RESEARCH ARTICLE

AMERICAN JOURNAL OF BOTANY

INVITED PAPER

*For the Special Issue: Evolutionary Insights from Studies of Geographic Variation*

# Mating system divergence affects the distribution of sequence diversity within and among populations of recently diverged subspecies of *Clarkia xantiana* (Onagraceae)[1]

James B. Pettengill, Ryan D. Briscoe Runquist, and David A. Moeller[2]

**PREMISE OF THE STUDY:** The population biology of outcrossing and self-fertilizing taxa is thought to differ because of the advantage that selfers have in colonizing unoccupied sites where mates and pollinators may be limiting (Baker's Law). This reduced tendency for outcrossers to colonize new sites, along with their greater dependence on pollinators to disperse pollen, has the potential to differently influence the genetic diversity and structure of outcrossing and selfing populations.

**METHODS:** We conducted a comparative population genetic study of two sister outcrossing and selfing subspecies of *Clarkia xantiana* that have very recently diverged. We used DNA sequence variation (>40 kb from eight nuclear loci) from large samples of individuals from 14 populations to assess geographic patterns of genetic diversity and make inferences about the demographic and colonization histories of each subspecies.

**KEY RESULTS:** We show that sequence variation is strongly reduced across all selfing populations. The demographic history of selfing populations exhibits recent colonization bottlenecks, whereas such bottlenecks are rarely observed for the outcrosser. The greater effect of genetic drift in the selfer has resulted in strong population genetic structure, but with no pattern of isolation by distance. By contrast, the stronger effect of gene flow in the outcrosser has resulted in considerably less structure, but a significant pattern of isolation by distance.

**CONCLUSIONS:** Taken together, our results suggest that selfing populations are not at migration–drift equilibrium, are affected by strong episodes of genetic drift during colonization, and experience little or no subsequent gene flow from other populations after those founder events.

**KEY WORDS** Baker's law; breeding system; demographic history; dispersal; colonization bottleneck; isolation by distance; floral evolution; mating system; self-fertilization; single nucleotide polymorphism

Mating system is one of the strongest predictors of population genetic diversity and structure in plants (Brown, 1979; Hamrick and Godt, 1990, 1996; Nybom, 2004; Glemin et al., 2006). The consequences of the evolutionary transition from outcrossing to selfing for genome-wide patterns of variation have been of particular interest as this shift is among the most common evolutionary changes in plant lineages (Stebbins, 1974; Holsinger, 2000; Barrett, 2002) and is often associated with speciation (Stebbins, 1957; Baker, 1959). The evolution of self-fertilization is expected to have diverse consequences for genetic variation and population structure (Charlesworth and Wright, 2001; Glemin et al., 2006; Wright et al., 2013). Selfing species are often derived from outcrossing progenitors, and transitions to self-fertilization are generally thought to be recent due to the low adaptive potential and high extinction rates of selfers (Dobzhansky, 1950; Igic et al., 2008; Goldberg et al., 2010; Igic and Busch, 2013). When divergence occurs with a strong bottleneck, drift can result in low effective population size ($N_e$) across the entire selfing taxon, and substantial periods of time are required for new mutations to accumulate. A high rate of self-fertilization *per se* is also expected to result in a loss of approximately half of the $N_e$ (Nordborg, 2000). Because selfing increases homozygosity, effective rates of recombination are reduced (Nordborg, 2000). This greater linkage disequilibrium can enhance hitchhiking effects due to positive or background selection and cause further losses of

---

[1] Manuscript received 2 April 2015; revision accepted 14 July 2015.

Department of Plant Biology, 1445 Gortner Avenue, University of Minnesota, Saint Paul, Minnesota 55108 USA

[2] Author for correspondence (e-mail: moeller@umn.edu), fax: (612) 625-1738

**doi:10.3732/ajb.1500147**

$N_e$ (Charlesworth et al., 1993). Altogether, the evolution of selfing is expected to cause significant losses of variation, especially from within populations (Charlesworth and Pannell, 2001; Glemin et al., 2006).

The ecology of selfing and outcrossing species can also have important consequences for patterns of population genetic variation. Selfers can readily establish new colonies and persist even when both mates and pollinators are unavailable or limited (Baker's Law: Baker, 1955, 1967; Moeller and Geber, 2005). The potential for gene flow through seed dispersal may differ dramatically among selfing taxa. Selfing taxa have been observed to exist in small, stable but highly fragmented (peripheral) populations in which long-distance dispersal events are thought to be rare (Vasek, 1968; Lewis, 1973; Vasek and Harding, 1976; Solbrig and Rollins, 1977; Koelling et al., 2011). Consistent with this observation, meta-analyses of allozyme data by Hamrick and Godt (1990, 1996) found that population structure was comparatively higher in selfers and that mating system was among the strongest predictors of the distribution of genetic variation within vs. among populations. By contrast, it has also been suggested that there is an association between selfing and high seed dispersal potential in weedy species (Baker, 1955, 1965, 1974; Stebbins, 1957; Jain, 1976; Lloyd, 1980; Price and Jain, 1981; Cheptou, 2012) where metapopulation dynamics (frequent extinction and colonization) homogenize variation among populations (Pannell and Barrett, 1998; Barrett and Pannell, 1999; but see Cheptou and Massol, 2009). In both scenarios, the dynamics of repeated colonization and the effects of genetic drift during colonization bottlenecks can strongly reduce within-population variation (Charlesworth and Pannell, 2001). In outcrossers, persistence following colonization events can be impeded by Allee effects (positive population density- or population size-dependent mating success: Allee, 1951; Dennis, 1989; Kunin and Iwasa, 1996; Hackney and McGraw, 2001; Moeller, 2004), which may lead to population extinction (Groom, 1998; Courchamp et al., 1999). While selfers can found new populations and persist with small population sizes even in the absence of dispersal, outcrossers tend to persist in large populations (Stebbins, 1950, 1957; Lloyd, 1980). By virtue of their larger populations, the outcrossers may also have a greater potential to export seeds potentially increasing gene flow among populations.

In addition to dispersal via seeds, pollinators of outcrossing plants disperse pollen as they forage, in some cases long distances (Broyles et al., 1994; Schulke and Waser, 2001; Jha and Dick, 2010), whereas selfers lack of reliance on pollinators for reproduction can limit the influx or export of alleles from populations due to pollen flow. For selfers, colonization of sites via few individuals coupled with limited pollen flow after establishment may act to elevate levels of differentiation among populations. Even though variation is low within populations, the pronounced differentiation among selfing populations may result in high levels of diversity species-wide, potentially comparable with sister outcrossing taxa (Charlesworth and Pannell, 2001; Glemin et al., 2006). Therefore, understanding the consequences of mating system divergence for population genetic structure requires intensively sampling individuals at local and geographic scales (Charlesworth and Pannell, 2001). In outcrossers, by contrast, populations may have greater connectivity due to both the frequent export of seeds and pollen. Overall, drift may be a more important force in determining the population structure of selfers, whereas patterns of gene flow may more strongly influence outcrossers' population structure.

The relative contribution of genetic drift and gene flow to population structure can be evaluated by examining the relationship between genetic distance and geographic distance. Under a stepping stone model of population structure (Kimura, 1953; Kimura and Weiss, 1964), genetic distance should increase monotonically with geographic distance across a region at equilibrium (Malécot, 1955). This isolation by distance relationship occurs because gene flow homogenizes populations in close proximity to one another but is weak at large distances due to limited dispersal. When this relationship is not observed (i.e., is flat), one can infer that populations are either panmictic across the observed region or not at migration–drift equilibrium and that historical factors have influenced spatial patterns of variation. Historical effects on population structure are common as geographic ranges often shift or are fragmented during periods of climate change. Nonetheless, the relative contribution of gene flow and drift can still be inferred when isolation by distance is rejected (Hutchison and Templeton, 1999). If non-equilibrium populations have high rates of gene flow, the flat relationship between genetic and geographic distance will have low variance because of the consistent homogenizing effects of gene flow. By contrast, if there is strong fragmentation and gene flow is very low, the flat relationship will have high variance because of the stochastic effects of drift in independently evolving populations (Hutchison and Templeton, 1999). Population structure in an expanding range margin will therefore be determined by the recency of the invasion and the magnitude of gene flow following invasion.

The population structure of outcrossers and selfers may differ especially because of the extent to which populations of each taxon are at migration–drift equilibrium. Selfers are nearly always derived from outcrossing progenitors, sometimes very recently (Foxe et al., 2009; Pettengill and Moeller, 2012a), and their geographic distributions may often reflect recent range expansion. The different extremes of spatial population dynamics possible in selfing taxa may lead to very different patterns of isolation by distance. When selfers are ruderal and disperse frequently across landscapes ("metapopulation dynamics"), population differentiation can be low even in the absence of pollen dispersal by pollinators (Pannell and Charlesworth, 2000). In addition, the relationship between genetic and geographic distance would have low variance because of the high rates of seed dispersal across the landscape. Whereas when selfers have fragmented distributions and infrequent seed dispersal (Cheptou and Massol, 2009), population genetic structure should be high, and the relationship between genetic and geographic distance would have high variance owing to the stochastic effects of drift on patterns of genetic variation among populations. Here, both low seed dispersal and pollen dispersal jointly contribute to population subdivision.

We examined the effects of the evolution of self-fertilization on patterns of within-population sequence variation and among-population structure using two incipient *Clarkia* species that differ in mating system. One goal of this paper is to explicitly consider sequence variation across large fractions of two focal species' ranges while also intensively sampling within populations. The population genetic effects of a transition in mating system are most apparent on within-population diversity, necessitating both high within-population and among-population sampling to understand the influence of mating system on sequence variation (Charlesworth and Pannell, 2001). *Clarkia xantiana* subsp. *xantiana* A.Gray (Onagraceae) is a self-compatible but predominantly outcrossing taxon (Moeller et al., 2011) endemic to the southern Sierra Nevada of

California, United States (Eckhart and Geber, 1999; Runions and Geber, 2000). Inbreeding can occur in subspecies *xantiana* via geitonogamous selfing or mating between close relatives, but we have detected little propensity for autonomous selfing in this taxon (Moeller, 2006; Moeller et al., 2011). Its sister taxon, *C. xantiana* subsp. *parviflora* (Eastw.) Harlan Lewis & P.H.Raven, is primarily selfing (Moeller, 2006; Briscoe Runquist and Moeller, 2013, 2014) and very recently derived from subspecies *xantiana* (ca. 65 000 yr bp; Pettengill and Moeller, 2012a). Subspecies *parviflora* readily autonomously self-pollinates in the field and in pollinator-free environments and selfing provides considerable reproductive assurance (Moeller, 2006). Selfing occurs concurrently with the period when outcrossing can occur (competing selfing) rather than selfing occurring as a delayed "fail safe" mechanism (Runions and Geber, 2000). Populations also vary geographically in autofertility, with higher rates of autofertility in western populations closest to the range of subspecies *xantiana* (Briscoe Runquist and Moeller, 2014). Therefore, it was also of interest to ask whether geographic variation in mating system traits within subspecies *parviflora* was associated with variation in sequence diversity.

In this study, we build on previous work that primarily examined the divergence history of the two incipient species, but where within-population sampling was not substantial enough to examine important questions about the effects of mating system evolution on sequence variation and the scale of population structure. First, we tested the simple prediction that effective population size is reduced and linkage disequilibrium elevated in populations of the recently derived selfing taxon relative to populations of the progenitor outcrossing taxon. We were particularly interested in examining variation in sequence diversity between subspecies as well as among populations within subspecies. Second, we tested the prediction that population structure is stronger in the selfer due to more pronounced random genetic drift (e.g., during colonization bottlenecks). We also examined whether each taxon exhibited a pattern of isolation by distance, which is expected if populations are at migration–drift equilibrium. Finally, we examined the demographic history of individual populations using coalescent models to test the prediction that individual selfing populations have more frequently been affected by recent population bottlenecks than outcrossing populations. This prediction directly follows from Baker's Law and the hypothesized reproductive assurance value of self-fertilization during colonization. We discuss our results in a geographic context because our populations were sampled across a west-to-east gradient spanning a large part of both taxon's ranges.

## MATERIALS AND METHODS

***Study system***—*Clarkia xantiana* subspecies have narrow geographic ranges that occupy the southern extent of the Sierra Nevada Range as well as parts of the Transverse Ranges (Fig. 1; Eckhart and Geber, 1999). The subspecies are parapatric with *xantiana* occupying foothills on the western slope of the Sierras and with *parviflora* occupying arid foothills to the east (hereafter "*xantiana*" and "*parviflora*"). This region is south of the high peaks of the Sierra Nevada, and therefore, no physical barrier prevents dispersal between the ranges of the two taxa. Subspecies *xantiana* occurs as far west as the San Joaquin Valley and *parviflora* as far east as the Mojave Desert (Fig. 1). Both taxa range as far south as Los Angeles County, but populations in the southern portion of the distribution are uncommon; therefore, our study focuses on the northern section of the range where most populations occur.

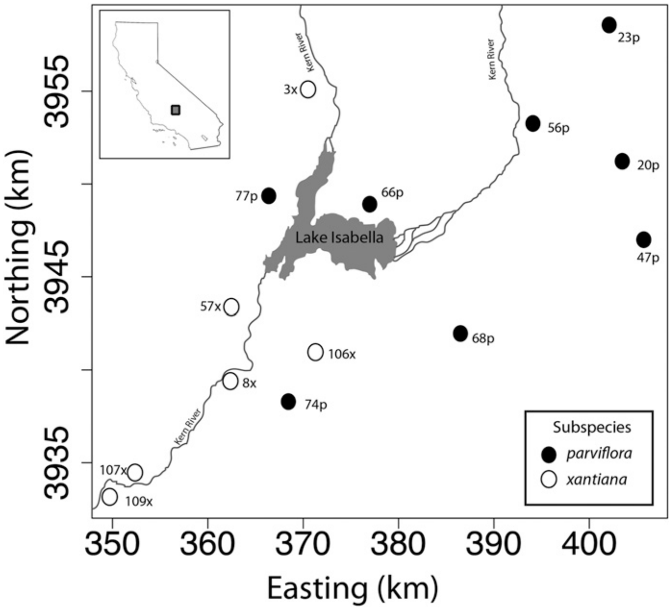

**FIGURE 1** Map showing the distribution of six *Clarkia xantiana* subsp. *xantiana* and eight subsp. *parviflora* populations sampled in this study and their numerical identifiers. The region of study is centered on Isabella Lake, Kern County, California, USA.

Subspecies *parviflora* comprises a monophyletic group derived from *xantiana* (Pettengill and Moeller, 2012a). Statistical phylogeographic analyses and paleoclimate modeling have indicated that the two taxa are in secondary sympatry, having come into contact due to range expansion following partial divergence in allopatry (Pettengill and Moeller, 2012b). *Xantiana* appears to have migrated north and to higher elevations following the last glacial maximum; *parviflora* migrated primarily west and up in elevation (Pettengill and Moeller, 2012b).

Both subspecies are annuals and have at least a short-lived seed bank (Eckhart et al., 2011). Unlike many annuals, neither subspecies is associated with human disturbance; both typically occur on steep foothill slopes where natural disturbance occurs and competition from other plants is reduced. The census size of *xantiana* populations tends to be large, numbering typically from thousands to several hundred thousand individuals (Eckhart et al., 2011; Moeller et al., 2011). *Parviflora* populations are typically smaller (D. A. Moeller, unpublished data) and can persist in sites where only tens of individuals appear in most years (Eckhart and Geber, 1999; Moeller and Geber, 2005). The density of individuals within populations also tends to be considerably lower in *parviflora* than *xantiana* populations (Briscoe Runquist et al., 2014; D. A. Moeller, unpublished data). Both subspecies occur typically in discrete colonies, but the spatial separation of populations is often quite different; *parviflora* populations tend to be more isolated across the landscape with greater distances separating known localities (D. A. Moeller, personal observation). The average precipitation is higher and the variance in precipitation lower in *xantiana*'s portion of the range (Eckhart et al., 2010, 2011). Owing to unpredictable precipitation, both subspecies can experience dramatic year-to-year variation in population size and reproductive output (Eckhart et al., 2011).

However, this also tends to be more pronounced in *parviflora*, where it is not uncommon for individual populations to yield zero reproducing individuals in a given year (D. A. Moeller, personal observation).

***Taxon sampling and DNA sequencing—***We sampled 136 individuals from eight *parviflora* populations and 130 individuals from six *xantiana* populations (Appendix S1, see Supplemental Data with the online version of this article). Sequence variation for *xantiana* has been reported previously in Moeller et al. (2011). Each individual represented a different maternal family. DNA was extracted using Qiagen DNeasy plant mini kits (Valencia, California, USA). We used PCR to amplify eight single-copy nuclear genomic regions (nDNA) from the 266 individuals (40 263 bp with an average length of 369 bp per locus). The PCR primers used to amplify nDNA loci were designed from EST sequences isolated from *Clarkia breweri* flower buds (see Moeller et al., 2011). PCR products were sequenced directly using the ABI BigDye v3.1 chemistry and an ABI 3730xl DNA sequencer. We inspected all chromatograms using the program Sequencher v4.8 (Gene Codes Corp., Ann Arbor, Michigan, USA), aligned sequences using the program Muscle (Edgar, 2004) based on the default settings, and determined haplotype phase using the program Phase v2.1 (Stephens and Donnelly, 2003; Stephens et al., 2001). Additional details on haplotype phasing can be found in Pettengill and Moeller (2012a).

***Inbreeding, effective population size, and linkage disequilibrium—***We estimated all statistics describing sequence variation for each individual population. First, we used the program GDA (Lewis and Zaykin, 2001) to obtain estimates of the inbreeding coefficient, *F*, for each population. To examine population variation in effective population size, we estimated two measures of $\theta = 4N_e\mu$: the number of segregating sites per sequenced site, $\theta_w$ (Watterson, 1975), and the pairwise differences in nucleotides between sequences, $\theta_\pi$ (Nei, 1987). We calculated Tajima's *D* (Tajima, 1989) for each population as a description of the site frequency spectrum and as a potential indicator of population size changes. We also calculated Fu's $F_s$ (Fu, 1997), a metric that describes the haplotype frequency spectrum and is sensitive to changes in population size (Ramos-Onsins and Rozas, 2002). We estimated linkage disequilibrium using the *ZnS* statistic, which describes the average squared correlation coefficient over all pairwise comparisons of polymorphic sites (Kelly, 1997). All of the descriptors of sequence diversity were calculated using the program Sites (Hey and Wakeley, 1997).

We tested whether population-specific values of each summary statistic differ between subspecies. All analyses were conducted in program R version 3.1 (R Development Core Team, 2011). We verified that the assumptions of normality of residuals and homogeneity of variance were satisfied for all analyses. The models included subspecies and locus crossed with subspecies as fixed-effect predictor variables and population as a random effect using the lmer function in the lme4 package in R (Bates et al., 2015). Significance of fixed effects was determined using the mixed function from the afex package (Singmann et al., 2015). The mixed function is an implementation of the lmer function that calculates type III-like *p* values for mixed models using the Kenward–Rogers approximation for degrees of freedom (Judd et al., 2012). For *ZnS*, we dropped the interaction of subspecies and locus from the model because it was highly collinear with subspecies. We also tested whether the variance in Tajima's *D* differed between subspecies because a high variance among loci within populations can indicate a history of population bottlenecks (Charlesworth et al., 2003; Wright and Gaut, 2005). To test for differences in variance between subspecies, we calculated the unbiased estimate of the coefficient of variation (CV) for each population using the locus-specific estimates for that population. These estimates were then used in a one-way ANOVA with CV as the dependent variable and subspecies as the independent predictor variable.

Because mating system varies among populations of *parviflora*, with eastern populations having lower autofertility and a higher probability of outcrossing (Briscoe Runquist and Moeller, 2014), we also quantified the relationship between geography and population genetic parameters. Geographic variation in population genetic parameters for *xantiana* have been previously reported (Moeller et al., 2011). To assess the linear relationship between $\theta_w$, $\theta_\pi$, *D*, *ZnS*, and geography, we conducted analyses that included the population UTM easting coordinate (a continuous variable that represents the predominant west-to-east geographic gradient) and locus. We also assessed nonlinear relationships by running a model with the variable already mentioned and the centered and squared value of the easting coordinate as an additional predictor variable; no quadratic variables were significant so we only report linear models. Population 66p was excluded from the geographic analysis of *D* and $F_S$ because we were only able to estimate a value for one locus due to a lack of variation in the others; population 74p was also excluded for the analysis of *ZnS* for the same reason.

***Population genetic structure—***We used two methods to determine the relative roles of gene flow and genetic drift in shaping patterns of genetic differentiation among populations within each taxon. First, we conducted an isolation by distance analysis by examining the relationship between $F_{ST}$ and pairwise geographic distances with Mantel tests implemented in the ecodist package of R (R Development Core Team, 2011). We conducted separate analyses for each subspecies. To determine whether the patterns differed between subspecies, we constructed 95% confidence intervals estimates for each Mantel statistic using 1000 bootstrap replicates and determined whether they were non-overlapping. A significant pattern of isolation by distance is consistent with migration–drift equilibrium. We also calculated the $R^2$ value from an ordinary least square analysis to characterize the overall variance in the relationship between $F_{ST}$ and geographic distance.

Because $F_{ST}$ is a relativized measure of genetic differentiation, its magnitude can be affected both by the degree of differentiation between populations as well as differences in diversity (Pannell and Charlesworth, 2000; Charlesworth et al., 2003). This effect can be problematic especially when comparing levels of differentiation between outcrossers and selfers where levels of within-population diversity often differ (Charlesworth and Pannell, 2001). Therefore, we also calculated an absolute metric of divergence ($\pi_T - \pi_S$) between pairs of populations within each taxon; this metric is unaffected by differences between populations in diversity. Here, divergence is measured by the difference between $\pi_T$ (pairwise sequence differences across both populations) and $\pi_S$ (pairwise sequence differences within the focal population). We tested whether absolute divergence ($\pi_T - \pi_S$) differed between subspecies using a model that included subspecies, population, and locus crossed with subspecies as fixed effects. Because divergence values are not necessarily independent, we determined the significance of predictor variables using a permutation test where absolute divergence values were randomized with respect to population pairs. Subspecies were deemed significantly different if the *F* value fell into the upper or lower 2.5% of the distribution of *F* values from 999 models run with permuted data.

Second, we used the program InStruct (Gao et al., 2007) to estimate the degree of population structure within each taxon. This Bayesian model-based clustering program does not incorporate any *a priori* information regarding the sampling location of individuals. Rather, the method assigns individuals to clusters based on optimizing the degree of Hardy–Weinberg equilibrium and linkage disequilibrium while also allowing for a degree of inbreeding within clusters, which makes InStruct well suited for partially selfing plant taxa. The optimal value of *K* (i.e., the number of clusters or subpopulations within each subspecies) was inferred using the deviance information criterion as implemented in InStruct. We visualized the results using the program distruct (Rosenberg, 2004).

***Historical demography***—We constructed Skyline plots (Drummond et al., 2005) to detect historical changes in population size using the program BEAST (Drummond and Rambaut, 2007). These analyses determine the effective population size, $N_e$, at different time intervals along the gene genealogies and, hence, can be used to assess whether there have been fluctuations in population size over time. Because the analyses are relatively parameter rich, we chose the HKY model of nucleotide substitution and linked trees as is required under the Skyline approach. We ran each analysis under a strict clock for $10^7$ generations with the first $10^6$ generations treated as burn-in. We monitored effective sample size (ESS) values to ensure sufficient mixing and convergence. We also evaluated variants of the Skyline approach (e.g., Extended Bayesian Skyline; Heled and Drummond, 2008), which are better able to detect fluctuations in historical population size by treating loci independently, but we could not achieve acceptable mixing/convergence under that model.

## RESULTS

***Inbreeding, effective population size, and linkage disequilibrium***—Inbreeding coefficients were significantly greater for *parviflora* than *xantiana*, with no overlap between the taxa (*parviflora*: mean = 0.60; *xantiana*: mean = 0.38; $t = 2.84$, $P = 0.015$). We observed consistently higher sequence variation ($N_e$) across individual *xantiana* populations relative to *parviflora* populations for both $\theta_w$ (Fig. 2A; Appendix S2 (see online Supplemental Data); subsp: $F_{1,12} = 178.8$, $P < 0.0001$; locus × subsp.: $F_{14,84} = 12.94$, $P < 0.0001$) and $\theta_\pi$ (Fig. 2B; Appendix S2; subsp. $F_{1,12} = 101.6$, $P < 0.0001$; locus × subsp.: $F_{14,84} = 9.4$, $P < 0.0001$). In *parviflora*, there was a trend toward lower diversity in western populations, but the relationship was not significant (Fig. 2A, B; $\theta_w$: $b = 5 \times 10^{-5}$, $P = 0.109$; $\theta_\pi$: $b = 7 \times 10^{-5}$, $P = 0.091$).

Tajima's *D* for individual populations was marginally significantly greater in *parviflora* than *xantiana* (Fig. 2C; Appendix S2; $F_{1,11.7} = 4.32$, $P = 0.06$). The average *D* was slightly positive in *parviflora* (0.27) and slightly negative in *xantiana* (−0.19). In *parviflora*, we did not find a significant relationship between *D* and longitude (Fig. 2C; *D*: $b = 9 \times 10^{-3}$, $P = 0.488$). In addition to differences in the average value of *D*, an elevated variance in *D* among loci can signify population bottlenecks (Charlesworth et al., 2003); however, we did not detect a significant difference in the coefficient of variation (CV) in *D* between subspecies (subsp.: $F_{1,11} = 1.9$, $P = 0.194$). Similar to *D*, Fu's $F_s$ was significantly lower in *xantiana* (mean = −5.48) compared with *parviflora* (mean = 1.75). Strongly negative values of $F_s$ are consistent with population expansion in *xantiana* (subsp.: $F_{1,11.4} = 28.1$, $P = 0.0002$; locus × sub.: $F_{14,60.5} = 2.1$, $P = 0.020$). There was no significant linear relationship between Fu's $F_s$ and geography in *parviflora* ($b = 3 \times 10^{-1}$, $P = 0.153$).

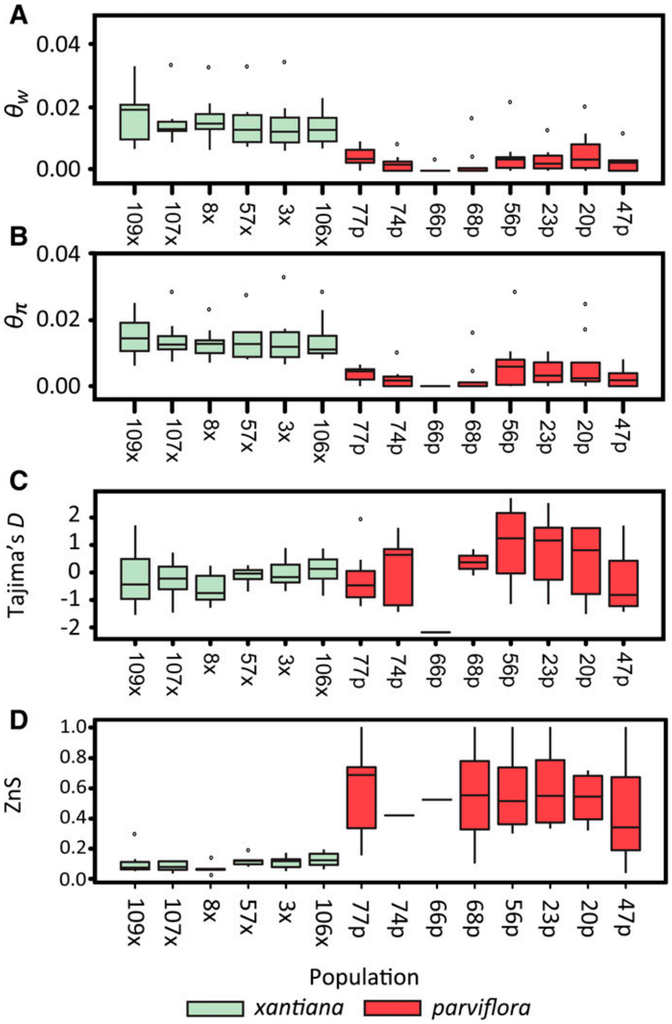

**FIGURE 2** Boxplots illustrating sequence variation between taxa and among populations within taxa. (A) $\theta_w$, (B) $\theta_\pi$, (C) Tajima's *D*, and (D) linkage disequilibrium (*ZnS*). Each boxplot reflects variation among loci within each population. Populations are arranged from west to east across both taxon's ranges.

Linkage disequilibrium (*ZnS*), was significantly greater across *parviflora* populations (mean = 0.56) compared with *xantiana* (mean = 0.10; Fig. 2D; subsp.: $F_{1,11.2} = 73.2$, $P < 0.0001$). There was no significant linear relationship between ZnS and geography in *parviflora* ($b = -1 \times 10^{-3}$, $P = 0.669$).

***Population genetic structure***—The correlation between genetic distance (measured as $F_{ST}$) and geographic distance was significant and positive across *xantiana* populations (Fig. 3A; Mantel $R = 0.638$, 95% CI: $0.497 < R < 0.877$; $P = 0.009$; $R^2 = 0.401$). This result is consistent with populations at migration–drift equilibrium. By contrast, we did not observe a significant isolation by distance

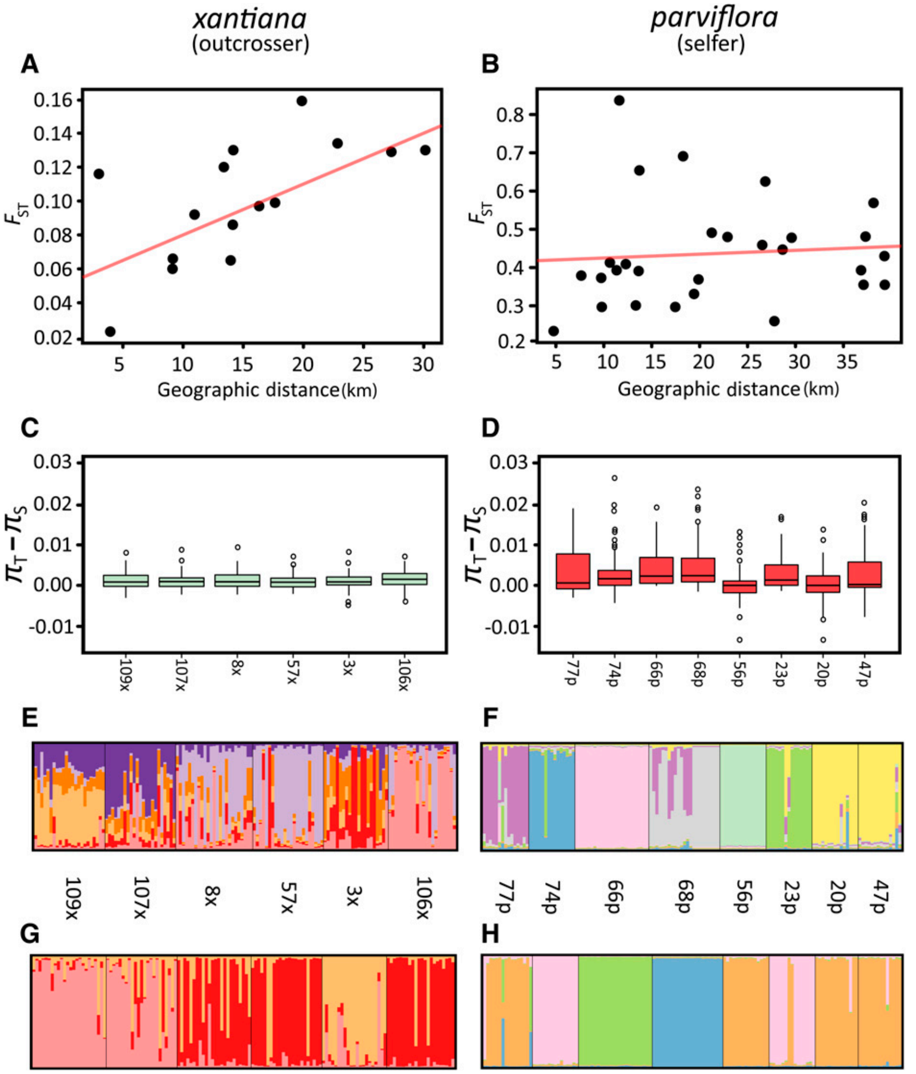

**FIGURE 3** The upper panels show patterns of isolation by distance for (A) *xantiana* and (B) *parviflora* as indicated by the regression of $F_{ST}$ on geographic distance (kilometers). Panels (C) and (D) show pairwise absolute divergence ($\pi_T - \pi_S$) among populations within each taxon for *xantiana* and *parviflora*, respectively. Here, boxplots represent all pairwise comparisons between the focal population and other populations within the taxon for each locus. Panels (E) and (G) show InStruct results for *xantiana* at $K = 6$ and $K = 3$, respectively. Panels (F) and (H) show InStruct results for *parviflora* at $K = 8$ and $K = 4$, respectively. In the InStruct figures, bars represent individuals and larger blocks depict a priori geographic populations.

relationship across *parviflora* even though the geographic region sampled was larger than for *xantiana* (Fig. 3B; Mantel $R = 0.081$, 95% CI: $-0.198 < R < 0.285$; $P = 0.677$). Here, the variance in $F_{ST}$ was high across geographic distances ($R^2 = 0.021$; Fig. 3B), suggesting that genetic drift has had a stronger effect than gene flow on spatial patterns of genetic differentiation. The 95% CI of the Mantel correlation coefficients did not overlap for the two subspecies, indicating significantly different relationships between genetic and geographic distance. It is also notable that the range of $F_{ST}$ values for each taxon did not overlap (Fig. 3A, B). The lowest pairwise $F_{ST}$ for *parviflora* was greater than the highest pairwise $F_{ST}$ for *xantiana*, suggesting that population differentiation is much stronger in *parviflora*.

In contrast to $F_{ST}$, we did not find strong evidence that the magnitude of population structure differed between subspecies using the absolute metric of genetic divergence ($\pi_T - \pi_S$; Fig. 3C, D). There was only a marginally significant difference between subspecies in absolute divergence (subsp.: $F = 3.72$, $P = 0.106$). Although the mean absolute divergence was rather different in each subspecies (*xantiana*: 0.0012; *parviflora*: 0.0027), the variance among estimates in *parviflora* was high (Fig. 3D).

Our analyses of population structure using the clustering method implemented in Instruct indicated support for six clusters in *xantiana* ($K = 6$) and eight clusters in *parviflora* ($K = 8$), both equivalent to the number of geographic populations sampled. Patterns of population structure, however, were quite different in the two subspecies. *Xantiana* exhibited a high degree of admixture among populations with some evidence of differentiation in the frequencies of clusters (Fig. 3E). When forcing $K = 3$ in *xantiana*, there was differentiation between the western (geographically central) populations and the eastern (peripheral) populations (Fig. 3G). In *parviflora*, there was strong differentiation among populations and little admixture (Fig. 3F). Geographic populations corresponded remarkably closely to genetic clusters. Only two populations, 20p and 47p, shared the same cluster and these are geographically most proximal to one another (4.74 km). When forcing $K = 4$ in *parviflora*, there remains a strong correspondence between genetic clusters and geographic populations with little admixture (Fig. 3H).

***Historical demography***—The Skyline plots showed that *xantiana* populations have deeper coalescence than *parviflora*, i.e., a longer time to the most recent common ancestor (note the different *x*-axes for *xantiana* and *parviflora* in Fig. 4). This pattern reflects both population history and species history (e.g., maintenance of ancestral polymorphism). In terms of effective population size changes through time, *parviflora* populations often exhibited signatures of population bottlenecks. Most populations had very little sequence variation followed by a sharp increase, suggestive of colonization bottlenecks (e.g., 66p, 56p, 23p, 74p). Several populations exhibit distinct dips in $N_e$ followed by some recovery (20p, 68p). By contrast, *xantiana* populations did not have evidence of bottlenecks. The four western populations appeared to have stable $N_e$ in recent time (109x, 107x, 8x, 57x). One peripheral population (3x) at lower elevation (855 m) had evidence of a recent decline in $N_e$, indicating that historical population sizes exceed contemporary ones. The other peripheral population (106x) at a higher elevation (1138 m) had a more consistent increase in $N_e$ through time compared with the other populations.

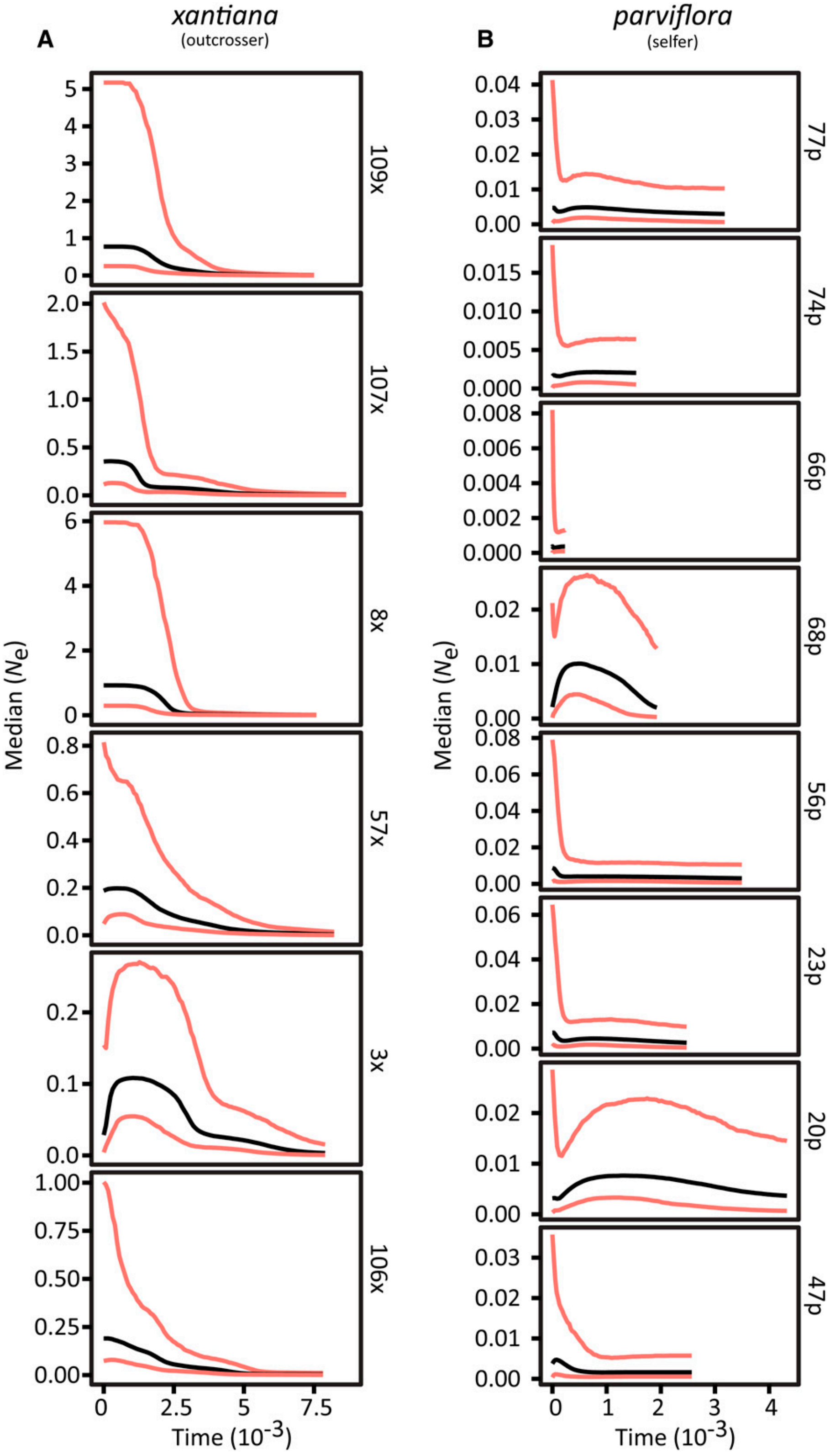

**FIGURE 4** Skyline plots depicting fluctuations in the effective population size ($N_e$) throughout the evolutionary history of each population for (A) *xantiana* and (B) *parviflora*. Black lines represent the median $N_e$; red lines represent the 95% credibility interval. The *x*-axis indicates time in units of mutations per site. Note the different scales on the *y*-axis in each panel and the *x*-axis between each subspecies.

## DISCUSSION

Mating system transitions are well known to have important consequences for the ecology and genetics of populations. In their review of mating systems and population structure, Charlesworth and Pannell (2001) argued that more studies of genetic variation (especially sequence variation) are needed that explicitly compare sister pairs of taxa that differ in mating system (e.g., St. Onge et al., 2011) and that characterize within-population variation rather than only species-wide patterns (e.g., allozyme studies: Wyatt et al., 1992; Koelling et al., 2011). In this study, we conducted a comparative population genetic study of sister *Clarkia xantiana* subspecies that differ strongly in mating system but have only recently diverged. Our results suggest that outcrossing populations have large effective population sizes and are at migration–drift equilibrium, exhibiting evidence of isolation by distance. Although we detected some population structure in the outcrosser, there is significant admixture among populations. By contrast, populations of the selfing subspecies are highly isolated, have experienced strong colonization bottlenecks, and contain a fraction of diversity observed species-wide. Despite examining populations across the entire west-to-east expanse of the taxon, we did not find evidence of isolation by distance. Rather, we find evidence that population structure has been influenced by recent range expansion and strong genetic drift among populations. Overall, our results indicate marked differences between subspecies in the scale of population structure and patterns of sequence variation despite their recent divergence.

A long history of marker-based studies (especially allozymes and microsatellites) has shown that within-population diversity is typically reduced in selfers relative to outcrossers and is among the strongest predictors of population structure (Hamrick and Godt, 1990, 1996; Schoen and Brown, 1991; Charlesworth and Pannell, 2001). These studies have found that selfers have roughly half of the diversity as outcrossers, but the difference is often higher in the smaller set of studies that compare sister selfers and outcrossers (Charlesworth and Pannell, 2001). When patterns of diversity are examined across entire species' ranges, selfers have lower diversity, but the differences are less pronounced (Hamrick and Godt, 1990; Charlesworth et al., 1997). This discrepancy is important because it reflects the potential for selfing taxa to accumulate comparable species-wide levels of diversity (e.g., due to high population subdivision) even though individual populations may contain little molecular variation (reviewed by Charlesworth and Pannell, 2001). Thus far, studies of DNA sequence variation intended to examine the effects of mating system evolution on genome variation typically employ sparse species-wide sampling, where one or few individuals are sampled from individual populations across species ranges; these studies typically find strong reductions in diversity in selfers (e.g., Foxe et al., 2009; Guo et al., 2009; Busch et al., 2011; Pettengill and Moeller, 2012a). This type of sparse sampling makes it challenging to determine the contribution of population structure to patterns of variation in outcrossers vs. selfers.

Our results show strong differences between subspecies and fairly consistent patterns among populations within subspecies. The outcrosser, *xantiana*, has approximately five times as much variation within populations as the selfer, *parviflora* (Fig. 2; Appendix S2); this magnitude is consistent with that observed by Charlesworth and Pannell (2001) in their review of studies of sister pairs of outcrossers and selfers. For *xantiana*, the mean population-specific estimates of diversity are very similar to taxon-wide estimates previously reported (Pettengill and Moeller, 2012a), whereas for *parviflora* mean population-specific estimates are approximately half as large as taxon-wide estimates (Pettengill and Moeller, 2012a). This discrepancy in *parviflora* highlights the importance of population subdivision in affecting range-wide patterns of sequence variation in selfers (Charlesworth and Pannell, 2001). These results are similar to other studies of sequence variation within and among populations, including outcrossers, where estimates of $N_e$ from species-wide samples were higher than estimates from within individual populations owing in part to population structure (e.g., Moeller et al., 2007; Arunyawat et al., 2007). Although a loss of diversity is likely to have occurred as a result of a divergence bottleneck associated with speciation (Pettengill and Moeller, 2012a), these results indicate that the population biology of outcrossers and selfers also contributes to levels of diversity found in populations. As expected from the high homozygosity and low variation observed in the selfing populations, linkage disequilibrium (LD) was considerably higher in the selfer than the outcrosser. The consistency of these patterns across populations and loci indicates that high rates of selfing and demographic factors (e.g., population isolation and drift) are important factors affecting the magnitude of LD genome-wide. These patterns are less consistent with the action of hitchhiking, alone, which may affect only a fraction of the genome. More extensive sampling of loci genome-wide is needed to quantify the scale over which LD decays in the two taxa.

Population genetic structure was much greater in the selfer than the outcrosser as quantified by the relativized metric of population differentiation, $F_{ST}$. This is most readily apparent in the InStruct plots, which show very little admixture among selfing populations but considerable admixture among outcrossing populations. This pattern is also apparent in the magnitude of pairwise $F_{ST}$ values among populations within taxa. The average pairwise $F_{ST}$ values are almost five times higher in the selfer, albeit with considerable variation within each taxon (Fig. 3). It is important to recognize that $F_{ST}$ is a relativized measure of differentiation and therefore is affected not only by the degree of divergence between populations but also by total diversity, which appears in the denominator of the $F_{ST}$ equation (Charlesworth et al., 2003). Therefore, $F_{ST}$ values could be higher in selfers even when divergence does not differ, simply because of the strongly reduced levels of diversity. To examine this possibility more closely, we calculated an absolute measure of divergence ($\pi_T - \pi_S$) and found a considerably less-pronounced difference between subspecies. There was a trend toward greater absolute divergence in the selfer, but this difference was not significant owing to the high variance among estimates in the selfer. It appears that in the selfer, population divergence occurs due to a small number of sequence differences that are sometimes fixed between populations; the very high differentiation in $F_{ST}$ can be attributed largely to low within-population sequence diversity.

Our results show that diversity in selfing populations is much less than half of that of outcrossing populations, making it clear that factors other than the mating system per se must differently influence outcrossing and selfing populations. Specifically, our results suggest that random drift is the dominant force shaping the genetic structure of selfing populations. The outcrosser (*xantiana*) exhibits a pattern of isolation by distance (IBD) despite considerable admixture among populations. By contrast, the selfer (*parviflora*) exhibits no pattern of IBD (i.e., not at migration–drift equilibrium) despite very strong differentiation among populations. For *parviflora*, the relationship between genetic and geographic distance is not only flat but has high variance ($R^2 = 0.007$), suggesting that the stochastic force of genetic drift is much stronger than gene flow in determining structure (Hutchison and Templeton, 1999). Given that differentiation is so pronounced among *parviflora* populations even over short distances (kilometers) and that some populations contain very little sequence variation, it follows that contemporary gene flow among the sampled populations is minimal or nonexistent. Interestingly, this pattern of IBD in the outcrosser but not its sister selfer has also been found in *Capsella grandiflora* and *C. rubella* (St. Onge et al., 2011). Much like our study, the geographic area sampled for the outcrosser was smaller than the selfer and the difference in IBD was still apparent.

One likely cause of differences in IBD between subspecies is the extent to which each relies on animal pollinators for reproduction. For *xantiana*, pollen dispersal may occasionally facilitate genetic exchange among populations because bees are required for outcrossing (Moeller, 2005; Moeller et al., 2012); however, bees are central place foragers and many forage over relatively short distances (Gathmann and Tscharntke, 2002; Greenleaf et al., 2007; Zurbuchen et al., 2010), whereas the lack of dependence on bees for reproduction in *parviflora* may minimize pollen flow over long distances. Another likely cause of high population differentiation is that *Clarkia* species have no known dispersal mechanism other than gravity (Lewis and Lewis, 1955). Finally, extensive searches for populations across each range have suggested that the density of *parviflora* populations across the landscape is lower, which may limit the connectivity of populations through pollen or seeds. Taken together, our results cast doubt on the idea that *parviflora* experiences metapopulation dynamics (frequent extinction and colonization) or has a greater capacity for seed dispersal, as has been observed for many selfing ruderals that track disturbance (Baker, 1974; Lloyd, 1980; Pannell and Barrett, 1998). Had this been the case, we expected that high levels of seed dispersal across the landscape would have resulted in significant admixture. Instead, our results suggest that any kind of long-distance dispersal for the selfer, whether by pollen or seed, is likely rare. Populations may persist even as small colonies because they have seed banks (Eckhart et al., 2011) and can reproduce without mates or pollinators.

Although migration of *parviflora* seeds over large distances is apparently rare, it must have occurred on occasion such that new populations were founded during historical range expansion. For example, our past work suggested that *parviflora* populations migrated from east to west across the range during the Holocene as the western portion of the range became climatically suitable (Pettengill and Moeller, 2012b). During range expansion, it appears that populations experienced strong colonization bottlenecks. Our examination of demographic history using Skyline plots revealed recent times to the most recent common ancestor, especially compared with *xantiana*, and sometimes pronounced reductions in sequence diversity, indicating population bottlenecks (e.g., 20p, 68p). Interestingly, some populations (e.g., 66p, 74p) have little sequence variation and very short coalescence times despite the fact that the number of breeding individuals in these populations is substantial. For example, population 66p, which has virtually no sequence variation in the loci sampled, is a relatively large *parviflora* population that spans more than a hectare of a hillside and contains variation in ecologically important traits (e.g., flower color). Collectively, these patterns suggest that selfing populations may have been colonized only recently during westward range expansion (Pettengill and Moeller, 2012b) and that subsequently little genetic diversity has been added to populations via mutation or gene flow.

## ACKNOWLEDGEMENTS

We are grateful to E. Beckman, R. Bier, S. Evans, J. Iverson, and J. Kopp, and for assistance with collecting seeds, rearing plants, and collecting sequence data. Funding was provided by the National Science Foundation (DEB-1025004) and a John Hall Memorial Research Award from the University of Minnesota.

## LITERATURE CITED

Allee, W. C. 1951. The social life of animals. Beacon, Boston, Massachusetts, USA.

Arunyawat, U., W. Stephan, and T. Städler. 2007. Using multilocus sequence data to assess population structure, natural selection, and linkage disequilibrium in wild tomatoes. *Molecular Biology and Evolution* 24: 2310–2322.

Baker, H. G. 1955. Self-compatibility and establishment after 'long-distance' dispersal. *Evolution* 9: 347–349.

Baker, H. G. 1959. Reproductive methods as factors in speciation in flowering plants. *Cold Springs Harbor Symposium on Quantitative Biology* 24: 177-191.

Baker, H. G. 1965. Characteristics and modes of origins of weeds. *In* H. G. Baker and G. L. Stebbins [eds.], The genetics of colonizing species, 147–168. Academic Press, New York, New York, USA.

Baker, H. G. 1967. Support for Baker's Law—as a rule. *Evolution* 21: 853–856.

Baker, H. G. 1974. The evolution of weeds. *Annual Review of Ecology and Systematics* 5: 1–24.

Barrett, S. C. H. 2002. Evolution of sex: The evolution of plant sexual diversity. *Nature Reviews. Genetics* 3: 274–284.

Barrett, S. C. H., and J. R. Pannell. 1999. Metapopulation dynamics and mating-system evolution in plants. *In* P. M. Hollingsworth, R. M. Bateman and R. J. Gornall [eds.], Molecular systematics and plant evolution, 74–99. Taylor & Francis, London, UK.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Statistical Software* 67: 1–48. doi:10.18637/jss.v067.i01

Briscoe Runquist, R. D., E. Chu, J. Iverson, and D. A. Moeller. 2014. Rapid evolution of reproductive isolation between incipient outcrossing and selfing *Clarkia* species. *Evolution* 68: 2885–2900.

Briscoe Runquist, R. D., and D. A. Moeller. 2013. Resource reallocation does not influence estimates of pollen limitation or reproductive assurance in *Clarkia xantiana* subsp. *parviflora. American Journal of Botany* 100: 1916–1921.

Briscoe Runquist, R. D., and D. A. Moeller. 2014. Floral and mating system divergence in secondary sympatry: Testing an alternative hypothesis to reinforcement in *Clarkia. Annals of Botany* 113: 223–235.

Brown, A. H. D. 1979. Enzyme polymorphism in plant populations. *Theoretical Population Biology* 15: 1–42.

Broyles, S. B., A. Schnabel, and R. Wyatt. 1994. Evidence for long-distance dispersal in milkweeds (*Asclepias exaltata*). *Evolution* 48: 1032–1040.

Busch, J. W., S. Joly, and D. J. Schoen. 2011. Demographic signatures accompanying the evolution of selfing in *Leavenworthia alabamica. Molecular Biology and Evolution* 28: 1717–1729.

Charlesworth, B., D. Charlesworth, and N. H. Barton. 2003. The effects of genetic and geographic structure on neutral variation. *Annual Review of Ecology, Evolution and Systematics* 34: 99–125.

Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* 134: 1289–1303.

Charlesworth, B., M. Nordborg, and D. Charlesworth. 1997. The effects of local selection, balancing polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genetical Research* 70: 155–174.

Charlesworth, D., and J. R. Pannell. 2001. Mating systems and population genetic structure in the light of coalescent theory. *In* J. Silvertown and J. Antonovics [eds.], Integrating ecology and evolution in a spatial context, 73–96. British Ecological Society, Blackwell Sciences, Oxford, UK.

Charlesworth, D., and S. I. Wright. 2001. Breeding systems and genome evolution. *Current Opinion in Genetics & Development* 11: 685–690.

Cheptou, P.-O. 2012. Clarifying Baker's Law. *Annals of Botany* 109: 633–641.

Cheptou, P.-O., and F. Massol. 2009. Pollination fluctuations drive evolutionary syndromes linking dispersal and mating system. *American Naturalist* 174: 46–55.

Courchamp, F., T. Clutton-Brock, and B. Grenfell. 1999. Inverse density dependence and the Allee effect. *Trends in Ecology & Evolution* 14: 405–410.

Dennis, B. 1989. Allee effects: Population growth, critical density, and the chance of extinction. *Natural Resource Modeling* 3: 481–538.

Dobzhansky, T. 1950. Evolution in the tropics. *American Scientist* 38: 209–221.

Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7: 214.

Drummond, A. J., A. Rambaut, B. Shapiro, and O. G. Pybus. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* 22: 1185–1192.

Eckhart, V. M., and M. A. Geber. 1999. Character variation and geographic distribution of *Clarkia xantiana* A. Gray (Onagraceae): Flowers and phenology distinguish two subspecies. *Madrono* 46: 117–125.

Eckhart, V. M., M. A. Geber, W. F. Morris, E. S. Fabio, P. Tiffin, and D. A. Moeller. 2011. The geoography of demography: Long-term demographic studies and species distribution models reveal a species border with limited adaptation. *American Naturalist* 178: S26–S43.

Eckhart, V. M., I. Singh, A. M. Louthan, A. J. Keledjian, A. Chu, D. A. Moeller, and M. A. Geber. 2010. Plant–soil water relations and species border of *Clarkia xantiana* subsp. *xantiana* (Onagraceae). *International Journal of Plant Sciences* 171: 749–760.

Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.

Foxe, J. P., T. Slotte, E. A. Stahl, B. Neuffer, H. Hurka, and S. I. Wright. 2009. Recent speciation associated with the evolution of selfing in *Capsella. Proceedings of the National Academy of Sciences, USA* 106: 5241–5245.

Fu, Y.-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking, and background selection. *Genetics* 147: 915–925.

Gao, H., S. Williamson, and C. D. Bustamante. 2007. A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics* 176: 1635–1651.

Gathmann, A., and T. Tscharntke. 2002. Foraging ranges of solitary bees. *Journal of Animal Ecology* 71: 757–764.

Glemin, S., E. Bazin, and D. Charlesworth. 2006. Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proceedings. Biological Sciences* 273: 3011–3019.

Goldberg, E. E., J. R. Kohn, R. Lande, K. A. Robertson, S. A. Smith, and B. Igic. 2010. Species selection maintains self-incompatibility. *Science* 330: 493–495.

Greenleaf, S. S., N. M. Williams, R. Winfree, and C. Kremen. 2007. Bee foraging ranges and their relationship to body size. *Oecologia* 153: 589–596.

Groom, M. 1998. Allee effects limit population viability of an annual plant. *American Naturalist* 151: 487–496.

Guo, Y.-A., J. S. Bechsgaard, T. Slotte, B. Neuffer, M. Lascoux, D. Weigel, and M. H. Schierup. 2009. Recent speciation of *Capsella rubella* from *Capsella grandiflora*, associated with loss of self-incompatibility and extreme bottleneck. *Proceedings of the National Academy of Sciences, USA* 106: 5246–5251.

Hackney, E. E., and J. B. McGraw. 2001. Experimental demonstration of the Allee effect in American ginseng. *Conservation Biology* 15: 129–136.

Hamrick, J. L., and M. J. W. Godt. 1990. Allozyme diversity in plants. *In* A. H. Brown, M. Clegg, A. L. Kahler, and B. S. Weir [eds.], Plant population genetics, breeding, and genetic resources, 43–63. Sinauer, Sunderland, Massachusetts, USA.

Hamrick, J. L., and M. J. W. Godt. 1996. Effects of life history traits on genetic diversity in plant species. *Proceedings. Biological Sciences* 351: 1291–1298.

Heled, J., and A. J. Drummond. 2008. Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology* 8: 289.

Hey, J., and J. Wakeley. 1997. A coalescent estimator of the population recombination rate. *Genetics* 145: 833–846.

Holsinger, K. E. 2000. Reproductive systems and evolution in vascular plants. *Proceedings of the National Academy of Sciences, USA* 97: 7037–7042.

Hutchison, D. W., and A. R. Templeton. 1999. Correlation of pairwise genetic and geographic distance measures: Inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* 53: 1898–1914.

Igic, B., and J. Busch. 2013. Is self-fertilization an evolutionary dead end? *New Phytologist* 198: 386–397.

Igic, B., R. Lande, and J. R. Kohn. 2008. Loss of self-incompatibility and its evolutionary consequences. *International Journal of Plant Sciences* 169: 93–104.

Jain, S. K. 1976. The evolution of inbreeding in plants. *Annual Review of Ecology and Systematics* 7: 469–495.

Jha, S., and C. W. Dick. 2010. Native bees mediate long-distance pollen dispersal in a shade coffee landscape mosaic. *Proceedings of the National Academy of Sciences, USA* 107: 13760–13764.

Judd, C. M., J. Westfall, and D. A. Kenny. 2012. Treating stimuli as a random factor in social psychology: A new comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology* 103: 54–69.

Kelly, J. K. 1997. A test of neutrality based on interlocus associations. *Genetics* 146: 1197–1206.

Kimura, M. 1953. "Stepping stone" model of population. *Annual Report of the National Institute of Genetics, Japan* 3: 62–63.

Kimura, M., and G. H. Weiss. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* 49: 561–576.

Koelling, V. A., J. L. Hamrick, and R. Mauricio. 2011. Genetic diversity and structure in two species of *Leavenworthia* with self-incompatible and self-compatible populations. *Heredity* 106: 310–318.

Kunin, W., and Y. Iwasa. 1996. Pollinator foraging strategies in mixed floral arrays: Density effects and floral constancy. *Theoretical Population Biology* 49: 232–263.

Lewis, H. 1973. The origin of diploid neospecies in *Clarkia. American Naturalist* 107: 161–170.

Lewis, H., and M. E. Lewis. 1955. The genus *Clarkia. University of California Publications in Botany* 20: 241–392.

Lewis, P. O., and D. Zaykin. 2001. Genetic data analysis: Computer program for the analysis of allelic data, version 1.1. Available at http://phylogeny.uconn.edu/software/ [accessed 17 November 2015].

Lloyd, D. G. 1980. Demographic factors and mating patterns in angiosperms. *In* O. T. Solbrig [ed.], Demography and evolution in plant populations, 67–88. University of California Press, Berkeley, California, USA.

Malécot, G. 1955. The decrease of relationship with distance. *Cold Springs Harbor Symposium on Quantitative Biology* 20: 52–53.

Moeller, D. A. 2004. Facilitative interactions among plants via shared pollinators. *Ecology* 85: 3289–3301.

Moeller, D. A. 2005. Pollinator community structure and sources of spatial variation in plant–pollinator interactions in *Clarkia xantiana* subsp. *xantiana. Oecologia* 142: 28–37.

Moeller, D. A. 2006. Geographic structure of pollinator communities, reproductive assurance, and the evolution of self-pollination. *Ecology* 87: 1510–1522.

Moeller, D. A., and M. A. Geber. 2005. Ecological context of the evolution of self-pollination in *Clarkia xantiana*: Population size, plant communities, and reproductive assurance. *Evolution* 59: 786–799.

Moeller, D. A., M. A. Geber, V. M. Eckhart, and P. Tiffin. 2012. Reduced pollinator service and elevated pollen limitation at the geographic range limit of an annual plant. *Ecology* 93: 1036–1048.

Moeller, D. A., M. A. Geber, and P. Tiffin. 2011. Population genetics and the evolution of geographic range limits in an annual plant. *American Naturalist* 178: S44–S61.

Moeller, D. A., M. I. Tenaillon, and P. Tiffin. 2007. Population structure and its effects on patterns of nucleotide polymorphism in teosinte (*Zea mays* ssp. *parviglumis*). *Genetics* 176: 1799–1809.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, New York, USA.

Nordborg, M. 2000. Linkage disequilibrium, gene trees and selfing: An ancestral recombination graph with partial self-fertilization. *Genetics* 154: 923–929.

Nybom, H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* 13: 1143–1155.

Pannell, J. R., and S. C. H. Barrett. 1998. Baker's Law revisited: Reproductive assurance in a metapopulation. *Evolution* 52: 657–668.

Pannell, J. R., and B. Charlesworth. 2000. Effects of metapopulation processes on measures of genetic diversity. *Philosophical Transactions of the Royal Society, B, Biological Sciences* 355: 1851–1864.

Pettengill, J. B., and D. A. Moeller. 2012a. Tempo and mode of mating system evolution between incipient *Clarkia* species. *Evolution* 66: 1210–1225.

Pettengill, J. B., and D. A. Moeller. 2012b. Phylogeography of speciation: Allopatric divergence and secondary contact between outcrossing and selfing *Clarkia. Molecular Ecology* 21: 4578–4592.

Price, S. C., and S. K. Jain. 1981. Are inbreeders better colonizers? *Oecologia* 49: 283–286.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramos-Onsins, S. E., and J. Rozas. 2002. Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* 19: 2092–2100.

Rosenberg, N. A. 2004. distruct: A program for the graphical display of population structure. *Molecular Ecology Notes* 4: 137–138.

Runions, C. J., and M. A. Geber. 2000. Evolution of the self-pollinating flower in *Clarkia xantiana* (Onagraceae). I. Size and development of floral organs. *American Journal of Botany* 87: 1439–1451.

Schoen, D. J., and A. H. Brown. 1991. Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. *Proceedings of the National Academy of Sciences, USA* 88: 4494–4497.

Schulke, B., and N. M. Waser. 2001. Long-distance pollinator flights and pollen dispersal between populations of *Delphinium nuttallianum. Oecologia* 127: 239–245.

Singmann, H., B. Bolker, J. Westfall, S. Højsgaard, J. Fox, M. A. Lawrence, and U. Mertens. 2015. Package 'afex': Analysis of factorial experiments. Available at http://www.psychologie.uni-freiburg.de/Members/singmann/R/afex [accessed 3 March 2015].

Solbrig, O. T., and R. C. Rollins. 1977. The evolution of autogamy in species of the mustard genus *Leavenworthia. Evolution* 31: 265–281.

St. Onge, K. R., T. Kallman, T. Slotte, M. Lascoux, and A. E. Palme. 2011. Contrasting demographic history and population structure in *Capsella rubella* and *Capsella grandiflora*, two closely related species with different mating systems. *Molecular Ecology* 20: 3306–3320.

Stebbins, G. L. 1950. Variation and evolution of plants. Columbia University Press, New York, New York, USA.

Stebbins, G. L. 1957. Self fertilization and population variability in the higher plants. *American Naturalist* 91: 337–354.

Stebbins, G. L. 1974. Flowering plants: Evolution above the species level. Belknap Press, Cambridge, Massachusetts, USA.

Stephens, M., and P. Donnelly. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics* 73: 1162–1169.

Stephens, M., N. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* 68: 978–989.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.

Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology* 7: 256–276.

Wyatt, R., E. A. Evans, and J. C. Sorenson. 1992. The evolution of self-pollination in granite outcrop species of *Arenaria* (Caryophyllaceae) VI. Electrophoretically detectable genetic variation. *Systematic Botany* 17: 201–209.

Vasek, F. C. 1968. The relationships of two ecologically marginal sympatric *Clarkia* populations. *American Naturalist* 102: 25–40.

Vasek, F. C., and J. Harding. 1976. Outcrossing in natural populations. V. *Evolution* 30: 403–411.

Wright, S. I., and B. S. Gaut. 2005. Molecular population genetics and the search for adaptive evolution in plants. *Molecular Biology and Evolution* 22: 506–519.

Wright, S. I., S. Kalisz, and T. Slotte. 2013. Evoutionary consequences of self-fertilization in plants. *Proceedings of the Royal Society, B, Biological Sciences* 280: 20130133.

Zurbuchen, A., L. Landert, J. Klaiber, A. Müller, S. Hein, and S. Dorn. 2010. Maximum foraging ranges in solitary bees: Only and few individuals have the capability to cover long foraging distances. *Biological Conservation* 143: 669–676.