THESIS

MATING SYSTEM TRANSITIONS IMPACT POPULATION STRUCTURE AND BIODIVERSITY IN *SOLANUM HABROCHAITES*

Submitted by

Chris Miller

Department of Biology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2018

Master's Committee:

Advisor: Patricia Bedinger

Chris Richards
Dan Sloan
Sarah Ward

ABSTRACT

MATING SYSTEM TRANSITIONS IMPACT POPULATION STRUCTURE

AND BIODIVERSITY IN *SOLANUM HABROCHAITES*

Although many plant species have evolved means of preventing self-fertilization, self-compatibility (SC), the ability to set self-seed using pollen and ovules of the same plant is exceedingly common. In the wild tomatoes (*Solanum* section *Lycopersicon*) and other Solanaceous species, plants have evolved a genetic mechanism for preventing self-fertilization. Although many species are entirely either SC or self-incompatible (SI, unable to self-fertilize), one species of wild tomato, *Solanum habrochaites*, is notable for having separate SI and SC populations.

In Chapter 1 of this thesis, I introduce mating system and self-incompatibility in *Solanum habrochaites*. Briefly, Solanaceous plants exhibit a specific type of gametophytic self-incompatibility controlled by cytotoxic, stylar-expressed S-RNases (and other factors) and pollen-expressed male resistance factors. At the northern species margin in southern and central Ecuador, *S. habrochaites* has undergone as least one SI → SC mating system transition. The loss of SI in this region coincides with a unique geographical feature, the Amotape Huancabamba Depression.

In Chapter 2, I explore the loss of SI at the northern species margin using population genetics and reproductive biology. By analyzing the population structure of these populations in combination with controlled crosses, protein expression, and S-RNase allele screening, I identified at least four SC groups resulting from independent transitions from SI→SC. I also identified a fifth SC group of populations which likely arose due to the interbreeding of two separately derived SC populations. Stylar

S-RNase protein expression can also be detected in this region, suggesting previously inactivated S-RNase genes in the parental groups may have become reactivated upon hybridization.

In Chapter 3, I present the analysis of reproductive characters and morphology of newly collected populations of *S. habrochaites*. In one SC group, I find evidence of the "selfing syndrome," a phenomenon in which SC populations are predicted to possess small flowers and unexserted stigmas compared to their SI counterparts. This syndrome was not detected in the other SC populations, however.

In Chapter 4, I describe two new "selfing" S-RNase alleles (*hab-7,* and *hab*-8) in different SC groups using degenerate primers and RNA-seq. *hab-7* likely could encode a functional S-RNase protein, but it is likely unable to function in the SI response due to very low gene expression. The other allele, *hab-8*, detected in a different SC group, cannot express a functional S-RNase protein due to a single nucleotide substitution that produces a premature stop codon.

Finally, in Chapter 5, I summarize my conclusions which support multiple SI→SC mating system transitions at the northern margin of *S. habrochaites* and review evidence that two distinct SC populations have interbred. I also suggest that some S-RNase alleles can be reversibly silenced and reactivated.

ACKNOWLEDGEMENTS

I would like to thank the many individuals and organizations who have helped me personally, academically, and financially over the course of my graduate studies. My first thanks go out to my wonderful advisor Patricia Bedinger for her mentorship and support going back to my time as an undergraduate. I would also like to thank my other committee members, Dan Sloan, Sarah Ward, and Chris Richards for their feedback, guidance, and overall direction in my studies. Other faculty members including Mark Simmons, David Steingraeber, and Jennifer Ackferfield have acted as key figures in my education. I thank them and the rest of the wonderful faculty members of the Biology department. This project was funded by the National Science Foundation. I would also like to thank PMPB for supporting me financially during my first year.

This work would not have been possible without the many incredible members of the Bedinger lab who aided in the generation of data for this project. Alex King spent many hours slaving over a computer screen, tirelessly measuring flowers using ImageJ. An REU student, Nina Paranjpe, and another former undergraduate, Lauren Nalezny helped screen countless accessions for *LhgSRN-1* using PCR. Laura Hantzis, Suzanne Royer, and You Soon Baek helped teach me everything I know about reproductive and molecular biology. I would like to extend the most special of thanks to Amanda Broz who essentially acted as a second advisor to me. Without her guidance, mentorship, and friendship this project truly would not have been possible.

I would like to thank my family for their support and the many times they made the trek to Fort Collins to see us (and to buy us dinner!). Finally, I thank my best friend, Sarah Lockfeld for her love, support, and understanding both during my graduate studies, and since we first met.

TABLE OF CONTENTS

CHAPTER 1:

Introduction to plant mating systems and *Solanum habrochaites*

Introduction

Natural selection acts on the standing genetic variation in a population. Higher levels of genetic variation allow for a greater potential for populations to evolve, especially when they encounter new biotic and abiotic environmental stresses (Fisher, 1930). In plants, self-pollination can lead to inbreeding depression in which significant decreases in genetic diversity reduce a population's adaptive potential and survival probability (Antonovics, 1976; Ellstrand and Elam, 1993; Sakai *et al.,* 2001). Many plants have evolved mechanisms to avoid inbreeding by diminishing the possibility of self-fertilization. Some species have evolved such that male and female reproductive structures are positionally isolated to prevent self-fertilization. For example, many pine species allocate development of pollen and ovules to different portions of the tree; ovulate cones are produced at the crown of the tree whereas pollen cones (microstrobili) develop closer to the ground (Ne'eman *et al.*, 2011). Because pines are wind-pollinated, pollen at the base of the tree is less likely to reach ovules near the crown. Many plants, including some date palms (*Phoenix* sp.), strawberries (*Fragaria* sp.), and junipers (*Juniperus* sp.) are dioecious, with separate male and female individuals, thus completely avoiding the potential negative effects of self-fertilization (Ortiz *et al.*, 2002; Renner, 2014). Some plants have evolved structural mechanisms to prevent self-fertilization within a single perfect (hermaphroditic) flower. For example, Darwin first described this phenomenon in which multiple floral forms can exist within a single species of primrose (*Primula* sp.) (Darwin, 1862). In thrum-type flowers, anthers develop below the stigma, in direct contrast to pin flowers in which the stigma lies below the anthers. The positional separation in thrum flowers reduces the probability of self-pollen landing on the same stigma, as the pollen cannot simply fall onto the stigma directly from the anthers. Other plant species such as breadfruit (*Artocarpus altilis*) and many

Apiaceae further reduce rates of self-fertilization by temporally separating pollen and ovule development (Cruden and Hermann-Parker, 1977; Lloyd and Webb, 1986).

While structural or temporal separation of gametes may diminish instances of self-pollination by lessening the likelihood of self-pollen interacting with stigmas from the same plant, the pollen from these plants is often still compatible with pistils if pollen is applied to stigmas. In contrast, many plant species have evolved genetically determined means of preventing inbreeding, including self-incompatibility (SI) which actively prevents self-pollen from delivering sperm to the ovules, thus forcing outcrossing. SI has arisen multiple times across a broad range of angiosperm families (Haring *et al.*, 1990; Takayama and Isogai, 2005; Sassa *et al.,* 2010; Lanaud *et al*., 2017, Bedinger *et al.*, 2017), suggesting a selective benefit to reducing inbreeding through this mechanism.

Although SI helps preserve genetic diversity through recombination, under some conditions there may be a selective advantage to self-compatibility (SC), especially during migration of species from one set of environmental conditions to another (Baker, 1955; Stebbins, 1957; Baker, 1967; Pannell *et al.*¸ 2015). According to Baker's law (Baker, 1955; Stebbins, 1957), SC populations have the advantage of reproductive assurance when colonizing new environments compared to their SI counterparts. In a SI species, at least three genetically diverse individuals would be required to establish a sufficient pool of genotypes capable of supporting an obligately outcrossing population. Even if enough genetic diversity existed in the migratory population to sustain outcrossing, these new environments may lack the appropriate pollinators to facilitate those crosses. In contrast, a single SC individual could conceivably colonize a novel environment without the need for other individuals or possibly even pollinators. As noted above, SC comes with costs associated with inbreeding depression, resulting in reduced long-term population viability when paired with selection or drift (Lande and Schemske, 1985; Ellstrand and Elam; 1993; Igic and Busch, 2013). Because marginal populations are inherently less genetically diverse than their progenitor populations due to the founder effect, marginal SC populations may be particularly

sensitive to drift. Given a small colonization population size (potentially even as low as a single individual in a SC colonist), outbreeding may not be a viable option due to lack of mate availability. Self-compatibility mitigates problems associated with mate availability, allowing these populations to become established beyond the original species distribution. However, the especially low effective population size of these marginal populations could lead to reduced adaptability and long-term population persistence. It should be noted that even small or infrequent introductions of novel genotypes could alleviate problems associated with low genetic diversity in marginal SC populations. This conflict between colony initialization strategies and long-term viability is especially important when considering the potential of a species to become invasive (Sakai *et al.*, 2010) but should be considered when discussing range expansion of any species. Despite the problems associated with SC, mating system transitions from SI→SC are common in plants (Moeller and Gebre, 2005) and some of the most successful invaders (e.g. *Bromus tectorum*) are obligately SC (Mack, 1981). In the 13-member clade encompassing tomato (*Solanum lycopersicum*) and its wild relatives, SI→SC transitions have occurred at least nine times (Igic *et al.*, 2008), further suggesting a possible adaptive value of SC in some circumstances.

The wild tomatoes (*Solanum* section *Lycopersicon*) offer the ability to simultaneously study reproductive barriers at the individual, population, and species level. Solanaceous species, including tomatoes, exhibit gametophytic SI in which self-pollen can germinate on the nutrient-rich surface of the stigma but pollen tubes are later killed in their journey through the stylar tissue towards the ovules (McClure *et al.*, 1989; Lee *et al.*, 1994; Bedinger *et al.*, 2011; Li and Chetelat, 2015; Baek *et al.*¸2015; Bedinger *et al.*, 2017). Cytotoxic S-RNase proteins secreted into transmitting tissue of styles are taken up by growing pollen tubes and, with the assistance of other stylar factors, degrade self-pollen tube RNA, resulting in death (Lee *et al.*, 1994; Huang *et al.*, 1994, McClure *et al.*, 1989, McClure *et* al., 1990, McClure *et al*. 1999). In a compatible cross, S-locus F-box (SLF) proteins act as resistance factors by

collaboratively recognizing and detoxifying all non-self S-RNases directly or indirectly through ubiquitin 26*S*-proteasome mediated degradation (Sijacic *et al.*, 2004; Kubo *et al.* 2010, Hua and Kao, 2008). This detoxification may occur directly in the pollen tube cytoplasm (Hua *et al.*, 2008; Liu *et al.*, 2014) or following vacuole sequestration of S-RNases (Goldraij *et al.*, 2006; McClure *et al.*, 2011). Self-pollen in an incompatible cross lacks the corresponding SLF proteins required for self-S-RNase detoxification, resulting in RNA degradation followed by arrest of growth and pollen tube death. In wild tomato species, genes encoding both male resistance SLF factors and female S-RNases are inherited as a tightly linked unit on Chromosome One, known as the S-locus.

Studies in our research group have demonstrated that SI pollen rejection mechanisms overlap with those involved in interspecies pollen rejection (Bedinger *et al*., 2011; Baek *et al*., 2016; Broz *et al*., 2017a). For example, cultivated tomato (*S. lycopersicum)* lacks expression of S-RNase and, another pistil SI factor, HT protein, and is unable to reject pollen of any closely related wild species (McClure *et al*., 1999; Covey *et al*., 2010). However, introduction of genes encoding these two pistil-expressed proteins creates an interspecific barrier (Tovar-Méndez *et al.*, 2014; Tovar-Méndez *et al.*, 2017). Pollen-expressed SLF proteins and non-S-locus encoded pollen SI factors, including other ubiquitin ligase proteins like Cullin I, are also involved in interspecific incompatibility (Li and Chetelat, 2010; 2014; 2015; Markova *et al.*, 2016). In general, SI→SC mating system transitions in the Solanaceae are thought to be initiated by the loss of pistil SI factors (Bedinger *et al.*, 2017). Pollen SI factors can then also be lost by relaxed purifying selection and mutation accumulation (Markova *et al*., 2016; Broz *et al*., 2017a). One species of wild tomato, *Solanum habrochaites*, exhibits a particularly diverse array of reproductive characters including intraspecific incompatibilities, interpopulation reproductive isolation, and interspecific barriers.

S. habrochaites (previously known as *Lycopersicon hirsutum*) is distributed across the western Andes of Ecuador and Peru, generally at high elevations but extending down to sea level in central

Ecuador (Peralta *et al.*, 2008; Grandillo *et al*., 2011). At the northern and southern species margins, independent transitions from SI→SC have occurred (Rick and Chetelat, 1991). As predicted, these marginal SC populations are less genetically variable than the SI populations of the center of the species distribution (Rick *et al.*¸1979; Sifres *et al*., 2010). Historically a distinction has been made between the Peruvian and Ecuadorean populations with the latter considered a separate subspecies, *S. habrochaites* ssp. *glabratum,* characterized by glabrous stems and being entirely SC. The Ecuadorean subspecies was also reported to be reproductively incompatible with SI populations (Martin, 1963). However, recent studies, including the work presented in this thesis, have shown these historical descriptions of *S. habrochaites* are not strictly accurate, especially at the northern species margin in Ecuador. We have shown that populations of southwestern Ecuador—north of the border but west of the Andes, near the town of Sozoranga—are demonstrably SI and can interbreed with other SI populations to the south in Peru (Broz *et al.*, 2017a). The SC populations at the northern margin vary in reproductive characters—the most northern population may even represent a case of ongoing, incipient speciation because they are unilaterally incompatible with SI populations due to additional loss of pollen factors after an initial pistil-side mutation leading to SC (Markova *et al*., 2016; Broz *et al*., 2017a). Some northern populations have also lost additional pistil factors involved in interspecific pollen tube rejection (Broz *et al.*, 2017a). Other SC populations retain both the ability to breed with SI populations broadly across the species range and to prevent interspecific hybridization (Broz *et al.*, 2017a).

The work described in this thesis identifies multiple SI→SC mating system transitions in northern *S. habrochaites*. One potential factor driving these transitions to SC are the significant changes in altitude and climate as *S. habrochaites* spread through a region near the Ecuador-Peru border that acts as a major generator of biological diversity, the Amotape-Huancabamba Depression (AHD). Also known as the Andes Depression, the AHD is a region of cordillera disruption spanning northern Peru and southern Ecuador (Sillitoe, 1974; Weigend, 2004) in which the mountains essentially “break up.” The

associated microhabitats of the AHD, with varying elevations and climates, act to produce a major biodiversity hotspot for both plants and animals (Berry, 1982; Weigend, 2002). We propose that as *S. habrochaites* spread north, encountering for the first time low-elevation, warm environments, selection for SC according to Baker's law led to multiple, independent mating system transitions at the northern species margin. In this work, I provide evidence supporting this idea by describing the population structure of *S. habrochaites*, by phenotypically characterizing newly collected wild populations in the region, and by identifying different putative "selfing" S-RNase alleles associated with different SC populations. Together, my data supports the role of mating system transitions as generators of biodiversity.

CHAPTER 2:
Reproductive traits are associated with population structure in *Solanum habrochaites*

In the Solanaceae, the machinery of the gametophytic S-RNase-based self-incompatibility (SI) system prevents self-fertilization (McClure *et al.*, 1989; Lee *et al.*, 1994; Bedinger *et al.*, 2011; Li and Chetelat, 2015; Baek *et al.*¸2015; Bedinger *et al.*, 2017). In the tomato clade, this machinery is also involved in preventing interspecific hybridization (Bedinger *et al.*, 2011; Tovar-Méndez, 2014; Li and Chetelat, 2015; Bedinger *et al.*, 2017). SI → SC (self-compatibility) mating system transitions are common at species margins when mate limitation can select for reproductive assurance by self-fertilization (Baker, 1955; Stebbins, 1967; Baker, 1967; Moeller and Gebre, 2005; Busch, 2009; Pannell *et al.* 2015). At both the northern and southern species margins of the wild tomato *Solanum habrochaites*, some populations have lost pistil SI factors, becoming self-compatible, and in some cases additional loss of pollen SI factors results in becoming unidirectionally reproductively isolated from SI populations (Martin, 1963; Markova *et al*., 2016, Broz *et al*., 2017a). Our results suggest that the phylogeographic history of *S. habrochaites* is intimately involved with the loss and generation of reproductive characters.

While multiple distinct groups have previously been described (Sifres *et al.,* 2011; Broz *et al.* 2017a), it was unknown exactly how many independent SI→SC transitions occurred at the northern species margin of *S. habrochaites* and how those transitions arose. In this chapter, I seek to answer these questions by describing the population structure of publicly available (tgrc.ucdavis.edu) populations of *Solanum habrochaites* and by reproductively characterizing these same populations. I hypothesize that the northernmost SI populations acted as progenitors to four distinct SC groups, two of which have interbred to form a fifth, admixed group. Because each of the groups detected by the population structure is also associated with a unique suite of reproductive characters, I propose that in northern *S. habrochaites*, mating system transitions act as generators of biodiversity.

Materials and Methods:

At Colorado State University, *S. habrochaites* plants were grown from seed obtained from the Tomato Genetic Resource Center at UC Davis (TGRC, tgrc.ucdavis.edu) in an irrigated outdoor common garden in summer of 2017. Seeds were first germinated in soil on a light shelf in the lab and then seedlings were transferred to larger one-gallon pots in the greenhouse. Plants were propagated using cuttings, and each cutting was transferred from a four-inch pot directly into the soil of the outdoor common garden after reaching approximately six inches in height in the greenhouse. In both the greenhouse and in the common garden plants were watered twice a day using a timed drip irrigation system. At least three genetically distinct individuals (each grown from a separate seed) of each accession were used for phenotyping when available, but hail damage and the failure of some plants to thrive sometimes resulted in fewer individuals being used. For the STRUCTURE analysis, each population is represented by a single genetic individual. Throughout this thesis, the words accession and population are used interchangeably and refer to a single geographic collection of *S. habrochaites*.

Mating system was determined using self-pollinations as previously described in Broz *et al.* (2017a). Briefly, open flowers from an inflorescence were removed to ensure no cross-fertilization had taken place before the remaining buds were protected from pollination by wild bumblebees using nylon mesh bags. As flowers opened, inflorescences were "buzzed" using a tooth polisher to mimic bee visitation and promote self-pollination. If production of self-fruit was observed, plants were recorded as SC. If plants failed to set self-fruit using this approach, hand pollinations were performed. In this case, flowers were emasculated the day before anthesis by removing the anther cone and pollinated the same day using self-pollen. In the event that fruit development was not observed, hand-pollinations were performed again and pollen tube growth in styles was evaluated. Styles were collected into fixative (3 EtOH : 1 glacial acetic acid) 48 hours post pollination. After 24 hours followed by overnight decolorization treatment with 5M NaOH, samples were stained with Aniline blue fluorochrome (ABF)

and examined using fluorescence microscopy to visualize pollen tubes in styles. When at least three pollen tubes could be visualized at the base of the style in multiple independent crosses, and fruits formed after hand-pollination, then a plant was considered SC.

Inter-population crosses using LA0407 as a pistil-tester and LA1777 as a pollen-tester were performed using the same methodology, but stigmas were pollinated 24 hours after emasculation to ensure that stigmas were mature. Expression of S-RNase and an additional pistil SI factor, HT protein, was assessed using anti-peptide antibodies specific to the two proteins (Covey *et al*., 2010; Chalivendra *et al*., 2013). Allele-specific primers (Table 2-1) for the previously characterized *LhgSRN-1* allele and two newly described S-RNase alleles (*hab-7, hab*-8) were used in PCR reactions, as described previously (Broz *et al.*, 2017a) and in Chapter 4.

Targeted Sanger sequencing for population structure

To enhance an ongoing restriction site-associated DNA sequencing (RAD-seq) population study initiated by Drs. Robert Last and Gaurav Moghe at Michigan State University, eighteen accessions of *S. habrochaites* from Ecuador and Northern Peru (Fig. 2-1) were analyzed using targeted Sanger sequencing (TSS). Twenty-four loci polymorphic loci were selected by STACKS from the original RAD-seq data based solely on the number of polymorphisms across all populations in the original RAD-seq dataset. Of the original twenty-four loci, only twenty-two are included in this analysis due to problems associated with highly repetitive, AT-rich sequences flanking the intended target region in one primer set and poor quality, nonspecific sequence results in another primer set. An initial nucleotide BLAST was performed using the *lyc4* (Aflitos *et al.*, 2014; http://www.tomatogenome.net/) *S. habrochaites* assembly to identify longer sequences for each of the 22 loci. These corresponding longer sequences were then used as a query in a second nucleotide BLAST to identity polymorphisms between the intended targets and similar loci present in the available *S. pennellii* and *S. lycopersicum* genomic

sequences (Table 2-1). Primers were designed with the aid of IDT Oligo Analyzer tool (Owczarzy *et al.*, 2008) to specifically amplify polymorphic regions of the top BLAST hit (the presumed intended target) and not potential off-target loci (Table 2-1). For all loci, appropriately sized PCR products were amplified using DNA of accession LA0407, purified (Zymo, Irvine,CA), and submitted for sequencing (Qunitara Bio, San Francisco, CA) to confirm that the intended targeted loci were amplified with each primer set. Sequenced products were aligned in MEGA7 (Kumar *et al.*, 2016) using Muscle (Edgar, 2004) with the original intended target, the corresponding sequence of LA0407 from the original RAD-seq dataset, and the top BLAST hits.

After confirming that the proper locus was amplified using each primer set, purified PCR products produced using DNA of a single individual from each of the 18 selected accessions were submitted for sequencing (Genewiz). Sequences were assembled into separate files using custom scripts and aligned with their consensus sequences using MEGA7. The diploid state of each locus of each population was determined by aligning the sequences for all populations, trimming off poor-quality sequences, and examining the set of trace files for each locus by eye for heterozygous base calls (Chromas Pro). If no ambiguous calls were present in the trace files, the individual was assumed to be homozygous at that locus.

In order to combine targeted Sanger sequences with the RAD-seq data, the trimmed Sanger sequences were BLASTed against the set of all RAD loci consensus sequences to identify each corresponding RAD-seq locus. Sequences representing the 22 loci were extracted from the total set of RAD data (51 samples) in the populations program of STACKS using a whitelist of loci identified by BLAST, combined with their targeted Sanger counterparts using custom scripts, aligned, manually inspected, and trimmed when necessary. PGDSpider (Lischer and Excoffier, 2015) was used to call allele variants for each locus; the separate matrices generated by PGDSpider were then combined to create the STRUCTURE (Pritchard *et al.* 2000) input matrix of allele variants for all 22 loci for the 69 total

samples (15 Sanger sequences from accessions not in the original experiment, three Sanger sequences from accessions included in the original RAD-seq dataset, and the 51 sequences from the RAD-seq samples) using a custom python script. STRUCTURE was run using default parameters with no prior population groups assumed for K=1-8 (three replicates per K) for 10,000 burn-in and 10,000 MCMC cycles (Fig. 2-2). The STRUCTURE results were extracted using STRUCTURE HARVESTER (Earl and vonHoldt, 2011) and replicate runs were combined using CLUMPP (Jakobsson and Rosenberg, 2007). All statistics (adegenet, pophelper, diveRsity), data analysis (pophelper), and plot generation (ggplot2, scatterpie) were performed using R 3.4.1 (R Core Team, 2017; Jombart and Ahmed 2011; Francis, 2017; Keenan *et al.*, 2013; Wickham, 2009; Yu, 2017).

Results:

Mating systems and molecular characters associated with mating system were assessed in previously uncharacterized Ecuadorean *S. habrochaites* populations available from the Tomato Genetics Resource Center and combined with data from previous studies (Covey *et al.* 2010; Broz *et al.* 2017a) to generate the most comprehensive survey of mating system and pistil SI components to date in *S. habrochaites* (Table 2-2). Most *S. habrochaites* populations in Ecuador are self-compatible (SC), except for self-incompatible (SI) populations in the most southwestern region, near the town of Sozoranga. In general, SC populations lack detectable expression of S-RNase, with the exception of at least some individuals in SC accessions PI251305 and PI390515, which express S-RNase as detectable by immunoblotting (Table 2-2, Fig. 2-3). S-RNase protein was detected in all individuals in mixed SI/SC populations (MP) and SI populations, as well as in SC populations at the southern species margin in Peru (LA1927, Covey *et al.*, 2010). HT protein, which functions in both SI (McClure *et al.*, 1999) and in interspecific pollen tube rejection (Tovar-Méndez *et al.*, 2014, Tovar-Mendez *et al.*, 2017), is expressed in all populations except in three near the town of Alausí in central Ecuador: (LA1223, PI251305 and some individuals of PI390515) and in one population (LA4655) near the town of Girón. The previously

characterized, non-expressed *LhgSRN-1 S-RNase* allele (Kondo *et al.*, 2002; Covey *et al.*, 2010; Broz, *et al.*, 2017a) was detected in the northern-most SC *S. habrochaites* populations in central coastal Ecuador (LA4656, LA1624, LA1625, LA0407, PI134417), and in some populations in mountainous regions of central Ecuador at the northern species margin (PI129157, LA1223, PI390515). As described in Chapter 4, we identified a low-expression *S-RNase* allele, *hab-7*, that is associated with SC populations generally found in a mountainous corridor in central Ecuador centered around the town of Loja. At its northern limit near the town of Alausí, populations containing the *hab-7* S-RNase allele overlap geographically with SC populations containing the *LhgSRN-1* allele (Table 2-2, Table 3-1). In fact, different individuals of LA2144, an accession collected in this region of overlap, contain either the *LghSRN-1* or the *hab-7 S-RNase* allele. SC populations from two other geographical regions—near the town of Girón (LA4654, LA4655) and near Cariamanga (LA2101, LA2860)— contain neither the *LhgSRN-1* nor the *hab-7* allele, suggesting that there are additional S-RNase allele(s) associated with SC in Ecuador, one of which (*hab-8*) is described further in Chapter 4.

In addition to assessing mating system and expression of SI genes, inter-population and interspecific reproductive barriers in previously untested accessions were evaluated by examining pollen tube growth in crosses using tester pistils and pollen, as shown in Fig. 2-4. Pollen and pistil testers used were those described in Broz *et al* (2017a). Pistils of SI accession LA1777 were used to test for loss of pollen-side SI resistance factors. Pollen tubes of most accessions tested were generally accepted by LA1777 pistils, with the exception of pollen tubes from LA4656, a population at the northern species margin— a finding consistent with previous results indicating the loss of pollen factors in the northernmost *S. habrochaites* populations (Markova *et al.*, 2016; Broz *et al.* 2017a). Pollen from SC accession LA0407, which lacks pollen factors necessary to overcome SI pistil barriers (Martin 1963; Baek *et al.*, 2015; Markova *et al.*, 2016; Broz *et al.*, 2017a), was used to test for loss of pistil function in the same accessions. We found that rejection of LA0407 pollen tube correlated with expression of S-RNase

proteins; pistils of SC accessions LA4656, LA4654/LA4655, LA2128, LA1252, and LA2860 lack S-RNase expression and generally accepted LA0407 pollen tubes, but S-RNase-expressing styles of LA2855, PI390515 and PI251305 typically rejected LA0407 pollen tubes. Pollen of cultivated tomato, rejected by SI and (most) SC populations of *S. habrochaites,* was used to test interspecific compatibility. Previously, only SC LA1223, collected near the town of Alausí, was found to accept interspecific pollen tubes (Broz *et al.*, 2017a), and this phenotype was hypothesized to be due to the loss of HT protein expression, since this factor is known to affect interspecific pollen tube growth (Tovar-Méndez *et al.*, 2014; Tovar-Mendez *et al.*, 2017). However, our findings with two populations from the same region indicate that HT loss does not strictly correlate with the loss of interspecific barriers; some pistils of PI390515 express HT protein but accept *S. lycopersicum* pollen tubes, and pistils of PI251305 do not express HT protein but reject interspecific pollen tubes (Table 2-2, Fig. 2-4).

Our reproductive phenotyping data indicate that there are at least two, and up to four, distinct SC groups at the northern species margin. Rick *et al*. (1979) detected lower genetic variability in SC populations using allozyme analysis, and a previous population study by Sifres *et al.*, using amplified fragment length polymorphisms (AFLPs) and simple sequence repeats (SSRs) detected three distinct groups in Ecuador (2011). We became aware of an ongoing population study using RAD-seq at Michigan State University that did not include some of the reproductive diversity of Ecuadorean populations of *S. habrochaites.* We wished to determine how northern SC populations relate to each other and to their possible SI ancestor populations using phylogeography, population genetics and the analysis of reproductive traits. In addition, we hoped to determine the number of independent transitions from SI→SC in this region and how those transitions arose. We selected 18 total accessions for targeted Sanger sequencing (TSS) of 22 polymorphic loci. Three of these 18 accessions were sampled in the original RAD-seq dataset and acted as controls to ensure we sequenced the intended target loci. The other 15 additional accessions from Ecuador were selected to more fully capture the entire distribution

of *S. habrochaites*. When these data were combined with the RAD-seq data for the same 22 loci, seven groups that correspond with multiple reproductive characters could reliably be identified throughout the *S. habrochaites* species range (Fig. 2-5, 2-6, 2-7). By comparison, the RAD-seq dataset alone was able to detect six distinct groups.

By including additional accessions, we achieved a higher resolution of groups in Ecuador, splitting Ecuadorean populations into three unique groups compared to the two Ecuadorean groups found using the original RAD-seq dataset. Although significantly more loci were sampled in the RAD-seq study, the higher density sampling in our analysis may have added the extra power necessary to discern this additional group. The northernmost coastal SC accessions of Ecuador (Ecuador region A) previously classified as the "SC-2" mating system group based on reproductive features (Broz *et al.* 2017a), cluster together (Fig. 2-5, orange bars) and contain the *LhgSRN-1* S-RNase allele. Notably, the northernmost, lowest elevation SI accession in Ecuador (LA2868), collected near the town of Arenillas, also clusters with the SC-2 mating system group, suggesting a possible SI progenitor-SC descendent relationship. Accessions in the previously classified "SC-1" mating system group (Broz *et al.*, 2017a) are generally from a mountainous corridor in central Ecuador (Ecuador region D). These accessions contain the *hab-7 S-RNase* allele and cluster together (Fig. 2-5, green bars; Fig. 2-8). The ranges of the SC-1 and SC-2 mating system groups overlap in Ecuador region B, and our data suggest that these two SC groups may interbreed in the region of overlap. For example, while accessions LA1223 and PI390515, collected in or near Alausí, contain the *LhgSRN-1* allele, accession PI251305, which was collected in the nearby town of Sibambe clusters with the SC-2 mating system group, and contains the *hab-7* allele. Further, accession LA2144, collected in Huigra, a town in the same region, includes individuals containing either the *hab-7* or the *LhgSRN-1* alleles.

Two other SC mating system groups are denoted as SC-X and SC-Y because they contain neither the *LghSRN-1* nor the *hab-7* S-RNase allele and do not group with each other or the other SC groups. The

SC-X mating system group consists of accessions collected near Girón (Ecuador region C) that share a unique polymorphism pattern (red, green and orange bars in Fig. 2-5). The SC-Y mating system group includes accessions from Ecuador region E that cluster with the SI populations in this same region, again suggesting an SI progenitor-SC descendent relationship (blue bars). I have also identified an S-RNase allele associated with this SC group (Chapter 4). The Peruvian populations cluster geographically, following the same patterns recovered in the original RAD-seq analysis, with the southern Peruvian SC mating system group ("SC-S"), clustering with nearby SI accessions in Peru region D. One SC accession from this region (LA1927) contains the expressed but low-activity *hab-6* S-RNase allele (Covey et al., 2010).

Discussion

We have discovered reproductive and genetic diversity in Ecuadorean populations of *S. habrochaites* that has been generated as the species migrated northward. Speciation and population spread in western South America is profoundly influenced by the Andes in terms of landscape, ecology, and climate. The uplifting of the Andes cordillera led to innumerable speciation events—either directly as a vicariant barrier, or by creating new environments and habitats for species to colonize (Leubert and Weigend, 2014). The Andes form a significant east-west barrier throughout most of their range, with the exception of the Amotape-Huancabamba Depression (AHD), a region of cordillera disruption spanning northern Peru and southern Ecuador (Sillitoe, 1974; Weigend, 2004). The AHD is a known hotspot of biological diversity (Berry, 1982; Weigend, 2002) and is of exceptional interest when considering that during northern migration of *S. habrochaites* (found 2000-3000 masl in most of its range) this is likely where the species first encountered low elevation, warmer habitats. This environmental transition is striking when examining the distribution of *S. habrochaites* in terms of temperature and elevation (Moyle,2008).

Migration through the AHD could also have promoted mating system transitions in *S. habrochaites.* Baker's law predicts that, as a population migrates into new environments, SC individuals have a reproductive advantage over their SI counterparts due to their ability to colonize under mate-limiting (including pollinator-limiting) conditions in which a single SC individual could readily spawn a new population (Baker, 1955; Stebbins, 1957; Baker, 1967; Pannell *et al.* 2015). SI → SC transitions can mitigate reproductive challenges associated with migration into new habitats, and are common at species margins (Richards, 1986; Barret 2002; Moeller and Gebre. 2005). According to previous studies, this kind of transition has occurred at least twice at the northern *S. habrochaites* species margin as SI populations migrated through the warmer, low-elevation climate of the AHD, (Sifres *et al.*, 2011; Broz *et al.,* 2017a). Sifres et al., (2011) detected three ecological and genetics-based groups in Ecuador which they designated as the A (Central Ecuador), B (Central Loja), and C (Western Loja – Piura). These groups generally correspond with our Ecuador A, Ecuador D, and Ecuador F groups respectively (Fig. 2-5). However, our simultaneous analysis of population structure and reproductive characters have allowed the differentiation of previously unidentified SC-X and SC-Y mating system groups in Ecuadorean *S. habrochaites* (Fig. 2-5). In fact, the data suggest that up to four independent mating system transitions occurred in this region (Fig. 2-5).

The northern-most SC-2 mating system group (orange bars, Fig. 2-5), associated with the *LhgSRN-1* S-RNase allele, may represent the oldest SI → SC transition, since these populations display a "selfing syndrome" phenotype that includes reduced flower size (Rick *et al.*, 1979; Sicard and Lenhard, 2011; Broz *et al.*, 2017a). For further discussion of morphology and reproductive barriers in Ecuadorean *S. habrochaites*, see Chapter 3. Further, these northernmost populations have accumulated pollen-side mutations that have created a unidirectional reproductive barrier between the SC-2 and ancestral SI populations (Markova *et al.*, 2016; Broz *et al.*, 2017a). Our analysis is somewhat equivocal regarding the history of SC-X populations near the town of Girón (LA4654, LA4655). Interestingly, Sifres *et al.* (2011)

tested LA4655 (also known as ECU0436) and found multiple potential group predictions by STRUCTURE for each of their marker types. It is still unknown if populations from the Girón region represent an unusually heterozygous population or are the result of continued breeding between populations to the north and south of them. However, the lack of either the *LhgSRN-1* or the *hab-7* S-RNase alleles (or the *hab-8* allele, see Chapter 4), combined with their unusual pattern of polymorphism suggest a separate mating system transition for the SC-X mating system group. The SC-Y mating system group in some ways reproductively resembles other SC populations in southern Ecuador but lacks the *hab*-7 S-RNase allele and is genetically more closely related to SI populations east of the Andes than SC-1 populations to the west. We believe the SC-Y populations may represent a fourth mating system transition in Ecuadorean *S. habrochaites*. This idea is further supported in Chapter 4 with the discovery of another possible "selfing" S-RNase allele, *hab-8.*

Both Rick *et al.* (1979) and Sifres *et al.* (2011) detected lower levels of heterozygosity at the northern and southern species margin, suggesting that centrally distributed SI populations acted as progenitors to the rest of the species distribution. We also found heterozygosity to generally be higher in SI populations and lower in SC populations (Table 2-3), although it should be noted this analysis only sampled one individual from each collection site. However, Sifres *et al.* reported the lowest heterozygosity in Central Ecuador (our Ecuador A group) and our analysis instead suggested the lowest levels of genetic diversity further south, in the Ecuador D group near the town of Loja. This somewhat puzzling result is likely the result of sampling error, but an ancestral bottleneck or sustained inbreeding would also result in a loss of genetic variation in this region.

Interestingly, our data suggest that the ranges of two independently derived SC Ecuadorean populations – the northernmost SC population containing the *LhsgSRN-1* allele and the central Ecuador SC population containing the *hab-7* allele – overlap near the town of Alausí, resulting in a zone of interbreeding. Consistent with this idea, the analysis of accession LA1266 from this region by Sifres *et al.*

(2011) produced ambiguous group predictions. Furthermore, allelic richness was higher in this region than in other northern SC populations (Table 2-3). If this population arose from the admixture of two distinct SC groups as hypothesized here, the combinations of alleles from both regions might result in a higher genetic diversity.

In some cases, our analysis suggests SI progenitor- SC descendent relationships between SI and SC populations, although it is understood that these relationships could also reflect recent admixture (Fig. 2-5, 2-6). For example, the northern-most SC-2 mating system group (in Ecuador regions A and B) clusters with the northern-most SI population (LA2868). Genetic similarities between Ecuadorean SI accessions and members of the SC-1 can also be detected. The SC-Y mating system group also genetically resembles SI populations in Southern Ecuador. Principal Components Analysis (PCA) analysis further suggests a relationship between Ecuadorean SI populations and the SC-X mating system group (Fig. 2-6), despite an absence of detectable genetic similarity in the STRUCTURE analysis (Fig. 2-5). Further research will be required to determine the precise number of mating system transitions that have occurred and their underlying causes at the northern *S. habrochaites* species margin.

Taken together, our results support the notion that a combination of migration through the AHD and the generation of multiple, independent SI→ SC mating system transitions acted as the key drivers that generated the remarkable reproductive and phenotypic diversity that we have observed at the northern species margin of *S. habrochaites*.

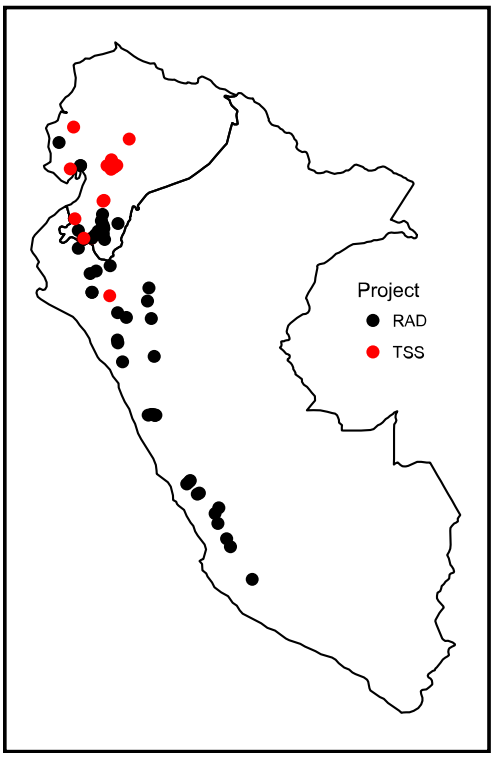

Fig. 2-1. Map of sampled accessions of S. habrochaites for Targeted Sanger Sequencing (TSS). Black points represent populations in the original RAD-seq experiment by MSU. All plotted points are publicly available from the TGRC or USDA seed banks.

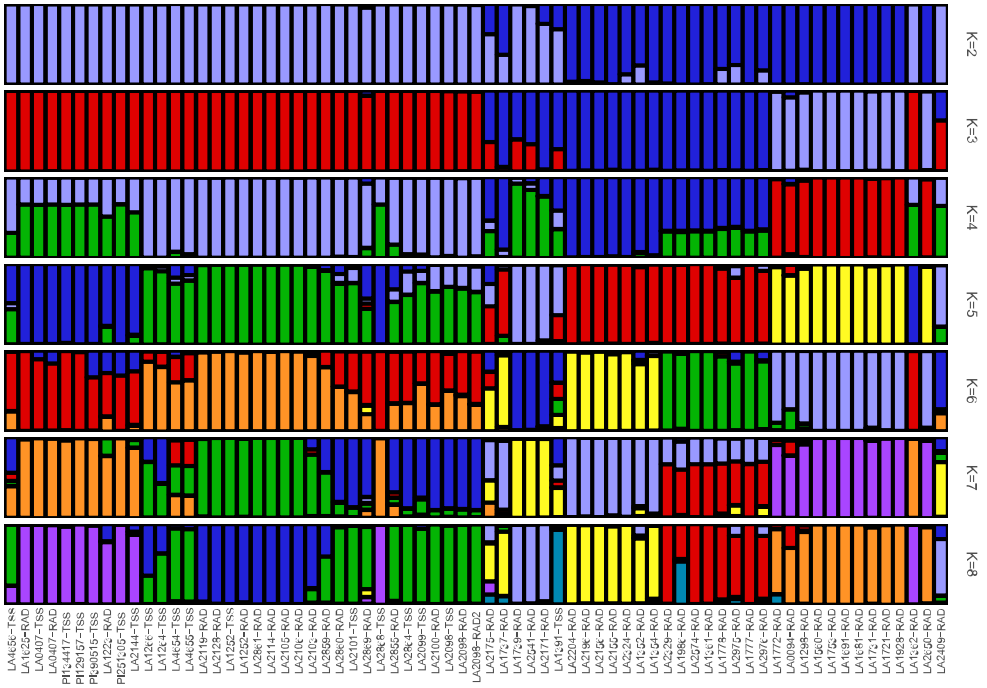

Fig. 2-2. STRUCTURE results for 2-8 predicted groups (K). Each bar represents a single individual of a population. Colors represent shared genetic similarity in the 22 sampled TSS loci. Accessions are sorted generally from North to South but retaining regional E-W associations. For analysis of K=7 alone, see Fig. 2-5.

Fig. 2-3. Immunoblots for S-RNase (top panel) and HT (bottom panel) proteins. The positive controls (+) are extracts from an SI population (LA1777) for S-RNase and LA2119 for HT. Accessions are indicated above horizontal lines and individual identifiers are oriented vertically below each accession. Each lane represents the stylar-extracted proteins of multiple flowers from a single individual.

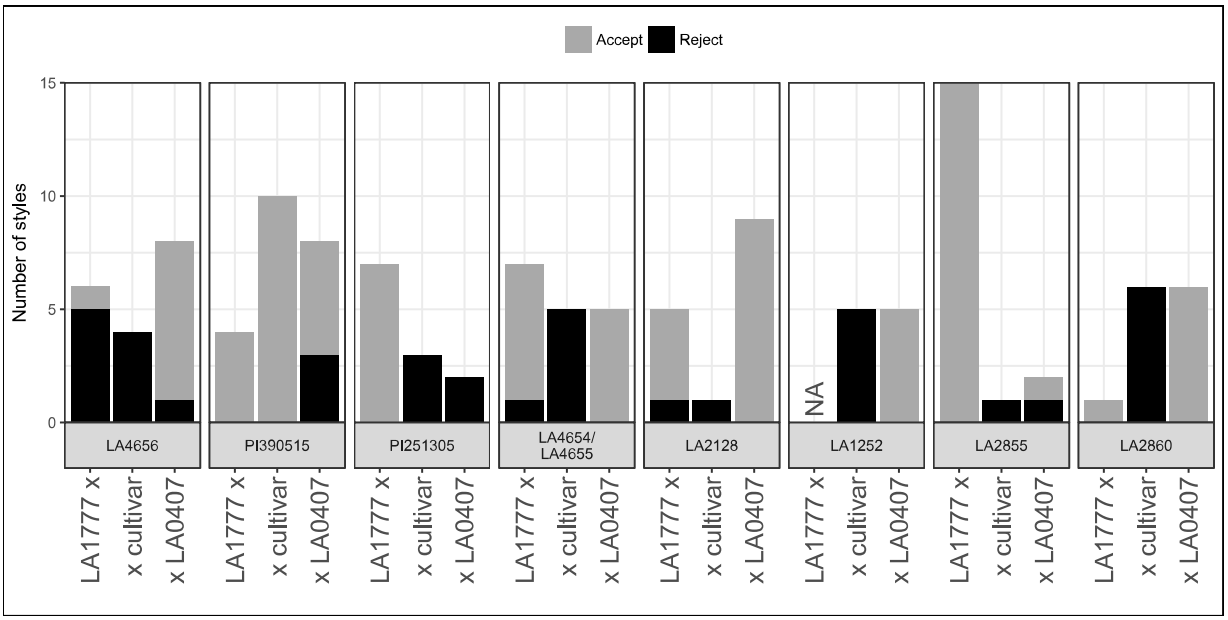

Fig. 2-4. Pollen tube growth in inter-population and interspecific test crosses. Pollen tube growth in pistils was assessed 48 hours after pollination in previously untested *S. habrochaites* accessions. "LA1777 x" crosses test for pollen function. "x cultivar" crosses test for interspecific pollen rejection. "x LA0407" crosses test for pistil function. "Accept" (gray) indicates that a minimum of 3 pollen tubes reached ovaries at the base of the style. "Reject" (black) indicates that pollen tubes grew but failed to reach ovaries. Each bar represents the sum of all crosses of each type for all individuals. Results from accessions LA4654 and LA4655 are combined due to low numbers of crosses with successfully germinated pollen and the very close spatial proximity of these accessions

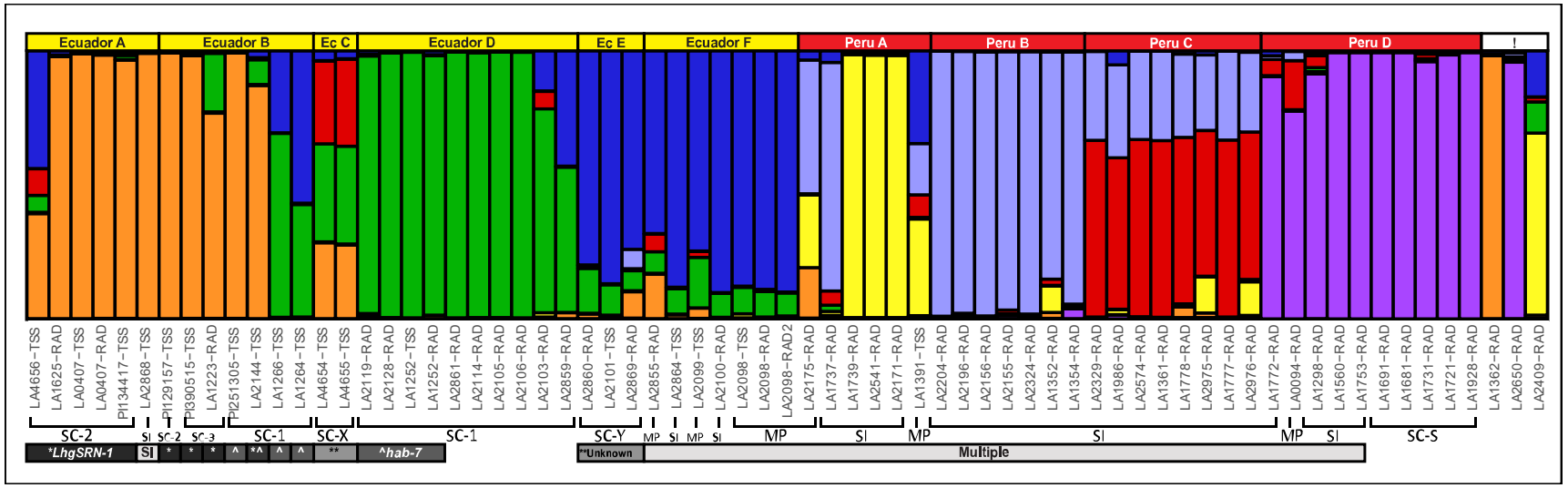

Fig. 2-5. STRUCTURE plot generated using 22 polymorphic loci with RAD-seq data (RAD) combined with targeted Sanger sequencing (TSS). Clustering of populations at K = 7 is shown, with accessions sorted generally from North to South but retaining regional relationships informed by the East/West division of the Andes. Regional divisions indicated in top bar include Ecuador A) Northern low-elevation, coastal SC populations. Ecuador B) Northern SC populations near the town of Alausí. Ecuador/Ec C) SC populations near the town of Girón. Ecuador D) SC populations in south-central Ecuador running N-S along the Loja corridor. Ecuador/Ec E) Predominantly SI populations west of the Andes near the town of Sozoranga. Peru A) Populations from the Amotape-Huancabamba Depression in Peru, between the Peru-Ecuador border and the Río Chamaya. Peru B) Between the Río Chamaya and the Río Chicama. Peru C) Between the Rio Moche and the Río Culebras. Peru D) South of the Río Chancay. (!): These populations disagree with stated geographical position and are consequently sorted separately. Brackets at the bottom of the plot indicate the mating system groups proposed by Broz *et al.* (2017a) with modifications based on additional findings in this work. SI = Self-incompatible, SC= Self-compatible, MP = mixed populations with SI and SC individuals. Bottom horizontal bars indicate whether the *LhgSRN-1* (*) or hab-6 (^) *S-RNase* allele was detected in the accession (both *hab-7* and *LhgSRN-1* were detected in LA2144). If accessions were screened for the *hab-7* and *LhgSRN-1 S-RNase* alleles but neither was detected, the allele was considered "Unknown" (**). Accessions lacking an allele bar were not tested for any *S-RNase* allele. SI and mixed SI/SC accessions possess multiple S-RNase alleles.

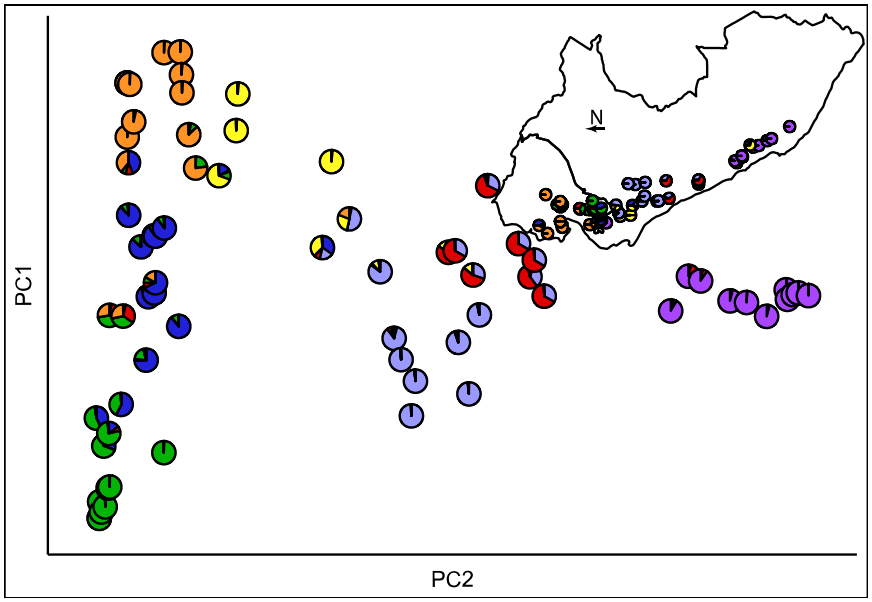

Fig. 2-6. Principal components analysis of allele states, generated from the STRUCTURE input matrix. Pies reflect the proportions and colors of the STRUCTURE plot for K = 7 with 22 loci (Fig. 2-5). These pies are also plotted according to geographical position with the map of Ecuador and Peru turned on its axis. This figure includes the three "misplaced" accessions.

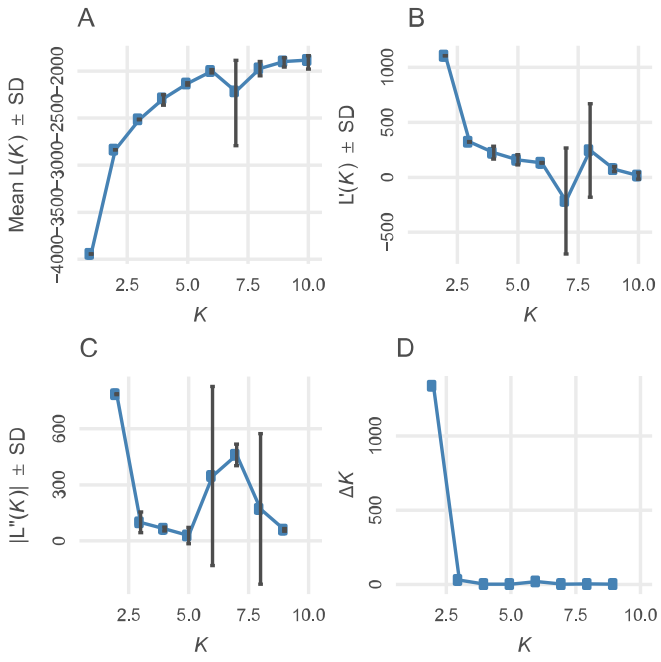

Fig. 2-7. Statistical analysis for the best estimate of K based using the Evanno method. However, because power is low with only 22 loci and the statistical estimates of K varies, K=7 was selected based on meaningful reproductive characters.

Table 2-1. Primers used in TSS and for S-RNase screening. The 5' → 3' sequence is shown for each primer pair along with the top BLAST hit by blastn.

Target/RAD locusID	Primer Name	Sequence(5'->3')	BLAST
77821	CM1-F	GTGTATGAAACTCTTGGACATTAGCCT	HG975443.1:26482777-26482934 Solanum pennellii
	CM2-R	GGATACGGAGAATTTCGTTGAGTAAC	chromosome ch04 complete genome
129683	CM3-F	GTAGCAATGGTAGAGGGAGGTATC	HG975442.1:606999-607147 Solanum pennellii
	CM4-R	CTAGTGGCAACCTGAACTATATCC	chromosome ch03 complete genome
18212	CM7-F	ACAAACTCTACTCCGTTGTCC	HG975447.1:70161129-70161379 Solanum pennellii
	CM8-R	TTGAAGTGGATAGGTTTGAGC	chromosome ch08 complete genome
46312	CM11-F	GTTGCTAGTATGTATCCAATAATTATG	HG975441.1:53247706-53248006 Solanum pennellii
	CM12-R	TGCAAAGGTGAAGAAAGCC	chromosome ch02, complete genome
5644	CM13-F	TGAAAACATAATATGCCTCAGG	HG975450.1:45046195-45046331 Solanum pennellii
	CM14-R	TTGTGCGTTTATTGTCTGAAG	chromosome ch11 complete genome
96307	CM17-F	AGTGATATTTAGTGGCGAATG	HG975450.1:65336264-65336471 Solanum pennellii
	CM18-R	AGCATAGAAAGGAGAGGAATC	chromosome ch11 complete genome
135142	CM19-F	TGAAATGTGCTGATGTTCATCC	HG975442.1:67105762-67105974 Solanum pennellii
	CM20-R	AGCTCGAATTAAGGAGTTCCAG	chromosome ch03 complete genome
19144	CM21-F	TAACTTGTGGTGACAGTGTCG	HG975445.1:49824838-49825029 Solanum pennellii
	CM22-R	ATGCACTCAAGGTTCGTTCTC	chromosome ch06 complete genome
32713	CM23-F	AAGCATAGGGATCATTGTTGTC	HG975451.1:27973064-27973251 Solanum pennellii
	CM24-R	GTGTCTATTTGCAATTATGTGTTTC	chromosome ch12 complete genome
74694	CM25-F	GTGGTGCATTCTTCCAATATCAC	HG975443.1:50601610-50601815 Solanum pennellii
	CM26-R	TCTTCCCATCCCACTTCAAAATTTC	chromosome ch09 complete genome
119691	CM27-F	TGTAGAGACAATGGTAGCATCC	HG975442.1:69209752-69209910 Solanum pennellii
	CM28-R	TCTTCCTCTTAGATTTGAACTTGC	chromosome ch03 complete genome
10871	CM29-F	ATGGGACAAGTCTTGCCTTTATG	HG975441.1:37216730-37216857 Solanum pennellii
	CM30-R	ATTTCTTTCTCTTCAAACAGTCCAC	chromosome ch02 complete genome
63672	CM31-F	ACAAATTGTCCACCCTCTTTAGC	HG975447.1:58548003-58548153 Solanum pennellii
	CM32-R	TGAGCACTGCTGGTTTCTTG	chromosome ch08 complete genome
85323	CM33-F	AGGCATCTCAGTCCCAAAG	HG975451.1:5833203-5833404 Solanum pennellii
	CM34-R	TCCCTAGAAAGATCCTCCGTATC	chromosome ch12 complete genome
45627	CM35-F	GCAATGATCATGGAAGAGTTGG	HG975442.1:71500263-71500432 Solanum pennellii
	CM36-R	TTTCGCCTCCTCTCTGC	chromosome ch03 complete genome
10450	CM41-F	TAGGTGAGGCGGACAAGTAATC	HG975451.1:66431277-66431396 Solanum pennellii
	CM42-R	GTACTCAACTAGCAGATTTTCCTAATAG	chromosome ch12 complete genome
54720	CM43-F	CTCCAAGCACCAAGATCAAATC	HG975451.1:27370093-27370272 Solanum pennellii
	CM44-R	GCCAAGCCAAGGGAAAATTG	chromosome ch12 complete genome
32628	CM52-F	AGGTTAAATCTCAAGGTGAATGTAGG	HG975449.1:33929749-33930049 Solanum pennellii
	CM53-R	GGTTGAAGCATCATCCTTTGGAAC	chromosome ch10, complete genome
53693	CM55-R	CTCAGAGGTATTATAGAAGCCCTTTAC	HG975443.1:37137091-37137391 Solanum pennellii
	CM56-F	GAAATGATTAAACCACAGCTAATGTCAG	chromosome ch04, complete genome
138956	CM58-F	CACGCAGAGATATATACCTTGGC	HG975443.1:1496305-1496425 Solanum pennellii
	CM47-R	GGATAAGAAGCATAGAATTGGTCAG	chromosome ch04 complete genome
63597	CM59-F	TCATCAATGTCATCCAAGTAGGAACAG	HG975450.1:59307011-59307202 Solanum pennellii
	CM60-R	TTGCTTCAGAAAGAGTGTATCTGC	chromosome ch11 complete genome
47097	CM61-F	CATATTTAGGACCACTTTACAG	HG975450.1:34788570-34788761 Solanum pennellii
	CM62-R	GTTGTAGGTATGCTTTTCTTG	chromosome ch11, complete genome
hab-7	*hab-7* _F	CTAGAGAATGTATAAGTCACAGATCC	D17325.1 Solanum peruvianum mRNA for S13-RNase, partial cds
	hab-7 _R	ATTCCCTTCCACTGGTTCC	
hab-8	*hab-8* _F	GATTAAACCACAGCTAATGTCAG	X76065.1 L.peruvianum mRNA for S-RNase S3
	hab-8 _R	GACGACAAGCAATCACC	

Table 2-2. Reproductive molecular characters of selected *S. habrochaites* accessions sorted by latitude from North to South. Unshaded cells represent Ecuadorean accessions and shaded cells represent Peruvian populations. Mating systems are those of TGRC (http://tgrc.ucdavis.edu/) unless otherwise designated. HT and S-RNase proteins were detected via immunoblotting of stylar extracts and *S-RNase* alleles were detected by PCR using allele-specific primers. Y = protein present, N = protein not present, Y/N = individuals segregating for presence of protein.

Accession	Latitude	Longitude	Mating System	HT protein	S-RNase protein	S-RNase Allele(s)
LA4656	-1.0488	-80.0902	SC	Y	N	*LhgSRN-1*
LA1624	-1.3	-80.58333	SC	Y*	N*	*LhgSRN-1* *
PI129157	-1.4	-78.45	SC	Y*	N*	*LhgSRN-1* *
LA1625	-1.5	-80.51667	SC	Y*	N*	*LhgSRN-1* *
LA1266	-2.011	-78.975	SC	Y*	N*	*hab-7*
PI134417	-2.13884	-79.42629	SC	Y	N	*LhgSRN-1*
LA1264	-2.167	-79.1	SC	Y*	N*	*hab-7*
PI390515	-2.17656	-78.81734	SC	Y/N	Y/N	*LhgSRN-1*
LA407	-2.18056	-79.88361	SC	Y*	N*	*LhgSRN-1* *
LA1223	-2.1959	-78.8506	SC	N*	N*	*LhgSRN-1* *
PI251305	-2.23802	-78.90471	SC	N	Y	*hab-7*
LA2144	-2.29	-78.9825	SC	-	-	*hab-7/LhgSRN-1*
LA4654	-3.2091	-79.1916	SC	Y	-	Unknown
LA4655	-3.2322	-79.218	SC	N	N	Unknown
LA2119	-3.62222	-79.23806	SC	Y*	N*	*hab-7*
LA2868	-3.7569	-80.0494	SI	Y*	Y*	Multiple
LA2128	-3.8936	-78.7802	SC	Y	N	*hab-7*
LA1252	-4	-79.21667	SC	Y	N	*hab-7*
LA2855	-4.1	-79.5	SI	Y	Y	Multiple
LA2106	-4.20333	-79.23	SC	Y*	N*	*hab-7*
LA2101	-4.332	-79.5625	SC	Y*	N*	*hab-8*
LA2864	-4.333	-79.783	SI	Y*	Y*	Multiple
LA2860	-4.3333	-79.55	SC	Y	N	Unknown
LA2099	-4.353	-79.802	MP*	Y*	Y*	Multiple
LA2098	-4.36583	-79.81278	MP*	Y*	Y*	Multiple
LA2175	-5.14167	-79.00833	MP*	Y*	Y*	Multiple
LA1391	-6.028	-79.022	MP*	Y*	Y*	Multiple
LA2314	-6.41667	-77.86667	SI	Y*	Y*	Multiple
LA1353	-7.36667	-78.8	SI	Y^	Y*	Multiple^
LA1777	-9.55	-77.66667	SI	Y^	Y*	Multiple^
LA1927	-14.45	-74.81667	SC	Y^	Y^	*hab-6* ^

*Broz et al., 2017

^Covey et al., 2010

Table 2-3. Analysis of heterozygosity using 22 TSS loci. N represents the average number of individuals from each assumed population sampled across all loci. Fractional values of n are due to missing data. He represents the expected heterozygosity and Ho the observed heterozygosity. Allelic richness is the average number of alleles per locus.

Region	n	Number of Alleles	He	Ho	Allelic Richness
Ecuador A	5.95	39	0.21	0.11	1.3
Ecuador B	6.91	50	0.32	0.1	1.42
Ecuador C	2	27	0.09	0.09	1.16
Ecuador D	9.64	32	0.09	0.02	1.1
Ecuador E	3	35	0.22	0.05	1.29
Ecuador F	6.86	58	0.34	0.22	1.56
Peru A	5.77	74	0.53	0.26	1.79
Peru B	6.59	62	0.38	0.2	1.55
Peru C	7.55	64	0.35	0.23	1.54
Peru D	9.86	44	0.2	0.06	1.26

CHAPTER 3:
Reproductive and morphological characterization of Ecuadorean *Solanum habrochaites*

Introduction

The Tomato Genetics Resource Center (TGRC, tgrc.ucdavis.edu) has acquired an excellent collection of tomato and wild tomato relative germplasm that is freely available, and these collections formed the basis of the population study of *S. habrochaites* presented in Chapter 2. However, our results revealed previously unrecognized variability in Ecuadorean populations, and regions of special interest with regard to reproductive traits were under-represented in the TGRC collections. Therefore, Dr. Bedinger undertook a field study in Ecuador as a Fulbright Scholar to study and collect *S. habrochaites* in these regions. By travelling to regions with known populations of *S. habrochaites* and additionally collecting between those regions, we hoped to determine the geographic boundaries of known reproductive syndromes.

In this chapter, I have characterized these newly collected populations of *S. habrochaites* reproductively using controlled pollinations, and morphologically using measurements of flowers, fruit, and leaves. For molecular analysis, collaborators at the University of Missouri screened populations for expression of S-RNase and HT protein using immunoblotting. I also screened these new populations for "selfing" S-RNase alleles using PCR, and these results will be presented in Chapter 4. In Chapter 2, I described the population structure of *S. habrochaites* and found evidence for multiple independent losses of SI at the northern species margin. I propose that as descendent SC populations become isolated from their SI progenitors they become morphologically and reproductively distinct. Populations of Ecuador group A, for example, produce smaller flowers than other groups, and are partially reproductively isolated from some other groups of *S. habrochaites*. Thus, the A group may be in the process of incipient speciation (Broz *et al.*, 2017a). The A group also inhabits a vastly different

environment compared to other populations in the range and are consequently most likely to exhibit divergent traits. By measuring characters associated with reproduction (e.g. perianth size, stigma exsertion, number of flowers per inflorescence, etc.), it may be possible to relatively estimate the degree to which SC populations are reproducing through self-fertilization. Together, these data further support the existence of multiple distinct SC groups of *S. habrochaites* in Ecuador.

Materials and Methods

Before I joined the lab, Drs. Patricia Bedinger, Bruce McClure, Roger Chetelat, and a number of other researchers worked in collaboration with the Universidad Técnica Particular de Loja (UTPL) to gather new collections of wild tomatoes from previously under-collected regions in Ecuador. Borders between geographic regions were of particular interest as, not only were these areas underrepresented in TGRC, the publicly available germplasm bank, but these geographic zones also represent sites of transitions in mating system from SI → SC (Chapter 2). For example, at the time of the collection trip, no accessions were available from near the town of Girón and relatively few were available near the town of Alausí. Dr. Bedinger and her team traveled between locations of known TGRC collections looking for new collections of *S. habrochaites*. At each collection site, flowers were collected and digital calipers were used to measure corolla diameter and stigma exsertion *in situ*. Flowers were photographed and the number of floral buds per inflorescence was tallied. Fruits were harvested and weighed from all available plants. Seeds were harvested by washing away excess fruit pulp, dried on a piece of filter paper, and occasionally the number of seeds per fruit was recorded. Following the trip, seeds were stored in paper coin-envelopes and shipped to Colorado State University (CSU). Soil samples from each site were collected but will be analyzed elsewhere. Herbarium samples from each site were also collected; one set remained in Ecuador and the other set was sent to CSU to be mounted. Eventually, these samples will be deposited in a public herbarium and seeds of new populations will be made available to the public. In total, 32 new populations of *S. habrochaites* were collected (Table 3-1) as

shown on the map in Fig. 3-1. Because other species were also collected but not characterized in this study, collection sites number for the 32 *S. habrochaites* accessions range from 1-45.

Shortly after I joined the lab as an undergraduate research assistant, seeds arrived in Colorado and I inspected and tallied the seeds from each packet. Unfortunately, due to a shortage of envelopes, some envelopes were improvised by hand in Ecuador and small holes in the corners of the envelopes resulted in the spillage of some seeds during shipping. Although we do not believe any loose seeds entered the packets, upon opening the shipping container, many seeds were found scattered freely throughout the seed box. To minimize the possibility of seed contamination, free seeds were discarded along with any seeds found attached to the tape holding the envelopes shut. However, as we have been aware of the possibility of seed contamination due to these defective envelopes since beginning these experiments, we have watched carefully over the years for phenotypic "odd balls" that lack traits (mating system, stem color, floral size and configuration, trichome density, etc.) of other plants from the region. We only found evidence of contamination for a single envelope of seeds. For safekeeping, the contents of each packet were divided. Some seeds from each site were sent to Dr. Roger Chetelat at the TGRC at UC Davis, some were stored in cryo-sealed bags for long-term storage at -20^{o}C at CSU, and the remainder were planted or stored at 4^{o}C.

Seedlings from seeds collected at each site were grown on a light shelf and later transferred to a greenhouse at Colorado State University. In the summers of 2015, 2016, and 2017, cuttings of plants were planted outside in a common garden at CSU (as described in Chapter 2), because many populations of *S. habrochaites* fail to produce flowers in the greenhouse. Tomatoes are readily propagated by cuttings, which allowed us to keep single individuals alive over multiple years. Seed from each collection site was produced using mass sibling crosses in which stigmas were pollinated using pollen pooled from at least three individuals from the same collection site. As described in Chapter 2, mating system of plants from each site was assessed (Table 3-1) and both inter-population and

interspecific crosses were performed to test for loss of pollen and pistil factors (Fig.3-3). Pollen from each plant was collected and stored at -80^{o}C for future experimentation. Typically, one flower opens each day on each of the branched inflorescences in the wild tomatoes (Chalivendra *et al*. 2013). The most recently opened flower is designated as 'Day 0' as it was the flower to open that day while a flower designated 'Day 1' opened the previous day. Proteins were extracted from styles into SDS sample buffer from flowers on the day of anthesis (Day 0), and up to two days after (Days 1 and 2) Dr. Alejandro Tovar-Mendez screened populations for S-RNase and HT proteins by immunoblotting (Table 3-1) at the University of Missouri using the same protocol and antibodies described in Chapter 2. DNA was isolated using a simple DNA extraction method. Briefly, a small piece of leaf tissue is ground in a buffer of 0.4M NaCl, 1% SDS, 25mM EDTA, and 0.2M Tris HCl pH 9.0, centrifuged to remove cellular debris, then precipitated with isopropanol and washed with ethanol. Populations were then screened for S-RNase alleles using PCR (Chapter 4).

Because leaf shape and size differ between species of tomato, we hypothesized the leaf form of different populations of *S. habrochaites* might differ, especially in coastal populations of Ecuador group A. Populations from this region differ significantly in environment and we thought it was possible plants growing in this region had evolved leaf adaptations to allow them to grow in these low-elevation, warm climates. In the July of 2015, Dr. Amanda Broz and I collected young leaves (third leaf down from the apical meristem) and the youngest fully-developed leaf from each plant. All leaves were collected within an hour before 9 a.m. to minimize variance associated with heat and water loss. The field weight was measured and leaves were stored in sliding-zipper storage bags with the petioles submerged in water. After rehydration for approximately four hours, turgid weight was assessed, and leaf area was calculated using a LiCor LI-3100C Area Meter. Leaves were dried in a drying oven and the dry weight of each leaf was noted. Using these measurements, we calculated specific water content (SWC)—the weight of water per cm^2 of leaf. Most crosses and mating system determinations were also performed during the

summer of 2015, in addition to other morphological analyses including bud- and open flower-based measures of inflorescence structure.

Because floral metrics could potentially be affected by environmental variability *in situ*, floral measurements were taken in a common garden. Even in a common garden, environmental variation in the field plot itself as well as the date of floral collection throughout the course of the growing season can affect morphological measurements. Therefore, I designed an experiment using the randomized complete block design in the summer of 2016. Regional groups for each collection site were designated based on previously known reproductive characters, and the population results reported in Chapter 2. University construction of the football stadium forced the use of a never-before used plot of land previously used as a rope climbing course. Furthermore, unusually high-water tables in the new field left some portions unsuitable for plant life. Instead, plants were randomly assigned to either of two smaller field plots such that each plot acted as a complete block with some plants from each region in Ecuador. Within plots, plants were randomly assigned a field position with a corresponding row and column, which were also treated as random effects. Although each genetic individual was not represented in each plot, each collection site was represented in both plots by multiple plants.

Measuring flowers with digital calipers, as was typically done *in situ* in Ecuador can be error-prone. For example, styles in newly opened flowers are often exserted less than 1mm from the anther cone and it is difficult to discern genuine differences in size using digital calipers. To increase the accuracy of measurements, flowers from plants grown in the common garden in the summer of 2016 were first preserved following the protocol set by Spooner and van den Berg (2001), using clear packing tape. The reproductive whorls were removed by snipping them at their base using forceps, the corolla was gently pressed face-down to the tape, the corolla lobes were rolled out to stick to the tape, and the calyx was removed by lightly pinching the base of the flower. All open flowers of three separate inflorescences were preserved for each plant. The result was a single sheet of paper with a row of

flowers from youngest to oldest from each inflorescence that was scanned at high resolution (1200dpi) and measured digitally using ImageJ.

Measurements included petal length (A), inter-petal distance (B), width (C), sepal length (E), anther length (F), and stigma exsertion (G). Corolla area was approximated by calculating the area of a 5-pointed star ($5AB*\sin(36^o)$). Acumen length (D) was also calculated ($A-\sqrt{B^2-C^2/4}$). Following (potato sources), floral shape was assessed using two widely used metrics—relative petal lobe depth (A/B), and relative petal constriction (D/C). Measurements taken are shown in Fig. 3-2.

A mixed model was used to detect significant differences between collection regions while accounting for sources of environmental variation and experimental blocks. Field designation (north or south plot), field position (row and column), flower collection date, and days post anthesis (day 0, 1, etc.) were used as random effects to detect significant ($p<0.05$) differences between geographical regions (modeled as a fixed effect) for each variable. Statistically significant groups are denoted with lowercase letters. Shared letters indicate non-significant relationships. Generalized linear models and were similarly used to detect significant differences between regions of collection sites for the other morphological observations (both *in situ* and common garden).

Results and Discussion

Site information

Locations of collection sites are shown in Fig. 3-1 with a brief description of each site in Table 3-1. Broadly, plants were collected from six major groups as identified in Chapter 2 and are generally color-coded according to the STRUCTURE plot (Fig.2-5). Groups like Ecuador B and E which were genetically similar to another group (or groups in the case of group B) but reproductively distinct are intermediately colored (orange + green = brown, blue + green = aqua). Collections sites encompass a broad range of altitudes from 171 masl to nearly 2500 masl.

Mating system determination, inter-population crosses and interspecific crosses

In the summers of 2015-2017 we performed controlled crosses to test for mating system, and loss of pollen and pistil factors involved in pollen-tube rejection (Fig.3-3). Results are generally consistent with previous findings (Markova *et al.*, 2016; Broz *et al.* 2017a, Chapter 2). Testing for loss of pollen resistance factors (Fig. 3-3A) revealed that pollen of plants from sites 31 and 33 (Ecuador SC-2 group A) is rejected by styles of SI LA1777, demonstrating the predicted unidirectional reproductive isolation of the SC-2 mating system group from SI populations. Surprisingly, LA1777 also rejects pollen of one population from Ecuador group C at site 20. Populations from this region have previously been reproductively uncharacterized but these results contrast with the findings of LA4654/LA4655 from the same region, populations that, as we suggested in Chapter 2, likely represent a fourth independent SI→SC mating system transition in Ecuador. The rejection of pollen from these sites by styles of LA1777 suggest there has been a loss of pollen resistance factors in at least some plants in these populations (Markova *et al.*, 2016). Our results of testing for the presence of pistil barriers (Fig. 3-3B) are consistent with previous results (Broz *et al.*, 2017a), in that only S-RNase expressing populations are generally capable of rejecting pollen of LA0407 (Table 3-1, Fig.3-3). In addition to SI accessions, some plants from near the town of Alausí (Ecuador group B) express S-RNase (Table 3-1), including site 26 (note that PI251305 was collected in the same region and also expresses S-RNase protein (Table 2-1), which also sometimes rejected LA0407 pollen. The ability to reject interspecific pollen can be influenced by HT protein (Tovar-Méndez *et al.*, 2014; Tovar-Méndez *et al.*, 2017). Previously the only known *S. habrochaites* population with a known mutational loss of HT-protein in this region was LA1223, whose styles do not reject interspecific pollen tubes (Broz *et al.*, 2017a). Our intensive sampling in this region revealed additional populations not expressing HT protein (all plants at site 25 and 26, and some plants at sites 27 and 42) which generally correlated with the loss of interspecific reproductive barriers (Fig. 3-3C, Table 3-1). However, plants from sites 26 and 42 that did not express HT were still able to reject

interspecific pollen. This is consistent with our results for accession PI251305 described in Chapter Two (Table 2-2) and research performed by undergraduate Nicole Irace for her Honor's thesis in the Bedinger Lab (unpublished), both of which implicate a factor in addition to HT that is involved in interspecific pollen tube rejection.

Perianth measurements

Floral displays influence outcrossing rates, as pollinators are more attracted to plants with larger and more numerous flowers than those possessing smaller floral displays (Ohashi and Yahara, 2001; Barrett 2002; Goodwillie *et al.*, 2009). Rick *et al.* (1979) noticed flowers of SC northern *S. habrochaites* were smaller and less vibrant, with less prominent stigma exsertion past the anthers. Later coined the "selfing syndrome," this phenomenon is an adaptive trait in selfing lineages, having occurred in parallel with SI→SC transitions (Sicard and Lenhard, 2011; Karron *et al*., 2012). Over time, the grandeur of floral displays of SC populations might therefore be predicted to decline in comparison to obligately outcrossing SI populations. However, we found that this was not the case for SI and four of the five different SC *S. habrochaites* groups (Fig. 3-4). Notably, flowers of SC Ecuador group C (Girón region) were even significantly larger in terms of corolla area and diameter *in situ* than those of the SI Ecuador group F (Sozoranga region), although no significant difference could be detected between this and the SI Ecuador F group when grown in a common garden. As previously noted but not quantified (Rick *et al.*, 1979) flowers from the northernmost margin of *S. habrochaites* along the coast (Ecuador group A) have significantly smaller flowers than other groups (Fig. 3-4). These flowers are significantly (~50%) smaller in area, and in diameter when grown both *in situ* and in a common garden compared to the populations with the largest flowers. Furthermore, populations from this region possess a markedly different shape. Corolla lobes of Ecuador A are relatively much deeper, and more narrowly constricted (Fig. 3-4), giving the corollas the appearance of a long-armed, five-pointed star—almost resembling a sea star in shape. In contrast, corollas of the other populations are broader, with wider petals and shallower lobes, more

closely resembling typical Solanaceous flowers. Although corolla size and shape differed between groups, sepal length remained relatively constant.

Stigma exsertion

The extent to which the style extends past the tip of the anther cone may act as an indicator of outcrossing, since this trait determines whether stigmas are exposed to pollen carried by pollinators; i.e. strongly exserted styles are thought to be more likely to engage in outcrossing. In contrast, stigmas that fail to exsert are much more likely to be self-pollinated as self-pollen, but not pollinators, can readily access them. Therefore, we would predict that SI populations would have more strongly exserted stigmas than SC populations of *S. habrochaites*. This outcrossing-stigma exsertion correlation has been demonstrated in another wild tomato species, *S. pimpinellifolium* (Rick *et al*., 1978). However, we again found that this prediction does not hold in *S. habrochaites*.

Because stigma exsertion may change over time as styles elongate or as the anther cone begins desiccating, I have treated the analysis of stigma exsertion using two methods for both common garden- and *in situ*- grown *S. habrochaites*. To detect broad patterns of stigma exsertion, the time since bud break was treated as a random effect in a mixed model, allowing for variance associated with age to be accounted for. This analysis is represented by the 'All days' plots (Fig. 3-5). Notably, the highest degree of stigma exsertion is not found in SI F populations as theory would predict, but rather in different SC populations. *In situ*, stigmas of three of the five SC populations were significantly more exserted than those of the SI populations and in a common garden, stigma exsertion is greater in one SC group than in the SI group while three other SC groups have equivalent exsertion.

When stigma exsertion is analyzed by floral age, variability between groups is observed but is not correlated with mating system, with the exception of the small, short-lived flowers of the northern-most SC populations (Ecuador group A) which have significantly less exserted styles, especially at bud

break in which they almost entirely lack exsertion. In the other SC groups, stigmas are exserted to a similar extent as is seen in SI group E, and stigmas are especially highly exserted in older flowers of SC populations near the towns of Girón (Ecuador group C) and Loja (Ecuador group D). The inverse relationship between anther length and stigma exsertion over time suggests some of the new exsertion in older flowers is due to stamen desiccation. However, the growth of protruding style exceeds the loss of anther length (Fig. 3-6). As the size of the ovary was not measured, it is unknown if further exertion of the stigma could be caused by development of the ovary.

Overall, neither corolla size nor stigma exsertion are greatly increased in SI populations compared to most SC populations, suggesting that these SC populations have either evolved recently or that pollen transfer, promoted by these floral structures, is ongoing regardless of mating system. While large floral displays may attract more pollinators, potentially increasing rates of outcrossing, more pollinator visits simultaneously increase geitonogamous (between flowers of the same individual) self-pollination (Harder and Barrett, 1995). Consequently, larger flowers could facilitate and even promote self-fertilization. Clearly more study is required to disentangle the effects of floral morphology on mating system, for example, one in which pollinator visitation (including geitonogamous visitations) is documented in SI and both large- and small- flowered SC populations.

Inflorescence structure

It is thought plants can increase their floral display (and thus be more attractive to pollinators) by either increasing the size of their flowers or by increasing the density of flowers per inflorescence (Harder and Barret, 1995; Ohashi and Yahara, 2001). Again, outcrossing populations dependent on pollinators are predicted to produce more densely flowered inflorescences than selfing populations that might lack the selective pressures to maintain “fancy,” pollinator-attracting, floral displays. To assess inflorescence structure, the number of buds on an inflorescence was assessed both *in situ* and in a

common garden (Fig. 3-7). In both environments, SC populations near the town of Girón (Ecuador group C) produced more buds per inflorescence than other SC populations but SI populations near Sozoranga (Ecuador group F) possessed the highest inflorescence bud density of all studied populations. Because bud density might not reflect density of open flowers, we also examined the number of open flowers on inflorescences in each population. SI Ecuador group F populations maintain on average one more open flower per inflorescence than the SC Ecuadorean *S. habrochaites*. At opposite ends of the range, inflorescences of Ecuador group A produce one-two open flowers per inflorescence, whereas inflorescences of SI group F exhibited at least 3 open flowers per inflorescence. Therefore, in this case, the predicted correlation of mating system and floral display is upheld.

Fruit Measurements

Ms. Tania Riofrio, a collaborator at UTPL, weighed and tabulated seed for some fruit collected *in situ* at the various collection sites (Fig. 3-8). Fruit from the Alausí area (group B) and the SI populations near Sozoranga (Ecuador group F) were significantly smaller than those of the other SC groups. The largest fruit collected were from plants of the Loja Corridor (group D), followed by the nearby Cariamanga (group E) and Girón (group C) collections. Seed count per fruit generally correlated with fruit weight, suggesting the increase in the size of the fruit may be due to the increased fecundity of these populations. However, fruit weight significantly varied between groups C and D by nearly 50%, despite an insignificant difference in seed count. Because fruit weight and seed set can be impacted by environmental variation, a future common garden experiment should reveal whether these differences are attributable to genetics, pollination rates, or nutrient abundance.

Leaf Measurements

Despite major differences in floral morphology and reproductive characters in the newly collected populations, we generally did not detect significant differences in leaf weight or area in this

study (Fig. 3-9). Both young and mature leaves from southern SC populations (groups D and C for each respectively) weigh more than leaves of northern SC populations. However, no significant leaf area differences were detected. Leaves of high-elevation, southern populations (groups C, E, and F) can hold more water per square centimeter than coastal northern populations (group A) as reflected by the specific water content (SWC). While this result may be attributable to differences in ecology and environment between sea-level and alpine Ecuador, it is unclear why coastal populations would hold less water. Relative water content (RWC), a common measure of cellular water deficit, was also calculated but not found to be significantly different between groups. As leaf measurements were not repeated in additional years or *in situ*, any differences in moisture content could also be attributable to relative shading of each plant, the position in the field, or some extraneous factor.

Conclusions

We previously described the population structure of Northern populations of *S. habrochaites* (Chapter 2), data that supports the idea that several independent SI → SC transitions have taken place at the northern species margin. The data presented in this chapter further supports the formation of multiple SC groups.

Using floral measurements, we tested the selfing syndrome hypothesis—that SC populations will possess markedly smaller flowers with unexserted stigmas as self-pollinating populations may face relaxed selection to maintain large, pollinator-attracting displays. In Ecuador group A, we found evidence of the selfing syndrome. These populations possess significantly smaller stigmas, and stigmas are almost entirely unexserted the time of anthesis However, in the other SC populations, we failed to detect the predicted smaller floral displays, suggesting either that selection has not had enough time to reduce the size of the floral displays, or that selection instead favors continued outcrossing. It should be noted, however, that SI populations may still possess increased floral displays compared to SC

populations as SI populations exhibit more densely packed inflorescences than those found in any of the SC groups. Future studies could holistically measure floral display size by accounting for other factors that contribute to attracting pollinators. In the case of group A, these changes in floral morphology may indicate this transition from SI→SC is one of the oldest, as other SC groups have generally not become morphologically distinguished from the SI group. Given this morphological differentiation and the unidirectional reproductive isolation between SC group A and SI *S. habrochaites* populations, as demonstrated previously (Broz *et al.*, 2017a), and in this thesis, the coastal group A populations may someday be described as a new variety of *S. habrochaites* or even as a separate species.

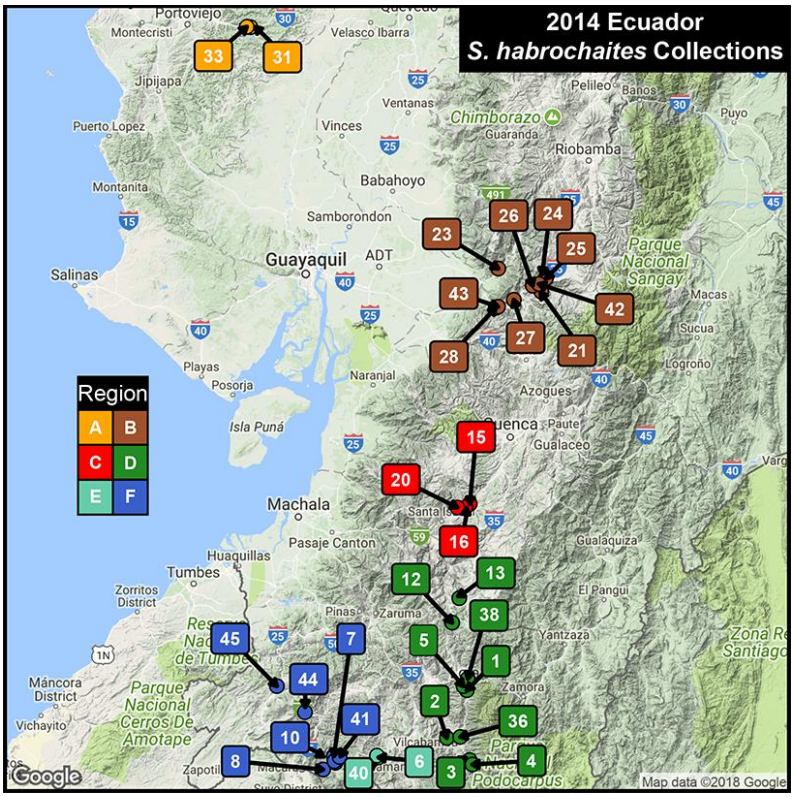

Fig. 3-1. **Map of 2014 Ecuador Collections of *S. habrochaites***. Each point represents a newly collected population. Colors generally reflect the Ecuadorean groups identified in the STUCTURE analysis (Fig. 2-5) with the SC-X group shown in red and the SC-Y group shown in aqua.

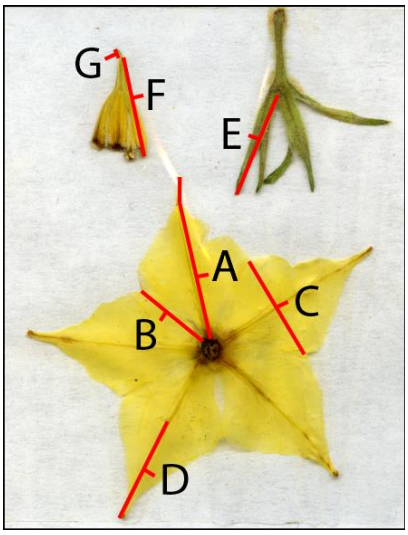

Fig. 3-2. Common garden digital floral measurements. A) Petal length along vein, B) distance to inter-petal constriction point, C) petal width across constriction points, D) calculated acumen length, E) sepal length, F) length of anther cone, filament excluded G) stigma exsertion –the length of style protruding from the anther cone. Images were scanned at 1200dpi and measured at 100% zoom, except G, which was measured at 150%. Measurements were taken using the segmented line tool in ImageJ.

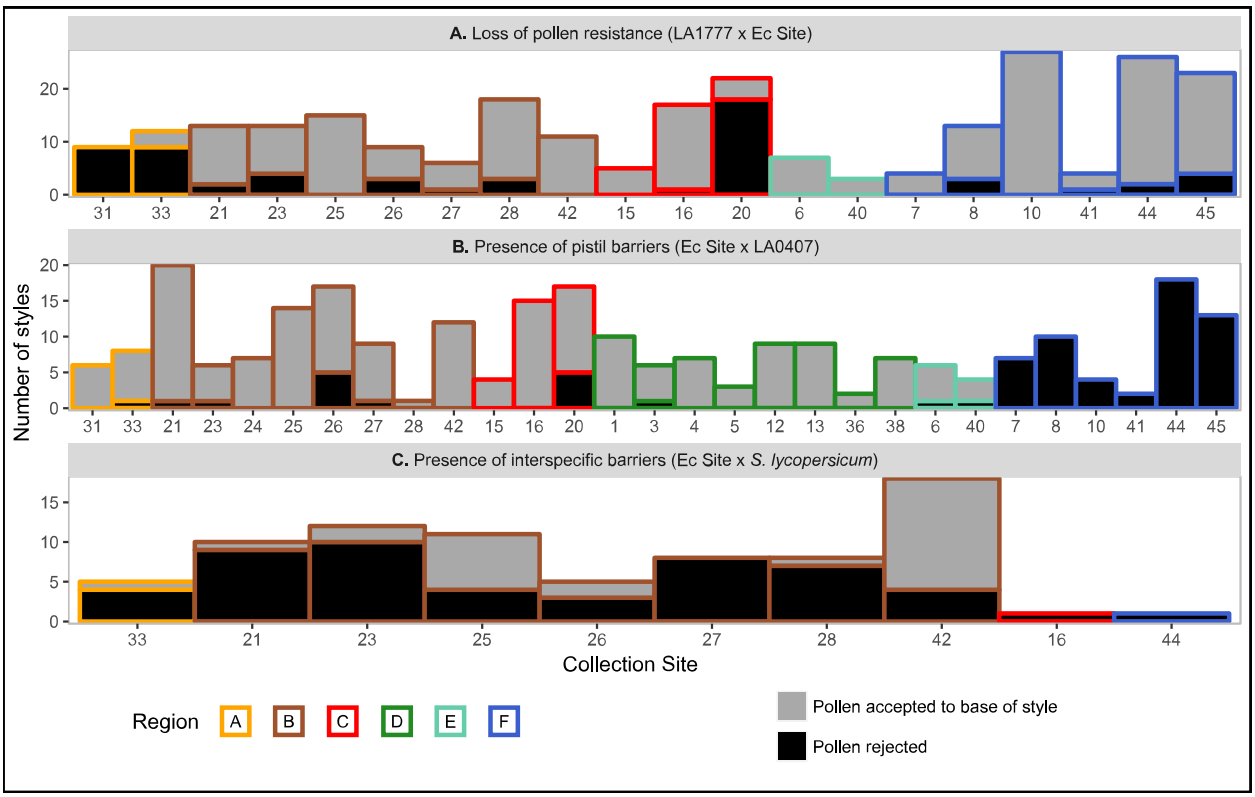

Fig. 3-3. **Pollen-tube growth of inter-population and inter-species test crosses for newly collected Ecuadorean *S. habrochaites***. Bars represent the collective number of crosses for the corresponding site using plants grown in a common garden at CSU. Bar outlines represent regions from Fig. 3-1. A) LA1777 is used as female to test for pollen-side mutations B) LA0407 is used as male to test for losses of pistil factors and C) Pollen of *S. lycopersicum* is used to test for interspecific barriers.

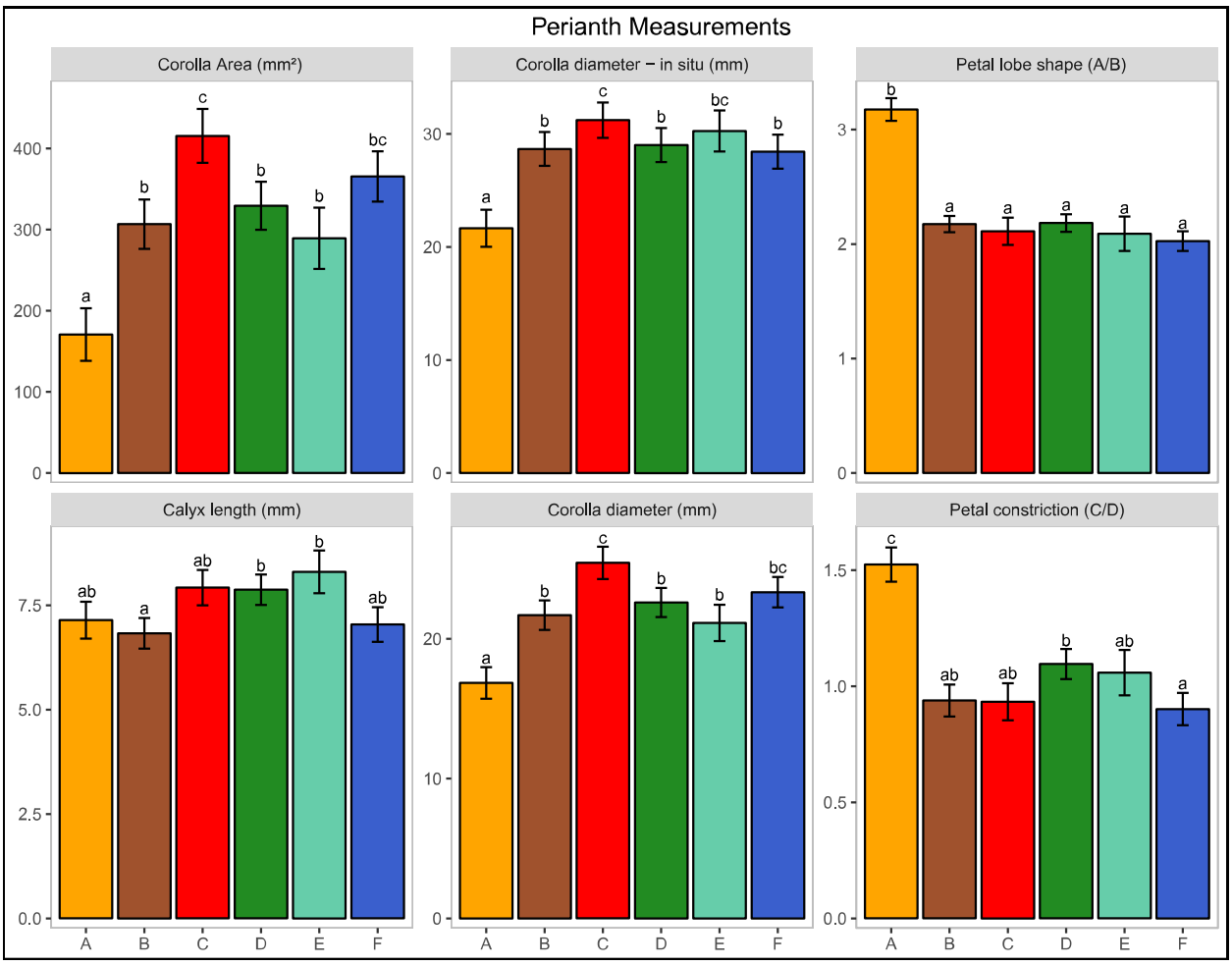

Fig. 3-4. Digital measurements of perianth whorls. Bars represent the lsmeans of each floral measurement (Fig. 3-2) for reproductive groups using the colors in Fig. 3-1 (Fig. 3-1, Table 3-1) predicted by the mixed linear model. Error bars represent standard error of the mean. Lowercase letters represent significant ($p < 0.05$) differences between groups. Shared letters indicate no significant difference. All measurements were performed on flowers from plants grown in a common garden except where noted.

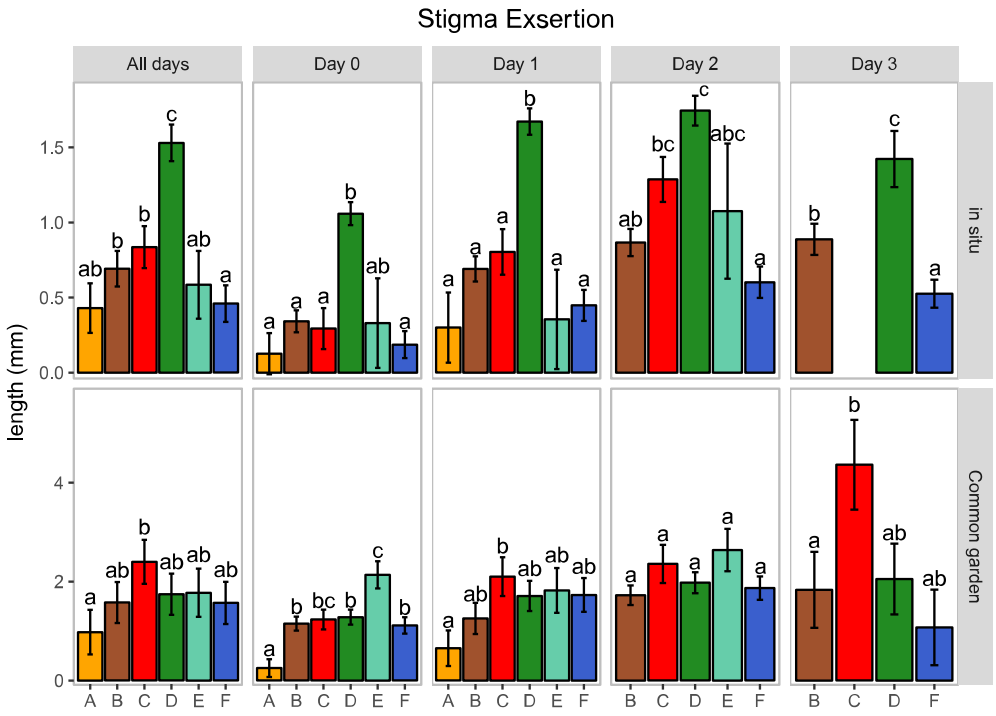

Fig. 3-5. Stigma exsertion of collections sites grown *in situ* and in a common garden. Bars represent the lsmeans and error bars the standard error of the mean. Shared lowercase letters represent no significant ($p<0.05$) difference between groups. Colors representing regions are the same as in Fig. 3-1.

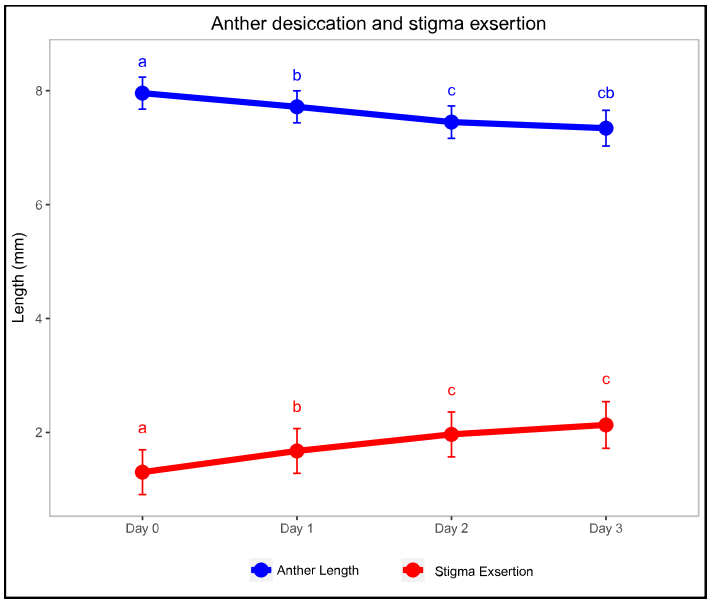

Fig. 3-6. Anther desiccation and stigma exsertion over time in common garden-grown flowers. Each point represents the lsmean of the flowers the day of opening (day 0), one day after opening (day 1), etc. for all measured populations. Error bars represent standard error of the mean. Shared letters are not significantly ($p<0.05$) different from one another.

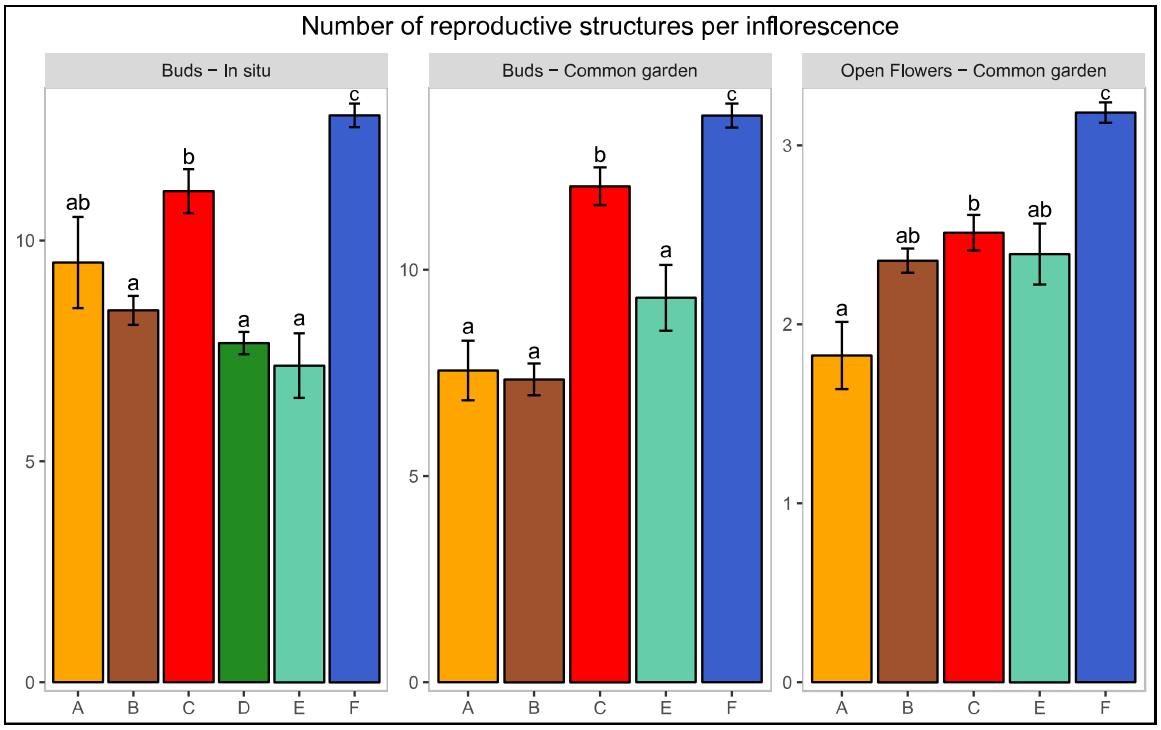

Fig. 3-7. Inflorescence structure of Ecuador collections by region. The total number of buds or open flowers per inflorescence were separately tabulated. Bars represent mean number of structures, error bars are standard error of the mean, and lowercase letters represent significant differences according to the linear model.

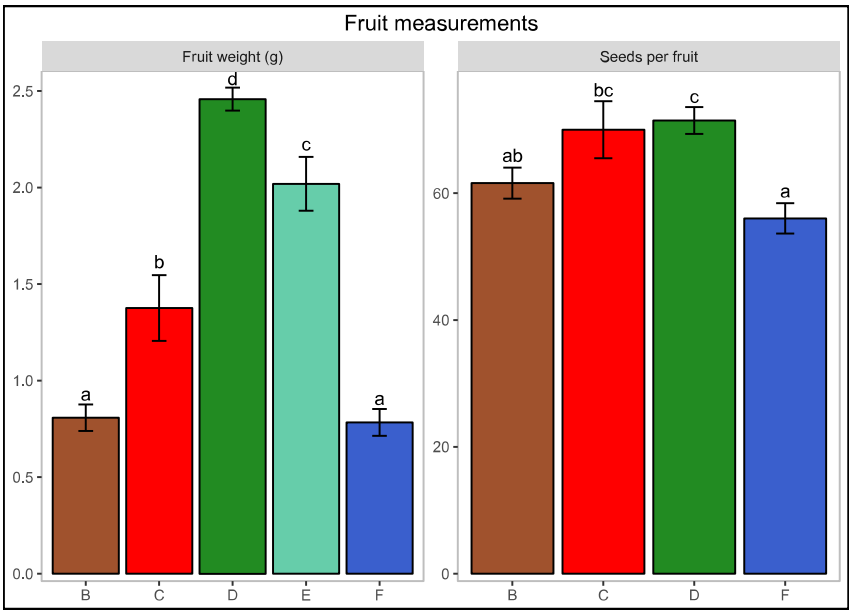

Fig. 3-8. Mean fruit weight and number of seeds per fruit. Fruit was collected *in situ* from collection sites and averaged across regions. Shared lowercase letters represent non-significant differences between groups.

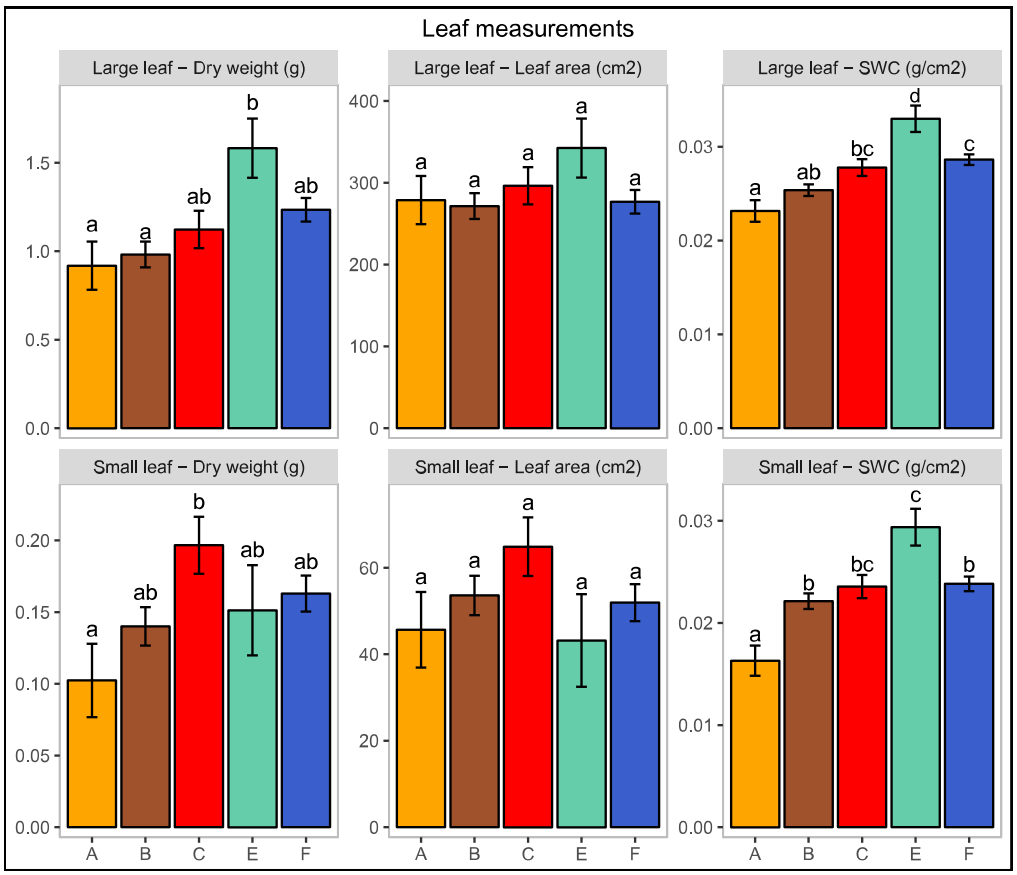

Fig. 3-9. Leaves measured in a common garden. Large leaf refers to the most recent, fully mature leaf and small leaf to the third leaf from the apical meristem. Dry weight was measured after leaves were dried in a drying oven for 48 hours. Specific water content (SWC) was calculated using the area and turgid and dry weights. Bars represent mean measurements for each region, error bars are standard error of the mean, and shared lowercase letters are non-significant differences ($p<.05$) between groups.

Table 3-1. Site details and reproductive characters of newly collected S. habrochaites collections. Sites are sorted generally from North to South but retaining regional identity established in Fig. 3-1. Mating system was determined using controlled crosses. S-RNase and HT protein were detected using antibody-specific immunoblotting of stylar-extracted proteins at the University of Missouri. Ratios represent the number of individuals from each collection site in which the protein was detected. The 'selfing' S-RNase alleles LhgSRN-1, hab-7, and hab-8 were detected using PCR of genomic DNA

Region	Site #	Nearest Town/Site	# Plants	Information about the site	Altitude (m)	Latitude	Longitude	Mating System	HT protein	S-RNase protein	LhgSRN-1	hab-7
A	31	San Placido - El Cruce 8120	3	Via San Placido - El Cruce. Near San Placido	234	-1.0728	-80.1692	SC		0/3	Y	N
A	33	Manchagrandi 8122	few	Via San Placido - El Cruce	171	-1.0714	-80.1850	SC		0/2	Y	N
B	28	19 km after Huigra before El Triunfo	2	Along road to Guayaquil 19 km past Huigra toward El Triunfo, on right, near small river	1246	-2.3294	-79.0689	SC	3/3	1/3	Y	Y/N
B	43	17.4 km from Huigra	~5	17.4 Km past Huigra going toward El Triufo, on right, may be continuous with site 28	1392	-2.3241	-79.0585	SC			N	Y
B	27	Huigra/Footbridge, railroad tracks	4	Huigra near foot bridge over Rio Chanchan south/west end of town at railroad tracks marker 600. One plant near river, other plants on cliff near railroad tracks; second visit one large plant almost to railroad tracks along dirt road	1320	-2.2932	-78.9877	SC*	1/2	1/2	Y/N	Y/N
B	21	4.3 km S of Zunag	27	4.3 Km south of Zunag, 2.3 Km from Alausi large population on right road side going toward Zunag; at later visit, site is filled with blooming lupines	2367	-2.2567	-78.8647	SC	4/4	0/4	Y	N
B	26	Sibambe	~15	Large pop along both sides of road 1.3 Km from Alausi on road to Huigra	<2374	-2.2303	-78.9031	SC*	0/4	3/4	Y/N	Y
B	42	Alausi/across rio, 8111	~10	5 Km past Alausi bridge on right going toward Huigra	2451, 2326	-2.2222	-78.8756	SC	1/3	0/3	Y	N
B	24	South Alausi	2	In Alausi, south part of town, by dirt road out of neighborhood on edge of corn field	2140	-2.1990	-78.8469	SC			N	N
B	25	North Alausi	~20	Up a foot trail, in north part of Alausi along Calle Colombia, at the entrance of the old town. Near home of Sr. Leoncio Yamasca. ~8 smaller plant near bottom of trail, more ~5 min walk up	2365	-2.1974	-78.8438	SC*	0/4	0/4	Y	N
B	23	Bucay/Chaguayaco	~20	Near Bucay at 13 Km toward Riobamba, thick vegetation, rain-forest-like, many insects	643	-2.1536	-79.0572	SC	2/2	1/2	N	N
C	20	Asunción	6 - 7	At Balneario - El Paraiso turn right onto dirt road, go straight ahead up to derrumbe, plants 1 - 2 Km toward Asunción; second visit one large plant near chain-link fence	1817	-3.2222	-79.2408	SC		0/5	N	N
C	16	Past Caledoneas	~10	Large site in road cut, plants on both sides of road, damaged by small landslides on second visit	1749	-3.2215	-79.2039	SC		0/3	N	N
C	15	El Pongo	2	5Km below Giron	1981	-3.2054	-79.1851	SC		0/1	N	
D	4	South of Yangana	6 - 8	1.2 km south of Yangana going uphill on right side	1924	-4.3704	-79.1735	SC			N	Y
D	3	North of Yangana	6	17 km south of Vilcabamba, 3.3 km north of Yangana on both sides of road on pass before town	1938	-4.3507	-79.1861	SC			N	Y
D	2	Vilcabamba/Santoru m	~10	In bushy area near home of Sr. Lorenzo Tacuri off of via A Quinara between Vilcabamaba and Quinara, 4 km from the Vilcamba-Yangana road	1410	-4.2517	-79.2880	NA				
D	36	Vilcabamba	4	Just past Madre Tierra on left one plant small flowers got 3 fruit then found more plants around 2-3 Km up the road on right side; second visit most plants destroyed by "Tigers"	1535	-4.2475	-79.2281	SC			N	Y
D	1	Loja/Botanical garden	5	3 plants in Jardin Botanica one along fence before the garden, and one plant across road, this sit is probably LA1252	2017	-4.0388	-79.1986	SC			N	Y
D	5	Loja/ Daniel Alvarez Lagoon	5	Above Cdla. Daniel Alvarez, on Calle Olivos off of Benjamin Carrion, vacant lot where horses are grazed	2198	-4.0205	-79.2133	SC			N	Y
D	38	Loja/Residencia de canciler	NA	On steep bank near where new greenhouses will be built	2021	-3.9871	-79.1944	SC			N	Y
D	12	San Lucas/school	2	46 km from Loja to San Lucas, in the school garden of school Buen Vivir, Sr. Victor Duchaicela	2439	-3.7368	-79.2640	SC			N	Y
D	13	Saraguro/ old Cuenca road	~20	Just past Saraguro on old road to Cuenca, in and near a cement block-producing operation, some plants in adjacent gulley, more on road bank going back towards town across from Samana Wasi hostel	2462	-3.6275	-79.2324	SC			N	Y
E	6	Past (south of) Cariamanga	1	Past (south of) Cariamanga along road side on right side going south before of San Pedro	1832	-4.3385	-79.5946	SC		0/2	N	N
E	40	Cariamanga/San Roque	1	7.2 Km past Cariamanga, 0.5 km past San Pedro turnoff	1786	-4.3398	-79.5944	SI^		2/2	N	Y
F	8	2.2 km Below (south of) Sabiango	~10	2.2 Km south of Sabiango, plants right near landslide, on both sides of road	572	-4.3930	-79.8425	SI		4/4	N	
F	10	Sozoranga-Sabiango Penjamo	~10	Big site between Sabiango and Sozoranga. 7 Km N from Sabiango, 6 km S from Sozoranga	1325	-4.3558	-79.7921	SI		7/7	Y/N	N
F	41	N of Sozoranga/Lopez garden	5	2 Km N of Sozoranga, 44 Km from Carimanga	1684 - 1709	-4.3367	-79.7690	SI^		1/1	N	
F	7	After (just south of) Sozoranga	4	1.2 km south of Sozoranga on road toward Sabiango, plants by stream near quebrada on right side of road	1564	-4.3331	-79.7947	SI		3/3	Y/N	N
F	44	Mollinomuna	>10	7 Km east of Celica, plants in pueblo near church and in garden, very foggy in afternoons	1545	-4.1370	-79.9245	SI		3/4	N	
F	45	Alamor	~10	On both sides of road below Alamor past roundabout going toward Puyanga, vetetation thick	1007	-4.0203	-80.0502	SI^		4/4	N	

CHAPTER 4:

Mutations in S-RNase genes lead to SI → SC mating system transitions

in Northern *Solanum habrochaites*

Introduction

In the tomato clade, at least nine SI→SC mating systems have occurred (Igic *et al.*, 2008), and the architecture of the gametophytic SI system in the Solanaceae suggests female factors are typically lost first (Bedinger *et al.* 2017), although there may be one case of pollen-side mutational loss leading to SC (Markova *et al.*, 2017). "Selfing" S-RNase alleles have been correlated with self-compatible mating systems both in entire species (Kondo et al., 2002) and in SC populations of mostly SI species (Kondo et al., 2002; Royo et al., 1994; Kowyama et al., 1994; Li and Chetelat, 2015). Previously, mating system transitions from SI→SC in *S. habrochaites* at its northern and southern species margins were found to result from independent mutations, because SI was recovered at a low frequency in F2 plants when the two SC lineages were crossed (Rick and Chetelat, 1991). Candidate "selfing" S-RNase alleles have been found that are associated with each of these two lineages. Covey *et al.* (2010) found an S-RNase allele (*hab-6*) that correlated with SI → SC mating system transition at the southern species margin. This allele is expressed but the protein is a low-activity S-RNase that may function in interspecific pollen-tube rejection but is unable to prevent self-fertilization (Broz, unpublished).

Kondo *et al.* (2002) described the *LhgSRN-1* allele which is correlated with SC in some Northern populations of *S. habrochaites* (Broz *et al.,* 2017a). Our results suggest not one, but as many as four mating system transitions may have independently arisen at the Northern species margin of *S. habrochaites* (Chapter Two). In this chapter, I present three different "selfing" S-RNase alleles (including two new alleles that I identified) that have likely lead to the independent losses of SI in Ecuador, and the potential functional progenitor alleles for these "selfing" S-RNase alleles. Furthermore, the geographical

distribution of these alleles further supports my hypothesis that Ecuador group B formed from the "hybridization" of two existing SC groups.

Materials and Methods

Isolation of hab-7 from SC-1 plants

The stylar transcriptome of SC accession LA2119 (Broz *et al.*, 2017b) was used to identify an S-RNase allele associated with some SC accessions from central Ecuador, including LA2119. The transcriptome was assembled *de novo* using Trinity (Grabherr *et al.*, 2011). Using a set of known S-RNase sequences as a BLAST query, the resulting *de novo* assembly was used as a BLAST database to discover potential *S-RNase* alleles. I recovered the entire coding region and entire genomic sequence of a potential *S-RNase* allele. Following the convention set by Covey *et al.* (2010), we have dubbed this S-RNase allele *hab-7*.

Isolation of *hab-8* from SC-Y plants

We used a PCR-based strategy using degenerate primers as devised by Kondo *et al.* (2002) for the isolation of S-RNase alleles. Briefly, we amplified unknown S-RNase sequences from the genomic DNA of LA2101 using degenerate primers (Covey *et al.*, 2010) and appropriately sized products were cleaned up (Qiagen) and ligated to pJET1.2. Colony PCR was performed, and the resulting PCR products were cleaned up (Zymo) and sequenced (Genewiz). Following sequencing, allele-specific primers were designed and synthesized (Table 2-1) and used to amplify sequences from gDNA of two individuals of LA2101 for direct sequencing.

Results and Discussion

LhgSRN-1

Kondo *et al.* (2002) previously reported the *LhgSRN-1* S-RNase allele is associated with SC in some populations of northern *S. habrochaites*. While it appeared that *LhgSRN-1* could potentially function as an S-RNase if expressed, it is believed the lack of S-RNase protein expression in these northern populations is the result of a small transposable element (a MITE family member consisting almost entirely of inverted repeat sequences) inserted into the promoter region of *LhgSRN-1* that prevented transcription (Covey *et al.* 2010). This MITE might directly interfere with the transcription apparatus or epigenetically via transposon silencing, which can affect transcription of nearby genes (Sigman and Slotkin, 2016). Future experiments such as an analysis of methylation patterns in SI and SC populations would be required to determine exactly how expression of this potentially functional S-RNase is downregulated.

The *LhgSRN-1* allele in the SC-2 mating system group is closely related to S-RNase sequences found not only in SC-2 mating system group accessions, but also in some SI accessions of *S. habrochaites*, including LA1391, LA2099, LA2868, Ecuador site 7 and Ecuador site 10. Alignments of the translated sequences of the encoded S-RNase proteins are presented in Fig. 4-1. We detected five amino acid variants involving only six amino acid sites within the proteins. Of these, two variants were detected only in SI populations from Ecuador and Peru. However, the other protein-level variants were found in both SI and SC populations (Fig.4-2). Potentially additional mutations in the nonfunctional regions of these related S-RNases might explain the functional differences between *LhgSRN-1* and its related alleles in SC and SI populations.

Surprisingly, Dr. Amanda Broz and I discovered that the MITE associated with down-regulation of the *LhgSRN-1* allele was present along with coding regions in SI plants of LA1391, LA2868, Ecuador 7

and Ecuador 10. Thus, these results suggest the *LhgSRN-1* "selfing allele", with its associated TE, likely has existed in the standing genetic variation of *S. habrochaites* for a long time, predating the migration of the species into Ecuador and the establishment of SC populations. We would expect that the functional loss of one S-RNase allele in an SI individual should result in half the self-pollen being accepted, given the mechanism of gametophytic SI in the Solanaceae. However, we found that individuals in Ecuador group F and Peruvian LA1391 that contain the *LhgSRN-1* allele, including the upstream MITE, are generally SI, because these individuals do not set self-fruit, and self-pollen tubes fail to reach the ovary (Fig. 4-3, Table 2-2, Table 3-1), with the exception of a single LA1391 individual in which self-pollen tubes occasionally reached the base of the style (Fig. 4-3). Lack of transposon silencing might explain why SI populations in Ecuador group F and northern Peru are not partially SC. As populations spread north through the AHD, the MITE may have become silenced, and selection for reproductive assurance would lead to rapid fixation of the silenced *LhgSRN-1* allele in migrating populations.

Of course, an alternative hypothesis would be that this allele may not be part of the S-locus. Placing *LhgSRN-1* at the S-locus would alleviate concerns that this is not a true S-RNase allele. We are currently conducting experiments with known S-RNase alleles combined with *LhgSRN-1* to determine if *LhgSRN-1* (and the other selfing alleles described in this thesis) segregate as would be expected for *bona-fide* S-locus encoded S-RNases. Another hypothesis would be that these populations became SC not due to the loss of female S-RNases typical of the other Solanaceae but rather by additional, unreported pollen-side mutations. This latter scenario may be less likely as their pollen functions on SI LA1777 tester pistils (Fig.3-3).

The hab-7 S-RNase allele

The *hab-7* allele that we found in SC-1 plants retains motifs unique to known S-RNases, including conserved sequences C1-C5 (Fig.4-3), while lacking those found in non-S locus S-like RNases (Vieira *et al.*, 2008), as well as the predicted single intron. This allele shares >99% amino acid identity with partial coding regions of *S. peruvianum* S13-RNase (Chung *et al.*, 1994; GenBank: BAA04147.1) and a *S. chilense* S-RNase (Zhao 2011 NCBI direct submission; GenBank: AEN95116.1). Compared to the two different *S-RNase* transcripts found in styles of an SI *S. habrochaites* accession LA1777 individual (FPKM = from 4604.9 and 14579.88, n = 1), *hab-7* transcripts in styles of the SC accession LA2119 are about 400- to 1200-fold less abundant (FPKM = 12.25, n = 1). Future studies will be conducted to precisely quantify expression of *hab-7* to determine if populations possessing the *hab-7* allele are SC because this seemingly functional S-RNase allele is not expressed at a high enough level to function in SI.

The hab-8 S-RNase allele

We recovered one potential S-RNase allele from SC-Y accession LA2101 from Ecuador region E that is 98% identical to both a *S. habrochaites* S19 pseudogene (365/371bp) and the functional *S. peruvianum* S3 S-RNase (364/371 nucleotides) (Fig. 4-4).

We believe *S. peruvianum* S3 (Royo *et al.*, 1994; X76065.1) represents the functional ancestral allele to both the *S. habrochaites* S19 (Zhao 2011 direct submission; JQ040816.1) and *hab-8* pseudogenes. When aligning these sequences, we find that the *hab-8* allele is pseudogenized by a non-synonymous base-pair substitution nonsense mutation (TGG/Trp → TAG/STOP) resulting in a premature stop codon. In contrast, the pseudogenization of S19 is the result of an insertion of a single base pair that results in a frame-shift with an eventual different premature stop codon.

Because of the high degree of sequence conservation with other known S-RNase alleles, we believe the *hab*-8 allele represents a genuine S-RNase selfing allele causing the SI→SC mating system

transition in populations near the town of Cariamanga (Ecuador group E, SC-Y). In contrast with the down-regulation of gene expression of potentially functional alleles seen with *LhgSRN-1* and *hab-7*, the *hab-8* mutation has produced a *bone fide* loss-of-function allele and would result in the loss of pistil barriers to self-pollen.

Distribution of alleles suggests a "hybrid" origin to Ecuador group B

Using primers specific to the *LhgSRN-1, hab-7* and *hab-8* sequences (Kondo *et al.*, 2002; Broz *et al.*, 2017a; Table 2-1), we screened publicly available and newly collected accessions of *S. habrochaites* by PCR and gel electrophoresis for "selfing" alleles (Table 2-2, Table 3-1). The mixed distribution of *hab-7* and *LhgSRN-1* in populations near the town of Alausí (group B) support our hypothesis of a 'hybrid' origin of group B as populations throughout this region possess either or both of these two alleles that otherwise are apparently separately at fixation in the two proposed parent populations (*LhgSRN-1* in group A/SC-2, *hab-7* in group D/SC-1).

Both the *LhgSRN-1* and *hab-7* selfing alleles could potentially encode a functional S-RNase protein if they were expressed. We reported multiple populations from group B expressing S-RNase protein as detectable by immunoblotting (Table 2-2, Table 3-1). When these independently evolved SC groups come in contact, it is possible these functional but lowly expressed S-RNase alleles are becoming transcriptionally reactivated.

Fig. 4-1. Alignment of *LhgSRN-1* amino acid sequences in different populations of *S. habrochaites*. Populations are sorted from North to South. Ecuador site 33, AB0725, LA0407 and LA0094 are SC. The remaining sequences are from SI populations. Conserved regions are highlighted in yellow. Differences in amino acid sequence are highlighted in magenta.

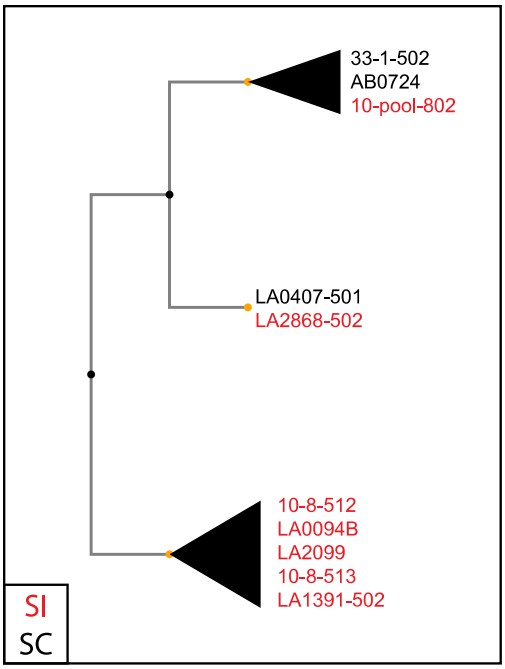

Fig. 4-2. Neighbor-joining tree of *LhgSRN-1*. The tree reflects relationships among *LhgSRN-1* amino acid sequences aligned in Fig. 4-1.

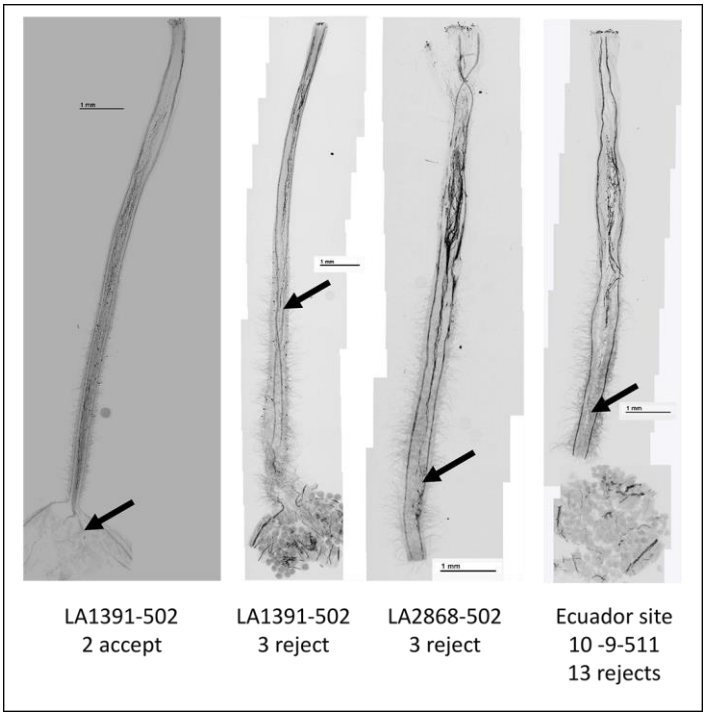

Fig. 4-3. Self-pollinations of *LhgSRN-1* positive SI populations. Growing pollen-tubes were stained using ABF and imaged using DAPI fluorescent microscopy. Arrows show the length of the longest pollen tube in the style. Of the SI individuals with *LhgSRN-1*, only LA1391 occasionally failed to reject self-pollen.

Fig. 4-4. Alignment of the *hab-7* S-RNase allele first detected in the transcriptome of LA2119. Because of the high rate (>99%) nucleotide conservation between the *S. peruvianum* S13 S-RNase and *hab-7*, only the amino acid sequence of *hab-7* (LA2119) is shown. Only one nucleotide difference was detected in the coding region (in red), a change from a Leucine in S13 to a Valine in *hab-7*. Conserved regions of S-RNases are highlighted in yellow.

Fig. 4-5. Nucleotide and amino acid alignment of *hab-8* and associated S-RNase alleles. The previously published *L. peruvianum* S3 S-RNase allele is 100% identical to an allele found in the pooled stylar transcriptome of LA2314. STOP and START codons are bolded and underlined. The newly described *hab-8* allele has a missense mutation highlighted in blue. The previously published S19 pseudogene has a single nucleotide deletion highlighted in yellow. The coding region is in red.

CHAPTER 5:

Conclusions

In this thesis, I provide evidence for multiple (up to four) independent SI→SC transitions at the northern species margin of *S. habrochaites* that not only allowed for efficient colonization of new habitats but have also contributed to the biodiversity of Ecuadorean populations (Table 5-1). Each SC group is genetically distinct according to population structure analysis, which showed that SC populations in Ecuador fall into four distinct mating system groups (SC-1, SC-2, SC-X and SC-Y, Chapter 2). Further supporting the idea of multiple independent mating system transitions, I have shown that these SC mating system groups possess different 'selfing' S-RNase alleles, two of which I discovered (*hab-7* and *hab-8,* Chapter 4). Although further experimentation will be required to fully characterize these alleles, I believe they may represent causal elements in transitions from SI→SC. My phenotypic studies of current Ecuadorean *S. habrochaites* populations (Chapter 3) have confirmed that only one SC mating system group in Ecuador (SC-2, Ecuador group A) exhibits reduced flower size and less stigma exsertion, "selfing syndrome" traits that are not exhibited in the other three distinct SC groups in Ecuador. The lack of selfing syndrome traits in the SC-1, SC-X and SC-Y mating system groups suggests either that these mating system transitions are relatively recent or, alternatively, that outcrossing may still be common in these SC populations, resulting in the maintenance of large flowers to attract insect pollinators. Another distinguishing feature of the SC-2 mating system group is its partial isolation from ancestor SI populations through the accumulation of pollen-side mutations (Markova *et al.*, 2016; Broz *et al.* 2017a).

My work has also demonstrated the potential for gene flow between distinct SC populations. In Ecuador region B, I showed that populations could contain either or both of the S-RNase allele associated with SC-1 (*hab-7*) or with SC-2 (*LhgSRN-1).* This suggests that two independent SC lineages

with overlapping geographical ranges have interbred, allowing for the expansion of the genetic diversity of SC populations (Chapter 2).

In addition to revealing multiple independent mating system transitions, two findings from my work have provided insight into the genetics underlying these transitions. Dr. Amanda Broz and I discovered that one known "selfing" S-RNase allele (*LhgSRN-1*), with its associated transposable element in the promoter region, can be found at low frequencies in SI populations as a component of the standing genetic variation. As a species' range expands during migration, selfing alleles may spread to rapid fixation following selection for reproductive assurance. This kind of process may have led to the establishment of the SC-2 mating system group A at the northernmost *S. habrochaites* species margin.

Second, I have detected S-RNase protein expression in some populations in Ecuador group B, in which two distinct SC populations that each lack S-RNase expression have interbred. These results suggest that previously "silenced" S-RNase alleles may have become reactivated for expression upon hybridization. Future experiments will focus on whether gene silencing machinery could be responsible for initial suppression of S-RNase expression to produce SC, and on mechanisms involved in the reactivation of S-RNase protein expression in this 'hybrid' region.

Table 5-1. Summary table of mating system groups.

Mating System Group	Region	*S-RNase* Allele	Interpopulation Barriers	Notes
SC-1	Ecuador D, some populations of Ecuador B	*hab-7*	None	Low expression of S-RNase
SC-2	Ecuador A, some populations of Ecuador B	*LhgSRN-1*	Pollen rejected by SI populations	Transposon results in low expression of S-RNase, selfing syndrome displayed
SC-X	Ecuador C	unknown	None	Populations lack all known S-RNase alleles
SC-Y	Ecuador E	*hab-8*	None	Loss of function S-RNase mutation

REFERENCES

AFLITOS, S., E. SCHIJLEN, H. DE JONG, D. DE RIDDER, S. SMIT, R. FINKERS, J. WANG, ET AL. 2014. Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. *Plant Journal* 80: 136–148.

ANTONOVICS, J. 1976. The Nature of Limits to Natural Selection. *Annals of the Missouri Botanical Garden* 63: 224.

BAEK, Y.S., P.A. COVEY, J.J. PETERSEN, R.T. CHETELAT, B. MCCLURE, and P.A. BEDINGER. 2015. Testing the SI x SC rule: Pollen-pistil interactions in interspecific crosses between members of the tomato clade (*Solanum* section *Lycopersicon*, Solanaceae). *American Journal of Botany* 102: 302–311.

BAEK, Y.S., S.M. ROYER, A.K. BROZ, P.A. COVEY, G. LOPEZ-CASADO, R. NUNEZ, P.J. KEAR, ET AL. 2016. Interspecific reproductive barriers between sympatric populations of wild tomato species (*Solanum* section *Lycopersicon*). *American Journal of Botany* 103: 1964–1978.

BAKER, H.G. 1955. Self-Compatibility and Establishment After "Long-Distance" Dispersal. *Evolution* 9: 347.

BAKER, H.G. 1967. Support for Baker's Law-As a Rule. *Evolution* 21: 853.

BARRETT, S.C.H. 2010. Understanding plant reproductive diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365: 99–109.

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

BEDINGER, P.A., A.K. BROZ, A. TOVAR-MENDEZ, and B. MCCLURE. 2017. Pollen-Pistil Interactions and Their Role in Mate Selection. *Plant Physiology* 173: 79–90.

BEDINGER, P.A., R.T. CHETELAT, B. MCCLURE, L.C. MOYLE, J.K.C. ROSE, S.M. STACK, E. VAN DER KNAAP, ET AL. 2011. Interspecific reproductive barriers in the tomato clade: opportunities to decipher mechanisms of reproductive isolation. *Sexual Plant Reproduction* 24: 171–187.

BERRY, P.E. 1982. The Systematics and Evolution of *Fuchsia* Sect. *Fuchsia* (Onagraceae). *Annals of The Missouri Botanical Garden* 69: 1–198.

BROZ, A.K., A.M. RANDLE, S.A. SIANTA, A. TOVAR-MÉNDEZ, B. MCCLURE, and P.A. BEDINGER. 2017A. Mating system transitions in *Solanum habrochaites* impact interactions between populations and species. *New Phytologist* 213: 440–454.

BROZ, A.K., R.F. GUERRERO, A.M. RANDLE, Y.S. BAEK, M.W. HAHN, and P.A. BEDINGER. 2017B. Transcriptomic analysis links gene expression to unilateral pollen-pistil reproductive barriers. *BMC Plant Biology* 17: 81.

CHALIVENDRA, S.C., G. LOPEZ-CASADO, A. KUMAR, A.R. KASSENBROCK, S. ROYER, A. TOVAR-MÈNDEZ, P.A. COVEY, ET AL. 2013. Developmental onset of reproductive barriers and associated proteome changes in stigma/styles of *Solanum pennellii*. *Journal of Experimental Botany* 64: 265–279.

CHUNG, I.-K., T. ITO, H. TANAKA, A. OHTA, H.G. NAN, and M. TAKAGI. 1994. Molecular diversity of three S-allele cDNAs associated with gametophytic self-incompatibility in *Lycopersicon peruvianum*. *Plant Molecular Biology* 26: 757–762.

COVEY, P.A., K. KONDO, L. WELCH, E. FRANK, S. SIANTA, A. KUMAR, R. NUÑEZ, ET AL. 2010. Multiple features that distinguish unilateral incongruity and self-incompatibility in the tomato clade. *The Plant Journal* 64: 367–378.

CRUDEN, R.W., and S.M. HERMANN-PARKER. 1977. Temporal Dioecism: An Alternative to Dioecism. *Evolution* 31: 863.

DARWIN, C. 1862. On the Two Forms, or Dimorphic Condition, in the Species of *Primula*, and on their remarkable Sexual Relations. *Journal of the Proceedings of the Linnean Society of London. Botany* 6: 77–96.

EARL, D.A., and B.M. VONHOLDT. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.

EDGAR, R.C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5: 113.

ELLSTRAND, N.C., and D.R. ELAM. 1993. Population genetic consequences of small population size: Implications for plant conservation. *Annu. Rev. Ecol. Syst* 24: 217–42.

FISHER, R.A. 1930. The genetical theory of natural selection. Oxford University Press.

FRANCIS, R.M. 2017. pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources* 17: 27–32.

GOLDRAIJ, A., K. KONDO, C.B. LEE, C.N. HANCOCK, M. SIVAGURU, S. VAZQUEZ-SANTANA, S. KIM, ET AL. 2006. Compartmentalization of S-RNase and HT-B degradation in self-incompatible *Nicotiana*. *Nature* 439: 805–810.

GOODWILLIE, C., R.D. SARGENT, C.G. ECKERT, E. ELLE, M.A. GEBER, M.O. JOHNSTON, S. KALISZ, ET AL. 2010. Correlated evolution of mating system and floral display traits in flowering plants and its implications for the distribution of mating system variation. *New Phytologist* 185: 311–321.

GRANDILLO, S., R. CHETELAT, S. KNAPP, D. SPOONER, I. PERALTA, M. CAMMARERI, O. PEREZ, ET AL. 2011. *Solanum* sect. *Lycopersicon*. In Wild Crop Relatives: Genomic and Breeding Resources, 129–215. Springer Berlin Heidelberg, Berlin, Heidelberg.

HARDER, L.D., and S.C.H. BARRETT. 1995. Mating cost of large floral displays in hermaphrodite plants. *Nature* 373: 512–515.

HARING, V., J.E. GRAY, B.A. MCCLURE, M.A. ANDERSON, and A.E. CLARKE. 1990. Self-incompatibility: a recognition system in plants. *Science* 250: 937–941.

HUA, Z.H., A. FIELDS, and T.H. KAO. 2008. Biochemical models for S-RNase-based self-incompatibility. *Molecular Plant* 1: 575–585.

HUA, Z., and T. KAO. 2008. Identification of major lysine residues of S $_3$ -RNase of *Petunia inflata* involved in ubiquitin–26S proteasome-mediated degradation in vitro. *The Plant Journal* 54: 1094–1104.

HUANG, S., H.S. LEE, B. KARUNANANDAA, and T.-H. KAO. 1994. Ribonuclease activity of *Petunia inflata* S proteins is essential for rejection of self-Pollen. *the Plant Cell Online* 6: 1021–1028.

IGIC, B., and J.W. 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. *Source: International Journal of Plant Sciences* 169: 93–104.

JAKOBSSON, M., and N.A. ROSENBERG. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806.

JOMBART, T., and I. AHMED. 2011. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* 27: 3070–3071.

KARRON, J.D., C.T. IVEY, R.J. MITCHELL, M.R. WHITEHEAD, R. PEAKALL, and A.L. CASE. 2012. New perspectives on the evolution of plant mating systems. *Annals of Botany* 109: 493–503.

KEENAN, K., P. MCGINNITY, T.F. CROSS, W.W. CROZIER, and P.A. PRODÖHL. 2013. diveRsity : An R package for the estimation and exploration of population genetics parameters and their associated errors R. B. O'Hara [ed.],. *Methods in Ecology and Evolution* 4: 782–788.

KOWYAMA, Y., C. KUNZ, I. LEWIS, E. NEWBIGIN, A.E. CLARKE, and M.A. ANDERSON. 1994. Self-compatibility in a *Lycopersicon peruvianum* variant (LA2157) is associated with a lack of style S-RNase activity. *Theoretical and Applied Genetics* 88: 859–864.

KUBO, K. -I., T. ENTANI, A. TAKARA, N. WANG, A.M. FIELDS, Z. HUA, M. TOYODA, ET AL. 2010. Collaborative Non-Self Recognition System in S-RNase-Based Self-Incompatibility. *Science* 330: 796–799.

KUMAR, S., G. STECHER, and K. TAMURA. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* 33: 1870–1874.

LANAUD, C., O. FOUET, T. LEGAVRE, U. LOPES, O. SOUNIGO, M.C. EYANGO, B. MERMAZ, ET AL. 2017. Deciphering the *Theobroma cacao* self-incompatibility system: from genomics to diagnostic markers for self-compatibility. *Journal of Experimental Botany* 68: 4775–4790.

LANDE, R., and D.W. SCHEMSKE. 1985. the Evolution of Self-Fertilization and Inbreeding Depression in Plants .1. Genetic Models. *Evolution; international journal of organic evolution* 39: 24–40.

LEE, H.-S., S. HUANG, and T. KAO. 1994. S proteins control rejection of incompatible pollen in *Petunia inflata*. *Nature* 367: 560–563.

LI, W., and R.T. CHETELAT. 2010. A Pollen Factor Linking Inter- and Intraspecific Pollen Rejection in Tomato. *Science* 330: 1827–1830.

LI, W., and R.T. CHETELAT. 2014. The Role of a Pollen-Expressed Cullin1 Protein in Gametophytic Self-Incompatibility in *Solanum*. *Genetics* 196: 439–442.

LI, W., and R.T. CHETELAT. 2015. Unilateral incompatibility gene ui1.1 encodes an S-locus F-box protein expressed in pollen of *Solanum* species. *Proceedings of the National Academy of Sciences* 112: 4417–4422.

LISCHER, H.E.L., and L. EXCOFFIER. 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28: 298–299.

LIU, W., J. FAN, J. LI, Y. SONG, Q. LI, Y. ZHANG, and Y. XUE. 2014. SCF/SLF-mediated cytosolic degradation of S-RNase is required for cross-pollen compatibility in S-RNase-based self-incompatibility in *Petunia hybrida*. *Frontiers in Genetics* 5: 1-15.

LLOYD, D.G., and C.J. WEBB. 1986. The avoidance of interference between the presentation of pollen and stigmas in angiosperms I. Dichogamy. *New Zealand Journal of Botany* 24: 135–162.

LUEBERT, F., and M. WEIGEND. 2014. Phylogenetic insights into Andean plant diversification. *Frontiers in Ecology and Evolution* 2: 1–17.

MARKOVA, D.N., J.J. PETERSEN, X. QIN, D.R. SHORT, M.J. VALLE, A. TOVAR-MENDEZ, B.A. MCCLURE, and R.T. CHETELAT. 2016. Mutations in two pollen self-incompatibility factors in geographically marginal populations of *Solanum habrochaites* impact mating system transitions and reproductive isolation. *American Journal of Botany* 103: 1847–1861.

MARKOVA, D.N., J.J. PETERSON, S.E. YAM, A. CORRAL, M.J. VALLE, W. LI, and R. CHETELAT. 2017. Evolutionary history of two pollen self-incompatibility factors reveals alternate routes to self-compatibility within *Solanum*. *American Journal of Botany (prepublication)*.

MARTIN, F.W. 1963. Distribution and Interrelationships of Incompatibility Barriers in the *Lycopersicon hirsutum* Humb. and Bonpl. complex. *Evolution* 17: 519–528.

McCLURE, B., B. MOU, S. CANEVASCINI, and R. BERNATZKY. 1999. A small asparagine-rich protein required for S-allele-specific pollen rejection in *Nicotiana*. *Proceedings of the National Academy of Sciences of the United States of America* 96: 13548–53.

McCLURE, B.A., J.E. GRAY, M.A. ANDERSON, and A.E. CLARKE. 1990. Self-incompatibility in *Nicotiana alata* involves degradation of pollen rRNA. *Nature* 347: 757–760.

McCLURE, B.A., V. HARING, P.R. EBERT, M.A. ANDERSON, R.J. SIMPSON, F. SAKIYAMA, and A.E. CLARKE. 1989. Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. *Nature* 342: 955–957.

McCLURE, B., F. CRUZ-GARCÍA, and C. ROMERO. 2011. Compatibility and incompatibility in S-RNase-based systems. *Annals of Botany* 108: 647–658.

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.

MOYLE, L.C. 2008. Ecological and evolutionary genomics in the wild tomatoes (*Solanum* Sect. *Lycopersicon*). *Evolution* 62: 2995–3013.

NE'EMAN, G., S. GOUBITZ, M.J.A. WERGER, and A. SHMIDA. 2011. Relationships between tree size, crown shape, gender segregation and sex allocation in *Pinus halepensis*, a Mediterranean pine tree. *Annals of botany* 108: 197–206.

OHASHI, K., and T. YAHARA. 2001. Behavioral responses of pollinators to variation in floral display size and their influences on the evolution of floral traits. *In* L. Chittka, and J. D. Thomson [eds.], Cognitive ecology of pollination, 274–296. Cambridge University Press, Cambridge, UK.

ORTIZ, P.L., M. ARISTA, and S. TALAVERA. 2002. Sex ratio and reproductive effort in the dioecious *Juniperus communis* subsp. *alpina* (Suter) Celak. (Cupressaceae) along an altitudinal gradient. *Annals of botany* 89: 205–11.

OWCZARZY, R., A. V TATAUROV, Y. WU, J.A. MANTHEY, K.A. MCQUISTEN, H.G. ALMABRAZI, K.F. PEDERSEN, ET AL. 2008. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic acids research* 36: W163-9.

PANNELL, J.R., J.R. AULD, Y. BRANDVAIN, M. BURD, J.W. BUSCH, P.-O. CHEPTOU, J.K. CONNER, ET AL. 2015. The scope of Baker's law. *New Phytologist* 208: 656–667.

PERALTA, I.E., D.M. SPOONER, and S. KNAPP. 2008. Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicoides*, sect. *Juglandifolia*, sect. *Lycopersicon* ; Solanaceae). American Society of Plant Taxonomists.

PRITCHARD, J.K., M. STEPHENS, and P. DONNELLY. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–59.

R CORE TEAM. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.r-project.org/.

RANDLE, A.M., J.B. SLYDER, and S. KALISZ. 2009. Can differences in autonomous selfing ability explain differences in range size among sister-taxa pairs of *Collinsia* (Plantaginaceae)? An extension of Baker's Law. *New Phytologist* 183: 618–629.

RENNER, S.S. 2014. The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. *American Journal of Botany* 101: 1588–1596.

RICK, C.M., and R.T. CHETELAT. 1991. The Breakdown of Self-Incompability in *Lycopersicum hirsutum*. *In* Hawkes, Lester, Nee, and Estrada [eds.], Solanaceae III: Taxonomy, Chemistry, Evolution, 253–256. Royal Botanic Gardens Kew and Linnaen Society of London.

RICK, C.M., J.F. FOBES, and S.D. TANKSLEY. 1979. Evolution of mating systems in *Lycopersicon hirsutum* as deduced from genetic variation in electrophoretic and morphological characters. *Plant Systematics and Evolution* 132: 279–298.

RICK, C.M., M. HOLLE, and R.W. THORP. 1978. Rates of cross-pollination in *Lycopersicon pimpinellifolium*: Impact of genetic variation in floral characters. *Plant Systematics and Evolution* 129: 31–44.

ROYO, J., C. KUNZ, Y. KOWYAMA, M. ANDERSON, A.E. CLARKE, and E. NEWBIGIN. 1994. Loss of a histidine residue at the active site of S-locus ribonuclease is associated with self-compatibility in *Lycopersicon peruvianum*. *Proceedings of the National Academy of Sciences* 91: 6511–6514.

SAKAI, A.K., F.W. ALLENDORF, J.S. HOLT, D.M. LODGE, J. MOLOFSKY, K.A. WITH, S. BAUGHMAN, ET AL. 2001. The population biology of invasive species. *Annual Review of Ecology and Systematics* 32: 305–332.

SICARD, A., and M. LENHARD. 2011. The selfing syndrome: A model for studying the genetic and evolutionary basis of morphological adaptation in plants. *Annals of Botany* 107: 1433–1443.

SIFRES, A., J. BLANCA, and F. NUEZ. 2011. Pattern of genetic variability of *Solanum habrochaites* in its natural area of distribution. *Genetic Resources and Crop Evolution* 58: 347–360.

SIGMAN, M.J., and R.K. SLOTKIN. 2016. The First Rule of Plant Transposable Element Silencing: Location, Location, Location. *The Plant Cell* 28: 304–313.

SIJACIC, P., X. WANG, A.L. SKIRPAN, Y. WANG, P.E. DOWD, A.G. MCCUBBIN, S. HUANG, and T. KAO. 2004. Identification of the pollen determinant of S-RNase-mediated self-incompatibility. *Nature* 429: 302–305.

SILLITOE, R.H. 1974. Tectonic segmentation of the Andes: Implications for magmatism and metallogeny. *Nature* 250: 542–545.

SPOONER, D.M., and R.. VAN DEN BERG. 2001. Quantitative Assessment of Corolla Shape Variation in Mexican *Solanum* Section *Petota*. *In* R. . van den Berg, G. W. W. Barendse, G. M. can der Weerden, and C. Mariani [eds.], Solanaceae V: Advances in Taxonomy and Utilization, 61–71. Nikmegen University Press.

STEBBINS, G.L. 1957. Self Fertilization and Population Variability in the Higher Plants. *The American Naturalist* 91: 337–354.

TAKAYAMA, S., and A. ISOGAI. 2005. Self-Incompatibility in Plants. *Annu. Rev. Plant Biol* 56: 467–89.

TOVAR-MÉNDEZ, A., A. KUMAR, K. KONDO, A. ASHFORD, Y.S. BAEK, L. WELCH, P.A. BEDINGER, and B.A. MCCLURE. 2014. Restoring pistil-side self-incompatibility factors recapitulates an interspecific reproductive barrier between tomato species. *The Plant Journal* 77: 727–736.

TOVAR-MÉNDEZ, A., L. LU, and B. MCCLURE. 2017. HT proteins contribute to S-RNase-independent pollen rejection in *Solanum*. *Plant Journal* 89: 718–729.

WEIGEND, M. 2004. Additional observations on the biogeography of the Amotape-Huancabamba zone in Northern Peru: Defining the South-Eastern limits. *Rev. peru. biol* 11: 127–134.

WICKHAM, H. 2009. ggplot2: Elegant Graphics for Data Analysis. http://ggplot2.org.

YU, G. 2017. scatterpie: Scatter Pie Plot. https://cran.r-project.org/package=scatterpie.