Genetic interaction between two unlinked loci underlies the loss of self-incompatibility in *Arabidopsis lyrata*

Yan Li[1,2*], Mark van Kleunen[2,3] & Marc Stift[2*]

[1]Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101 China

[2]Ecology, Department of Biology, University of Konstanz, Universitätsstrasse 10, D-78457 Konstanz, Germany

[3]Zhejiang Provincial Key Laboratory of Plant Evolutionary Ecology and Conservation, Taizhou University, Taizhou 318000, China

***Authors for correspondence:**

Yan Li; yan.li@uni-konstanz.de

Marc Stift; marcstift@gmail.com

Abstract

As the first step towards the evolution of selfing from obligate outcrossing, identifying the key mutations underlying the loss of self-incompatibility is of particular interest. However, our current knowledge is primarily based on sequence-based comparisons between selfing species and their self-incompatible relatives, which makes it hard to distinguish causal from secondary mutations. To by-pass this problem, we inferred the genetic basis of the loss of self-incompatibility by intercrossing plants from twelve geographically interspersed outcrossing and selfing populations of North-American *Arabidopsis lyrata* and determining the breeding system of 1,580 progeny. Self-incompatibility was not restored after crosses between different self-compatible populations. Equal frequencies of self-compatible and self-incompatible progeny emerged from crosses between parents with different breeding systems. We propose a two-locus genetic model for the loss of self-incompatibility in which specific *S*-locus haplotypes (S_1 and S_{19}) are associated with loss of self-incompatibility through their interaction with an unlinked modifier.

Keywords: breakdown of self-incompatibility, breeding system, evolution of selfing, mating system transition, segregation, *S*-locus, *S*-locus modifier

Introduction

To avoid the negative consequences of self-fertilization, about half of the angiosperms have some form of self-incompatibility[1, 2]. For example, Brassicaceae have a sporophytic self-incompatibility system, which renders plants self-incompatible through recognition and rejection of self-pollen. Two tightly linked recognition genes (the "female" gene *S*-locus Receptor Kinase – *SRK*, and the "male" gene *S*-locus Cystein Rich – *SCR*) encode stigma- and pollen proteins respectively, together forming what is commonly referred to as the *S*-locus. If stigma- and pollen proteins have matching specificities (as would be the case with self-pollination), pollen tubes cannot grow and preclude fertilization[3, 4, 5]. Self-incompatibility frequently breaks down[6, 7], and many extant self-compatible and selfing species have evolved from self-incompatible ancestors[8, 9, 10, 11, 12]. Since the loss of self-incompatibility is the first step towards the evolution of selfing, it is of particular interest to understand its underlying genetic basis.

Given the pivotal role of self-recognition in self-incompatibility, a loss-of-function mutation at the *S*-locus could represent the primary cause of the loss of self-incompatibility. However, mutations at any of the other genes (unlinked to the *S*-locus) involved in modifying self-recognition or encoding downstream components would also render plants self-compatible[13]. Direct evidence for a primary role of the *S*-locus was obtained by a gain-of-function study that transferred a functional *S*-locus from *Arabidopsis lyrata* into *A. thaliana*, which led to complete restoration of self-incompatibility in a few accessions[14, 15, 16]. In other accessions, transformation had no effect or only led to transient self-incompatibility, indicating that these accessions must have carried loss-of-function mutations both at the *S*-locus and for additional genetic elements required for functional self-incompatibility[15]. This makes it impossible to distinguish primary from secondary mutations, because after the initial loss of self-incompatibility, relaxed selection is expected to cause pseudogenization of any locus with an exclusive function in self-incompatibility. Therefore, the finding that extant selfing species with a fully developed selfing syndrome tend to have a non-functional *S*-locus (e.g., *Capsella rubella*[17]); and *A. thaliana*[18] does not necessarily provide evidence that mutations at the *S*-locus caused the loss of self-incompatibility. The

likelihood of secondary mutations obviously increases with divergence time, which makes study systems with more recent transitions to self-compatibility more suitable to identify the mutations underlying the transitions.

Here, we make use of the intraspecific breeding- and mating system variation in North American *A. lyrata*. Since only a few *A. lyrata* populations have lost self-incompatibility and transitioned to high selfing rates[19, 20], without evolution of a selfing syndrome[21] or purging of genetic load[22, 23], self-compatibility and selfing are likely of relatively recent origin. Joint clustering of selfing and outcrossing populations in multiple genetic backgrounds[20] and an association of self-compatibility with homozygosity for two different *S*-haplotypes, S_1 and S_{19}[24], have provided indirect evidence for multiple origins of selfing in this species. However, direct evidence is still lacking, and it is unknown whether there is a functional link between the S_1 and S_{19} haplotypes and self-compatibility. Absence of non-synonymous mutations in *SRK* and *SCR* for haplotype S_1 in selfing populations[25] suggests that loss-of-function mutations at self-recognition genes cannot explain the loss of self-incompatibility. Based on segregation patterns in an F_2 family derived from a single cross between a self-incompatible plant and a self-compatible S_1S_1 plant, it was hypothesized that a recessive modifier-locus unlinked to the *S*-locus causes self-compatibility in S_1-homozygotes of *A. lyrata*[24]. However, besides the single cross, there is still little empirical data to support this "modifier-hypothesis". Moreover, it has remained unclear whether the putative modifier is specific to S_1, or could also explain self-compatibility in backgrounds with other *S*-haplotypes such as S_{19}.

To address this, we did intra- and inter-population crosses using self-incompatible plants from six predominantly outcrossing populations and self-compatible plants from six predominantly selfing populations of North-American *A. lyrata*. First, we tested whether self-incompatibility was restored in progeny resulting from crosses between self-compatible parents from different selfing populations through genetic complementation. This would be expected if two different recessive loss-of-function mutations caused the loss of self-incompatibility in different selfing populations. Second, we tested whether progeny resulting from crosses between parents with different breeding systems were always SI

through functional dominance. This would be expected if a loss-of-function mutation underlies self-compatibility. Alternatively, mixtures of self-incompatible and self-compatible progeny would indicate a multigenic genetic basis of self-compatibility, for example through interaction of *S*-haplotypes and unlinked modifiers, such as proposed for the S_1-haplotype[24]. Crosses within and between outcrossing populations served as controls for crosses between breeding systems, and also to test the prediction of the modifier-hypothesis that crosses between self-incompatible partners can give rise to a low frequency of self-compatible progeny, matching previous observations of a low frequency of self-compatible individuals in otherwise self-incompatible populations[24].

Material and methods

Source plant material and crossing design

To study the inheritance of self-compatibility, we used seed material from 12 North American *Arabidopsis lyrata* populations (kindly provided by Barbara Mable, University of Glasgow) with contrasting breeding- and mating systems. According to genotyping of progeny arrays and self-pollinations for these 12 populations (Table S1), six populations (hereafter SI populations) mainly consist of self-incompatible (SI) plants, and have multi-locus outcrossing rates over 80%. The remaining six populations (hereafter SC populations, either $S_{19}S_{19}$ homozygotes or S_1S_1 homozygotes) have high frequencies of self-compatible (SC) plants (four populations 100%, one 88% and one 50%), and much lower outcrossing rates (five of them less than 30%, one population [TSSA] 41%; see Table S1)[20, 24].

In March 2013, to test whether self-incompatibility could be restored through complementation, we randomly selected 18 focal plants (three plants labelled A, B and C from each of the six SC populations; Fig. 1). We used the selected parental "SC" plants to produce progeny through between-population crosses ($BP_{♀SC \times ♂SC}$) among all plants labelled A, and the same for all plants labelled B and C. As controls, we made crosses within SC populations ($WP_{♀SC \times ♂SC}$) among the A, B and C plants,

respectively. To test whether self-compatibility as a trait behaves recessively, we also randomly selected 18 focal plants (three plants labelled A, B and C from each of the six SI populations; Fig. 1). We used these to produce progeny through crosses between breeding systems ($BBS_{♀SI \times ♂SC}$ and $BBS_{♀SC \times ♂SI}$) among all plants labelled A, and the same for the plants labelled B, and C. As controls for SI populations, we made crosses within SI populations ($WP_{♀SI \times ♂SI}$) among the A, B and C plants. To test whether SC plants can arise when crossing more distantly related SI plants, we also made crosses between SI populations ($BP_{♀SI \times ♂SI}$) among all plants labelled A, and the same for the plants labelled B, and C (Fig. 1). In August 2014, we replicated the complete crossing design (as summarized in Fig. 1 for the A, B, and C parental plants) with 36 different parental focal plants (three plants labelled D, E and F from each of the 12 populations).

In principle, our design would have generated 13 seed families per parental focal plant: 5 from crosses with plants from populations with the same breeding system ($BP_{♀SC \times ♂SC}$ or $BP_{♀SI \times ♂SI}$), 6 from crosses with plants from populations with a different breeding system ($BBS_{♀SI \times ♂SC}$ or $BBS_{♀SC \times ♂SI}$) and 2 from crosses with plants from the same population ($WP_{♀SC \times ♂SC}$ or $WP_{♀SI \times ♂SI}$). In total, for the 72 (36+36) parental plants, the complete design would thus have yielded 936 (468+468) seed families. However, not all crosses resulted in seeds, so that we obtained 898 seed families (446+452).

As not all populations are completely fixed for one or the other breeding system, we assessed the breeding system for all parental plants by performing at least six self-pollinations per plant on at least three different days. For the "SC populations", this confirmed that all 36 plants but one (genotype D of population TSSA) were indeed SC (Table S2). For the "SI populations", plants were either completely SI or leaky SI. Following criteria set in Refs [19, 20], the phenotype leaky SI is used for cases where self-pollination led to fruits with seeds, often a variable number among replicates, but fewer seeds per fruit than after cross-pollination with a compatible pollen donor.

To avoid accidental self-pollination when making crosses, we had emasculated potential recipient flowers of SC plants in the bud-stage before anther dehiscence. As a control, we left at least one emasculated flower without cross-pollen. In rare cases where this control led to fruit development (which

would indicate accidental self-pollination), we discarded any fruits from flowers on the same plant cross-pollinated on the same day. To obtain at least one developed fruit for each recipient-donor combination, we did up to three cross pollinations per combination. If all three cross-pollinations did not result in fruit set, we considered the parents to be cross-incompatible. All fruits were collected when mature (4–7 weeks after pollination) and stored under dry and cool conditions until further use.

Growth of progeny

To determine their breeding system, we sowed progeny from all 898 seed families in a growth chamber with a light period of 16 h per day, keeping temperature between 17°C and 21°C (day) and at 15°C (night) and humidity above 50%. For practical reasons, we sowed in five batches such that one seed per seed family was sown at one time, either for the 446 families derived from the A, B and C parents (batch 1, 2 and 5), or for the 452 families from the D, E and F parents (batch 3 and 4). Additionally, to increase the sample size for between-breeding-system crosses ($BBS_{♀SI \times ♂SC}$ and $BBS_{♀SC \times ♂SI}$), for this particular cross-type, we sowed seeds from all parental plants (A-F) in a sixth batch, and from parental plants A-C in a seventh and eighth batch. Batches were sown between July 2015 and February 2018 (Table S3), by placing single seeds on moistened peat-based nutrient-poor substrate (Einheitserde und Humuswerke Gebr. Patzer GmbH & Co., Sinntal, Germany) randomly assigned to a 2.5×3.2×11.0 cm cell in 54-cell QuickPot trays (QP 54 T/11, Herkuplast Kubern GmbH, Ering/Inn, Germany). Seedlings were transplanted into individual square polypropylene pots (7×7×6.5 cm, Pöppelmann, Germany) filled with the same substrate as for germination. We watered twice a week and fertilized weekly with 50 ml of 0.1% Scotts Universol® blue solution (Everris International B. V., Waardenburg, Netherlands).

Determination of the breeding system

Although not all seeds in each batch germinated and some plants did not flower within a few months from sowing, we could perform self-pollinations on a total of 1607 progeny (Table S3) from 653 out of 898

seed families representing all 144 (12×12) population combinations. On each plant that flowered, we performed ten self-pollinations on at least five different days by rubbing a ripe anther from a donor plant on the stigma of the recipient flower.

After pollination, fruits elongate to accommodate the developing seeds, and attain their final length one to two weeks after pollination[19, 26]. Fruit length is a good proxy of seed number (Fig. S1). Therefore, as seeds can only be counted reliably at least four weeks after pollination, we used fruit length at two weeks as a proxy of seed set to enable a higher throughput and allow screening of more plants. Fruits without any developing seeds do not elongate and thus roughly maintain the length of the ovary (FL_{zero}), whereas fruits with full seed set elongate to a maximum fruit length (FL_{max}). Values of FL_{zero} and FL_{max} might vary among individual plants. Therefore, after pollination and fertilization, we measured the fruit lengths at least two weeks after self-pollination, and subsequently expressed the degree of self-compatibility relative to FL_{zero} and FL_{max}. As it was logistically not feasible to perform control emasculations (to obtain FL_{zero} directly) and reference outcrosses (to obtain FL_{max} directly) for each progeny in the design, we used the available information from the parental plants to calculate expected FL_{zero} and FL_{max} values for each progeny. As there is considerable within and among population variation in the maximum fruit lengths, we first calculated FL_{max} for each parent individually by taking the median fruit length resulting from the four used pollen donor types (self, within population, between population, between breeding system), and taking the maximum of these medians:

$$FL_{parent_{max}} = \max\left(median(FL_{self}), median(FL_{WP}), median(FL_{BP}), median(FL_{BBS})\right)$$

Then, given that progeny fruit length is inherited additively (Table S4), we calculated the average expected FL_{max} for each progeny

$$FL_{max_{Progeny}} = \frac{1}{2}\left(FL_{max_{maternal\ plant\ ♀}} + FL_{max_{paternal\ plant ♂}}\right)$$

There is limited variation for the length of fruits without developing seeds within populations, but there are differences between populations[21]. To account for this, we calculated a population specific FL_{zero} as twice the mean pistil length in non-pollinated flowers for each population reported in[21]. Then, again

assuming an additive contribution of the maternal and paternal parent of each progeny, we calculated the average expected FL_{zero} fruit length for each progeny:

$$FL_{zero_{progeny}} = \frac{1}{2}\left(FL_{zero_{maternal\,population♀}} + FL_{zero_{paternal\,population♂}}\right)$$

Based on these and the average fruit length resulting from self-pollination, we calculated an index of self-compatibility (SC) for each progeny:

$$\text{SC-index} = \frac{FL_{progeny} - FL_{zero_{progeny}}}{FL_{max_{progeny}} - FL_{zero_{progeny}}}$$

In principle, this index ranges from 0 (complete self-incompatibility) to 1 (complete self-compatibility), but is not mathematically bound by 0 and 1 due to variation in the estimates of the different component parameters.

Progeny that do not produce any elongated fruits (SC-index = 0) after self-pollination are considered to be SI, although in principle female or male sterility would give a similar outcome. To discern truly SI progeny from ones with sterility, the female and male fertility of all apparent SI progeny was tested by using them as donor (to test male fertility) and recipient (to test female fertility) in crosses with up to two haphazardly chosen unrelated progeny (only SI progeny were used if testing male fertility of the progeny). This was done for all 356 progeny that appeared SI in the first four batches. Twenty-seven out of the 356 tested progeny did not produce seeds with either partner. Although such progeny are not necessarily sterile, as cross-incompatibility could also explain lack of seed formation after crossing crosses, we conservatively excluded the 27 potentially sterile progeny from further analyses (see Table S3).

Statistical analyses

To test for differences in SC-index between cross-types, we used linear mixed effects models implemented in the *lme* function of the 'nlme' package[27] in R 3.4.3[28]. With the SC-index as the dependent variable, the model fixed part included cross-type ($WP_{♀SI \times ♂SI}$, $WP_{♀SC \times ♂SC}$, $BP_{♀SI \times ♂SI}$, $BP_{♀SC \times ♂SC}$, $BBS_{♀SI \times ♂SC}$ and $BBS_{♀SC \times ♂SI}$), and the model random part included maternal population,

maternal individual (nested in maternal population), paternal population, paternal individual (nested in parental population) and batch number. To enable using a Gaussian error distribution and ensure an appropriate normality and homogeneity of model residuals, we transformed the SC-index, which ranged from -0.39 to 2.71, by adding 1.39 to all values and subsequent natural log-transformation. To account for heterogeneity of variance (i.e. differences in variance between the different cross-types), the model included a VarIdent variance structure that allowed each level of the factor cross type to have a different variance[29].

To compare the mean SC-index among the different cross-types, we specified a matrix defining a series of 13 custom linear comparisons between different cross-type combinations (contrasts C1 to C13) and used the *glht* function in the *multcomp* package[30] to perform z-tests corrected for multiple comparisons (Table 1). First, we tested whether progeny from between- and within-population crosses had a different SC-index, both for SI populations (C1: $BP_{♀SI \times ♂SI}$ *vs* $WP_{♀SI \times ♂SI}$) and for SC populations (C2: $BP_{♀SC \times ♂SC}$ *vs* $WP_{♀SC \times ♂SC}$). Second, we tested whether the SC-index of progeny from between-population crosses differed between SC and SI populations (C3: $BP_{♀SI \times ♂SI}$ *vs* $BP_{♀SC \times ♂SC}$). Third, we tested whether the SC-index of progeny from between-breeding-system crosses depended on the direction of the cross (C4: $BBS_{♀SI \times ♂SC}$ *vs* $BBS_{♀SC \times ♂SI}$). When C4 was not significant, we merged $BBS_{♀SI \times ♂SC}$ and $BBS_{♀SC \times ♂SI}$, and did further contrasts to test for phenotypic additivity (C5: BBS *vs* BP), complete phenotypic dominance of self-incompatibility (C6: BBS *vs* $BP_{♀SI \times ♂SI}$) or complete phenotypic dominance of self-compatibility (C7: BBS *vs* $BP_{♀SC \times ♂SC}$). When C4 was significant, we did not merge $BBS_{♀SI \times ♂SC}$ and $BBS_{♀SC \times ♂SI}$ and used separate tests for phenotypic additivity (C8: $BBS_{♀SI \times ♂SC}$ *vs* BP and C9: $BBS_{♀SC \times ♂SI}$ *vs* BP), phenotypic dominance of self-incompatibility (C10: $BBS_{♀SI \times ♂SC}$ *vs* $BP_{♀SI \times ♂SI}$ and C11: $BBS_{♀SC \times ♂SI}$ *vs* $BP_{♀SI \times ♂SI}$,), and phenotypic dominance of self-compatibility (C12: $BBS_{♀SI \times ♂SC}$ *vs* $BP_{♀SC \times ♂SC}$ and C13:$BBS_{♀SC \times ♂SI}$ *vs* $BP_{♀SC \times ♂SC}$). The SC-index values were bimodally distributed with peaks around 0 (corresponding to self-incompatibility) and 1 (corresponding to self-compatibility) (Fig. 2). Based on this, we assigned a phenotype to individual plants based on their SC-index value. Conservatively, we assigned the

phenotype “SC” to all progeny with an SC-index>0.75 and “SI” to all progeny with an SC-index<0.25. Progeny with intermediate values (SC-index between 0.25 and 0.75) were thus not phenotyped, and excluded from the analyses below. To test whether self-compatibility could be explained by one or two loci, we formulated one- and two-locus models in which each locus had a single allele required for self-compatibility. Alleles required for self-compatibility could either be recessive or dominant to other allelic states, and were assumed to be fixed in SC populations (i.e., the SC cross partner was always assumed to be homozygous for the involved allele or alleles). For outcrossing populations, we used exact binomial goodness-of-fit tests to evaluate extreme scenarios in which the alleles were completely absent or completely fixed, and a “best-fit” scenario. The latter corresponds to the allele-frequency scenario that, following Hardy-Weinberg principles, best predicted the observed frequencies of SC and SI progeny from between-breeding-system crosses. For these scenarios, we also assessed the goodness-of-fit for SC and SI frequencies in progeny from crosses between SI plants within populations ($WP_{♀SI \times ♂SI}$) and between populations ($BP_{♀SI \times ♂SI}$).

Results

Crosses within the self-compatible breeding system (♀SC × ♂SC crosses)

For crosses within SC populations ($WP_{♀SC \times ♂SC}$), the SC-index of progeny had a unimodal distribution with a median of 1.16. The vast majority of progeny (80 out of 86) had an SC-index above the conservative threshold of 0.75, and was thus phenotyped as SC (Fig. 2, Table S5). The same pattern emerged for crosses between SC populations ($BP_{♀SC \times ♂SC}$), with an even higher SC-index median of 1.30 (significant effect of C2 in Table 1, Fig. 2), and with 233 out of 242 progeny phenotyped as SC (Table S5). Thus, overall, most progeny from crosses between SC parents showed no evidence for restoration of self-incompatibility (Fig. 3). Of the 15 cases with an SC-index below the 0.75 threshold, 14 had values above the 0.25 threshold for self-incompatibility, and could thus not be phenotyped unambiguously. Only one had an SC-index below 0.25 and was thus phenotyped as SI (Fig. 3C, Table S5). This single exception involved parents from the SC populations PTP and TSSA, which otherwise produced unambiguously SC progeny (Fig. S2).

Crosses within the self-incompatible breeding system (♀SI × ♂SI crosses)

For crosses within SI populations ($WP_{♀SI \times ♂SI}$), the SC-index of progeny had a unimodal distribution with a median of 0.04. The majority of progeny (51 out of 79) had an SC-index below the threshold of 0.25, and was thus phenotyped as SI (Fig. 2, Table S5). However, a considerable number of progeny (9 out of 79) had an SC-index above 0.75 and was thus phenotyped as SC. The remaining 19 progeny had intermediate SC-index values and could thus not be phenotyped unambiguously according to our (conservative) thresholds (Table S5). Similar patterns emerged for crosses between SI populations ($BP_{♀SI \times ♂SI}$, no significant effect of C1 in Table 1). With a median SC-index of -0.03 (Fig. 2), 171 out of 215 progeny was phenotyped as SI, 8 as SC and the remaining 36 intermediate (Table S5). Thus, overall, most progeny from crosses between SI parents were SI (Fig. 3). However, SC progeny emerged in crosses within four out of six SI populations (MAN, PCR, PIN, SBD), and in several crosses between SI populations (eight different population-combinations; Fig. 3D).

Crosses between plants with a different breeding system (♀SC × ♂SI and ♀SI × ♂SC crosses)

The range of SC-index values of progeny from between-breeding-system crosses ($BBS_{♀SI \times ♂SC}$ and $BBS_{♀SC \times ♂SI}$) included the complete spectrum from completely SI to completely SC. Although on average intermediate to the SC-index of progeny from crosses within breeding systems ($BP_{♀SC \times ♂SC}$ and $BP_{♀SI \times ♂SI}$; no significant effect of C5 in Table 1), most between-breeding-system progeny could be phenotyped as either SC (477 out of 958) or SI (382 out of 958) (Table S5). Consequently, the SC-index showed a bimodal distribution with a median of 0.74 and peaks at c. 0 and c. 1.2 (Fig. 2B). The remaining 99 progeny had intermediate SC-index values and could thus not be phenotyped unambiguously according to our (conservative) thresholds (Table S5). The cross direction did not have a significant effect on the average SC-index value of progeny from between-breeding-system crosses (no significant effect of C4 in Table 1; Fig. 2), although the proportion of SC plants was higher when the mother was SC (Table S5). The patterns were not specific to any cross-combination, as the vast majority of population-combinations (31 out of 36) resulted in both SI and SC progeny (Fig. 3C; Fig. 3D). At the seed family level, 103 out of 266 families segregated for breeding system, i.e., they contained both SI progeny (SC-index<0.25) and SC progeny (SC-index>0.75) (Fig S2).

Evaluation of one and two-locus models to explain the loss of self-incompatibility

The observed relatively equal frequencies of SC and SI plants strongly speak against a single-locus genetic basis with completely additively interacting alleles, as this should have resulted in a unimodal distribution of SC-index values after between-breeding-system crosses. Similarly, single-locus models with allelic dominance of the allele conferring self-compatibility are unlikely, as these should have resulted in phenotypic dominance of self-compatibility, which could be statistically rejected (significant contrasts C6 and C7 in Table 1; significant deviation from single-locus dominant model in Table 2). The only single-locus model that could fit our observations would be one where the mutation acts recessively and has a relatively high frequency of 0.555 in SI populations (Table 2). However, for that scenario to

work, crosses between SI parents should have yielded a much higher frequency (30.8%) of SC individuals than observed both within and between populations (15% and 4.5%, respectively; Table 2).

Based on the overall observed frequencies of SC and SI plants after between-breeding-system crosses, two-locus models with dominance of both alleles required for self-compatibility could be rejected regardless of the assumed allele frequencies in outcrossing populations. Two-locus models in which one or both alleles act recessively could predict our data if the product of the frequency of the alleles required for SC equaled 0.555 (for the recessive-recessive models), or if the frequency of the recessive allele equaled 0.555 (for the dominant-recessive models).

Based on the observed frequencies of SC and SI plants after crosses between SI plants, the recessive-recessive model could be rejected, because the frequency of SC individuals (15% for within population crosses; 4.5% for between population crosses) did not fit the prediction of 30.8% SC progeny (Table 2). The dominant-recessive model could predict our data at the inferred recessive allele-frequency of 0.555, if the required dominant allele had a frequency of 0.284 (for crosses within populations) or 0.076 (for crosses between populations).

Discussion

By determining the breeding system of 1,580 progeny from crosses within and between selfing and outcrossing populations of North American *Arabidopsis lyrata*, we obtained unique insights in the genetic basis of the loss of self-incompatibility in a species with a relatively recent (intraspecific) transition to selfing. Our first main finding is that most progeny from crosses between SC parents from six different selfing populations were also SC. This shows that the loss of self-incompatibility in the different selfing populations of *A. lyrata* cannot be explained by different recessive loss-of-function mutations. Our second main finding is that crosses between parents with different breeding systems yielded an even segregation ratio of breeding-system phenotypes (i.e., equal frequencies of SC and SI progeny). Taken together, our findings show that single-locus models cannot explain the genetic basis of the breakdown of self-

incompatibility in North American *A. lyrata*. Instead, we propose that at least two loci are involved in self-compatibility: the *S*-locus and one or more modifier-loci. More specifically, we propose that plants carrying the dominant *S*-haplotype S_{19} are rendered SC if they are homozygous for a recessive modifier-allele, and that plants homozygous for the recessive *S*-haplotype S_1 are rendered SC if they carry a dominant modifier-allele.

A single loss-of-function mutation in any of the genes required for self-incompatibility provides the simplest potential genetic mechanism that would explain the breakdown of self-incompatibility[13]. As it is improbable for the same mutation to happen twice, under a scenario of multiple origins of self-compatibility, one would then expect that different loss-of-function mutations underlie self-compatibility in different selfing populations. Indeed, different loss-of-function mutations at the *S*-locus have been found in different *A. thaliana* haplogroups[31]. Since loss-of-function mutations are usually recessive[32], one would thus expect restoration of self-incompatibility through genetic complementation when the genomes of two SC plants with different recessive loss-of-function mutations are combined. Our results, however, do not support this scenario for *A. lyrata*. When crossing SC plants from six selfing populations with three different population genetic backgrounds[20], self-incompatibility was never restored. This lack of genetic complementation could either imply that one and the same recessive loss-of-function mutation causes self-compatibility in the different selfing backgrounds, or that the loss of self-incompatibility has a more complex genetic basis involving more loci.

Our finding of roughly equal frequencies of SI and SC progeny from between-breeding-system crosses ($BBS_{♀SI \times ♂SC}$ and $BBS_{♀SC \times ♂SI}$) and the emergence of some SC progeny from crosses between SI plants ($WP_{♀SI \times ♂SI}$, $BP_{♀SI \times ♂SI}$) supports the involvement of at least two loci in conferring self-compatibility in *A. lyrata*. This could reflect that a genetic interaction underlies the loss of self-incompatibility, but could also be explained by secondary mutations accumulating after a single primary

causal mutation. Such accumulation is expected due to relaxed purifying selection on genes involved in self-incompatibility after a transition to selfing, especially in selfing species that have diverged from their outcrossing ancestors sufficiently long ago for a selfing syndrome to evolve[33]. The resulting molecular signatures of secondary mutations in *A. thaliana*[18] and *C. rubella*[17] make it difficult to distinguish primary from secondary mutations in such species[13]. However, in the case of *A. lyrata*, secondary mutations are much less likely given the recent intraspecific origin of selfing in this species. Therefore, the involvement of two loci in *A. lyrata* more likely reflects a genetic interaction in which an allele of one locus confers self-compatibility by interacting with one or more alleles of another locus.

Interactions of modifiers with the *S*-locus are common features of homomorphic self-incompatibility systems[34, 35, 36]. Modifiers are particularly important in regulation of the sporophytic self-incompatibility of the Brassicaceae[13, 37, 38]. Based on this idea, it was proposed that in *A. lyrata*, an allele for a modifier-locus confers self-compatibility in plants homozygous for the *S*-locus haplotype S_1[24]. However, the empirical support for this model was thin, as it was only based on segregation ratios in a single F_2 family derived from one ♀SI × ♂SC cross. Moreover, although two selfing populations are fixed for the S_1 haplotype, four of the selfing populations are fixed for the S_{19} haplotype[24]. Our comprehensive crossing design, involving pairwise crosses between representatives from all six known selfing populations and six outcrossing populations, provides the first rigorous test of the modifier-hypothesis. The observed segregation of SC and SI cross-progeny in the first generation after between-breeding-system crosses, and the emergence of SC progeny after crosses between SI plants can all be explained by the hypothesis. Together, this provides strong support that self-compatibility in all selfing populations of North American of *A. lyrata* is conferred by a genetic interaction between modifier-alleles and the *S*-locus haplotypes S_1 and S_{19}.

The appearance of SC progeny in crosses between SI parents from different populations suggests that one or more of the modifier-alleles segregate in outcrossing populations. Maintenance of modifier-alleles conferring partial or complete self-compatibility is expected as a means to provide reproductive assurance under conditions where costs of complete self-incompatibility are high[39, 40, 41]. As the S_{19}-

haplotype is dominant and thus relatively rare (10%) in our six SI populations[24], the modifier allele interacting with S_{19} should act recessively, and have a relatively high frequency in outcrossing populations (>50%) to explain the observed equal frequencies of SI and SC phenotypes in between-breeding-system crosses. Unlike the dominant S_{19}, S_1 is recessive to any other *S*-haplotype[42], and thus generally has a high frequency (79%) in outcrossing populations (e.g., Refs [24, 43]). Thus, unless there is segregation distortion against S_1[44], this would mean that about 60% of the plants in outcrossing populations are homozygous for S_1. Since SC plants are relatively rare in outcrossing populations, this implies that the modifier-allele conferring self-compatibility in S_1 homozygotes should have a relatively low frequency in outcrossing populations (much lower than the inferred frequency of the modifier conferring self-compatibility in plants carrying S_{19}). Based on this, we propose that S_1 and S_{19} do not interact with the same modifier-allele. This would also explain why between-breeding-system crosses yielded different frequencies of SC phenotypes depending on whether the SC parent came from a selfing population fixed for S_1 or from a population fixed for S_{19} (Fig. S3; Table S5). It remains to be tested whether the two modifier-alleles are variants of the same locus, or belong to independent loci, and whether they can also interact with other *S*-locus haplotypes.

Acknowledgements

We thank Barbara Mable for providing seeds from natural populations, Sina Konitzer-Glöckner, Katya Mamonova and Nadja Köhler for help with crossing, and Heinz Vahlenkamp, Karoline Jetter, Yanjie Liu, Samuel Fernandes and Beate Rüter for help with the experiment. YL was funded by a scholarship from the China Scholarship Council. MS was supported by grant 388824194 from the German Research Foundation (Deutsche Forschungsgemeinschaft).

Author contributions

MS conceived the study. MS and MvK designed the experiment. YL executed the experiment. YL analyzed the data with inputs from MS and MvK. YL and MS wrote the paper with input from MvK.

Competing interests

Authors declare no competing interests

References

1. De Nettancourt D. Incompatibility and incongruity in wild and cultivated plants. Springer Science & Business Media (2001).

2. Razanajatovo M, *et al*. Plants capable of selfing are more likely to become naturalized. *Nat. Commun.* **7**, 13313 (2016).

3. Takayama S, Isogai A. Self-incompatibility in plants. *Annu. Rev. Plant Biol.* **56**, 467-489 (2005).

4. Ivanov R, Fobis-Loisy I, Gaude T. When no means no: guide to Brassicaceae self-incompatibility. *Trends Plant Sci.* **15**, 387-394 (2010).

5. Fujii S, Kubo K, Takayama S. Non-self- and self-recognition models in plant self-incompatibility. *Nat. Plants* **2**, 16130 (2016).

6. Stebbins GL. Flowering plants: evolution above the species level. (1974).

7. Wright SI, Kalisz S, Slotte T. Evolutionary consequences of self-fertilization in plants. *Proc. Biol. Sci.* **280**, 20130133 (2013).

8. Barrett SC. The evolution of plant sexual diversity. *Nat. Rev. Genet.* **3**, 274-284 (2002).

9. Igic B, Lande R, Kohn JR. Loss of self-incompatibility and its evolutionary consequences. *Int. J. Plant Sci.* **169**, 93-104 (2008).

10. Mable BK. Genetic causes and consequences of the breakdown of self-incompatibility: case studies in the Brassicaceae. *Genet. Res.* **90**, 47-60 (2008).

11. Vekemans X, Poux C, Goubet PM, Castric V. The evolution of selfing from outcrossing ancestors in Brassicaceae: what have we learned from variation at the *S*-locus? *J. Evol. Biol.* **27**, 1372-1385 (2014).

12. Shimizu KK, Tsuchimatsu T. Evolution of selfing: recurrent patterns in molecular adaptation. *Annu. Rev. Ecol. Evol. Syst.* **46**, 593-622 (2015).

13. Nasrallah JB. Plant mating systems: self-incompatibility and evolutionary transitions to self-fertility in the mustard family. *Curr. Opin. Genet. Dev.* **47**, 54-60 (2017).

14. Nasrallah ME, Liu P, Nasrallah JB. Generation of self-incompatible *Arabidopsis thaliana* by transfer of two *S*-locus genes from *A. lyrata*. *Science* **297**, 247-249 (2002).

15. Nasrallah ME, Liu P, Sherman-Broyles S, Boggs NA, Nasrallah JB. Natural variation in expression of self-incompatibility in *Arabidopsis thaliana*: implications for the evolution of selfing. *Proc. Natl. Acad. Sci.* USA **101**, 16070-16074 (2004).

16. Boggs NA, Nasrallah JB, Nasrallah ME. Independent S-locus mutations caused self-fertility in *Arabidopsis thaliana*. *PLoS Genet.* **5**, e1000426 (2009).

17. Guo YL, *et al*. Recent speciation of *Capsella rubella* from *Capsella grandiflora*, associated with loss of self-incompatibility and an extreme bottleneck. *Proc. Natl. Acad. Sci.* USA **106**, 5246-5251 (2009).

18. Tsuchimatsu T, *et al.* Patterns of polymorphism at the self-incompatibility locus in 1,083 *Arabidopsis thaliana* genomes. *Mol. Biol. Evol.* **34**, 1878-1889 (2017).

19. Mable BK, Schierup MH, Charlesworth D. Estimating the number, frequency, and dominance of *S*-alleles in a natural population of *Arabidopsis lyrata* (Brassicaceae) with sporophytic control of self-incompatibility. *Heredity* **90**, 422-431 (2003).

20. Foxe JP, Stift M, Tedder A, Haudry A, Wright SI, Mable BK. Reconstructing origins of loss of self-incompatibility and selfing in North American *Arabidopsis lyrata*: a population genetic context. *Evolution* **64**, 3495-3510 (2010).

21. Carleial S, van Kleunen M, Stift M. Small reductions in corolla size and pollen: ovule ratio, but no changes in flower shape in selfing populations of the North American *Arabidopsis lyrata*. *Oecologia* **183**, 401-413 (2017).

22. Carleial S, van Kleunen M, Stift M. Relatively weak inbreeding depression in selfing but also in outcrossing populations of North American *Arabidopsis lyrata*. *J. Evol. Biol.* **30**, 1994-2004 (2017).

23. Li Y, van Kleunen M, Stift M. Sibling competition does not magnify inbreeding depression in North American *Arabidopsis lyrata*. *Heredity* https://doi.org/10.5061/dryad.6474895 (2019).

24. Mable BK, *et al.* What causes mating system shifts in plants? *Arabidopsis lyrata* as a case study. *Heredity* **118**, 110 (2017).

25. Goubet PM, *et al.* Contrasted patterns of molecular evolution in dominant and recessive self-incompatibility haplotypes in *Arabidopsis*. *PLoS Genet.* **8**, e1002495 (2012).

26. Mable BK, Robertson AV, Dart S, Di Berardo C, Witham L. Breakdown of self-incompatibility in the perennial *Arabidopsis lyrata* (Brassicaceae) and its genetic consequences. *Evolution* **59**, 1437-1448 (2005).

27. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. nlme: Linear and nonlinear mixed effects models. (2013).

28. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2013).

29. Zuur A, Ieno EN, Walker N, Saveliev AA, Smith GM. Mixed effects models and extensions in ecology with R. Springer Science & Business Media (2009).

30. Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. *Biom. J.* **50**, 346-363 (2008).

31. Shimizu KK, Shimizu-Inatsugi R, Tsuchimatsu T, Purugganan MD. Independent origins of self-compatibility in *Arabidopsis thaliana*. *Mol. Ecol.* **17**, 704-714 (2008).

32. Nanjundiah V. Why are most mutations recessive? *J. Genet.* **72**, 85 (1993).

33. Sicard A, Lenhard M. The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants. *Ann. Bot.* **107**, 1433-1443 (2011).

34. Nasrallah ME, Kandasamy MK, Nasrallah JB. A genetically defined trans-acting locus regulates *S*-locus function in *Brassica*. *Plant J.* **2**, 497-506 (1992).

35. Levin DA. The evolutionary significance of pseudo-self-fertility. *Am. Nat.* **148**, 321-332 (1996).

36. Tsukamoto T, *et al.* Breakdown of self incompatibility in a natural population of *Petunia axillaris* (Solanaceae) in Uruguay containing both self-incompatible and self compatible plants. *Sex Plant Reprod.* **12**, 6-13 (1999).

37. Liu P, Sherman-Broyles S, Nasrallah ME, Nasrallah JB. A cryptic modifier causing transient self-incompatibility in *Arabidopsis thaliana*. *Curr. Biol.* **17**, 734-740 (2007).

38. Indriolo E, Safavian D, Goring DR. The ARC1 E3 Ligase Promotes Two Different Self-Pollen Avoidance Traits in *Arabidopsis*. *Plant Cell* **26**, 1525-1543 (2014).

39. Vallejo-Marin M, Uyenoyama MK. On the evolutionary costs of self-incompatibility: incomplete reproductive compensation due to pollen limitation. *Evolution* **58**, 1924-1935 (2004).

40. Busch JW, Schoen DJ. The evolution of self-incompatibility when mates are limiting. *Trends Plant Sci.* **13**, 128-136 (2008).

41. Harder LD, Richards SA, Routley MB. Effects of reproductive compensation, gamete discounting and reproductive assurance on mating-system diversity in hermaphrodites. *Evolution* **62**, 157-172 (2008).

42. Prigoda NL, Nassuth A, Mable BK. Phenotypic and genotypic expression of self-incompatibility haplotypes in *Arabidopsis lyrata* suggests unique origin of alleles in different dominance classes. *Mol. Biol. Evol.* **22**, 1609-1620 (2005).

43. Schierup MH, Mable BK, Awadalla P, Charlesworth D. Identification and characterization of a polymorphic receptor kinase gene linked to the self-incompatibility locus of *Arabidopsis lyrata*. *Genetics* **158**, 387-399 (2001).

44. Bechsgaard J, Bataillon T, Schierup MH. Uneven segregation of sporophytic self-incompatibility alleles in *Arabidopsis lyrata*. *J. Evol. Biol.* **17**, 554-561 (2004).

Table 1 Between cross-type comparisons of mean SC-index

Contrasts	Estimate	z	P
C1: $BP_{♀SI \times ♂SI}$ *vs* $WP_{♀SI \times ♂SI}$	-0.06	-2.15	0.20
C2: $BP_{♀SC \times ♂SC}$ *vs* $WP_{♀SC \times ♂SC}$	**0.04**	**3.53**	**0.004**
C3: $BP_{♀SI \times SI}$ *vs* $BP_{♀SC \times ♂SC}$	**-0.63**	**-19.11**	**<0.0001**
C4: $BBS_{♀SI \times ♂SC}$ *vs* $BBS_{♀SC \times ♂SI}$	-0.08	-2.11	0.22
C5 BBS *vs* BP	-0.01	-0.94	0.91
C6: BBS *vs* $BP_{♀SI \times ♂SI}$	**0.33**	**14.56**	**<0.0001**
C7: BBS *vs* $BP_{♀SC \times ♂SC}$	**-0.30**	**-14.39**	**<0.0001**
C8: $BBS_{♀SI \times ♂SC}$ *vs* BP			
C9: $BBS_{♀SC \times ♂SI}$ *vs* BP			
C10: $BBS_{♀SI \times ♂SC}$ *vs* $BP_{♀SI \times ♂SI}$		Tests are irrelevant[a]	
C11: $BBS_{♀SC \times ♂SI}$ *vs* $BP_{♀SI \times ♂SI}$			
C12: $BBS_{♀SI \times ♂SC}$ *vs* $BP_{♀SC \times ♂SC}$			
C13: $BBS_{♀SC \times ♂SI}$ *vs* $BP_{♀SC \times ♂SC}$			

Notes: Contrasts were specified using a contrast matrix and significances evaluated using post-hoc z-tests. Negative estimates indicate that the first cross type in the contrast had a lower mean SC-index value than the second one, and positive estimates indicate the opposite. Significant effects are highlighted in bold. WP: Within Population. BP: Between Population. BBS: Between Breeding System (by definition also between populations). Subscripts indicate the breeding system (SI: Self-Incompatible; SC: Self-Compatible) of the maternal (♀) and paternal (♂) cross-partners.

[a] Given that the SC-index of progeny did not depend on the maternal/paternal breeding system (i.e., C4 was not significant), contrasts C8-C13 are irrelevant and thus not reported. However, as they were part of the design and could not be omitted *a priori*, significance levels of individual contrasts have conservatively been corrected for all 13 contrasts.

Table 2 Evaluation of one- and two locus genetic models to predict breeding system segregation ratios

Genetic model	Genotypes conferring SC[a]	Genotypes conferring SI	Crosses between breeding systems ($BBS_{♀SI \times ♂SC}$ and $BBS_{♀SC \times ♂SI}$)		[b]Crosses within SI breeding system ($WP_{♀SI \times ♂SI}$ and $BP_{♀SI \times ♂SI}$)	
			Scenarios for frequency in SI populations of allele(s) required for SC and associated predicted percentage of SC plants	Evaluation of model fit to observed frequencies[c]: 477 SC (55.5%) 382 SI (44.5%)	Predicted frequency of SC plants in progeny	Evaluation of model fit to observed frequencies[c]: WP: 9 SC (15.0%) 51 SI (85.0%) BP: 8 SC (4.5%); 171 SI (95.5%)
Single-locus dominant	AA Aa	aa	Any [A] 100% SC	Reject (**P<2.2^{-16}**)		
Single-locus recessive	aa	AA Aa	[a] = 0 0% SC	Reject (**P<2.2^{-16}**)		
			Best-fit[d]*:* [a] = 0.555 55.5% SC	Fits (P=1)	30.8% SC	WP: Reject (**P=0.007378**) BP: Reject (**P<2.2^{-16}**)
			[a] = 1 100% SC	Reject (**P<2.2^{-16}**)		
Two-locus dominant-dominant	AABB, AABb, AaBB, AaBb	AAbb, Aabb, aaBB, aaBb, aabb	Any [A] and [B] 100% SC	Reject (**P<2.2^{-16}**)		
Two-locus recessive-recessive	aabb	AABB, AABb, AAbb, AaBB, AaBb, Aabb, aaBB, aaBb	[a]=0 or [b]=0 0% SC	Reject (**P<2.2^{-16}**)		
			Best-fit[d]*:* [a]*[b]=0.555 55.5% SC	Fits (P=1)	[a]*[b]=0.555 30.8% SC	WP: Reject (**P=0.007378**) BP: Reject (**P<2.2^{-16}**)
			[a]=1 and [b]=1 100% SC	Reject (**P<2.2^{-16}**)		

Table 2 *continued*

Genetic model	Genotypes conferring SC[a]	Genotypes conferring SI	Crosses between breeding systems ($BBS_{♀SI \times ♂SC}$ and $BBS_{♀SC \times ♂SI}$)		[b]Crosses within SI breeding system ($WP_{♀SI \times ♂SI}$ and $BP_{♀SI \times ♂SI}$)	
			Scenarios for frequency in SI populations of allele(s) required for SC and associated predicted percentage of SC plants	Evaluation of model fit to observed frequencies[c]: 477 SC (55.5%) 382 SI (44.5%)	Predicted frequency of SC plants in progeny	Evaluation of model fit to observed frequencies[c]: WP: 9 SC (15.0%) 51 SI (85.0%) BP: 8 SC (4.5%); 171 SI (95.5%)
Two-locus dominant-recessive	<u>AAbb,</u> Aabb	AABB, AABb, AaBB, AaBb aaBB, aaBb, aabb	Any [A] and [b]=0 0% SC	Reject (**P<2.2^{-16}**)		
			Best-fit[d]*:* Any [A] and [b]=0.555 55.5% SC	Fits (P=1)	[A]=0, [b]=0.555 0% SC WP: *Best-fit*[d]*:* [A]=0.284, [b]=0.555; 15.0% SC BP: *Best-fit*[d]*:* [A]=0.076, [b]=0.555 4.5% SC [A]=1, [b]=0.555 30.8% SC	WP: Reject (**P<2.2^{-16}**) BP: Reject (**P<2.2^{-16}**) WP: Fits (P=1) BP: Fits (P=1) WP: Reject (**P=0.007378**) BP: Reject (**P<2.2^{-16}**)
			Any [A] and [b]=1 100% SC	Reject (**P<2.2^{-16}**)		

Abbreviations: WP: Within Population; BP: Between Population, BBS: Between Breeding System; SC: Self-Compatible; SI: Self-Incompatible;

[a] We assume the allele(s) needed for selfing are fixed in selfing (SC) populations due to high selfing rates, and thus that the SC cross-partner had the underlined homozygote genotype.

[b] These crosses were only analyzed if segregation ratios in BBS progeny (previous column) could not be rejected.

[c] Model fit was evaluated by exact binomial goodness-of-fit tests. Separate frequencies for the S_1S_1 and $S_{19}S_{19}$ selfing backgrounds are given in Table S5.

[d] Frequency that provides best fit to observed SI/SC frequencies after crosses between-breeding-system ($BBS_{♀SI \times ♂SC}$ and $BBS_{♀SC \times ♂SI}$), within SI populations ($WP_{♀SI \times ♂SI}$) or between SI populations ($BP_{♀SI \times ♂SI}$)

Fig. 1 Schematic representation of the crossing design

We generated within-population (WP, grey), between-population (BP, orange; always within a breeding system) and between-breeding-system (BBS, red; always between populations) crosses using three parents (A, B, C) from six outcrossing (SI) populations (SI-1 to SI-6) and six selfing (SC) populations (SC-1 to SC-6). Solid borders indicate crosses where the maternal parent was SI, dashed borders where the maternal parent was SC. We duplicated the design with three further plants (D, E, F; not shown in this scheme) for each population.

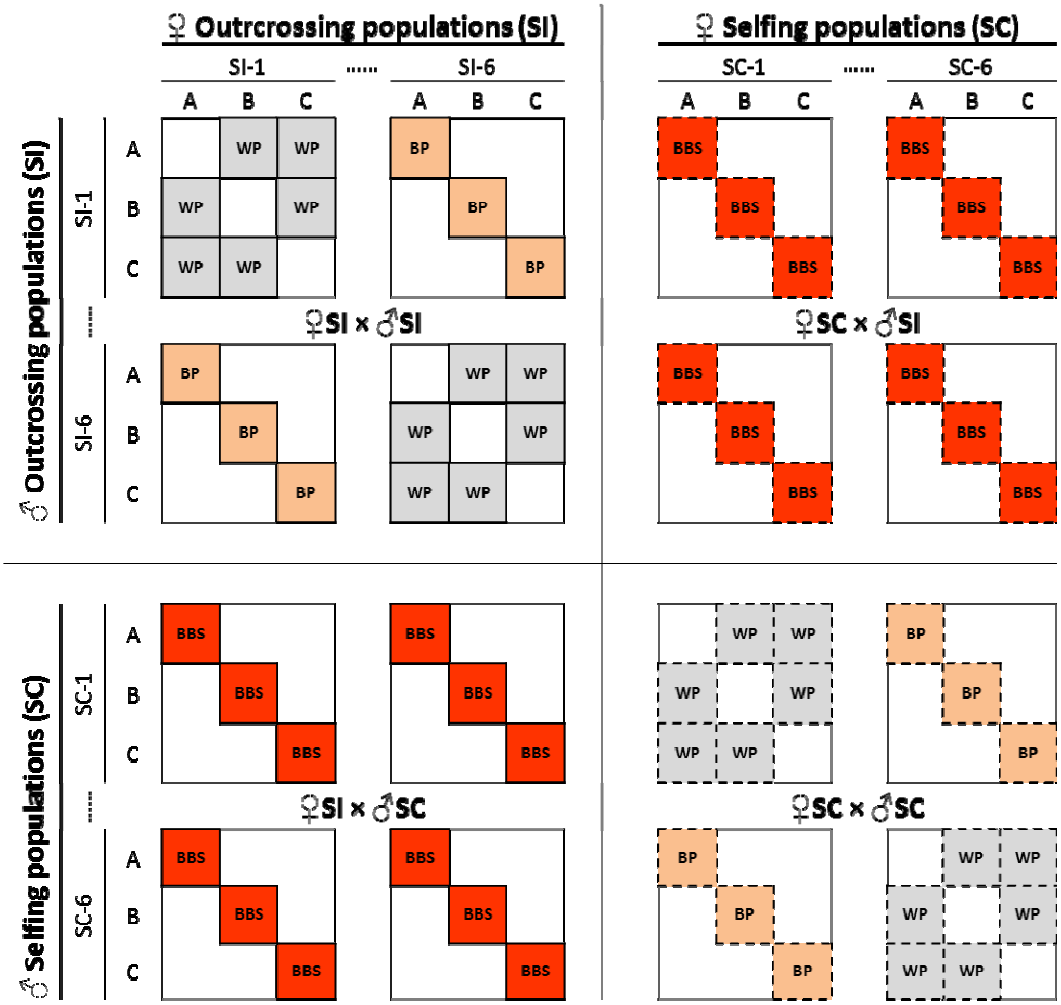


Fig. 2 SC-index values for progeny from crosses within and between breeding systems

A) Box-and whisker plot. Boxes delimit the interquartile range (IQR) with the median indicated as a solid line. Whiskers extend to 1.5×IQR or to the lowest/highest data point within 1.5×IQR. Points beyond the whiskers' limits represent outliers. B) Density plot. Color scheme used: grey for within-population (WP) crosses; orange for between-population (BP) crosses; red for between-breeding-system (BBS) crosses (by design also between populations). Solid borders and lines are used for cross-types with self-incompatible maternal parents (♀SI) and dashed borders and lines for cross-types with self-compatible maternal parents (♀SC). Note the bimodal shape for the BBS cross-type with peaks matching those observed for the cross-types from between-SI-populations crosses and between-SC-populations crosses.

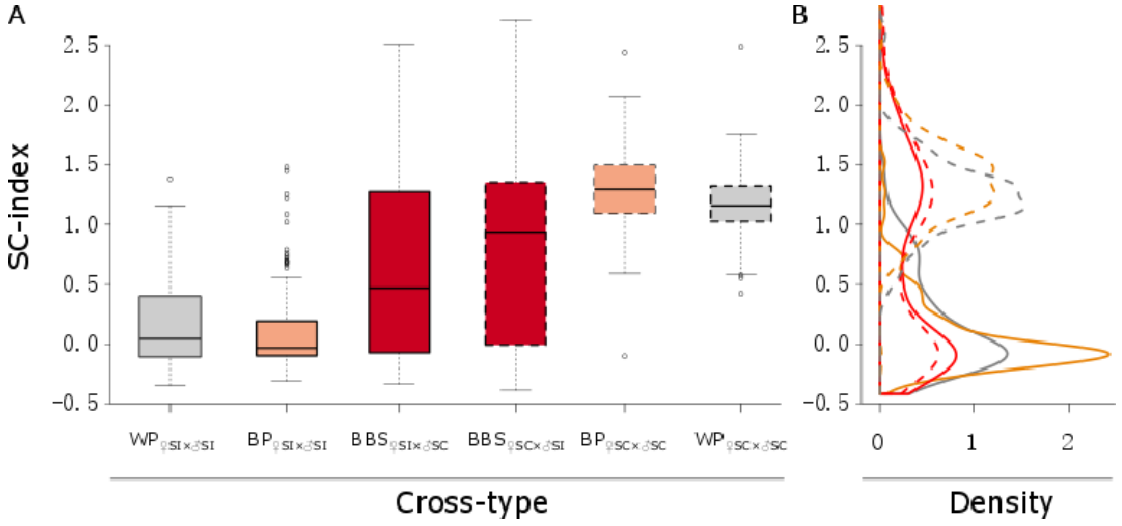
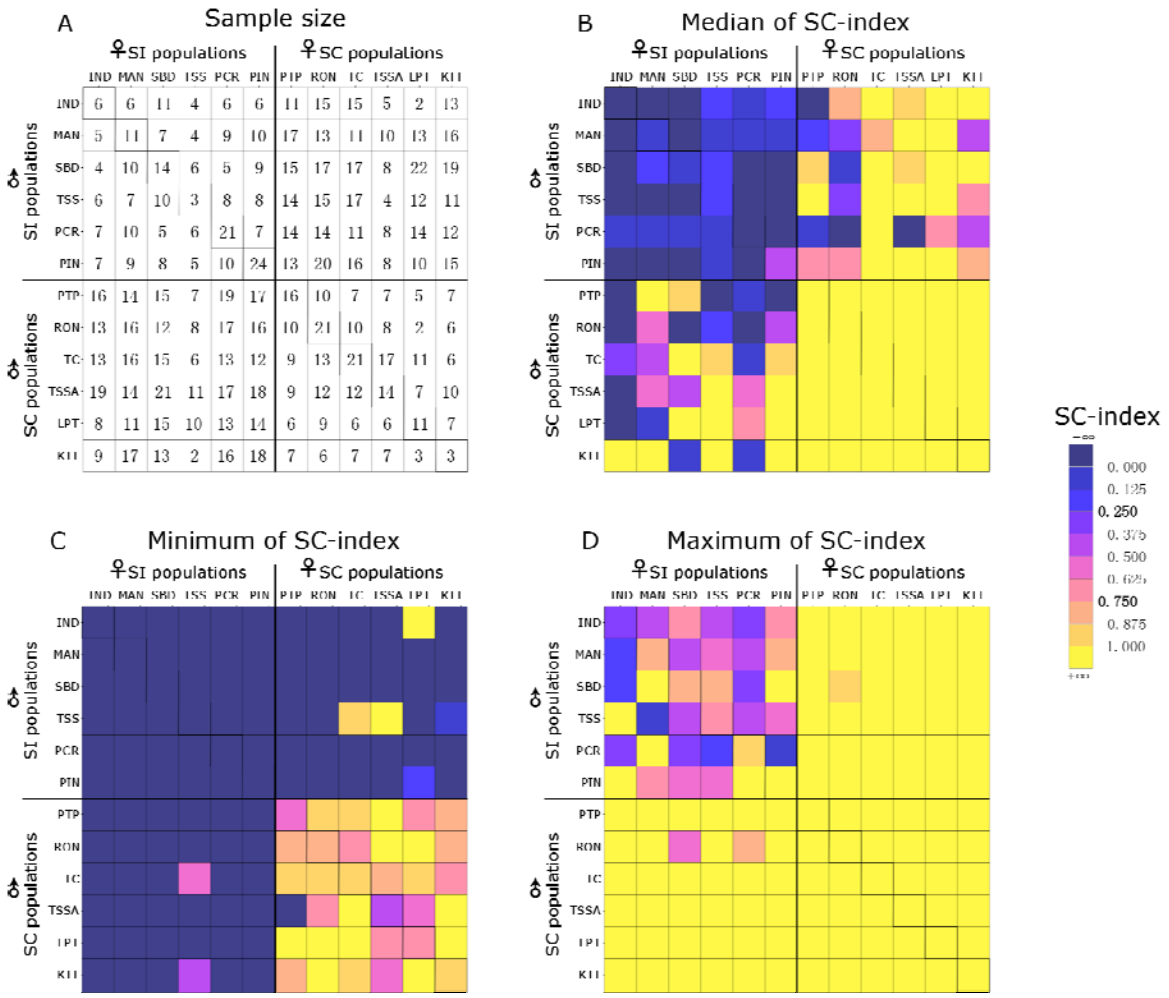


Fig. 3 Summary statistics for the SC-index for within and between population crosses

A) Cell-specific number of progeny for which the breeding system phenotype was determined; B) Heat-map of median SC-index; C) Heat-map of minimum SC-index; D) Heat-map of maximum SC-index. Population codes correspond to the ones in Table S1. Diagonal cells marked with an extra outline represent progeny from within-population crosses. Other squares represent each possible combination of populations in two directions (with the maternal population shown at the top). The 36-cell top-left quadrat and 36-cell bottom-right quadrat represent crosses between plants with the same breeding system (♀SI × ♂SI and ♀SC × ♂SC, respectively). The 36-cell top-right quadrat and 36-cell bottom-left quadrat represent crosses between plants with different breeding systems (♀SC × ♂SI and ♀SI × ♂SC, respectively).



Supporting information

Table S1 Background information for the six selfing and six outcrossing *A. lyrata* populations from which we used seeds

Population code	Location, Lake, State/Province, Country	Population coordinates		Predominant breeding system (proportion of SC individuals) [a]	Outcrossing rate T_m [a]
		Latitude	Longitude		
IND	Indiana Dunes National Lakeshore, Lake Michigan, Indiana, USA	N 41˚37'17"	W 87˚12'44"	SI (0.14)	0.99
MAN	Manitoulin Island, Georgian Bay, Ontario, Canada	N 47˚39'54"	W 82˚15'52"	SI (0.13)	0.83
SBD	Sleeping Bear Dunes National Lakeshore, Lake Michigan, Michigan, USA	N 44˚56'20"	W 85˚52'13"	SI (0.00)	0.94
TSS	Tobermory Singing Sands, BPNP, Lake Huron, Ontario, Canada	N 45˚11'33"	W 81˚35'02"	SI (0.00)	0.91
PCR	Port Crescent State Park, Lake Huron, Michigan, USA	N 44˚00'15"	W 83˚04'26"	SI (0.00)	0.98
PIN	Pinery Provincial Park, Lake Huron, Ontario, Canada	N 43˚16'08"	W 81˚49'53"	SI (0.00)	0.84
PTP	Point Pelee National Park, Lake Erie, Ontario, Canada	N 41˚55'40"	W 82˚30'51"	SC (1.00)	0.09
RON	Rondeau Provincial Park, Lake Erie, Ontario, Canada	N 42˚15'41"	W 81˚50'47"	SC (1.00)	0.28
TC	Tobermory Cliffs, Bruce Peninsula National Park, Georgian Bay, Ontario, Canada	N 45˚14'30"	W 81˚31'03"	SC (0.88)	0.18
TSSA	Tobermory Singing Sands Alvar, BPNP, Lake Huron, Ontario, Canada	N 45˚11'27"	W 81˚35'26"	SC (0.50)	0.41
LPT	Long Point Provincial Park, Lake Erie, Ontario, Canada	N 42˚34'47"	W 80˚23'15"	SC (1.00)	0.13
KTT	Kitty Todd State Nature Preserve, -, Ohio, USA	N 41˚37'14"	W 83˚47'15"	SC (1.00)	0.31

Notes: Seeds from natural populations were kindly provided by Barbara Mable allowing us to grow the parents of our crossing design (Fig. 1).
[a] Values from Ref [20]

Table S2 Fruit lengths and corresponding SC-index values after self-pollination of the 72 parental plants in our crossing design (Fig. 1).

		Set 1				Set2			
Mating system	Population code	Plant code	Flower ID	Fruit length (mm)	SC-index	Plant code	Flower ID	Fruit length (mm)	SC-index.
SI	IND	IND-A	37	5.98	0.00	IND-D	1	5.83	-0.01
			38	5.03	-0.07		2	10.09	0.20
			39	6.24	0.02		3	6.03	0.00
			40	10.7	0.36		4	14.85	0.43
			41	8.35	0.18		5	5.92	0.00
			42	5.98	0.00		6	8.55	0.12
		IND-B	37	5	-0.09	IND-E	1	4.12	-0.11
			38	4.01	-0.19		2	3.25	-0.16
			39	3.85	-0.20		3	3.22	-0.16
			40	3.9	-0.20		4	3.04	-0.17
			41	3.15	-0.27		5	NA	NA
			42	4.28	-0.16		6	NA	NA
		IND-C	1	8.53	0.14	IND-F	1	3.13	-0.16
			2	7.89	0.11		2	3.12	-0.16
			3	3.75	-0.13		3	4.04	-0.11
			4	3.74	-0.13		4	3.15	-0.16
			5	4.18	-0.10		5	4.15	-0.10
			6	4.02	-0.11		6	4.27	-0.10
	MAN	MAN-A	40	7.87	0.07	MAN -D	1	NA	NA
			41	7.15	0.04		2	NA	NA
			42	5.44	-0.02		3	NA	NA
			43	13.57	0.29		4	NA	NA
			44	13.98	0.30		5	5.05	-0.05
			45	5.29	-0.03		6	5.04	-0.06
		MAN-B	43	5.22	-0.06	MAN -E	1	4.26	-0.12
			44	4.74	-0.10		2	4.15	-0.12
			45	5.44	-0.04		3	NA	NA
			46	5.92	-0.01		4	NA	NA
			47	5.16	-0.06		5	6.28	0.02
			48	13.08	0.54		6	NA	NA
		MAN-C	1	4.67	-0.05	MAN-F	1	7.05	0.05
			2	4.38	-0.06		2	5.47	-0.02
			3	5.21	-0.03		3	5.15	-0.04
			4	4.92	-0.04		4	5.03	-0.05
			5	4.7	-0.05		5	5.91	0.00
			6	4.94	-0.04		6	8.69	0.13
	PCR	PCR-A	40	14.29	0.55	PCR-D	1	4.17	-0.12
			41	15.58	0.64		2	6.9	0.06
			42	15.32	0.62		3	5.14	-0.06
			43	11.42	0.36		4	5.72	-0.02
			44	16.42	0.70		5	4.99	-0.06
			45	18.17	0.81		6	6.37	0.02
			46	16.64	0.71				
			47	14.59	0.57				
			48	12.47	0.43				
			49	8.91	0.19				
		PCR-B	43	6.27	0.02	PCR-E	1	5.12	-0.05
			44	5.71	-0.02		2	5.71	-0.02
			45	5.59	-0.03		3	7.2	0.07
			46	5.9	-0.01		4	4.63	-0.08
			47	6.05	0.00		5	5.28	-0.04
			48	5.79	-0.01		6	15.09	0.51
		PCR-C	1	5.62	-0.02	PCR-F	1	6.5	0.03
			2	12.74	0.37		2	5.71	-0.01
			3	11.51	0.30		3	5.24	-0.04
			4	24.83	1.02		4	6.21	0.01
			5	21.89	0.86		5	6.03	0.00
			6	21.04	0.81		6	6	0.00

Table S2 *continued*

		Set 1				Set2			
Mating system	Population code	Plant code	Flower ID	Fruit length (mm)	SC-index	Plant code	Flower ID	Fruit length (mm)	SC-index
SI	PIN	PIN-A	43	4.69	-0.12	PIN-D	1	5.57	-0.02
			44	10.87	0.46		2	7.56	0.09
			45	3.68	-0.22		3	11.19	0.29
			46	4.38	-0.15		4	6.27	0.01
			47	7.08	0.10		5	10.8	0.26
			48	4.11	-0.18		6	11.5	0.30
		PIN-B	37	7.89	0.18	PIN-E	1	5.22	-0.05
			38	4.96	-0.10		2	5.59	-0.03
			39	4.43	-0.15		3	6.38	0.02
			40	4.23	-0.17		4	12.1	0.38
			41	4.52	-0.14		5	5.14	-0.05
			42	3.22	-0.27		6	6.72	0.04
		PIN-C	1	13.78	0.53	PIN-F	1	0	-0.33
			2	8.42	0.17		2	0	-0.33
			3	7.59	0.11		3	0	-0.33
			4	7.6	0.11		4	0	-0.33
			5	10.63	0.32		5	0	-0.33
			6	13.83	0.54		6	0	-0.33
	SBD	SBD-A	19	9.55	0.29	SBD-D	1	7.18	0.06
			20	5.74	-0.02		2	6.92	0.05
			21	4.96	-0.08		5	6.22	0.01
			46	14.59	0.69		6	5.68	-0.02
			47	4.64	-0.11		27	5.41	-0.03
			48	6.01	0.00		28	6.25	0.01
		SBD-B	37	6.12	0.01	SBD-E	1	15.89	0.58
			38	4.17	-0.11		2	8.46	0.14
			39	4.02	-0.12		3	5.08	-0.05
			40	5.33	-0.04		4	6.86	0.05
			41	4.29	-0.11		5	9.57	0.21
			42	5.34	-0.04		6	13.35	0.43
		SBD-C	1	7.48	0.07	SBD-F	1	6.13	0.01
			2	4.57	-0.07		2	5.63	-0.02
			3	4.56	-0.07		3	14.45	0.42
			4	5.17	-0.04		4	6.04	0.00
			5	4.32	-0.08		5	6	0.00
			6	4.81	-0.06		6	10.06	0.20
	TSS	TSS-A	19	4.85	-0.09	TSS-D	1	6.73	0.04
			20	3.06	-0.23		2	6.72	0.04
			21	3.98	-0.16		3	7.09	0.06
			45	4.28	-0.14		4	5.04	-0.05
			46	4.6	-0.11		5	7.12	0.06
			47	3.82	-0.17		6	6.66	0.04
		TSS-B	43	5.12	-0.14	TSS-E	1	5.97	0.00
			44	3.04	-0.48		2	9.88	0.44
			45	3.91	-0.34		3	9.86	0.44
			46	5.08	-0.15		4	12.52	0.74
			47	4.25	-0.29		5	9.71	0.42
			48	0	-0.98		6	6.87	0.10
		TSS-C	1	4.72	-0.12	TSS-F	1	13.86	0.81
			2	8.23	0.21		2	14.19	0.84
			3	10.04	0.38		3	11.29	0.54
			4	10.05	0.38		4	6.94	0.10
			5	4.89	-0.10		5	11.1	0.52
			6	0	-0.56		6	9.09	0.32

Table S2 *continued*

		Set 1				Set2			
Mating system	Population code	Plant code	Flower ID	Fruit length (mm)	SC-index.	Plant code	Flower ID	Fruit length (mm)	SC-index
SC	KTT	KTT-A	1	17.05	0.53	KTT-D	1	25.6	0.95
			2	22.06	0.77		2	25.29	0.93
			3	15.65	0.46		3	24.09	0.87
			4	22.56	0.79		4	21.86	0.77
			5	26.24	0.97		5	26.6	1.00
			6	20.24	0.68		6	26.08	0.97
		KTT-B	1	15.83	0.53	KTT-E	1	33.61	1.05
			2	20.61	0.79		2	32.59	1.01
			3	23.5	0.95		3	33.45	1.04
			4	21.18	0.82		4	32.33	1.00
			5	27.16	1.14		5	32.85	1.02
			6	24.3	0.99		6	28.71	0.86
		KTT-C	1	23.65	0.99	KTT-F	1	NA	NA
			2	19.49	0.76		2	NA	NA
			3	23.5	0.98		3	NA	NA
			4	22.78	0.94		4	NA	NA
			5	14.85	0.50		5	NA	NA
			6	22.49	0.92		6	NA	NA
	LPT	LPT-A	1	31.06	1.01	LPT-D	1	33.8	1.53
			2	27.24	0.86		2	36.36	1.67
			3	36.14	1.21		3	36.05	1.65
			4	35.46	1.19		4	34.23	1.55
			5	25.53	0.79		5	33.98	1.54
			6	NA	NA		6	27.63	1.19
		LPT-B	1	49.39	1.60	LPT-E	1	41.06	1.59
			2	36.96	1.14		2	40.85	1.58
			3	34.04	1.03		3	38.6	1.48
			4	35.24	1.08		4	36.05	1.36
			5	38.67	1.20		5	37.09	1.41
			6	21.73	0.58		6	35.7	1.35
		LPT-C	1	25.75	1.01	LPT-F	1	37.17	1.21
			2	NA	NA		2	34.58	1.11
			3	NA	NA		3	37.9	1.24
			4	NA	NA		4	41.68	1.39
			5	NA	NA		5	40.26	1.33
			6	19.75	0.70		6	36.19	1.17
	PTP	PTP-A	1	20.21	0.96	PTP-D	1	32	1.08
			2	21.46	1.04		2	36.22	1.25
			3	22.68	1.13		3	33.52	1.14
			4	28.78	1.54		4	33.33	1.13
			5	29.62	1.59		5	34.03	1.16
			6	21.79	1.07		6	34.72	1.19
		PTP-B	1	22.86	0.86	PTP-E	1	28.88	0.92
			2	24.67	0.95		2	30.22	0.98
			3	22.35	0.83		3	34.13	1.13
			4	32.96	1.37		4	34.26	1.14
			5	29.74	1.21		5	27.09	0.85
			6	NA	NA		6	28.28	0.90
		PTP-C	1	24.26	0.83	PTP-F	1	30.71	1.05
			2	29.46	1.07		2	30.45	1.04
			3	25.64	0.89		3	32.48	1.13
			4	26.73	0.94		4	32.91	1.15
			5	31.07	1.14		5	30.8	1.06
			22	35.96	1.36		6	34.91	1.23

Table S2 *continued*

		Set 1				Set2			
Mating system	Population code	Plant code	Flower ID	Fruit length (mm)	SC-index	Plant code	Flower ID	Fruit length (mm)	SC-index.
SC	RON	RON-A	1	21.56	0.94	RON-D	1	29.03	1.09
			2	21.58	0.94		2	28.57	1.07
			3	17.43	0.69		3	26.54	0.97
			4	NA	NA		4	30	1.13
			5	25.63	1.18		5	22.84	0.80
			6	19.36	0.80		6	25.48	0.92
		RON-B	1	31.02	1.10	RON-E	1	28.78	1.36
			2	33.22	1.19		2	30.5	1.46
			3	33.34	1.20		3	24.42	1.10
			4	32.45	1.16		4	28.7	1.35
			5	31.18	1.11		5	26.21	1.20
			6	NA	NA		6	27.36	1.27
		RON-C	1	24.32	1.45	RON-F	1	26.53	1.03
			2	19.48	1.07		2	20.54	0.73
			3	18.4	0.98		3	28.92	1.15
			4	22.11	1.28		4	20.4	0.72
			5	20.5	1.15		5	26.86	1.05
			6	16.39	0.82		6	26.64	1.04
	TC	TC-A	1	16.67	0.74	TC-D	1	30.12	1.00
			2	22.61	1.15		2	32.61	1.10
			3	25.75	1.37		3	27.62	0.89
			4	28.82	1.58		4	28.09	0.91
			5	26.12	1.39		5	28.24	0.92
			6	17.98	0.83		6	29.25	0.96
		TC-B	1	28.89	1.08	TC-E	1	24.82	0.76
			2	29.75	1.12		2	27.03	0.85
			3	22.54	0.78		3	33.23	1.10
			4	21.11	0.71		4	32.79	1.08
			5	18.95	0.61		5	36.63	1.24
			6	32.16	1.24		6	30.85	1.00
		TC-C	1	36.57	1.19	TC-F	1	29.42	0.76
			2	27.14	0.82		2	34.19	0.91
			3	18.28	0.48		3	33.53	0.89
			4	30.99	0.97		4	39.31	1.08
			5	38.82	1.27		5	37.35	1.01
			6	28.18	0.86		6	43.13	1.20
	TSSA	TSSA-A	1	15.12	0.88	TSSA-D	1	13.5	0.49
			2	15.74	0.94		2	13.97	0.52
			3	17.35	1.10		3	9.05	0.20
			4	19.01	1.26		4	8.07	0.14
			5	22.24	1.57		5	13.5	0.49
			6	7.18	0.11		6	0	-0.39
		TSSA-B	1	15.14	1.10		36	9.13	0.21
			2	15.41	1.13		37	8.1	0.14
			3	14.66	1.04		38	10.7	0.31
			4	7.56	0.19	TSSA-E	1	12.61	0.54
			5	12.43	0.77		2	21.48	1.25
			6	12.89	0.83		3	20.56	1.18
			7	16.85	1.30		4	14.45	0.68
			8	9.76	0.45		5	13.82	0.63
			9	10.86	0.58		6	14.03	0.65
		TSSA-C	1	16.03	0.95	TSSA-F	1	12.18	0.54
			2	13.76	0.74		2	12.35	0.55
			3	12.75	0.64		3	13.56	0.66
			4	17.01	1.04		4	18.15	1.05
			5	NA	NA		5	17.55	1.00
			19	20.45	1.37		6	18.65	1.10

Note: Population codes correspond to the ones in Table S1.

Table S3 Timeline and progeny numbers for eight batches of cross-progeny

Source	Batch number								Totals
	1	2	3	4	5	[a] 6	[b] 7	[a] 8	
Time schedule									
Sowing month	7.2015	10.2015	9.2016	2.2017	8.2017	8.2017	1.2018	2.2018	-
First pollination	9.2015	1.2016	11.2016	4.2017	10.2017	10.2017	4.2018	4.2018	-
Last pollination	3.2016	7.2016	4.2017	9.2017	4.2018	4.2018	8.2018	9.2018	-
Seed and germination									
Parental set	A, B, C	A, B, C	D, E, F	D, E, F	A, B, C	A, B, C, D, E, F	A, B, C	A, B, C	-
Number of families with at least one seed	446	446	452	452	430	211	269	204	2,437
Number of germinated seeds	419	427	204	206	326	155	164	138	2,039
Self-Pollination									
Number of plants that could be self-pollinated	276	196	202	206	319	145	153	110	1,607
Total number of self-pollinations (in principle, 10 self-pollinations per plant)	2,564	1,855	2,045	2,137	3,247	1,503	1,483	1,146	15,980
Sterility testing									
Number of tested plants	116	62	95	85	-	-	-	-	358
Fertile	114	55	87	75	-	-	-	-	331 (93.0%)
Completely sterile	0	4	3	3	-	-	-	-	10 (2.8%)
Only female sterile	0	0	5	7	-	-	-	-	12 (3.4%)
Only male sterile	2	3	0	0	-	-	-	-	5 (1.4%)

Notes: In each batch, we sowed one seed per seed-family

[a] This batch only included progeny from between-breeding-system crosses.

[b] This batch included progeny from crosses between-breeding-system and within-population, but not between populations with the same breeding system.

Table S4 Linear mixed model analysis to test the effect of maternal and paternal fruit length and their interaction on progeny fruit length after self-pollination in selfing populations

Source	Progeny fruit length after selfing (Gaussian)				
Fixed effects	df	Estimate	SE	χ^2	P
Maternal fruit length after selfing (MFL)	1	0.32	0.22	**4.97**	**0.026**
Paternal fruit length after selfing (PFL)	1	0.32	0.22	**8.50**	**0.004**
MFL:PFL	1	-0.00	0.01	0.10	0.756
Random effects			SD		
Maternal population	1		1.59		
Paternal population	1		1.02		
Maternal plant (nested in Maternal population)	1		1.70		
Paternal plant (nested in Paternal population)	1		1.40		
Batch code	1		0.84		
Residual	219		4.22		

Notes: We tested the significance of the fixed effects by comparing models with and without the tested term using likelihood-ratio tests. Significant effects are indicated in bold. We individually calculated fruit length of progeny plants and parental plants based on the median fruit length of multiple fruits per plant prior to analysis.

Table S5 Breeding system phenotype frequencies based on SC-index values for cross-progeny for each cross-type

Cross-type	Breeding system phenotype			Totals
	Self-compatible (SC) (SC-index ≥ 0.75)	Self-incompatible (SI) (SC-index ≤ 0.25)	Not phenotyped (0.25 < SC-index < 0.75)	
Crosses within breeding system, within population (WP) and between population (BP)				
$WP_{♀SI \times ♂SI}$	9	51	19	79
$BP_{♀SI \times ♂SI}$	8	171	36	215
$WP_{♀SC \times ♂SC}$	80	0	6	86
$BP_{♀SC \times ♂SC}$	233	1	8	242
Crosses between breeding systems (BBS)				
For crosses involving all six selfing populations				
$BBS_{♀SI \times ♂SC}$	216	219	56	491
$BBS_{♀SC \times ♂SI}$	261	163	43	467
BBS_{total}	**477**	**382**	**99**	**958**
Subset of crosses involving SC populations fixed for S_1 (RON and PTP)				
$BBS_{♀SI \times ♂SC}$	48	98	24	170
$BBS_{♀SC \times ♂SI}$	67	87	24	178
BBS_{total}	**115**	**185**	**48**	**348**
Subset of crosses involving SC populations fixed for S_{19} (TC, TSSA LPT ,and KTT)				
$BBS_{♀SI \times ♂SC}$	168	122	31	321
$BBS_{♀SC \times ♂SI}$	194	78	17	289
BBS_{total}	**362**	**200**	**48**	**610**

Notes: Plants with an intermediate SC-index (0.25 < SC-index < 0.75) could not be phenotyped unambiguously.

Fig. S1 Relationship between fruit length (mm) and seed number per fruit by maternal population. For the subset of the fourth and fifth two batches, we counted seeds per fruit for self-pollinated fruits, to test whether fruit length was a good predictor of seed set. Population codes correspond to those in Table S1). Best-fit regression lines are shown along with their equation and proportion of explained variance R^2.

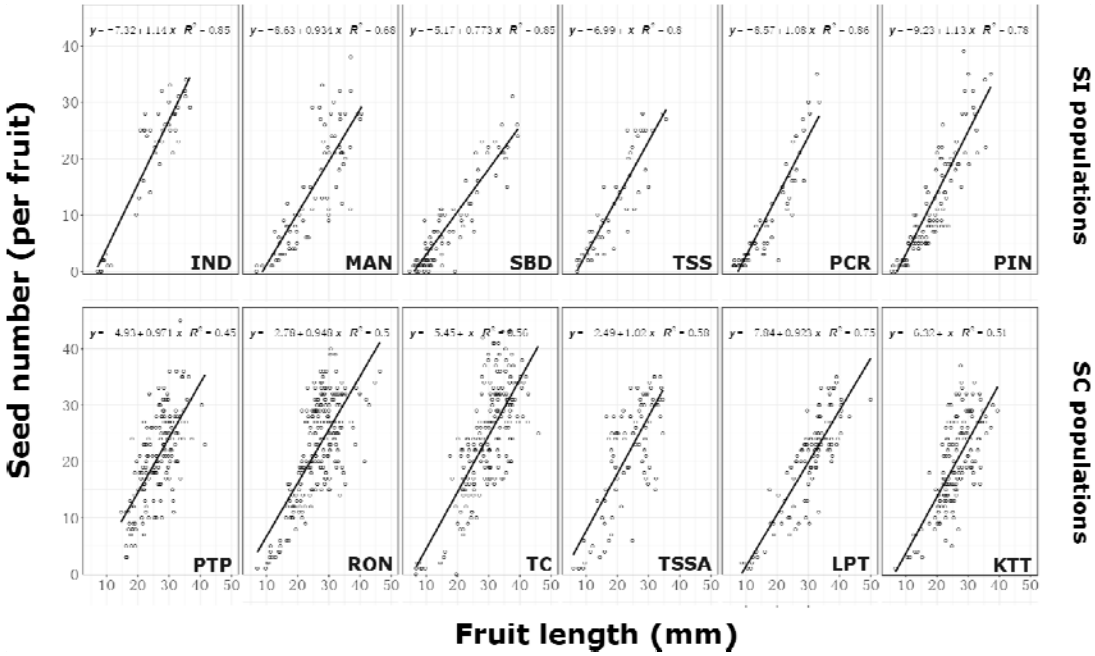


Fig. S2 SC-index values for individual cross-progeny by seed family.

Crosses between populations were always done between parents with the same plant code (i.e. plants coded A were always crossed with another plant coded A, plants coded B with other plants coded B, et cetera, cf. crossing design in Fig. 1). Population coding corresponds to the coding in Table S1. Diagonal cells marked with an extra outline represent progeny from within-population crosses which resulted from crosses among plants coded A, B and C and among plants coded D, E and F.

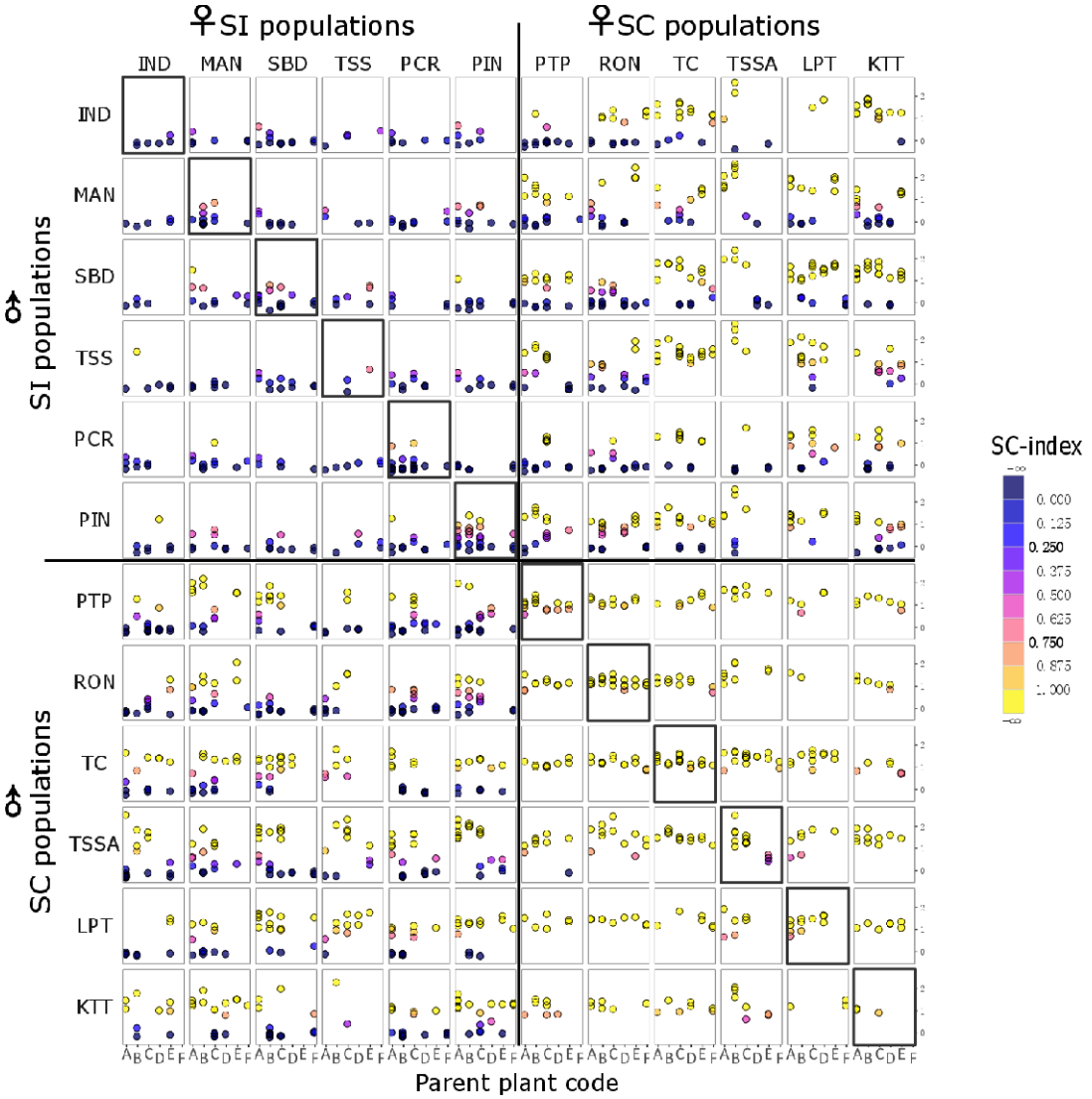
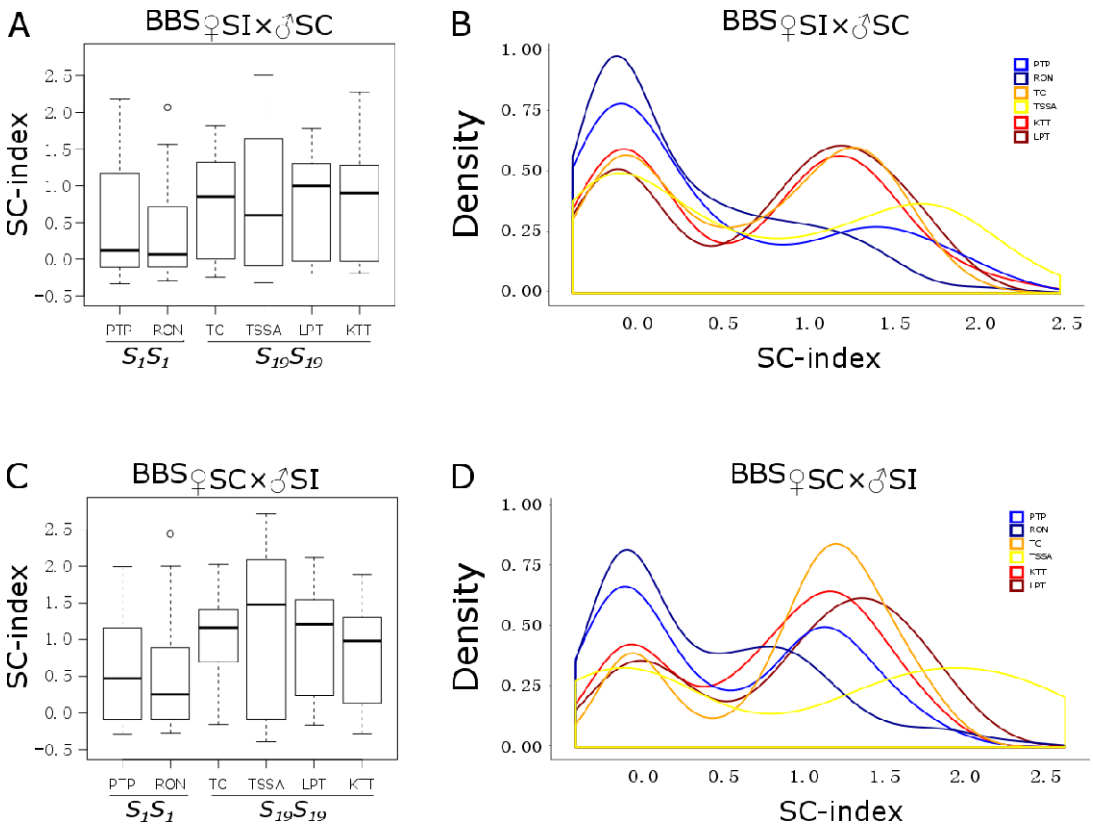


Fig. S3 SC-index values for progeny from crosses between breeding systems for the six SC populations grouped by *S*-haplotype.

Box plots (A, C) show the interquartile range (IQR) around the median (black horizontal line), with whiskers extending to 1.5×IQR or to the lowest/highest data point within 1.5×IQR. Points beyond whiskers' limits represent outliers. Density plots (B, D). The *S*-haplotypes of SC parents from PTP and RON are S_1S_1 homozygotes, and from the other four populations $S_{19}S_{19}$.