

Genetic variation and dispersal in *Penstemon hirsutus*
and *P. tenuiflorus*

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Anna Katherine Becker
May 2013

Copyright © 2012 by Anna Becker
All rights reserved.

Acknowledgements

For plant materials, I thank Wayne Barger, Mike Vincent, James Beck, Dwayne Estes and Wayne Elisens. For help troubleshooting molecular methods, I thank Joe May. For nothing, I thank Randall Small. For field, lab, analytical and writing assistance, I thank Steve Furches.

Abstract

Studying plant-pollinator relationships essential for understanding angiosperm evolution. In the large endemic genus *Penstemon* (Plantaginaceae), shifts in pollination syndrome are proposed to be important for explaining taxonomic and morphological diversity (Wilson et al., 2004; Wolfe et al., 2006). However, little work has been done to determine the relationship between morphological and genetic divergence within pollination syndromes. This study utilized genetic data to explore whether divergence in corolla morphology among nine closely related, bee pollinated *Penstemon* species was consistent with pollinator-driven selection.

Bee pollinated species in *Penstemon* subsection *Penstemon* are often divided into two morphological groups based on inflation of the corolla throat. This trait has been proposed to be an important target for pollinator selection (Pennell, 1935; Clements, 1995). Consistent with this theory, phylogenetic analyses of the nuclear gene *GBSSI* indicated that genetic divergence among species was consistent with morphological divergence. However, a similar pattern was reflected by chloroplast data (noncoding region *rps16-trnK*), indicating divergence likely occurred during a period of geographic isolation. Therefore, pollinator-driven selection on corolla morphology does not appear to be the primary cause of divergence among morphological groups. Instead, gene flow patterns indicate that pollinator selection may serve to reinforce pre-existing divergence between groups.

Population genetic analyses utilized microsatellites in addition to *GBSSI* and *rps16-trnK* and focused specifically on two species that share a high degree of morphological similarity. *Penstemon hirsutus* and *P. tenuiflorus* have small average population sizes, fragmented geographic distributions and are pollinated by insects (primarily bees) with short foraging ranges. Patterns consistent with geographic isolation of populations were observed in *P. hirsutus*, but *P. tenuiflorus* nuclear data reflected much lower levels of genetic structure. Although *P. hirsutus* and *P. tenuiflorus* are highly similar, these findings imply that pollinator dynamics differ between species. Therefore, a trait other than corolla throat inflation may be important for explaining differences in pollinator-driven selection. Furthermore, *P. tenuiflorus* may represent an example of Slatkin's paradox. Higher than expected pollen-mediated gene flow among populations of *P. tenuiflorus* suggest that long-distance foragers such as moths may be more important than previously realized.

Table of Contents

Introduction	1
Materials and Methods	3
<i>Plant materials, extraction and amplification</i>	3
<i>Phylogenetic analyses</i>	4
<i>Population genetic analyses</i>	5
Results	7
<i>Phylogenetic analyses</i>	7
<i>Population genetic analyses</i>	8
Discussion	10
<i>Phylogenetic analyses</i>	10
<i>Population genetic analyses</i>	11
<i>Conclusions</i>	12
References	14
Appendices	18
<i>Appendix I</i>	19
<i>Appendix II</i>	32
<i>Appendix III</i>	41
Vita	43

List of Tables

Table 1. Population locality information.....	20
Table 2. <i>rps16-trnK</i> summary statistics for <i>Penstemon hirsutus</i> and <i>P. tenuiflorus</i>	25
Table 3. <i>GBSSI</i> summary statistics for <i>Penstemon hirsutus</i> and <i>P. tenuiflorus</i>	26
Table 4. Microsatellite summary statistics for <i>Penstemon hirsutus</i> and <i>P. tenuiflorus</i>	27
Table 5. AMOVA of <i>Penstemon hirsutus</i> genetic data.	28
Table 6. AMOVA of <i>Penstemon tenuiflorus</i> genetic data.....	29
Table A1. Voucher information and GenBank accession numbers.....	33
Table A2. Species level summary statistics estimated from <i>rps16-trnK</i> and <i>GBSSI</i> data.....	34
Table A3. Estimated null allele frequencies per population and microsatellite locus.....	36
Table A4. <i>rps16-trnK</i> population-level summary statistics for <i>Penstemon hirsutus</i> and <i>P.</i> <i>tenuiflorus</i>	37
Table A5. <i>GBSSI</i> population-level summary statistics for <i>Penstemon hirsutus</i> and <i>P.</i> <i>tenuiflorus</i>	38
Table A6. Microsatellite summary statistics per locus for <i>Penstemon hirsutus</i> and <i>P.</i> <i>tenuiflorus</i>	39
Table A7. Microsatellite population-level summary statistics for <i>Penstemon hirsutus</i> and <i>P.</i> <i>tenuiflorus</i>	40

List of Figures

Figure 1. Distribution of <i>rps16-trnK</i> haplogroups	21
Figure 2. <i>rps16-trnK</i> phylogram	22
Figure 3. Distribution of <i>GBSSI</i> haplogroups.....	23
Figure 4. <i>GBSSI</i> phylogram.....	24
Figure 5. STRUCTURE bar graph for <i>Penstemon hirsutus</i>	30
Figure 6. STRUCTURE bar graph for <i>Penstemon tenuiflorus</i>	31
Figure A1. <i>GBSSI</i> phylogram excluding recombinant alleles.....	35

Introduction

While shifts in pollination syndrome have been recognized as playing important roles in the diversification of angiosperms, the influence of shifting pollinator assemblages has received less attention. Fluctuation in the frequencies of pollinator species due to geography or other factors can result in changes in selection pressure that lead to phenotypic and genetic divergence (Kay and Sargent, 2009). In the diverse North American genus *Penstemon* (Plantaginaceae), attention has been focused on divergence between bird (ornithophilous) and insect (entomophilous) pollinated species (Wilson et al., 2004; Castellanos et al., 2006). Pollination by insects (entomophily, primarily by bees) appears to be the ancestral state, with hummingbird pollination representing a derived state (Pennell, 1935). It has been hypothesized that shifts from bee to bird pollination led to greater taxonomic diversity within the genus. However, entomophily has been retained throughout much of the genus (greater than 80% of approximately 280 species) suggesting dynamics within a pollination syndrome may be of greater importance for generating diversity. This study explored whether divergence in corolla morphology among nine *Penstemon* species characterized by the same pollination syndrome was consistent with pollinator-driven selection by examining patterns of genetic divergence.

The species of *Penstemon* subsection *Penstemon* are primarily distributed east of the Mississippi River and have overlapping geographic ranges (Pennell, 1935; USDA NRCS, 2011). Pollinators include several species of large and small bees and other insects (Pennell, 1935; Clements, 1995). These morphologically similar species are identified primarily based on corolla morphology. Species are divided into two morphological groups based on inflation of the corolla throat. This trait may be an important target for pollinator-driven selection (Pennell, 1935; Clements, 1995). In species such as *P. calycosus*, the throat of the corolla is described as being inflated. In contrast, species such as *P. australis* have narrowed corolla throats and ridged corolla floors. *Penstemon hirsutus* and *P. tenuiflorus* represent an extreme in narrow corolla morphology, in which the corolla is effectively closed by an upward curve in the corolla floor. Species within morphological groups are typically distinguished by more subtle and/or variable characteristics such as length of the corolla and sepals, degree of pubescence, and stem or corolla coloration.

Formal studies comparing the pollination biology of species with inflated (IC) versus narrow corolla (NC) morphology are lacking (although see Wilson et al., 2004 for a more general discussion). However, it has been proposed that this trait may affect pollinator assemblages and reflect variation in pollinator body size (Pennell, 1935; Crosswhite and Crosswhite, 1966; Wilson et al., 2004).

Two different approaches were used to gain insight into the importance of pollinators as drivers of morphological and genetic divergence. First, genetic data were collected from nine species utilizing a non-coding chloroplast region (*rps16-trnK*) and a low copy number nuclear gene (granule-bound starch synthase, or *GBSSI*). Data were analyzed using phylogenetic methods to investigate whether divergence between morphologically defined groups of species (IC versus NC) was reflected in genetic data. Although infrageneric relationships have yet to be resolved, there is evidence that members of subsection *Penstemon* are closely related and have diverged recently (Pennell, 1935; Albach et al., 2005; Wolfe et al., 2006). The non-coding chloroplast region *rps16-trnK* was thus expected to reflect recent shared ancestry among species regardless of morphological group. In contrast, nuclear genes such as *GBSSI* are inherited via pollen and therefore provide information about pollinators. It was predicted that if pollinator-driven selection on corolla morphology was an important driver of divergence between IC and NC groups, patterns observed in *GBSSI* data would follow morphological divisions.

Second, microsatellite data were collected from *Penstemon hirsutus* and *P. tenuiflorus*. Microsatellite data, in conjunction with chloroplast and *GBSSI* data, were analyzed using population genetic methods to examine the relationship between morphological and genetic patterns at a finer scale. Although relatively little is known about seed dispersal in *Penstemon*, it is generally attributed to gravity (barochory) (Clements et al., 2002a). Barochory is associated with relatively localized gene flow (Duminil et al., 2007). Therefore, chloroplast population genetic data were expected to reflect geographically restricted patterns of gene flow, such as highly structured genetic variation and private haplotypes. Pollination syndrome has been more extensively studied than seed dispersal. The majority of *Penstemon* taxa are pollinated by generalist insects (primarily bees) that have relatively small foraging ranges (Pennell, 1935; Crosswhite and Crosswhite, 1966; Osborne et al., 2008). Pollen-mediated gene flow facilitated by short-range insects generally follows a stepping-stone pattern in which geographically proximal populations are connected (Ibrahim et al., 1996). Therefore, microsatellite and *GBSSI* population genetic data were expected to reflect moderately structured genetic variation and fewer private alleles than chloroplast data. Because corolla throat inflation does not vary between morphologically similar *P. hirsutus* and *P. tenuiflorus*, pollinator-driven selection on this trait was not expected to vary between species; observing similar population genetic patterns between these two species would be consistent with this expectation.

Materials and Methods

This study focused on nine *Penstemon* species: *P. australis* Small, *P. brevisepalus* Pennell, *P. calycosus* Small, *P. canescens* (Britt.) Britt., *P. digitalis* Nutt. ex Sims, *P. hirsutus* (L.) Willd., *P. laevigatus* (L.) Ait., *P. pallidus* Small and *P. tenuiflorus* Pennell. These *Penstemon* species are protandrous, self-compatible and outcrossing (Crosswhite and Crosswhite, 1966; Clements, 1995). These species are herbaceous perennials that produce basal rosettes from rhizomes in late winter to early spring, followed by one or more erect stems with opposite, sessile to clasping leaves (Pennell, 1935; Clements, 1995). These self-compatible outcrossers reach sexual reproductive maturity after three to five years and are capable of clonal reproduction via rhizomatous offshoots within the first year (Clements, 1995). Terminal inflorescences of panicles or racemes are present from spring to mid-summer and are pollinated primarily by bees, but wasps and other insects are also common (Pennell, 1935; Clements, 1995). While relatively little is known about seed dispersal in *Penstemon*, most species are described as being predominantly barochorus (Clements et al., 2002a; b).

Penstemon hirsutus and *P. tenuiflorus* have highly similar morphologies and life history traits. They are most often distinguished from one another according to corolla color (white in *P. tenuiflorus*, lavender to white in *P. hirsutus*) and stem and leaf pubescence (Pennell, 1935; Clements, 1995). They are among the few *Penstemon* species for which life history traits and other characteristics have been thoroughly investigated (Koelling, 1964; Clements, 1995).

Plant materials, extraction and amplification

Samples were collected from populations of *Penstemon australis*, *P. brevisepalus*, *P. canescens*, *P. calycosus*, *P. digitalis*, *P. hirsutus*, *P. laevigatus*, *P. pallidus* and *P. tenuiflorus* (Table 1). Samples of *Chionophila jamesii* Benth., *P. roseus* G.Don and *P. apateticus* Straw were also collected to serve as outgroups during phylogenetic analyses. GPS coordinates were collected for each population and voucher specimens for previously undocumented populations were deposited at the University of Tennessee Herbarium (TENN) (Tables 1 and A1). Two mid-cauline leaves were collected from each accession and

stored on ice until transferred into a -80°F freezer for long-term storage. DNA was extracted using DNEASY MINI PLANT KITS (Qiagen, Valencia CA).

The chloroplast intergenic spacer *rps16-trnK* was amplified using primers and protocols from Shaw *et al* (2005) using the following optimized conditions: 90°C for 5min, 30 cycles (94°C for 30sec, 53°C for 30sec, 70°C for 90sec), 72°C for 5min. The nuclear granule-bound starch synthase gene (*GBSSI*) was amplified between the third and eighth exons (primers: Small, unpublished; Appendix III) under the following conditions: 94°C for 5min, 30 cycles (94°C for 30sec, 62°C for 30sec, 72°C for 60sec), 5min at 72°C. Amplification for sequencing followed recommended protocols for BIGDYE CYCLE SEQUENCING KIT (Applied Biosystems, Foster City CA). Sequencing was performed in an AVANT 3600 GENETIC ANALYZER (Applied Biosystems) and electropherograms were analyzed using SEQUENCHER (GeneCodes Corporation, Ann Arbor MI). Chloroplast and nuclear sequences were deposited in GENBANK (Table A1).

Five microsatellite loci were amplified using primers and protocols from Kramer and Fant (2007). As primers were designed using *Penstemon rostriflorus* template DNA (Kramer and Fant, 2007), modified amplification techniques were utilized to minimize the potential for null alleles. Singleplex techniques and/or lower-stringency PCR conditions were utilized when duplex reactions failed. For accessions in which amplification persistently failed, quality of template DNA was assessed by evaluating amplification success rates for other microsatellite loci, *rps16-trnK* and *GBSSI*. Fragment length analyses were performed in an AVANT 3100 GENETIC ANALYZER and fragment sizes were determined using PEAK SCANNER (Applied Biosystems). Alleles were binned using automated methods (FLEXIBIN, Amos *et al.*, 2006) with adjustment as necessary for marginal allele sizes. Tests for linkage disequilibrium and deviation from Hardy-Weinberg equilibrium were performed using GENEPOP (Raymond and Rousset, 1995).

Phylogenetic analyses

Sequences were aligned using MAFFT (Katoh and Toh, 2008) and MESQUITE (Maddison and Maddison, 2011). Redundant sequences were collapsed into haplotypes using ALTER (Glez-Peña *et al.*, 2010). Haplotypic phase of heterozygous *GBSSI* genotypes were inferred using PHASE (Stephens *et al.*, 2001; Stephens and Scheet, 2005) in DNASP (Librado and Rozas, 2009). The recombination model was utilized (MR0); 150 iterations were performed with a thinning interval of two and a burn-in period of 100 iterations. The algorithm was run five times in order to assess convergence and variance. The final run was five times the number of iterations and burn-in length.

Intragenic recombination, which has been documented in *GBSSI* (Okagaki and Weil, 1997; Mason-Gamer et al., 1998), can bias analyses (Posada and Crandall, 2002; Woolley et al., 2008). Therefore, the four-gamete test (Hudson and Kaplan, 1985) and the PDM (Husmeier and Wright, 2001) and DSS (McGuire and Wright, 2000) methods (TOPALI, Milne et al., 2009) were utilized to detect recombination among sequences.

Bayesian phylogenetic analyses were performed using MRBAYES (Ronquist et al., 2012) via CIPRES (Miller et al., 2010). Models of molecular evolution were evaluated using jMODELTEST (Guindon and Gascuel, 2003; Woolley et al., 2008) and Bayes factors (Kass and Raftery, 1995). *rps16-trnK* indels were coded with SEQSTATE (Müller, 2005) using simple indel coding (Simmons and Ochoterena, 2000) and manual methods (per recommendations: Simmons et al., 2007; Ochoterena, 2009). Indels were partitioned prior to analysis. *rps16-trnK* sequence data was analyzed using the JC model, and coded indels were analyzed using variable coding and gamma rates settings. Analysis of *rps16-trnK* data utilized the following parameters: two runs of five chains, 5 000 000 generations, 0.12 temperature, sampling every 1000 generations, four chain swap attempts every five generations and 33% relative burn-in.

GBSSI intron and exon regions were identified by alignment with published exon and mRNA sequences (GENBANK, Benson et al., 2011), evaluated independently for models of molecular evolution (all exons: JC model; all introns: HKY+G model) and partitioned prior to analyses. Analysis of the *GBSSI* data set utilized the following parameters: two runs of four chains, 5 000 000 generations, 0.12 temperature, sampling every 1000 generations, default chain swapping, 45% relative burn-in. A subsampled *GBSSI* data set that excluded recombinant alleles was analyzed independently and utilized the following parameters: two runs of four chains, 5 000 000 generations, 0.12 temperature, sampling every 1000 generations, default chain swapping, 40% relative burn-in. Majority-rule consensus trees were visualized in MESQUITE.

Population genetic analyses

Microsatellite data were evaluated for excess homozygosity associated with the presence of null alleles using MICRO-CHECKER (Van Oosterhout et al., 2004). To assess relative bias per population and locus, null alleles frequencies were estimated using the individual inbreeding model implemented in INEST (Chybicki and Burczyk, 2009).

rps16-trnK, *GBSSI* and microsatellite summary statistics were calculated and AMOVA were performed using ARLEQUIN (Excoffier et al., 1992), DNASP and GENODIVE (Meirmans and Van

Tienderen, 2004). Rarefied allelic richness was estimated from microsatellites using HP-RARE (Kalinowski, 2005).

Bayesian clustering analyses were performed using STRUCTURE (Pritchard et al., 2000; Falush et al., 2003) via BIOPORTAL (Kumar et al., 2009) to identify distinct genetic populations and evaluate population structure. In total, five different analyses were performed using eight data sets. The first analysis focused on the original data sets, which consisted of all available microsatellite data points (i.e. - all individuals for which amplification was successful in at least one locus): (1) *P. hirsutus* data (99 individuals, 11 populations), and (2) *P. tenuiflorus* data (80 individuals, 10 populations). The second through fifth analyses were conducted in order to assess the effects of missing data on results.

For the second analysis, data sets were generated by subsampling the original data sets to include only those individuals with $\leq 40\%$ missing data. The third analysis utilized data sets generated by subsampling the original data sets to include only those loci with $\leq 15\%$ missing data (*Pen02*, *Pen04*, *Pen05*). For the fourth analysis, data sets were generated by adding data from *rps16-trnK*, *GBSSI* and two additional microsatellite loci (*Pen24* and *Pen25*, which were not incorporated into other analyses due to high levels of missing data) to the original microsatellite data sets. The fifth analysis focused on the original microsatellite data sets, but utilized an alternative analysis setting recommended by the authors for data sets that include missing data due to null alleles.

Analyses one through four were performed using “standard” settings for co-dominant data: correlated allele model, population as locality prior, 16 iterations per *K* value (*Penstemon hirsutus* data sets, *K* = 1-18; *P. tenuiflorus* data sets, *K* = 1-18) of 500 000 generations following a burn-in periods of 250 000 generations. The fifth analysis also was performed using these settings, however data were treated as dominant rather than co-dominant marker data.

For each data set and analysis method, the most likely number of genetic clusters (*K*) among individuals was estimated using the methods of Evanno et al. (2005) as implemented in STRUCTURE HARVESTER (Earl and VonHoldt, 2012). In addition, Bayes factors (Kass and Raftery, 1995) were utilized to evaluate whether subsampling to minimize missing data or treating data as dominant was beneficial (in other words, yielded higher probabilities than those observed for original data sets and standard analysis methods).

Results

Phylogenetic analyses

The aligned *rps16-trnK* data set was 933 bp long and contained 46 unique haplotypes (see Table A2 for summary statistics). Two major haplogroups identified by nucleotide and indel polymorphisms were consistent with inflated (IC) and narrow corolla (NC) throat morphological groups, respectively (Figure 1). The NC haplogroup contained two minor haplogroups, NC-1 and NC-2. The only exception was a single *Penstemon canescens* accession (haplotype 2050, population N3) that was morphologically consistent with the NC group but genetically consistent with the IC group. This was reflected in the gene tree produced by Bayesian phylogenetic analysis, which contained two major clades consistent with IC and NC haplogroups (Figure 2). Haplotypes were shared among species within each haplogroup: one haplotype was shared by two IC species and eleven haplotypes were shared by three NC species. *Penstemon brevisepalus* accessions formed a derived clade with strong support (haplotypes 1050 and 1060). However, sampling was limited to a single population; interspecific haplotype sharing would likely be observed if additional populations were sampled.

The aligned *GBSSI* data set was 922 bp long. Fifty-eight unique alleles were observed, nine of which were shared by two to five species (Table A2). Two major haplogroups were differentiated by amino acid and intron nucleotide polymorphisms (Figure 3). As was observed during analysis of chloroplast data, these major haplogroups corresponded to corolla morphology. Three minor haplogroups were identified within the major IC haplogroup (IC-1 through IC-3), and five minor haplogroups were identified within the major NC haplogroup (NC-1 through NC-5). The *GBSSI* gene tree indicated IC and NC haplogroups have diverged (Figure 4). Similar to the *rps16-trnK* tree, the IC haplogroup was not monophyletic due to the placement of a single *Penstemon canescens* accession (allele 250) in a position that was not consistent with morphology. *GBSSI* allele 250 was the only allele shared between IC and NC species (populations N3 and Y1).

One accession sampled from *Penstemon hirsutus* population H6 was heterozygous for alleles (490 and 491) that exhibited unusually high levels of sequence divergence relative to other ingroup alleles (Figure 4). These sequences were included in phylogenetic analyses, but excluded from all other analyses.

Signals corresponding to recombination among *GBSSI* alleles were minimized by excluding sixteen alleles, of which three were private to *Penstemon tenuiflorus*, one was shared between *P. hirsutus*

and *P. tenuiflorus*, and twelve were observed in *P. digitalis*, *P. calycosus* and *P. laevigatus*. The effects of recombination on tree topology are shown in Figure A1. The relationships among NC haplotypes were not affected by recombination. However, IC haplotypes no longer formed a derived group when recombinant alleles were excluded from the analysis.

Population genetic analyses

Tests for linkage disequilibrium among microsatellite loci were not significant (Fisher's exact test, each pairwise locus $p \geq 0.1434$). Populations exhibiting significant deviation from Hardy-Weinberg equilibrium at microsatellite loci included six populations at *Pen02* (three *Penstemon hirsutus*, three *P. tenuiflorus*; $p \leq 0.0278$), three populations at *Pen04* (all *P. hirsutus*; $p \leq 0.0300$), four populations at *Pen05* (three *P. hirsutus*, one *P. tenuiflorus*; $p \leq 0.0298$), two populations at *Pen21* (both *P. hirsutus*; $p \leq 0.0221$), and four populations at *Pen23* (three *P. hirsutus*, one *P. tenuiflorus*; $p \leq 0.0107$).

Statistical analyses of microsatellite data indicated null alleles were likely present at all loci (Table A3). Mean estimated frequency of null alleles per locus was 0.0733 for *Penstemon hirsutus* and 0.0783 for *P. tenuiflorus*. Estimated null allele frequencies per population per locus were not significantly different between species (U-test, $p = 0.2983$). Although 99% of accessions were successfully amplified using *rps16-trnK* and *GBSSI* primers, the majority of failed microsatellite PCR reactions were not recoverable using modified conditions. This may indicate that at least some of the failed amplifications were due to null alleles (i.e. - polymorphism within primer binding sites), but issues with template DNA quality or other factors could not be ruled out as contributing to missing microsatellite data.

Summary statistics for *rps16-trnK*, *GBSSI* and microsatellite loci are provided in Tables 2 – 4 (population level summary statistics in Tables A4 – A7), and AMOVA results are provided in Tables 5 and 6. *rps16-trnK* statistics and AMOVA results indicated that *Penstemon tenuiflorus* populations were more genetically diverse than *P. hirsutus* populations, but that population genetic structure was similar between species. *GBSSI* and microsatellite summary statistics and AMOVA results indicated that levels of genetic diversity were similar between species, but that *P. hirsutus* populations were more highly structured than *P. tenuiflorus* populations.

STRUCTURE is widely utilized because it does not require complex data, results are easy to interpret, and analyses are robust to common issues such as missing data (Falush et al., 2003). While the authors caution the model is not expected to be robust when data are missing due to null alleles, trials utilizing alternative sampling and analysis methods indicate the model may be quite robust in some

circumstances. Compared to standard analyses of the original microsatellite data sets, neither standard analyses of alternative data sets nor analyses of the original data using the “dominant marker” setting affected the most likely K value for *Penstemon hirsutus* or *P. tenuiflorus*. Furthermore, posterior probabilities of inferred K values did not significantly differ regardless of data set or analysis method (Bayes factors for all comparisons: $0.75 > \text{BF} > 1.05$).

Analysis of *Penstemon hirsutus* data indicated the most likely number of clusters was eight ($K = 8$). Clusters were largely consistent with naturally occurring populations (Figures 5 and 6). Missing data and small sample size could have contributed to levels of admixture within populations H9, H10 and H11. However, these populations exhibited allele sharing at chloroplast and *GBSSI* markers as well. The most likely number of clusters for *P. tenuiflorus* was two ($K = 2$). Individuals in populations T1 and T5 were primarily assigned to $K1$ and individuals in other populations were primarily assigned to $K2$. However, individuals in population T6 were admixed.

Discussion

Pollination biology is believed to be important for explaining the diversity and distribution of species in the large endemic genus *Penstemon* (Wilson et al., 2004; Wolfe et al., 2006). Among species belonging to section *Penstemon* subsection *Penstemon*, corolla throat inflation is thought to reflect selection by pollinator assemblages in which body size varies (Pennell, 1935; Crosswhite and Crosswhite, 1966). Although phylogenetic analyses of *GBSSI* sequence data were consistent with this idea, evidence from non-coding chloroplast data indicated pollinator-driven selection on corolla morphology is not likely the primary cause of genetic divergence. In addition, nuclear markers indicated that levels of pollen-mediated gene flow were much higher among *Penstemon tenuiflorus* populations than *P. hirsutus* populations.

Phylogenetic analyses

Phylogenetic analyses of *GBSSI* data revealed genetic divergence among *Penstemon* species that was consistent with variation in corolla morphology and divergence resulting from pollinator selection on corolla morphology. Chloroplast divergence was also consistent with morphological divergence. However, since chloroplast genetic data reflect seed dispersal patterns (Borsch and Quandt, 2009), these findings suggest species in the IC and NC morphological groups likely experienced different biogeographic histories. Therefore, geographic isolation over time likely played an important role in the divergence of morphological groups.

Although morphological divergence appears to be primarily attributable to geographic isolation, pollinator-driven selection on corolla throat inflation may have served to reinforce pre-existing phenotypic and genetic divergence. Chloroplast haplotypes and *GBSSI* alleles were not shared between IC and NC species, even when populations occurred at the same geographic location (Figures 1 – 4). This pattern is striking because many *Penstemon* species have been shown to hybridize successfully – a characteristic long recognized by horticulturalists (Lindgren and Schaaf, 2007). Hybridization between species from different sections and/or with different pollination syndromes (ornithophily and entomophily, e.g.) has been shown to produce fertile hybrid offspring (Lindgren and Schaaf, 2007). Therefore, recently diverged species that share a single pollination syndrome and demonstrate limited

genetic divergence might be predicted to experience gene flow in nature. However, a single *GBSSI* allele (250) represents the only example of sharing between IC and NC species (*P. calycosus* and *P. canescens*). In contrast, several chloroplast haplotypes and *GBSSI* alleles were shared among species within each morphological group (Figures 2 and 4).

Population genetic analyses

Penstemon hirsutus and *P. tenuiflorus* have small average population sizes, fragmented geographic distributions, gravity-dispersed seeds and insect pollinators with relatively short foraging ranges. Pollinator assemblages for each species are believed to be highly similar, with bees serving as primary pollinators (Pennell, 1935; Clements, 1995; Crosswhite and Crosswhite, 1966). Therefore, patterns of genetic divergence consistent with strong genetic drift and geographic isolation were expected, and patterns were expected to be similar between species.

Chloroplast results were largely consistent with expectations for barochory and indicated that genetic structure among populations was similar in *Penstemon hirsutus* and *P. tenuiflorus* (Tables 2, 5 and 6). Patterns were consistent with divergence over time due to geographic isolation of populations. However, *GBSSI* and microsatellite patterns deviated from expectations. Genetic structure among *P. tenuiflorus* populations was much lower than among *P. hirsutus* populations (Tables 3 – 6, Figures 5 and 6). Greater connectivity among *P. tenuiflorus* populations implies levels of pollen-mediated gene flow are higher in this species than in *P. hirsutus*. Pollination in *P. tenuiflorus* may therefore represent an example of Slatkin's paradox (Yu et al., 2010). Slatkin's paradox refers to situations in which population differentiation is lower than expected based on estimated migration (Mallet, 2001). This can occur when effective migration distances of pollinators are underestimated, or when important long-distance pollinators are unknown. This could indicate that pollinator assemblages vary between *P. hirsutus* and *P. tenuiflorus* despite being highly similar in morphology and other life history traits. For *P. tenuiflorus*, long-distance foragers such as moths may be more important than previously believed.

Kramer et al. (2011) investigated the influence of pollination syndrome on gene flow among populations in three species of *Penstemon* in the Great Basin region. They found that hummingbird pollination lead to long-distance gene flow and low differentiation among populations, while pools of generalist insects (primarily bees) resulted in isolation of populations. In this study, patterns observed in *P. hirsutus* nuclear data are consistent with bee pollination. *Penstemon hirsutus* STRUCTURE analyses indicated eight well-defined clusters with relatively limited admixture. However, *P. tenuiflorus*

populations exhibited connectivity more closely resembling that of hummingbird pollinated *P. rostriflorus* populations (Tables 3, 4 and 6 and Figures 3 and 6).

Since the current study focused on the southern portion of the *Penstemon hirsutus* range, a broader study encompassing the entire geographic range may provide a greater understanding of the apparent differences in population genetic structure between *P. hirsutus* and *P. tenuiflorus*. This may help determine whether the greater connectivity observed among *P. tenuiflorus* populations in nuclear data is attributable to differences in pollinators. If so, variation in corolla color or other characteristics yet to be recognized by investigators likely represent more important targets of pollinator selection than corolla throat inflation.

Conclusions

Little work has been done to determine the relationship between morphological and genetic divergence within pollination syndromes in *Penstemon*. This study examined nine species with similar morphologies and found that while genetic divergence was correlated with morphological divergence, differences are likely due to geographic isolation rather than pollinator differences. However, gene flow patterns suggest that pollinators serve to reinforce pre-existing divergence through selection on morphological traits. Until now, inflation of the corolla throat was thought to be the most important target of pollinator selection, thus explaining morphological divergence in *Penstemon* subsection *Penstemon*. However, this study demonstrated that species with similar corolla morphology experience different pollinator dynamics. Therefore, characteristics that have yet to be recognized by investigators may represent the most important targets of pollinator selection.

In the past, human perception of corolla traits has directed the focus of pollination biology studies. However, a study by Ollerton et al. (2009) concluded that traditional pollination syndromes often neglect the primary targets of pollinator selection. As recent studies demonstrate, the traits that are important to pollinators may not be obvious to humans. In one example, the increase in relative humidity produced by the evaporation of nectar was found to be a more important floral cue for foraging hawk moths than corolla color, shape or scent (Von Arx et al., 2012). Another study on *Penstemon digitalis* demonstrated that the concentration of linalool in nectar, which affects floral scent, is under stronger selection than corolla color or size (Parachnowitsch et al., 2013). In combination with the findings presented here, these studies emphasize that human perception is not a reliable means of identifying

which traits are the most important targets of pollinator selection. Applying this knowledge to future research will advance our understanding of the *Penstemon* genus and of angiosperm evolution in general.

References

- Albach, D., H. Meudt, and B. Oxelman. 2005. Piecing together the “new” Plantaginaceae. *American Journal of Botany* 92: p.297–315.
- Amos, W., J.I. Hoffman, A. Frodsham, L. Zhang, S. Best, and A.V.S. Hill. 2006. Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes* 7: p.10–14.
- Benson, D.A., I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, and E.W. Sayers. 2011. GenBank. *Nucleic Acids Research* 39: p.D32–7.
- Borsch, T., and D. Quandt. 2009. Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. *Plant Systematics and Evolution* 282: p.169–199.
- Castellanos, M.C., P. Wilson, S.J. Keller, A.D. Wolfe, and J.D. Thomson. 2006. Anther evolution: pollen presentation strategies when pollinators differ. *The American Naturalist* 167: p.288–96.
- Chybicki, I.J., and J. Burczyk. 2009. Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity* 100: p.106–13.
- Clements, R.K. 1995. Studies on the taxonomy and general ecology of the two closely-related species *Penstemon tenuiflorus* Pennell and *P. hirsutus* (L.) Willd. (Scrophulariaceae section Graciles). University of Kentucky.
- Clements, R.K., J. Baskin, and C. Baskin. 2002a. The Comparative Biology of the Two Closely-Related Species *Penstemon tenuiflorus* Pennell and *P. hirsutus* (L.) Willd. (Scrophulariaceae, Section Graciles): III. Ecological Life Cycle, Growth Characteristics, and Flowering Requirements. *Castanea* 67: p.161–176.
- Clements, R.K., J.M. Baskin, and C.C. Baskin. 2002b. The Comparative Biology of the Two Closely-Related Species *Penstemon tenuiflorus* Pennell and *P. hirsutus* (L.) Willd. (Scrophulariaceae, Section Graciles): IV. Effects of Shade, Drought, and Soil Type on Survival and Growth. *Castanea* 67: p.177–187.
- Crosswhite, F.S., and C.D. Crosswhite. 1966. Insect pollinators of *Penstemon* series Graciles (Scrophulariaceae) with notes on *Osmia* and other Megachilidae. *American Midland Naturalist* 76: p.450–467.
- Duminil, J., S. Fineschi, A. Hampe, and P. Jordano. 2007. Can Population Genetic Structure Be Predicted from Life-History Traits? *The American Naturalist* 169: p.662–672.
- Earl, D., and B. VonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: p.359–361.
- Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: p.2611–20.
- Excoffier, L., P.E. Smouse, and J.M. Quattro. 1992. Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics* 131: p.479–491.
- Falush, D., M. Stephens, and J.K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: p.1567–87.
- Glez-Peña, D., D. Gómez-Blanco, M. Reboiro-Jato, F. Fdez-Riverola, and D. Posada. 2010. ALTER: program-oriented conversion of DNA and protein alignments. *Nucleic Acids Research* 38: p.W14–W18.
- Guindon, S., and O. Gascuel. 2003. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology* 52: p.696–704.
- Hudson, R.R., and N.L. Kaplan. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* 111: p.147–64.
- Husmeier, D., and F. Wright. 2001. Probabilistic divergence measures for detecting interspecies recombination. *Bioinformatics* 17 Suppl 1: p.S123–31.
- Ibrahim, K.M., R.A. Nichols, and G.M. Hewitt. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Journal of Heredity* 77: p.282–291.
- Kalinowski, S.T. 2005. HP-RARE 1.0: a Computer Program for Performing Rarefaction on Measures of Allelic Richness. *Molecular Ecology Notes* 5: p.187–189.
- Kass, R., and A.E. Raftery. 1995. Bayes factors. *Journal of the American Statistical Association* 90: p.773–795.
- Katoh, K., and H. Toh. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* 9: p.286–98.

- Kay, K.M., and R.D. Sargent. 2009. The Role of Animal Pollination in Plant Speciation: Integrating Ecology, Geography, and Genetics. *Annual Review of Ecology, Evolution, and Systematics* 40: p.637–656.
- Koelling, A.C. 1964. Taxonomic studies in *Penstemon deamii* and its allies. University of Illinois, Urbana.
- Kramer, A.T., and J.B. Fant. 2007. Isolation and characterization of microsatellite loci in *Penstemon rostriflorus* (Plantaginaceae) and cross-species amplification. *Molecular Ecology Notes* 7: p.998–1001.
- Kramer, A.T., J.B. Fant, and M. V. Ashley. 2011. Influences of landscape and pollinators on population genetic structure: Examples from three *Penstemon* (Plantaginaceae) species in the Great Basin. *American Journal of Botany* 98: p.109–21.
- Kumar, S. et al. 2009. AIR: A batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. *BMC Bioinformatics* 10: p.357.
- Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: p.1451–2.
- Lindgren, D.T., and D.M. Schaaf. 2007. *Penstemon*: A summary of interspecific crosses. *HortScience* 42: p.494–498.
- Maddison, W.P., and D.R. Maddison. 2011. Mesquite: a modular system for evolutionary analysis.
- Mallet, J. 2001. Gene Flow. In I. Woiwod, D. Reynolds, and C. Thomas [eds.], *Insect Movement: Mechanisms and Consequence*, 337–360. CABI Publishing, CAB International, Oxon, UK.
- Mason-Gamer, R.J., C.F. Weil, and E.A. Kellogg. 1998. Granule-bound starch synthase: structure, function, and phylogenetic utility. *Molecular Biology and Evolution* 15: p.1658–73.
- McGuire, G., and F. Wright. 2000. TOPAL 2.0: improved detection of mosaic sequences within multiple alignments. *Bioinformatics* 16: p.130–4.
- Meirmans, P.G., and P.H. van Tienderen. 2004. Genotype and Genodive: Two Programs for the Analysis of Genetic Diversity of Asexual Organisms. *Molecular Ecology Notes* 4: p.792–794.
- Miller, M., W. Pfeiffer, and T. Schwartz. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE)*, 1–8. New Orleans, LA.
- Milne, I., D. Lindner, M. Bayer, D. Husmeier, G. McGuire, D.F. Marshall and F. Wright. 2009. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics* 25: p.126–127.
- Müller, K. 2005. SeqState: primer design and sequence statistics for phylogenetic DNA datasets. *Applied Bioinformatics* 4: p.65–69.
- Ochoterena, H. 2009. Homology in coding and non-coding DNA sequences: a parsimony perspective. *Plant Systematics and Evolution* 282: p.151–168.
- Okagaki, R.J., and C.F. Weil. 1997. Analysis of Recombination Sites Within the Maize waxy Locus. *Genetics* 147: p.815–821.
- Ollerton, J., R. Alarcón, N.M. Waser, M.V. Price, S. Watts, L. Cranmer, A. Hingston, C.I. Peter, and J. Rotenberry. 2009. A global test of the pollination syndrome hypothesis. *Annals of Botany* 103: p.1471–80.
- Van Oosterhout, C., W.F. Hutchinson, D.P.M. Wills, and P. Shipley. 2004. Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. *Molecular Ecology Notes* 4: p.535–538.
- Osborne, J.L., A.P. Martin, N.L. Carreck, J.L. Swain, M.E. Knight, D. Goulson, R.J. Hale, and R.A. Sanderson. 2008. Bumblebee flight distances in relation to the forage landscape. *Journal of Animal Ecology* 77: p.406–15.
- Parachnowitsch, A., R.C.F. Burdon, R.A. Raguso, and A. Kessler. 2013. Natural selection on floral volatile production in *Penstemon digitalis*: Highlighting the role of linalool. *Plant Signaling & Behavior* 8(1).
- Pennell, F.W. 1935. The Scrophulariaceae of eastern temperate North America. The Academy of Natural Sciences of Philadelphia, Philadelphia.
- Posada, D., and K.A. Crandall. 2002. The effect of recombination on the accuracy of phylogeny estimation. *Journal of Molecular Evolution* 54: p.396–402.
- Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype

- data. *Genetics* 155: p.945–59.
- Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: p.248.
- Ronquist, F., M. Teslenko, P. van der Mark, D.L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M.A. Suchard, and J.P. Hulsbeck. 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology* 63: p.539–542.
- Simmons, M.P., K. Müller, and A.P. Norton. 2007. The relative performance of indel-coding methods in simulations. *Molecular Phylogenetics and Evolution* 44: p.724–40.
- Simmons, M.P., and H. Ochoterena. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* 49: p.369–81.
- Stephens, M., and P. Scheet. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *American Journal of Human Genetics* 76: p.449–62.
- Stephens, M., N.J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* 68: p.978–89.
- USDA NRCS. 2011. The PLANTS Database. Available at: <http://plants.usda.gov> [Accessed November 7, 2011].
- Wilson, P., M.C. Castellanos, J.N. Hogue, J.D. Thomson, and W.S. Armbruster. 2004. A multivariate search for pollination syndromes among penstemons. *Oikos* 2: p.345–361.
- Wolfe, A.D., C.P. Randle, S.L. Datwyler, J.J. Morawetz, N. Arguedas, and J. Diaz. 2006. Phylogeny, taxonomic affinities, and biogeography of *Penstemon* (Plantaginaceae) based on ITS and cpDNA sequence data. *American Journal of Botany* 93: p.1699–713.
- Woolley, S.M., D. Posada, and K.A. Crandall. 2008. A comparison of phylogenetic network methods using computer simulation. *PloS One* 3: p.e1913.
- von Arx, M., J. Goyret, and E. Al. 2012. Floral humidity as a reliable sensory cue for profitability assessment by nectar-foraging hawkmoths. *Proceedings of the National Academy of Sciences* 109: p.9471–9476.
- Yu, H., J.D. Nason, X. Ge, and J. Zeng. 2010. Slatkin's Paradox: when direct observation and realized gene flow disagree. A case study in *Ficus*. *Molecular Ecology* 19: p.4441–53.

Appendices

Appendix I

Table 1. Population locality information.

Species	Pop. ID	N	County	State	Latitude	Longitude
<i>P. australis</i> Small	PA1	1	Emanuel	GA	32.3758°N	82.3575°W
	PA2	1	Rockcastle	KY	37.2998°N	84.1619°W
	PA3	2	Picket	TN	36.5624°N	85.0790°W
<i>P. brevisepalus</i> Pennell	B	5	Knox	TN	35.9687°N	83.6873°W
<i>P. calycosus</i> Small	Y1	2	Maury	TN	36.5772°N	86.9210°W
	Y4	1	Barren	KY	36.8788°N	86.0453°W
	Y5	1	Sumner	TN	36.4647°N	86.4562°W
<i>P. canescens</i> (Britt.) Britt.	N1	1	Rockcastle	KY	37.8594°N	84.4367°W
	N2	1	Lee	AL	32.6100°N	85.4300°W
	N3	1	Cherokee	AL	34.3885°N	85.5910°W
	N4	1	Dekalb	AL	34.5293°N	85.5934°W
	N5	1	Picket	TN	36.5624°N	85.0790°W
<i>P. digitalis</i> Nutt. ex Sims	D2	3	Franklin	AL	34.4886°N	87.5320°W
	D4	1	Christian	KY	37.0248°N	87.4093°W
	D5	20	Greene	PA	39.9716°N	80.1246°W
<i>P. hirsutus</i> (L.) Willd.	H1	10	Montgomery	TN	36.5016°N	87.3235°W
	H2	10	Nicholas	KY	38.4164°N	84.0020°W
	H3	10	Franklin	KY	38.1880°N	84.8797°W
	H4	10	Franklin	KY	38.2808°N	84.7975°W
	H5	11	Franklin	KY	38.2797°N	84.7919°W
	H6	11	Smith	TN	36.2678°N	85.9631°W
	H7	8	Jackson	TN	36.4109°N	85.6445°W
	H8	5	Meigs	TN	35.4365°N	84.9792°W
	H9	16	Jessamine	KY	37.8223°N	84.7223°W
	H10	6	Fayette	KY	37.8862°N	84.3402°W
	H11	2	Pendleton	WV	38.7042°N	79.4157°W
<i>P. laevigatus</i> (L.) Ait.	L1	6	Gadsden	FL	30.7012°N	84.8336°W
	L3	1	Fayette	KY	37.8594°N	84.4367°W
<i>P. pallidus</i> Small	P1	3	Rockcastle	KY	37.8594°N	87.4367°W
	P2	1	Picket	TN	36.5624°N	85.0790°W
<i>P. tenuiflorus</i> Pennell	T1	13	Maury	TN	35.6634°N	86.8407°W
	T2	10	Bibb	AL	33.0792°N	86.9802°W
	T3	10	Lawrence	AL	34.4859°N	87.5005°W
	T4	10	Rutherford	TN	36.0735°N	86.5225°W
	T5	10	Maury	TN	35.5874°N	86.8975°W
	T6	10	Wilson	TN	36.0916°N	86.3309°W
	T7	13	Bibb	AL	33.0707°N	87.1296°W
	T8	1	Marshall	TN	35.6063°N	86.7416°W
	T9	13	Muhlenberg	KY	37.2223°N	87.1113°W
	T10	4	Sumner	TN	36.4638°N	86.4559°W
	T11	6	Davidson	TN	36.1012°N	86.5295°W
	T12	1	Dallas	AL	32.3182°N	87.1269°W

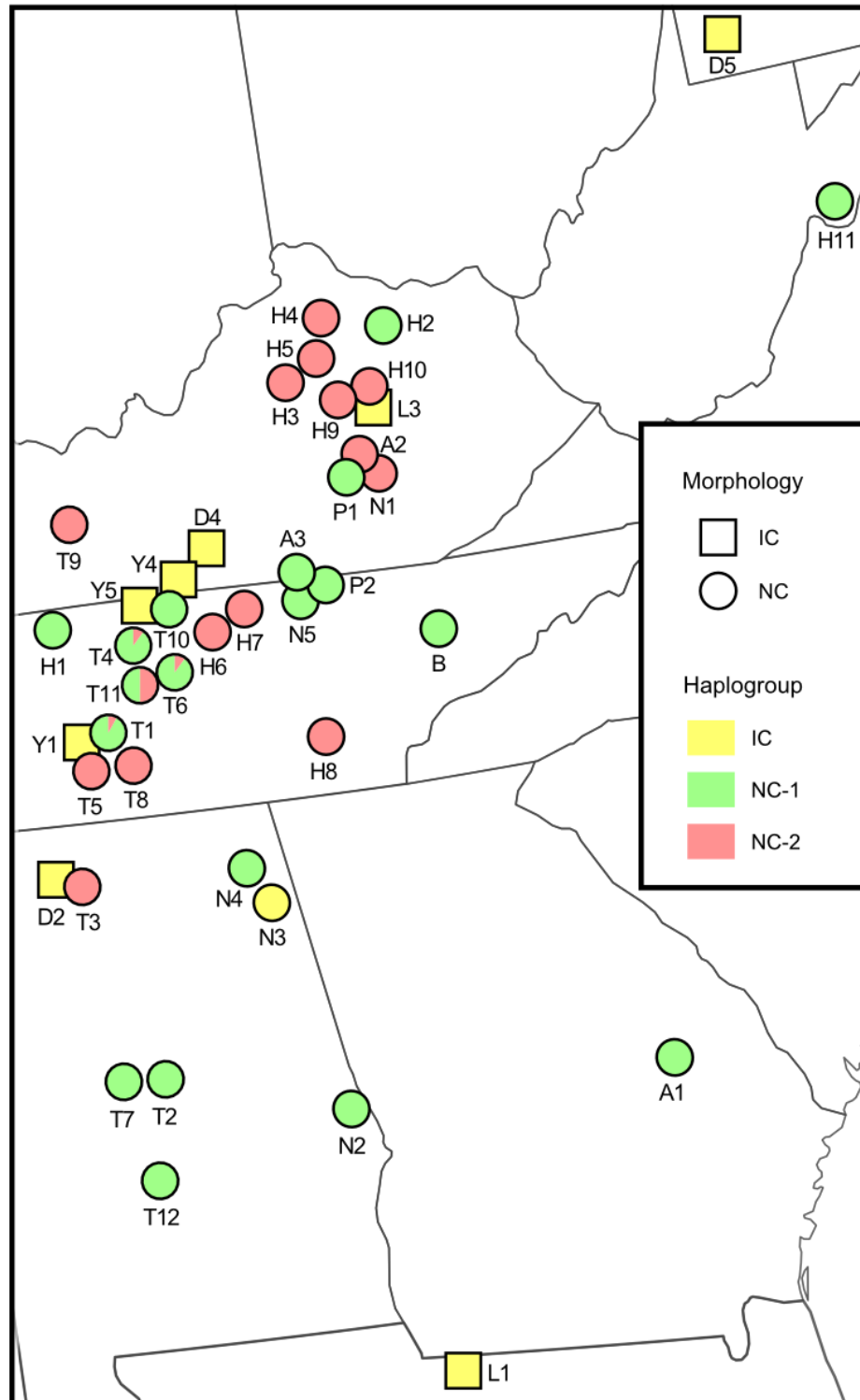


Figure 1. Distribution of *rps16-trnK* haplogroups. Populations of inflated and narrow corolla throat species are represented by circles or squares, respectively (see Table 1 for population identifiers). Relative frequency of each haplogroup within a population is represented as a pie chart.

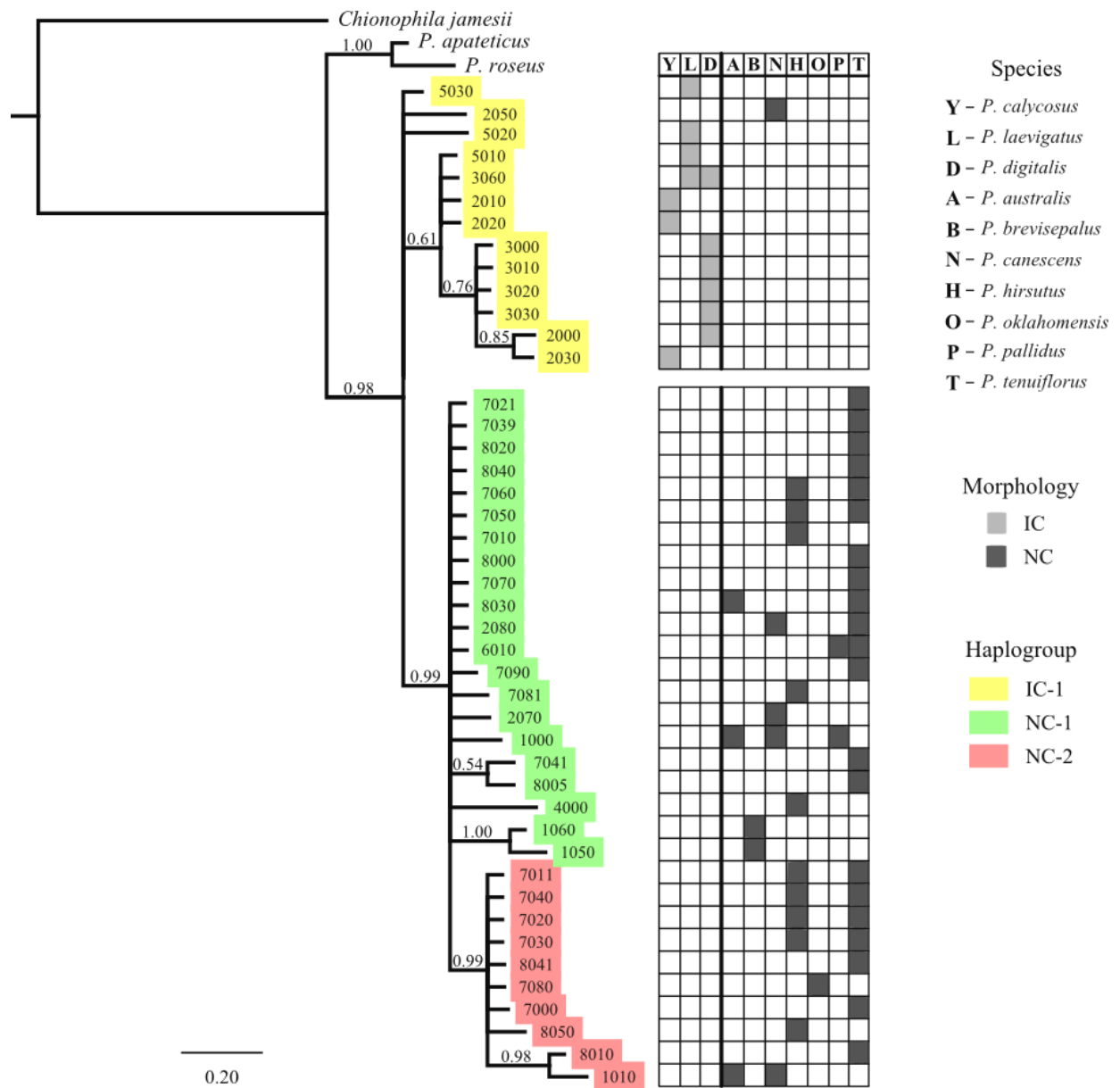


Figure 2. *rps16-trnK* phylogram. Tip labels represent haplotypes. Posterior probabilities are indicated on branches. Branch lengths are proportional to changes per site. The distribution of haplotypes and haplogroups among species and morphological groups is indicated in the table by gray squares.

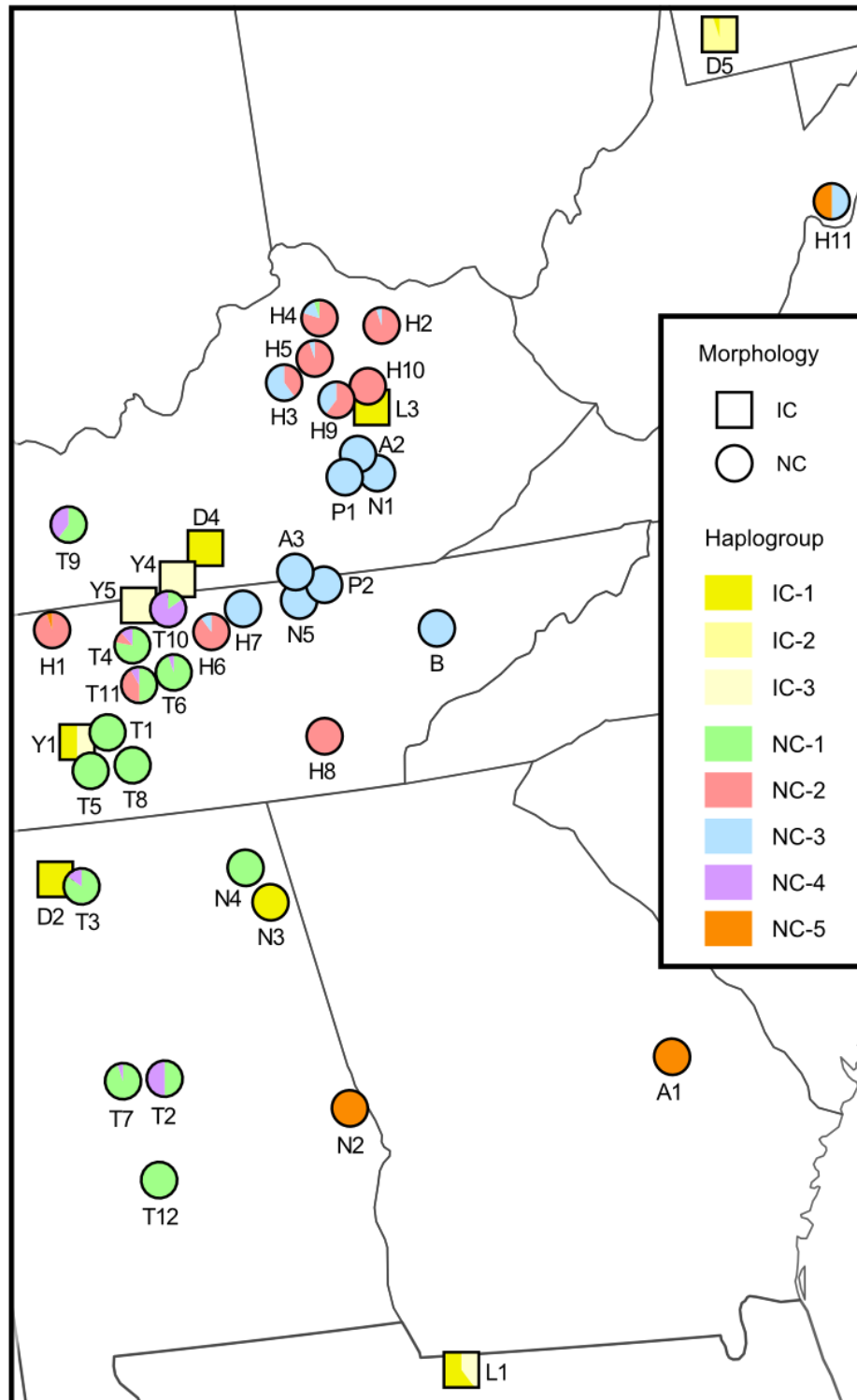


Figure 3. Distribution of *GBSSI* haplogroups. Populations of inflated and narrow corolla throat species are represented by circles or squares, respectively (see Table 1 for population identifiers). Relative frequency of each haplogroup within a population is represented as a pie chart.

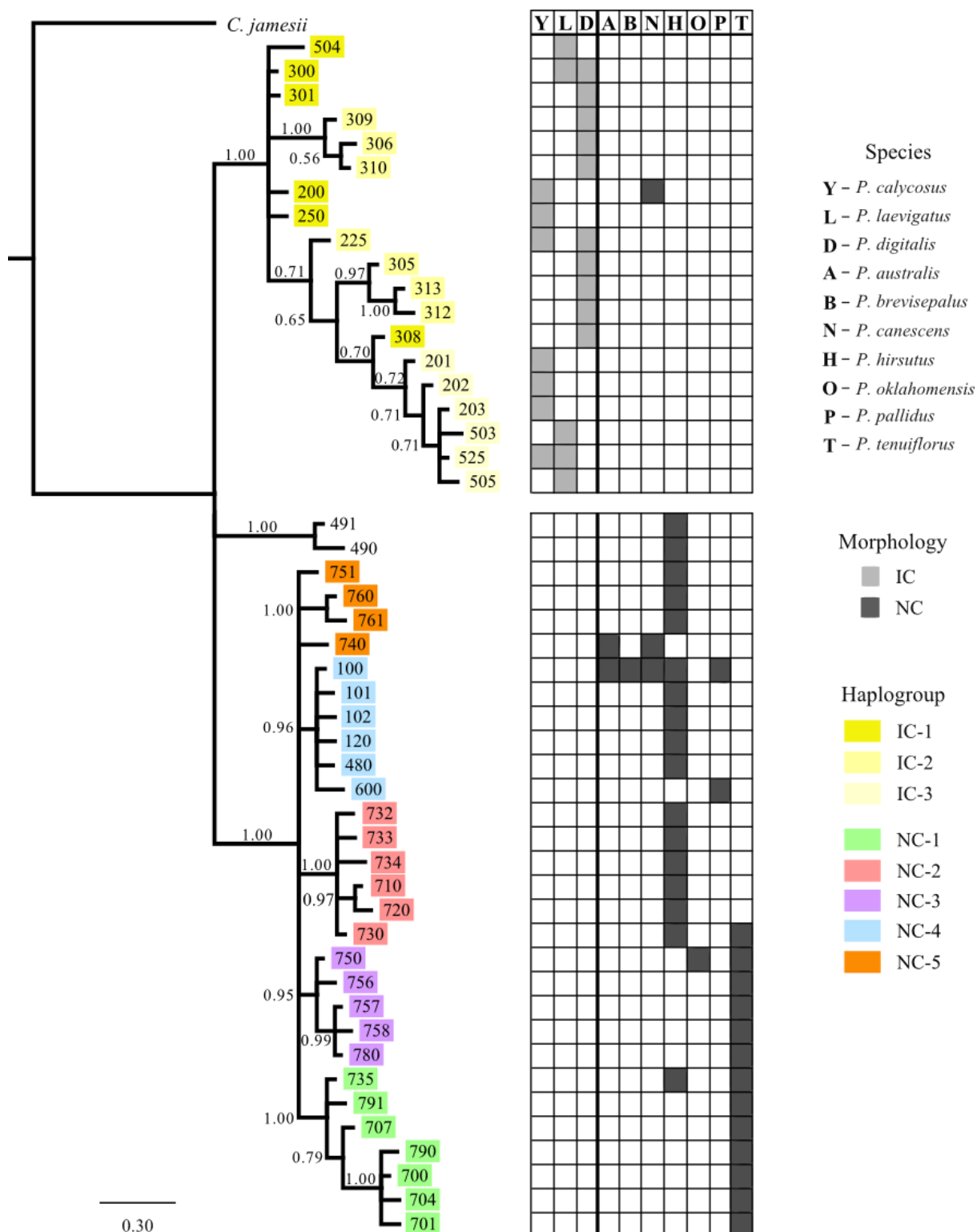


Figure 4. GBSSI phylogram. Tip labels represent haplotypes. Posterior probabilities are indicated on branches. Branch lengths are proportional to changes per site. The distribution of haplotypes and haplogroups among species and morphological groups is indicated in the table by gray squares.

Table 2. *rps16-trnK* summary statistics for *Penstemon hirsutus* and *P. tenuiflorus*.

	<i>P. hirsutus</i>	<i>P. tenuiflorus</i>
N	99 (11)	101 (12)
h	10	21
h _p	4	14
H _T	0.791 ±0.030	0.924 ±0.009
F _{ST}	0.691*	0.716*
k _T	1.55 ±0.94	2.27 ±1.26
Φ _{ST}	0.821*	0.835*

N = number of individuals (number of populations)

h = number of unique haplotypes

h_p = number of private haplotypes

H_T = gene diversity

k_T = mean number pairwise nucleotide differences between haplotypes

* $P \leq 0.05$

Table 3. GBSSI summary statistics for *Penstemon hirsutus* and *P. tenuiflorus*.

	<i>P. hirsutus</i>	<i>P. tenuiflorus</i>
N	98	100
A	15	14
A _p	12	11
H _O	0.378	0.320
H _T	0.566	0.489
F _{IS}	0.074	0.150*
F _{ST}	0.301*	0.251*
k _T	2.03 ±1.15	2.58 ±1.39
Φ _{IS}	0.103	0.160
Φ _{ST}	0.405*	0.264*

N = number of accessions

S = number of polymorphic nucleotide sites

S_p = number of private (single species) polymorphic sites

π = mean nucleotide diversity per allele

A = number of unique alleles

A_p = number of private (single species) alleles

H_O = observed heterozygosity

H_T = total expected heterozygosity

k_T = mean number of pairwise nucleotide differences

**P* ≤ 0.05

Table 4. Microsatellite summary statistics for *Penstemon hirsutus* and *P. tenuiflorus*.

	<i>P. hirsutus</i>	<i>P. tenuiflorus</i>
N	75.80 ±12.28	59.40 ±15.57
missing	0.23	0.41
A	26.20 ±8.32	23.80 ±8.53
A _p	5.80	4.60
R	20.39	20.28
R _p	7.92	7.81
H _O	0.603 ±0.162	0.665 ±0.088
H _T	0.908 ±0.053	0.904 ±0.053
F _{IS}	0.176*	0.142*
F _{ST}	0.216*	0.149*
K _T	2.75	2.63

N = mean number of sampled individuals per locus

missing = mean proportion missing data per locus

A = mean number of alleles per locus

A_p = mean number of private (single species) alleles per locus

R = mean rarefied allelic richness per locus

R_p = mean rarefied private (single species) allelic richness per locus

H_O = mean observed heterozygosity per locus

H_T = mean total expected heterozygosity per locus

* $P \leq 0.05$

Table 5. AMOVA of *Penstemon hirsutus* genetic data.**(a)** *rps16-trnK*

source	Φ -statistic				<i>F</i> -statistic			
	DF	SS	VC	%V	DF	SS	VC	%V
among populations	10	62.83	0.692*	82.13	10	27.25	0.293*	69.14
within populations	88	13.25	0.151	17.87	88	11.50	0.131	30.86
Total	98	76.08	0.843		98	38.75	0.423	

(b) *GBSSI*

source	Φ -statistic				<i>F</i> -statistic			
	DF	SS	VC	%V	DF	SS	VC	%V
among populations	10	83.99	0.438*	40.54	10	17.620	0.088*	30.13
among individuals, within populations	87	61.67	0.066	6.12	87	19.064	0.015	5.20
within individuals	98	53.50	0.577*	53.34	98	18.500	0.189*	64.68
Total	195	202.16	1.081		195	55.184	0.292	

(c) microsatellite loci

source	<i>F</i> -statistic		
	SS	VC	%V
among populations	84.904	0.503*	21.62
among individuals, within populations	140.059	0.321*	13.79
within individuals	114.500	1.503*	64.59
Total	339.463	2.326	

SS: sum of squares

VC: variance component

DF: degrees of freedom

%V: percent variance

* $P \leq 0.05$

Table 6. AMOVA of *Penstemon tenuiflorus* genetic data.**(a) *rps16-trnK***

source	Φ -statistic				<i>F</i> -statistic			
	DF	SS	VC	%V	DF	SS	VC	%V
among populations	11	167.35	1.806*	83.52	11	33.66	0.355*	71.58
within populations	89	31.72	0.356	16.48	89	12.54	0.141	28.42
Total	100	199.07	2.162		100	46.20	0.496	

(b) *GBSSI*

source	Φ -statistic				<i>F</i> -statistic			
	DF	SS	VC	%V	DF	SS	VC	%V
among populations	11	72.219	0.350*	26.43	11	13.644	0.063*	25.05
among individuals, within populations	88	99.541	0.156	11.73	88	19.036	0.028*	11.22
within individuals	100	82.000	0.820*	61.84	100	16.000	0.160*	63.74
Total	199	256.760	1.326		199	48.680	0.251	

(c) microsatellite loci

source	<i>F</i> -statistic		
	SS	VC	%V
among populations	46.723	0.345*	14.85
among individuals, within populations	100.416	0.281*	12.09
within individuals	91.500	1.696*	73.06
Total	238.638	2.322	

SS: sum of squares

VC: variance component

DF: degrees of freedom

%V: percent variance

* $P \leq 0.05$

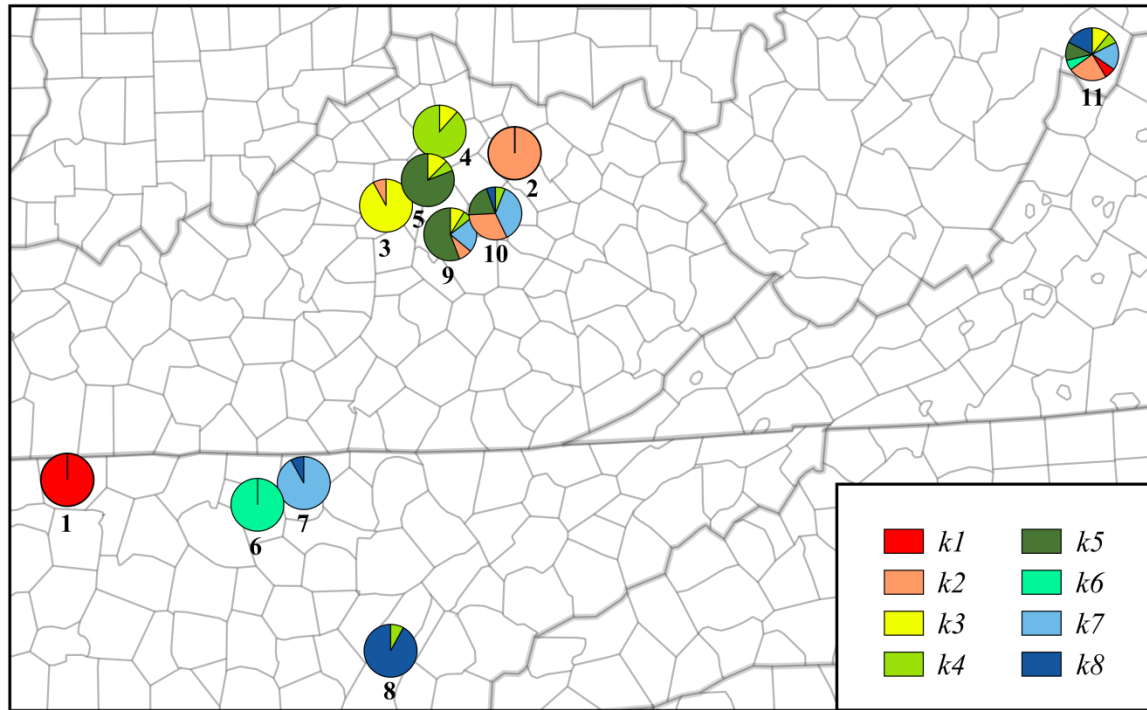
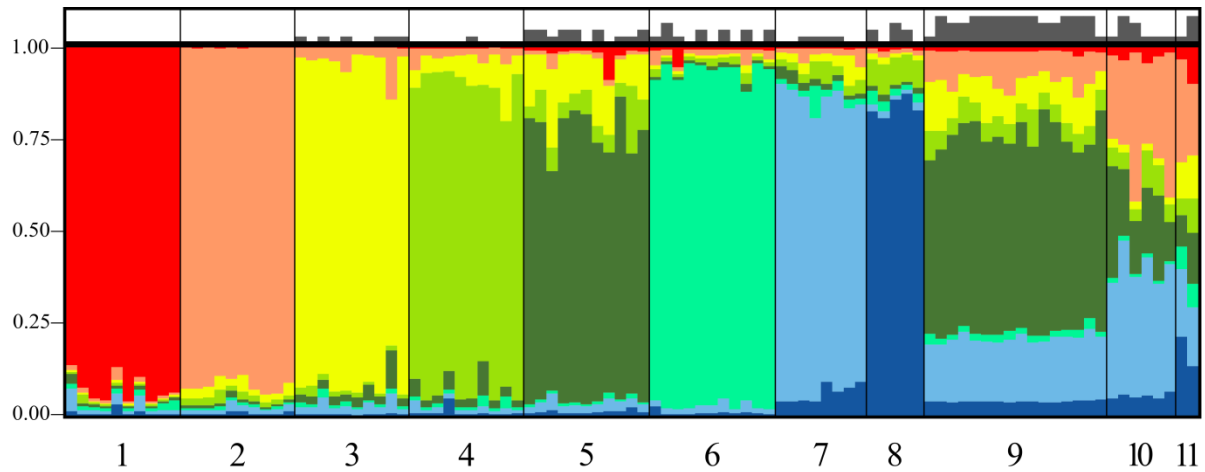


Figure 5. STRUCTURE bar graph for *Penstemon hirsutus*. Bayesian assignment probabilities of individuals to genetic clusters ($K = 8$) are represented in the bar graph. Gray bars above the STRUCTURE bar graph represent the proportion of missing microsatellite data for each individual. The geographic distribution of genetic clusters among populations is represented in the map below the bar graph.

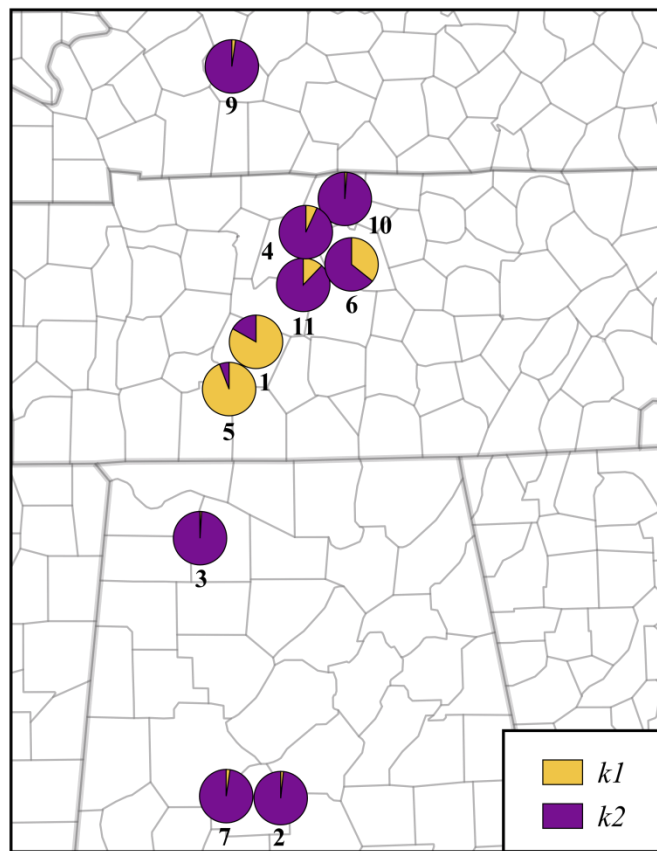
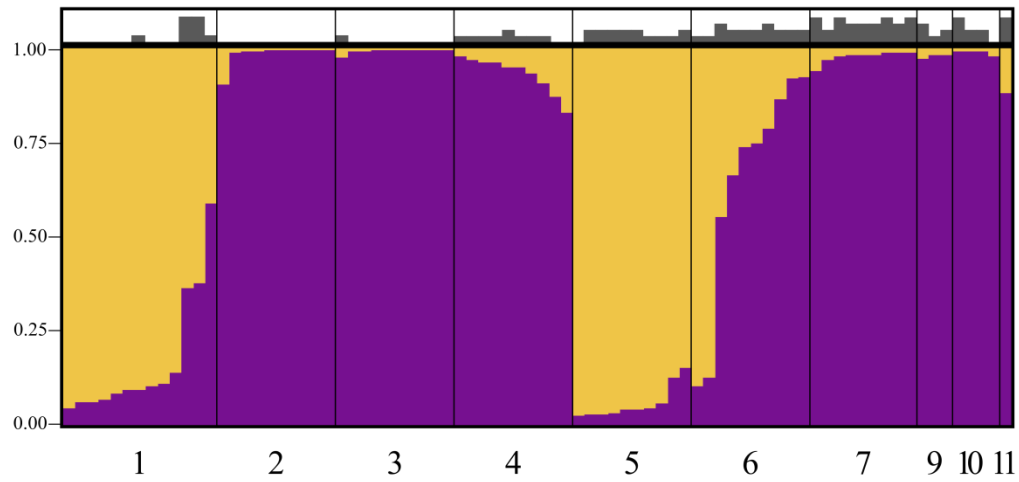


Figure 6. STRUCTURE bar graph for *Penstemon tenuiflorus*. Bayesian assignment probabilities of individuals to genetic clusters ($K = 2$) are represented in the bar graph. Gray bars above the STRUCTURE bar graph represent the proportion of missing microsatellite data for each individual. The geographic distribution of genetic clusters among populations is represented in the map below the bar graph.

Appendix II

Table A1. Voucher information and GenBank accession numbers.

Pop.	voucher	<i>rps16-trnK</i> GenBank no.	<i>GBSSI</i> GenBank no.
A1	<i>A. Furches</i> 1010 (TENN)	JX967032	JX985367
A2	<i>H. Ahles</i> 279064 (NCU)	JX967033	JX985360
A3	<i>A. Furches</i> 1114 (TENN)	JX967034	JX985360
B	<i>A. Furches</i> 1107 (TENN)	JX967035, JX967036	JX985361
Y1	<i>A. Furches</i> 1019 (TENN)	JX967037, JX967039	JX985362, JX985363
Y4	<i>A. Furches</i> 1119 (TENN)	JX967038	JX985364
Y5	<i>A. Furches</i> 1121 (TENN)	JX967039	JX985365
N1	<i>A. Furches</i> 1103 (TENN)	JX967040	JX985366
N2	<i>A. Furches</i> 1110 (TENN)	JX967041	JX985367
N3	<i>A. Furches</i> 1111 (TENN)	JX967042	JX985368
N4	<i>A. Furches</i> 1112 (TENN)	JX967043	N/A
N5	<i>A. Furches</i> 1115 (TENN)	JX967044	JX985366
D2	<i>A. Furches</i> 1016 (TENN)	JX967045	JX985369
D4	<i>A. Furches</i> 1124 (TENN)	JX967046	JX985370
D5	<i>M. Vincent</i> 14751 (TENN)	JX967047 - JX967050	N/A
H1	<i>E. Palmer</i> 17628 (NY)	JX967051	JX985372, JX985371, JX985378
H2	<i>A. Furches</i> 1022 (TENN)	JX967052 - JX967054	JX985375, JX985378
H3	<i>A. Furches</i> 1023 (TENN)	JX967055, JX967056	JX985376 - JX985379
H4	<i>A. Furches</i> 1024 (TENN)	JX967055	JX985379, JX985378
H5	<i>A. Furches</i> 1026 (TENN)	JX967056	JX985375, JX985378, JX985380
H6	<i>A. Furches</i> 1027 (TENN)	JX967057, JX967055	JX985381, JX985375, JX985378
H7	<i>A. Furches</i> 1028 (TENN)	JX967057 - JX967059	JX985376, JX985384, JX985383
H8	<i>A. Dattilo</i> 0910 (TENN)	JX967055	JX985378
H9	<i>A. Furches</i> 1100 (TENN)	JX967055, JX967058	JX985379, JX985386, JX985385, JX985378
H10	<i>A. Furches</i> 1101 (TENN)	JX967055	JX985378, JX985373
H11	<i>M. Vincent</i> 14754 (TENN)	JX967060	JX985376, JX985374
L1	<i>A. Furches</i> 1011 (TENN)	JX967061 - JX967063	JX985387 - JX985390
L3	<i>A. Furches</i> 1123 (TENN)	JX967064	N/A
O		JX967070	JX985391
P1	<i>A. Furches</i> 1102 (TENN)	JX967066	JX985392, JX985393
P2	<i>A. Furches</i> 1116 (TENN)	JX967067	JX985392
T1	<i>D. Estes</i> 00093 (TENN)	JX967068 - JX967071	JX985397
T2	<i>A. Furches</i> 1014 (TENN)	JX967072, JX967073	JX985396 - JX985399
T3	<i>A. Furches</i> 1017 (TENN)	JX967074	JX985397, JX985400, JX985399, JX985401
T4	<i>A. Furches</i> 1018 (TENN)	JX967074 - JX967077	JX985397, JX985403, JX985400, JX985402
T5	<i>A. Furches</i> 1020 (TENN)	JX967078	JX985397, JX985403
T6	<i>A. Furches</i> 1021 (TENN)	JX967074, JX967079	JX985397, JX985400
T7	<i>A. Furches</i> 1113 (TENN)	JX967080	JX985397, JX985399
T8	<i>V. McNeilus</i> 98-186 (NY)	JX967081	JX985397
T9	<i>A. Furches</i> 1118 (TENN)	JX967082	JX985397, JX985400, JX985399, JX985405, JX985404
T10	<i>A. Furches</i> 1120 (TENN)	JX967083, JX967084	JX985400, JX985394
T11	<i>H. DeSelm</i> s.n. (TENN)	JX967085 - JX967087, JX967075	JX985397, JX985400, JX985402, JX985395
T12	<i>W. Barger</i> s.n. (TENN)	JX967088	JX985397

A = *P. australis* SmallB = *P. brevisepalus* PennellY = *P. calycosus* SmallN = *P. canescens* (Britt.) Britt.D = *P. digitalis* Nutt. ex SimsH = *P. hirsutus* (L.) Willd.L = *P. laevigatus* (L.) Ait.O = *P. oklahomensis* PennellP = *P. pallidus* SmallT = *P. tenuiflorus* Pennell

Table A2. Species level summary statistics estimated from *rps16-trnK* and *GBSSI* data.

	Species	<i>rps16-trnK</i>						<i>GBSSI</i>					
		N	S	S _p	π	h	h _p	N	S	S _p	π	A	A _p
IC	<i>P. calycosus</i>	4 (3)	6 (6)	0	0.0023 ±0.0019	3	3	4 (3)	14	0	0.0068 ±0.0041	7	4
	<i>P. digitalis</i>	22 (3)	7 (7)	0	0.0008 ±0.0007	5	5	17 (3)	14	3	0.0064 ±0.0036	15	13
	<i>P. laevigatus</i>	7 (2)	9 (9)	1 (0)	0.0023 ±0.0017	4	3	4 (1)	14	0	0.0075 ±0.0044	8	6
NC	<i>P. australis</i>	4 (3)	7 (5)	0	0.0032 ±0.0026	3	0	4 (3)	3	0	0.0013 ±0.0010	2	0
	<i>P. brevisepalus</i>	3 (1)	3 (2)	3 (0)	0.0015 ±0.0016	2	2	5 (1)	0	0	0.0000 ±0.0000	1	0
	<i>P. canescens</i>	5 (5)	14 (10)	2 (1)	0.0046 ±0.0032	5	3	4 (4)	17	0	0.0074 ±0.0044	3	0
	<i>P. hirsutus</i>	99 (11)	12 (10)	4 (2)	0.0018 ±0.0012	10	4	98 (11)	20	14	0.0020 ±0.0013	15	12
	<i>P. pallidus</i>	4 (2)	3 (2)	0	0.0012 ±0.0012	2	0	4 (2)	2	2	0.0009 ±0.0008	2	1
	<i>P. tenuiflorus</i>	101 (12)	15 (14)	1 (0)	0.0026 ±0.0016	21	14	100 (12)	16	9	0.0026 ±0.0015	14	11

IC: inflated corolla throat

NC: narrow corolla throat

N: number of individuals (number of populations)

S: number of polymorphic nucleotide sites (number of polymorphic indel sites)

S_p: number of private polymorphic nucleotide sites (private indel sites)

π: mean nucleotide diversity per haplotype or allele

h, A: number of haplotypes or alleles

h_p, A_p: number of private haplotypes or alleles

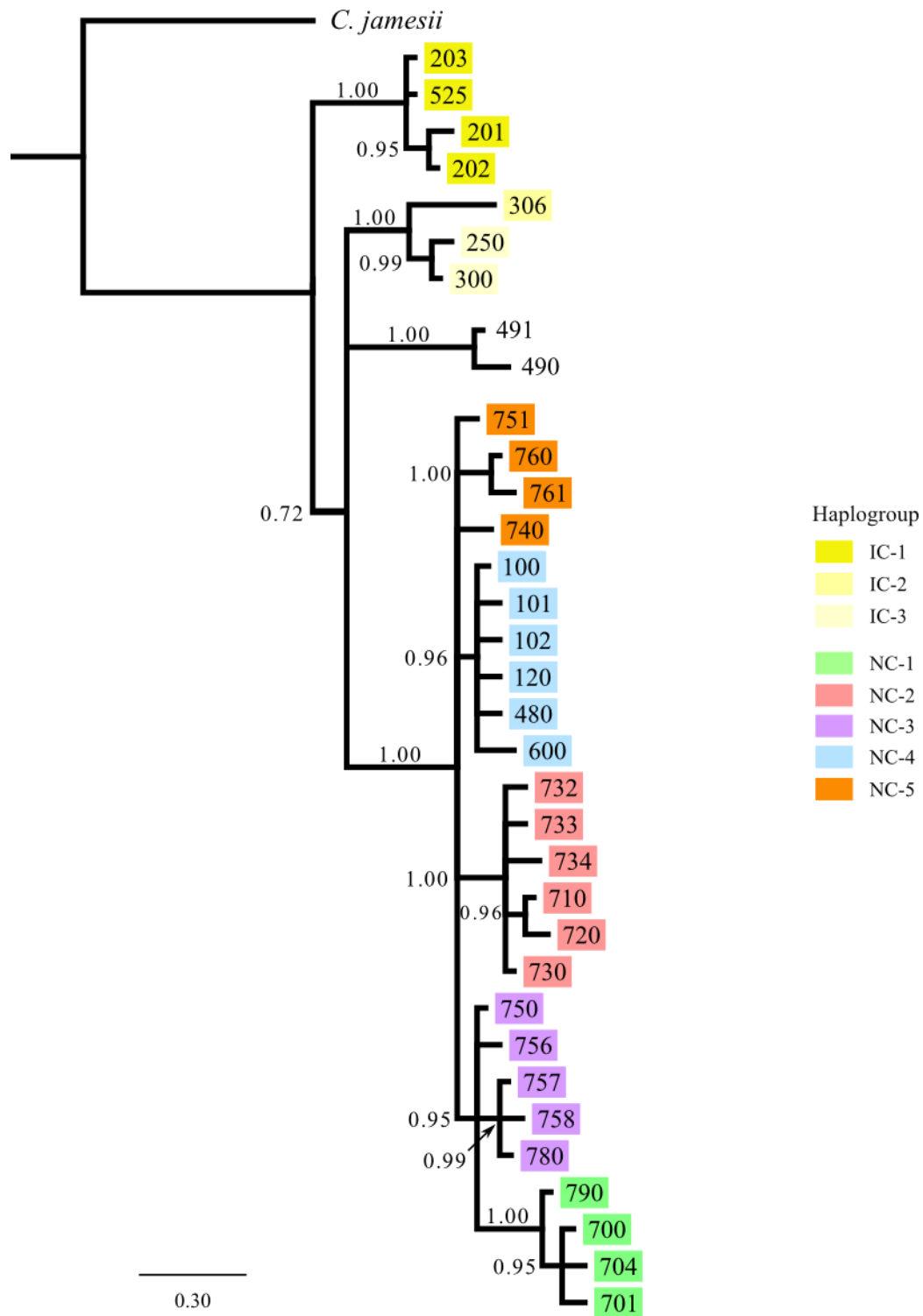


Figure A1. GBSSI phylogram excluding recombinant alleles. Tip labels represent alleles. Posterior probabilities are indicated on branches. Branch lengths are proportional to changes per site. The distribution of haplotypes and haplogroups among species and morphological groups is indicated in Figure 4.

Table A3. Estimated null allele frequencies per population and microsatellite locus.

ID	<i>Pen02</i>		<i>Pen04</i>		<i>Pen05</i>		<i>Pen21</i>		<i>Pen23</i>		Overall
H1	0.1534	(0.0926)	0.0661	(0.0557)	0.0844	(0.0695)	0.1299	(0.0937)	0.2975	(0.1281)	0.1463
H2	0.0687	(0.0582)	0.1267	(0.0816)	0.0727	(0.0620)	0.0773	(0.0643)	0.0509	(0.0464)	0.0793
H3	0.1557	(0.0918)	0.0994	(0.0749)	0.1781	(0.1016)	0.1213	(0.0999)	0.1959	(0.1329)	0.1501
H4	0.0606	(0.0536)	0.0553	(0.0497)	0.0852	(0.0701)	0.1166	(0.0859)	0.1879	(0.1113)	0.1011
H5	0.2004	(0.1011)	0.2039	(0.1391)	0.0709	(0.0592)	0.1839	(0.1258)	0.1453	(0.1075)	0.1609
H6	0.0838	(0.0687)	0.1365	(0.0906)	0.1781	(0.1084)	0.1689	(0.1246)	0.3923	(0.1640)	0.1919
H7	0.1341	(0.0969)	0.1214	(0.0859)	0.1328	(0.0948)	0.2674	(0.1478)	0.2503	(0.1245)	0.1812
H8	0.2691	(0.1862)	0.3017	(0.1933)	0.3018	(0.1957)	0.1982	(0.1486)	0.3806	(0.1948)	0.2903
H9	0.1115	(0.0950)	0.1258	(0.0942)	0.0390	(0.0364)	0.2060	(0.1610)	0.5528	(0.2750)	0.2070
H10	0.1627	(0.1222)	0.2085	(0.1264)	0.1735	(0.1251)	0.2285	(0.1518)	0.4152	(0.2396)	0.2377
H11	0.3230	(0.2005)	0.3364	(0.2297)	0.3305	(0.2259)	0.3341	(0.2280)	-	-	0.3310
<i>Mean</i>	0.1566		0.1620		0.1497		0.1847		0.2869		0.1888
T1	0.0580	(0.0512)	0.0861	(0.0676)	0.0681	(0.0531)	0.2044	(0.1197)	0.0722	(0.0587)	0.0978
T2	0.0494	(0.0446)	0.0498	(0.0463)	0.0690	(0.0577)	0.0994	(0.0756)	0.2138	(0.0987)	0.0963
T3	0.0501	(0.0459)	0.0794	(0.0659)	0.0764	(0.0637)	0.0632	(0.0550)	0.1455	(0.0901)	0.0829
T4	0.0894	(0.0687)	0.1214	(0.0818)	0.1126	(0.0793)	0.1666	(0.1022)	0.4881	(0.2295)	0.1956
T5	0.0835	(0.0672)	0.1547	(0.0907)	0.0921	(0.0760)	0.1900	(0.1375)	0.3304	(0.2261)	0.1701
T6	0.1610	(0.0982)	0.0921	(0.0697)	0.2049	(0.1098)	0.3171	(0.1999)	-	-	0.1938
T7	0.3396	(0.1706)	0.2144	(0.1284)	0.2033	(0.1606)	0.5544	(0.2757)	0.3151	(0.2034)	0.3253
T9	0.4120	(0.2423)	0.4473	(0.2065)	0.4143	(0.2453)	0.5470	(0.2754)	0.3300	(0.2251)	0.4301
T10	0.3047	(0.1726)	0.3299	(0.1840)	0.3204	(0.2007)	0.3134	(0.2015)	0.5519	(0.2742)	0.3641
<i>Mean</i>	0.1704		0.1737		0.1711		0.2640		0.3038		0.2145

(Standard error)

Table A4. *rps16-trnK* population-level summary statistics for *Penstemon hirsutus* and *P. tenuiflorus*.

Pop.	N	S	h	h_p	H_s	k_s
H1	10	0	1	1	0.000	0
H2	10	3	3	1	0.644	0.64 ±0.54
H3	10	1	2	0	0.200	0.20 ±0.27
H4	10	0	1	0	0.000	0
H5	11	0	1	0	0.000	0
H6	11	1	2	0	0.327	0.33 ±0.36
H7	8	7	3	1	0.607	1.11 ±0.80
H8	5	0	1	0	0.000	0
H9	16	2	2	0	0.525	0.53 ±0.46
H10	6	0	1	0	0.000	0
H11	2	0	1	1	0.000	0
Mean	9.00	1.27	1.64	0.36	0.209	0.25
T1	13	4	4	2	0.654	0.77 ±0.60
T2	10	2	2	2	0.533	1.07 ±0.77
T3	10	0	1	0	0.000	0
T4	10	5	4	1	0.533	0.93 ±0.70
T5	10	0	1	1	0.000	0
T6	10	3	2	1	0.200	0.60 ±0.52
T7	13	0	1	1	0.000	0
T8	1	0	1	0	0.000	0
T9	13	0	1	0	0.000	0
T10	4	2	2	2	0.500	0.50 ±0.52
T11	6	8	4	2	0.867	1.73 ±1.17
T12	1	0	1	0	0.000	0
Mean	8.42	2.00	2.00	1.00	0.274	0.47

N: number of individuals sampled

S: number of polymorphic nucleotide sites

S_p : number of private polymorphic sites

h: number of unique haplotypes

h_p : number of private haplotypes

H_s : gene diversity

k_s : mean number pairwise nucleotide differences between haplotypes

Table A5. GBSSI population-level summary statistics for *Penstemon hirsutus* and *P. tenuiflorus*.

Pop.	N	S	S _p	k _s	A	A _p	H _O	H _S	F _{IS}	Φ _{IS}
H1	10	6	1	1.19 ±0.80	3	1	0.400	0.468	0.153	0.167
H2	10	4	0	0.40 ±0.39	2	0	0.100	0.100	0.000	0.000
H3	10	5	0	2.12 ±1.23	3	0	0.600	0.542	-0.113	-0.191
H4	10	8	4	1.68 ±1.03	4	1	0.400	0.432	0.077	-0.140
H5	11	5	0	0.46 ±0.42	3	0	0.182	0.178	-0.026	0.000
H6	10	6	1	1.45 ±0.92	3	1	0.600	0.611	0.018	0.600
H7	8	2	2	0.47 ±0.43	3	2	0.250	0.433	0.440	0.482
H8	5	0	0	0.00	1	0	0.000	0.000	-9.999	-9.999
H9	16	5	1	2.32 ±1.30	3	1	0.500	0.548	0.091	0.032
H10	6	1	1	0.41 ±0.40	2	1	0.500	0.409	-0.250	-0.250
H11	2	4	3	2.50 ±1.69	3	2	0.500	0.833	0.500	0.857
Mean	8.91	4.18	1.18	1.18 ±0.89	2.73	0.82	0.367	0.414		
T1	13	0	0	0.00	1	0	0.000	0.000	-9.999	-9.999
T2	10	8	1	3.74 ±1.97	3	1	0.800	0.653	-0.241	-0.186
T3	10	8	0	2.35 ±1.34	5	1	0.400	0.442	0.100	0.155
T4	10	12	1	3.56 ±1.90	6	1	0.556	0.739	0.259	0.215
T5	10	1	0	0.10 ±0.18	2	0	0.100	0.100	0.000	0.000
T6	10	6	0	0.10 ±0.50	2	0	0.100	0.100	0.000	0.000
T7	13	7	0	0.54 ±0.46	2	0	0.077	0.077	0.000	0.000
T8	1	0	0	0.00	1	0	0.000	0.000	-9.999	-9.999
T9	13	7	1	3.08 ±1.65	7	2	0.692	0.803	0.143	0.099
T10	4	6	0	1.50 ±1.01	3	0	0.250	0.464	0.500	0.000
T11	6	10	0	5.00 ±2.61	4	0	0.333	0.697	0.545	0.631*
T12	1	0	0	0.00	1	0	0.000	0.000	-9.999	-9.999
Mean	8.42	5.42	0.25	1.66 ±1.81	3.08	0.42	0.276	0.340		

N: number of sampled individuals per population

S: number of polymorphic nucleotide sites

S_p: number of private (single population within one species) polymorphic sitesk_s: mean number of nucleotide differences between pairs of haplotypes ±standard deviation

A: number of alleles

A_p: number of private (single population within one species) allelesH_O: observed heterozygosityH_S: expected heterozygosity

*P ≤ 0.05

Table A6. Microsatellite summary statistics per locus for *Penstemon hirsutus* (a) and *P. tenuiflorus* (b).**(a) *P. hirsutus***

	<i>Pen02</i>	<i>Pen04</i>	<i>Pen05</i>	<i>Pen21</i>	<i>Pen23</i>
size (bp)	166-246	168-268	157-237	116-164	154-192
N	83	78	91	61	66
missing	0.16	0.21	0.08	0.38	0.33
A	30	36	30	18	17
A _p	5	11	6	5	2
R	22.82	27.05	22.60	14.90	14.58
R _p	9.03	11.25	7.91	8.05	3.34
H _O	0.6718	0.6803	0.6648	0.6697	0.2495
H _T	0.8162	0.7874	0.7445	0.7502	0.5117
null	0.0368	0.0371	0.0452	0.0403	0.2071

(b) *P. tenuiflorus*

	<i>Pen02</i>	<i>Pen04</i>	<i>Pen05</i>	<i>Pen21</i>	<i>Pen23</i>
size (bp)	156-250	152-268	151-203	122-166	150-216
N	68	73	69	51	36
missing	0.33	0.28	0.32	0.50	0.64
A	31	32	24	11	21
A _p	6	7	4	2	4
R	25.04	25.71	19.55	10.09	21.00
R _p	11.25	9.91	4.86	3.25	9.76
H _O	0.6250	0.5527	0.6385	0.4153	0.5708
H _T	0.7861	0.8417	0.7224	0.6266	0.9182
null	0.0398	0.0743	0.0472	0.1165	0.1139

N = mean number of sampled individuals per locus

missing = mean proportion missing data per locus

k_T = mean number of allelic differences between all pairs of haplotypes

A = mean number of alleles per locus

A_p = mean number of private alleles per locus

R = mean rarefied allelic richness per locus

R_p = mean rarefied private allelic richness per locusH_O = mean observed heterozygosity per locusH_T = mean total expected heterozygosity per locus

null = mean estimated frequency of null alleles per locus

Table A7. Microsatellite population-level summary statistics for *Penstemon hirsutus* and *P. tenuiflorus*.

ID	N	miss.	null	A	A _p	H _O	H _S	F _{IS}	K _S
H1	10.00	0.00	0.146	5.80	1.25	0.560	0.683	0.181*	3.38
H2	10.00	0.00	0.079	6.60	1.00	0.840	0.826	-0.018	4.13
H3	8.80	0.12	0.150	5.80	1.00	0.531	0.662	0.137	2.53
H4	9.80	0.02	0.101	5.00	0.00	0.700	0.655	-0.095	3.12
H5	8.00	0.27	0.161	6.00	1.00	0.555	0.751	0.076	2.26
H6	8.40	0.24	0.192	6.60	2.33	0.472	0.809	0.267*	2.54
H7	7.20	0.10	0.181	7.20	1.33	0.566	0.869	0.259*	3.48
H8	3.60	0.28	0.290	1.80	1.00	0.200	0.283	-0.043	0.58
H9	5.20	0.68	0.207	6.75	0.00	0.958	0.866	-1.000	0.77
H10	3.80	0.37	0.238	5.00	1.33	0.520	0.733	-0.058	1.74
H11	1.00	0.50	0.331	3.00	1.00	0.500	1.000	-1.000	1.33
Mean	6.89	0.23	0.189	5.41	1.02	0.582	0.740		2.35
T1	11.00	0.15	0.098	8.00	1.33	0.711	0.737	-0.133	2.66
T2	10.00	0.00	0.096	8.60	1.00	0.820	0.853	0.039	4.26
T3	9.80	0.02	0.083	7.20	1.00	0.833	0.813	-0.053	3.91
T4	8.20	0.18	0.196	7.00	0.00	0.551	0.895	0.190*	3.30
T5	7.00	0.30	0.170	5.50	0.00	0.617	0.637	-0.107	2.09
T6	6.00	0.40	0.194	7.50	2.33	0.581	0.908	0.214*	2.26
T7	3.00	0.77	0.325	4.25	0.00	0.563	0.913	-0.318	0.60
T9	1.80	0.86	0.430	1.67	2.00	0.000	0.333	0.667	0.87
T10	2.40	0.40	0.364	3.50	2.00	0.396	0.958	0.400	1.57
T11	0.20	0.97	N/A	0.20	1.00	0.000	0.000	-9.999	0.00
Mean	5.94	0.41	0.215	5.34	1.07	0.507	0.740		2.15

N = mean number of sampled individuals per locus

miss. = mean proportion missing data per locus

k_S = mean number of allelic differences between pairs of haplotypes

A = mean number of alleles per locus

A_p = mean number of private alleles per locus

prop. (A_p) = proportion total alleles that are private alleles

H_O = mean observed heterozygosity per locus

H_S = mean total expected heterozygosity per locus

null = mean estimated frequency of null alleles per locus

*P ≤ 0.05

Appendix III

GBSSI primer sequences

X3F 5' – ACTGTYCGRTTCTTCCAC – 3'

X8R 5' – TCACCRGAWACAAGCTCCTC – 3'

Vita

Anna earned a Bachelor of Science degree from the University of Tennessee in 2007.