IDENTIFICATION OF CANDIDATE GENES FOR SELF-COMPATIBILITY IN A DIPLOID POPULATION OF POTATO DERIVED FROM PARENTS USED IN GENOME SEQUENCING

Brenda Elaine Arnold

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Horticulture

Richard Veilleux, Chairman
Jason Holliday
Joel Shuman

September 9, 2013
Blacksburg, VA

Keywords: Solanaceae, self-incompatibility, *S*-locus, small nucleotide polymorphism (SNP)

IDENTIFICATION OF CANDIDATE GENES FOR SELF-COMPATIBILITY IN A DIPLOID POPULATION OF POTATO DERIVED FROM PARENTS USED IN GENOME SEQUENCING

Brenda Elaine Arnold

Abstract

Gametophytic self-incompatibility limits the ability to derive inbred lines of potato through self-pollination and is prevalent in diploid potato. Within a population of F_1 hybrids between two genotypes used in potato genome sequencing, we observed fruit set on many greenhouse-grown plants. Subsequently, after controlled self-pollinations, we confirmed fruit set in 32 of 103 F_1 plants. Our goal was to identify genes responsible for self-compatibility in this population and to advance selfed progeny to develop highly homozygous inbred lines. The F_1 population was genotyped using a single nucleotide polymorphism (SNP) array. Polymorphic and robust SNPs were analyzed by Fisher's Exact Test to identify allelic states segregating with the self-compatible phenotype. Filtering 1966 SNPs to retain only those with p-values less than 0.0001 yielded 95 highly significant SNPs, with all SNPs on anchored scaffolds located on chromosome 12. Candidate genes encoding for multiple notable proteins including an S-protein homologue were identified near highly significant SNPs on the Potato Genome Browser. Seeds obtained after self-pollination of self-compatible individuals were used to advance the population for three generations. SNP chip genotyping of the S_3 generation revealed entirely different SNPs segregating for self-compatibility on nine different chromosomes. Comparison of the allelic state of SNPs in the F_1 and S_3 generations revealed a heterozygosity reduction by 80%, with fixation of many SNPs including those surrounding the S-protein homologue. We conclude that the genes responsible for segregation of self-compatibility in the S_3 generation are different from those in the F_1 generation.

Acknowledgments
I would like to thank the following people for their role in supporting me through my graduate research and studies:

- My advisor Dr. Veilleux for his inspiration, motivation, and patience. He has spent a great amount of time meeting with me to discuss my project and share ideas. He has my respect and I appreciate the open communication he offers all of his students.
- My committee members, professors, and professional colleagues for their encouragement, input, and constructive criticism.
- My lab team for their support and positive attitude.
- Jeff Burr for his advice and help at the greenhouse.
- The Horticulture Department along with the Latham community for support and inspiration.
- Collaborators at Michigan State University for their contributions.
- Jonathan Stallings and Jing Niu from the Laboratory for Interdisciplinary Statistical Analysis for their statistical expertise and contributions.
- The National Science Foundation and the College of Agriculture and Life Sciences at Virginia Tech that have funded my research and made it financially possible for me to achieve my educational goals.

Table of contents

List of figures

List of tables

Chapter 1. Introduction

The DRH population of potato

Historically, potato has been a challenging crop to study at the genetic level due to its ploidy level (2*n*=4*x*=48) and high heterozygosity. However, despite these obstacles the potato genome was recently sequenced using two diploid genotypes, *S. tuberosum* Group Phureja DM 1-3 516 R44 (DM) and *S. tuberosum* group Tuberosum RH89-039-16 (RH) (The Potato Genome Sequencing Consortium 2011). DM is a homozygous doubled monoploid (2*n*=2*x*=24) derived from embryos from anther culture of a *Solanum tuberosum* Group Phureja accession (Paz and Veilleux 1999). RH is a heterozygous diploid that resulted from a cross between a *S. tuberosum* dihaploid and *S. tuberosum* x *S. tuberosum* Group Phureja hybrids (Van Os et al. 2006). In sequencing the genome DM helped to simply the project with its homozygosity while RH added relevancy as it represented cultivated potato (The Potato Genome Sequencing Consortium 2011). In contrast to the slow progress that potato researchers have struggled with for years, the new genomic and transcriptomic tools available through the sequencing of DM and RH provide great opportunities for exploring this important agricultural staple. Furthermore, thousands of molecular markers have recently become available through the creation of the Infinium 8303 Potato Array genotyping single nucleotide polymorphisms (SNPs) (Hamilton et al. 2011; Solanaceae Coordinated Agricultural Project). The male-fertility of RH was used to take advantage of the opportunity to cross DM with RH to generate the DRH segregating population that can be used for studies to better understand potato genetics (Felcher et al. 2012). Indeed the utility of the DRH population has already been demonstrated through the construction of linkages maps highlighting regions of segregation distortion (Felcher et al. 2012).

An introduction to self-incompatibility

Self-compatibility can be defined as the ability of a fertile plant, upon pollinating itself, to form zygotes (de Nettancourt 1977). As with any other biological mechanism, plants exhibit variation in methods of determining compatibility preferences. As a generalization, plants take advantage of two alternative mechanisms for determining self-incompatibility: gametophytic or sporophytic self-incompatibility. Pollen identities in sporophytic self-incompatibility systems are controlled by the genotype of the plant producing

the pollen while gametophytic self-incompatibility systems are controlled by the genotype of the microspores. Sporophytic self-incompatibility occurs in Brassicaceae, whereas gametophytic self-incompatibility is distributed in Solanaceae, Rosaceae, and Scrophulariaceae (Takayama and Isogai 2005). Aside from elementary classifications regarding sporophytic and gametophytic self-incompatibility, there remain numerous other specifications to aid in categorizing different mechanisms for determining self-incompatibility. For instance, if the floral structure of a plant influences the self-incompatibility response, then that plant is labeled heteromorphic, whereas if floral structure does not alter the self-incompatibility response that plant is regarded as homomorphic (Lewis 1949). The number of S-loci present within a population of self-incompatible plants also aid in characterization, with monofactorial, bi-factorial, and polyfactorial classifications indicating increasing number of S-loci, respectively. In general, members of Solanaceae exhibit a homomorphic, monofactorial, gametophytic system, with some exceptions (de Nettancourt 1977).

Self-incompatibility is a fundamental trait to analyze because the level of heterozygosity of the progeny is directly impacted by whether the parent crosses with divergent, familiar, or identical (self) genotypes. Self-compatible plants favor inbreeding and increased homozygosity, while self-incompatible plants favor out-crossing and increased heterozygosity.

Unlike many tetraploid cultivars which are self-compatible, diploid potato is frequently limited by self-incompatibility (Cipar et al. 1964). Despite the fact that tetraploid potato has historically been used for cultivation due to relatively superior vigor, diploid potato breeding schemes are much easier to manipulate (Chase 1963). As with many other members of Solanaceae, diploid potato self-incompatibility is determined by a gametophytic self incompatibility system (Cipar et al. 1964; Pandey 1962). The inability of diploid potato to successfully self has created a substantial roadblock to developing stocks and tester lines which could be valuable tools to breeders and researchers alike (Jacobs et al. 1995). Since the high level of heterozygosity in potato allows for deleterious recessive alleles of various genes to be passed on from generation to generation, it would be beneficial to generate homozygous germplasm devoid of these harmful alleles. Advancement of homozygosity in potato is more quickly achieved by selfing diploids than tetraploids (Bartlett and Haldane 1934).

Traditionally, potato is propagated vegetatively; however efforts to develop true potato seed programs have demonstrated that sexual propagation is a potential alternative. At the International Potato Center [Centro Internacional de la Papa (CIP)], multiple breeding programs have been attempted using true potato seed (Pallais 1991). One important advantage of propagating potato via seeds instead of tubers is that most viruses which typically are destructively transmitted through tubers are not passed on from mother plants to progeny when sexually propagated seeds are used (Pallais 1991). If produced in a reliable method, seed propagation could potentially be more affordable than tuber propagation due to lower rates of infection (Simmonds 1997). Recommendations to select for low seed dormancy and high seed vigor have been suggested to promote the applicability of true potato seed breeding (Pallais 1991). Other benefits of a seed propagation program are that potato seeds take up less room than tubers and are easier to store due to less strict postharvest requirements. Given that diploids are typically self-incompatible, cross pollinations to produce seed are more common (Simmonds 1997). However, in order to maintain homozygous lines by a sexual propagation program it is necessary to work with self-compatible genotypes to avoid recombination events that promote heterozygosity.

Self-incompatibility in nature

The evolution of plants in the wild can be greatly influenced by self-incompatibility as it plays an important role in determining mate availability. When certain mate limiting situations arise, such as during colonization periods or after catastrophic disasters, self-compatibility creates an advantage in allowing plants to reproduce when sexual partners may be unavailable. Self-compatibility is also advantageous when pollinators such as birds or insects are not readily available to carry pollen from one location to another. Pollinators can be limited by factors such as population density, time of flowering, extreme temperatures, and high altitudes (Lloyd 1992). However, once plants are well established self-compatibility is less desirable as it encourages inbreeding which frequently results in inferior phenotypes due to inbreeding depression. In this case, self-incompatibility encourages cross pollination, leading to more recombination and increased variation to promote fitness. Given that both self-incompatibility and self-compatibility exhibit unique advantages it is difficult to predict the compatibility state from which plants originally evolved. Even as molecular research progresses and more models are developed to explain self-incompatibility, it is still not clear what the default condition is (McClure et al. 2011). Self-

compatibility has been suggested to lead to an evolutionary dead end due to the build-up of deleterious alleles and challenges to adaptation (Igic and Busch 2013; Takebayashi and Morrell 2001). Despite the drawbacks that selfing presents, cases have nevertheless been proposed in which there are still opportunities for those self-compatible phenotypes to revert to self-incompatible phenotypes. Sometimes the case may be that these self-incompatible and self-compatible phenotypes are manifestations of pseudo-self-compatibility, as will be discussed later.

Self-incompatibility in breeding and research

Through years of breeding different crops, humans have selected for self-incompatibility and self-compatibility in different situations. Inbred lines, which are useful for achieving homozygosity to select for alleles for particular traits of interest, are most quickly achieved through self-pollinations. Inbred lines are often established for heterosis breeding programs in which it is also necessary to maintain the capacity to cross pollinate. Recently, self-compatible diploids were generated to successfully develop an F_1 hybrid potato breeding scheme (Lindhout et al. 2011). The ability to manipulate which genotypes a plant can or cannot cross is a powerful tool for developing novel and improved varieties.

Inbreeding depression and heterosis

Inbreeding depression is the weakening of the phenotype of the progeny relative to the phenotypes of the parents due to relatedness of the parents. As studied in diploid potato, selfing may lead to inbreeding depression due to fixation of methylation sites which suppress gene expression (Nakamura and Hosaka 2010). According to de Nettancourt (1977), inbreeding can be used to select for self-compatibility but this phenotype will not be permanently maintained possibly due to a need for a threshold level of heterozygosity. Indicators of inbreeding depression in potato include but are not limited to, lack of flower buds, bud dropping, no pollen shed, low seed set, and male sterility (Birhman and Hosaka 2000). Inbreeding depression in diploid potato has previously been noted to be most severe in the first selfed (S_1) generation compared to the second (S_2) and third (S_3) selfed generations (De Jong and Rowe 1971). A study of true potato seed propagation comparing 4x-2x hybrid, first and second generation open-pollinated (OP), and first and second generation selfed families revealed that inbreeding depression plays a significant role in reduction of OP fruit set, flowering, vigor, and yield (Shonnard and Peloquin 1991).

Inbreeding depression in tetraploid *Solanum tuberosum* Group Andigena has been shown to affect fruit and seed set with many that resulted in reduced pollen production and viability (Golmirzaie et al. 1998).

While inbreeding depression results in weaker phenotypes with successive generations, heterosis results in superior phenotypes than the parents. Heterosis, or hybrid vigor, describes the increased fitness of the progeny resulting from the recombination of genes due to the crossing of different parents. Genetically diverse inbred lines are intercrossed in heterosis breeding schemes with the expectation that the progeny of the cross will be even stronger than the inbred lines used to create that cross.

The *S* locus

The ability of pollen and pistil specific proteins to recognize each other controls the compatibility condition within gametophytic systems (de Nettancourt 1977; East and Mangelsdorf 1925; McClure et al. 2011). Genes encoding for pollen and pistil recognition factors are found at the same multiallelic locus, the *S*-locus. According to research in a variety of families, the *S*-locus has the potential to span more than 100 kb (Lai et al. 2002). In potato the *S*-locus is found on chromosome 1, which is consistent with studies identifying *S*-locus position in *Lycopersicon* (Gebhardt et al. 1991; Jacobs et al. 1995). In addition to determining self-incompatibility, *S*-alleles have also been implicated in playing a role in unilateral incompatibility, that is the ability of an intercross to succeed in only one direction (Eijlander et al. 2000). Indeed the *S*-locus has been regarded as such an important subject that markers for these alleles have been deemed classical genetic markers in potato (Jacobs et al. 1995).

Proteins encoded by genes at the *S*-locus exhibit a certain degree of specificity allowing the plant to distinguish between self and cross pollinations (Bredemeijer and Blaas 1981). Prior to full functional and biochemical characterization of these genes and the proteins they encode, much research was devoted to exploring *S*-allele variation in order to uncover the basis of *S*-locus specificity. Polymorphism of the *S*-locus is believed to be an ancient event, as even within the same species, *S. chacoense*, only about 60% homology between two *S*-alleles was observed (Xu et al. 1990). As a broad generalization from research of homology within *Solanum tuberosum*, at least three different proposed classes of *S* alleles have evolved maintaining regions of conserved and variable domains (Kaufmann et al. 1991). At least 41 conserved regions in *S* alleles have been identified among species of Solanaceae such as *Nicotiana*

alata, *Petunia inflata*, and *Solanum chacoense* (Ioerger et al. 1991). In contrast, two hypervariable regions, HVa and HVb, consist of variable residues postulated to contribute the specificity necessary for pollen recognition (Ioerger et al. 1991). The variation necessary for conferring specificity of *S*-alleles requires only small dissimilarities, as sensitivity to differences as subtle as only ten amino acid changes has been observed (Saba-El-Leil et al. 1994).

Given that differences in the hypervariable regions of the *S*-locus help determine which genotypes a plant can cross with, it is important to understand how these variations originate. New *S*-alleles are proposed to evolve from mutations at a site on the *S*-locus known as the specificity segment (de Nettancourt 1977). According to a tripartite structure model this specificity segment is one of three other segments that are linked, including a pollen activity segment and a stylar activity segment (de Nettancourt 1977; Lewis 1960). Experimental radiation of the *S*-locus has shed light on the effect of mutations of *S* alleles (de Nettancourt 1977). These radiation-induced mutations could induce either permanent or reversible mutations, implying that sometimes a stable allele could be mutated into an unstable allele (Lewis 1951). Through studying artificially induced mutations we gain a better insight into the effect that natural mutations have in shaping populations by altering compatibility status.

Pistil specific components

The pistil specific component of the *S*-locus has been identified as a ribonuclease and appropriately labeled S-RNase (McClure et al. 1989). Support for S-RNase activity in Solanaceae was further provided by identifying shared conserved residues with fungal ribonucleases (Ioerger et al. 1991). A cytotoxic model was developed in *Nicotiana alata* in which the S-RNase activity of the stylar *S*-alleles degrades pollen rRNA in pollen tubes growing inside the style of self-incompatible systems (McClure et al. 1990). Based on *in-vitro* studies it was determined that S-RNases were able to enter pollen tubes intact (Gray et al. 1991). By creating mutants to eliminate the ribonuclease activity of S-RNase Huang (1994) demonstrated that the ribonuclease activity indeed was necessary for pollen rejection and thus a self-incompatible response.

Pollen specific components

In the efforts to classify the components of gametophytic self-incompatibility it was not until over 10 years after the discovery of the pistil specific component S-RNase that the complementary pollen component was characterized. The lack of pollen *S*-allele deletions implied that pollen viability relies upon the pollen *S*-gene (Golz et al. 2001). The pollen determinant is tightly linked to the *S*-locus and has been identified as an F-box protein now commonly referred to as SLF (*S*-locus F-box) (Lai et al. 2002). The association of SLF with the previously indentified S-RNases helps to complete the picture, at least on the most basic level, of gametophytic self-incompatibility (Sijacic et al. 2004).

SLFs belong to the broader protein family of F-box proteins. These F-box proteins play a critical role in the cell cycle by facilitating degradation via ubiquitin ligase (Chen et al. 2012). Phosphorylation of serine and threonine residues allows the marking of proteins for ubiquitin mediated degradation (Craig and Tyers 1999). Ubiquitin degradation is carried out by three integral features: E1 a ubiquitin activating enzyme, E2 a ubiquitin conjugating enzyme, and E3 a ubiquitin protein ligase (Hershko et al. 1983). One important class of E3 protein ligases is the SCF complex that is comprised of Skp1p, Cdc53p (cullin), and an F-box protein (Patton et al. 1998). According to the F-box hypothesis, the protein Skp1p functions to bring together different kinds of F-box proteins to aid in the degradation process (Patton et al. 1998). The Skp1 gene was previously identified as using the F-box motif to aid in cell cycle regulation through proteolysis (Bai et al. 1996). Protein-protein interaction domains such as WD40 repeats and leucine rich repeats (LLR) are common elements of F-box proteins and are used for gathering substrates to the SCF complex (Craig and Tyers 1999). While SLF is proposed to form SCF complexes, SCF complexes are not limited to self-incompatibility functions and are used by organisms to regulate a wide variety of pathways (Craig and Tyers 1999; Qiao et al. 2004). In fact, in arabidopsis nearly 700 F-box proteins have been identified and it has been claimed to be one of the largest super-families in plants (Gagne et al. 2002). Following this trend, numerous SLF-like genes have been identified in Solanaceae as well (Wheeler and Newbigin 2007).

As applied to the gametophytic self-incompatibility system, SLF targets S-RNase for degradation, thereby preventing the S-RNase protein from degrading the formation of pollen tubes. The compatibility condition of a plant depends on whether or not SLF and S-RNase interact with each other. This

interaction is influenced by preferential attraction of SLF to non-self S-RNases (Hua and Kao 2006). Since there are many types of SLF genes present, it is possible that certain SLF genes can bind to more than one type of S-RNase (Kubo et al. 2010). One model suggests that there are three domains of the SLF gene (Hua et al. 2007). In this model one of the domains, FD2, encourages strong interactions between non-self SLF and S-RNases, while the other two domains, FD1 and FD3, serve as negative regulators. These domains are known as S-RNase binding domains (SBD) and S-RNase binding regulating domains (SBRD), respectively (Hua et al. 2007).

Competitive interaction

Since gametophytic self-incompatibility is based on recognition between pollen and pistil specific components the number of alleles at these loci has the potential to alter the compatibility condition. When the haploid state (pollen) of an organism has the capacity to be heterozygous, as is the case for tetraploids but not diploids, it is possible to code for two alternative recognition factors to encounter the pistil produced S-RNase. This phenomenon, in which self-incompatibility is defeated by increasing the ploidy level of a diploid to a tetraploid, is known as the heteroallelic pollen effect (de Nettancourt 1977; McClure et al. 2011). Competitive interaction was further confirmed by experimentation of induced mutations of the pollen specific S-allele which demonstrated a correlation between allele duplication and induced self-compatibility (Golz et al. 1999).

Modifier genes

In addition to genes found at the S locus there are also modifier genes found elsewhere throughout the genome. Modifier genes are those genes which do not serve a purpose in the recognition process but which do have the capacity to alter the compatibility condition (Tao and Iezzoni Scientia Horticulturae2010). Upon continuous selfing these modifier genes experience segregation distortion as the self-compatibility phenotype is positively selected for (de Nettancourt 1977).

One example of an important modifier gene is HT-B which encodes for proteins in the pistil. While HT exhibits two alternative types, A and B, only B is associated with altering the compatibility phenotype (O'Brien et al. 2002). In some models of gametophytic self-incompatibility the S-RNases that enter pollen tubes are compartmentalized into vacuoles to prevent interaction with the pollen tube RNA. In order for

the S-RNases to encounter the pollen tube RNA it is necessary for these vacuoles to somehow be degraded. Under the self-incompatible condition HT-B is able to disrupt these S-RNase containing vacuoles and thereby mediate pollen tube degradation (Goldraij et al. 2006). The degradation of HT-B itself may therefore result in self-compatibility as it hinders S-RNase from escaping compartmentalization (Goldraij et al. 2006). Although the degradation of HT-B has been suggested to aid in self-compatibility it is not necessarily required to achieve this phenotype (McClure et al. 2011). In some instances partial self-compatibility can be obtained by low expression of HT-B (Puerta et al. 2009).

Another modifier gene encodes for the SBP1 protein, which interacts with both SLF and S-RNase (Hua and Kao 2006). Early work in Petunia revealed that when SBP1 binds with S-RNase it elicits a self-incompatibility response (Sims and Ordanic 2001). Further work in *S. chacoense* involving yeast-two-hybrid assays revealed an SBP1 RING-finger protein which interacted specifically with the HV portion of S-RNase (O'Brien et al. 2004).

Yet another example of a modifier gene is a dominant switch gene that was identified in tomato that, when a corresponding dominant *S*-allele is present, allows the self-incompatible phenotype to be expressed (Martin 1968). It has been proposed that a mutation in this switch gene may aid the evolution from the self-incompatible to the self-compatible condition (Martin 1968).

Finally, probably one of the most well studied modifier genes, at least in potato, is the *S*-locus inhibitor (or *Sli*) gene identified in a mutant *S. chacoense* genotype (Hosaka and Hanneman 1998a). The *Sli* gene acts sporophytically to induce self-compatibility, that is to say, the genotype of the pollen donor parent determines the compatibility condition (Hosaka and Hanneman 1998a). Located on chromosome 12, the *Sli* gene is nonallelic to the *S*-locus and functions to block the pollen *S*-allele interaction (Hosaka and Hanneman 1998b).

Research efforts on the *Sli* gene have significantly advanced the study of how to manipulate self-incompatibility. As *S. chacoense* is a wild diploid weed, the *Sli* gene is limited in usefulness to breeders unless considerable backcrossing is performed to return to cultivated types. Multiple efforts have been made to breed the *Sli* gene into self-incompatible stock to confer self-compatibility (Birhman and Hosaka 2000; Lindhout et al. 2011; Phumichai et al. 2005). For example, the *Sli* gene was bred into a cultivated diploid population consisting of a mixture of *S. tuberosum* Groups Phureja and Stenotomum germplasm

in order to confer self-compatibility (Birhman and Hosaka 2000). Self pollinations on individuals from this same population were used to reach the S_5 generation with 90% homozygosity (Phumichai et al. 2005). While observing a slower decline in heterozygosity than expected, Phumichai (2005) did not find specific heterozygous loci which implied that heterozygosity may not necessarily be required for a plant to be able to survive and maintain self-compatible potential. This finding helped to dispel previous notions that homozygosity of *Sli*, as well as homozygosity of the *S*-locus in general, was deleterious due to linkage with disadvantageous recessive genes (Hosaka and Hanneman 1998a; Simmonds 1966). However, complete homozygosity through inbreeding remains an unachieved goal that begs further work to be done before conclusive results can be made. More recently Lindhout et al. (2011) used the *Sli* gene bred into elite diploid germplasm to prove the feasibility of an F_1 hybrid potato breeding program which selects for self-compatibility and tuber quality at the same time.

Pseudo-self-compatibility

Incomplete self-incompatibility is sometimes referred to as pseudo self-compatibility. Pseudo self-compatible pollen is less likely to succeed in fertilization compared to self-compatible pollen, yet given the opportunity it can lead to different levels of seed set (Lloyd and Schoen 1992). While it was previously proposed that pseudo-self compatibility was due to different numbers of S alleles present, research has also indicated that modifier genes may play a role in pseudo-self compatibility (de Nettancourt 1977; Levin 1996). In the evolution from self-incompatibility to self-compatibility, pseudo-self compatibility can exist as an evolutionary intermediate phenotype (Levin 1996).

Seed versus fruit set

When breeding for self-compatibility in order to increase homozygosity it is most advantageous to use plants that produce many fruits with high seed set in order to increase the chance of selecting progeny that have inherited the mechanisms necessary to overcome incompatibility. Both of these traits, fruit set and seed set, are influenced by how well a plant takes advantage of self-compatibility and also by how sensitive it is to the effects of inbreeding depression. Both fruit and seed share similar constraints at the nutrient, hormonal, transcriptional, and metabolic levels (Ruan et al. 2012). Regardless of these shared elements of development it should be noted that seed development more heavily relies on the success of

pollinations given that parthenocarpic fruit occasionally develops without the occurrence of pollination (Ruan et al. 2012). Since it is possible for empty fruits to form without viable seeds inside, one should not assume that fruit set alone is a suitable indicator of degree of self-compatibility.

An important feature that regulates seed set in many plants is a class of proteins known as cyclin dependent kinase (CDK) inhibitors. CDK inhibitors can alter nuclear division during pollen development which in turn affects seed set (Zhou et al. 2002). Through valuable research performed on arabidopsis, Iwakawa et al. (2006) discovered a specific cyclin-dependent kinase, CDKA; 1 which encouraged cell division during development of the male gametophyte, embryo, and endosperm. In particular CDKA; 1 is responsible for cell division in the generative (as opposed to the vegetative) cell. The importance of CDKA; 1 was demonstrated by studying a loss of function mutant which yielded smaller seed size than normal. CDKs may also regulate fruit development in addition to seed development as a tomato WEE1 kinase, which negatively regulates CDK, plays a role in determining cell size in fruit development (Gonzalez et al. 2007).

Environmental factors that influence self-incompatibility

In addition to the alleles at the *S*-locus and modifier genes environmental factors play a role in determining self-incompatibility. As is the case with many other studies of biological systems, the variety of environmental factors that alter compatibility complicates efforts to elucidate the genetic components that control this trait. Studies of diploid *S. tuberosum* dihaploid genotypes indicated that the self-incompatibility barrier may not be fully functional in early floral stages due to the fact that bud pollinations on expected self-incompatible plants occasionally yielded fruit with low seed set (Eijlander et al. 1997). On the other hand, research in the weedy species *Solanum carolinense* revealed that self-pollen tube growth was greater in older flowers, suggesting a weakening of self-incompatibility with increased age (Stephenson et al. 2003).

Similar to floral age, the time of season during which pollinations are made also can influence the compatibility condition. For instance, self-incompatibility has been reported to break down at the end of the flowering season in some *Nicotiana* and *Petunia* species (East 1934; Yasuda 1934). The presence of OP fruit set on plants which previously were resistant to fruit set upon hand pollination further suggests the possibility that the self-incompatible response changes over time (Stone et al. 2006).

In addition self-incompatibility may be altered by stigma size and pollen moisture which influence if pollinations are successful (de Nettancourt 1977). Temperatures between 32 and 60°C have the potential to break down self-incompatibility (de Nettancourt 1977). Still other factors include light, weather, and population density (Lloyd and Schoen 1992). Due to varying conditions, autonomous fruit set in nature is not entirely based on the genetic predisposition to self-incompatibility (Lloyd and Schoen 1992).

Measuring self-incompatibility

As self-incompatibility has been a popular topic of research over the years many different methods have been employed for detecting and measuring the trait of self-incompatibility. In some cases, emasculations were performed one day before bud opening and self-pollinations were performed one day after bud opening (Birhman and Hosaka 2000). In this same study, fertility was tested by screening pollen stainability with acetocarmine stain and by performing bulk pollinations (Birhman and Hosaka 2000). Different variables that have been used to measure self-compatibility include, but are not limited to, seed set, fruit set, stylar S-RNase production, and pollen tube growth (Birhman and Hosaka 2000; Lloyd and Schoen 1992; Stone et al. 2006; Van Gastel and De Nettancourt 1975). In general self-compatibility is more readily confirmed than self-incompatibility. Since even self-compatible plants do not always positively respond to every self-pollination performed it is dangerous to conclude that a plant is self-incompatible simply because it does not set fruit after one or two self-pollinations. To avoid such dubious phenotyping, Birhman and Hosaka (2000) disregarded giving plants the designation of “self-incompatible” until more than five self-pollinations were attempted.

In terms of evaluating self-incompatibility, demonstrations have shown that this historically qualitative trait can in fact be measured on both a qualitative and a quantitative basis (Stone et al. 2006). By observing previous research of self-incompatibility there are multiple warnings that should be recognized when evaluating this trait. For example, in the process of detecting self-incompatibility one must take into consideration the occurrence of OP fruits. OP, or spontaneous, fruits are those fruits that appear on plants that are not the result of intentionally performed self-pollinations. The difficulty lies in distinguishing whether such fruits are truly the result of self-compatibility or the result of unintentional insect cross pollinations. In the case where there is a clear distinction between those plants producing OP fruit and

those that do not, especially if the plants are in the same greenhouse, it can generally be concluded that such OP fruits are the results of self-compatibility (Stone et al. 2006).

Another warning addresses the occurrence of pollen-part spontaneous mutation rates which, in *Nicotiana alata*, have been reported to range from 0.2 to 0.4 mutations per million pollen grains (de Nettancourt 1977). Spontaneous mutation rates have the potential to yield false positive or false negative results when measuring the compatibility phenotype of a plant. One should also be cautious of confusing self-incompatibility with a loss of fertility (Birhman and Hosaka 2000). While these two traits are similar in terms of phenotypic measurements, they are not necessarily the same. Finally, it is beneficial to measure self-incompatibility over the course of multiple generations instead of simply considering one generation alone. As de Nettancourt (1977) points out, when measuring mutations at the *S*-locus it is dangerous to only consider those frequencies of mutations in the first generation.

Goals

To date, much research has been accomplished to better understand the gametophytic self-incompatibility system in members of the Solanaceae family. We propose to take this research even further by exploring the DRH population as a new and highly informative set of germplasm to more fully understand this trait of interest. The DRH population offers an excellent opportunity for studying the trait of self-incompatibility in potato. The diploid level of DRH, along with the homozygosity of the maternal parent, will aid in simplifying this research to a manageable level. The genomic and transcriptomic data that have already been generated will allow us to tap into a vast pool of data to gain a more comprehensive understanding of how self-incompatibility is determined. Furthermore, the molecular markers already developed for this population are ideal for studies of such quantitative traits. By using the DRH population we hope to greatly increase the opportunities that have already begun to be realized in previously studied potato species.

The aim of this study is to discover specific genes, or haplotypes, which segregate with the trait of self-compatibility and which are responsible for this phenotype in the DRH population. We hope to examine both the first generation and continuously inbred generations. Identifying self-compatibility genes in DRH will increase the usefulness of what has already proven to be instrumental diploid germplasm to potato geneticists. The results of this study should bring us closer to producing a true breeding homozygous

diploid which will be an advantageous tool for many later projects. Given the superior molecular resources available for DRH we believe such a homozygous line will be a suitable model for future research of potato.

References

Bai C, Sen P, Hofmann K, Ma L, Goebl M, Harper JW, Elledge SJ (1996) *SKP1* connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. Cell 86:263-274

Bartlett MS, Haldane JBS (1934) The theory of inbreeding in autotetraploids. J Genet 29:175-180

Birhman RK, Hosaka K (2000) Production of inbred progenies of diploid potatoes using an *S*-locus inhibitor (*Sli*) gene, and their characterization. Genome 43:495-502

Bredemeijer GMM, Blaas J (1981) S-Specific proteins in styles of self-incompatible *Nicotiana alata*. Theor Appl Genet 59:185-190

Chase SS (1963) Analytic breeding in *Solanum tuberosum* L. - A scheme utilizing parthenotes and other diploid stocks. Can J Genet Cytol 5:359-363

Chen G, Zhang B, Liu LJ, Li Q, Zhang YE, Xie Q, Xue YB (2012) Identification of a ubiquitin-binding structure in the S-locus F-box protein controlling S-RNase-based self-incompatibility. J Genet Genom 39:93-102

Cipar MS, Peloquin SJ, Hougas RW (1964) Inheritance of incompatibility in hybrids between *Solanum tuberosum* haploids and diploid species. Euphytica 13:163-172

Craig KL, Tyers M (1999) The F-box: a new motif for ubiquitin dependent proteolysis in cell cycle regulation and signal transduction. Prog Biophys Mol Biol 72:299-328

De Jong H, Rowe PR (1971) Inbreeding in cultivated diploid potatoes. Potato Res 14:74-83

de Nettancourt D (1977) Incompatibility in angiosperms. Springer-Verlag, Berlin: 2-134

East E (1934) Norms of pollen-tube growth in incompatible matings of self-sterile plants. Proc Natl Acad Sci USA 20:225-230

East E, Mangelsdorf A (1925) A new interpretation of the hereditary behavior of self-sterile plants. Proc Natl Acad Sci USA 11:166-171

Eijlander R, Ramanna MS, Jacobsen E (1997) Selection of vigorous and fertile S-homo- and heterozygous tester clones from self-incompatible diploid potato, *Solanum tuberosum* L. Euphytica 97:97-111

Eijlander R, ter Laak W, Hermsen JGT, Ramanna MS, Jacobsen E (2000) Occurrence of self-compatibility, self-incompatibility and unilateral incompatibility after crossing diploid *S. tuberosum* (SI) with *S. verrucosum* (SC): I. Expression and inheritance of self-compatibility. Euphytica 115:127-139

Felcher KJ, Coombs JJ, Massa AN, Hansey CN, Hamilton JP, Veilleux RE, Buell CR, Douches DS (2012) Integration of two diploid potato linkage maps with the potato genome sequence. PLoS One 7:e36347

Gagne JM, Downes BP, Shiu SH, Durski AM, Vierstra RD (2002) The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in *Arabidopsis*. Proc Natl Acad Sci USA 99:11519-11524

Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufmann H, Thompson RD, Bonierbale MW, Ganal MW, Tanksley SD, Salamini F (1991) RFLP maps of potato and their alignment with the homoeologous tomato genome. Theor Appl Genet 83:49-57

Goldraij A, Kondo K, Lee CB, Hancock CN, Sivaguru M, Vazquez-Santana S, Kim S, Phillips TE, Cruz-Garcia F, McClure B (2006) Compartmentalization of S-RNase and HT-B degradation in self-incompatible *Nicotiana*. Nature 439:805-810

Golmirzaie AM, Ortiz R, Atlin GN, Iwanaga M (1998) Inbreeding and true seed in tetrasomic potato. I. Selfing and open pollination in Andean landraces (*Solanum tuberosum* Gp. *Andigena*). Theor Appl Genet 97:1125-1128

Golz JF, Oh HY, Su V, Kusaba M, Newbigin E (2001) Genetic analysis of *Nicotiana* pollen-part mutants is consistent with the presence of an S-ribonuclease inhibitor at the S locus. Proc Natl Acad Sci USA 98:15372-15376

Golz JF, Su V, Clarke AE, Newbigin E (1999) A molecular description of mutations affecting the pollen component of the *Nicotiana alata* S locus. Genetics 152:1123-1135

Gonzalez N, Gévaudant F, Hernould M, Chevalier C, Mouras A (2007) The cell cycle-associated protein kinase WEE1 regulates cell size in relation to endoreduplication in developing tomato fruit. Plant J 51:642-655

Gray JE, McClure BA, Bönig I, Anderson MA, Clarke AE (1991) Action of the style product of the self-incompatibility gene of *Nicotiana alata* (*S*-RNase) on in vitro-grown pollen tubes. Plant Cell 3:271-283

Hamilton JP, Hansey CN, Whitty BR, Stoffel K, Massa AN, Van Deynze A, De Jong WS, Douches DS, Buell CR (2011) Single nucleotide polymorphism discovery in elite north american potato germplasm. BMC Genomics 12

Hershko A, Heller H, Elias S, Ciechanover A (1983) Components of ubiquitin-protein ligase system - resolution, affinity purification, and role in protein breakdown. J Biol Chem 258:8206-8214

Hosaka K, Hanneman RE (1998a) Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 1. Detection of an S locus inhibitor (Sli) gene. Euphytica 99:191-197

Hosaka K, Hanneman RE (1998b) Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 2. Localization of an S locus inhibitor (Sli) gene on the potato genome using DNA markers. Euphytica 103:265-271

Hua ZH, Kao TH (2006) Identification and characterization of components of a putative *Petunia* S-locus F-box-containing E3 ligase complex involved in S-RNase-based self-incompatibility. Plant Cell 18:2531-2553

Hua ZH, Meng XY, Kao TH (2007) Comparison of *Petunia inflata S*-locus F-box protein (Pi SLF) with Pi SLF-like proteins reveals its unique function in S-RNase-based self-incompatibility. Plant Cell 19:3593-3609

Huang S, Lee HS, Karunanandaa B, Kao TH (1994) Ribonuclease activity of *Petunia inflata* S-proteins is essential for rejection of self-pollen. Plant Cell 6:1021-1028

Igic B, Busch JW (2013) Is self-fertilization an evolutionary dead end? New Phytol 198:386-397

Ioerger TR, Gohlke JR, Xu B, Kao TH (1991) Primary structural features of the self-incompatibility protein in solanaceae. Sex Plant Reprod 4:81-87

Iwakawa H, Shinmyo A, Sekine M (2006) Arabidopsis *CDKA;1*, a *cdc2* homologue, controls proliferation of generative cells in male gametogenesis. Plant J 45:819-831

Jacobs JME, Vaneck HJ, Arens P, Verkerkbakker B, Hekkert BTL, Bastiaanssen HJM, Elkharbotly A, Pereira A, Jacobsen E, Stiekema WJ (1995) A genetic map of potato (*Solanum tuberosum*) integrating molecular markers, including transposons, and classical markers. Theor Appl Genet 91:289-300

Kaufmann H, Salamini F, Thompson RD (1991) Sequence variability and gene structure at the self-incompatibility locus of *Solanum tuberosum*. Mol Gen Genet 226:457-466

Kubo K, Entani T, Takara A, Wang N, Fields AM, Hua ZH, Toyoda M, Kawashima S, Ando T, Isogai A, Kao T, Takayama S (2010) Collaborative non-self recognition system in S-RNase-based self-incompatibility. Science 330:796-799

Lai Z, Ma WS, Han B, Liang LZ, Zhang YS, Hong GF, Xue YB (2002) An F-box gene linked to the self-incompatibility (S) locus of *Antirrhinum* is expressed specifically in pollen and tapetum. Plant Mol Biol 50:29-42

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

Lewis D (1949) Incompatibility in flowering plants. Biological Reviews 24:472-496

Lewis D (1951) Structure of the incompatibility gene. Heredity 5:399-414

Lewis D (1960) Genetic control of specificity and activity of the S antigen in plants. Proc Royal Society; Biol Sci 151:468-477

Lindhout P, Meijer D, Schotte T, Hutten RCB, Visser RGF, van Eck HJ (2011) Towards F_1 hybrid seed potato breeding. Potato Res 54:301-312
Lloyd DG (1992) Self- and cross-fertilization in plants .2. The selection of self-fertilization. Int J Plant Sci 153:370-380
Lloyd DG, Schoen DJ (1992) Self- and cross-fertilization in plants .1. Functional dimensions. Int J Plant Sci 153:358-369
Martin FW (1968) The behaviour of *Lycopersicon* incompatibility alleles in an alien genetic milieu. Genetics 60:101-109
McClure B, Cruz-García F, Romero C (2011) Compatibility and incompatibility in S-RNase-based systems. Ann Bot 108:647-658
McClure BA, Gray JE, Anderson MA, Clarke AE (1990) Self-incompatibility in *Nicotiana alata* involves degradation of pollen rRNA. Nature 347:757-760
McClure BA, Haring V, Ebert PR, Anderson MA, Simpson RJ, Sakiyama F, Clarke AE (1989) Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. Nature 342:955-957
Nakamura S, Hosaka K (2010) DNA methylation in diploid inbred lines of potatoes and its possible role in the regulation of heterosis. Theor Appl Genet 120:205-214
O'Brien M, Kapfer C, Major G, Laurin M, Bertrand C, Kondo K, Kowyama Y, Matton DP (2002) Molecular analysis of the stylar-expressed *Solanum chacoense* small asparagine-rich protein family related to the HT modifier of gametophytic self-incompatibility in *Nicotiana*. Plant J 32:985-996
O'Brien M, Major G, Chantha SC, Matton DP (2004) Isolation of S-RNase binding proteins from *Solanum chacoense*: identification of an SBP1 (RING finger protein) orthologue. Sex Plant Reprod 17:81-87
Pallais N (1991) True potato seed: Changing potato propagation from vegetative to sexual. Hortscience 26:239-241
Pandey KK (1962) Incompatibility in Some *Solanum* Crosse. Genetica 33:24-30
Patton EE, Willems AR, Tyers M (1998) Combinatorial control in ubiquitin-dependent proteolysis: don't Skp the F-box hypothesis. Trends Genet 14:236-243
Paz MM, Veilleux RE (1999) Influence of culture medium and *in vitro* conditions on shoot regeneration in *Solanum phureja* monoploids and fertility of regenerated doubled monoploids. Plant Breeding 118:53-57
Phumichai C, Mori M, Kobayashi A, Kamijima O, Hosaka K (2005) Toward the development of highly homozygous diploid potato lines using the self-compatibility controlling *Sli* gene. Genome 48:977-984
Puerta AR, Ushijima K, Koba T, Sassa H (2009) Identification and functional analysis of pistil self-incompatibility factor *HT-B* of *Petunia*. J Exp Bot 60:1309-1318
Qiao H, Wang HY, Zhao L, Zhou JL, Huang J, Zhang YS, Xue YB (2004) The F-box protein AhSLF-S-2 physically interacts with S-RNases that may be inhibited by the ubiquitin/26S proteasome pathway of protein degradation during compatible pollination in antirrhinum. Plant Cell 16:582-595
Ruan YL, Patrick JW, Bouzayen M, Osorio S, Fernie AR (2012) Molecular regulation of seed and fruit set. Trends Plant Sci 17:656-665
Saba-El-Leil MK, Rivard S, Morse D, Cappadocia M (1994) The *S11* and *S13* self-incompatibility alleles in *Solanum chacoense* Bitt. are remarkably similar. Plant Mol Biol 24:571-583
Shonnard GC, Peloquin SJ (1991) Performance of true potato seed families .1. Effect of level of inbreeding. Potato Res 34:397-407
Sijacic P, Wang X, Skirpan AL, Wang Y, Dowd PE, McCubbin AG, Huang S, Kao TH (2004) Identification of the pollen determinant of S-RNase-mediated self-incompatibility. Nature 429:302-305
Simmonds NW (1966) Linkage to the S-locus in diploid potatoes. Heredity 21:473-479

Simmonds NW (1997) A review of potato propagation by means of seed, as distinct from clonal propagation by tubers. Potato Res 40:191-214
Sims TL, Ordanic M (2001) Identification of a S-ribonuclease-binding protein in *Petunia hybrida*. Plant Mol Biol 47:771-783
Solanaceae Coordinated Agricultural Project. http://solcap.msu.edu/. Accessed: 11 September, 2013
Stephenson AG, Travers SE, Mena-Ali JI, Winsor JA (2003) Pollen performance before and during the autotrophic-heterotrophic transition of pollen tube growth. Philos T Roy Soc B 358:1009-1017
Stone JL, Sasuclark MA, Blomberg CP (2006) Variation in the self-incompatibility response within and among populations of the tropical shrub *Witheringia solanacea* (Solanaceae). Am J Bot 93:592-598
Takayama S, Isogai A (2005) Self-incompatibility in plants. Annual review of plant biology. Annual Reviews, Palo Alto, pp 467-489
Takebayashi N, Morrell PL (2001) Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. Am J Bot 88:1143-1150
Tao R, Iezzoni AF (Scientia Horticulturae2010) The S-RNase-based gametophytic self-incompatibility system in *Prunus* exhibits distinct genetic and molecular features. Sci Hortic-Amsterdam 124:423-433
The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189-U194
Van Gastel AJG, De Nettancourt D (1975) The effects of different mutagens on self-incompatibility in Nicotiana alata Link and Otto. Heredity 34:381-392
Van Os H, Andrzejewski S, Bakker E, Barrena I, Bryan GJ, Caromel B, Ghareeb B, Isidore E, de Jong W, van Koert P, Lefebvre V, Milbourne D, Ritter E, van der Voort J, Rousselle-Bourgeois F, van Vliet J, Waugh R, Visser RGF, Bakker J, van Eck HJ (2006) Construction of a 10,000-marker ultradense genetic recombination map of potato: Providing a framework for accelerated gene isolation and a genomewide physical map. Genetics 173:1075-1087
Wheeler D, Newbigin E (2007) Expression of 10 S-class *SLF-like* genes in *Nicotiana alata* pollen and its implications for understanding the pollen factor of the S locus. Genetics 177:2171-2180
Xu BB, Mu JH, Nevins DL, Grun P, Kao TH (1990) Cloning and sequencing of cDNAs encoding two self-incompatibility associated proteins in *Solanum chacoense*. Mol Gen Genet 224:341-346
Yasuda S (1934) Physiological research on self-incompatibility in *Petunia violacea*. Imperial College of Agriculture and Forestry 20:1-95
Zhou YM, Wang H, Gilmer S, Whitwill S, Keller W, Fowke LC (2002) Control of petal and pollen development by the plant cyclin-dependent kinase inhibitor ICK1 in transgenic *Brassica* plants. Planta 215:248-257

Chapter 2. Self-compatibility in the DRH population

Introduction

Diploid potato (2*n*=2*x*=24) is essential to potato genetic studies due to the simplicity it offers over traditionally cultivated highly heterozygous tetraploids (2*n*=4*x*=48). One of the drawbacks to diploid potato propagation, however, is that the production of inbred lines through self-pollination is typically limited by gametophytic self-incompatibility (Cipar et al. 1964). This limitation is of particular concern currently as two diploid genotypes, *S. tuberosum* Group Phureja DM 1-3 516 R44 (DM) and *S. tuberosum* group Tuberosum RH89-039-16 (RH) were recently used to sequence the potato genome (The Potato Genome Sequencing Consortium 2011). While RH, as an exception to the usual condition, is self-compatible, DM cannot be classified due to unsuccessful self or cross-pollinations as pollinator in various studies. These two genotypes have been crossed to generate the hybrid F_1 DRH segregating population which has already proved to be a valuable tool to potato geneticists (Felcher et al. 2012). Self-incompatibility is restrictive because it hinders the selective advancement of homozygosity in a population. While alternative methods of reaching homozygosity exist, such as doubling haploids derived through anther culture, some genotypes are recalcitrant to these techniques. More importantly these techniques do not allow the researcher to impose selection for specific phenotypic qualities along the way. Given the profusion of unfavorable alleles present in potato, these one-step techniques for reaching homozygosity often fail to eliminate unfavorable alleles and thereby result in genotypes of inferior fitness or limited fertility (Lindhout et al. 2011). Through selfing it is possible to advance homozygosity while selecting for the fittest and most fertile phenotypes at the same time. Many popular breeding schemes such as heterosis breeding, sexual polyploidization, and recombinant inbred line breeding rely on the ability to develop vigorous homozygous germplasm that can be sexually propagated. Unfortunately there are no such homozygous inbred lines currently available for cultivated potato.

In solanaceous crops self-incompatibility is determined by a gametophytic system in which the genotype of the pollen controls the compatibility response. Recognition is necessary between pollen and pistil proteins in order for the plant to distinguish between self and cross pollinations (de Nettancourt 1977; East and Mangelsdorf 1925; McClure et al. 2011). Both pollen and pistil recognition proteins are

encoded by the same multiallelic locus, the *S*-locus. Self-recognition is made possible through protein specificity encoded by *S*-locus genes (Bredemeijer and Blaas 1981). The variation necessary to confer this specificity can be as subtle as only ten amino acid dissimilarities (Saba-El-Leil et al. 1994). Previous studies using RFLP markers and stylar glycoproteins mapped the *S*-locus to chromosome 1 of potato (Gebhardt et al. 1991; Jacobs et al. 1995).

The pistil specific component of the *S*-locus is known as S-RNase due to its ribonuclease activity (McClure et al. 1989). According to the cytotoxic model S-RNase functions to degrade rRNA of pollen tubes growing inside the styles of self-incompatible genotypes (McClure et al. 1990). Indeed the ribonuclease activity of S-RNase is necessary for self-incompatibility as established by the inability of S-RNase mutants to reject pollen tubes (Huang et al. 1994).

As the complimentary recognition *S*-locus factor to S-RNase, the pollen specific component is an *S*-locus F-box commonly referred to as SLF (Lai et al. 2002). SLFs are members of the F-box family, one of the largest super-families in the plant kingdom (Gagne et al. 2002). The primary goal of F-box proteins is to regulate the cell cycle by facilitating degradation via ubiquitin ligase (Chen et al. 2012). Three critical enzymes, E1 (a ubiquitin activating enzyme), E2 (a ubiquitin conjugating enzyme), and E3 (a ubiquitin protein ligase), are engaged in the ubiquitin degradation process (Hershko et al. 1983). SLFs are proposed to form SCF complexes, which constitute an important class of E3 ubiquitin protein ligases (Patton et al. 1998; Qiao et al. 2004). The purpose of SLFs in the compatibility reaction is to target S-RNase for degradation in order to preserve pollen tube RNA. Studies have demonstrated that SLF is preferentially attracted to non-self S-RNases, which ultimately gives self S-RNases an advantage in destroying pollen tubes (Hua and Kao 2006). According to one model SLF consists of three domains: an S-RNase binding domain which promotes binding to non-self S-RNases, and two S-RNase binding regulatory domains which serve as negative regulators (Hua et al. 2007).

Due to the reliance of gametophytic self-compatibility on the recognition of S-RNase and SLF, the number of alleles at the *S*-locus has a profound effect on the possible interactions. If the SLF alleles are heterozygous then it is possible to code for more than one recognition factor to target S-RNase, thus increasing the chance of pollen tube survival. The lowest ploidy level at which SLF can maintain heterozygosity is tetraploid since the haploid cells of a tetraploid have two alleles per locus. The

phenomenon of overcoming self-incompatibility through increased ploidy level is referred to as the heteroallelic pollen effect (de Nettancourt 1977; McClure et al. 2011). This phenomenon explains why self-compatibility is more common among tetraploid than diploid potatoes.

In addition to S-RNase and SLF there are other genes outside the *S* locus known as modifier genes which support the self-compatibility reaction through altering the compatibility condition without direct involvement in the recognition process (Tao and Iezzoni Scientia Horticulturae2010). One of the most notable modifier genes identified is the *S*-locus inhibitor (*Sli*) discovered in a mutant of the wild diploid *S. chacoense* (Hosaka and Hanneman 1998a). Through sporophytic action the *Sli* gene, located on chromosome 12 and nonallelic to the *S*-locus, is able to induce self-compatibility in otherwise self-incompatible genotypes via blockage of pollen *S*-allele interaction (Hosaka and Hanneman 1998a, b). Researchers have succeeded in breeding the *Sli* gene into self-incompatible diploid potato germplasm to confer self-compatibility (Birhman and Hosaka 2000; Lindhout et al. 2011; Phumichai et al. 2005).

In addition to the variety of genetic elements responsible for determining self-compatibility there also remains considerable environmental variation. Floral age has been noted to influence the likelihood of observing a self-compatible phenotype as the mechanisms for determining self-incompatibility are more likely to be dysfunctional in young and old flowers (Eijlander et al. 1997; Stephenson et al. 2003). Furthermore the season alters the compatibility condition as self-incompatibility is more likely to break down later in the season rather than earlier (East 1934; Stone et al. 2006; Yasuda 1934). Additional important environmental factors to consider include, but are not limited to: temperature, stigma size, pollen moisture, light, weather, and population density (de Nettancourt 1977; Lloyd and Schoen 1992). Altogether these factors should serve as cautionary reminders to any researcher studying this important trait that there are many uncontrollable elements to take into consideration.

The purpose of this project was to understand the unexpected occurrence of self-compatible phenotypes within the DRH population. In a previous season fruit set following self-pollination was observed among several but not all greenhouse grown DRH F_1 genotypes (unpublished data). The observed segregation among the DRH F_1 hybrids makes this population ideal for studying self-compatibility using the abundance of newly available genomic tools. Specifically, the recently developed 8303 Infinium single nucleotide polymorphism marker (SNP) array developed by the Solanaceae

Coordinated Agriculture Project (SolCap) (Hamilton et al. 2011; Solanaceae Coordinated Agricultural Project) allowed us to nominate candidate genes from gene models on the Potato Genome Browser in regions that co-segregated with SNPs that were significantly associated with the self-compatibility phenotype.

Materials and methods

Plant material

The DRH population was developed by cross-pollination between DM 1-3 516 R44 (DM) as stylar parent and RH89-039-16 (RH) as pollen parent. DM has not been used successfully as a pollinator although it has considerable stainable pollen using acetocarmine stain (unpublished). RH was observed to set fruit after self-pollination under our growing conditions; hence, self-compatibility which segregated in the DRH F_1 progeny was believed to reflect a particular set of alleles or haplotypes inherited from the paternal genotype.

Production of inbred lines

Unique F_1 genotypes (n=103) of the DRH population were planted during the fall season from either tubers or young sprouts (Fig. 1). Each genotype was replicated to yield a total 206 individuals. All plants were grown in 15.2 cm pots filled with ProMix Growing Medium (©Premier Tech Ltd.) in soil beds in a greenhouse under a 16-h photoperiod for approximately 5 months. On flowering a single anther per bud of the majority of genotypes was removed, pollen extruded on a glass slide and stained with acetocarmine. Screening was performed at x40 using a Nikon Alphaphot YS compound microscope to estimate the approximate percentage of aborted pollen for each clone. When possible, screening was repeated using anthers from different buds of the same genotype to determine the average percentage of aborted pollen.

Self-pollinations were performed by hand on all DRH genotypes that flowered. On genotypes with abundant flowering a total of 25 pollinations was performed per genotype, not exceeding more than five pollinations per day per genotype. On all other genotypes with limited flowering as many pollinations as possible were performed, still limiting the number of pollinations per day per genotype to five. The number of pollinations was spread out over multiple days to compensate for the daily environmental fluctuations which contributed to unpredictable rates of pollination success. Genotypes producing fruit from self-

pollinations were recorded, along with the number of fruit produced per number of pollinations performed. After approximately 6 weeks the ripened fruit were harvested and seeds extracted and counted to determine the approximate number of seeds per fruit for each genotype.

Tubers from the F_1 genotypes were harvested, vernalized for 6 months at 4°C, and planted for a second fall season (approximately 5 months) in the greenhouse under a 16-h photoperiod to verify the compatibility condition of genotypes for which previous records were inconsistent or for which less than 15 pollinations (not resulting in fruit set) had been previously performed. Once plants began to flower the same methods applied in the first planting of the F_1 generation for self-pollination and recording fruit and seed count were used. By repeating the experiment in two different seasons we hoped to eliminate some variation affecting our phenotypic characterizations due to seasonal variation.

In order to increase the homozygosity of DRH and to further test the genes that co-segregated with self-compatibility in the F_1 generation, we planted the seeds from selfed fruit of the RH parent and five different DRH F_1 genotypes (DRH 16, 67, 76, 84, and 90) in a Conviron controlled climate growth chamber. After germination, roughly 35 S_1 seedlings from each F_1 genotype were transplanted to 15.2 cm pots on gravel beds in a greenhouse under a 16-h photoperiod during the fall season for approximately 6 months to study the first selfed, or S_1, generation. Thirty-five S_1 RH seedlings and eight DRH F_1 seedlings were also planted to serve as controls. Once plants began to flower, the same methods applied in the F_1 generation for self-pollination and recording fruit and seed count were used for the S_1 generation. Due to poor growth and fertility (only one fruit was produced from all of the S_1 genotypes derived from the RH parent) we discontinued working on the RH family.

Seeds from selfed fruit representing all five DRH S_1 families were planted in a Conviron controlled climate growth chamber under 250 µmol/m^2/s of fluorescent light. After germination, a total of 160 seedlings was transplanted to deep-pots and remained in the growth chamber for approximately 5 months under a 14-h photoperiod to encourage a short growth habit to prevent overcrowding. The growth chamber was used in order to advance our population by another generation during the summer of 2012 when greenhouse conditions would have been unsuitable for adequate flowering and fruit set in this cool season crop. Once plants began to flower the same methods applied in the F_1 generation for self-pollination and recording fruit and seed count were used for the S_2 generation, with the exception that the

total number of pollinations was no longer limited to 25. This modification in protocol reflected our hope to compensate for predicted reductions in fruit and seed set in the further inbred lines.

Due to the inability to identify self-compatible individuals in two of the families in the S_2 generation only seeds from selfed fruit of three DRH S_2 families (DRH 16, 76, and 90) were planted in a controlled climate growth chamber to advance to the next generation. A total of 458 seedlings germinated and was transplanted to a greenhouse in 15.2 cm pots in soil beds set to grow under a 16-h photoperiod during the fall season for approximately 6 months. Once plants began to flower the same methods applied in the F_1 generation for self-pollination and recording fruit and seed count were repeated for the S_3 generation, with the exception that the total number of pollinations was no longer limited to 25. Pollen screening via acetocarmine staining was performed on selected genotypes so that the approximate percentage of aborted pollen could be calculated.

DNA extraction

DNA was extracted from selected plants in each inbred generation using a modified hexadecyltrimethylammonium bromide (CTAB) protocol similar to that described by Doyle and Doyle (1987). Approximately 3 g of freeze-dried young leaf tissue was finely ground and suspended in 500 µL extraction buffer [96.3% stock solution (150 mM Tris-HCL, 1M NaCl, 15 mM EDTA), 1.5% CTAB, 1.2% β-mercaptoethanol, 1.0% PVP-40] and incubated at 65°C for 30-60 min. An aqueous DNA phase was separated from extraneous material by a phenol:chloroform:isoamyl alcohol solution (25:24:1) [alt. chloroform:isoamyl alcohol (24:1)]. DNA was washed with a mixture of isopropanol and 3 M sodium acetate (alt. 100% ethanol and 5 M sodium chloride). A final washing step of 70% ethanol was performed before drying the DNA and resuspending it in 100-120 µL TE with 20 µg/ml RNase.

Infinium 8303 array

A single nucleotide polymorphism (SNP) marker array was recently developed using an Infinium platform (Illumina, San Diego, CA), yielding 8,303 high confidence markers for potato (Hamilton et al. 2011; Solanaceae Coordinated Agricultural Project). Genotyping was performed using an Illumina iScan Reader with the Infinium HD Assay Ultra and allele calls were made using the program GenomeStudio (Illumina, Inc., San Diego, CA). All 103 DRH F_1 individuals used in this study were previously genotyped using the

8303 SNP array (Felcher et al. 2012). Additionally 92 individuals from the S_3 generation (46 self-compatible and 46 self-incompatible selections) were genotyped using the 8303 SNP array.

Before proceeding to statistical analyses, we used a variety of filters to remove poor quality and uninformative SNPs in each population. SNPs which were previously identified by the Solanaceae Coordinated Agriculture Project as poor quality for three-cluster custom calling in GenomeStudio along with SNPs which mapped to two or more loci on the chromosomes were removed (Solanaceae Coordinated Agricultural Project). If missing calls were reported for greater than or equal to ten percent of the population then those SNPs were also removed. The parental lines DM and RH were genotyped and used to remove SNPs which yielded unexpected or inconsistent reads for these control lines. Finally, SNPs for which there was no segregation observed in the population were also removed.

Statistical analyses

Segregation patterns for both fruit set and seed set were used to identify SNPs significantly correlated with the self-compatibility phenotype. Genotypes for which no pollinations were performed, or for which no fruit production was observed and less than five pollinations were performed, were eliminated from the analysis to avoid possible erroneous categorization of phenotypes. A bivariate fit of fruit-per-pollination rate versus seeds per fruit was conducted to determine whether or not these two traits were significantly correlated. For the trait of fruit set categories, we assigned individuals into three classes of compatibility: if a genotype produced zero fruits per pollination it was categorized as “low”, if a genotype produced more than zero but less than 0.25 fruits per pollination it was categorized as “medium”, and finally if a genotype produced equal to or more than 0.25 fruits per pollination it was categorized as “high”. These three categories were created because we believed that simply scoring genotypes as “self-compatible” and “self-incompatible” might be misleading due to the occurrence of some genotypes which only produced few fruits after many pollinations. Be creating categories which differentiated between “medium” and “high” fruit-per-pollination rates we hoped to reduce the confusion of chance breakdown of self-incompatibility due to environmental conditions with the genetic factors which control this trait.

The statistics program R was used to test the correlation of each individual SNP with each phenotypic category (“low”, “medium, and “high”) using Fishers’ Exact Test (Team 2010). Only those SNPs with p-value ≤ 0.0001 were retained in the analysis. SNPs which met this criteria were analyzed by contingency

table analysis in JMP 10 in order to identify the potential for Type 1 errors in the dataset (SAS). For each SNP if an allelic state was represented by five or less individuals and that same allelic state was the sole cause for a SNP to be significant, then that SNP was discarded. SNPs were also removed when two or more allelic states each represented less than 50% of a category ("low", "medium", or "high"). These filters helped to ensure than only well represented allelic states that consistently segregated with a trait would lead to designating a SNP as significant. SNPs located on unmarked or unanchored scaffolds were handled separately due to the inability to map these SNPs in order to search for candidate genes.

As an alternative approach to measuring the degree of self-compatibility by fruit set, we also analyzed seed set for those genotypes that produced fruit. To avoid confusing SNPs segregating for fruit set with SNPs segregating for seed set, we used only data from genotypes in the "high" category for this analysis. That is to say, all self-incompatible plants, which by nature have a value of zero seed per fruit, were removed from the data set along with those genotypes which produced very few fruits possibly due to chance breakdown of self-incompatibility. Seed counts for each individual fruit were considered as unique events in order to capture the full range of potential seed set within a given genotype. The program JMP 10 was used to perform one-way ANOVA on the number of seed per fruit per each individual SNP. Only those SNPs with a p-value ≤ 0.0001 were retained in the analysis. SNPs located on unmarked or unanchored scaffolds were handled separately due to the inability to map these SNPs for the purpose of searching for candidate genes.

The program HaploView was used to predict groups exhibiting similar haplotype trends within each chromosome based on the haplotypes of individuals at all loci identified as significant for fruit set or seed set (Barrett et al. 2005). This facilitated efforts to identify which groups of SNPs segregated in similar patterns.

Preferential transmission of RH alleles in the F_1 progeny was tested for by identifying those SNPs which were 100% heterozygous and for which the allelic state of RH was heterozygous. By nature of the design of the SNP chip only RH SNPs which share at least one allele with DM were included. Since DM is homozygous all F_1 progeny that are completely heterozygous at a particular SNP must favor the alternative allele (not shared by DM) donated by RH.

In order to predict whether genotypes assigned to the “medium” category were truly self-compatible the list of significant SNPs for fruit set was used to determine whether alleles at theses specific SNPs segregated more or less with the alleles associated with self-compatibility. If more than 75% of the alleles matched those alleles correlated with self-compatibility then that plant was considered a truly self-compatible plant. On the other hand, if more than 75% of the significant alleles matched those alleles correlated with self-incompatibility then the fruit on those genotypes was considered to have arisen from breakdown of self-incompatibility.

Potato Genome Sequencing Consortium’s Genome Browser

The Potato Genome Sequencing Consortium’s Genome Browser (v2.1.11) was used to search for candidate genes within the genomic regions of those SNPs most highly correlated with self-compatibility. Regions both upstream and downstream of these SNPs were investigated, taking into account the possibility that distant genes may be segregating along with these markers due to linkage disequilibrium (D'Hoop et al. 2010). Once we identified unique candidate genes the Potato eFP Browser was used to search for corresponding Gene ID codes to aid in visualizing the relative expression differences of that particular gene in both the DM and RH parental lines (Massa et al. 2011; The Potato Genome Sequencing Consortium 2011; Winter et al. 2007). Genes for which no expression differences were recorded in the floral or fruit tissues were discarded from the list of acceptable candidates.

Bi-allelic discrimination assays

Prior to genotyping the S_3 generation using the 8303 SNP array, we performed bi-allelic discrimination assays to genotype specific loci in both the S_1 and S_2 generations in order to determine whether or not SNPs significantly correlated with self-compatibility in the F_1 generation maintained significance in the inbred lines. Immature leaf tissue samples of the DRH S_1 population were collected for DNA from 17 plants producing fruits from self-pollinations and also from 15 plants that did not produce fruit from self-pollinations. To ensure that both phenotypes were represented in each family, multiple tissue samples for both fruit-producing and non-fruit-producing plants were collect within each family. Likewise in the S_2 generation tissue was collected for DNA from all 13 fruit producing plants and also from 14 plants that did not produce fruit. Ten allele specific primers were designed around SNPs in order to genotype the allelic state of these specific loci in the inbred lines (KBioScience) (Table S1). Of the ten SNPs used as primer

targets, six were Infinium High Confidence SNPs from the SolCap 8303 array and were chosen because these SNPs were located in regions of highly significant SNPs for self-fruit set and that segregated within different haplotype blocks. The other four SNPs used were RH SNPs documented on the Genome Browser and were chosen due to location of these SNPS within or very near candidate genes for self-compatibility. Calling dyes for the primers were FAM and VIC while ROX was used for a passive reference. Parental lines DM and RH were used as control samples. Applied Biosystems software (ABI 7300 and ABI 7500 systems) was used to analyze the results of real time (RT) PCR reads. Conditions for amplification can be found in Table S2.

Haplotype analysis

Due to segregation of significant SNPs on multiple chromosomes in the S_3 generation a cluster analysis was performed to identify overall trends among genotypes. Dendrograms were generated in JMP 10 using the Ward method of hierarchical clustering. Separate cluster analyses were performed for each family (DRH 16, 76, and 90) from which the S_3 inbred lines were derived to avoid confusing haplotype patterns reminiscent of the families with haplotype patterns segregating for the compatibility condition. Only genotypes for which confident phenotype calls were made were used in this analysis. For the self-incompatible genotypes this excluded plants that received less than 25 pollinations. For the self-compatible genotypes this excluded any plants classified in the “medium” category as well as plants which received less than five pollinations. These qualifications were enforced in an effort to single out only reliably self-compatible and self-incompatible genotypes. Cluster calls were made based on a representative set of nine SNPs, all of which were significant for fruit set in the S_3 generation. Each of the nine SNPs represents a different chromosome on which significant fruit set SNPs were located.

Using the same representative genotypes and SNPs from the cluster analysis, we performed contingency tests of each possible one-to-one combination of the nine individual SNPs (36 unique possible interactions, excluding the interaction of a SNP with itself). For each SNP if an allelic state was represented by five or less individuals and that same allelic state was the sole cause for a SNP to be significant, then that SNP was discarded. SNPs were also removed when two or more allelic states each represented less than 75% of a category (interaction with the two or three possible allelic states of the other SNP being tested). These filters helped to ensure than only well represented allelic states of one

SNP that consistently segregated with well represented allelic states of the alternative SNP would lead to designating an interaction as significant.

Homozygosity trends in the inbred lines

We performed comparative analyses using SNP genotyping data from both the F_1 and S_3 generations in order to uncover the trends imposed on the genome by selecting for self-compatibility. Before proceeding to comparative homozygosity analyses, we used a variety of filters to remove poor quality and uninformative SNPs in each population. SNPs which were previously identified by the Solanaceae Coordinated Agriculture Project as poor quality for three-cluster custom calling in GenomeStudio along with SNPs which mapped to two or more loci on the chromosomes were removed (Solanaceae Coordinated Agricultural Project). If missing calls were reported for greater than or equal to ten percent of the population then those SNPs were also removed. The parental lines DM and RH were genotyped and used to remove SNPs which yielded unexpected or inconsistent reads for these control lines. However SNPs with no segregation in the population were retained due to the valuable information these SNPs provided for heterozygosity studies.

To test for overall reduction of heterozygosity imposed by inbreeding, allele calls in both the F_1 and S_3 SNP data sets were re-recorded as simply homozygous or heterozygous. These reads were then used to calculate the overall percent of heterozygous alleles in both generations. On a more individual level, the percentage of heterozygous SNPs was calculated for each SNP. If the level of heterozygosity at a locus was greater in the S_3 generation than in the F_1 generation, then those SNPs were discarded. SNPs which were previously identified to be predisposed to genotyping errors of inflated heterozygosity were also removed (unpublished data).The percentage of heterozygosity at each SNP was used to create overlay plots of the F_1 and S_3 generations in order to find regions of retained heterozygosity in the inbred lines.

As an additional approach to study the reduction of heterozygosity in the S_3 generation, we calculated the frequency of fixed alleles in both the F_1 and S_3 generations. This was accomplished by analyzing those SNPs which showed no segregation in the population and comparing those SNPs in both generations to find new fixed alleles.

The program TASSEL 4.0 was used to generate a kinship matrix of the DRH F_1 progeny against the DM and RH parental genotypes in order to identify the relative spread of homozygosity among the F_1

genotypes. SNPs were subjected to a filter alignment with a minimum count of one and a minor frequency of 0.05 and minor SNP states were removed. Individuals were scored on a scale of zero to two. This allowed for differentiating which genotypes were more or less homozygous.

Results

Phenotypes observed across the DRH generations

In each generation several genotypes performed poorly and therefore could not be included in the analysis. This group included plants that produced buds with poor pollen shed, buds that did not open, or did not reach the flowering stage altogether due to slow maturity or early death. Pollinations were not possible for those plants and therefore we lost the ability to distinguish between compatibility phenotypes. Overall, the rate of flowering was lower in the inbred lines than in the F_1 generation (Table 1). While the percent of genotypes that produced selfed fruit in each population remained rather constant from generation to generation (range = 32% to 35%) the percent of genotypes that were classified in the "high" category exhibited a greater decline in the inbred lines (Table 1). The percent of genotypes in the "high" fruit set category dropped from 69% in the F_1 generation to 38% percent in the S_2 generation. However, with a percent of genotypes in the "high" category reaching 53% in the S_3 generation it appears that the inbred lines began to oppose the trend of decreased self-fruit set. Among genotypes in all generations, seed set counts varied greatly even within the same genotypes (Table 1). Of particular interest, though, is that fact that the greatest seed set (361 seeds per fruit) was observed in the S_3 generation with more than double the number of seeds observed in the F_1 generation. In the F_1 generation the fruit set varied from 1 to 74 fruits per genotype and the seed set varied from 1 to 171 seeds per fruit. In the S_3 generation the fruit set varied from 1 to 13 fruits per genotype and the seed set varied from 1 to 361 seeds per fruit. Although a slightly significant correlation existed between fruit and seed set in the F_1 generation, this correlation did not hold true in the further inbred lines (Table 1). Among all F_1 genotypes (both self-compatible and self-incompatible) screened for aborted pollen, none were found to have an average of greater than 90% aborted pollen grains. Of the 12 S_3 genotypes screened for aborted pollen four genotypes were found to have an average aborted pollen rate greater than 90%. Of these four genotypes one produced fruit on self-pollination and was considered self-compatible, one was inconclusive, and the other two were identified as self-incompatible.

Of the 98 F_1 DRH genotypes pollinated 7 yielded inconclusive incompatibility phenotypes across plantings in different seasons and therefore were eliminated, leaving 91 genotypes to be included in statistical analyses. Fruit and seed set for all four generations are shown in Table 1. Even though we recorded data for the number of fruit set per pollination, we often observed spontaneous fruit set on several genotypes. Seeds produced by the S_2 generation were bulked for all the fruit produced per individual genotype before counting, therefore average seed counts per genotype rather than individual seed counts per fruit were recorded for this generation.

SNP chip filtering in the F_1 generation

The filters described to remove poor and uninformative SNPs in the F_1 population reduced the number of available SNPs for analysis from 8,303 to 1,966. A total of 95 SNPs was identified as significantly correlated with F_1 fruit set (p-value <0.0001), seven of which were on unanchored scaffolds and the remaining 88 SNPs all located on chromosome 12 (Table S3, Fig. 2). Among these 88 SNPs, 11 distinct haplotype blocks were identified (Fig. 3). All 88 SNPs exhibited segregation distortion ranging from 1.8:1 to 3.8:1. Three candidate genes uniquely associated with self-compatibility were located within an 8 Mb of region of chromosome 12 that included a class-S F-box protein (PGSC0003DMG400008762), a style-specific self-incompatibility putative modifier protein HT-A1 (PGSC0003DMG400010793), and an S-protein homologue (PGSC0003DMG400008637) (Table 2).

Ten of 91 genotypes in the F_1 generation were classified in the “medium” category for fruit set. Of these ten genotypes six segregated with the self-incompatible haplotype more than 75% of the time and three segregated with the self-compatible haplotype more than 75% of the time (Table 3). For one genotype (DRH-093) SNPs segregated equally with the self-incompatible and self-compatible haplotypes.

For F_1 seed set a total of 45 significant SNPs was identified (p-value <0.0001), two located on unanchored scaffolds, the remaining 43 SNPs located on chromosomes 2, 4, and 12 (Table S3). On chromosome 2 a total of 30 SNPs segregated into at least five different haplotype blocks (Fig. 3). On chromosome 4 a total of 8 SNPs were identified in the same haplotype block (Fig. 3). On chromosome 12 a total of five SNPs was identified within a 2 Mb region with no distinct haplotype blocks. Only one candidate gene was found that was uniquely associated with self-compatibility, a Class-S F-box protein

(PGSC0003DMG400041057) located on chromosome 2 (Table 2, Fig. S1). The SNPs for seed set located on chromosome 12 were also present in the list of SNPs for fruit set on chromosome 12.

A total of 36 SNPs in the F_1 generation was found to be 100% heterozygous and for which the control RH reads were heterozygous (Table S4). These SNPs spanned all chromosomes in potato except chromosome 7. The one heterozygous SNP on chromosome 12 was located around 49.9 Mb which was within a region of significant SNPs for fruit set.

Bi-allelic discrimination assays

Of the ten SNPs that we selected from either the F_1 generation analysis or from location within candidate genes for verification in 32 genotypes (17 compatible and 15 incompatible) of the S_1 generation by using allele-specific primers, one (solcap_snp_c1_2689) showed no segregation of allelic states. This SNP was located within the region of the most significant SNPs for fruit set in the F_1 generation. Of the other nine SNPs examined only one SNP (RH_snp_2389842) continued to be significantly associated with fruit set in the S_1 generation (p-value < 0.0015). This SNP was located within the coding region of the S-protein homologue identified on chromosome 12. Genotyping of the S_2 generation confirmed the fixation of alleles at solcap_snp_c1_2689. None of the other nine SNPs tested were significant in the S_2 generation (p-value > 0.01).

SNP chip filtering in the S_3 generation

Using the filters described to remove poor and uninformative SNPs in the S_3 population, we reduced the number of available SNPs for analysis from 8,303 to 2,160. For S_3 fruit set a total of 33 significant SNPs was identified (p-value < 0.0001), two of which were located on unanchored scaffolds and the remaining 31 located on chromosomes 1, 2, 4, 6, 7, 8, 9, 11, and 12 (Table S3). Only one significant SNP each was located on chromosomes 1, 2, 7, and 9. A range of three to ten SNPs was located on each of chromosomes 4, 6, 8, 11, and 12. According to HaploView analysis the SNPs on these chromosomes did not segregate into multiple haplotype blocks (Fig. 3). Across all chromosomes a total of two candidate genes uniquely associated with self-compatibility was identified nearby significant SNPs. Located within a 2 Mb region on chromosome 6 these included an S-protein (PGSC0003DMG400026738) and an S-protein homologue (PGSC0003DMG400037544). Of the 33 significant SNPs associated with selfed fruit set in the S_3 generation, only those located on chromosome 6 were also marginally significant in the F_1

generation (0.03 < p-values < 0.07). Contingency analysis of the possible interactions among the nine representative SNPs on each chromosome harboring fruit set SNPs in the S_3 generation revealed 17 highly significant interactions (Fig. 4). All nine SNPs were involved in at least one significant interaction, and one SNP was involved in up to six significant interactions.

Cluster analysis of S_3 genotypes from the DRH 16 family could not be performed due to the lack of high confidence self-incompatible genotypes available to represent this family. However, cluster analyses using all nine representative SNPs for fruit set in the S_3 generation were done for the DRH 76 and DRH 90 families, both of which could be separated into two clusters each (Fig. 5). Within the DRH 76 family one cluster encompassed seven self-compatible genotypes and one self-incompatible genotype and the second cluster encompassed one self-compatible genotype and eight self-incompatible genotypes. Within the DRH 90 family one cluster encompassed nine self-compatible genotypes and three self-incompatible genotypes and the second cluster encompassed zero self-compatible genotypes and eight self-incompatible genotypes.

For S_3 seed set a total of 108 significant SNPs (p-value < 0.0001) was identified, three located on unanchored scaffolds, the remaining 105 located on chromosomes 1, 3, 4, 5, 8, 9, 11, and 12 (Table S3, Fig. S2). Only one significant SNP was located on each of chromosomes 5, 11, and 12. SNPs on chromosomes 1, 3, and 4 did not segregate into more than one haplotype block. However, SNPs on chromosomes 8 and 9 segregated into four and five haplotype blocks, respectively (Fig. 3). Six candidate genes uniquely associated with self-compatibility were identified, including an S-class F-box protein (PGSC0003DMG400024834) on chromosome 4 and five different S-protein homologues on chromosome 9 (PGSC0003DMG400011582, PGSC0003DMG400035767, PGSC0003DMG400046953, PGSC0003DMG400011584, and PGSC0003DMG400041960).

Comparison of the F_1 and S_3 SNPs

The overall percent of heterozygous loci decreased from 0.24 in the F_1 generation to 0.05 in the S_3 generation (Fig. 6). This resulted in an overall reduction of heterozygosity by 80% and a homozygosity level of 95% in the S_3 generation. The expectation for genes that differed at both alleles between DM and RH (AA x BB) would be 6.25% heterozygosity in the S_3 whereas loci with a shared allele between DM and RH (AA x AB) would be 12.5%. Despite the approach to homozygosity for most of the SNP loci, there

remained some 220 highly heterozygous loci (51-98% retained heterozygosity from the F_1 generation) even after three generations of inbreeding (Table S5). These SNPs spanned all 12 chromosomes, frequently segregating into clusters on each individual chromosome. While chromosomes 5 and 7 each had only one SNP with highly retained heterozygosity, chromosome 2 had up to 114 (Fig. 6, Table S5). Of the 220 SNPs, six on chromosome 12 were significant for selfed fruit set in the F_1 generation. Two of the six SNPs were also significant for selfed seed set in the F_1 generation. We also found five of the 220 highly heterozygous SNPs which corresponded to the F_1 SNPs that exhibited 100% preferential transmission of the alternative RH allele not shared by DM.

A portion of increased homozygosity was due to 430 loci with fixed alleles identified in the S_3 generation that were not previously fixed in the F_1 generation (Table S6). These fixed alleles were shared among all three families (DRH 16, 67, and 90) and spread over all chromosomes except chromosome 2. Of the 430 common fixed alleles, 428 derived from DM and 2 derived from RH. Of these 430 fixed alleles, eight represented significant SNPs adjacent to candidate genes in the F_1 generation. All eight of these SNPs were located on chromosome 12 in two distinct regions, 55640932-57805264 and 58642405-58689924. All eight of these SNPs were significantly associated with fruit set in the F_1 generation, and one was significant for both fruit and seed set in the F_1 generation. Two of these eight SNPs were fixed for RH alleles, while six were fixed for DM alleles. Within each family the number of fixed alleles was even greater. Not including those fixed alleles that were shared among all three families, there were 1,499 fixed alleles in the S_3 lines of DRH 16, 859 fixed alleles in the S_3 lines of DRH 76, and 1,401 fixed alleles in the S_3 lines of DRH 90. Of the 1,499 alleles fixed in DRH 16, 890 were fixed for the DM allele and 609 were fixed for the alternative allele contributed by RH. Of the 859 alleles fixed in DRH 76, 838 were fixed for the DM allele and 21 were fixed for the alternative allele contributed by RH. Of the 1,401 alleles fixed in DRH 90, 1104 were fixed for the DM allele and 297 were fixed for the alternative allele contributed by RH.

A bivariate fit of fruit per pollination values by the kinship values generated by TASSEL showed that the relative homozygosity of the F_1 progeny relative to DM and RH was not significantly correlated with fruit set (p-values 0.6248 and 0.3582).

Discussion

Self-compatibility in the F_1 generation

The segregation pattern for self-compatible plants in the F_1 generation does not follow a one-to-one Mendelian ratio expected for segregation of a single gene in a homozygous x heterozygous diploid cross. Segregation distortion has been reported to be common in DRH; therefore the deviation from one-to-one segregation was not entirely unexpected (Felcher et al. 2012). As a further explanation of the results we hypothesize that the trait of self-compatibility in the DRH population is under polygenic control.

The wide range of both fruit set and seed set observed support the idea that self-compatibility is not a strictly qualitative trait. While some genotypes readily produced fruit after only a few pollinations others required many pollinations before a single fruit would form. Given that occasionally self-incompatibility breaks down over time or under certain environmental conditions we believe that genotypes exhibiting a low fruit-per-pollination rate were the result of a breakdown of self-incompatibility barriers. By creating a "middle" category to separate the genotypes with low fruit set from those with high fruit set we hoped to avoid faulty classification of phenotypes that might have skewed the results.

While fruit and seed set were weakly correlated in the F_1 generation these traits do not necessarily measure self-compatibility to the same degree. Since seed set trends observed within genotypes were inconsistent, it is possible that environmental factors played a role in determining seed set. The number of seeds in a fruit is directly related to the number of pollen grains which are placed on the stigma, travel down the style, and fertilize the ovules. While self-incompatibility factors certainly influence the growth of the pollen tube there are many other factors to consider. Not every pollination event involves the same number of pollen grains, so some styles even within the same genotype receive a greater number of pollen grains and therefore a greater number of opportunities for seed development. Temperature and moisture also influence the survival of the pollen while it is sitting on the stigma. Furthermore the fertility of a plant, which is easy to confuse with compatibility, may alter the seed set. While self-incompatibility specifically affects pollen tube growth after self-pollination, infertility affects all sexual propagation alike whether self- or cross-pollinations. By examining the frequency of aborted pollen grains in our germplasm, we hoped to have decreased the likelihood of confounding infertility with self-incompatibility; however variation for fertility may still have influenced our results.

Since seed set was inconsistent within genotypes we chose to rely more heavily on fruit set for identifying candidate genes for self-compatibility. Also, given that the degree of fruit set variation decreased from the F_1 to the S_3 generation, while the degree of seed set variation increased, we believe that our efforts to select for the self-compatible phenotype have had a greater impact on the trait of fruit set compared to seed set. Previous studies have commonly used fruit set resulting from self-pollination as an indicator for degree of self-compatibility (Birhman and Hosaka 2000; Lloyd and Schoen 1992). Nevertheless one should be cautious in assuming that fruit set implies self-compatibility since other factors such as temperature, resource competition, and position of a flower within an inflorescence can influence the success of fruit development (Diggle 1995; Sato et al. 2000). For the purpose of our research we consider fruit set to be an indicator, not an absolute confirmation, of self-compatibility.

Results from the kinship matrix generated for the F_1 population indicate that overall no particular F_1 genotype was more or less likely to be self-compatible due to degree of homozygosity relative to DM or RH. While a few SNPs were identified as having preferential inheritance from the RH allele not shared by DM in a homozygous by heterozygous (AA x AB) cross, they did not influence the statistical analysis since all reads with no segregation were eliminated due to being uninformative for segregation studies.

Candidate genes in the F_1 generation

Despite having found a total of 95 significant SNPs for fruit set in the F_1 generation, we acknowledge the fact that many of these SNPs were predicted to segregate into haplotype blocks or were otherwise located in close proximity to each other indicating that linkage disequilibrium may have registered many SNPs as significant due to linkage with the same causal genes. Of the possible candidates that were identified by searching the genome near regions of significant SNPs many genes were selected which had general annotated functions that could play a role in many other processes besides determining self-compatibility. These include but are not limited to F-box family proteins, ubiquitin protein ligases, receptor kinases, serine/threonine protein kinases, and anther or ovule specific proteins. Such genes were identified for possible involvement in an SCF complex as part of an SLF element. However, given that most of these candidates belong to large gene families that play a role in global cell regulation functions it is quite possible that they are not responsible for the segregation in self-compatibility and are simply linked to the true causal locus. Genes possibly influencing seed set, such as cyclin dependent kinase

inhibitors, were also identified nearby significant SNPs. Once again, although these genes may play a role in determining seed set, they are not necessarily specific to seed set resulting from self-pollination.

While significant SNPs were spread out over chromosome 12 and it is likely that more than one gene was at work in determining self-compatibility, we considered the S-protein homologue which is located within the region of the top five most significant SNPs to be the leading candidate gene for self-compatibility in the DRH F_1 population. This gene is most greatly expressed in the stamens as would be expected for a gene which alters self-compatibility. Further support for the qualification of the S-protein homologue as a superior candidate is the fact that SNPs surrounding this gene became fixed in the inbred lines. Since each generation could only be propagated from parents which successfully produced fruit and seed we have been applying a selective pressure on the DRH population for the self-compatibility phenotype. The presence of fixed alleles surrounding the S-protein homologue indicates that this gene may be a necessary component of self-compatibility in order to reach the further inbred generations.

Within the S-protein homologue there is a plant self-incompatibility S1 domain noted for its similarity to the stigma-specific self-incompatibility determinant in *Papaver rhoeas*. This protein in *P. rhoeas* was discovered to be the female determinant in an entirely novel class of gametophytic self-incompatibility distinct from the system known in Solanaceae (Foote et al. 1994). As opposed to using ribonucleases to target the destruction of self-pollen tubes, the S-protein of *P. rhoeas* was proposed to inhibit pollen tube growth via a type of cytosolic calcium signal (Franklin-Tong et al. 1993). Genes sharing homology to the stigma specific S-protein in *P. rhoeas* have been identified in *Arabidopsis thaliana* and named S-protein homologues (Ride et al. 1999). These S-protein homologues in *A. thaliana* might not be restricted to self-compatibility but may serve a broader purpose in signaling (Ride et al. 1999). The S-protein homologue of interest to us in the DRH population has a protein match of 50.4 percent similarity and 81 percent coverage to an S-protein homologue in *A. thaliana* (AT5G12060.1). Although the S-protein homologue identified in the DRH population as a candidate for self-compatibility displays greatest expression in the stamens, expectations are that it would be expressed in the stigmatic tissue consistent with the S-protein in *P. rhoeas*. While there is a slight possibility that the DRH population is taking advantage of a similar gametophytic self-incompatibility system as *P. rhoeas*, as opposed to the traditional S-RNase system

known in Solanaceae, expression patterns indicate that the S-protein homologue in DRH is not a stigmatic determinant and may be functioning in a unique way.

The localization of the gene determining self-compatibility in the DRH population on chromosome 12 brings up several questions regarding previously identified compatibility genes in potato. Although the *S*-locus was previously localized on chromosome 1 of potato this conclusion was reached based on germplasm different from DRH. It is therefore not entirely impossible that the *S*-locus is in a different genomic region in the DRH population. Alternatively, the S-protein homologue may not be an S allele but rather a modifier, such as the *Sli* gene identified in *S. chacoense*. Indeed the *S. chacoense Sli* gene was localized the distal end of chromosome 12, as is the S-protein homologue in DRH. One notable similarity between the *P. rhoeas* S-protein and the *S. chacoense Sli* gene is that both function as pollen inhibitors. As the *Sli* gene has been reported to be nonallelic to the *S*-locus is possible that the S-protein homologue functions in a similar manner. More research needs to be done to further investigate these possibilities to better understand the true role of the S-protein homologue.

Identification of candidate genes for self-compatibility via segregation of SNPs for seed set in the DRH F_1 population yielded less conclusive results. The seed set analysis was necessarily restricted to the self-compatible genotypes and therefore represented only a subset of the F_1 population. Since a few significant SNPs were found nearby the region on chromosome 12 harboring the S-protein homologue it is possible that both seed set and fruit set share a gene located in this region and are both at least partially controlled by the same element. Alternatively it is possible that two separate genes control each trait but are simply located close to each other. It is also possible that SNPs for seed set segregated in this region due to a nearby cyclin dependent kinase inhibitor that was found rather than due to a self-compatibility gene.

Changes observed in the inbred lines

The phenotypic trends from one generation to the next revealed the effects of inbreeding depression from repeated selfing. Compared to the F_1 generation the inbred lines showed much greater rates of plants that could not be pollinated due to buds with poor pollen shed, buds that did not open, or plants that did not reach the budding stage altogether due to slow maturity or early death. While these rates fluctuated among the inbred generations, the most extreme reduction in flowering rate occurred during the transition

from the F_1 to the S_1 generation. These results are in agreement with previous records which claim that the most drastic changes in inbreeding depression in diploid potato occurred after the first selfing event (De Jong and Rowe 1971). The low rate of flowering in the inbred lines should be considered in any discussion of self-compatibility. Since phenotypic classifications cannot be determined for these genotypes, it remains unknown whether or not these genotypes exhibited allelic states favoring self-compatibility. By necessity these genotypes could not be included in the statistical analyses; however, their removal imposed a bias.

While fruit set rates remained rather constant in all four generations, the quick decrease in percentage of plants producing “high” fruit set in the S_1 and S_2 generations was likely another indicator of inbreeding depression or else reflected reshuffling of either alleles at modifier genes that controlled an alternative route to self-compatibility. In the S_3 generation the slight increase in percentage of plants producing “high” fruit set likely reflected the fact that we were selecting for self-compatibility, and that despite inbreeding depression the alleles favoring the self-compatible phenotype were slowly becoming more favored.

Over the four generations seed set continued to be unpredictable within genotypes. Over the course of inbreeding, the fruit set and seed set became less correlated, indicating that there were multiple factors that controlled these traits. Surprisingly, seed set in the S_3 generation surpassed that of the previous generations. All fruit with seed count greater than previously recorded were found on genotype SC-03, which was calculated to be approximately 95% homozygous. This finding demonstrates that through selection it is possible to find high seed producing selfed plants with a highly homozygous background.

Through examination of the segregation of alleles in the S_1 and S_2 generations we determined that many representative SNPs that were highly significant in the F_1 generation were no longer significant in the inbred lines. Upon discovering this lack of significance we formulated two hypotheses to explain these findings. The first hypothesis was that recombination during the selfing caused SNPs which were previously linked to causal genes in the F_1 generation to no longer be linked to these same genes in the inbred lines. If this were true, then the same candidate genes could be responsible for self-compatibility in the inbred lines, but these genes would no longer have co-segregated with the same SNPs. The second hypothesis was that an entirely different mechanism for determining self-compatibility occurred in the inbred lines. Based on segregation of alleles at only a few loci of interest we could not discriminate

between these two hypotheses. By performing a second genome-wide SNP analysis on the S_3 generation though we were able to assess allelic states on a more global level and therefore acquire a clearer picture of the segregation patterns in the inbred lines.

New mechanism for determining self-compatibility in the S_3 generation

Analysis of SNP segregation in the S_3 generation revealed many fixed alleles including those surrounding the primary gene of interest for fruit set on chromosome 12 in the F_1 generation. However, despite selecting for this particular region segregation for self-compatible phenotypes was still observed in the S_3 generation. Since the significant SNPs for fruit set in the S_3 generation did not correspond to any of the previously identified regions in the F_1 generation it is evident that there were different genes determining self-compatibility in the inbred lines. This finding supports the second hypothesis that the inbred lines exhibited an alternative method for determining self-compatibility. While the candidate genes from the F_1 generation were important in the S_3 lines as evidenced by the fixation of alleles, there remained another level of genetic control.

By identifying significant interactions among multiple combinations of representative SNPs on different chromosomes we have concluded that there are numerous genes with possible epistatic relationships that help determine the compatibility condition in the S_3 generation. Cluster analysis revealed haplotypes for both self-compatible and self-incompatible phenotypes. However, since not all clusters consisted of exclusively self-compatible or exclusively self-incompatible genotypes these haplotype trends are flexible and there was not a specific single combination of SNPs necessary for fruit production. Furthermore within each cluster there existed multiple haplotypes which could yield the self-compatible phenotype and multiple haplotypes which could yield the self-incompatible phenotype. We hypothesize therefore that even among these highly homozygous genotypes there remain a variety of gene combinations that can lead to self-compatibility. Further selfing and a continued decrease in heterozygosity would provide a means by which to narrow down specific genes in these interactions that are necessary for continued fruit set in a nearly homozygous background.

Regarding the two candidate genes identified on chromosome 6, the S-protein and S-protein homologue, perhaps these genes were partially aiding in the self-compatibility reaction in the F_1 generation due to the weakly significant SNPs surrounding these genes in the F_1 data set. Even if the S-

protein homologue was the primary determinant of self-compatibility there is still the possibility that other genes modified the compatibility reaction. After the primary determinant of self-compatibility has been fixed these modifying genes were brought to light in the segregation for self-compatibility in the inbred lines. If it is true that the inbred lines are taking advantage of a novel mechanism for determining self-compatibility, then our current knowledge of possible genes to search for based on previous literature is rendered much less useful than in the F_1 generation. While a few specific candidate genes for self-compatibility were identified nearby the most significant SNPs, it is possible that entirely unknown genes are responsible for self-compatibility at this point.

Reduction in heterozygosity

One of the goals of the project was to generate highly homozygous self-fertile inbred lines to develop a tool for future genetic studies. As demonstrated by the reduced percent of heterozygous SNPs in the S_3 generation it can be concluded that the inbred lines were overall more homozygous than the F_1 generation. While traditional methods for calculating the rate of heterozygosity reduction would predict a faster decline in heterozygosity (87.5%), the rates observed in this study (80%) are not entirely unexpected. While selfing diploid potato up to the S_5 generation using *Sli*, Phumichai (2005) also observed a slow reach of homozygosity, possibly due to selection for seed germination and vegetative vigor. Also in agreement with the findings of Phumichai (2005) is the lack of loci in the S_3 generation exhibiting uniform heterozygosity across all genotypes. This supports the concept that heterozygosity is not required for self-compatibility and that achieving a homozygous, self-propagating inbred line is not unachievable. We expect that many loci of maintained heterozygosity may yield a fitness advantage or otherwise prove beneficial in the selection for self-compatibility. The results showing that different families attained different numbers of unique fixed alleles demonstrates that not all families are achieving homozygosity at the same rate. This effect could possibly be due to selection of a different combination of genes or haplotypes for self-compatibility among the families. The tendency of more alleles to become fixed for DM alleles rather than the alternative allele derived from RH indicated that the fixation of alleles might not be solely due to neutral selection, but that a preference exists for the DM alleles. Of the 3 families in the S_3 generation, DRH 16 had the highest tolerance for fixation of RH alleles, demonstrating that the degree of similarity to the RH genotype was not consistent among families. Further selfing of the

inbred DRH lines and progression to increased homozygosity will help elucidate whether or not a minimum level of heterozygosity is needed to maintain the self-compatibility phenotype in the DRH population.

Potential sources of error

One of the greatest potential sources of error in the experiment is the presence of environmental variation since the genotype of a plant is not the sole determining factor of a plant's compatibility condition. The age of the plants, the growing season, and the weather conditions during pollination all play a role in self-compatibility. Therefore false-positive results for fruit and seed set may have arisen from a breakdown in self-incompatibility. While caution was taken to perform pollinations on separate days to help compensate for these environmental factors, it is still important to take these factors into account.

The presence of open pollinated fruits also posed a potential source of error in the experiment as we cannot be entirely sure that these fruit were the result of self pollinations as opposed to cross pollinations. Plants that consistently spontaneously produced fruit were generally considered self-compatible, even if a self-pollination had not been performed, as has been assumed in previous studies of plants under controlled conditions (Stone et al. 2006). Since open pollinated fruit generally appeared on plants which were already deemed self-compatible due to fruit production from intentional self-pollination and since open pollinated fruit generally appeared consistently on some plants but not on others it is likely that these fruit were the result of self-compatibility.

Conclusion

Through genome wide SNP analysis of the F_1 DRH generation we identified chromosome 12 as harboring the gene or genes responsible for the segregation of self-pollinated fruit set. Currently we believe that the S-protein homologue on the distal end of chromosome 12 was the primary gene responsible for the segregation of self-compatibility due to the highly significant SNPs found nearby and the fact that two of these SNPs became fixed for the RH alleles from 57,064,746 bp to 57,695,670 bp surrounded by fixed DM alleles from 55,640,932 bp to 55,982,568 bp and from 57,805,264 bp to 58,689,924 bp on either side of the S-protein homologue in subsequent generations under the selection pressure for the self-compatible phenotype. We also conclude that the S-protein homologue alone was not enough to elicit self-compatibility, but that a combination of other genes was also necessary. We identified regions

spanning nine chromosomes in which such genes may be found, and we believe that epistatic interactions among these genes are necessary to yield the self-compatible phenotype.

While a great reduction in heterozygosity occurred from the process of inbreeding for three generations, regions of heterozygosity still remained throughout the genome. While it remains to be determined whether a minimum level of heterozygosity is necessary to maintain self-compatibility, efforts to continue advancing the inbred generations in order to reduce heterozygosity will further clarify which genes are necessary for self-compatibility and will help in the production of more reliably self-compatible diploid potatoes.

References

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-265

Birhman RK, Hosaka K (2000) Production of inbred progenies of diploid potatoes using an *S*-locus inhibitor (*Sli*) gene, and their characterization. Genome 43:495-502

Bredemeijer GMM, Blaas J (1981) S-Specific proteins in styles of self-incompatible *Nicotiana alata*. Theor Appl Genet 59:185-190

Chen G, Zhang B, Liu LJ, Li Q, Zhang YE, Xie Q, Xue YB (2012) Identification of a ubiquitin-binding structure in the S-locus F-box protein controlling S-RNase-based self-incompatibility. J Genet Genom 39:93-102

Cipar MS, Peloquin SJ, Hougas RW (1964) Inheritance of incompatibility in hybrids between *Solanum tuberosum* haploids and diploid species. Euphytica 13:163-172

D'Hoop BB, Paulo MJ, Kowitwanich K, Sengers M, Visser RGF, van Eck HJ, van Eeuwijk FA (2010) Population structure and linkage disequilibrium unravelled in tetraploid potato. Theor Appl Genet 121:1151-1170

De Jong H, Rowe PR (1971) Inbreeding in cultivated diploid potatoes. Potato Res 14:74-83

de Nettancourt D (1977) Incompatibility in angiosperms. Springer-Verlag, Berlin: 2-134

Diggle PK (1995) Architectural effects and the interpretation of patterns of fruit and seed development. Annu Rev Ecol Syst 26:531-552

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull:11-15

East E (1934) Norms of pollen-tube growth in incompatible matings of self-sterile plants. Proc Natl Acad Sci USA 20:225-230

East E, Mangelsdorf A (1925) A new interpretation of the hereditary behavior of self-sterile plants. Proc Natl Acad Sci USA 11:166-171

Eijlander R, Ramanna MS, Jacobsen E (1997) Selection of vigorous and fertile S-homo- and heterozygous tester clones from self-incompatible diploid potato, *Solanum tuberosum* L. Euphytica 97:97-111

Felcher KJ, Coombs JJ, Massa AN, Hansey CN, Hamilton JP, Veilleux RE, Buell CR, Douches DS (2012) Integration of two diploid potato linkage maps with the potato genome sequence. PLoS One 7:e36347

Foote HCC, Ride JP, Franklintong VE, Walker EA, Lawrence MJ, Franklin FCH (1994) Cloning and Expression of a Distinctive Class of Self-Incompatibility (S) Gene from *Papaver rhoeas* L. Proceedings of the National Academy of Sciences of the United States of America 91:2265-2269

Franklin-Tong VE, Ride JP, Read ND, Trewavas AJ, Franklin FCH (1993) The self-incompatibility response in *Papaver rhoeas* is mediated by cytosolic free calcium. Plant J 4:163-177

Gagne JM, Downes BP, Shiu SH, Durski AM, Vierstra RD (2002) The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in *Arabidopsis*. Proc Natl Acad Sci USA 99:11519-11524

Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufmann H, Thompson RD, Bonierbale MW, Ganal MW, Tanksley SD, Salamini F (1991) RFLP maps of potato and their alignment with the homoeologous tomato genome. Theor Appl Genet 83:49-57

Hamilton JP, Hansey CN, Whitty BR, Stoffel K, Massa AN, Van Deynze A, De Jong WS, Douches DS, Buell CR (2011) Single nucleotide polymorphism discovery in elite north american potato germplasm. BMC Genomics 12

Hershko A, Heller H, Elias S, Ciechanover A (1983) Components of ubiquitin-protein ligase system - resolution, affinity purification, and role in protein breakdown. J Biol Chem 258:8206-8214

Hosaka K, Hanneman RE (1998a) Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 1. Detection of an S locus inhibitor (Sli) gene. Euphytica 99:191-197

Hosaka K, Hanneman RE (1998b) Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 2. Localization of an S locus inhibitor (Sli) gene on the potato genome using DNA markers. Euphytica 103:265-271

Hua ZH, Kao TH (2006) Identification and characterization of components of a putative *Petunia* S-locus F-box-containing E3 ligase complex involved in S-RNase-based self-incompatibility. Plant Cell 18:2531-2553

Hua ZH, Meng XY, Kao TH (2007) Comparison of *Petunia inflata S*-locus F-box protein (Pi SLF) with Pi SLF-like proteins reveals its unique function in S-RNase-based self-incompatibility. Plant Cell 19:3593-3609

Huang S, Lee HS, Karunanandaa B, Kao TH (1994) Ribonuclease activity of *Petunia inflata* S-proteins is essential for rejection of self-pollen. Plant Cell 6:1021-1028

Jacobs JME, Vaneck HJ, Arens P, Verkerkbakker B, Hekkert BTL, Bastiaanssen HJM, Elkharbotly A, Pereira A, Jacobsen E, Stiekema WJ (1995) A genetic map of potato (*Solanum tuberosum*) integrating molecular markers, including transposons, and classical markers. Theor Appl Genet 91:289-300

Lai Z, Ma WS, Han B, Liang LZ, Zhang YS, Hong GF, Xue YB (2002) An F-box gene linked to the self-incompatibility (S) locus of *Antirrhinum* is expressed specifically in pollen and tapetum. Plant Mol Biol 50:29-42

Lindhout P, Meijer D, Schotte T, Hutten RCB, Visser RGF, van Eck HJ (2011) Towards F_1 hybrid seed potato breeding. Potato Res 54:301-312

Lloyd DG, Schoen DJ (1992) Self- and cross-fertilization in plants .1. Functional dimensions. Int J Plant Sci 153:358-369

Massa AN, Childs KL, Lin H, Bryan GJ, Giuliano G, Buell CR (2011) The transcriptome of the reference potato genome *Solanum tuberosum* group Phureja clone DM1-3 516R44. PLoS One 6:e26801

McClure B, Cruz-García F, Romero C (2011) Compatibility and incompatibility in S-RNase-based systems. Ann Bot 108:647-658

McClure BA, Gray JE, Anderson MA, Clarke AE (1990) Self-incompatibility in *Nicotiana alata* involves degradation of pollen rRNA. Nature 347:757-760

McClure BA, Haring V, Ebert PR, Anderson MA, Simpson RJ, Sakiyama F, Clarke AE (1989) Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. Nature 342:955-957

Patton EE, Willems AR, Tyers M (1998) Combinatorial control in ubiquitin-dependent proteolysis: don't Skp the F-box hypothesis. Trends Genet 14:236-243

Phumichai C, Mori M, Kobayashi A, Kamijima O, Hosaka K (2005) Toward the development of highly homozygous diploid potato lines using the self-compatibility controlling *Sli* gene. Genome 48:977-984

Qiao H, Wang HY, Zhao L, Zhou JL, Huang J, Zhang YS, Xue YB (2004) The F-box protein AhSLF-S-2 physically interacts with S-RNases that may be inhibited by the ubiquitin/26S proteasome pathway of protein degradation during compatible pollination in antirrhinum. Plant Cell 16:582-595

Ride JP, Davies EM, Franklin FCH, Marshall DF (1999) Analysis of *Arabidopsis* genome sequence reveals a large new gene family in plants. Plant Mol Biol 39:927-932

Saba-El-Leil MK, Rivard S, Morse D, Cappadocia M (1994) The *S11* and *S13* self-incompatibility alleles in *Solanum chacoense* Bitt. are remarkably similar. Plant Mol Biol 24:571-583

SAS JMP. Version 10. SAS Institute Inc., Cary, NC

Sato S, Peet MM, Thomas JF (2000) Physiological factors limit fruit set of tomato (*Lycopersicon esculentum* Mill.) under chronic, mild heat stress. Plant Cell Environ 23:719-726

Solanaceae Coordinated Agricultural Project. http://solcap.msu.edu/. Accessed: 11 September, 2013

Stephenson AG, Travers SE, Mena-Ali JI, Winsor JA (2003) Pollen performance before and during the autotrophic-heterotrophic transition of pollen tube growth. Philos T Roy Soc B 358:1009-1017

Stone JL, Sasuclark MA, Blomberg CP (2006) Variation in the self-incompatibility response within and among populations of the tropical shrub *Witheringia solanacea* (Solanaceae). Am J Bot 93:592-598

Tao R, Iezzoni AF (Scientia Horticulturae2010) The S-RNase-based gametophytic self-incompatibility system in *Prunus* exhibits distinct genetic and molecular features. Sci Hortic-Amsterdam 124:423-433

Team RDC (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189-U194

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007) An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS One 2

Yasuda S (1934) Physiological research on self-incompatibility in *Petunia violacea*. Imperial College of Agriculture and Forestry 20:1-95

Figures

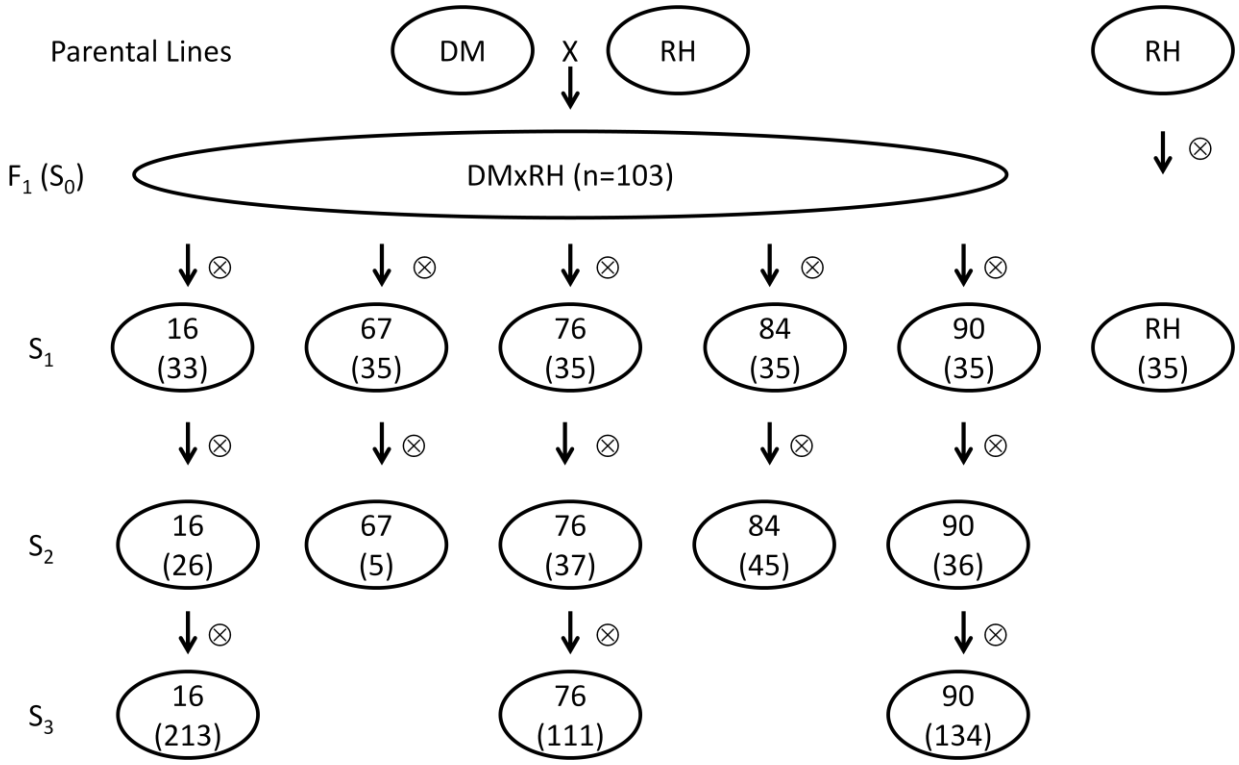
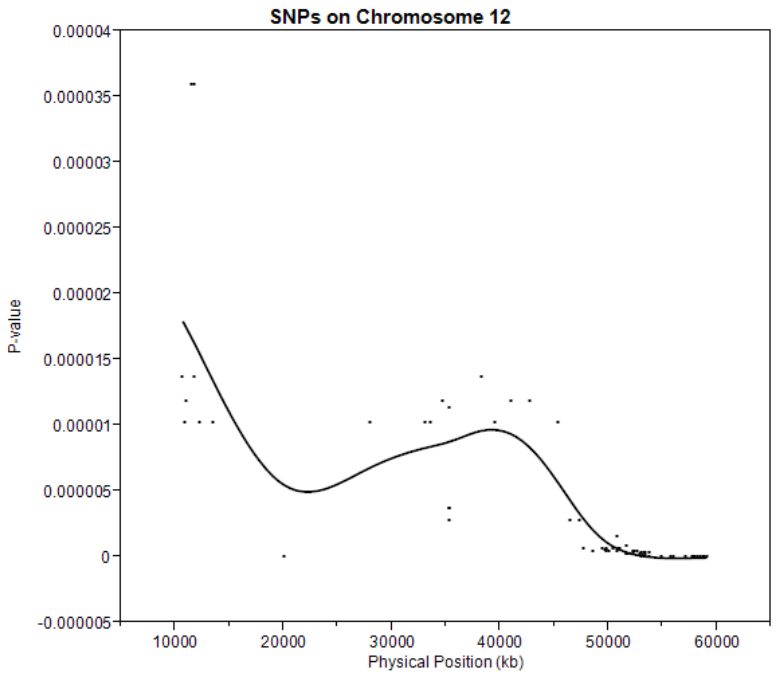


Fig. 1 Pedigree of DRH families used for self-incompatibility study

Inbred series of diploid potato generated from a cross between DM (female parent) and RH (male parent). Successive selfing in each generation yielded increasingly inbred progeny. Families in which no seed was successfully obtained from self-pollinations were not advanced to the next generation. Each generation is designated on the left. Numbers within parentheses indicate the number of individual genotypes within each family. The ⊗ symbol designates self-pollinations


SNPs on Chromosome 12
P-value
0.00004
0.000035
0.00003
0.000025
0.00002
0.000015
0.00001
0.000005
0
-0.000005
10000
20000
30000
40000
50000
60000
Physical Position (kb)

A

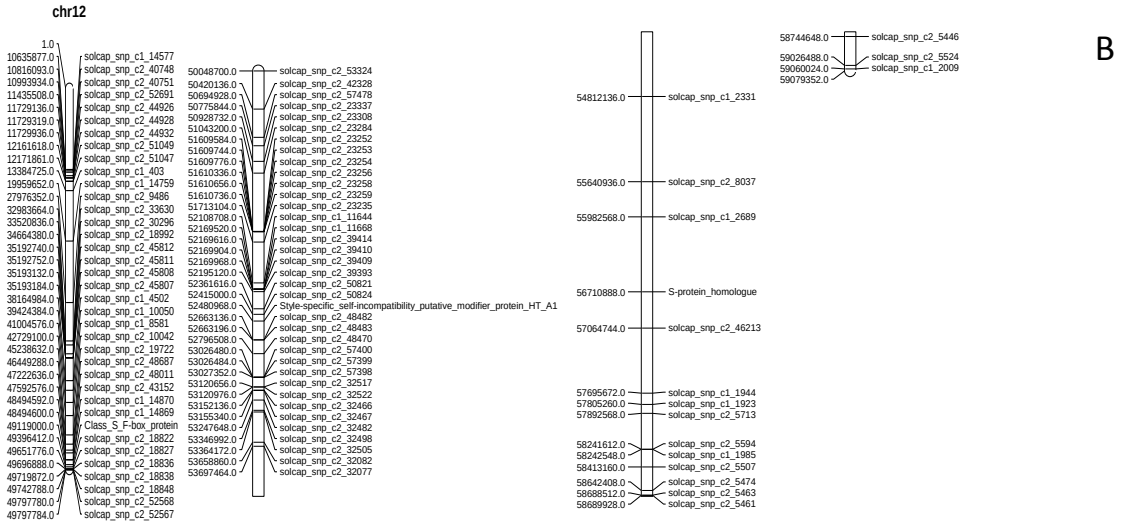
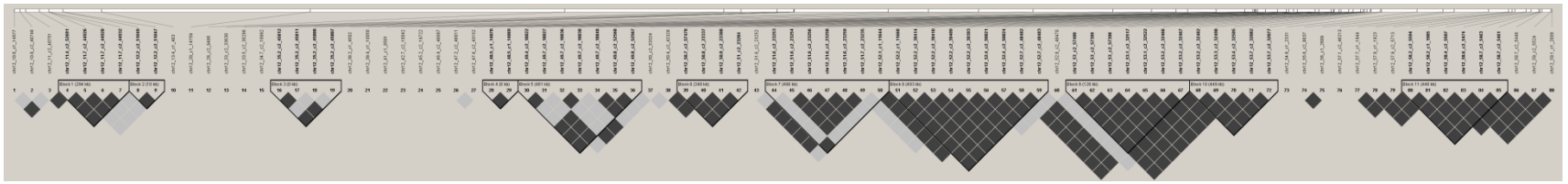


Fig. 2 Significant SNPs for fruit set in DRH F_1

A. Significant SNPs (p-value <0.0001) for fruit set in the F_1 generation plotted according to physical position along chromosome 12. B. Ideogram designating physical location on chromosome 12 of all significant SNPs for fruit set in the F_1 generation along with self-compatibility candidate genes identified. Positions along the chromosome are organized from top to bottom while segments of the chromosome are organized left to right

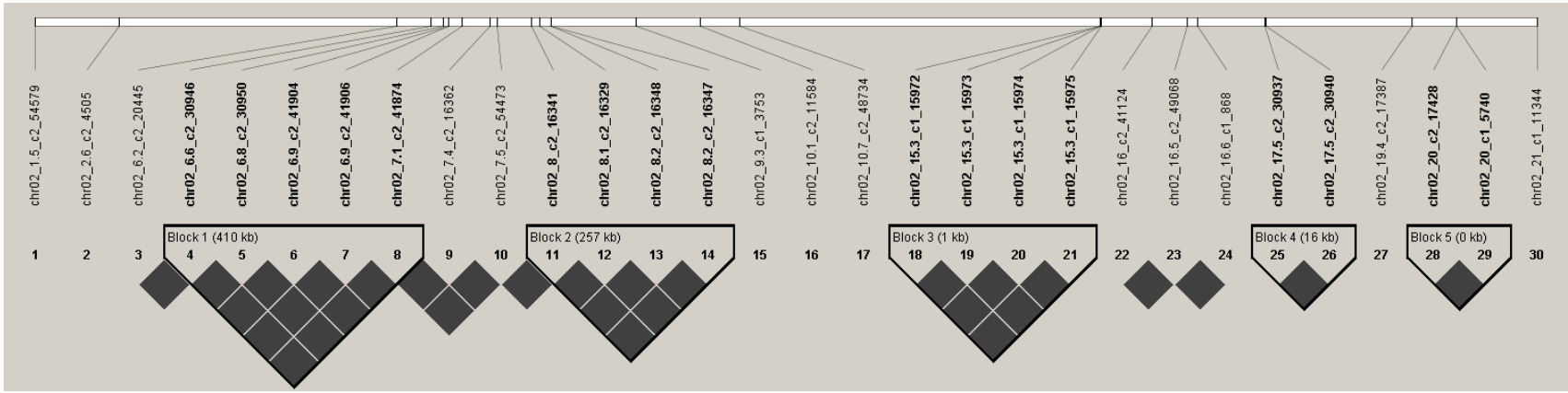
A.

Chromosome 12

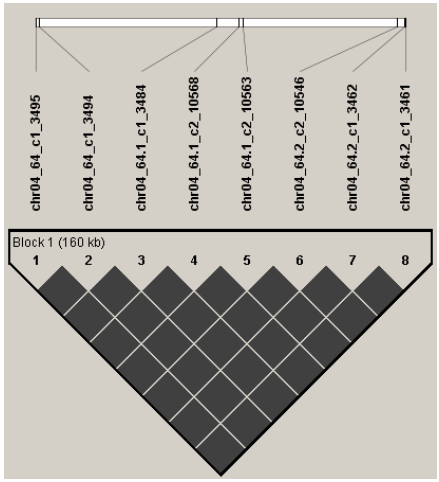


B.

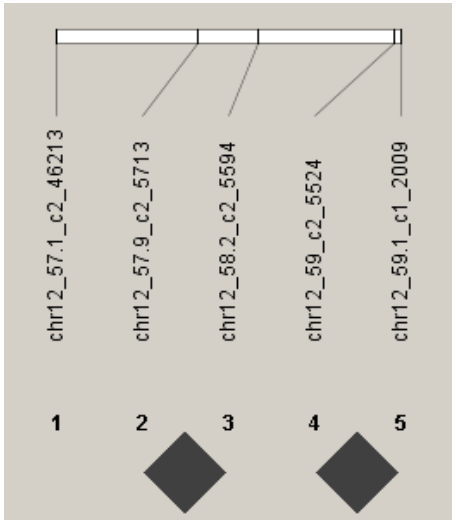
Chromosome 2



Chromosome 4

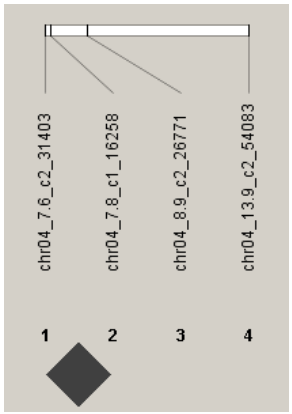


Chromosome 12

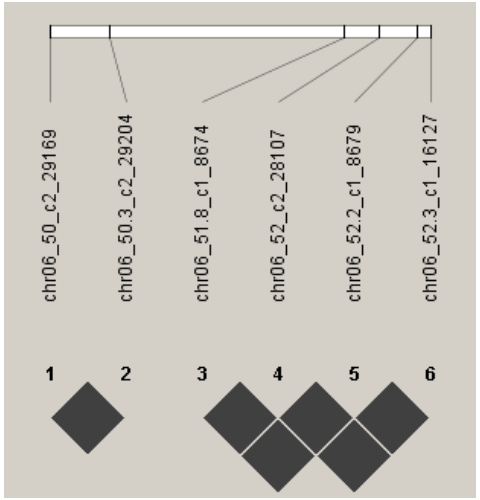


C.

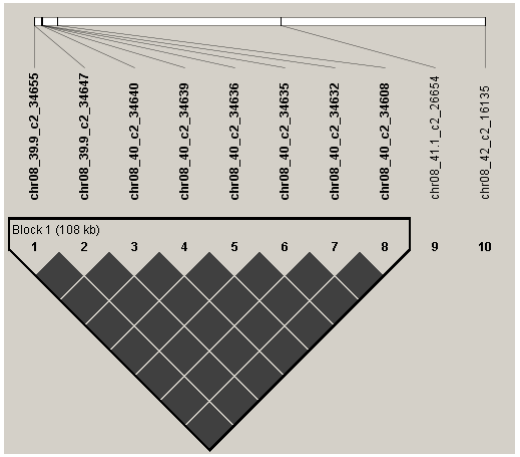
Chromosome 4



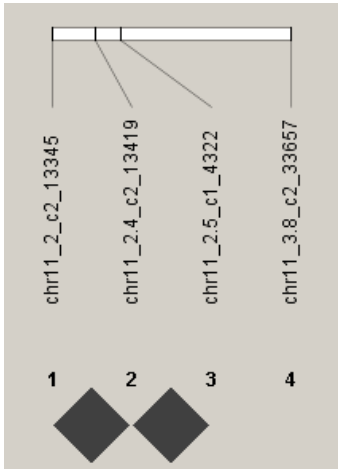
Chromosome 6



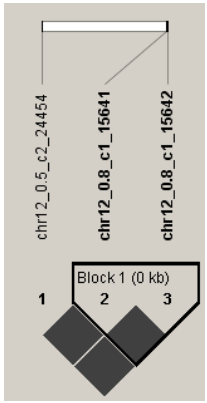
Chromosome 8



Chromosome 11

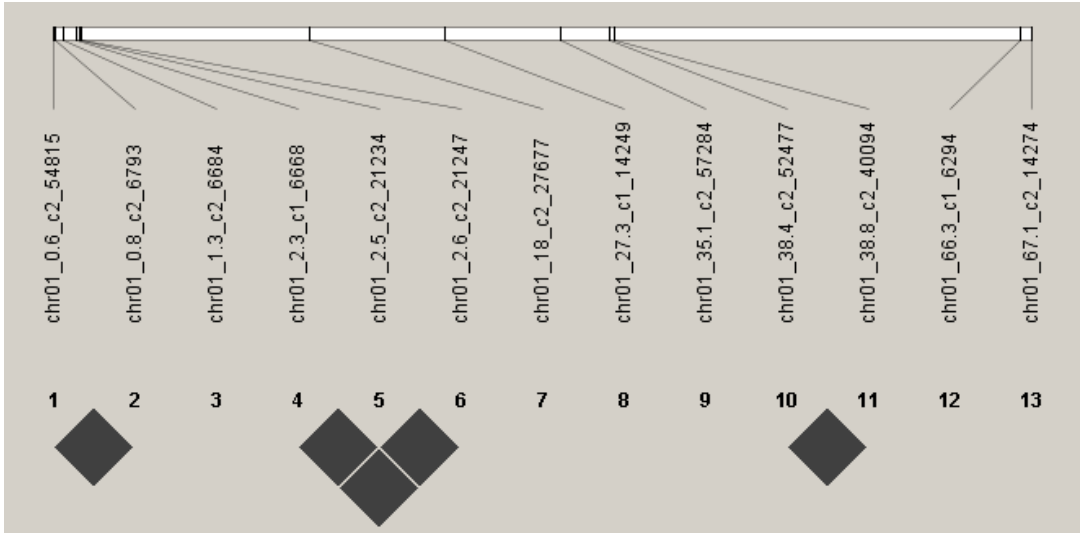


Chromosome 12

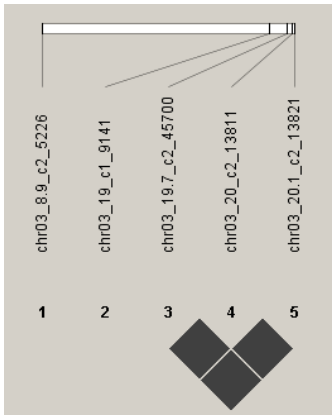


D.

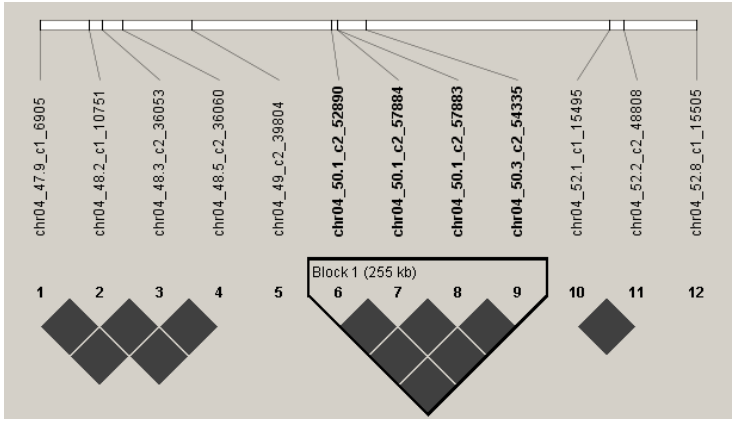
Chromosome 1



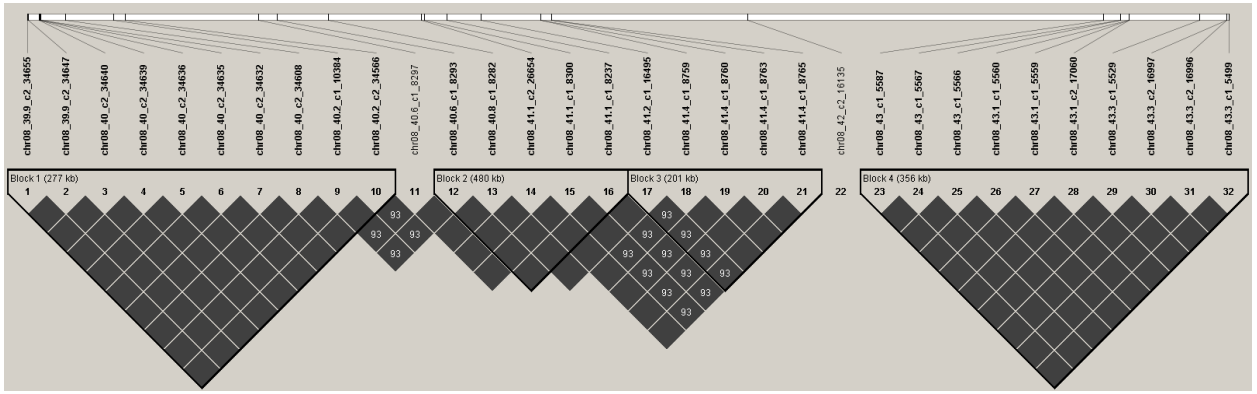
Chromosome 3



Chromosome 4



Chromosome 8



Chromosome 9

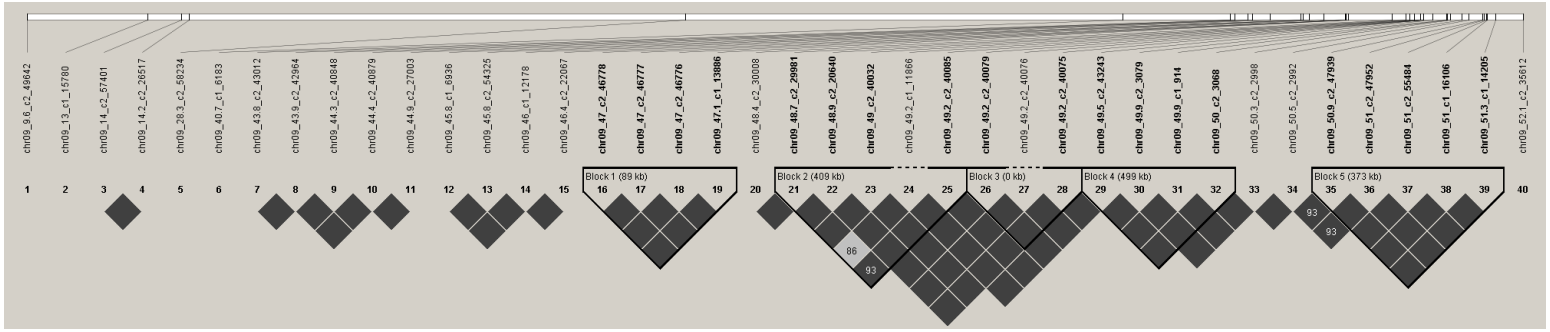


Fig. 3 Haplotype blocks of SNPs significantly correlated with fruit and seed set

Haplotypes blocks of significant SNPs correlated with fruit and seed set. Haplotypes blocks are indicated by bold triangles below SNP names. Dark gray diamonds indicate strong evidence of linkage disequilibrium while light gray diamonds indicate uninformative relationships. White bars above SNP names indicate physical position on the chromosome. Haplotype blocks were modeled separately for each chromosome harboring SNPs significant for fruit set in the F_1 generation (A), seed set in the F_1 generation (B), fruit set in the S_3 generation (C), and seed set in the S_3 generation (D).

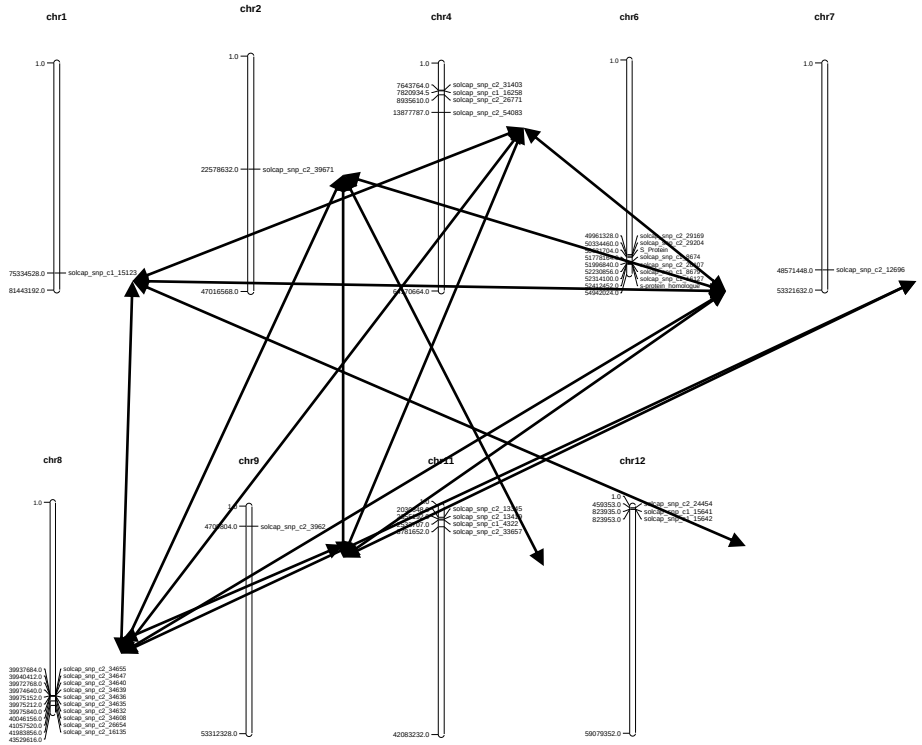


Fig. 4 Interactions among significant SNPs for fruit set in DRH S_3

Ideogram designating physical location on chromosomes 1, 2, 4, 6, 7, 8, 9, 11, and 12 of all significant SNPs for fruit set in the S_3 generation along with self-compatibility candidate genes identified. Positions along the chromosome are organized from top to bottom. Lines between chromosomes designate significant interactions which have been identified according to contingency analysis

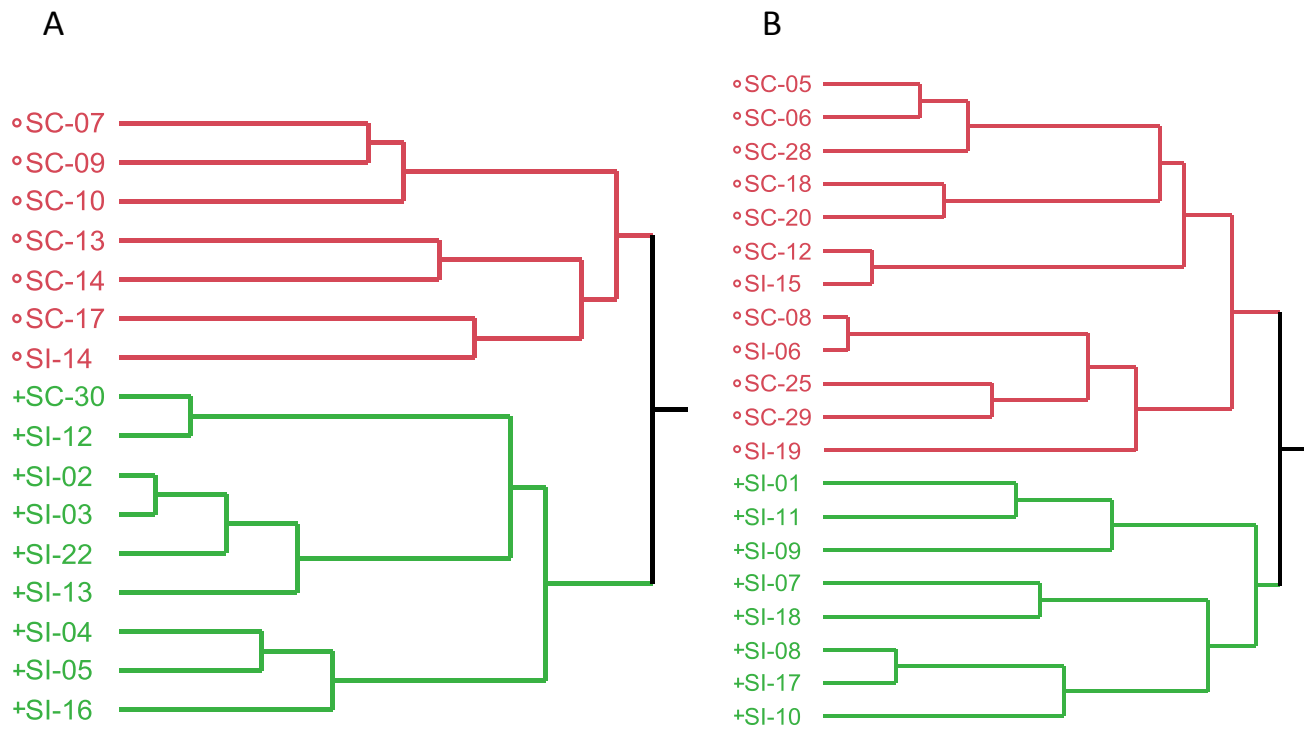
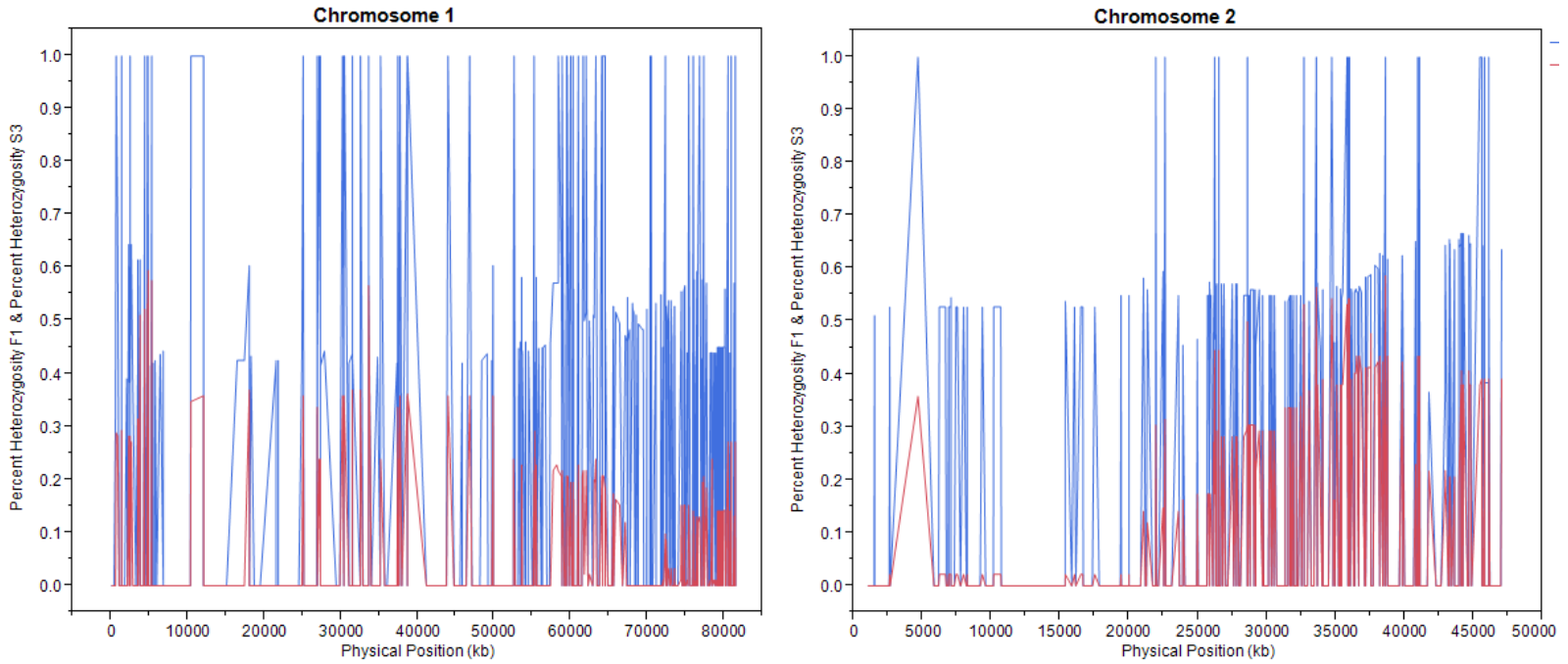
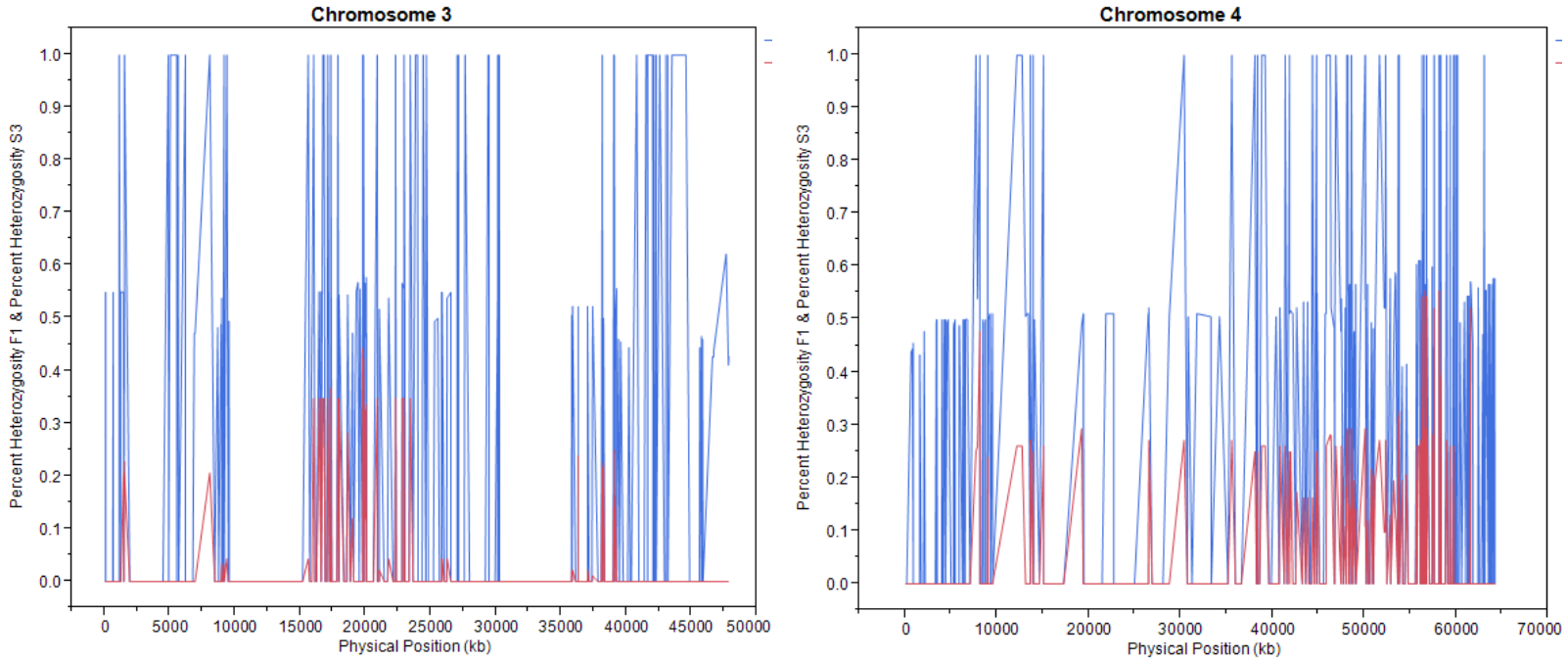
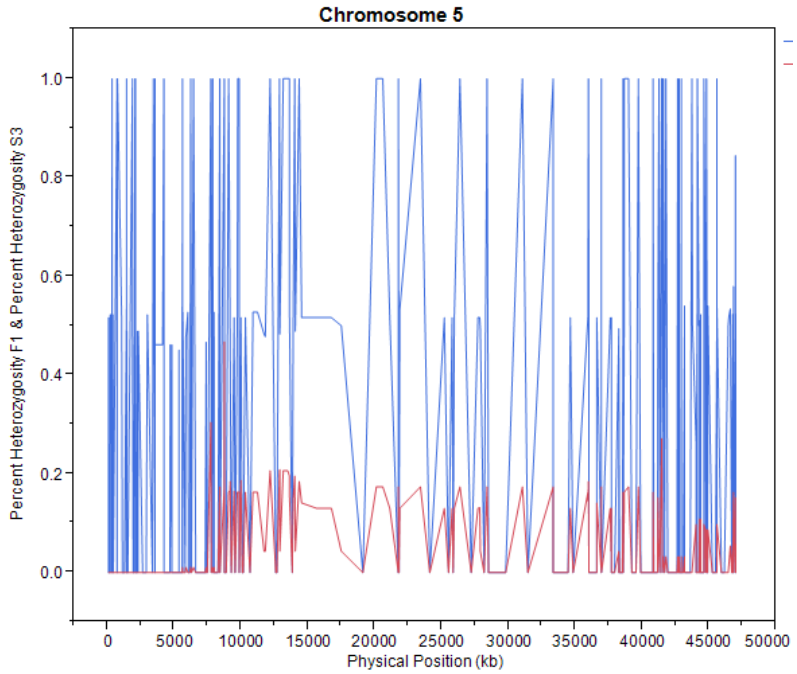


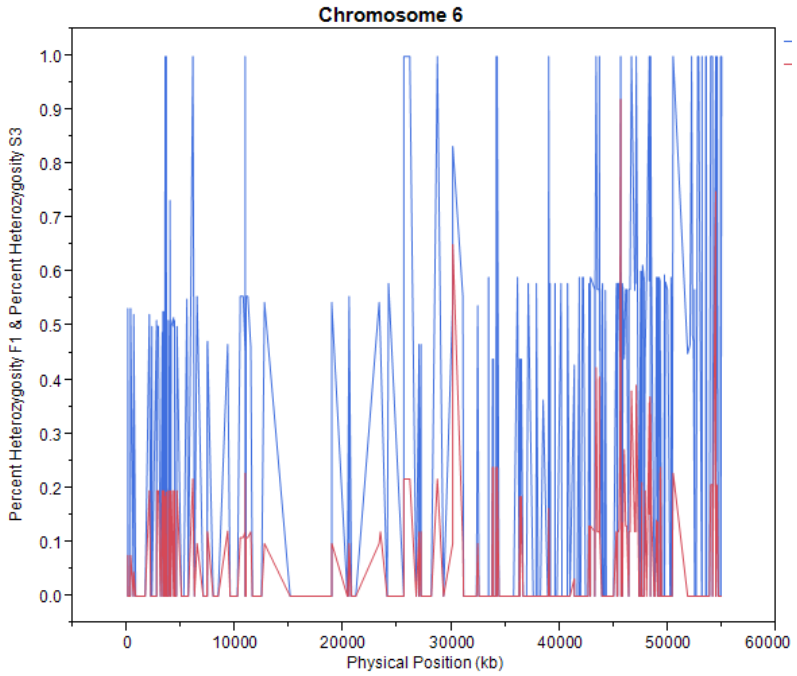
Fig. 5 Cluster analysis of haplotypes of two DRH S_3 families

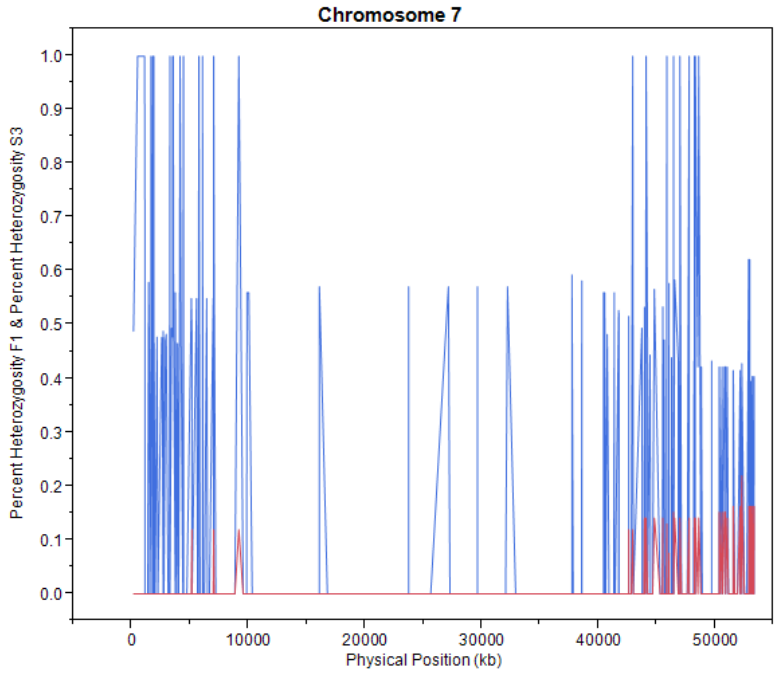
Dendrogram of S_3 genotypes. Clustering based off of allelic states at nine loci representing the nine chromosomes on which significant SNPs for fruit set were found in the S_3 generation. Both self-compatible (SC) and self-incompatible (SI) genotypes are indicated to the left. A. Dendrogram of S_3 genotypes derived from the DRH 76 family. B. Dendrogram of S_3 genotypes derived from the DRH 90 family. Circles and cross-marks designate separate clusters according to haplotype trends

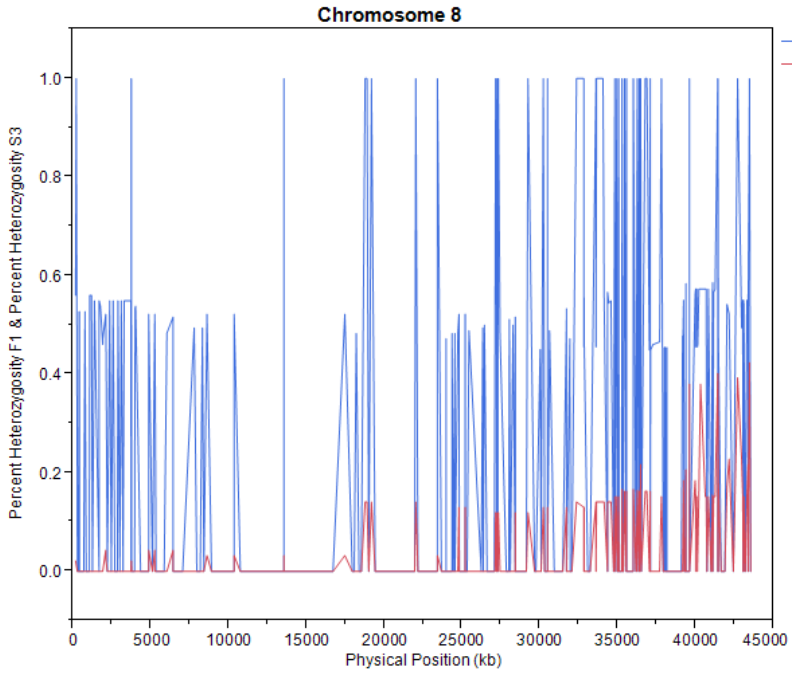

Chromosome 1
Percent Heterozygosity F1 & Percent Heterozygosity S3
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10000
20000
30000
40000
50000
60000
70000
80000
Physical Position (kb)
Chromosome 2
Percent Heterozygosity F1 & Percent Heterozygosity S3
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
5000
10000
15000
20000
25000
30000
35000
40000
45000
50000
Physical Position (kb)

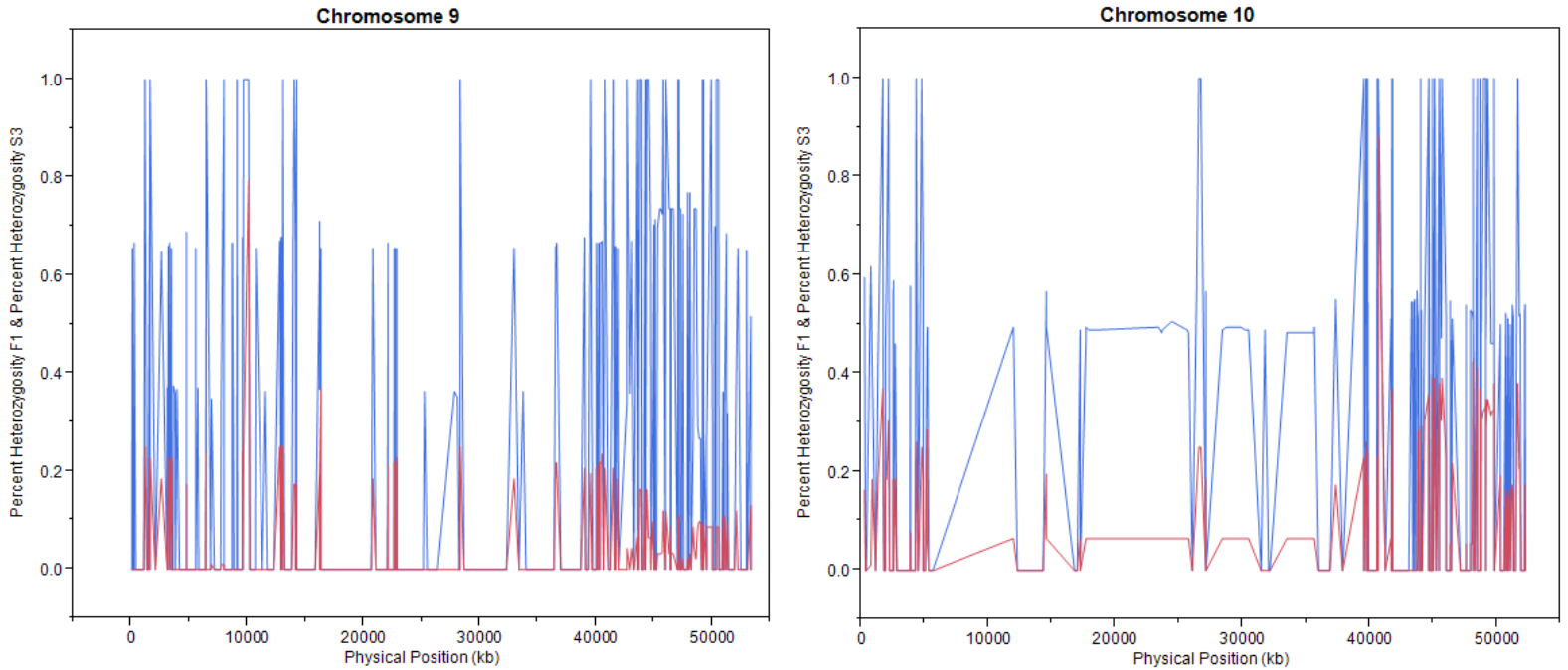

Chromosome 3
Percent Heterozygosity F1 & Percent Heterozygosity S3
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
5000
10000
15000
20000
25000
30000
35000
40000
45000
50000
Physical Position (kb)
Chromosome 4
Percent Heterozygosity F1 & Percent Heterozygosity S3
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10000
20000
30000
40000
50000
60000
70000
Physical Position (kb)


Chromosome 5
Percent Heterozygosity F1 & Percent Heterozygosity S3
1.0
0.8
0.6
0.4
0.2
0.0
0
5000
10000
15000
20000
25000
30000
35000
40000
45000
50000
Physical Position (kb)


Chromosome 6
Percent Heterozygosity F1 & Percent Heterozygosity S3
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10000
20000
30000
40000
50000
60000
Physical Position (kb)


Chromosome 7
Percent Heterozygosity F1 & Percent Heterozygosity S3
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10000
20000
30000
40000
50000
Physical Position (kb)


Chromosome 8
Percent Heterozygosity F1 & Percent Heterozygosity S3
1.0
0.8
0.6
0.4
0.2
0.0
0
5000
10000
15000
20000
25000
30000
35000
40000
45000
Physical Position (kb)


Chromosome 9
Percent Heterozygosity F1 & Percent Heterozygosity S3
1.0
0.8
0.6
0.4
0.2
0.0
0
10000
20000
30000
40000
50000
Physical Position (kb)
Chromosome 10
Percent Heterozygosity F1 & Percent Heterozygosity S3
1.0
0.8
0.6
0.4
0.2
0.0
0
10000
20000
30000
40000
50000
Physical Position (kb)

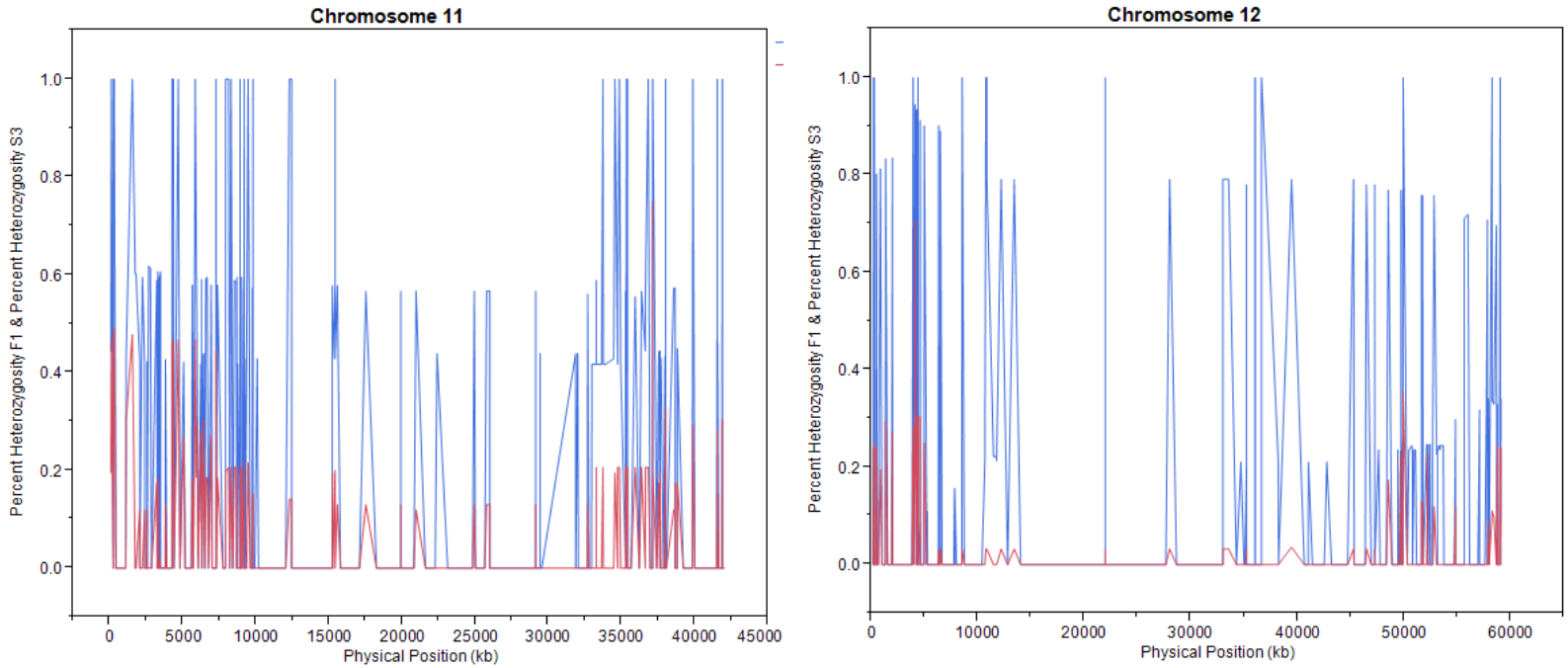


Fig. 6 Reduction of heterozygosity via inbreeding of DRH

Overlay plots for chromosomes 1-12 displaying the reduction in heterozygosity observed from the F_1 to the S_3 generation. Blue lines represent heterozygosity ratios in the F_1 generation while red lines represent heterozygosity ratios in the S_3 generation. Loci mapped based on physical position (kp)

Tables

Table 1 Observed phenotypes in the DRH population over four generations

Generation	Number of genotypes planted	Flowering rate	Fruit present	High fruit set	Seed set range	Correlation of fruit to seed[b]
F_1	103	95%	35%	69%	1-171	p-value = 0.0175
S_1	173	40%	35%	44%	1-111	p-value = 0.1399
S_2	149	37%	32%	38%	1-86[a]	p-value = 0.3176
S_3	458	45%	33%	53%	1-361	p-value = 0.0624

[a] Average seed count

[b] Bivariate fit of the number of fruits-per-pollination by the number of seeds per fruit

Table 2 Candidate genes for self-compatibility in the DRH population

Gene annotation	Gene ID[a]	Generation	Trait[b]	Chr[c]	Nearest significant SNP	Distance to SNP (bp)	p-value
Class S F-box protein	PGSC0003DMG400008762	F_1	Fruit	12	solcap_snp_c2_18822	277412	6.13E-07
Style-specific self-incompatibility putative modifier protein HT A1	PGSC0003DMG400010793	F_1	Fruit	12	solcap_snp_c2_50824	65967	3.85E-07
S-protein homologue	PGSC0003DMG400008637	F_1	Fruit	12	solcap_snp_c2_46213	353860	7.13E-12
Class S F-box protein	PGSC0003DMG400041057	F_1	Seed	2	solcap_snp_c2_17387	74440	<0.0001
S-protein	PGSC0003DMG400026738	S_3	Fruit	6	solcap_snp_c2_29204	297243	0.00096
S-protein homologue	PGSC0003DMG400037544	S_3	Fruit	6	solcap_snp_c1_16127	98353	0.00096
Class S F-box protein	PGSC0003DMG400024834	S_3	Seed	4	solcap_snp_c2_54335	770062	<0.0001
S-protein homologue	PGSC0003DMG400011582	S_3	Seed	9	solcap_snp_c2_43243	44453	<0.0001
S-protein homologue	PGSC0003DMG400035767	S_3	Seed	9	solcap_snp_c2_43243	50877	<0.0001
S-protein homologue	PGSC0003DMG400046953	S_3	Seed	9	solcap_snp_c2_43243	119609	<0.0001
S-protein homologue	PGSC0003DMG400011584	S_3	Seed	9	solcap_snp_c2_43243	126158	<0.0001
S-protein homologue	PGSC0003DMG400041960	S_3	Seed	9	solcap_snp_c2_47939	11297	<0.0001

[a] Gene ID according to convention of the Potato Genome Sequencing Consortium (http://solanaceae.plantbiology.msu.edu/cgi-bin/gbrowse/potato_dm_v_2_1_11/)

[b] Trait used to identify segregation patterns of significant SNPs, “fruit” refers to number of fruits-per-pollination while “seed” refers to number of seeds per fruit

[c] Chromosome on which gene of interest is located

Table 3 Haplotypes of DRH F_1 genotypes in the "medium" category

SNP	DRH-007	DRH-011	DRH-016	DRH-024	DRH-032	DRH-047	DRH-090	DRH-091	DRH-093	DRH-164
solcap_snp_c1_14577	TT	TC	TC	TT	TT	TC	TT	TT	TT	TT
solcap_snp_c2_40748	TC	CC	CC	TC	TC	CC	TC	TC	TC	TC
solcap_snp_c2_40751	TT	TC	TC	TT	TT	TC	TT	TT	TT	TT
solcap_snp_c2_52691	AA	AG	AG	AA	AA	AG	AA	AA	AA	AA
solcap_snp_c2_44926	CC	TC	TC	CC	CC	TC	CC	CC	CC	CC
solcap_snp_c2_44928	GG	AG	AG	GG	GG	AG	GG	GG	GG	GG
solcap_snp_c2_44932	TT	TC	TC	TT	TT	TC	TT	TT	TT	TT
solcap_snp_c2_51049	AC	CC	CC	AC	AC	CC	AC	AC	AC	AC
solcap_snp_c2_51047	TG	GG	GG	TG	TG	GG	TG	TG	TG	TG
solcap_snp_c1_403	TG	GG	GG	TG	TG	GG	TG	TG	TG	TG
solcap_snp_c1_14759	CC	TC	TC	CC	CC	CC	TC	CC	TC	CC
solcap_snp_c2_9486	AG	GG	GG	AG	AG	GG	AG	AG	AG	AG
solcap_snp_c2_33630	TC	TT	TT	TC	TC	TT	TC	TC	TC	TC
solcap_snp_c2_30296	TC	TT	TT	TC	TC	TT	TC	TC	TC	TC
solcap_snp_c2_18992	AA	AC	AC	AA	AA	AC	AA	AA	AA	AA
solcap_snp_c2_45812	CC	TC	TC	CC	CC	TC	CC	CC	CC	CC
solcap_snp_c2_45811	AA	AG	AG	AA	AA	AG	AA	AA	AA	AA
solcap_snp_c2_45808	TG	GG	GG	TG	TG	GG	TG	TG	TG	TG
solcap_snp_c2_45807	GG	AG	AG	GG	GG	AG	GG	GG	GG	GG
solcap_snp_c1_4502	AA	AG	AG	AA	AA	AG	AA	AA	AA	AA
solcap_snp_c1_10050	TC	CC	CC	TC	TC	CC	TC	TC	TC	TC
solcap_snp_c1_8581	GG	AG	AG	GG	GG	AG	GG	GG	GG	GG
solcap_snp_c2_10042	CC	TC	TC	CC	CC	TC	CC	CC	CC	CC
solcap_snp_c2_19722	TC	TT	TT	TC	TC	TT	TC	TC	TC	TC
solcap_snp_c2_48687	AG	GG	GG	AG	AG	GG	AG	AG	AG	AG
solcap_snp_c2_48011	AG	AA	AA	AG	AG	AA	AG	AG	AG	AG
solcap_snp_c2_43152	CC	TC	TC	CC	CC	TC	CC	CC	CC	CC
solcap_snp_c1_14870	TC	TT	TT	TC	TC	TT	TC	TC	TC	TC
solcap_snp_c1_14869	AG	AA	AA	AG	AG	AA	AG	AG	AG	AG
solcap_snp_c2_18822	AA	AG	AG	AA	AA	AG	AA	AA	AA	AA
solcap_snp_c2_18827	GG	AG	AG	GG	GG	AG	GG	GG	GG	GG
solcap_snp_c2_18836	AG	GG	GG	AG	AG	GG	AG	AG	AG	AG
solcap_snp_c2_18838	AG	GG	GG	AG	AG	GG	AG	AG	AG	AG
solcap_snp_c2_18848	CC	TC	TC	CC	CC	TC	CC	CC	CC	CC
solcap_snp_c2_52568	AA	AC	AC	AA	AA	AC	AA	AA	AA	AA
solcap_snp_c2_52567	TT	TC	TC	TT	TT	TC	TT	TT	TT	TT
solcap_snp_c2_53324	AG	GG	GG	AG	AG	GG	AG	AG	AG	AG
solcap_snp_c2_42328	CC	AC	AC	CC	CC	AC	CC	CC	CC	CC
solcap_snp_c2_57478	TT	TC	TC	TT	TT	TC	TT	TT	TT	TT

Table 3 Haplotypes of DRH F_1 genotypes in the "medium" category

solcap_snp_c2_23337	CC	TC	TC	CC	CC	TC	CC	CC	CC	CC
solcap_snp_c2_23308	AA	AG	AG	AA	AA	AG	AA	AA	AA	AA
solcap_snp_c2_23284	GG	AG	AG	GG	GG	AG	GG	GG	GG	GG
solcap_snp_c2_23252	AC	CC	CC	AC	AC	CC	AC	AC	CC	AC
solcap_snp_c2_23253	TT	AT	AT	TT	TT	AT	TT	TT	AT	TT
solcap_snp_c2_23254	AG	GG	GG	AG	AG	GG	AG	AG	GG	AG
solcap_snp_c2_23256	AC	AA	AA	AC	AC	AA	AC	AC	AA	AC
solcap_snp_c2_23258	TC	CC	CC	TC	TC	CC	TC	TC	CC	TC
solcap_snp_c2_23259	TC	CC	CC	TC	TC	CC	TC	TC	CC	TC
solcap_snp_c2_23235	AC	CC	CC	AC	AC	CC	AC	AC	CC	AC
solcap_snp_c1_11644	GG	AG	AG	GG	GG	AG	GG	GG	AG	GG
solcap_snp_c1_11668	AA	AG	AG	AA	AA	AG	AA	AA	AG	AA
solcap_snp_c2_39414	GG	TG	TG	GG	GG	TG	GG	GG	TG	GG
solcap_snp_c2_39410	GG	AG	AG	GG	GG	AG	GG	GG	AG	GG
solcap_snp_c2_39409	AA	AG	AG	AA	AA	AG	AA	AA	AG	AA
solcap_snp_c2_39393	GG	AG	AG	GG	GG	AG	GG	GG	AG	GG
solcap_snp_c2_50821	GG	TG	TG	GG	GG	TG	GG	GG	TG	GG
solcap_snp_c2_50824	TT	TC	TC	TT	TT	TC	TT	TT	TC	TT
solcap_snp_c2_48482	AA	AG	AG	AA	AA	AG	AA	AA	AG	AA
solcap_snp_c2_48483	GG	AG	AG	GG	GG	AG	GG	GG	AG	GG
solcap_snp_c2_48470	AG	AA	AA	AG	AG	AA	AG	AG	AA	AG
solcap_snp_c2_57400	TT	TC	TC	TT	TT	TC	TT	TT	TC	TT
solcap_snp_c2_57399	CC	TC	TC	CC	CC	TC	CC	CC	TC	CC
solcap_snp_c2_57398	CC	TC	TC	CC	CC	TC	CC	CC	TC	CC
solcap_snp_c2_32517	GG	AG	AG	GG	GG	AG	GG	GG	AG	GG
solcap_snp_c2_32522	GG	AG	AG	GG	GG	AG	GG	GG	AG	GG
solcap_snp_c2_32466	CC	AC	AC	CC	CC	AC	CC	CC	AC	CC
solcap_snp_c2_32467	TT	TC	TC	TT	TT	TC	TT	TT	TC	TT
solcap_snp_c2_32482	AA	AG	AG	AA	AA	AG	AA	AA	AG	AA
solcap_snp_c2_32498	TT	TC	TC	TT	TT	TC	TT	TT	TC	TT
solcap_snp_c2_32505	GG	AG	AG	GG	GG	AG	GG	GG	AG	GG
solcap_snp_c2_32082	TT	TC	TC	TT	TT	TC	TT	TT	TC	TT
solcap_snp_c2_32077	TT	TC	TC	TT	TT	TC	TT	TT	TC	TT
solcap_snp_c1_2331	AA	AG	AG	AA	AA	AG	AG	AA	AG	AA
solcap_snp_c2_8037	TC	TT	TT	TC	TC	TT	--	TC	TT	TC
solcap_snp_c1_2689	TC	TT	TT	TC	TC	TC	--	TC	TT	TC
solcap_snp_c2_46213	GG	AG	AG	GG	GG	GG	AG	GG	AG	GG
solcap_snp_c1_1944	GG	AG	AG	GG	GG	GG	AG	GG	AG	GG
solcap_snp_c1_1923	TC	CC	CC	TC	TC	TC	--	TC	CC	TC
solcap_snp_c2_5713	TT	AT	AT	TT	TT	TT	AT	TT	AT	TT
solcap_snp_c2_5594	GG	AG	AG	GG	GG	GG	AG	GG	AG	GG

Table 3 Haplotypes of DRH F_1 genotypes in the "medium" category

solcap_snp_c1_1985	TT	TC	TC	TT	TT	TT	TC	TT	TC	TT
solcap_snp_c2_5507	CC	CC	TC	CC	CC	CC	TC	CC	TC	CC
solcap_snp_c2_5474	AG	AG	AA	AG	AG	AG	--	AG	AA	AG
solcap_snp_c2_5463	TC	TC	TT	TC	TC	TC	--	TC	TT	TC
solcap_snp_c2_5461	AG	AG	GG	AG	AG	AG	--	AG	GG	AG
solcap_snp_c2_5446	TT	TT	TC	TT	TT	TT	TC	TT	TC	TT
solcap_snp_c2_5524	CC	CC	TC	CC	CC	CC	TC	CC	TC	CC
solcap_snp_c1_2009	TT	TT	TC	TT	TT	TT	TC	TT	TC	TT
Number of loci with "high" allelic state	3	78	85	3	3	70	14	3	44	3
Number of loci with "low" allelic state	85	10	3	85	85	18	68	85	44	85
Phenotype conclusion	Low	High	High	Low	Low	High	Low	Low	Un-known	Low

Dark gray alleles correspond to the “low” genotypes, while light gray alleles correspond to the “high” genotypes

Appendix

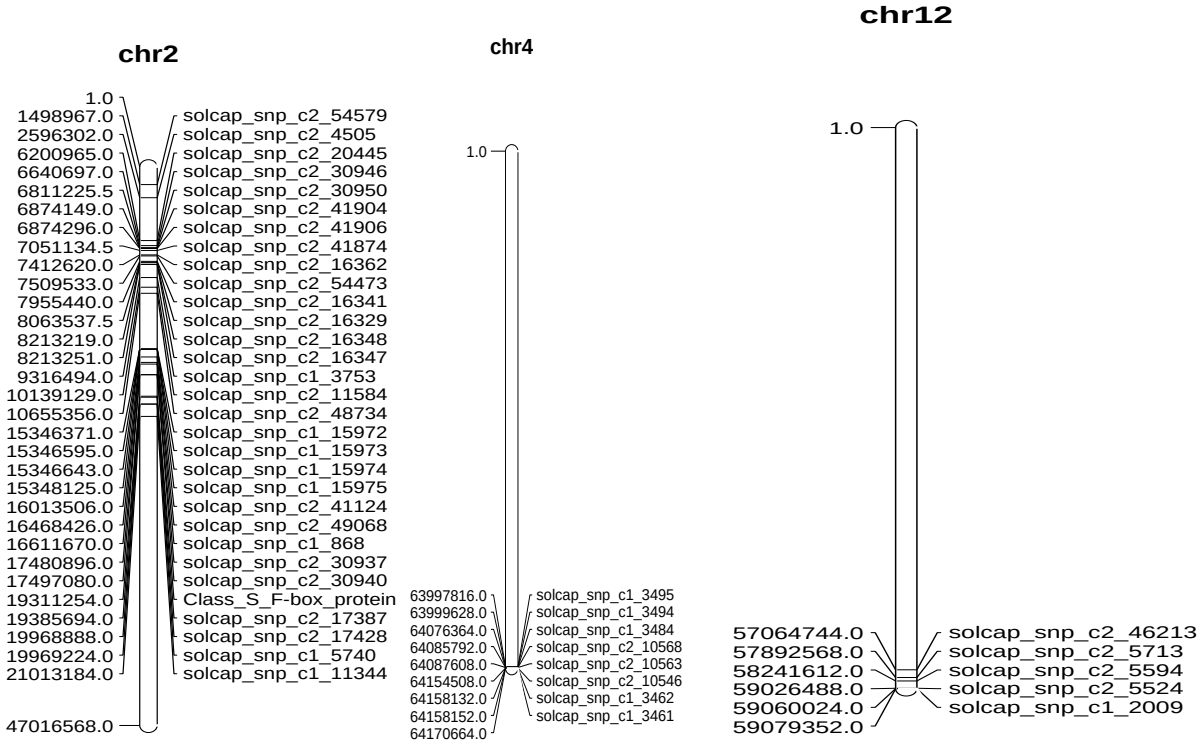


Fig. S1 Significant SNPs and candidate genes for seed set in DRH F_1

Ideogram designating physical location on chromosomes 2, 4, and 12 of all significant SNPs for seed set in the F_1 generation along with self-compatibility candidate genes identified. Positions along the chromosome are organized from top to bottom.

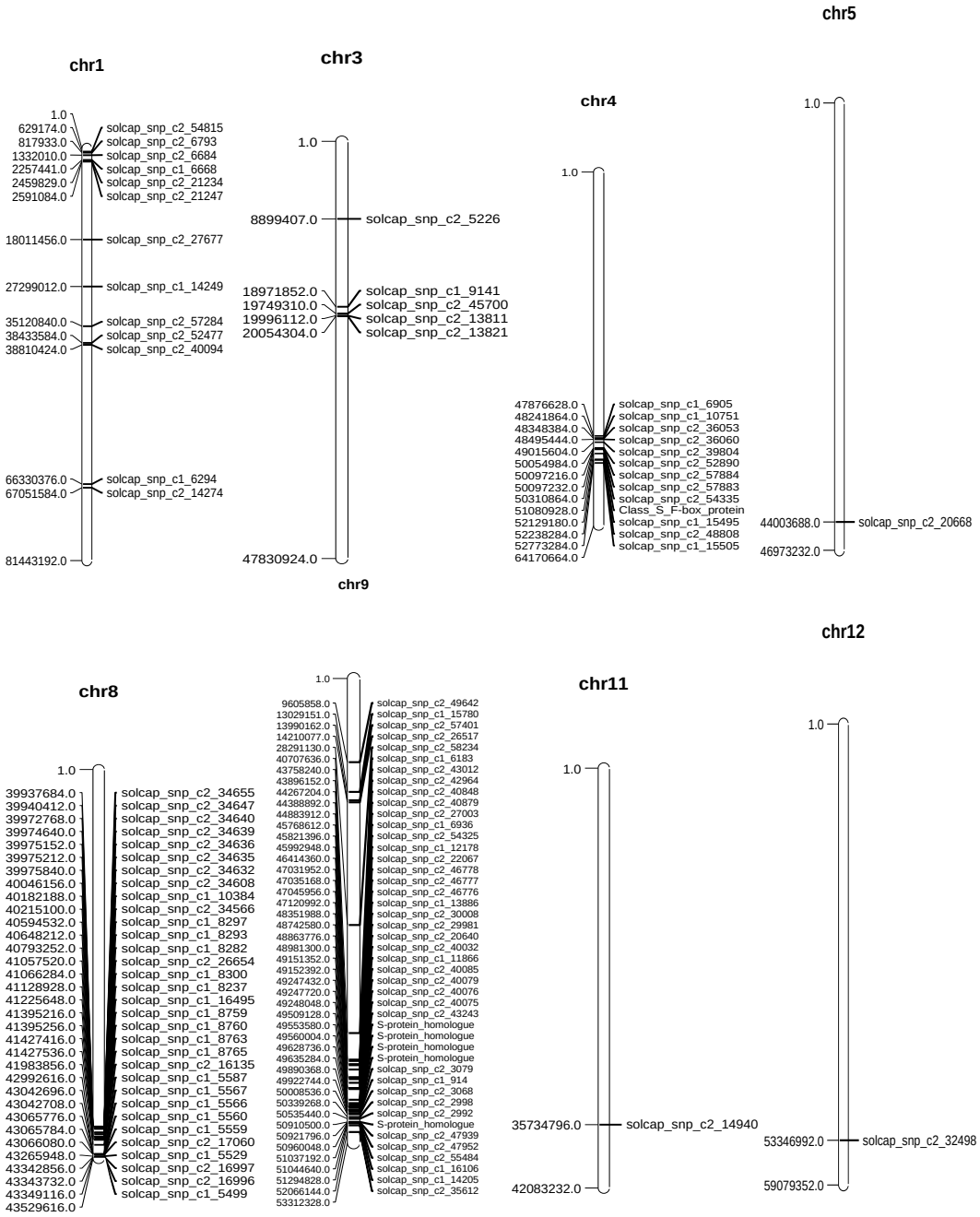


Fig. S2 Significant SNPs and candidate genes for seed set in DRH S_3

Ideogram designating physical location on chromosomes 1, 3, 4, 5, 8, 9, 11, and 12 of all significant SNPs for seed set in the S_3 generation along with self-compatibility candidate genes identified. Positions along the chromosome are organized from top to bottom

Table S1 Allele specific primers			
Primers	SNP Targeted	Allele	Sequence
1	solcap_snp_c1_2689	T/C	CTTTCCGTATTACTTATTGCTGCCGGAGTTCTTATCGCAGCTCTTGGAGA[T/C]TTTTCCTTTGAT CTTTTTGGATACAGTTTGGCCTTTATTTCTGTTTTCTT
2	solcap_snp_c2_23308	C/T	TTCCGATGTAATGTAGCCTTTTCTACCAGAATGTCTCTCTTTATAGTTTG[C/T]TGTTTATGGGAT CTATCAGCTTATCTGTTCAGCAAGATCTTGTTGGTTTC
3	solcap_snp_c2_40748	G/A	AAGATCAACAATCTTCTTCTACAACTCCGAATCTCTCAACAATCATGATC[G/A]AATTCCAAAAGG TAACAGTCCTAAACCCCAAGAAAAACTAGAAGATCTCA
4	solcap_snp_c2_50824	T/C	CAGAAAACAGAATTAAGTACTACATTTTAAGCCCTTTCACCCCTAATACG[T/C]CTAGCTAATTGC ATATCCTTAGGCATAATCGTGACTCTTTTAGCATGAAT
5	solcap_snp_c2_5713	A/T	AAGAAGGGTGTCCCTAAATTTGCTGAAGATGGTATGGATGATGTTGTTGT[A/T]GAAGCACAAAC TTGTGAGTATAGTTTTAATTATTATAATAAATTGGATTT
6	solcap_snp_c2_57398	T/C	CTCAACAAAAACTCAAGGTACTTTATAGTTTCGGCTATATGGGGAGCTGG[T/C]GGCGGAGGCG TGAGGCTTGCTAATCTCGGAAATCAAGGTCAAAACGATTG
7	RH_snp_2361608	A/G	CAAGCAAAGCTAAGGCTTTTTCAACTCTATTTTTCTTGCAATACCCATCAATAAGTATTGAATAAG TAAAAGAGCTTGGAGCAACACCATTTGCTTTCATTCT[A/G]TCAAACCAAGATGAAACCTCGGAT ATAGGCGCTTTCGACTCAAACAGGGATTTTATAACAGTATTATATGTCACAACATTCGGTTTACA GTTTAATGACTCCATTTCATTAAACAA
8	RH_snp_2363183	G/A	ACGATTTAAATTAATCTCAAAAACTTGGTACACTCTCATTCAATCCTCAACATTTATAAATATCCA TCTCAACG[G/A]TACTACAATACCAAAGAAAGAATTCATCGTTTGTAGCCATTCCATTAAANNNNN NNNNNNNNNNNNNNNNNNNNNNNNATCGATTCTTTATAGTGACGATCACGATAATATTAATATC ATTTTACCAGATTTAGATCTGCCATATTTGTCTTTTACTCC
9	RH_snp_2374701	C/A	TTTAAAGTATTGTTTCACTATAAATAGGAGAGGAATGCATATTATTTATTTGATCATCAACAAATT AAATATCAATAAACTCAAACAAAATGGCATTCAAGGCAAATAT[C/A]NNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNGGTTATTGCAAGGGAAATGGTTGAGGGTAAGTTGTTTTAA TTGTAGTTTTAAGTACTAATTACACTTTCATATGACAAATTAAATTAAGTGCATAATGAA
10	RH_snp_2389842	T/C	GCAATGTATTTCTTCCCTTTATCATCTTCCAGAAAAAATCCTATATCCGTTATTGTCCATTTGCAT AAACGAGTATCGGGATTGAGTGGTGCATTTTTAATACATCCATGGAAGACATTANNNNNNNNNN NNNNNNN[T/C]TCTTTTAAGCCCCACCAAAAATGACAAAAAAATAATGTGTTAGAAAAATATCCCT CTTTAAATGTCCATTCAAATTGGT
[a]Sequences were submitted to KbioScience for primer design. Not all nucleotides included in the sequence were necessarily incorporated			

Table S2 Conditions for real-time PCR

Step	Denaturation	Time	Annealing	Time	Cycles
1	94°C	15 min	--	--	1
2	94°C	20 sec	65°C	60 sec	1
3	94°C	20 sec	64.2°C	60 sec	1
4	94°C	20 sec	63.4°C	60 sec	1
5	94°C	20 sec	62.6°C	60 sec	1
6	94°C	20 sec	61.8°C	60 sec	1
7	94°C	20 sec	61.0°C	60 sec	1
8	94°C	20 sec	60.2°C	60 sec	1
9	94°C	20 sec	59.4°C	60 sec	1
10	94°C	20 sec	58.6°C	60 sec	1
11	94°C	20 sec	57.8°C	60 sec	1
12	94°C	20 sec	57.0°C	60 sec	35

Table S3 Significant SNPs for fruit and seed set in the DRH F_1 and S_3 populations

Gen	Trait	SNP	P-value	Chr	Location	Alleles	Hap block[a]
F_1	Fruit	solcap_snp_c1_14577	1.40E-05	chr12	10635877	T/C	
F_1	Fruit	solcap_snp_c2_40748	1.00E-05	chr12	10816093	T/C	
F_1	Fruit	solcap_snp_c2_40751	1.20E-05	chr12	10993934	T/C	
F_1	Fruit	solcap_snp_c2_52691	3.60E-05	chr12	11435508	A/G	Block 1
F_1	Fruit	solcap_snp_c2_44926	3.60E-05	chr12	11729136	T/C	Block 1
F_1	Fruit	solcap_snp_c2_44928	1.40E-05	chr12	11729319	A/G	Block 1
F_1	Fruit	solcap_snp_c2_44932	1.40E-05	chr12	11729936	T/C	Block 1
F_1	Fruit	solcap_snp_c2_51049	1.00E-05	chr12	12161618	A/C	Block 2
F_1	Fruit	solcap_snp_c2_51047	1.00E-05	chr12	12171861	T/G	Block 2
F_1	Fruit	solcap_snp_c1_403	1.00E-05	chr12	13384725	T/G	
F_1	Fruit	solcap_snp_c1_14759	3.80E-10	chr12	19959652	T/C	
F_1	Fruit	solcap_snp_c2_9486	1.00E-05	chr12	27976352	A/G	
F_1	Fruit	solcap_snp_c2_33630	1.00E-05	chr12	32983664	T/C	
F_1	Fruit	solcap_snp_c2_30296	1.00E-05	chr12	33520836	T/C	
F_1	Fruit	solcap_snp_c2_18992	1.20E-05	chr12	34664381	A/C	
F_1	Fruit	solcap_snp_c2_45812	3.70E-06	chr12	35192739	T/C	Block 3
F_1	Fruit	solcap_snp_c2_45811	1.10E-05	chr12	35192753	A/G	Block 3
F_1	Fruit	solcap_snp_c2_45808	2.70E-06	chr12	35193132	T/G	Block 3
F_1	Fruit	solcap_snp_c2_45807	3.70E-06	chr12	35193185	A/G	Block 3
F_1	Fruit	solcap_snp_c1_4502	1.40E-05	chr12	38164983	A/G	
F_1	Fruit	solcap_snp_c1_10050	1.00E-05	chr12	39424382	T/C	
F_1	Fruit	solcap_snp_c1_8581	1.20E-05	chr12	41004577	A/G	
F_1	Fruit	solcap_snp_c2_10042	1.20E-05	chr12	42729099	T/C	
F_1	Fruit	solcap_snp_c2_19722	1.00E-05	chr12	45238632	T/C	
F_1	Fruit	solcap_snp_c2_48687	2.70E-06	chr12	46449287	A/G	
F_1	Fruit	solcap_snp_c2_48011	2.70E-06	chr12	47222636	A/G	
F_1	Fruit	solcap_snp_c2_43152	6.10E-07	chr12	47592575	T/C	
F_1	Fruit	solcap_snp_c1_14870	4.40E-07	chr12	48494590	T/C	Block 4
F_1	Fruit	solcap_snp_c1_14869	4.40E-07	chr12	48494602	A/G	Block 4
F_1	Fruit	solcap_snp_c2_18822	6.10E-07	chr12	49396411	A/G	Block 5
F_1	Fruit	solcap_snp_c2_18827	5.20E-07	chr12	49651778	A/G	Block 5
F_1	Fruit	solcap_snp_c2_18836	4.40E-07	chr12	49696887	A/G	Block 5
F_1	Fruit	solcap_snp_c2_18838	4.40E-07	chr12	49719870	A/G	Block 5

Table S3 Significant SNPs for fruit and seed set in the DRH F_1 and S_3 populations

F_1	Fruit	solcap_snp_c2_18848	5.20E-07	chr12	49742789	T/C	Block 5
F_1	Fruit	solcap_snp_c2_52568	6.10E-07	chr12	49797780	A/C	Block 5
F_1	Fruit	solcap_snp_c2_52567	6.10E-07	chr12	49797786	T/C	Block 5
F_1	Fruit	solcap_snp_c2_53324	4.40E-07	chr12	50048700	A/G	
F_1	Fruit	solcap_snp_c2_42328	6.10E-07	chr12	50420138	A/C	
F_1	Fruit	solcap_snp_c2_57478	1.60E-06	chr12	50694930	T/C	Block 6
F_1	Fruit	solcap_snp_c2_23337	4.40E-07	chr12	50775845	T/C	Block 6
F_1	Fruit	solcap_snp_c2_23308	6.10E-07	chr12	50928733	A/G	Block 6
F_1	Fruit	solcap_snp_c2_23284	6.10E-07	chr12	51043202	A/G	Block 6
F_1	Fruit	solcap_snp_c2_23252	2.30E-07	chr12	51609585	A/C	
F_1	Fruit	solcap_snp_c2_23253	8.40E-07	chr12	51609743	A/T	Block 7
F_1	Fruit	solcap_snp_c2_23254	2.30E-07	chr12	51609775	A/G	Block 7
F_1	Fruit	solcap_snp_c2_23256	2.30E-07	chr12	51610338	A/C	Block 7
F_1	Fruit	solcap_snp_c2_23258	2.30E-07	chr12	51610655	T/C	Block 7
F_1	Fruit	solcap_snp_c2_23259	2.30E-07	chr12	51610737	T/C	Block 7
F_1	Fruit	solcap_snp_c2_23235	2.30E-07	chr12	51713106	A/C	Block 7
F_1	Fruit	solcap_snp_c1_11644	3.90E-07	chr12	52108708	A/G	Block 7
F_1	Fruit	solcap_snp_c1_11668	3.90E-07	chr12	52169519	A/G	Block 8
F_1	Fruit	solcap_snp_c2_39414	3.90E-07	chr12	52169618	T/G	Block 8
F_1	Fruit	solcap_snp_c2_39410	3.90E-07	chr12	52169906	A/G	Block 8
F_1	Fruit	solcap_snp_c2_39409	3.90E-07	chr12	52169970	A/G	Block 8
F_1	Fruit	solcap_snp_c2_39393	2.70E-07	chr12	52195122	A/G	Block 8
F_1	Fruit	solcap_snp_c2_50821	3.90E-07	chr12	52361617	T/G	Block 8
F_1	Fruit	solcap_snp_c2_50824	3.90E-07	chr12	52414999	T/C	Block 8
F_1	Fruit	solcap_snp_c2_48482	3.90E-07	chr12	52663134	A/G	Block 8
F_1	Fruit	solcap_snp_c2_48483	3.90E-07	chr12	52663197	A/G	Block 8
F_1	Fruit	solcap_snp_c2_48470	2.30E-07	chr12	52796508	A/G	
F_1	Fruit	solcap_snp_c2_57400	2.70E-07	chr12	53026479	T/C	Block 9
F_1	Fruit	solcap_snp_c2_57399	2.70E-07	chr12	53026485	T/C	Block 9
F_1	Fruit	solcap_snp_c2_57398	1.70E-08	chr12	53027354	T/C	Block 9
F_1	Fruit	solcap_snp_c2_32517	8.10E-08	chr12	53120655	A/G	Block 9
F_1	Fruit	solcap_snp_c2_32522	5.50E-08	chr12	53120976	A/G	Block 9
F_1	Fruit	solcap_snp_c2_32466	8.10E-08	chr12	53152137	A/C	Block 9
F_1	Fruit	solcap_snp_c2_32467	8.10E-08	chr12	53155339	T/C	Block 9

Table S3 Significant SNPs for fruit and seed set in the DRH F_1 and S_3 populations

F_1	Fruit	solcap_snp_c2_32482	2.70E-07	chr12	53247648	A/G	Block 10
F_1	Fruit	solcap_snp_c2_32498	8.10E-08	chr12	53346991	T/C	Block 10
F_1	Fruit	solcap_snp_c2_32505	2.70E-07	chr12	53364172	A/G	Block 10
F_1	Fruit	solcap_snp_c2_32082	2.70E-07	chr12	53658860	T/C	Block 10
F_1	Fruit	solcap_snp_c2_32077	1.10E-08	chr12	53697469	T/C	Block 10
F_1	Fruit	solcap_snp_c1_2331	1.10E-10	chr12	54812137	A/G	
F_1	Fruit	solcap_snp_c2_8037	5.30E-13	chr12	55640932	T/C	
F_1	Fruit	solcap_snp_c1_2689	1.80E-12	chr12	55982568	T/C	
F_1	Fruit	solcap_snp_c2_46213	7.10E-12	chr12	57064746	A/G	
F_1	Fruit	solcap_snp_c1_1944	1.30E-11	chr12	57695670	A/G	
F_1	Fruit	solcap_snp_c1_1923	8.40E-12	chr12	57805264	T/C	
F_1	Fruit	solcap_snp_c2_5713	1.20E-10	chr12	57892572	A/T	
F_1	Fruit	solcap_snp_c2_5594	3.10E-10	chr12	58241614	A/G	Block 11
F_1	Fruit	solcap_snp_c1_1985	6.70E-10	chr12	58242547	T/C	Block 11
F_1	Fruit	solcap_snp_c2_5507	7.90E-10	chr12	58413157	T/C	Block 11
F_1	Fruit	solcap_snp_c2_5474	1.30E-10	chr12	58642405	A/G	Block 11
F_1	Fruit	solcap_snp_c2_5463	1.30E-10	chr12	58688510	T/C	Block 11
F_1	Fruit	solcap_snp_c2_5461	1.30E-10	chr12	58689924	A/G	Block 11
F_1	Fruit	solcap_snp_c2_5446	7.90E-10	chr12	58744648	T/C	
F_1	Fruit	solcap_snp_c2_5524	6.60E-10	chr12	59026489	T/C	
F_1	Fruit	solcap_snp_c1_2009	6.60E-10	chr12	59060025	T/C	
F_1	Seed	solcap_snp_c2_54579	0.0001	chr02	1498967	T/G	
F_1	Seed	solcap_snp_c2_4505	0.0001	chr02	2596302	T/C	
F_1	Seed	solcap_snp_c2_20445	0.0001	chr02	6200965	T/C	
F_1	Seed	solcap_snp_c2_30946	0.0001	chr02	6640697	T/C	Block 1
F_1	Seed	solcap_snp_c2_30950	0.0001	chr02	6811226	A/C	Block 1
F_1	Seed	solcap_snp_c2_41904	0.0001	chr02	6874149	T/C	Block 1
F_1	Seed	solcap_snp_c2_41906	0.0001	chr02	6874296	A/C	Block 1
F_1	Seed	solcap_snp_c2_41874	0.0001	chr02	7051134	T/G	Block 1
F_1	Seed	solcap_snp_c2_16362	0.0001	chr02	7412620	A/G	
F_1	Seed	solcap_snp_c2_54473	0.0001	chr02	7509533	A/G	
F_1	Seed	solcap_snp_c2_16341	0.0001	chr02	7955440	A/C	Block 2
F_1	Seed	solcap_snp_c2_16329	0.0001	chr02	8063538	A/C	Block 2
F_1	Seed	solcap_snp_c2_16348	0.0001	chr02	8213219	A/C	Block 2

Table S3 Significant SNPs for fruit and seed set in the DRH F_1 and S_3 populations

F_1	Seed	solcap_snp_c2_16347	0.0001	chr02	8213251	T/C	Block 2
F_1	Seed	solcap_snp_c1_3753	0.0001	chr02	9316494	T/C	
F_1	Seed	solcap_snp_c2_11584	0.0001	chr02	10139129	A/G	
F_1	Seed	solcap_snp_c2_48734	0.0001	chr02	10655356	T/C	
F_1	Seed	solcap_snp_c1_15972	0.0001	chr02	15346371	A/G	Block 3
F_1	Seed	solcap_snp_c1_15973	0.0001	chr02	15346596	T/C	Block 3
F_1	Seed	solcap_snp_c1_15974	0.0001	chr02	15346644	A/C	Block 3
F_1	Seed	solcap_snp_c1_15975	0.0001	chr02	15348125	A/G	Block 3
F_1	Seed	solcap_snp_c2_41124	0.0001	chr02	16013506	A/G	
F_1	Seed	solcap_snp_c2_49068	0.0001	chr02	16468426	T/C	
F_1	Seed	solcap_snp_c1_868	0.0001	chr02	16611670	T/C	
F_1	Seed	solcap_snp_c2_30937	0.0001	chr02	17480896	A/G	Block 4
F_1	Seed	solcap_snp_c2_30940	0.0001	chr02	17497079	A/G	Block 4
F_1	Seed	solcap_snp_c2_17387	0.0001	chr02	19385694	A/C	
F_1	Seed	solcap_snp_c2_17428	0.0001	chr02	19968887	A/T	Block 5
F_1	Seed	solcap_snp_c1_5740	0.0001	chr02	19969225	T/G	Block 5
F_1	Seed	solcap_snp_c1_11344	0.0001	chr02	21013184	A/G	
F_1	Seed	solcap_snp_c1_3495	0.0001	chr04	63997821	T/C	Block 1
F_1	Seed	solcap_snp_c1_3494	0.0001	chr04	63999627	T/C	Block 1
F_1	Seed	solcap_snp_c1_3484	0.0001	chr04	64076369	T/C	Block 1
F_1	Seed	solcap_snp_c2_10568	0.0001	chr04	64085790	T/A	Block 1
F_1	Seed	solcap_snp_c2_10563	0.0001	chr04	64087612	C/G	Block 1
F_1	Seed	solcap_snp_c2_10546	0.0001	chr04	64154514	A/G	Block 1
F_1	Seed	solcap_snp_c1_3462	0.0001	chr04	64158130	T/A	Block 1
F_1	Seed	solcap_snp_c1_3461	0.0001	chr04	64158147	A/G	Block 1
F_1	Seed	solcap_snp_c2_46213	0.0001	chr12	57064746	A/G	
F_1	Seed	solcap_snp_c2_5713	0.0001	chr12	57892572	A/T	
F_1	Seed	solcap_snp_c2_5594	0.0001	chr12	58241614	A/G	
F_1	Seed	solcap_snp_c2_5524	0.0001	chr12	59026489	T/C	
F_1	Seed	solcap_snp_c1_2009	0.0001	chr12	59060025	T/C	
S_3	Fruit	solcap_snp_c1_15123	0.000486	chr01	75334524	A/G	
S_3	Fruit	solcap_snp_c2_39671	0.000872	chr02	22578633	T/C	
S_3	Fruit	solcap_snp_c2_31403	0.00038	chr04	7643764	A/G	
S_3	Fruit	solcap_snp_c1_16258	8.82E-05	chr04	7820934	T/C	

Table S3 Significant SNPs for fruit and seed set in the DRH F_1 and S_3 populations

S_3	Fruit	solcap_snp_c2_26771	0.000496	chr04	8935610	T/C	
S_3	Fruit	solcap_snp_c2_54083	0.00038	chr04	13877787	A/G	
S_3	Fruit	solcap_snp_c2_29169	0.000957	chr06	49961326	A/G	
S_3	Fruit	solcap_snp_c2_29204	0.000957	chr06	50334461	T/C	
S_3	Fruit	solcap_snp_c1_8674	0.000957	chr06	51778165	T/C	
S_3	Fruit	solcap_snp_c2_28107	0.000957	chr06	51996842	T/C	
S_3	Fruit	solcap_snp_c1_8679	0.000957	chr06	52230856	A/G	
S_3	Fruit	solcap_snp_c1_16127	0.000957	chr06	52314099	A/G	
S_3	Fruit	solcap_snp_c2_12696	0.000503	chr07	48571448	A/G	
S_3	Fruit	solcap_snp_c2_34655	0.000951	chr08	39937684	A/C	Block 1
S_3	Fruit	solcap_snp_c2_34647	0.000951	chr08	39940413	A/C	Block 1
S_3	Fruit	solcap_snp_c2_34640	0.000633	chr08	39972767	T/A	Block 1
S_3	Fruit	solcap_snp_c2_34639	0.000633	chr08	39974638	T/C	Block 1
S_3	Fruit	solcap_snp_c2_34636	0.000633	chr08	39975154	A/G	Block 1
S_3	Fruit	solcap_snp_c2_34635	0.000633	chr08	39975211	A/C	Block 1
S_3	Fruit	solcap_snp_c2_34632	0.000633	chr08	39975838	T/C	Block 1
S_3	Fruit	solcap_snp_c2_34608	0.000633	chr08	40046156	T/G	Block 1
S_3	Fruit	solcap_snp_c2_26654	0.000629	chr08	41057522	A/G	
S_3	Fruit	solcap_snp_c2_16135	0.000282	chr08	41983858	T/C	
S_3	Fruit	solcap_snp_c2_3962	6.48E-05	chr09	4709804	A/G	
S_3	Fruit	solcap_snp_c2_13345	2.87E-05	chr11	2030648	T/C	
S_3	Fruit	solcap_snp_c2_13419	2.87E-05	chr11	2355192	A/G	
S_3	Fruit	solcap_snp_c1_4322	2.87E-05	chr11	2533707	T/C	
S_3	Fruit	solcap_snp_c2_33657	1.18E-05	chr11	3781652	T/C	
S_3	Fruit	solcap_snp_c2_24454	0.000823	chr12	459353	T/C	
S_3	Fruit	solcap_snp_c1_15641	8.03E-05	chr12	823935	A/G	Block 1
S_3	Fruit	solcap_snp_c1_15642	8.03E-05	chr12	823953	A/G	Block 1
S_3	Seed	solcap_snp_c2_54815	0.0001	chr01	629174	T/C	
S_3	Seed	solcap_snp_c2_6793	0.0001	chr01	817933	A/G	
S_3	Seed	solcap_snp_c2_6684	0.0001	chr01	1332010	T/G	
S_3	Seed	solcap_snp_c1_6668	0.0001	chr01	2257441	T/C	
S_3	Seed	solcap_snp_c2_21234	0.0001	chr01	2459829	A/G	
S_3	Seed	solcap_snp_c2_21247	0.0001	chr01	2591084	T/C	
S_3	Seed	solcap_snp_c2_27677	0.0001	chr01	18011456	A/T	

Table S3 Significant SNPs for fruit and seed set in the DRH F_1 and S_3 populations

S_3	Seed	solcap_snp_c1_14249	0.0001	chr01	27299013	A/G	
S_3	Seed	solcap_snp_c2_57284	0.0001	chr01	35120841	A/G	
S_3	Seed	solcap_snp_c2_52477	0.0001	chr01	38433584	T/G	
S_3	Seed	solcap_snp_c2_40094	0.0001	chr01	38810424	A/G	
S_3	Seed	solcap_snp_c1_6294	0.0001	chr01	66330376	A/G	
S_3	Seed	solcap_snp_c2_14274	0.0001	chr01	67051585	T/G	
S_3	Seed	solcap_snp_c2_5226	0.0001	chr03	8899407	A/G	
S_3	Seed	solcap_snp_c1_9141	0.0001	chr03	18971851	A/G	
S_3	Seed	solcap_snp_c2_45700	0.0001	chr03	19749310	G/C	
S_3	Seed	solcap_snp_c2_13811	0.0001	chr03	19996111	A/C	
S_3	Seed	solcap_snp_c2_13821	0.0001	chr03	20054304	A/G	
S_3	Seed	solcap_snp_c1_6905	0.0001	chr04	47876627	T/C	
S_3	Seed	solcap_snp_c1_10751	0.0001	chr04	48241862	T/A	
S_3	Seed	solcap_snp_c2_36053	0.0001	chr04	48348385	T/G	
S_3	Seed	solcap_snp_c2_36060	0.0001	chr04	48495445	T/C	
S_3	Seed	solcap_snp_c2_39804	0.0001	chr04	49015604	T/C	
S_3	Seed	solcap_snp_c2_52890	0.0001	chr04	50054983	T/C	Block 1
S_3	Seed	solcap_snp_c2_57884	0.0001	chr04	50097215	A/G	Block 1
S_3	Seed	solcap_snp_c2_57883	0.0001	chr04	50097231	A/T	Block 1
S_3	Seed	solcap_snp_c2_54335	0.0001	chr04	50310866	T/C	Block 1
S_3	Seed	solcap_snp_c1_15495	0.0001	chr04	52129179	T/C	
S_3	Seed	solcap_snp_c2_48808	0.0001	chr04	52238283	T/C	
S_3	Seed	solcap_snp_c1_15505	0.0001	chr04	52773285	A/C	
S_3	Seed	solcap_snp_c2_20668	0.0001	chr05	44003690	A/G	
S_3	Seed	solcap_snp_c2_34655	0.0001	chr08	39937684	A/C	Block 1
S_3	Seed	solcap_snp_c2_34647	0.0001	chr08	39940413	A/C	Block 1
S_3	Seed	solcap_snp_c2_34640	0.0001	chr08	39972767	T/A	Block 1
S_3	Seed	solcap_snp_c2_34639	0.0001	chr08	39974638	T/C	Block 1
S_3	Seed	solcap_snp_c2_34636	0.0001	chr08	39975154	A/G	Block 1
S_3	Seed	solcap_snp_c2_34635	0.0001	chr08	39975211	A/C	Block 1
S_3	Seed	solcap_snp_c2_34632	0.0001	chr08	39975838	T/C	Block 1
S_3	Seed	solcap_snp_c2_34608	0.0001	chr08	40046156	T/G	Block 1
S_3	Seed	solcap_snp_c1_10384	0.0001	chr08	40182187	A/G	Block 1
S_3	Seed	solcap_snp_c2_34566	0.0001	chr08	40215099	A/G	Block 1

Table S3 Significant SNPs for fruit and seed set in the DRH F_1 and S_3 populations

S_3	Seed	solcap_snp_c1_8297	0.0001	chr08	40594533	T/C	
S_3	Seed	solcap_snp_c1_8293	0.0001	chr08	40648211	T/G	Block 2
S_3	Seed	solcap_snp_c1_8282	0.0001	chr08	40793252	A/G	Block 2
S_3	Seed	solcap_snp_c2_26654	0.0001	chr08	41057522	A/G	Block 2
S_3	Seed	solcap_snp_c1_8300	0.0001	chr08	41066285	A/G	Block 2
S_3	Seed	solcap_snp_c1_8237	0.0001	chr08	41128926	A/G	Block 2
S_3	Seed	solcap_snp_c1_16495	0.0001	chr08	41225646	T/C	Block 3
S_3	Seed	solcap_snp_c1_8759	0.0001	chr08	41395216	T/G	Block 3
S_3	Seed	solcap_snp_c1_8760	0.0001	chr08	41395254	C/G	Block 3
S_3	Seed	solcap_snp_c1_8763	0.0001	chr08	41427416	T/C	Block 3
S_3	Seed	solcap_snp_c1_8765	0.0001	chr08	41427535	A/G	Block 3
S_3	Seed	solcap_snp_c2_16135	0.0001	chr08	41983858	T/C	
S_3	Seed	solcap_snp_c1_5587	0.0001	chr08	42992615	T/C	Block 4
S_3	Seed	solcap_snp_c1_5567	0.0001	chr08	43042698	A/G	Block 4
S_3	Seed	solcap_snp_c1_5566	0.0001	chr08	43042707	A/G	Block 4
S_3	Seed	solcap_snp_c1_5560	0.0001	chr08	43065776	T/C	Block 4
S_3	Seed	solcap_snp_c1_5559	0.0001	chr08	43065782	T/C	Block 4
S_3	Seed	solcap_snp_c2_17060	0.0001	chr08	43066082	T/G	Block 4
S_3	Seed	solcap_snp_c1_5529	0.0001	chr08	43265949	A/G	Block 4
S_3	Seed	solcap_snp_c2_16997	0.0001	chr08	43342857	T/C	Block 4
S_3	Seed	solcap_snp_c2_16996	0.0001	chr08	43343731	A/C	Block 4
S_3	Seed	solcap_snp_c1_5499	0.0001	chr08	43349116	A/G	Block 4
S_3	Seed	solcap_snp_c2_49642	0.0001	chr09	9605858	A/G	
S_3	Seed	solcap_snp_c1_15780	0.0001	chr09	13029151	T/C	
S_3	Seed	solcap_snp_c2_57401	0.0001	chr09	13990162	T/C	
S_3	Seed	solcap_snp_c2_26517	0.0001	chr09	14210077	A/C	
S_3	Seed	solcap_snp_c2_58234	0.0001	chr09	28291127	A/C	
S_3	Seed	solcap_snp_c1_6183	0.0001	chr09	40707636	A/C	
S_3	Seed	solcap_snp_c2_43012	0.0001	chr09	43758241	T/C	
S_3	Seed	solcap_snp_c2_42964	0.0001	chr09	43896153	A/C	
S_3	Seed	solcap_snp_c2_40848	0.0001	chr09	44267204	A/G	
S_3	Seed	solcap_snp_c2_40879	0.0001	chr09	44388893	A/C	
S_3	Seed	solcap_snp_c2_27003	0.0001	chr09	44883913	T/C	
S_3	Seed	solcap_snp_c1_6936	0.0001	chr09	45768611	T/C	

Table S3 Significant SNPs for fruit and seed set in the DRH F_1 and S_3 populations

S_3	Seed	solcap_snp_c2_54325	0.0001	chr09	45821395	A/G	
S_3	Seed	solcap_snp_c1_12178	0.0001	chr09	45992948	T/G	
S_3	Seed	solcap_snp_c2_22067	0.0001	chr09	46414360	A/G	
S_3	Seed	solcap_snp_c2_46778	0.0001	chr09	47031950	T/G	Block 1
S_3	Seed	solcap_snp_c2_46777	0.0001	chr09	47035169	A/C	Block 1
S_3	Seed	solcap_snp_c2_46776	0.0001	chr09	47045957	A/G	Block 1
S_3	Seed	solcap_snp_c1_13886	0.0001	chr09	47120991	T/C	Block 1
S_3	Seed	solcap_snp_c2_30008	0.0001	chr09	48351988	A/C	
S_3	Seed	solcap_snp_c2_29981	0.0001	chr09	48742579	A/G	Block 2
S_3	Seed	solcap_snp_c2_20640	0.0001	chr09	48863776	T/C	Block 2
S_3	Seed	solcap_snp_c2_40032	0.0001	chr09	48981299	T/C	Block 2
S_3	Seed	solcap_snp_c1_11866	0.0001	chr09	49151351	T/C	
S_3	Seed	solcap_snp_c2_40085	0.0001	chr09	49152390	T/C	Block 2
S_3	Seed	solcap_snp_c2_40079	0.0001	chr09	49247431	T/C	Block 3
S_3	Seed	solcap_snp_c2_40076	0.0001	chr09	49247722	A/T	
S_3	Seed	solcap_snp_c2_40075	0.0001	chr09	49248049	C/G	Block 3
S_3	Seed	solcap_snp_c2_43243	0.0001	chr09	49509127	A/C	Block 4
S_3	Seed	solcap_snp_c2_3079	0.0001	chr09	49890367	A/G	Block 4
S_3	Seed	solcap_snp_c1_914	0.0001	chr09	49922744	A/G	Block 4
S_3	Seed	solcap_snp_c2_3068	0.0001	chr09	50008535	A/G	Block 4
S_3	Seed	solcap_snp_c2_2998	0.0001	chr09	50339269	A/G	
S_3	Seed	solcap_snp_c2_2992	0.0001	chr09	50535441	A/G	
S_3	Seed	solcap_snp_c2_47939	0.0001	chr09	50921796	A/C	Block 5
S_3	Seed	solcap_snp_c2_47952	0.0001	chr09	50960050	A/G	Block 5
S_3	Seed	solcap_snp_c2_55484	0.0001	chr09	51037194	A/C	Block 5
S_3	Seed	solcap_snp_c1_16106	0.0001	chr09	51044639	A/C	Block 5
S_3	Seed	solcap_snp_c1_14205	0.0001	chr09	51294827	A/G	Block 5
S_3	Seed	solcap_snp_c2_35612	0.0001	chr09	52066143	A/G	
S_3	Seed	solcap_snp_c2_14940	0.0001	chr11	35734796	T/C	
S_3	Seed	solcap_snp_c2_32498	0.0001	chr12	53346991	T/C	

[a] Hap blocks represent the haplotype blocks calculated per chromosome modeling the likelihood of linkage among SNPs

Table S4 SNPs with preferential inheritance of heterozygous RH alleles in the F_1 progeny

SNP	Chr	Location	DM	RH	Favored allele[a]
solcap_snp_c1_3860	chr01	63999651	TT	TC	C
solcap_snp_c1_15425	chr01	69132119	GG	AG	A
solcap_snp_c1_6465	chr02	4644953	TT	TC	C
solcap_snp_c2_17648	chr03	39045532	GG	AG	A
solcap_snp_c2_17650	chr03	39045594	CC	AC	A
solcap_snp_c2_46653	chr04	1046357	AA	AG	G
solcap_snp_c2_44620	chr04	8073332	TT	TC	C
solcap_snp_c2_44619	chr04	8073341	TT	TC	C
solcap_snp_c2_11732	chr05	2078142	AA	AG	G
solcap_snp_c2_37052	chr05	7709869	TT	TC	C
solcap_snp_c2_43519	chr05	8723035	TT	TC	C
solcap_snp_c2_43516	chr05	8730423	AA	AG	G
solcap_snp_c2_58105	chr05	41429346	GG	TG	T
solcap_snp_c2_5857	chr06	45586184	CC	AC	A
solcap_snp_c2_24066	chr06	49232604	AA	AC	C
solcap_snp_c1_15061	chr06	54377021	GG	AG	A
solcap_snp_c2_34178	chr08	3731043	AA	AG	G
solcap_snp_c2_28608	chr08	36438900	AA	AG	G
solcap_snp_c2_10962	chr09	6400422	TT	TC	C
solcap_snp_c2_38520	chr09	10047137	AA	AC	C
solcap_snp_c2_31391	chr09	24704883	TT	TC	C
solcap_snp_c2_4410	chr09	28715755	GG	AG	A
solcap_snp_c2_40076	chr09	49247722	AA	AT	T
solcap_snp_c2_919	chr10	1561206	CC	AC	A
solcap_snp_c2_58128	chr10	4657967	TT	TC	C
solcap_snp_c2_40830	chr10	40655396	CC	AC	A
solcap_snp_c2_41766	chr10	43981129	CC	AC	A
solcap_snp_c2_12267	chr11	8390679	TT	TG	G
solcap_snp_c2_12266	chr11	8390712	CC	AC	A
solcap_snp_c2_32975	chr11	9173275	AA	AG	G
solcap_snp_c2_3681	chr11	37091701	TT	TG	G
solcap_snp_c2_52554	chr12	49891318	AA	AG	G
solcap_snp_c1_8974	no hit		AA	AG	G
solcap_snp_c2_37343	no hit		TT	TC	C
solcap_snp_c2_40077	no hit		GG	AG	A
solcap_snp_c2_58273	no hit		GG	AG	A

[a] Loci with exclusive transmission to the F_1 progeny of the alternative RH allele not shared by DM in homozygous by heterozygous (AA x AB) cross

Table S5 S_3 generation SNPs with high retained heterozygosity

SNP	Chr	Location (bp)	F_1 heterozygosity	S_3 heterozygosity	Percent retained heterozygosity
solcap_snp_c1_6083	1	3456625	0.62	0.32	0.51
solcap_snp_c2_51812	1	3761637	0.62	0.51	0.83
solcap_snp_c2_51811	1	3761730	0.62	0.51	0.83
solcap_snp_c2_51810	1	3761845	0.62	0.51	0.83
solcap_snp_c2_55009	1	4361129	1.00	0.52	0.52
solcap_snp_c2_45058	1	4714970	1.00	0.53	0.53
solcap_snp_c2_45071	1	4754523	1.00	0.54	0.54
solcap_snp_c1_13318	1	4798733	1.00	0.60	0.60
solcap_snp_c2_27887	1	5271985	1.00	0.58	0.58
solcap_snp_c2_27683	1	18007512	0.60	0.37	0.61
solcap_snp_c2_27682	1	18007594	0.60	0.37	0.61
solcap_snp_c1_16425	1	30338879	0.60	0.36	0.59
solcap_snp_c2_50013	1	33573498	1.00	0.57	0.57
solcap_snp_c2_50011	1	33574077	0.60	0.55	0.92
solcap_snp_c2_54811	1	49795853	0.60	0.36	0.59
solcap_snp_c2_32371	1	55254524	0.42	0.29	0.69
solcap_snp_c2_42	1	78376797	0.44	0.24	0.54
solcap_snp_c1_12320	2	26097416	0.47	0.27	0.58
solcap_snp_c1_12345	2	26134464	0.47	0.27	0.58
solcap_snp_c1_12354	2	26142190	0.47	0.29	0.62
solcap_snp_c1_12329	2	26340920	0.57	0.29	0.51
solcap_snp_c2_51113	2	28299164	0.55	0.28	0.51
solcap_snp_c2_51115	2	28299386	0.55	0.28	0.51
solcap_snp_c2_534	2	28495182	0.55	0.29	0.53
solcap_snp_c1_13912	2	28517709	0.55	0.29	0.53
solcap_snp_c1_13923	2	28522091	0.55	0.29	0.53
solcap_snp_c2_46904	2	28556953	0.55	0.30	0.55
solcap_snp_c2_46908	2	28563862	0.55	0.30	0.55
solcap_snp_c2_46909	2	28564159	0.55	0.30	0.55
solcap_snp_c1_13910	2	28625062	0.55	0.30	0.55
solcap_snp_c1_13911	2	28656448	0.55	0.30	0.55
solcap_snp_c2_46885	2	28678732	0.55	0.30	0.55
solcap_snp_c2_46887	2	28678872	0.55	0.30	0.55
solcap_snp_c2_46898	2	28781437	0.56	0.30	0.54
solcap_snp_c1_7469	2	28998397	0.56	0.30	0.54
solcap_snp_c1_7430	2	29114760	0.56	0.30	0.54
solcap_snp_c2_23170	2	29419753	0.56	0.29	0.52
solcap_snp_c2_23190	2	29662096	0.55	0.29	0.53

Table S5 S_3 generation SNPs with high retained heterozygosity

solcap_snp_c2_23192	2	29694231	0.55	0.29	0.53
solcap_snp_c1_12169	2	30155024	0.55	0.29	0.53
solcap_snp_c2_41534	2	30205384	0.55	0.29	0.53
solcap_snp_c2_41541	2	30225061	0.55	0.29	0.53
solcap_snp_c1_13233	2	30348495	0.55	0.29	0.53
solcap_snp_c2_44982	2	30519243	0.55	0.29	0.53
solcap_snp_c2_44777	2	31304246	0.54	0.34	0.63
solcap_snp_c2_44778	2	31304543	0.54	0.34	0.63
solcap_snp_c2_44768	2	31518751	0.54	0.34	0.63
solcap_snp_c2_51990	2	31649320	0.55	0.34	0.61
solcap_snp_c2_51986	2	31654622	0.55	0.34	0.61
solcap_snp_c1_15178	2	31730861	0.55	0.34	0.61
solcap_snp_c2_13035	2	31900701	0.55	0.34	0.61
solcap_snp_c2_48784	2	32119855	0.55	0.34	0.61
solcap_snp_c2_47037	2	32430568	0.55	0.36	0.65
solcap_snp_c1_12264	2	32623491	0.55	0.36	0.65
solcap_snp_c1_12257	2	32663234	1.00	0.53	0.53
solcap_snp_c2_33141	2	33046224	0.54	0.37	0.69
solcap_snp_c2_40155	2	33563831	1.00	0.57	0.57
solcap_snp_c2_40169	2	33637834	0.57	0.38	0.66
solcap_snp_c2_40170	2	33637915	0.57	0.38	0.66
solcap_snp_c2_40172	2	33637981	0.56	0.38	0.68
solcap_snp_c2_50391	2	34030793	0.56	0.39	0.70
solcap_snp_c2_17973	2	34681416	1.00	0.54	0.54
solcap_snp_c2_17954	2	34780896	0.56	0.38	0.68
solcap_snp_c2_17858	2	35050799	0.57	0.38	0.67
solcap_snp_c1_5881	2	35320327	0.56	0.38	0.68
solcap_snp_c1_5845	2	35410920	0.54	0.38	0.70
solcap_snp_c2_49495	2	35469514	0.56	0.38	0.68
solcap_snp_c1_12377	2	35793401	1.00	0.53	0.53
solcap_snp_c2_42244	2	35821300	1.00	0.53	0.53
solcap_snp_c2_42169	2	35890077	0.57	0.39	0.69
solcap_snp_c2_42172	2	35890669	1.00	0.54	0.54
solcap_snp_c1_16170	2	35926249	0.57	0.39	0.69
solcap_snp_c1_12373	2	35982087	1.00	0.54	0.54
solcap_snp_c1_12375	2	35982537	0.56	0.39	0.70
solcap_snp_c1_12382	2	36106124	0.56	0.39	0.70
solcap_snp_c2_42241	2	36127738	0.56	0.39	0.70
solcap_snp_c1_10492	2	36342244	0.56	0.40	0.72
solcap_snp_c1_10491	2	36355040	0.56	0.40	0.72
solcap_snp_c2_35165	2	36434429	0.56	0.40	0.72

Table S5 S_3 generation SNPs with high retained heterozygosity

solcap_snp_c2_35139	2	36498372	0.56	0.44	0.78
solcap_snp_c2_35113	2	36572842	0.56	0.40	0.72
solcap_snp_c2_35147	2	36681729	0.57	0.44	0.77
solcap_snp_c1_15466	2	36837937	0.56	0.40	0.72
solcap_snp_c2_53037	2	37139799	0.58	0.41	0.71
solcap_snp_c2_53036	2	37139882	0.58	0.41	0.71
solcap_snp_c2_53034	2	37140041	0.58	0.41	0.71
solcap_snp_c2_53033	2	37140155	0.58	0.41	0.71
solcap_snp_c1_11955	2	37289197	0.58	0.41	0.71
solcap_snp_c1_11972	2	37514094	0.59	0.41	0.70
solcap_snp_c2_40635	2	37525091	0.59	0.41	0.70
solcap_snp_c2_40636	2	37525154	0.59	0.47	0.80
solcap_snp_c2_40637	2	37525191	0.59	0.48	0.81
solcap_snp_c2_42126	2	37820798	0.60	0.41	0.69
solcap_snp_c2_42127	2	37820900	0.60	0.41	0.69
solcap_snp_c2_42128	2	37821044	0.61	0.41	0.68
solcap_snp_c2_42130	2	37821806	0.61	0.41	0.68
solcap_snp_c2_25766	2	38044925	0.60	0.42	0.71
solcap_snp_c1_8091	2	38169610	0.63	0.43	0.69
solcap_snp_c1_8118	2	38392342	0.63	0.42	0.68
solcap_snp_c2_25897	2	38584192	1.00	0.59	0.59
solcap_snp_c2_25179	2	38746120	0.62	0.43	0.70
solcap_snp_c1_7964	2	38746669	0.62	0.43	0.70
solcap_snp_c2_7549	2	39799732	0.63	0.42	0.68
solcap_snp_c2_7565	2	39819907	0.62	0.42	0.69
solcap_snp_c1_7268	2	41735435	0.37	0.22	0.59
solcap_snp_c2_27268	2	42909848	0.36	0.18	0.51
solcap_snp_c2_27271	2	42910402	0.36	0.18	0.51
solcap_snp_c1_4847	2	44068435	0.65	0.38	0.59
solcap_snp_c2_15018	2	44070055	0.66	0.38	0.58
solcap_snp_c2_15021	2	44070475	0.65	0.38	0.59
solcap_snp_c1_4849	2	44070716	0.67	0.38	0.57
solcap_snp_c2_15040	2	44101397	0.67	0.41	0.61
solcap_snp_c2_15041	2	44104103	0.66	0.38	0.58
solcap_snp_c2_15043	2	44104241	0.63	0.38	0.60
solcap_snp_c2_15046	2	44104569	0.66	0.38	0.58
solcap_snp_c2_15048	2	44104716	0.67	0.38	0.57
solcap_snp_c1_4859	2	44147075	0.65	0.38	0.59
solcap_snp_c1_4862	2	44147324	0.67	0.38	0.57
solcap_snp_c1_4873	2	44196130	0.66	0.38	0.58
solcap_snp_c1_4885	2	44239696	0.67	0.38	0.57

Table S5 S_3 generation SNPs with high retained heterozygosity

solcap_snp_c1_7873	2	44646848	0.65	0.38	0.59
solcap_snp_c1_7872	2	44646861	0.64	0.41	0.64
solcap_snp_c1_7871	2	44646980	0.66	0.38	0.57
solcap_snp_c1_7867	2	44647995	0.64	0.38	0.59
solcap_snp_c1_7848	2	44743260	0.65	0.38	0.59
solcap_snp_c2_47197	2	45589059	0.64	0.39	0.61
solcap_snp_c2_47202	2	45594389	0.64	0.39	0.61
solcap_snp_c1_10607	2	45634393	0.64	0.39	0.61
solcap_snp_c2_35686	2	45775706	0.65	0.39	0.60
solcap_snp_c2_38555	2	46079361	0.64	0.39	0.61
solcap_snp_c1_5895	2	47016567	0.64	0.39	0.61
solcap_snp_c1_6349	3	16405332	0.55	0.35	0.63
solcap_snp_c2_55471	3	16568330	0.55	0.35	0.63
solcap_snp_c1_6352	3	17931555	0.54	0.35	0.64
solcap_snp_c1_9153	3	18598841	0.54	0.28	0.52
solcap_snp_c2_46603	3	19920935	0.57	0.33	0.58
solcap_snp_c1_4444	3	20025586	0.58	0.34	0.58
solcap_snp_c1_9025	3	22771416	0.57	0.35	0.61
solcap_snp_c1_9014	3	22921047	0.56	0.35	0.63
solcap_snp_c2_29645	3	22923600	0.56	0.35	0.62
solcap_snp_c2_51300	4	19152487	0.49	0.29	0.60
solcap_snp_c2_31359	4	26486070	0.52	0.27	0.52
solcap_snp_c2_31361	4	26487839	0.52	0.27	0.53
solcap_snp_c2_1505	4	46319893	0.52	0.28	0.54
solcap_snp_c2_43735	4	56074368	0.42	0.27	0.65
solcap_snp_c2_43748	4	56106621	0.43	0.27	0.64
solcap_snp_c2_39333	4	56253681	1.00	0.54	0.54
solcap_snp_c2_25284	4	56474686	1.00	0.55	0.55
solcap_snp_c1_7990	4	56480035	1.00	0.55	0.55
solcap_snp_c2_26688	4	56579593	0.57	0.29	0.52
solcap_snp_c2_26758	4	56746571	1.00	0.54	0.54
solcap_snp_c1_8328	4	56768932	0.58	0.29	0.51
solcap_snp_c1_6758	4	57479018	0.44	0.25	0.57
solcap_snp_c2_21572	4	57564878	0.46	0.25	0.55
solcap_snp_c2_21573	4	57564882	0.44	0.25	0.57
solcap_snp_c2_21574	4	57565798	1.00	0.52	0.52
solcap_snp_c2_21577	4	57570858	0.46	0.25	0.55
solcap_snp_c2_21579	4	57570962	0.44	0.25	0.56
solcap_snp_c2_55796	4	58101142	1.00	0.55	0.55
solcap_snp_c2_55793	4	58101357	1.00	0.55	0.55
solcap_snp_c2_55791	4	58101543	1.00	0.55	0.55

Table S5 S_3 generation SNPs with high retained heterozygosity

solcap_snp_c2_55785	4	58102214	1.00	0.55	0.55
solcap_snp_c2_55784	4	58102247	1.00	0.55	0.55
solcap_snp_c2_55783	4	58102411	1.00	0.55	0.55
solcap_snp_c2_55782	4	58102420	1.00	0.55	0.55
solcap_snp_c2_55777	4	58102702	1.00	0.55	0.55
solcap_snp_c2_55775	4	58103660	1.00	0.55	0.55
solcap_snp_c2_55772	4	58103847	1.00	0.55	0.55
solcap_snp_c2_55768	4	58104035	1.00	0.55	0.55
solcap_snp_c2_12917	4	61515081	0.57	0.52	0.91
solcap_snp_c2_43504	5	8733830	0.46	0.30	0.64
solcap_snp_c1_16550	6	30072209	0.83	0.65	0.78
solcap_snp_c2_56059	6	33744644	0.44	0.24	0.54
solcap_snp_c2_51758	6	33823946	0.43	0.24	0.55
solcap_snp_c2_51764	6	33824505	0.43	0.24	0.55
solcap_snp_c2_51769	6	33824911	0.44	0.24	0.54
solcap_snp_c2_5869	6	43631557	0.46	0.28	0.62
solcap_snp_c2_5858	6	45518761	0.44	0.27	0.61
solcap_snp_c2_5857	6	45586184	1.00	0.92	0.92
solcap_snp_c2_5864	6	45922705	0.44	0.27	0.62
solcap_snp_c2_24064	6	49230071	0.44	0.23	0.51
solcap_snp_c2_24082	6	49272759	0.45	0.24	0.53
solcap_snp_c1_15061	6	54377021	1.00	0.75	0.75
solcap_snp_c2_28854	7	52270034	0.40	0.22	0.54
solcap_snp_c1_10391	8	40282338	0.57	0.38	0.66
solcap_snp_c1_10390	8	40282365	0.57	0.38	0.66
solcap_snp_c2_34604	8	40293938	0.57	0.38	0.66
solcap_snp_c2_10961	9	6400460	0.35	0.24	0.69
solcap_snp_c2_38520	9	10047137	1.00	0.79	0.79
solcap_snp_c2_22761	9	16309099	0.66	0.37	0.56
solcap_snp_c2_32784	10	5149659	0.48	0.29	0.60
solcap_snp_c2_40830	10	40655396	1.00	0.88	0.88
solcap_snp_c2_28666	10	43785376	0.51	0.28	0.55
solcap_snp_c2_28665	10	43785987	0.51	0.28	0.55
solcap_snp_c2_28761	10	44280631	0.50	0.30	0.61
solcap_snp_c1_7189	10	45447139	0.45	0.30	0.68
solcap_snp_c1_7165	10	45610275	0.47	0.30	0.64
solcap_snp_c2_15528	10	48096702	0.51	0.37	0.73
solcap_snp_c2_48126	10	48777816	0.48	0.30	0.63
solcap_snp_c2_29746	10	49130637	0.47	0.29	0.61
solcap_snp_c1_12644	10	49532516	0.46	0.32	0.68
solcap_snp_c1_12619	10	49563012	0.46	0.32	0.70

Table S5 S_3 generation SNPs with high retained heterozygosity

solcap_snp_c2_44154	10	49762143	0.46	0.33	0.71
solcap_snp_c2_23944	11	62230	0.44	0.27	0.61
solcap_snp_c2_35856	11	1096666	0.44	0.30	0.69
solcap_snp_c2_57429	11	5022374	0.42	0.27	0.64
solcap_snp_c2_23915	11	5939690	0.42	0.29	0.70
solcap_snp_c2_21058	11	6203753	0.42	0.27	0.66
solcap_snp_c2_20952	11	6364839	0.43	0.30	0.70
solcap_snp_c2_20970	11	6398089	0.44	0.30	0.69
solcap_snp_c1_2228	11	6892437	0.44	0.27	0.61
solcap_snp_c2_3681	11	37091701	1.00	0.75	0.75
solcap_snp_c2_31337	12	4071804	0.95	0.71	0.75
solcap_snp_c2_31338	12	4074404	0.95	0.71	0.75
solcap_snp_c2_18827	12	49651778	0.23	0.23	0.98
solcap_snp_c1_11668	12	52169519	0.25	0.23	0.92
solcap_snp_c2_39409	12	52169970	0.25	0.23	0.92
solcap_snp_c2_5446	12	58744648	0.33	0.24	0.74
solcap_snp_c2_5524	12	59026489	0.34	0.24	0.70
solcap_snp_c1_2009	12	59060025	0.34	0.24	0.70

Table S6 Loci with fixed alleles in the S_3 generation

SNP	Chr	Location	DM	RH	Fixed alleles
solcap_snp_c1_2424	chr01	532390	CC	GC	CC
solcap_snp_c2_21078	chr01	2001049	AA	AG	AA
solcap_snp_c1_6674	chr01	2325968	AA	AC	AA
solcap_snp_c2_21227	chr01	2454815	CC	AC	CC
solcap_snp_c1_6704	chr01	2460933	GG	AG	GG
solcap_snp_c2_21235	chr01	2461926	GG	GA	GG
solcap_snp_c2_21236	chr01	2461945	GG	AG	GG
solcap_snp_c1_6123	chr01	2799716	CC	GC	CC
solcap_snp_c2_19261	chr01	3425136	CC	GC	CC
solcap_snp_c2_51808	chr01	3761986	AA	AC	AA
solcap_snp_c2_51806	chr01	3765562	AA	AG	AA
solcap_snp_c2_55012	chr01	4363144	CC	GC	CC
solcap_snp_c2_56125	chr01	4659424	GG	GC	GG
solcap_snp_c2_45056	chr01	4845413	CC	GC	CC
solcap_snp_c2_45064	chr01	4869706	GG	GC	GG
solcap_snp_c1_13289	chr01	4872679	GG	AG	GG
solcap_snp_c2_27882	chr01	5200836	AA	AG	AA
solcap_snp_c2_27884	chr01	5200938	AA	AG	AA
solcap_snp_c2_27885	chr01	5200977	GG	AG	GG
solcap_snp_c2_27899	chr01	5432675	AA	AG	AA
solcap_snp_c2_27918	chr01	5716932	AA	AG	AA
solcap_snp_c1_15323	chr01	6376505	AA	AG	AA
solcap_snp_c2_45625	chr01	6776901	AA	AC	AA
solcap_snp_c1_14956	chr01	16417609	CC	GC	CC
solcap_snp_c1_3693	chr01	17353216	AA	AC	AA
solcap_snp_c2_27680	chr01	18007757	GG	GC	GG
solcap_snp_c1_8541	chr01	18216831	GG	GC	GG
solcap_snp_c2_55618	chr01	21514658	GG	GC	GG
solcap_snp_c2_2873	chr01	21778683	GG	GC	GG
solcap_snp_c2_54308	chr01	24863843	GG	GC	GG
solcap_snp_c1_14647	chr01	26864819	GG	GA	GG
solcap_snp_c1_14648	chr01	26864892	CC	AC	CC
solcap_snp_c2_49720	chr01	26865021	CC	GC	CC
solcap_snp_c2_49723	chr01	26866837	AA	AG	AA
solcap_snp_c2_49724	chr01	26866854	CC	AC	CC
solcap_snp_c2_49726	chr01	26866972	CC	GC	CC
solcap_snp_c1_14654	chr01	26896631	GG	AG	GG
solcap_snp_c1_14261	chr01	27293213	CC	GC	CC
solcap_snp_c1_14248	chr01	27299025	CC	GC	CC

Table S6 Loci with fixed alleles in the S_3 generation

solcap_snp_c1_16169	chr01	27801489	GG	AG	GG
solcap_snp_c2_14939	chr01	31006712	GG	GG	GG
solcap_snp_c2_35218	chr01	31468483	GG	GC	GG
solcap_snp_c2_35221	chr01	31470481	GG	GC	GG
solcap_snp_c1_12858	chr01	34742341	CC	AC	CC
solcap_snp_c1_8908	chr01	37272787	GG	AG	GG
solcap_snp_c1_12926	chr01	37383003	AA	AG	AA
solcap_snp_c2_49867	chr01	45783793	CC	GC	CC
solcap_snp_c2_49042	chr01	48326922	GG	GC	GG
solcap_snp_c1_8358	chr01	49035726	GG	GC	GG
solcap_snp_c2_22545	chr01	49651360	GG	GC	GG
solcap_snp_c2_22531	chr01	49656404	GG	AG	GG
solcap_snp_c2_53347	chr01	53132378	AA	AG	AA
solcap_snp_c2_35480	chr01	53364418	CC	GC	CC
solcap_snp_c2_35503	chr01	53519516	CC	GC	CC
solcap_snp_c2_35536	chr01	53621301	GG	AG	GG
solcap_snp_c2_35537	chr01	53622308	GG	AG	GG
solcap_snp_c2_20798	chr01	54056807	GG	AG	GG
solcap_snp_c2_20799	chr01	54056840	AA	AG	AA
solcap_snp_c2_20895	chr01	54431892	CC	AC	CC
solcap_snp_c1_12118	chr01	55761383	AA	AG	AA
solcap_snp_c2_38404	chr01	56225790	CC	GC	CC
solcap_snp_c2_38405	chr01	56229019	CC	GC	CC
solcap_snp_c2_38428	chr01	56315308	GG	GC	GG
solcap_snp_c2_50903	chr01	56599476	AA	AG	AA
solcap_snp_c2_46195	chr01	57294821	GG	GC	GG
solcap_snp_c2_14589	chr01	59318119	GG	GC	GG
solcap_snp_c2_16039	chr01	59758200	AA	AG	AA
solcap_snp_c2_20508	chr01	60259599	CC	GC	CC
solcap_snp_c1_15782	chr03	44191	GG	AG	GG
solcap_snp_c2_54288	chr03	51153	CC	AC	CC
solcap_snp_c2_36232	chr03	635111	GG	GG	GG
solcap_snp_c2_36228	chr03	684851	GG	AG	GG
solcap_snp_c2_51389	chr03	1199021	AA	AG	AA
solcap_snp_c1_6348	chr03	15865760	GG	GC	GG
solcap_snp_c2_20347	chr03	15874080	AA	AG	AA
solcap_snp_c2_54785	chr03	22985036	GG	GG	GG
solcap_snp_c1_5812	chr03	39356741	GG	GG	GG
solcap_snp_c2_17770	chr03	39535422	CC	AC	CC
solcap_snp_c1_7096	chr03	40161213	CC	GC	CC
solcap_snp_c2_37100	chr03	45601665	CC	GC	CC
solcap_snp_c2_37119	chr03	45725401	AA	GA	AA

Table S6 Loci with fixed alleles in the S_3 generation

solcap_snp_c2_37121	chr03	45727779	AA	AG	AA
solcap_snp_c2_37139	chr03	45805743	GG	GC	GG
solcap_snp_c2_9531	chr03	46564231	GG	GC	GG
solcap_snp_c2_9498	chr03	46662640	GG	GC	GG
solcap_snp_c2_9497	chr03	46662711	AA	AG	AA
solcap_snp_c1_7722	chr03	47791222	GG	AG	GG
solcap_snp_c2_578	chr03	47830924	AA	AG	AA
solcap_snp_c2_56989	chr04	586240	AA	AC	AA
solcap_snp_c2_56991	chr04	586420	GG	AG	GG
solcap_snp_c2_54463	chr04	663212	GG	GC	GG
solcap_snp_c2_56206	chr04	710302	AA	AG	AA
solcap_snp_c1_16291	chr04	727721	GG	GC	GG
solcap_snp_c1_11242	chr04	824482	AA	AG	AA
solcap_snp_c2_39322	chr04	1529818	GG	AG	GG
solcap_snp_c1_16358	chr04	2002436	GG	GC	GG
solcap_snp_c2_31734	chr04	3324742	GG	AG	GG
solcap_snp_c2_31719	chr04	3347797	GG	AG	GG
solcap_snp_c2_31712	chr04	3379356	GG	GG	GG
solcap_snp_c2_21915	chr04	3955470	AA	AG	AA
solcap_snp_c2_21842	chr04	4181223	AA	AG	AA
solcap_snp_c2_21858	chr04	4340263	AA	AG	AA
solcap_snp_c1_3707	chr04	4516442	GG	AG	GG
solcap_snp_c2_11534	chr04	4619456	AA	AG	AA
solcap_snp_c2_11549	chr04	4853343	CC	AC	CC
solcap_snp_c2_11427	chr04	5210381	GG	AG	GG
solcap_snp_c2_11432	chr04	5232017	CC	GC	CC
solcap_snp_c2_11435	chr04	5290306	GG	GG	GG
solcap_snp_c2_11487	chr04	5371345	AA	AC	AA
solcap_snp_c2_11489	chr04	5373006	CC	GC	CC
solcap_snp_c2_11490	chr04	5373686	AA	AG	AA
solcap_snp_c2_48871	chr04	5817865	CC	GC	CC
solcap_snp_c2_48867	chr04	5839869	CC	GC	CC
solcap_snp_c2_48866	chr04	5839959	CC	GC	CC
solcap_snp_c2_48865	chr04	5839983	GG	AG	GG
solcap_snp_c2_48864	chr04	5840120	GG	GC	GG
solcap_snp_c2_48863	chr04	5840137	CC	GC	CC
solcap_snp_c2_48861	chr04	5841559	GG	GC	GG
solcap_snp_c2_48854	chr04	5847541	CC	GC	CC
solcap_snp_c1_11722	chr04	6259961	GG	GC	GG
solcap_snp_c2_39624	chr04	6354222	GG	GG	GG
solcap_snp_c2_39594	chr04	6511940	AA	GA	AA
solcap_snp_c2_39597	chr04	6512127	AA	AG	AA

Table S6 Loci with fixed alleles in the S_3 generation

solcap_snp_c2_47320	chr04	6707642	AA	AC	AA
solcap_snp_c2_26858	chr04	8401238	GG	GG	GG
solcap_snp_c2_26845	chr04	8461508	GG	AG	GG
solcap_snp_c2_26800	chr04	8686432	GG	GC	GG
solcap_snp_c2_26795	chr04	8687273	CC	GC	CC
solcap_snp_c2_26794	chr04	8687360	AA	AG	AA
solcap_snp_c2_26793	chr04	8687531	AA	AC	AA
solcap_snp_c1_8347	chr04	8687949	GG	GG	GG
solcap_snp_c2_26792	chr04	8688518	GG	GC	GG
solcap_snp_c1_8353	chr04	9008187	AA	AC	AA
solcap_snp_c2_29468	chr04	9177863	GG	GC	GG
solcap_snp_c2_38246	chr04	9457412	CC	GC	CC
solcap_snp_c2_38247	chr04	9457473	AA	AG	AA
solcap_snp_c2_10461	chr04	13052212	GG	GG	GG
solcap_snp_c1_3436	chr04	13225401	GG	GC	GG
solcap_snp_c2_11107	chr04	13362745	AA	AG	AA
solcap_snp_c1_16643	chr04	14017130	GG	GG	GG
solcap_snp_c2_55707	chr04	19354530	CC	GC	CC
solcap_snp_c2_55709	chr04	19369749	CC	GC	CC
solcap_snp_c2_38379	chr04	21810688	CC	AC	CC
solcap_snp_c2_32900	chr04	22656507	AA	AG	AA
solcap_snp_c2_32901	chr04	22656518	GG	AG	GG
solcap_snp_c2_32904	chr04	22668757	CC	GC	CC
solcap_snp_c1_9837	chr04	22670992	CC	GC	CC
solcap_snp_c1_9477	chr04	26422129	AA	AG	AA
solcap_snp_c2_31358	chr04	26486034	CC	GC	CC
solcap_snp_c2_31360	chr04	26486172	AA	GA	AA
solcap_snp_c2_58366	chr04	28706962	GG	GG	GG
solcap_snp_c2_44554	chr04	30765793	GG	AG	GG
solcap_snp_c2_37598	chr04	31711158	GG	AG	GG
solcap_snp_c2_37596	chr04	31711693	GG	AG	GG
solcap_snp_c2_49995	chr04	33237951	CC	GC	CC
solcap_snp_c1_9111	chr04	34176938	GG	GG	GG
solcap_snp_c1_3319	chr04	40340185	CC	AC	CC
solcap_snp_c2_16722	chr04	41820065	CC	AC	CC
solcap_snp_c2_16718	chr04	42164205	AA	AG	AA
solcap_snp_c1_6036	chr04	44234876	GG	GG	GG
solcap_snp_c1_15977	chr04	44779459	GG	GC	GG
solcap_snp_c2_54533	chr04	45698591	AA	AG	AA
solcap_snp_c1_15762	chr04	45804342	GG	GG	GG
solcap_snp_c2_48694	chr04	46675432	GG	AG	GG
solcap_snp_c2_48692	chr04	46734716	GG	AG	GG

Table S6 Loci with fixed alleles in the S_3 generation

solcap_snp_c2_48693	chr04	46736735	GG	GC	GG
solcap_snp_c2_52036	chr04	47018654	GG	GG	GG
solcap_snp_c2_57079	chr04	47393688	CC	GC	CC
solcap_snp_c2_45035	chr04	47456994	GG	GC	GG
solcap_snp_c2_55795	chr04	58101212	CC	CC	CC
solcap_snp_c2_34937	chr04	60390283	AA	AC	AA
solcap_snp_c2_12981	chr04	61157969	GG	GC	GG
solcap_snp_c2_12928	chr04	61453515	GG	GC	GG
solcap_snp_c1_4171	chr04	61467478	GG	GC	GG
solcap_snp_c2_12924	chr04	61476288	GG	GG	GG
solcap_snp_c1_4140	chr04	61780336	CC	AC	CC
solcap_snp_c2_35970	chr04	62863012	AA	AG	AA
solcap_snp_c2_35942	chr04	63010178	GG	AG	GG
solcap_snp_c1_10670	chr04	63011483	CC	GC	CC
solcap_snp_c1_10668	chr04	63012168	GG	GC	GG
solcap_snp_c2_23776	chr05	65193	CC	AC	CC
solcap_snp_c2_23804	chr05	71209	AA	GA	AA
solcap_snp_c2_23669	chr05	190926	AA	AC	AA
solcap_snp_c2_23678	chr05	224801	CC	AC	CC
solcap_snp_c2_23713	chr05	288467	GG	GG	GG
solcap_snp_c2_23729	chr05	325030	GG	AG	GG
solcap_snp_c2_23739	chr05	389174	CC	GC	CC
solcap_snp_c2_23740	chr05	389266	CC	GC	CC
solcap_snp_c2_23741	chr05	389308	CC	GC	CC
solcap_snp_c1_7621	chr05	391001	CC	GC	CC
solcap_snp_c1_7622	chr05	391031	AA	GA	AA
solcap_snp_c2_23743	chr05	391812	GG	GC	GG
solcap_snp_c2_23829	chr05	694151	CC	GC	CC
solcap_snp_c2_23843	chr05	705734	GG	GG	GG
solcap_snp_c2_33511	chr05	1401063	CC	GC	CC
solcap_snp_c2_33513	chr05	1401419	GG	GC	GG
solcap_snp_c2_33516	chr05	1413720	AA	AC	AA
solcap_snp_c2_33517	chr05	1413732	AA	AC	AA
solcap_snp_c2_33522	chr05	1414020	AA	AG	AA
solcap_snp_c2_33535	chr05	1437829	GG	AG	GG
solcap_snp_c1_10042	chr05	1438051	CC	GC	CC
solcap_snp_c2_52070	chr05	1853441	GG	AG	GG
solcap_snp_c2_11731	chr05	2078804	CC	AC	CC
solcap_snp_c2_11727	chr05	2079217	GG	AG	GG
solcap_snp_c1_3840	chr05	2961084	GG	GG	GG
solcap_snp_c2_47646	chr05	5562568	GG	GC	GG
solcap_snp_c2_47611	chr05	5613941	GG	AG	GG

Table S6 Loci with fixed alleles in the S_3 generation

solcap_snp_c2_47284	chr05	5995560	GG	AG	GG
solcap_snp_c1_11078	chr05	7938605	GG	AG	GG
solcap_snp_c2_54372	chr05	44089217	GG	AG	GG
solcap_snp_c2_54371	chr05	44098293	GG	AG	GG
solcap_snp_c2_42542	chr05	44213287	AA	AG	AA
solcap_snp_c2_8508	chr05	45558556	GG	AG	GG
solcap_snp_c2_8515	chr05	45558994	CC	GC	CC
solcap_snp_c2_8521	chr05	45559357	GG	GC	GG
solcap_snp_c2_24311	chr06	5482650	GG	GC	GG
solcap_snp_c1_12409	chr06	24124211	GG	AG	GG
solcap_snp_c2_43093	chr06	33378919	AA	AG	AA
solcap_snp_c2_33375	chr06	36052643	GG	GG	GG
solcap_snp_c2_52390	chr06	37041529	GG	AG	GG
solcap_snp_c2_49053	chr06	37797806	GG	GG	GG
solcap_snp_c2_49052	chr06	37797840	CC	GC	CC
solcap_snp_c2_37762	chr06	39057655	CC	GC	CC
solcap_snp_c2_46171	chr06	39520903	GG	GC	GG
solcap_snp_c2_16863	chr06	40072110	GG	GC	GG
solcap_snp_c2_47782	chr06	40676545	GG	AG	GG
solcap_snp_c1_15371	chr06	41766936	GG	GC	GG
solcap_snp_c2_16804	chr06	42061087	AA	AG	AA
solcap_snp_c2_16793	chr06	42113554	GG	GG	GG
solcap_snp_c2_16785	chr06	42130443	AA	AG	AA
solcap_snp_c2_54194	chr06	43861021	GG	AG	GG
solcap_snp_c1_15755	chr06	43861533	GG	AG	GG
solcap_snp_c2_54191	chr06	43911541	CC	GC	CC
solcap_snp_c2_31873	chr06	44174749	GG	GG	GG
solcap_snp_c2_43960	chr07	102420	GG	GG	GG
solcap_snp_c2_46111	chr07	1829920	GG	GG	GG
solcap_snp_c2_46110	chr07	1829983	GG	GC	GG
solcap_snp_c2_46107	chr07	1836511	CC	GC	CC
solcap_snp_c2_46100	chr07	1838559	GG	AG	GG
solcap_snp_c2_46081	chr07	1862875	GG	AG	GG
solcap_snp_c2_52905	chr07	2114063	GG	AG	GG
solcap_snp_c2_36872	chr07	2521171	GG	GC	GG
solcap_snp_c2_36869	chr07	2557728	CC	GC	CC
solcap_snp_c2_36859	chr07	2575602	GG	AG	GG
solcap_snp_c1_10974	chr07	2645272	GG	GG	GG
solcap_snp_c2_36838	chr07	2646784	CC	GC	CC
solcap_snp_c2_36835	chr07	2647468	GG	GC	GG
solcap_snp_c2_36833	chr07	2649020	AA	AG	AA
solcap_snp_c2_36818	chr07	2678451	GG	GC	GG

Table S6 Loci with fixed alleles in the S_3 generation

solcap_snp_c2_36882	chr07	2839957	GG	GG	GG
solcap_snp_c1_8186	chr07	2908663	AA	AG	AA
solcap_snp_c2_26197	chr07	3363569	GG	AG	GG
solcap_snp_c2_26296	chr07	3396876	GG	AG	GG
solcap_snp_c2_26145	chr07	3450008	CC	GC	CC
solcap_snp_c2_26154	chr07	3458322	CC	GC	CC
solcap_snp_c2_26166	chr07	3474373	GG	GG	GG
solcap_snp_c2_26182	chr07	3537096	AA	AG	AA
solcap_snp_c2_43607	chr07	3752682	CC	GC	CC
solcap_snp_c2_43574	chr07	3879539	GG	GC	GG
solcap_snp_c2_46379	chr07	4182885	GG	AG	GG
solcap_snp_c1_16227	chr07	4386467	GG	AG	GG
solcap_snp_c1_16221	chr07	4388800	AA	AG	AA
solcap_snp_c2_55833	chr07	4390187	CC	GC	CC
solcap_snp_c2_55831	chr07	4390352	CC	GC	CC
solcap_snp_c2_53197	chr07	4520482	AA	AC	AA
solcap_snp_c1_15485	chr07	5750279	AA	AG	AA
solcap_snp_c2_58112	chr07	40724780	GG	GA	GG
solcap_snp_c2_19746	chr07	44177848	AA	AC	AA
solcap_snp_c2_19748	chr07	44184190	CC	GC	CC
solcap_snp_c2_19787	chr07	44397124	CC	GC	CC
solcap_snp_c2_19803	chr07	44434077	CC	AC	CC
solcap_snp_c2_35100	chr07	46234254	GG	GG	GG
solcap_snp_c2_35058	chr07	46414678	GG	GA	GG
solcap_snp_c2_26003	chr07	46817851	GG	AG	GG
solcap_snp_c2_26006	chr07	46818484	AA	AG	AA
solcap_snp_c2_26008	chr07	46818557	AA	AG	AA
solcap_snp_c2_26011	chr07	46819043	GG	GC	GG
solcap_snp_c2_26014	chr07	46822436	CC	CG	CC
solcap_snp_c2_26015	chr07	46822685	GG	AG	GG
solcap_snp_c1_8686	chr07	48163662	GG	AG	GG
solcap_snp_c2_12720	chr07	48500388	CC	GC	CC
solcap_snp_c2_12601	chr07	48761616	AA	AG	AA
solcap_snp_c2_12598	chr07	48761838	GG	AG	GG
solcap_snp_c2_12597	chr07	48761853	GG	GA	GG
solcap_snp_c2_12578	chr07	48790114	GG	GC	GG
solcap_snp_c2_42758	chr07	49676106	CC	GC	CC
solcap_snp_c2_42760	chr07	49676336	GG	GC	GG
solcap_snp_c2_42762	chr07	49676492	GG	AG	GG
solcap_snp_c1_8854	chr08	384925	GG	GG	GG
solcap_snp_c2_29046	chr08	403148	AA	GA	AA
solcap_snp_c2_29015	chr08	756540	AA	AG	AA

Table S6 Loci with fixed alleles in the S_3 generation

solcap_snp_c1_9779	chr08	1067135	GG	AG	GG
solcap_snp_c1_9785	chr08	1179922	GG	GG	GG
solcap_snp_c2_32667	chr08	1218944	CC	GC	CC
solcap_snp_c1_9786	chr08	1220894	CC	GC	CC
solcap_snp_c2_24407	chr08	1665188	AA	AC	AA
solcap_snp_c1_15756	chr08	1898015	AA	AC	AA
solcap_snp_c2_54204	chr08	1901593	CC	AC	CC
solcap_snp_c2_30053	chr08	5249494	AA	GA	AA
solcap_snp_c2_30052	chr08	5249558	GG	GA	GG
solcap_snp_c1_15451	chr08	6017567	CC	GC	CC
solcap_snp_c2_52856	chr08	6020920	GG	GA	GG
solcap_snp_c2_52857	chr08	6022566	GG	AG	GG
solcap_snp_c2_49541	chr08	7770995	GG	GG	GG
solcap_snp_c1_5713	chr08	8300725	AA	AG	AA
solcap_snp_c1_14542	chr08	18181335	AA	AG	AA
solcap_snp_c2_49245	chr08	23950636	CC	AC	CC
solcap_snp_c1_12162	chr08	24327297	AA	AG	AA
solcap_snp_c2_41463	chr08	24514668	CC	GC	CC
solcap_snp_c2_56491	chr08	24705338	CC	GC	CC
solcap_snp_c1_838	chr08	25418116	GG	AG	GG
solcap_snp_c2_44331	chr08	26305549	GG	AG	GG
solcap_snp_c1_13043	chr08	26412608	GG	GG	GG
solcap_snp_c2_44335	chr08	26435309	AA	AG	AA
solcap_snp_c2_44334	chr08	26439800	AA	AG	AA
solcap_snp_c2_47468	chr08	27089659	CC	AA	CC
solcap_snp_c2_51372	chr08	27187626	CC	GC	CC
solcap_snp_c2_51374	chr08	27227035	AA	AG	AA
solcap_snp_c2_51370	chr08	27278313	AA	AG	AA
solcap_snp_c1_15045	chr08	27278652	CC	CG	CC
solcap_snp_c1_12166	chr08	27410482	GG	AG	GG
solcap_snp_c2_51329	chr08	28006950	GG	GC	GG
solcap_snp_c2_45770	chr08	28245261	GG	GC	GG
solcap_snp_c2_45751	chr08	28386854	CC	GC	CC
solcap_snp_c2_15834	chr08	29989855	GG	AG	GG
solcap_snp_c2_48182	chr08	30596949	GG	GC	GG
solcap_snp_c2_48184	chr08	30597181	AA	AG	AA
solcap_snp_c1_14271	chr08	30597561	GG	AG	GG
solcap_snp_c2_18918	chr08	31893468	GG	AG	GG
solcap_snp_c2_50153	chr08	32790084	GG	GG	GG
solcap_snp_c2_50150	chr08	32817651	CC	GC	CC
solcap_snp_c2_51052	chr08	33576455	GG	AG	GG
solcap_snp_c1_10397	chr08	37002746	AA	AC	AA

Table S6 Loci with fixed alleles in the S_3 generation

solcap_snp_c2_34709	chr08	37060418	GG	GG	GG
solcap_snp_c2_34698	chr08	37078384	GG	AG	GG
solcap_snp_c2_34717	chr08	37235593	AA	AC	AA
solcap_snp_c2_57476	chr08	37660078	AA	AC	AA
solcap_snp_c2_36747	chr08	37960972	GG	GC	GG
solcap_snp_c2_36748	chr08	37960983	GG	AG	GG
solcap_snp_c2_36749	chr08	38029997	CC	GC	CC
solcap_snp_c2_36779	chr08	38125327	GG	AG	GG
solcap_snp_c2_19144	chr08	39132248	AA	AG	AA
solcap_snp_c1_6052	chr08	39153909	GG	AG	GG
solcap_snp_c2_34634	chr08	39975274	GG	GG	GG
solcap_snp_c2_34590	chr08	40083266	CC	GC	CC
solcap_snp_c2_34587	chr08	40106266	CC	GC	CC
solcap_snp_c2_40086	chr09	49152324	GG	GC	GG
solcap_snp_c2_3021	chr09	50232550	AA	AG	AA
solcap_snp_c2_47950	chr09	50959925	GG	GC	GG
solcap_snp_c2_48041	chr09	51220125	GG	GC	GG
solcap_snp_c1_10579	chr09	52212398	AA	AG	AA
solcap_snp_c2_37960	chr10	46388790	AA	AG	AA
solcap_snp_c2_37961	chr10	46388841	GG	GC	GG
solcap_snp_c1_9499	chr11	15218007	CC	AC	CC
solcap_snp_c2_31472	chr11	15220676	AA	AC	AA
solcap_snp_c2_31444	chr11	15386201	AA	AC	AA
solcap_snp_c1_16141	chr11	22355078	CC	GC	CC
solcap_snp_c2_41084	chr11	29406549	GG	GC	GG
solcap_snp_c2_30361	chr11	31833113	CC	GC	CC
solcap_snp_c2_29435	chr11	31970415	GG	GC	GG
solcap_snp_c1_4378	chr11	32941597	CC	GC	CC
solcap_snp_c2_13613	chr11	33177152	GG	GC	GG
solcap_snp_c2_13627	chr11	33240507	CC	GC	CC
solcap_snp_c2_13629	chr11	33240687	AA	AG	AA
solcap_snp_c2_13632	chr11	33242504	AA	AG	AA
solcap_snp_c2_13633	chr11	33261131	CC	GC	CC
solcap_snp_c2_13634	chr11	33265279	GG	GC	GG
solcap_snp_c2_13636	chr11	33349313	AA	AC	AA
solcap_snp_c1_4371	chr11	33598950	GG	GC	GG
solcap_snp_c2_13590	chr11	33722886	CC	GC	CC
solcap_snp_c1_4359	chr11	33882885	GG	GC	GG
solcap_snp_c2_46858	chr11	34424525	CC	GC	CC
solcap_snp_c2_56243	chr11	34694629	CC	GC	CC
solcap_snp_c2_45206	chr11	34910782	AA	AC	AA
solcap_snp_c2_14947	chr11	35243976	GG	GC	GG

Table S6 Loci with fixed alleles in the S_3 generation

solcap_snp_c1_4822	chr11	35245870	GG	AG	GG
solcap_snp_c1_1318	chr11	36592986	GG	GC	GG
solcap_snp_c1_6964	chr11	37664069	GG	AG	GG
solcap_snp_c2_22165	chr11	37684326	CC	AC	CC
solcap_snp_c2_3753	chr11	37938261	AA	AC	AA
solcap_snp_c2_3747	chr11	37955334	GG	GG	GG
solcap_snp_c1_7766	chr12	239176	GG	GC	GG
solcap_snp_c2_24561	chr12	243865	GG	GA	GG
solcap_snp_c2_24553	chr12	244963	AA	AG	AA
solcap_snp_c1_8641	chr12	4235029	GG	GC	GG
solcap_snp_c2_27931	chr12	4560751	AA	AC	AA
solcap_snp_c2_27929	chr12	4560853	GG	AG	GG
solcap_snp_c2_27926	chr12	4566912	AA	AG	AA
solcap_snp_c1_13762	chr12	5169621	CC	GC	CC
solcap_snp_c2_46285	chr12	5222200	GG	GC	GG
solcap_snp_c2_8037	chr12	55640932	GG	GC	GG
solcap_snp_c1_2689	chr12	55982568	GG	GC	GG
solcap_snp_c2_46213	chr12	57064746	GG	AG	AA
solcap_snp_c1_1944	chr12	57695670	GG	AG	AA
solcap_snp_c1_1923	chr12	57805264	CC	GC	CC
solcap_snp_c2_5474	chr12	58642405	AA	AG	AA
solcap_snp_c2_5463	chr12	58688510	GG	GC	GG
solcap_snp_c2_5461	chr12	58689924	GG	AG	GG
solcap_snp_c1_15241	no hit		GG	GC	GG
solcap_snp_c1_16079	no hit		GG	AG	GG
solcap_snp_c1_16349	no hit		AA	AG	AA
solcap_snp_c1_7569	no hit		GG	AG	GG
solcap_snp_c2_2333	no hit		CC	GC	CC
solcap_snp_c2_23511	no hit		CC	GC	CC
solcap_snp_c2_23593	no hit		CC	GC	CC
solcap_snp_c2_23596	no hit		GG	AG	GG
solcap_snp_c2_23602	no hit		CC	GC	CC
solcap_snp_c2_23603	no hit		CC	AC	CC
solcap_snp_c2_34103	no hit		GG	GG	GG
solcap_snp_c2_39627	no hit		GG	AG	GG
solcap_snp_c2_41106	no hit		GG	GG	GG
solcap_snp_c2_46329	no hit		GG	AG	GG
solcap_snp_c2_46468	no hit		AA	AG	AA
solcap_snp_c2_51546	no hit		CC	GC	CC
solcap_snp_c2_51639	no hit		GG	AG	GG
solcap_snp_c2_51647	no hit		GG	AG	GG
solcap_snp_c2_51650	no hit		CC	GC	CC

Table S6 Loci with fixed alleles in the S_3 generation

solcap_snp_c2_52209	no hit	GG	AG	GG
solcap_snp_c2_52800	no hit	GG	AG	GG
solcap_snp_c2_52815	no hit	AA	AG	AA
solcap_snp_c2_53198	no hit	GG	GC	GG
solcap_snp_c2_53200	no hit	GG	GG	GG
solcap_snp_c2_53903	no hit	GG	AG	GG
solcap_snp_c2_54378	no hit	GG	GC	GG
solcap_snp_c2_54760	no hit	GG	GC	GG
solcap_snp_c2_54804	no hit	CC	AC	CC
solcap_snp_c2_55700	no hit	GG	GC	GG
solcap_snp_c2_56714	no hit	GG	GC	GG
solcap_snp_c2_57482	no hit	AA	AG	AA
solcap_snp_c2_57483	no hit	CC	AC	CC