hill Herbarium (NCU). For *A. angustifolia*, we collected fresh leaf tissue from a population in Diamantina, Minas Gerais, Brazil (18°16′31.4″S, 43°41′08.2″W) and dried it in silica gel. We deposited a voucher specimen at the Herbarium of the Universidade Federal de Minas Gerais (BHCB). We extracted DNA using a DNeasy kit (Qiagen, Redwood City, CA) following the manufacturer's instructions. Paired-end libraries were prepared with a TruSeq kit (Illumina, San Diego, CA) and sequenced on an Illumina HiSeq platform (2 × 150 bp). We trimmed adapters and low-quality regions with

Trimmomatic v0.33 (Bolger et al. 2014) and assembled the cleaned reads with SPAdes v3.15.4 (Bankevich et al. 2012), followed by a finishing step with Zanfona (Kieras et al. 2021). We assessed assembly completeness with BUSCO v6.1.0 against the eudicotyledons_odb12 dataset (Tegenfeldt et al. 2025).

Results and Data Availability

The median and mean ages of the sequenced herbarium specimens were 12 and 17 years, respectively (Table 1), with the oldest specimen being *A. divaricata*, collected in 1970. Sequencing produced a mean of 258,036,581 raw reads per sample, with the lowest being 143,798,190 for *A. tarijensis* and the highest being 322,938,654 for *A. kingsii*. The accession codes for the raw reads and the assembled genomes are provided in Table 1. Although DNA was extracted from historical herbarium specimens (except for *A. angustifolia*), most genomes showed high completeness, with 24 species exceeding 50% (Fig. 2). These results highlight the value of herbaria as genomic resources, especially for rare and hard-to-collect species.

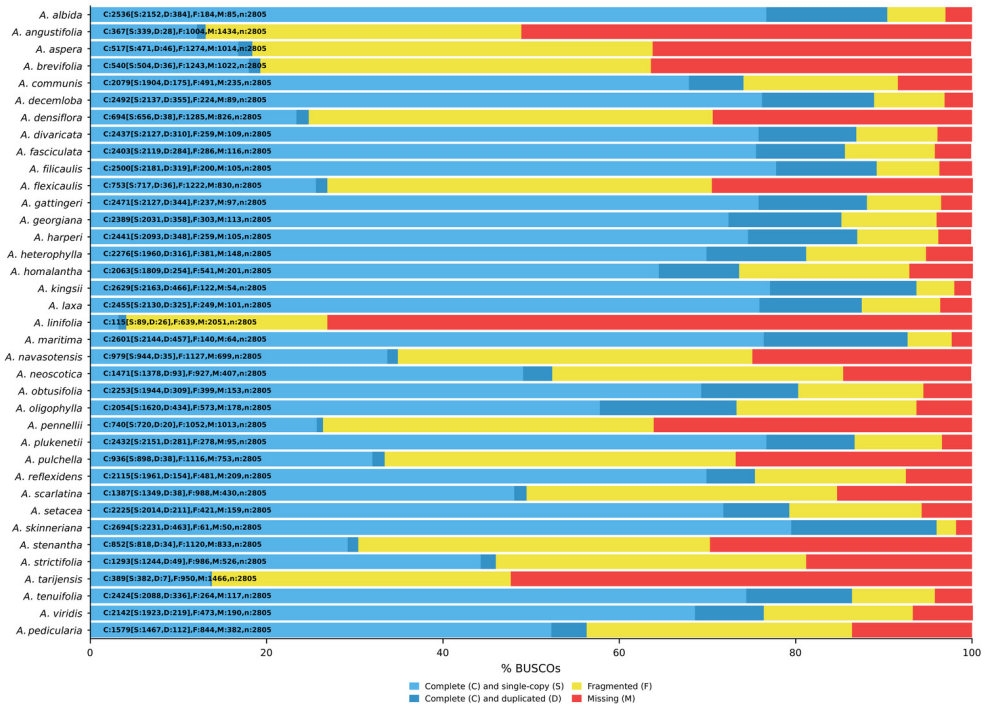


Figure 2. BUSCO completeness scores for the 37 genomes assembled in this study.

Table 1. Species included in this study, with their associated voucher information, collection year, and NCBI accession numbers for raw reads and assembled genomes. Asterisks (*) indicate species whose sequences were assembled from silica-dried leaf tissue. All other species were sampled from herbarium specimens.

Species	Voucher	Collection Year	SRA Accession	Genome Accession
Agalinis albida	FLAS116623	1972	SRR35318411	GCA_056108385.1
*Agalinis angustifolia**	BHCB224033	2025	SRR39261250	GCA_062052791.1
Agalinis aspera	MO5805413	2004	SRR39261308	GCA_062052793.1
Agalinis brevifolia	MO6065669	2006	SRR39261414	GCA_062052789.1
Agalinis communis	FLAS284087	2022	SRR35319423	GCA_056108445.1

Species	Voucher	Collection Year	SRA Accession	Genome Accession
Agalinis decemloba	NCU140243	2013	SRR35920870	GCA_056108405.1
Agalinis densiflora	BRIT617674	2019	SRR35921124	GCA_056644815.1
Agalinis divaricata	FLAS294906	1970	SRR35316942	GCA_056108425.1
Agalinis fasciculata	FLAS291960	2023	SRR35317428	GCA_056108605.1
Agalinis filicaulis	LSU117977	2007	SRR35920842	GCA_056523485.1
Agalinis flexicaulis	FLAS290370	2010	SRR35921130	GCA_056902695.1
Agalinis gattingeri	NCU450759	2018	SRR35921129	GCA_056108485.1
Agalinis georgiana	FLAS235345	2002	SRR35318202	GCA_056319905.1
Agalinis harperi	LSU220531	2023	SRR35920872	GCA_056523525.1
Agalinis heterophylla	LSU220561	2021	SRR35317045	GCA_056592735.1
Agalinis homalantha	BRIT617677	2019	SRR35317913	GCA_056319925.1
Agalinis kingsii	FLAS277787	2011	SRR35318410	GCA_056108625.1
Agalinis laxa	NCU19864	1999	SRR35317928	GCA_056108465.1
Agalinis linifolia	NCU452905	2022	SRR35317046	GCA_056108565.1
Agalinis maritima	LSU221519	2023	SRR35318150	GCA_056108645.1
Agalinis navasotensis	BRIT23261	1992	SRR35318151	GCA_056319705.1
Agalinis neoscotica	NCU38058	1990	SRR35318201	GCA_056591835.1
Agalinis obtusifolia	NCU452911	2022	SRR35317882	GCA_056319665.1
Agalinis oligophylla	BRIT713665	2020	SRR35920873	GCA_056108505.1
Agalinis pennelii	MO4037815	1991	SRR39263670	GCA_062057985.1
Agalinis plukenetii	FLAS242274	2014	SRR35920845	GCA_056108545.1
Agalinis pulchella	FLAS235361	2000	SRR35921125	GCA_056108585.1
Agalinis reflexidens	FLAS240887	1984	SRR35319746	GCA_056108525.1
Agalinis scarlatina	FLAS294704	1984	SRR35318210	GCA_056319645.1
Agalinis setacea	NCU4538800	2021	SRR35942161	GCA_056108685.1
Agalinis skinneriana	MO5813713	2004	SRR39263945	GCA_062052788.1
Agalinis stenantha	FLAS240889	1984	SRR35319745	GCA_056319945.1
Agalinis strictifolia	BRIT486317	2016	SRR35920866	GCA_056108705.1
Agalinis tarijensis	MO6565268	2007	SRR39263655	GCA_062057984.1
Agalinis tenuifolia	NCU436723	2022	SRR35923989	GCA_056523405.1
Agalinis viridis	LSU220075	2022	SRR35921010	GCA_056523325.1
Aureolaria pedicularia	NCU454411	2020	SRR35943154	GCA_056523725.1

Acknowledgments

We thank the Missouri Botanical Garden Herbarium (MO), Philecology Herbarium (BRIT), Shirley C. Tucker Herbarium at Louisiana State University (LSU), University of Florida Herbarium (FLAS), and University of North Carolina at Chapel Hill Herbarium (NCU) for access to the specimens sampled. We also thank the Federal University of Minas Gerais Herbarium (BHCB) for allowing us to deposit vouchers of *A. angustifolia*. Plant material of *A. angustifolia* was collected under SISBIO/ICMBio license 96764-1 and exported under SisGen registration A2898C1.

Funding

Sequencing was funded by Iridian Genomes (grant #IRGEN_RG_2021-1345). Fieldwork for the collection of *A. angustifolia* was supported by University of Arkansas start-up funds for M.L.

Published: November 10, 2026 EST.

REFERENCES

Bankevich, A., S. Nurk, D. Antipov, et al. 2012. "SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing." *Journal of Computational Biology* 19: 455–77. https://doi.org/10.1089/cmb.2012.0021.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30: 2114–20. https://doi.org/10.1093/bioinformatics/btu170.

Kieras, M., K. O'Neill, and S. Pirro. 2021. "Zanfona, a Genome Assembly Finishing Tool for Paired-End Illumina Reads." https://github.com/zanfona734/zanfona.

Latvis, M., V. C. Souza, D. C. Tank, P. S. Soltis, and D. E. Soltis. 2024. "An Evolutionary Framework for *Agalinis* (Orobanchaceae; The False Foxgloves) Reveals a Rapid South American Radiation That Includes *Esterhazya*." *Systematic Botany* 49: 455–81. https://doi.org/10.1600/036364424X17095905880222.

Pezzi, P. H., and M. Latvis. 2026. "Chromosome-Level Reference Genome of the Beach False Foxglove, *Agalinis Fasciculata* (Orobanchaceae)." *Genome Biology and Evolution* 18: evag116.

Pezzi, P. H., P. S. Soltis, D. E. Soltis, and M. Latvis. 2026. "'Dispersification' of *Agalinis* (Orobanchaceae) Into South America Is Associated With Hummingbird Pollination and Perennial Life-History Shifts." *Journal of Biogeography* 53: e70143. https://doi.org/10.1111/jbi.70143.

Souza, V. C., J. De Paula-Souza, and M. Latvis. 2025. "(3108) Proposal to Conserve *Agalinis*, nom. Cons. against the Additional Name *Esterhazya* (*Orobanchaceae*)." *TAXON* 74: 994–95. https://doi.org/10.1002/tax.70025.

Tegenfeldt, F., D. Kuznetsov, M. Manni, M. Berkeley, E. M. Zdobnov, and E. V. Kriventseva. 2025. "OrthoDB and BUSCO Update: Annotation of Orthologs with Wider Sampling of Genomes." *Nucleic Acids Research* 53: D516–22. https://doi.org/10.1093/nar/gkae987.