bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Identification of candidate genes for self-compatibility in perennial ryegrass (*Lolium perenne* L.)

Authors

Claudio Cropano[1,2], Chloé Manzanares[1], Steven Yates[1], Dario Copetti[1,3], Javier Do Canto[4], Thomas Lübberstedt[5], Michael Koch[2] and Bruno Studer[1*]

[*] Corresponding author. E-mail: bruno.studer@usys.ethz.ch

[1] Molecular Plant Breeding, Institute of Agricultural Sciences, ETH Zurich, Universitaetstrasse 2, 8092 Zurich, Switzerland

[2] Deutsche Saatveredelung AG, Weissenburger Str. 5, 59557 Lippstadt, Germany

[3] Department of Evolutionary Biology and Environmental Studies, University of Zurich, 8057 Zurich, Switzerland

[4] Instituto Nacional de Investigación Agropecuaria (INIA), Ruta 5km 386, 45000 Tacuarembó, Uruguay

[5] Department of Agronomy, Iowa State University, 716 Farm House Ln, 50011 Ames, USA

Key message

A previously reported QTL conferring self-compatibility to perennial ryegrass was fine mapped to a 0.26 cM locus, corresponding to an estimated ~ 3 Mb physical region containing 57 predicted protein-encoding genes.

Abstract

Self-incompatibility (SI) is a genetic mechanism preventing self-pollination in approximately 40% of plant species. Two multiallelic loci, called *S* and *Z*, control the gametophytic SI system of the grass family (Poaceae), which contains all major forage grasses. Loci independent from *S* and *Z* have been reported to disrupt SI and lead to self-compatibility (SC). A locus causing SC in perennial ryegrass (*Lolium perenne* L.) was previously mapped on linkage group (LG) 5 in a F_2 population segregating for SC. Using a subset of the same population (n=68), we first performed low-resolution quantitative trait locus (QTL) mapping to exclude the presence of additional, previously undetected contributors to SC. The previously reported QTL on LG 5 explained 38.4% of the phenotypic variation,

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

and no significant contribution from other genomic regions was found. This was verified by the presence of significantly distorted markers in the region overlapping with the QTL. Second, we fine mapped the QTL to 0.26 cM using additional 2,056 plants and 23 novel sequence-based markers. Using an Italian ryegrass (*Lolium multiflorum* Lam.) genome assembly as reference, the markers flanking SC were estimated to span a ~3 Mb region encoding for 57 predicted genes. Among these, seven genes were proposed as relevant candidate genes based on their annotation and function described in previous studies. Our work is a step forward to identify SC genes in forage grasses and provides diagnostic markers for marker-assisted introgression of SC into elite germplasm.

Key words: self-incompatibility (SI), self-compatibility (SC), perennial ryegrass (*Lolium perenne* L.), quantitative trait locus (QTL) analysis, segregation distortion, fine mapping, candidate genes

Introduction

Several species of the grass family (Poaceae) are economically valuable forage crops and represent a fundamental component of our grasslands. Most forage grasses are obligate outcrossers due to a gametophytic self-incompatibility (SI) system that prevents inbreeding (Cornish et al. 1979). Unlike most self-incompatible species, SI in grasses is controlled by two multiallelic loci, called *S* and *Z*. Their existence is known since more than half a century (Hayman 1956; Lundqvist 1956) but, despite several mapping efforts (Voylokov et al. 1998; Thorogood and Armstead 2002; Bian et al. 2004; Hackauf and Wehling 2005; Kakeda et al. 2008; Shinozuka et al. 2010), the identity of the genes involved is unknown. To date, two genes encoding for proteins with a DUF247 domain on linkage group (LG) 1 and 2 are the most promising candidates for *S* and *Z*, respectively, in perennial ryegrass (*Lolium perenne* L.) (Shinozuka et al. 2010; Manzanares et al. 2016a).

There is a continued need to accelerate breeding in forage grasses for traits such as biomass yield and quality, seed yield, and disease resistance (Capstaff and Miller 2018). However, the genetic gain achieved in self-incompatible grasses with population breeding strategies has lagged behind compared to inbred crops such as maize (*Zea mays* L.), wheat (*Triticum aestivum* L.), and rice (*Oriza sativa* L.) (Laidig et al. 2014). This is partially caused by the inability to develop F_1-hybrid cultivars and exploit the hybrid vigour that occurs when two genetically distant (and usually

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

highly homozygous) parents are crossed. Thus, the possibility to overcome SI and exploit self-compatibility (SC) to transition forage grasses to an inbred line-based breeding system is gaining traction (Do Canto et al. 2016; Herridge et al. 2020). Inbreeding by self-pollination fixes favourable genetic variants and purges deleterious alleles with high efficiency, resulting in highly homozygous inbred lines (Jansky et al. 2016). Inbred lines belonging to different heterotic groups can be crossed to systematically assemble desirable combinations of genes and maximize heterosis.

Several routes to SC have been described in grasses: SC can arise from mutations at *S* and/or *Z* that disrupt the initial self/non-self recognition between pollen and stigma (Do Canto et al. 2016). Mutations in loci non-allelic to *S* and *Z* can also lead to SC, by interrupting the downstream cascade triggered by the initial self-recognition (Do Canto et al. 2016). Such putative mutations have been found in rye (*Secale cereale* L.) and perennial ryegrass (Voylokov et al. 1993; Egorova et al. 2000; Thorogood et al. 2005; Arias-Aguirre et al. 2013, Do Canto et al. 2018; Slatter et al. 2020). A detailed molecular understanding of how SC arises in SI grasses is lacking, but several studies have acknowledged that calcium (Ca^{2+}) signaling plays an essential role in the recognition and/or inhibition of self-pollen (Yang et al. 2009; Chen et al. 2019). In fact, SI in grasses can be partially overcome by treating self-pollinated stigmas with chemical reagents affecting Ca^{2+} channeling across cell membranes (Klaas et al. 2011).

Two consecutive studies provided evidence on the location of a SC source segregating in a F_2 population of perennial ryegrass. Using *in vitro* pollinations, Arias-Aguirre et al. (2013) observed a 1:1 segregation into two phenotypic SC classes: plants showing a 50% pollen compatibility where half of the self-pollen germinated and grew a pollen tube upon contact with the stigma, and 100% SC where all self-pollen showed a compatible reaction. It was concluded that SC was caused by a putative mutation in a major gametophytic gene mapped to a 16 cM locus on LG 5. Do Canto et al. (2018) fine mapped the position of the locus to a 1.6 cM region. Despite providing evidence of one major locus explaining the SC variation, these two studies did not exclude the presence of additional major effects on other LGs. In addition, the genetic region co-segregating with SC in Do Canto et al. (2018) was inferred based on genome sequence information from rice and Brachypodium (*Brachypodium distachyon* (L.) P. Beauv.) and still considerably large, making it difficult to identify putative candidate genes for functional validation.

Building on these previous studies, we aimed at providing additional evidence for the quantitative trait locus (QTL) on LG 5 being the sole cause for SC variation in this population. An additional objective was to increase the genetic mapping resolution achieved by Do Canto et al. (2018) substantially, to facilitate transitioning from a genetic to a physical map in ryegrass, and to select and prioritize candidate genes in the target region. A better understanding of the genetic causes of SC can help biologists to determine the pathways involved in the SI system of grasses. In

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

addition, it paves the way for shifting the breeding of perennial grasses from population-based to pure line-based F_1 hybrid strategies.

Material and Methods

Plant material

This study was entirely based on the perennial ryegrass F_2 population segregating for SC described in Arias-Aguirre *et al.* (2013) and Do Canto *et al.* (2018). The population was obtained by self-pollinating a single F_1 individual originating from the initial cross between a self-compatible genotype (selfed for five generations) and a self-incompatible individual of the VrnA mapping population (Jensen *et al.*, 2005). A subset of 74 individuals of the population used in Do Canto *et al.* (2018) was used as starting material for linkage map construction, low-resolution QTL mapping and the analysis of segregation distortion. A larger set of 2,056 individuals, in addition to the 1,248 reported in Do Canto *et al.* (2018), was used for the fine mapping.

DNA extraction

All genomic DNA used in this study was extracted from powdered frozen leaf tissue using the Mag-Bind® Plant DNA DS 96 Kit (Omega Bio-tek, Inc. Norcross, GA, USA) on a 96-well plate KingFisher Flex Purification System (Thermo Fischer Scientific, Waltham, MA, USA) following manufacturer's recommendations. Genomic DNA was visualized on a 1% agarose gel and quantified with a NanoDrop 8000 spectrophotometer (Thermo Fisher Scientific,Waltham, MA, USA).

QTL mapping and analysis of segregation distortion

Phenotyping for self-compatibility

The SC phenotypic data of the 74 F_2 individuals was obtained from the *in vitro* pollination assays data reported in Do Canto *et al.* (2018). Briefly, the number of compatible pollen grains, usually translucent and with a bright pollen tubes growing towards the style, was counted on at least ten stigmas per plant. The data for each plant consisted of the overall mean percentage of self-compatible pollen, using the stigmas as subsamples. A chi-square goodness of fit test was performed to test if the phenotypic variation deviated from the hypothesis of SC being under the control of a single gametophytic gene. In that case, only heterozygote and homozygote plants for the SC allele (corresponding to the 50 and 100% phenotypic class) are expected in a 1:1 ratio in the F_2.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Genotyping-by-sequencing, data processing and DNA variant calling

A genotyping-by-sequencing (GBS) library was prepared following the protocol reported in Begheyn *et al.* (2018) starting from genomic DNA of 74 individuals of the F_2 population and sequenced using 125 bp single-end reads on an Illumina HiSeq 2500 platform at the Functional Genomics Center Zurich (FGCZ), Zurich, Switzerland. Sequencing reads were demultiplexed using Sabre (https://github.com/najoshi/sabre), allowing no mismatches. After demultiplexing, all adapters and barcodes were removed from the reads which were then trimmed to 100 bp. Variant calling was performed with the GBSmode pipeline (https://github.com/stevenandrewyates/GBSmode) using the Italian ryegrass (*Lolium multiflorum* Lam.) genome assembly as a reference (Copetti *et al.*, 2021).

Genetic linkage map construction

A genetic linkage map was constructed using the R package R/qtl (Broman *et al.*, 2003). Single nucleotide polymorphisms (SNP) resulting from the GBS variant calling were filtered for a 5% MAF, < 40% missing values, and distorted segregation from the expected 1:2:1 ratio in a F_2 population. Individuals showing < 40% missing genotype calls were also filtered. In addition, pairs of individuals with more than 90% matching genotypes were identified and one was removed. As a result, 473 high-quality SNPs and 68 individuals were kept for the construction of the genetic map. Markers were grouped in LGs with the *formLinkageGroups* function with a maximum recombination rate of 0.35 and minimum $-\log_{10}(p\text{-value})$ (LOD) threshold of 6. The initial marker order and genetic distances were established using the Kosambi mapping function ($d = (1/4)\ln(1+2r/1-2r)$), where d is the mapping distance and r is the recombination frequency (Kosambi, 1943). Singletons and double recombinations inflating the map size were identified and corrected using a graphical genotyping approach (Young and Tanksley, 1989). All singletons not followed by at least four markers of the same haplotype were turned into missing values and marker order and genetic distances were re-calculated. This process was repeated until all singletons and double recombinations within a window of four markers were removed. Linkage group numbers were assigned based on the physical position of their markers on the Italian ryegrass assembly (Copetti *et al.*, 2021).

QTL mapping and analysis of segregation distortion

The R package R/qtl was used to perform QTL mapping (Broman *et al.*, 2003). Missing values in the genotype table were filled with the *fill.geno* function and QTL mapping was carried out using the *mqmscan* function, which scans

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

the genome with a multiple QTL model based on haplotype dominance and considers additive effects. After running 1,000 permutations with an assumed genotyping error rate of 0.05, a LOD of 2.81 was set as the QTL significance threshold. Segregation distortion was calculated using a chi-square goodness of fit test to identify markers that significantly differed from the 1:2:1 expected segregation. A Bonferroni-corrected significance threshold was set at a LOD of 3.97.

Fine mapping

Marker development for fine mapping

To detect novel DNA sequence polymorphisms for marker development, 10 ng of genomic DNA from 50 individuals belonging to the collection of 2,056 F_2 individuals were pooled and sequenced at low coverage in 2x150 bp mode on an S4 flow cell lane of an Illumina Novaseq instrument at FGCZ. The paired-end reads were mapped to an Italian ryegrass reference assembly (Copetti *et al.*, 2021) using bowtie2 (v3.5.1) (Langmead and Salzberg, 2012) and alignments were converted to BAM format with SAMtools (v1.10) (Li *et al.*, 2009). A BAM file containing reads that aligned on the scaffolds delimited by the markers flanking the SC locus in Do Canto *et al.* (2018) (G05_065 and G05_095) was imported into the Integrative Genomics Viewer software (Robinson *et al.* 2011) and polymorphisms were visually inspected. To develop PCR-based markers, a 50-120 bp of consensus sequence spanning a single polymorphism showing the expected 50% frequency was extracted and used as a template for primer design. Primers were designed with Primer-BLAST (Ye *et al.*, 2012) using 59-61 °C as optimal melting temperature (Tm). Primer sets with the lowest self- and self 3′-complementarity (self-binding affinity) were chosen and tested for specificity. Similarly, BAM files of 25 genotypes with a 50% SC phenotype and 25 genotypes with a 100% SC phenotype used for GBS were also used as a template for PCR-based marker development.

Large-scale genotyping based on High Resolution Melting Analysis

Genotyping was performed with High Resolution Melting Analysis (HRMA) on a LightCycler® 480 System (Roche, Basel, Switzerland) using the same conditions as Do Canto *et al*, (2018). To identify additional individuals carrying a recombination, 2,056 F_2 plants were genotyped with the markers G05_065 and G06_096. The recombinants were further genotyped with 23 newly validated polymorphisms (Supplementary Table 1). The genotypes of G05_065 and G06_096 and of the new markers were combined into a local linkage map using the OneMap software (Margarido *et al.*, 2007). Among the new markers, seven non co-segregating markers were used to obtain marker order and genetic distances by estimating the multipoint likelihood of all the possible orders using the *compare* function. The Kosambi mapping function was used for estimation of mapping distances. The markers within the SC locus were positioned

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

by BLASTN on the Italian ryegrass genome assembly to obtain physical distances and extract information of the annotated protein-coding genes.

Phenotyping of the recombinants for self-compatibility

The individuals showing recombination between G05_065 and G06_096 were phenotyped using *in vitro* pollinations. Eight to ten mature virgin pistils were dissected from florets between 08:00 a.m. and 09:00 a.m. and placed on a petri dish containing a medium consisting of 2% agarose, 10% sucrose, and 100 ppm boric acid. From the same plants, four to five inflorescences with flowering florets were enclosed in paper bags prior to anthesis to collect pollen. Around noon, the paper bags were shaken and the pollen that fell at the bottom of the bag was sprinkled on the stigmas in the petri dishes. After a minimum of three hours, the pollinated stigmas were detached from the ovary with a razor blade, moved to a microscope slide, and submerged in a few drops of a staining solution containing 0.2% aniline blue and 2% K_3PO_4 (Martin, 1959). After covering with a cover slide, the samples were allowed to stain overnight. Pollen growth was detected using a Leica DM12000M UV light microscope (Leica Microsystems, Wetzlar, Germany). Plants were classified as either of the two SC phenotype classes segregating in the population (100% or 50% SC) according to the proportion of pollen grains showing a compatible reaction (Figure 1). In 100% SC plants, all pollen grains are translucent and displayed a clear pollen tube growing towards the style. In the 50% SC plants, nearly half of the pollen grains showed a short, thickened and bright blue pollen tube, with arrested growth upon contact with the stigma.

Identification of candidate genes in the genomic region co-segregating with self-compatibility

MCScanX (Wang *et al.*, 2012) was used to detect collinearity of the genes extracted from Italian ryegrass predicted in the genomic region co-segregating with SC with perennial ryegrass (Byrne *et al.*, 2015), barley (*Hordeum vulgare* L.) (Mascher *et al.*, 2017), rice (Ouyang *et al.*, 2007), and *Arabidopsis thaliana* (The Arabidopsis Genome Initiative, 2000) genome references by aligning their proteomes with BLASTP (Camacho *et al.*, 2009) at default parameters. BLASTP hits were retained, if they had an e value lower than $1e^{-5}$ and if they showed at least 80% and 60% (with perennial ryegrass, rice and barley, respectively) similarity to an Italian ryegrass gene. For *Arabidopsis thaliana*, the gene with highest similarity was chosen. Candidate genes were prioritized primarily based on functional annotation. Additional insights on their involvement in SC/SI were based on organ-specific gene expression data available on Genevestigator® (Zimmermann *et al.*, 2005) for *Arabidopsis thaliana*, barley and rice orthologs. The pollen and

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

stigma transcriptome data reported in Byrne *et al.* (2015) was used to assess if the perennial ryegrass genes present in the SC locus were differentially expressed during compatible/incompatible reactions.

Results

QTL mapping

Phenotypic segregation

Analysis of the *in vitro* pollination data for the 68 individuals from Do Canto et al. (2018) grouped the plants in two clear phenotypic classes (Figure 2a). The first class consisted of 43 plants with a mean compatibility of 90.5 (SD 11). Thirty plants were assigned to the second phenotypic class with a compatibility mean of 49% (SD 11.6). A chi-square test of goodness of fit showed no significant difference between observed and expected phenotypes of a monogenic trait under gametophytic control segregating in a 1:1 ratio in a F_2 generation (p=0.128).

Genetic linkage and QTL mapping

Out of 10,487 SNPs resulting from the GBS variant calling, 473 passed the stringent quality filtering and were used for linkage map construction and QTL mapping. The distribution of the markers across the seven LGs ranged from 28 to 107, with an average of one marker every 1.8 cM and a total map length of 860.1 cM (Figure 2b, Table 1). The SC locus mapped to a single QTL on LG 5 with a maximum LOD value of 7.165 explaining 38.4% of the phenotypic variance (Figure 2c). An effect plot for the most significant marker, m1042, determined the contribution of the allelic states (A, H, B) to the respective phenotypes (Figure 2d). In the homozygous state for the allele contributed by the SC parent (AA), m1042 was associated with an increase of > 40% in self-compatible pollen in an *in vitro* self-pollination assay. The QTL found coincided with the QTL reported in Do Canto *et al.* (2018) as confirmed by the close physical proximity of m1042 and the two flanking markers, reported by Do Canto *et al.* (2018), on the Italian ryegrass reference assembly (data not shown).

Segregation distortion analysis

Analysis of segregation distortion performed using the genetic map employed for QTL mapping identified a region on LG 5 deviating from the 1:2:1 marker segregation expected in a F_2 population. The region showed a 1:1 segregation in two genotypic classes (AA and AB) and overlapped with the region harboring the significant QTL for SC (Figure 2c). Similarly, four co-segregating markers (m1609, m113, m1780, m1440) at the end of LG 2 showed a significantly distorted segregation.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Fine mapping and identification of candidate genes

To further dissect the SC locus, a total of 2,056 F_2 plants were genotyped with the two markers G05_065 and G06_096 known to flank the locus in Do Canto *et al.* (2018). Thirty plants with recombination events between G05_065 and G06_096 were identified, nine of which survived and were further genotyped with the 23 newly developed markers (Supplementary Table 2). These additional genotyping data were used to generate a local linkage map and to estimate the genetic distances between those markers (Figure 3). The locus delimited by G05_065 and G06_096 spanned 1.44 cM, comparable to the genetic distance (1.6 cM) found by Do Canto *et al.* (2018).

Of the nine recombinants utilized for the calculation of genetic distances, five flowered and underwent *in vitro* pollinations, narrowing the region determining SC to 0.26 cM between markers SF28 and SF61. The position of the fine mapped locus was supported by a recombination in the interval between SF3 and SF28 found in P5E10 and by one individual with a recombination delimited by marker SF61 in P11C9 (Figure 3). Within the narrower region, one additional plant (P20H12) carried a recombination between SF3 and SF97, dividing the SC locus in two regions of 0.09 cM and 0.17 cM. However, P20H12 did not produce viable pollen, preventing phenotyping and the assignment of the SC co-segregating locus into one of the two sub-regions.

SF28 and SF61 mapped on two consecutive scaffolds of the Italian ryegrass genome assembly. This orthologous region spanned approximately 3 Mb and contained at least one sequence gap. In the interval, 57 protein-coding genes were predicted (Supplementary Table 3). The marker SF28 overlapped with *Lmu01_3856G0000030*, while SF61 mapped close to *Lmu01_702G0000540*, the most proximal gene inside the SC locus. The alignment of its proteome to the perennial ryegrass gene model revealed the presence of 27 highly similar genes distributed on 17 scaffolds. In barley, the region orthologous to the locus co-segregating with SC was approximately 3.6 Mb in size and contained 27 genes with *HORVU5Hr1G023880* and *HORVU5Hr1G024410* limiting the region on the SF28 and SF61 side, respectively. Fourteen genes had hits to the Italian ryegrass genes of the SC locus (Supplementary Table 3). The ortholog of the SC locus in rice was located between *LOC_Os12g36760* and *LOC_Os12g35670* on chromosome 12 and contained 76 genes, 16 of which had similarity to the genes predicted in Italian ryegrass. These orthologous regions from four different species allowed the compilation of a set of candidate genes based on their annotation, described function and information on organ-specific expression (Table 2).

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Discussion

The identification of loci non-allelic to *S* and *Z* suggests that the grass SI system relies on a network of genes needed for the recognition, signal transmission and rejection or acceptance of self-pollen. How these loci interact with each other and their function in the SI downstream signaling cascade is yet to be determined. However, the SC locus studied here acts epistatically over *S* and *Z*: A pollen grain carrying the SC allele overcomes SI regardless of the composition at *S* and *Z*, and induces a 1:1 segregation in two SC classes: 50% and 100% SC (Arias-Aguirre *et al.*, 2013; Do Canto *et al.*, 2018). Genetically, this segregation pattern is only explained by the action of a single gametophytic gene. Under such a hypothesis, only the heterozygote and homozygote for the SC allele are transmitted to the F_2 generation at a 1:1 ratio, resulting in the 50 and 100% pollen compatibility phenotypes. In our study, we observed the 1:1 phenotypic segregation pattern and, surveying a newly developed genome-wide linkage map, we found that the single QTL on LG 5 was not interacting with other genomic regions, confirming the single-locus inheritance.

Additional evidence was provided by the presence of a major region with distorted segregation in correspondence of the QTL on LG 5. In plants with a functional SI system, the segregation of a SC gene in certain crosses can lead to segregation distortion (Thorogood *et al.*, 2005). In our population, male gametes in the F_1 plants carrying the non-SC allele are not transmitted to the F_2 generation, causing a distorted segregation of markers linked to the SC causal gene. This has also been demonstrated recently by Slatter *et al.* (2020), where the QTL underlying the SC locus on LG 6 in perennial ryegrass overlapped with maximum marker segregation distortion. This has relevant implications for future studies aimed at discovering novel SC sources, as the identification of distorted regions can be used as a strategy to map their position in the genome. Mapping on the basis of segregation distortion represents a valid alternative to the low-throughput phenotyping provided by *in vitro* pollinations assays and it is not dependent on the duration of the flowering period. In addition, as no phenotyping is needed, it allows the screening of larger populations resulting in higher mapping resolution. A less relevant, yet significantly distorted region was also present on LG 2. There are several described biological reasons underlying marker segregation distortion (Xu *et al.*, 1997), thus it is difficult to identify its cause in this case.

After low-resolution QTL mapping, fine mapping resulted in the allocation of the locus to a 0.26 cM region between the markers SF28 and SF61, significantly reducing the extent of the previously described locus (Do Canto *et al.*, 2018) to a sub-cM region spanning ~ 3 Mb on the Italian ryegrass reference assembly. Importantly, there were nine additional markers that co-segregated with SF28, SF97, and SF61 due to lack of recombination events in the nine

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

individuals that we identified. Such markers will be a valuable resource to further dissect genetically the SC locus, should additional F_2 individuals be screened.

In the absence of a contiguous reference genome for perennial ryegrass, we cannot precisely define the genes present in the interval co-segregating with SC. However, given the high genome synteny with Italian ryegrass, barley and rice, it is relevant to assess and compare the genes annotated in the orthologous interval of their reference genomes as a proxy to suggesting candidate genes. On the Italian ryegrass reference assembly, the interval contains 57 protein-coding genes and we foresee to prioritize some of them as candidates responsible for SC based on their predicted function. The region showed conservation compared to other grasses, spanning a similar physical size on barley chromosome 5 and containing 27 genes. In rice, the SC locus ortholog was on chromosome 12. The wealth of functional information available for rice, barley, and *Arabidopsis thaliana* genes and similarity to other plant genes with known function allowed prioritization of genes for future identification of the SC determinant in perennial ryegrass.

Preliminary studies on the molecular determination of SI and SC suggested that Ca^{2+} signalling is involved in the recognition and/or inhibition of incompatible/compatible pollen. In fact, two independent expression studies registered an enrichment of transcripts during the SI response predicted to code for proteins containing calcium-binding domains in perennial ryegrass and sheepgrass (*Leymus chinensis* Trin.) (Yang *et al.*, 2009; Chen *et al.*, 2019). Two genes in the fine mapped SC locus region (*Lmu01_3856G0000420* and *Lmu01_3856G0000430*) encode for proteins with calcium-dependent functions. Their orthologs in *Arabidopsis thaliana*, *At2G18750* and *At1G09170*, encode for a calmodulin-binding protein and a kinesin motor protein with a calponin (calcium-binding protein) domain, respectively. No *At2G18750* expression in floral tissues has been described so far. However, *At1G09170* is co-expressed in mature pollen with several genes involved in mediating microtubule organization during pollen tube polar growth (Schneider and Persson, 2015). Interestingly, their orthologs in perennial ryegrass were reported to be up-regulated during compatible reactions (Byrne *et al.*, 2015), making them strong candidates for causing SC in our population.

Another relevant candidate gene is *Lmu01_3856G0000440*, whose function was annotated as a leucine-rich repeat extensin (LRX). Its ortholog in *Arabidopsis thaliana* (*At2G15880*) shows pollen-specific expression, and the rice ortholog (*LOC_Os12g356710)* shows high expression during pollen germination. Leucine-rich repeat extensins are cell wall proteins and part of a signalling system that transfers extracellular signals from the cell wall to the

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

cytoplasm in order to regulate cell growth in vegetative tissue and in the pollen tube (Herger *et al.*, 2019). In several recent studies in *Arabidopsis thaliana*, *lrx* mutants show severe defects in pollen germination and pollen tube growth (Fabrice *et al.*, 2018; Sede *et al.*, 2018; Wang *et al.*, 2018). However, manipulating Ca^{2+} availability alleviates these defects, suggesting that LRX proteins can regulate Ca^{2+}-related processes. In carrying out their regulatory function, LRX proteins interact with short cysteine-rich peptides known as RALFs (rapid alkanization factors). Interestingly, a gene encoding a RALF-like protein (*Lmu01_702G0000440)* is also present in the SC co-segregating region. RALFs induce a rapid alkalinization of extracellular space by increasing the cytoplasmic Ca^{2+} concentration that leads to calcium-dependent signalling events. They control several plant stress response mechanisms, growth, and development (Murphy and De Smet, 2014), and pollen-specific RALFs regulating pollen tube growth are described in tomato (*Solanum lycopersicum* L.) (Covey *et al.*, 2010) and *Arabidopsis thaliana* (Mecchia *et al.*, 2017; Moussu *et al.*, 2020). Both LRX and RALF-like-encoding genes are strong candidates for causing the inhibition of self-pollen tube. They could be a fundamental component of the process that inhibits self-pollen tube by transducing SI signals to activate the downstream pathway that modulates the Ca^{2+} gradient necessary for pollen tube growth. A disruptive mutation in such genes could interrupt the inhibitory mechanism, leading to SC.

Three other genes located in the narrower interval encode for F-box proteins (*Lmu01_3856G0000060*, *Lmu01_3856G0000100*, *Lmu01_3856G0000370*). *F-box* genes constitute one of the largest gene families in plants, involved in degradation of cellular proteins (Lechner *et al.*, 2006). F-box proteins form a subunit of the Skp1-cullin-F-box (SCF) complex that confers specificity for a target substrate to be degraded and are involved in important biological processes such as embryogenesis, floral development, plant growth and development, biotic and abiotic stress, hormonal responses and senescence, among others (Gupta *et al.*, 2015). *S*-locus *F-box* genes (*SFB* or *SLF*, depending on the family) are known as the male determinants of *S*-RNase based gametophytic SI systems (i.e. Rosaceae, Solanaceae, Scrophulariaceae, Rubiaceae). In these systems, the female determinant encodes for transmembrane *S*-RNase in the stigma with high cytotoxic activity for the pollen tube. In a non-self-pollination event, *SFBs* identify and degrade non-self *S*-RNases allowing fertilisation to occur unimpeded via the ubiquitin-26S proteasome pathway. Whereas in a self-pollination, self *S*-RNases evade degradation and exert cytotoxicity inside the pollen tube to inhibit its growth (Hua and Kao, 2008; McClure *et al.*, 2011). An *F-box* gene could also play a role in the complex downstream pathway that leads to rejection of self-pollen in grasses and a disruptive mutation might lead SC by halting its inhibiting function.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

The genes discussed here are plausible candidates to cause SC, selected on the basis of their annotation and function described in previous studies. However, it is not possible to exclude the remaining genes in the genome region co-segregating with SC as potential candidates. Only additional, complementary analyses will allow the identification and confirmation of the molecular determinant of SC. Among these, targeted sequencing and sequence comparison of such genes in our SC population and diverse SI genotypes can contribute to reducing the number of candidate genes. Furthermore, techniques such as TILLING (Targeting Induced Local Lesions in Genomes) offer opportunities to screen for point mutations in candidate genes in populations mutagenized by chemical treatment and evaluate if those lead to a SC phenotype (Manzanares *et al.*, 2016b). Also, the recent advances in genome editing through CRISPR/Cas9 and its implementation in ryegrass (Zhang *et al.*, 2020) makes it an attractive alternative to achieve targeted gene inactivation of SC candidate genes to unequivocally determine their function.

Conclusions

In this study, we provide additional evidence of a major QTL on LG 5 being solely responsible for SC in a perennial ryegrass population and demonstrate that segregation distortion can be an efficient strategy to identify loci conferring SC. Additionally, we developed new markers and fine mapped the previously described region from 1.6 cM down to 0.26 cM. Genome synteny to other grass species was used to determine a 3 Mb ortholog region in Italian ryegrass and to find the gene content in the region co-segregating with SC. From a total of 57 genes, seven encode for functionally relevant proteins and will be prioritized for validation as genetic determinants of SC. Furthermore, the genetic markers developed for fine mapping are tightly linked to the SC gene and can readily be converted into high throughput genotyping assays. They provide a convenient molecular tool to transfer SC to elite germplasm for the development of inbred lines-based breeding strategies.

Acknowledgments

The authors wish to acknowledge Ingrid Stoffel-Studer and Dr. Timothy Sykes from the Molecular Plant Breeding group at ETH Zurich for assisting in DNA extraction and GBS library preparation, the Functional Genomic Center Zurich for sequencing service, the Zurich-Basel Plant Science Center for professional management of the PlantHUB programme and Nic Boerboom from DSV Zaden Nederland B.V for the helpful comments on the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Electronic supplementary material

ESM_1.xlsx: Genotype and mapping data from 68 F2 individuals segregating for self-compatibility (SC), formatted as input for the R/qtl package. The first three rows indicate the marker names, their respective linkage group and genetic position in centimorgan. The column on the left indicates the averaged SC phenotypes from Do Canto et al., (2018) for each genotype.

ESM_2.xlsx: Marker name, type of polymorphism, primer sequences and source of the markers used for High Resolution Melting Analysis.

ESM_3.xlsx: High Resolution Melting Analysis of the markers G05_065 and G06_096 for the 2,056 F2 perennial ryegrass individuals screened for the identification of recombinants (A= homozygous for self-compatibility allele, B= homozygous for self-incompatibility allele, H=heterozygous)

ESM_4.xlsx: High Resolution Melting Analysis of the markers used for fine mapping of the nine F2 individuals with at least one recombination event between G05_065 and G06_096 (A= homozygous for the self-compatibility allele, B=homozygous for the self-incompatibility allele, H=heterozygous).

ESM_5.xlsx: Genes predicted in the Italian ryegrass orthologous region of the fine mapped self-compatibility locus and their orthologs in perennial ryegrass, barley, rice and *Arabidopsis thaliana* as resulting from BLASTP analysis.

Figure captions

Fig 1 The two self-compatibility (SC) phenotypic classes segregating in the F_2 population as visible in an *in vitro* pollination assay. a) In genotypes showing 100% SC, all pollen grains form a pollen tube that is able to elongate within the stigma reaching the ovary. b) In genotypes showing 50% SC, only half of the pollen grains penetrate the stigma and elongate a pollen tube (indicated by the red arrows). For the remaining half, pollen tube growth is aborted upon contact with the stigma and followed by a rapid deposition of callose at the pollen tube tip (white arrows). Bars: 400 µm.

Fig 2 QTL analysis for self-compatibility (SC). a) The bi-modal phenotypic segregation for the SC phenotype of the 68 individuals used for the generation of the linkage map, QTL analysis and segregation distortion analysis. Data distributions and means (indicated by the vertical dotted lines) are displayed in dark grey for individuals with a 50% SC phenotype and light grey for the 100% self-compatible individuals. Small solid vertical lines above the horizontal

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

axis indicate single data points. b) Genetic linkage map displaying the distribution of the 473 markers across the seven linkage groups of perennial ryegrass used for the QTL analysis. c) QTL analysis for SC (black line) and analysis of segregation distortion (red dotted line). Significance thresholds are indicated by the black and red horizontal lines for the QTL analysis and segregation distortion analysis, respectively. d) Marker effect plot for the most significant marker. AA represents the homozygous genotype for the SC phenotype, while AB and BB represent the heterozygous and homozygous genotypes for the self-incompatible phenotype, respectively. Plots show mean and standard deviation for the SC phenotype.

Fig 3 Fine mapping of the self-compatibility (SC) locus on linkage group 5. The top bar shows the genetic order and distances in centimorgan (cM) of the nine informative markers used for the calculation of genetic distances. For SF68, SF28 and SF3, the number of co-segregating markers is indicated in brackets. Genotype at each marker (vertical bars) and haplotype configuration (horizontal bars) for the nine individuals screened with G05_065 and G06_096 and the newly developed markers are shown. The column on the right-side reports the phenotypes, when available. The dashed vertical lines at SF28 and SF61 limit the SC locus.

Tables

Table 1: Descriptive statistics of the genotyping-by-sequencing-based genetic linkage map developed from 68 F_2 individuals of perennial ryegrass segregating for self-compatibility. Length, average and maximum spacing for each linkage group are indicated in centimorgan (cM).

Linkage group	Number of markers	Length (cM)	Average spacing (cM)	Maximum spacing (cM)
1	28	44.7	1.7	5.3
2	89	135.3	1.5	11.8
3	79	183.3	2.3	22.7
4	107	237.8	2.2	44.2
5	43	50	1.2	7.5
6	52	85	1.7	15.9
7	75	124	1.7	11.5
Overall	**473**	**860.1**	**1.8**	**44.2**

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Table 2: List of the seven prioritized candidate genes co-segregating with self-compatibility, their annotation and orthologs in perennial ryegrass (*Lolium perenne* L.), rice (*Oryza sativa* L.), barley (*Hordeum vulgare* L.) and *Arabidopsis thaliana*.

Italian ryegrass	Perennial ryegrass	Rice	Barley	*Arabidopsis thaliana*	Functional annotation
Lmu01_3856G0000060	*maker-scaffold_1442|ref0021156-exonerate_est2genome-gene-0.1-mRNA-1*	-	-	-	Protein of unknown function (DUF295), F-box-like protein
Lmu01_3856G0000100	*maker-scaffold_10104|ref0010435-exonerate_est2genome-gene-0.3-mRNA-1*	*HORVU5Hr1G024060.1*	*LOC_Os12g36670.1*	*AT5G27920*	F-box family protein
Lmu01_3856G0000370	*maker-scaffold_2321|ref0036192-exonerate_est2genome-gene-0.1-mRNA-1*	-	-	*AT1G13570*	F-box/RNI-like superfamily protein
Lmu01_3856G0000420	*maker-scaffold_792|ref0008797-exonerate_est2genome-gene-0.1-mRNA-1*	*HORVU5Hr1G024400.1*	-	*AT2G18750*	Calmodulin-binding protein
Lmu01_3856G0000430	*maker-scaffold_792|ref0008797-exonerate_est2genome-gene-0.2-mRNA-1*	*HORVU5Hr1G024390.9*	*LOC_Os12g36100.1*	*AT1G09170*	P-loop nucleoside triphosphate hydrolases superfamily protein with Calponin Homology domain
Lmu01_3856G0000440	-	*HORVU5Hr1G024380.4*	*LOC_Os12g35710.1*	*AT2G15880*	Leucine-rich repeat extensin (LRX)
Lmu01_702G0000440	-	*HORVU5Hr1G024410*	*LOC_Os12g35670*	*AT4G15800*	Rapid ALkalinization Factor (RALF) protein

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Declarations

Funding

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme Marie Skłodowska-Curie grant agreement No 722338 - PlantHUB and by the Swiss National Science Foundation (SNSF Professorship PP00P2_138983).

Conflict of interest

The authors declare no conflict of interest to declare that are relevant to the content of this article.

Availability of data and material

All data analyzed during this study are included in this published article and its supplementary information files. All raw sequencing data (deplexed GBS and pooled sequencing) are available at the Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) under BioProject: PRJNA724991.

Author contribution statement

Study conception: CC, CM, MK, TL and BS. Generation of plant material: JDC, TL. GBS library preparation, sequencing and bioinformatics: CC, CM and SY. Genetic map development, QTL mapping and analysis of segregation distortion: CC. Marker development, genotyping and marker data analysis: CC, CM, SY and DC. Comparative genomic analysis: CC and DC. Funding acquisition: MK, BS. Manuscript writing: CC, CM, DC and BS. All authors read and approved the final version of the manuscript.

Literature cited

Arias-Aguirre A, Studer B, Do Canto J, et al (2013) Mapping a New Source of Self-fertility in Perennial Ryegrass (*Lolium perenne* L.) . Plant Breed Biotechnol 1:385–395. https://doi.org/10.9787/pbb.2013.1.4.385

Begheyn RF, Yates SA, Sykes T, Studer B (2018) Genetic Loci Governing androgenic capacity in perennial ryegrass (*Lolium perenne* L.). G3 Genes, Genomes, Genet 8:1897–1908. https://doi.org/10.1534/g3.117.300550

Bian XY, Friedrich A, Bai JR, et al (2004) High-resolution mapping of the S and Z loci of *Phalaris coerulescens*. Genome 47:918–930. https://doi.org/10.1139/G04-017

Broman KW, Wu H, Sen Ś, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

19:889–890. https://doi.org/10.1093/bioinformatics/btg112

Byrne SL, Nagy I, Pfeifer M, et al (2015) A synteny-based draft genome sequence of the forage grass *Lolium perenne*. Plant J 84:816–826. https://doi.org/10.1111/tpj.13037

Camacho C, Coulouris G, Avagyan V, et al (2009) BLAST+: Architecture and applications. BMC Bioinformatics 10:1–9. https://doi.org/10.1186/1471-2105-10-421

Capstaff NM, Miller AJ (2018) Improving the yield and nutritional quality of forage crops. Front Plant Sci 9:1–18. https://doi.org/10.3389/fpls.2018.00535

Chen S, Jia J, Cheng L, et al (2019) Transcriptomic analysis reveals a comprehensive calcium- and phytohormone-dominated signaling response in leymus chinensis self-incompatibility. Int J Mol Sci 20:1–24. https://doi.org/10.3390/ijms20092356

Copetti D, Yates S, Vogt MM, et al (2021). Evidence for high intergenic sequence variation in heterozygous Italian ryegrass (*Lolium multiflorum* Lam.) genome revealed by a high-quality draft diploid genome assembly. bioRxiv. https://doi.org/10.1101/2021.05.05.442707

Cornish MA, Hayward MD, Lawrence MJ (1979) Self-incompatibility in ryegrass. Heredity 43:95–106. https://doi.org/10.1038/hdy.1983.5

Covey PA, Subbaiah CC, Parsons RL, et al (2010) A pollen-specific RALF from tomato that regulates pollen tube elongation. Plant Physiol 153:703–715. https://doi.org/10.1104/pp.110.155457

Do Canto J, Studer B, Frei U, Lübberstedt T (2018) Fine mapping a self-fertility locus in perennial ryegrass. Theor Appl Genet 131:817–827. https://doi.org/10.1007/s00122-017-3038-6

Do Canto J, Studer B, Lubberstedt T (2016) Overcoming self-incompatibility in grasses: a pathway to hybrid breeding. Theor Appl Genet 129:1815–1829. https://doi.org/10.1007/s00122-016-2775-2

Egorova IA, Peneva TI, Baranova OA, Voylokov A V. (2000) Analysis of linkage between biochemical and morphological markers of rye chromosomes 1R, 2R, and 5R and mutations of self-fertility at the main incompatibility loci. Russ J Genet 36:1423–1430. https://doi.org/10.1023/A:1009075408921

Fabrice TN, Vogler H, Draeger C, et al (2018) LRX proteins play a crucial role in pollen grain and pollen tube cell wall development. Plant Physiol 176:1981–1992. https://doi.org/10.1104/pp.17.01374

Gupta S, Garg V, Kant C, Bhatia S (2015) Genome-wide survey and expression analysis of F-box genes in chickpea. BMC Genomics 16:1–15. https://doi.org/10.1186/s12864-015-1293-y

Hackauf B, Wehling P (2005) Approaching the self-incompatibility locus Z in rye (Secale cereale L.) via comparative genetics. Theor Appl Genet 110:832–845. https://doi.org/10.1007/s00122-004-1869-4

Hayman D (1956) The Genetical Control of Incompatibility in *Phalaris coerulescens* Desf. Aust J Biol Sci 9:321. https://doi.org/10.1071/bi9560321

Herger A, Dünser K, Kleine-Vehn J, Ringli C (2019) Leucine-Rich Repeat Extensin Proteins and Their Role in Cell Wall Sensing. Curr Biol 29(17):R851–R858. https://doi.org/10.1016/j.cub.2019.07.039

Herridge RP, Macknight RC, Brownfield LR (2020) Prospects for F1 hybrid production in ryegrass. New Zeal J Agric Res 63(3):405-415. https://doi.org/10.1080/00288233.2018.1559867

Hua Z, Kao TH (2008) Identification of major lysine residues of S3-RNase of *Petunia inflata* involved in ubiquitin-26S proteasome-mediated degradation in vitro. Plant J 54:1094–1104. https://doi.org/10.1111/j.1365-313X.2008.03487.x

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815. https://doi.org/10.1038/35048692

Robinson JT, Thorvaldsdóttir H, Winckler W, et al (2011) Integrative genomics viewer. Nat Biotechnol 29:24–26. https://doi.org/10.1038/nbt.1754

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Jansky SH, Charkowski AO, Douches DS, et al (2016) Reinventing potato as a diploid inbred line-based crop. Crop Sci 56:1412–1422. https://doi.org/10.2135/cropsci2015.12.0740

Jensen LB, Andersen JR, Frei U, et al (2005) QTL mapping of vernalization response in perennial ryegrass (*Lolium perenne* L.) reveals co-location with an orthologue of wheat *VRN1*. Theor Appl Genet 110:527–536. https://doi.org/10.1007/s00122-004-1865-8

Kakeda K, Ibuki T, Suzuki J, et al (2008) Molecular and genetic characterization of the S locus in *Hordeum bulbosum* L., a wild self-incompatible species related to cultivated barley. Mol Genet Genomics 280:509–519. https://doi.org/10.1007/s00438-008-0383-9

Klaas M, Yang B, Bosch M, et al (2011) Progress towards elucidating the mechanisms of self-incompatibility in the grasses: Further insights from studies in *Lolium*. Ann Bot 108:677–685. https://doi.org/10.1093/aob/mcr186

Kosambi DD (1943) The estimation of map distances from recombination values. Ann Hum Genet 12(1):172–175. https://doi.org/10.1111/j.1469-1809.1943.tb02321.x

Laidig F, Piepho HP, Drobek T, Meyer U (2014) Genetic and non-genetic long-term trends of 12 different crops in German official variety performance trials and on-farm yield trends. Theor Appl Genet 127:2599–2617. https://doi.org/10.1007/s00122-014-2402-z

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923

Lechner E, Achard P, Vansiri A, et al (2006) F-box proteins everywhere. Curr Opin Plant Biol 9:631–638. https://doi.org/10.1016/j.pbi.2006.09.003

Li H, Handsaker B, Wysoker A, et al (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lundqvist, A (1956). Self-incompatibility in rye. I. Genetic control in the diploid. Hereditas 44, 193–256. doi: 10.1111/j.1601-5223.1958.tb03481.x

Manzanares C, Barth S, Thorogood D, et al (2016a) A gene encoding a DUF247 domain protein cosegregates with the *S* self-incompatibility locus in perennial ryegrass. Mol Biol Evol 33:870–884. https://doi.org/10.1093/molbev/msv335

Manzanares C, Yates S, Ruckle M, et al (2016b) TILLING in forage grasses for gene discovery and breeding improvement. N Biotechnol 33:594–603. https://doi.org/10.1016/j.nbt.2016.02.009

Margarido GRA, Souza AP, Garcia AAF (2007) OneMap: Software for genetic mapping in outcrossing species. Hereditas 144:78–79. https://doi.org/10.1111/j.2007.0018-0661.02000.x

Martin FW (1959) Staining and observing pollen tubes in the style by means of fluorescence. Biotech Histochem 34:125–128. https://doi.org/10.3109/10520295909114663

Mascher M, Gundlach H, Himmelbach A, et al (2017) A chromosome conformation capture ordered sequence of the barley genome. Nature 544:427–433. https://doi.org/10.1038/nature22043

McClure B, Cruz-García F, Romero C (2011) Compatibility and incompatibility in S-RNase-based systems. Ann Bot 108:647–658. https://doi.org/10.1093/aob/mcr179

Mecchia MA, Santos-Fernandez G, Duss NN, et al (2017) RALF4/19 peptides interact with LRX proteins to control pollen tube growth in Arabidopsis. Science (80-) 358:1600–1603. https://doi.org/10.1126/science.aao5467

Moussu S, Broyart C, Santos-Fernandez G, et al (2020) Structural basis for recognition of RALF peptides by LRX proteins during pollen tube growth. Proc Natl Acad Sci 117:7494–7503. https://doi.org/10.1073/pnas.2000100117

Murphy E, De Smet I (2014) Understanding the RALF family: A tale of many species. Trends Plant Sci 19:664–671. https://doi.org/10.1016/j.tplants.2014.06.005

Ouyang S, Zhu W, Hamilton J, et al (2007) The TIGR Rice Genome Annotation Resource: Improvements and new

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.29.437398; this version posted May 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

features. Nucleic Acids Res 35:8–11. https://doi.org/10.1093/nar/gkl976

R Core Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Schneider R, Persson S (2015) Connecting two arrays: The emerging role of actin-microtubule cross-linking motor proteins. Front Plant Sci 6:1–7. https://doi.org/10.3389/fpls.2015.00415

Sede AR, Borassi C, Wengier DL, et al (2018) Arabidopsis pollen extensins LRX are required for cell wall integrity during pollen tube growth. FEBS Lett 592:233–243. https://doi.org/10.1002/1873-3468.12947

Shinozuka H, Cogan NOI, Smith KF, et al (2010) Fine-scale comparative genetic and physical mapping supports map-based cloning strategies for the self-incompatibility loci of perennial ryegrass (*Lolium perenne* L.). Plant Mol Biol 72:343–355. https://doi.org/10.1007/s11103-009-9574-y

Slatter LM, Barth S, Manzanares C, et al (2020) A new genetic locus for self-compatibility in the outcrossing grass species perennial ryegrass (*Lolium perenne*). Ann Bot XX:1–8. https://doi.org/10.1093/aob/mcaa140

Studer B, Asp T, Frei U, Hentrup S, Meally H et al (2008) Expressed sequence tag-derived microsatellite markers in perennial ryegrass (*Lolium perenne* L.). Mol Breed 21:533–548. https://doi.org/10.1007/s11032-007-9148-0

Thorogood D, Armstead I (2002) Self-incompatibility in ryegrass 12. Genotyping and mapping the *S* and *Z* loci of *Lolium perenne* L. Heredity 385–390. https://doi.org/10.1038/sj/hdy/6800071

Thorogood D, Armstead IP, Turner LB, et al (2005) Identification and mode of action of self-compatibility loci in Lolium perenne L. Heredity 94:356–363. https://doi.org/10.1038/sj.hdy.6800582

Voylokov A V, Korzun V, Börner A (1998) Mapping of three self-fertility mutations in rye (*Secale cereale* L.) using RFLP, isozyme and morphological markers. Theor Appl Genet 97:147–153. https://doi.org/10.1007/s001220050879

Voylokov A V, Fuong FT, Smirnov VG (1993) Genetic studies of self-fertility in rye (*Secale cereale* L.). 1. The identification of genotypes of self-fertile lines for the *Sf alleles* of self-incompatibility genes. Theor Appl Genet 87:616–618. https://doi.org/10.1007/BF00221887

Wang X, Wang K, Yin G, et al (2018) Pollen-expressed leucine-rich repeat extensins are essential for pollen germination and growth. Plant Physiol 176:1993–2006. https://doi.org/10.1104/pp.17.01241

Wang Y, Tang H, Debarry JD, et al (2012) MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40:1–14. https://doi.org/10.1093/nar/gkr1293

Xu Y, Zhu L, Xiao J, et al (1997) Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). Mol Gen Genet 253:535–545. https://doi.org/10.1007/s004380050355

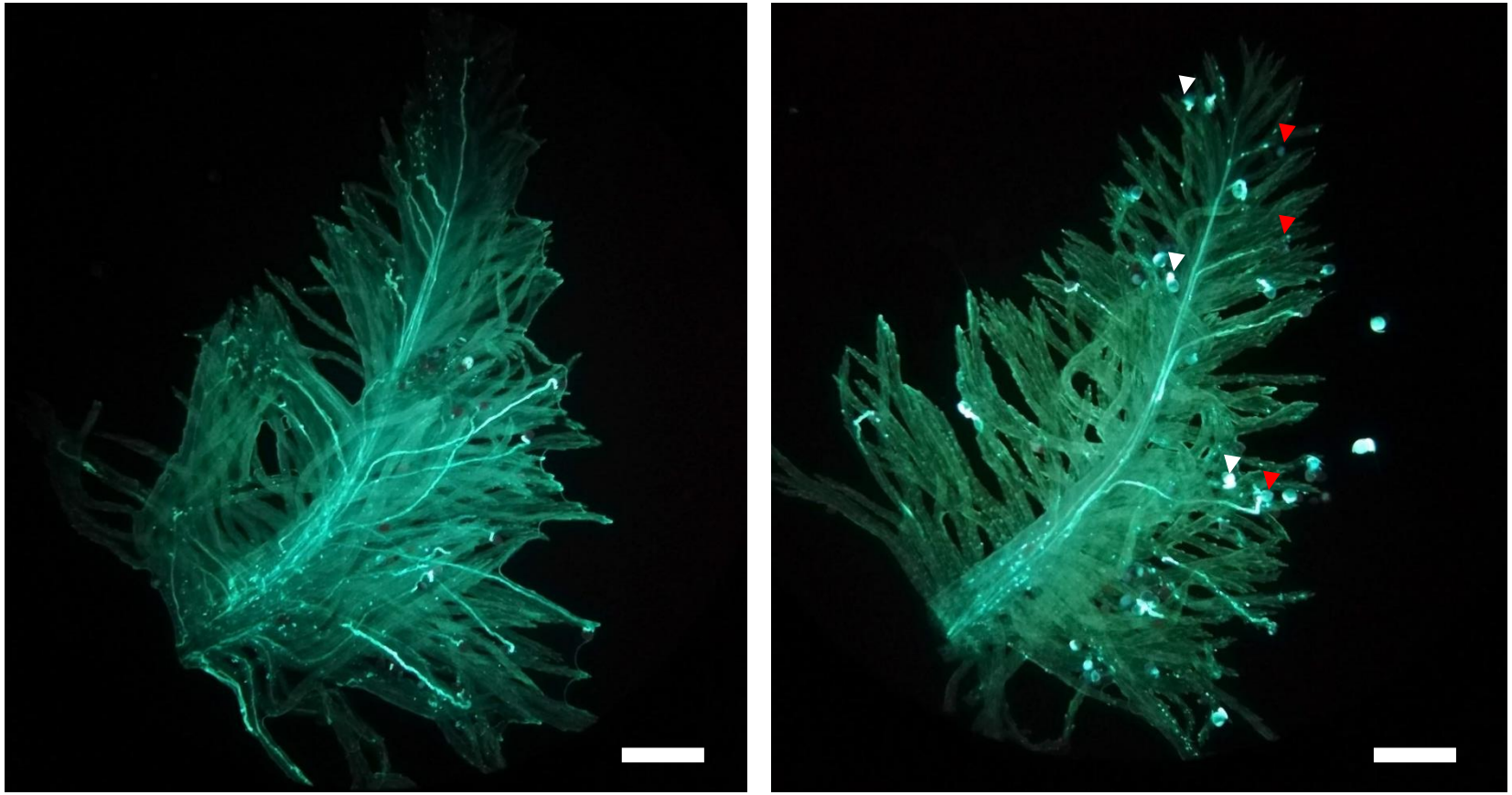
Yang B, Thorogood D, Armstead IP, et al (2009) Identification of genes expressed during the self-incompatibility response in perennial ryegrass (*Lolium perenne* L.). Plant Mol Biol 70:709–723. https://doi.org/10.1007/s11103-009-9501-2

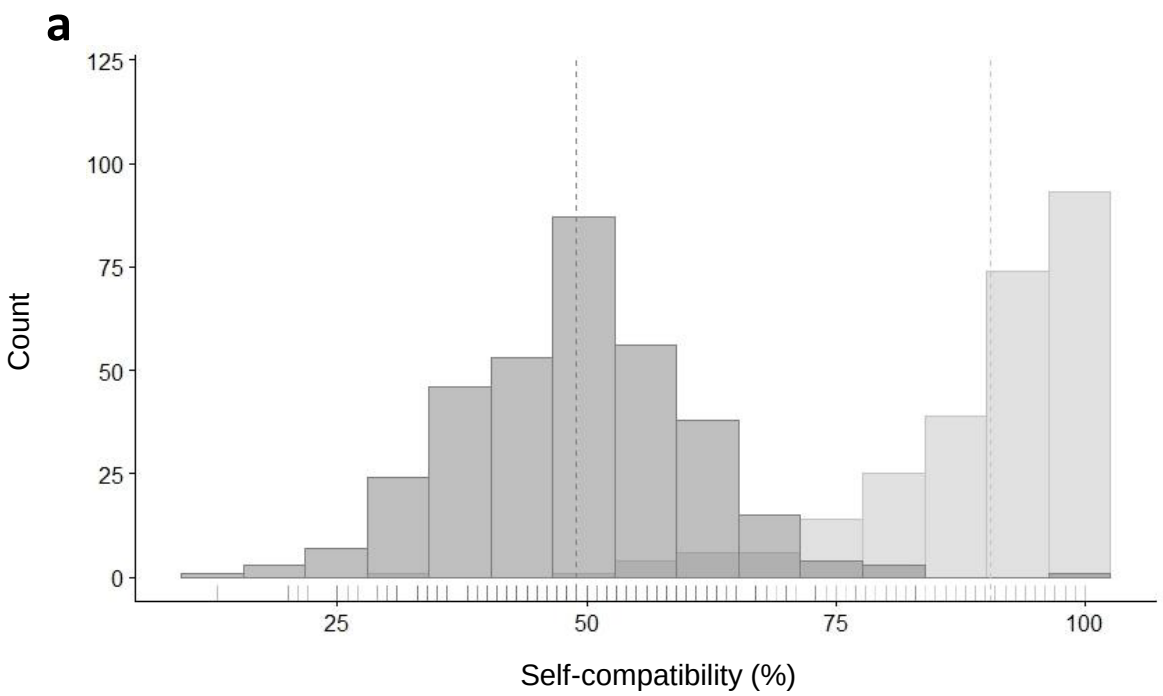
Ye J, Coulouris G, Zaretskaya I, et al (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13:134. https://doi.org/10.1186/1471-2105-13-134

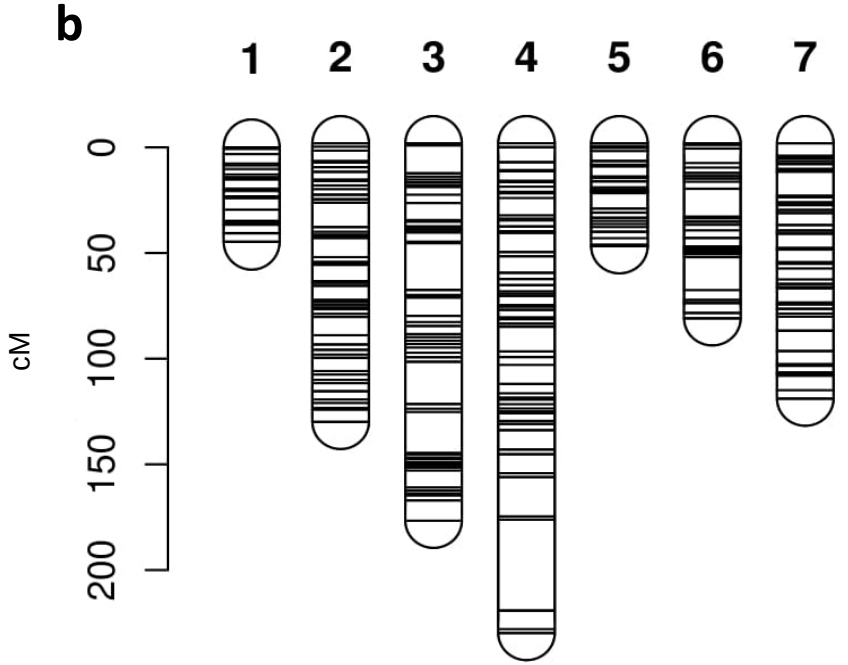
Young ND, Tanksley SD (1989) Restriction fragment length polymorphism maps and the concept of graphical genotypes. Theor Appl Genet 77:95–101. https://doi.org/10.1007/BF00292322

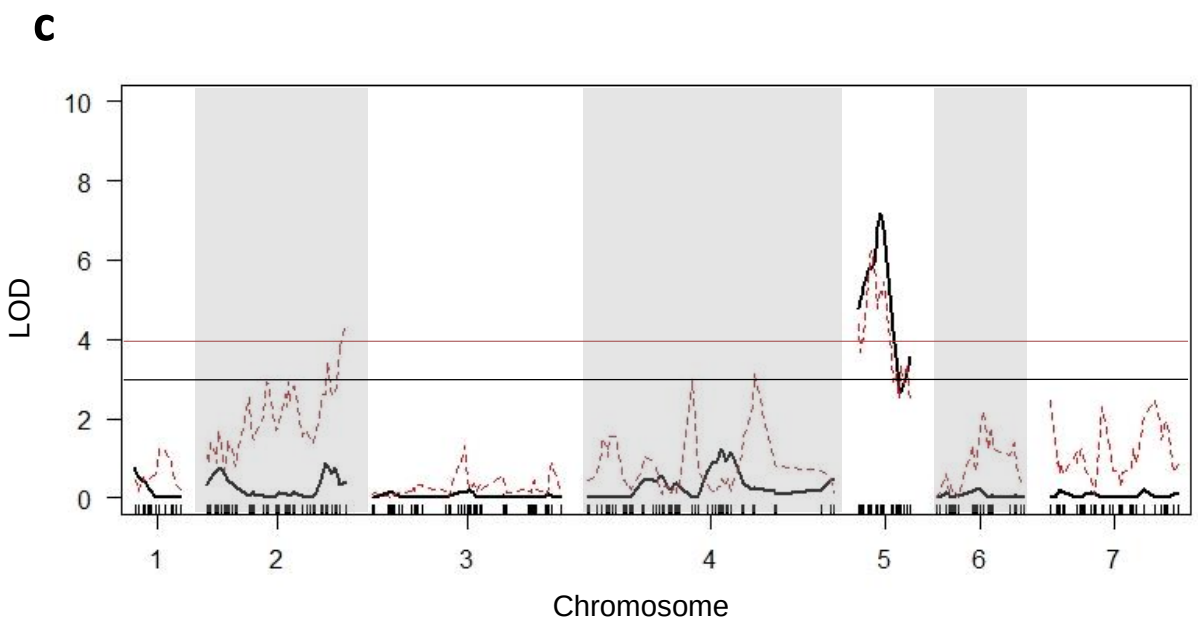
Zhang Y, Ran Y, Nagy I, et al (2020) Targeted mutagenesis in ryegrass (*Lolium* spp.) using the CRISPR/Cas9 system. Plant Biotechnol J 18:1854–1856. https://doi.org/10.1111/pbi.13359

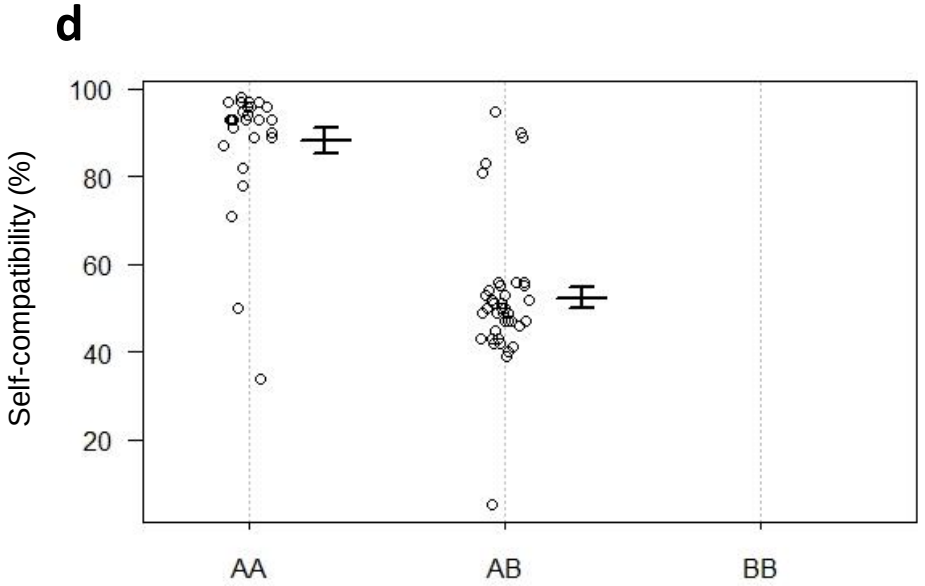
Zimmermann P, Hennig L, Gruissem W (2005) Gene-expression analysis and network discovery using Genevestigator. Trends Plant Sci 10:407–409. https://doi.org/10.1016/j.tplants.2005.07.003

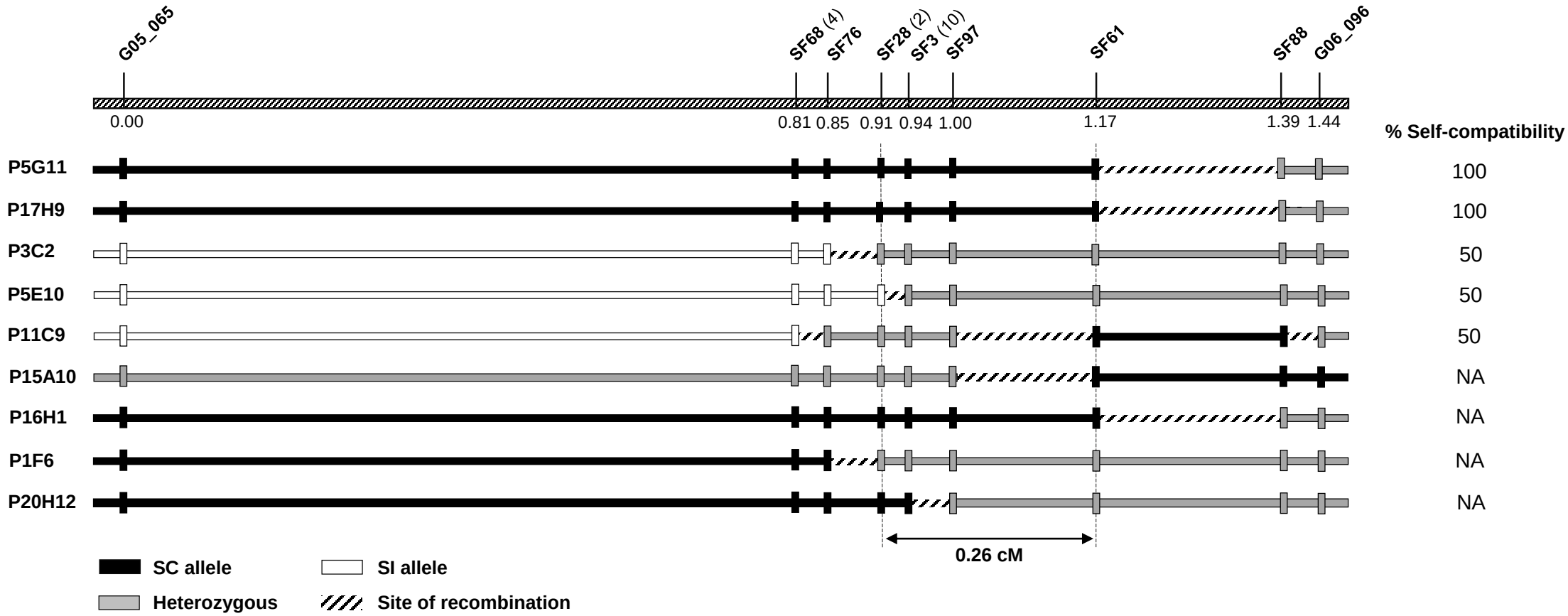



a
Count
125
100
75
50
25
0
25
50
75
100
Self-compatibility (%)


b
1
2
3
4
5
6
7
cM
0
50
100
150
200


c
LOD
10
8
6
4
2
0
1
2
3
4
5
6
7
Chromosome


d
Self-compatibility (%)
100
80
60
40
20
AA
AB
BB


G05_065
SF68 (4)
SF76
SF28 (2)
SF3 (10)
SF97
SF61
SF88
G06_096
0.00
0.81 0.85 0.91 0.94 1.00
1.17
1.39 1.44
% Self-compatibility
P5G11 100
P17H9 100
P3C2 50
P5E10 50
P11C9 50
P15A10 NA
P16H1 NA
P1F6 NA
P20H12 NA
0.26 cM
SC allele
SI allele
Heterozygous
Site of recombination