

# Phylogenomics of willows (*Salix* L.) in Alaska

Campbell O. Webb<sup>1,\*</sup>

J. Gordon Burleigh<sup>2</sup>

Stefanie M. Ickert-Bond<sup>1</sup>

1: Univ. of Alaska Museum of the North, Fairbanks, AK

2: Univ. of Florida Dept. of Biology, Gainesville, FL

\* <cowebb@alaska.edu>

## Goals

Willows in Alaska are diverse (50+ species) and ecologically important. Many *Salix* species are allopolyploid, suggesting extensive reticulate evolution. Previous molecular data and methods have not in general been able to resolve phylogenetic relationships in *Salix* with high statistical support, but phylogenomic approaches have been more successful [9]. Using target-enrichment sequencing we aim to i) discover the evolutionary relationships among Alaskan willow taxa, and ii) contribute to the revision of higher-level classification of *Salix*. This poster presents preliminary results.

## Methods

**Sampling:** We obtained dried leaf material from 122 specimens at Univ. of Alaska Museum of the North (ALA); almost all had been identified by George Argus. 43 taxa were included.

**Sequencing:** Library preparation, target enrichment using the GoFlag angiosperm 408 probe set [3], and sequencing on an Illumina HiSeq 3000 were done by RAPiD Genomics (Gainesville, FL). Contigs were assembled by iterative bait assembly using Bridger [4]. In the GoFlag probe set, each target locus is an exon from a single or low copy nuclear gene. For each locus we built alignments for the target exons plus flanking regions.

**Data filtering:** we i) dropped any locus for a sample for which sequence variants had been found for that sample (occurring in 9% of original loci), ii) dropped the 38 loci that showed >2 variants for any samples, iii) removed all loci with data from <90% of samples, and iv) removed any samples with data from <90% of loci. The final dataset had 120 samples and 341 loci (total alignment length of 523,246 BP).

**Phylogenetic analysis:** Maximum likelihood in IQTree2 [7] to generate trees for each locus, for which a species tree was then produced with Astral [6]. In another run we included a legume outgroup. *S. pentandra* was sister to the other willow taxa. The phylogeny was re-rooted on this sample, and annotated using FigTree [8]. Future analyses will use network analyses to explore reticulate relationships.

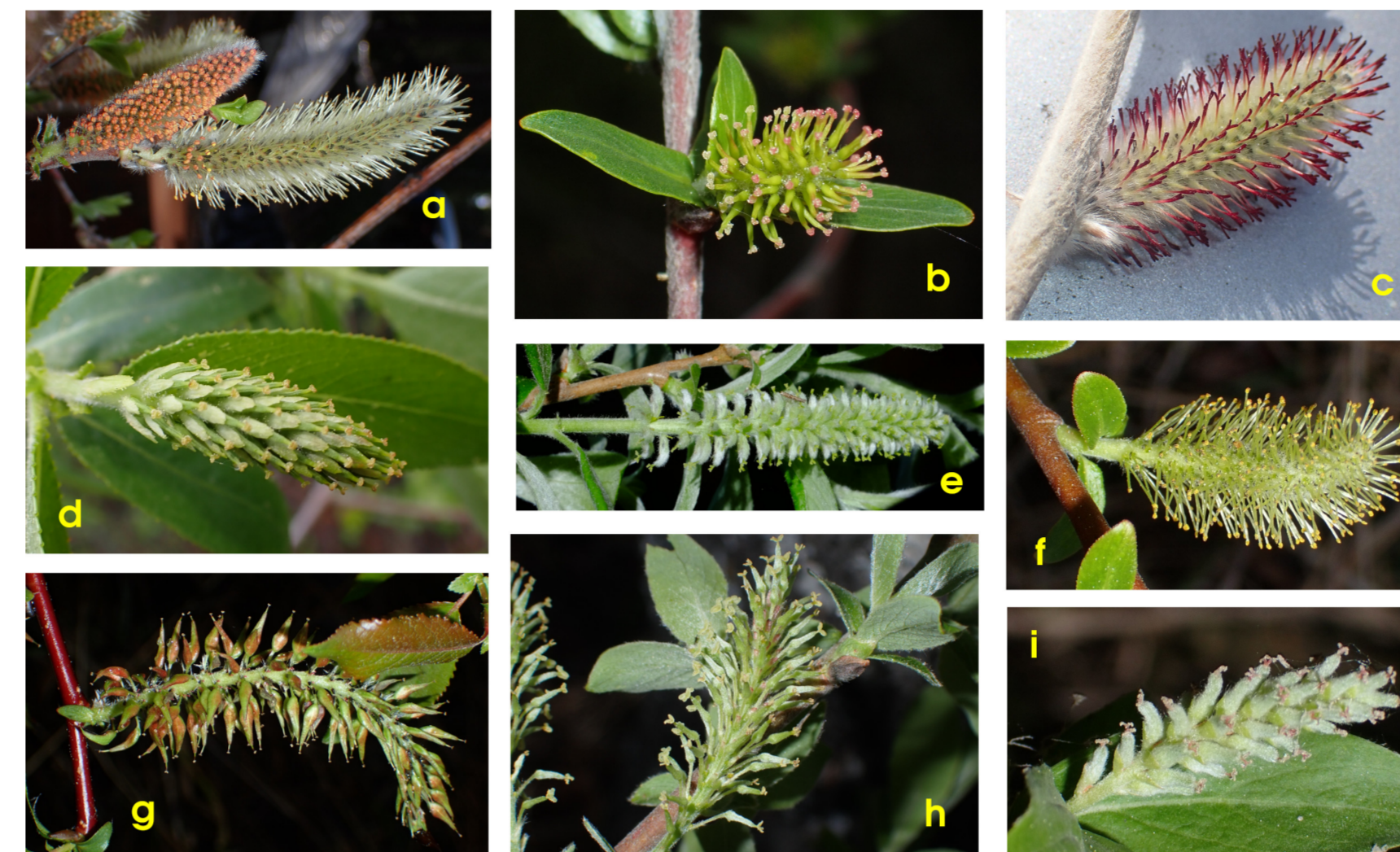


Figure 1 – A selection of Alaska willow species: (a) *S. pulchra* ♂, (b) *S. pseudomyrsinites* ♀, (c) *S. alaxensis* ♀, (d) *S. lasiandra* ♀, (e) *S. glauca* ♀, (f) *S. myrtillofolia* ♂, (g) *S. pseudomonticola* ♀, (h) *S. bebbiana* ♀, (i) *S. niphoclada* ♀. Photographed near North Pole, AK, by Webb.

## Findings

- The phylogeny was largely robust to choices about data (only-exon vs. exon-plus-flanking) and method (IQTree2 ML on concatenated matrix vs. Astral species tree).
- Our results support the patterns found previously [5, 9] of i) a clade of species (“Clade II”) from mixed sections in subgenera *Vetrix* and *Chamaetia* (the former generally being shrub willows and the latter dwarf willows), with ii) *S. reticulata* and *S. interior/exigua* being the outer groups within this clade.
- 22 out of the 32 species with more than one sample are monophyletic in the tree. Overall, more species in section *Vetrix* were monophyletic than in section *Chamaetia*.
- The lack of species monophyly in the clade containing *S. glauca* may reflect the evolutionary porosity of that species. Argus [1] describes it thus: “a high polyploid consisting of tetraploids, hexaploids, and octoploids. The morphological links between *S. glauca* and [other sections] suggest that this species may [...] have evolved through complex hybridization and allopolyploidy with members of those groups [...]. The placement of sect. *Glaucæ* in subg. *Chamaetia* is arbitrary. It could equally well have been placed in subg. *Vetrix*.”
- More sampling is needed to draw conclusions about the monophyly of most sections, but section *Hastatae* may form a natural group.
- The sister relationship of *S. setchelliana* and *S. interior* has not been reported before, but is consistent with field characteristics.

- Argus (1997) doi:10.2307/25096638
- Argus (2010) Flora of North America <https://floranorthamerica.org/Salix>
- Breinholt et al. (2021) doi:10.1002/aps3.11406
- Chang et al. (2015) doi:10.1186/s13059-015-0596-2
- Lauron-Moreau et al. (2015) doi:10.1371/journal.pone.0121965
- Mirarab et al. (2015) doi:10.1093/bioinformatics/btv234
- Nguyen et al. (2015) doi:10.1093/molbev/msu300
- Rambaut et al. (v.1.4.4 = 2018) <https://github.com/rambaut/figtree>
- Wagner et al. (2020) doi:10.3389/fpls.2020.01077

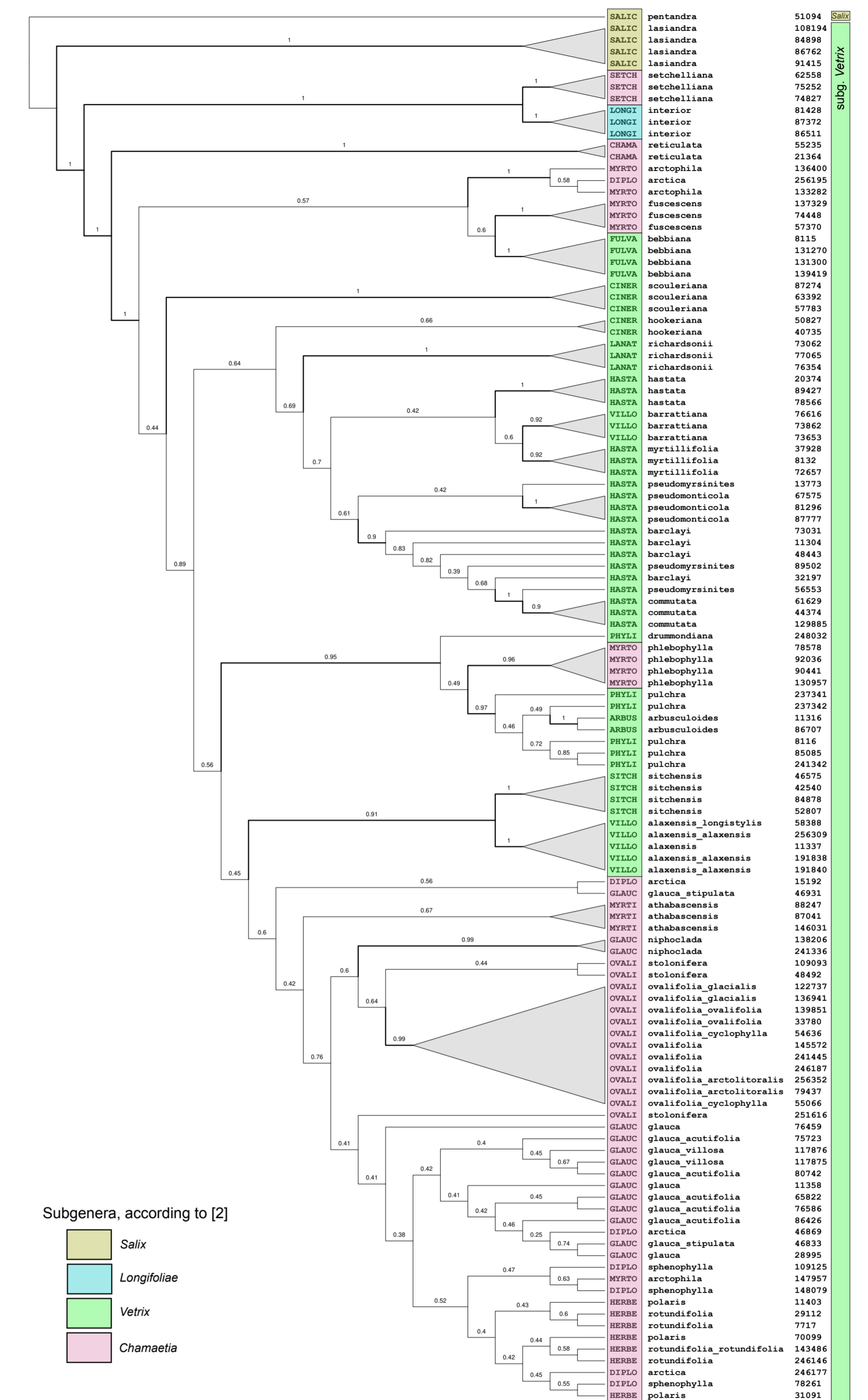


Figure 2 – Cladogram of 119 Alaskan *Salix* samples from 43 taxa. Branch annotation is Astral support value; branches with support  $\geq 0.9$  have wider lines. Monophyletic species are drawn with a triangular clade. Sections and subgenera according to [2] are indicated respectively by the five letter codes preceding species names and the colored boxes around them. Section codes: ARBUS = *Arbuscella*, CHAMA = *Chamaetia*, CINER = *Cinerella*, DIPLO = *Diplodictyae*, FULVA = *Fulvae*, GLAUC = *Glaucæ*, HASTA = *Hastatae*, HERBE = *Herbella*, LANAT = *Lanatae*, LONGI = *Longifoliae*, MYRTI = *Myrtilloides*, MYRTO = *Myrtoisalis*, OVALI = *Ovalifoliae*, PHYLI = *Phylifoliae*, SALIC = *Salicaster*, SETCH = *Setchelliana*, SITCH = *Sitchenses*, VILLO = *Villosae*. The colored bar at the right indicates the likely new subgenus for each species according to [5]. The numbers after species names are the catalog numbers in the Arctos database (prefix with <https://arctos.database.museum/guid/UAM:Herb:> to view online).

Thanks to The National Park Service for funding (Shared Beringian Heritage Program grant to Ickert-Bond).