Determining self-incompatibility genotypes in Belgian wild cherries

B. DE CUYPER,* T. SONNEVELD†‡ and K.R. TOBUTT†

**Institute for Forestry and Game Management, Gaverstraat 4, B-9500 Geraardsbergen, Belgium, †East Malling Research, New Road, East Malling, Kent ME19 6BJ, United Kingdom, ‡University of Nottingham, Plant Science Division, School of Biosciences, Sutton Bonington Campus, Loughborough, Leicestershire LE12 5RD, United Kingdom*

Keywords: *Prunus avium*, wild cherry, incompatibility genotypes, consensus primers, allele-specific primers, seed orchards

Correspondence address: Bart De Cuyper, Institute for Forestry and Game Management, Gaverstraat 4, B-9500 Geraardsbergen, Belgium. Fax: 0032 54 436160; E-mail: bart.decuyper@lin.vlaanderen.be

Running title: Incompatibility genotypes in wild cherry

Abstract

The self-incompatibility (*S*) genotypes of a collection of 65 Belgian accessions of wild cherry, selected within two populations and planted in a seed orchard, were determined using PCR-methods. Initially, DNA extracts were amplified with consensus primers that amplify across the second intron of the *S*-ribonuclease gene which shows considerable length polymorphism. The provisional genotypes deduced were checked with the appropriate allele-specific primers for the known alleles S_1 to S_{16}. Putative new alleles were subjected to PCR with consensus primers amplifying across the first intron. Six new alleles, S_{17} to S_{22}, were thus indicated on the basis of the estimated lengths of the first and second intron PCR-products. Examples of these alleles were partially sequenced and were indeed mutually distinct and different from the known alleles. The incompatibility genotypes of all 65 accessions were determined and one triploid individual was found. Seventeen alleles were detected in all. Allele frequencies differed between samples and the expected total number of alleles in the underlying populations was estimated. The wild cherry populations differed significantly with respect to allelic frequencies from sweet cherry cultivars; alleles S_4 and S_5, which are moderately frequent in sweet cherry, were absent from the wild cherry accessions. The knowledge of the *S*-genotypes will be useful for studying the gene flow within the seed orchard and these approaches should also be informative in wild populations.

Introduction

The cherry, *Prunus avium* L., occurs naturally in much of temperate Eurasia, showing a vast distribution area stretching from Ireland to the Caspian Sea and from southern Scandinavia to North Africa (Schütt *et al.* 1995). It is cultivated for its fruit and for its high grade timber which is mainly used for panelling and cabinet-making. For the purposes of this paper we will refer to fruit forms as sweet cherry, and to timber and wild forms as wild cherry.

P. avium is allogamous. This trait is governed by a single multi-allelic *S*-locus with gametophytic action (Crane and Lawrence 1929) which controls self-incompatibility, and cross-incompatibility between some individuals. The *S*-locus codes for stylar ribonucleases (Bošković and Tobutt 1996, Bošković *et al*. 1997). When the *S*-allele of the pollen grain matches one of the *S*-alleles of the style these ribonucleases inhibit the growth of the pollen tube in the style. Various *S*-alleles have been cloned and shown to contain several highly conserved regions and allele-specific and consensus primers have been designed to amplify across each of the two introns (Sonneveld *et al.* 2001, Sonneveld *et al*. 2003). The introns, in particular the second intron, show considerable length polymorphism, allowing most of the alleles to be distinguished according to the size of the PCR-amplification products.

Determination of incompatibility genotypes has been restricted largely to sweet cherries, for which compatibility data are important to pomologists recommending appropriate pollinators and to breeders designing crosses. Matthews and Dow (1969) collated conclusions from crossing tests at the John Innes Institute and elsewhere to assign genotypes to some 145 cultivars. Genotypes published recently, including those determined by separating and characterising stylar ribonucleases (Bošković and Tobutt 2001) or by amplification with allele-specific primers (Sonneveld *et al.* 2001), have been compiled into a 'Harmonization Table' (Tobutt *et al*. 2001). In this, some 165 cultivars were assigned to 22 incompatibility

66 groups, besides a set of ten cultivars of unique genotype, involving combinations of the
67 alleles S_1 to S_7, S_9 to S_{14} and S_x. Sonneveld *et al.* (2003) have genotyped eight more self-
68 incompatible cultivars, have 'retired' S_{11}, labelled S_x as S_{16} and proposed two additional
69 groups to be included in a revised Harmonization Table (Tobutt *et al.* in press).
70 Just three items which may be regarded as wild cherry rather than sweet, are in this list.
71 These include the timber selection Orleans 171 ($S_{10}S_{11} = S_7S_{10}$), the mazzard rootstock F1/3
72 (S_1S_2) and the rootstock 'Charger' (S_1S_7), which is a seedling of F1/3. The only substantial
73 study of incompatibility alleles in wild cherry is that of Berger (1963). In a partial diallel
74 cross of ten trees he found nine to be intercompatible and deduced the presence of a
75 minimum of ten alleles on the questionable assumption that all the compatible crosses were
76 semi-compatible and had the same allele in common.
77 In recent years there has been an increasing interest in Europe in wild cherry as a potentially
78 valuable hardwood timber tree for planting in woodlands and as an alternative to poplar for
79 the afforestation of abandoned farmland. This policy generates a strong demand for high
80 quality forest reproductive material and thus for the selection and/or creation of genetically
81 superior basic material. As wild cherry generally occurs as individual trees or small clusters,
82 often clonal, scattered throughout mixed forests, the potential for selecting natural seed stands
83 is limited. Consequently, most of the European wild cherry breeding programmes aim to
84 establish seed orchards. Besides the assessment of the combining ability for adaptive traits
85 such as vigour, morphology, phenology and disease resistance, the selection of clones for the
86 seed orchards should pay attention to *S*-genotypes.
87 *P. avium* is entomophilous; bumble bees (*Bombus* spp.) in particular act as pollen vectors
88 (Heinrich 1976). As bumble bees tend to forage on small groups of trees, leaving a scent
89 mark on the flowers they visit, a patch-like pollination pattern is observed in populations of
90 wild cherry (Goulson *et al.* 1998, S. Mariette, personal communication). Therefore, when

91 planning a seed orchard, neighbouring trees should be cross-compatible to help maximise
92 seed set. The development of reliable primers, allele-specific and consensus, for detecting
93 incompatibility alleles by PCR (Sonneveld *et al.* 2001, Sonneveld *et al.* 2003) provides the
94 opportunity for genotyping the mother trees used for the establishment of seed orchards.
95 This study, the first to use molecular techniques to determine incompatibility alleles in wild
96 cherry, was carried out in a Belgian multiclonal plantation to provide data on tree *S*-
97 genotypes. These are needed for the assessment of gene flow and for the spatial design of
98 future seed orchards. The work led to the detection and partial sequencing of new alleles and
99 interesting data on allelic frequencies.

Material and methods

Plant material

The wild cherries studied were a set of 65 accessions, originating from two populations in the southern part of Belgium (Table 2): 26 accessions were selected as mature phenotypically superior trees (i.e. "plus" trees) in a 72-year-old forest stand near Anor (AN) and 39 accessions as juvenile clones within the half-sib progenies from various plus trees growing in a mixed forest stand near Solre-Saint-Géry (SSG). In 1988, grafted replicates of these accessions had been planted in a comparative multiclonal trial, which now constitutes the certified seed orchard of 'Mommedeel'.

Ten standard cultivars were included as references for the known alleles S_1 to S_{16} (Sonneveld *et al.* 2003).

DNA extraction

Genomic DNA was extracted from young leaves using the DNeasy® Plant Mini Kit (Qiagen). For wild cherry, no more than 5 mg of lyophilised leaf material should be processed. Higher amounts of starting material result in inefficient lysis of the plant material, low DNA purity and thus no or incomplete amplification with the second intron consensus primers, a problem which could be rectified by diluting the DNA samples in a 0.1 % TE buffer.

For some repeat analyses conducted in the winter, buds were collected and DNA was extracted following the procedure described by Sonneveld *et al.* (2001). DNA yield was quantified for each accession.

Genotyping approach

Genotyping the accessions involved several stages. First the samples were amplified with the consensus primers for the second intron. The size of the PCR-amplification products was estimated and preliminary deductions concerning the putative presence of one or two known alleles for each accession were tested by using the appropriate allele-specific primers. For genotypes that were not fully resolved, confirmation of the presence of a new allele was obtained by using fluorescent consensus primers for the first intron (Sonneveld *et al.* submitted) which yield PCR-products that can be sized on an automated sequencer resulting in accurate size determinations of the first intron product of the new alleles. Finally, representatives of new alleles were partially sequenced.

Consensus PCR and sizing – second intron

Amplifications for the second intron with the consensus primers PaConsII-F1 + PaConsII-R1 and detection of PCR products followed the protocols given in Sonneveld *et al.* (2003). Two 1 Kb Plus DNA ladders (Invitrogen) were loaded at the left and right side of the gel and a third ladder was loaded in the central part of the gel to enhance the accuracy of the estimates of the molecular weight (i.e. number of bp) of the DNA bands.

Rather than relying on a mere visual assessment of band sizes, these were estimated using a regression equation which expresses the linear relation between the natural logarithm of the size or length l (bp) of a band and its migration distance d measured from an arbitrary horizontal base line: $\ln(l) = \alpha * d + \beta$. In order to estimate the values of the parameters α and β, the base line is drawn through two bands of adjacent 1 Kb Plus ladders with a size larger then the size of any of the accession bands (Figure 1). To get a fairly accurate estimate of the

α value, the calculation of the regression equation should be based on at least sevens pairs of d_i and l_i values. In order to genotype an accession for the *S*-alleles present, the sizes of the two PCR-products were calculated by substituting their migration distance in the regression equation of the nearest ladder. Estimated band sizes were then compared with those of the standard cultivars; bands not approximating to known alleles were provisionally ascribed to new alleles.

Allele-specific PCR

To check the predictions from the consensus primers, DNA samples of accessions having bands matching or nearly matching known alleles were subjected to a PCR-reaction with appropriate allele-specific primers (Sonneveld *et al.* 2001, Sonneveld *et al.* 2003). Protocols for S_1 and S_3 to S_6 followed Sonneveld *et al.* (2001), except for the annealing temperature of the PCR which was adapted according to the recommendations of Sonneveld *et al.* (2003). Protocols for S_2 and S_7 to S_{16} followed Sonneveld *et al.* (2003). In each allele-specific PCR-reaction, a positive standard (i.e. a sweet cherry cultivar carrying the target *S*-allele) and a negative standard (i.e. a sweet cherry not carrying the target *S*-allele) were included, together with distilled water as the negative control. Samples not amplifying as expected were retested, with the inclusion of a PCR internal control, based on the use of primers for the PAL gene, recently recommended by Sonneveld *et al.* (2003) which allows false negatives to be identified. If alleles still failed to amplify, although the internal control produced a band, they were considered to be new.

Consensus PCR and sizing – first intron

Samples thought to have new alleles were amplified with the primer pair PaConsI-F1 + PaConsI-R2 with the forward primer carrying a fluorescent label (Sonneveld *et al.* submitted). The PCR-products were sized on an ABI 3100 automated capillary sequencer (Applied Biosystems) to gain accurate size information of the first intron product.

Partial sequencing of presumptive new alleles

For the purpose of sequencing, one representative accession for each putative new allele was chosen (Table 1). Sequencing of the new alleles was restricted to the part amplified by the consensus primers for the second intron, i.e. from conserved region C2 to conserved region C5 (Sonneveld *et al*. 2003). After amplification, the PCR-product was excised from the agarose gel and purified using the MinElute™ Gel Extraction Kit (Qiagen). Before sequencing, the purified PCR-product was re-amplified with the consensus primers and the excision and purification steps were repeated. After sequencing, the location of the intron/exon splicing sites, marked by the conserved GT/AG consensus sequence, was determined.

Amino acid sequences translated from the coding regions and nucleotide sequences of the second intron of the six alleles were compared using the DNAStar software (DNAStar Inc.). On the basis of the alignments, the percentage similarity of the partial amino acid sequences was calculated from the BLOSUM62 substitution matrix (Henikoff and Henikoff 1992) using the MultAlin software (Corpet 1988). For the analysis, residues from position 116 onward were discarded because of indels found in this variable region.

Allele frequencies

The modified chi-square test recommended by Davies in Campbell and Lawrence (1981) was used to test the null hypothesis of equal allele frequencies in the underlying populations. The expected total number of alleles in the populations was calculated using the E2 estimator developed by O'Donnell and Lawrence (1984) for populations showing unequal allelic frequencies.

Results

Amplification with consensus primers for the second intron

Amplification with the consensus primers for the second intron gave two bands for all accessions, except for one individual, presumably triploid, for which three bands were observed. The estimated band sizes ranged from 570 bp to 2430 bp (Table 2). Through visual alignment of the DNA bands on the gels, the alleles of all 65 accessions were classed into 13 groups, A to M. When the size of all bands was estimated graphically, each group was characterised by a median band size, the lower and upper quartiles, the range and the occurence of any outliers (Figure 2). After comparison with the known size of the genomic PCR-products for the second intron of the alleles S_1 to S_{16}, each size class appeared to correspond with one or two known *S*-alleles (e.g. size class A = S_6), except for classes G and J for which no match was found, thus indicating the presence of at least two new alleles.

Confirmation with allele-specific primers

When the DNA samples of the accessions belonging to a certain size class were subjected to a PCR-reaction using the specific primers developed for the matching known *S*-alleles, the presence of the known alleles was confirmed for all bands in classes A, B, C, E, F, H, K and M. In classes D and I, although the band sizes indicated the possible presence of alleles S_9 and S_{12} respectively, amplification with primers specific for these alleles was not obtained for all accessions, although the internal PAL-control amplified successfully. Testing the accessions in size class L against the S_2-specific primers, which was inspired by the proximity of the L-boxplot to that of size class K and by the decreasing accuracy of the

estimate of the number of base pairs with increasing band size, gave amplification for one of the three accessions. As five size classes (D, G, I, J, L) include alleles which were not amplified with any of the available allele-specific primers, the occurrence of at least five new alleles differing with respect to the length of the second intron, was indicated. However, in case alleles with similar size of the second intron PCR-product are not identical, their identity was checked with the consensus primers for the first intron.

Characterising putative new alleles with consensus primers for the first intron

When all accessions in size classes G and J and the accessions of classes D, I and L that did not amplify with any of the available allele-specific primers were tested with fluorescent primers for the first intron, all alleles within a class showed the same fluorescent peaks except for class G (Table 1). Within class G, four accessions yielded the same first intron product size (~ 342 bp) and were assumed to carry the same allele whereas accession AN10, the outlier in class G, is distinct with a different product size (~ 427 bp). Incidentally, in the three accessions that had S_{13} in addition to a new allele, the amplification of S_{13} was very weak or failed.

Sequencing of presumptive new alleles

Six putative new alleles were thus found in 19 accessions (Table 1) and characterised with respect to the size of the first and second intron. They are numbered S_{17} to S_{22}, in increasing order of size of the second intron.

Genomic sequences of all six '*S*-RNase' alleles from C2 to C5 were successfully obtained. Alignment of the deduced partial amino acid sequences of the alleles S_1 to S_{22} confirmed the

presence of the conserved regions C3 and RC4. In addition, it revealed considerable differences in the hypervariable region RHV, as expected, and just upstream of the conserved region C5 (Figure 3). It should be noted that far more length variation is found in the indel region prior to C5 than previously apparent in sweet cherry alleles. It is clear that the alleles S_{17} to S_{22} are not only different from the known alleles but are also mutually distinct. Amino acid similarity between new and known alleles ranged from 63.7 % (S_9 *vs.* S_{20}) to 82.1 % (S_2 *vs.* S_{21}). Comparison of the six new alleles revealed identities varying from 66.1 % (S_{20} *vs.* S_{22}) to 80.5 % (S_{19} *vs.* S_{22}).

Intron sequences could not be aligned, except for the two ends of the second intron including the splicing sites, 13 bp both at the 5' end and the 3' end. The A+T percentage of the second intron ranges from 65.9 % to 74.7 %, compared to 53.7 to 57.9 % for the coding regions.

The partial genomic sequences of the six new alleles, S_{17} – S_{22}, have been submitted to the EMBL database (DDBJ/EMBL/GenBank AJ862656, AJ862657, AJ862658, AJ862659, AJ863119 and AJ863120 respectively).

Genotypes and allele frequencies

The incompatibility genotypes of all 65 accessions were thus resolved (Table 2). One accession (AN57) showed three clear DNA bands for the second intron and so appeared to be triploid. This was confirmed by cytometric analysis (data not shown).

Seventeen alleles were detected in all, 14 in the 26 trees sampled from the Anor population and 13 in the 39 accessions from the Solre-Saint-Géry population. The frequency distribution of the *S*-alleles differs between the Solre-Saint-Gèry and Anor accessions (Figure 4). Although the expected number of alleles in each class is rather low, this difference is supported by a chi-square test ($\chi^2 = 74.54$; df = 16; $P<0.001$). Alleles S_6 and S_{13}, which are

very frequent alleles in the Solre-Saint-Gèry population, appear to be completely absent from the Anor-accessions, whereas alleles S_3 and S_{14}, which are very frequent in Anor, are rare in Solre-Saint-Gèry. Allele S_9 occurs in Anor but is absent from Solre-Saint-Géry. Of the new alleles, only S_{17} and S_{18} appear to be common to both populations.

The hypothesis of equal allele frequency can be rejected, both in the Solre-Saint-Géry population ($\chi^2 = 39.27$; df = 12; $P<0.001$) and in the Anor population ($\chi^2 = 34.67$; df = 13; $P<0.001$). Taking into account the unequal allele frequencies in both populations, the expected total number of alleles calculated with the E2 estimator is 14 and 16 in the Solre-Saint-Géry and Anor populations respectively. However, these may be underestimates as the sampling was not random. Accessions could indeed be related as they were selected as phenotypically similar plus trees.

Discussion

The combination of consensus and allele-specific *S*-ribonuclease primers allowed the *S*-genotyping of all 65 accessions. Both alleles came from S_1 to S_{16} in 46 accessions, 19 individuals carry one new allele and none showed the presence of two new alleles. One accession appeared to be triploid, carrying three *S*-alleles. Convincing evidence, backed up with sequence data, is given for the occurrence of six new alleles which are labelled from S_{17} on.

Consensus and allele-specific primers

The usefulness of the consensus primers, in particular those developed for the second intron, lies in the drastic reduction of the number of PCR-reactions required for the subsequent confirmation of the *S*-alleles present. Without the information yielded by the consensus primers, genotyping the 65 wild cherry accessions for 13 *S*-alleles could involve up to 845 PCR-reactions. In reality, 65 second intron consensus PCR's meant that just 136 allele-specific PCR's and 19 first intron consensus PCR's were required to resolve all genotypes, i.e. identifying known as well as new alleles. In 82 % of the cases, the most likely *S*-allele, as indicated by the size of the second intron, was later confirmed to be the one present. Known alleles with second introns of the same size were consistently identified as being identical with allele-specific primers in all 112 examples, supporting the assumption of a one-to-one correspondence between *S*-alleles and second intron lengths in cherry. Minor intra-allelic size variation of the second intron has been reported in almond for the allele S_1 (Ma and Oliveira 2001), but not in cherry.

Although non-amplification with any of the available allele-specific primers indicates the occurrence of a new allele, no information about the number of new alleles can be obtained in the absence of consensus amplification data. Use of a combination of first and second intron consensus primers provides a fairly reliable indication of the number of new alleles.
Finally, the use of consensus primers can detect polyploid individuals as these are marked by the appearance of a clear multiple band pattern when running the PCR-products on an agarose gel.

Confirmation of new alleles

The genotypes of accessions thought to carry new alleles on the basis of failure to amplify with allele-specific primers were clarified with the use of fluorescent primers amplifying the first intron. Assuming *S*-alleles are distinguished by the length of the first intron and the second intron, six new alleles were indicated. Sequencing confirmed they were mutually distinct and different from the alleles S_1 to S_{16}. Interestingly, the partial amino acid sequence of S_{21} perfectly matched that of the allele recently reported as S_{25} in a Spanish cultivar of sweet cherry by Wünsch and Hormaza (2004) (DDBJ/EMBL/Genbank AY259113).
The detection of six new *S*-alleles makes the development of additional allele-specific primers desirable. As the consensus primers for the second intron are positioned in the highly conserved regions C2 (forward primer) and C5 (reverse primer), a large span of the *S*-allele has been amplified, including the highly variable region RHV and the intron. This sequence thus offers sufficient information for designing allele-specific primers. After alignment of these partial sequences with those of the 13 known alleles $S_1 – S_{16}$, differences can be used for the design of specific primers for the six new alleles $S_{17} – S_{22}$.

Allele frequencies

The allele frequencies both in the Anor and the Solre-Saint-Géry populations were shown to be unequal. Campbell and Lawrence (1981) discussed two hypotheses to account for such unequal allele frequencies. First, some alleles may be subject to selection over and above that associated with incompatibility, resulting either from pleiotropic effects of the *S*-gene or from its close linkage with genes that affect fitness. Second, unequal allele frequencies may be due to founder effects, if the populations are in a state of disequilibrium. As the most common alleles are different in the two populations, the latter hypothesis is plausible but *S*-linked genes that affect fitness remain a possible explanation if different couplings are present in the different populations. Different common alleles certainly rule out pleiotropic effect.

The relative occurrence of *S*-alleles in the 65 wild cherry accessions constituting the seed orchard 'Mommedeel' was compared to the occurrence of incompatibility alleles in 67 sweet cherry cultivars identified by analysing stylar ribonucleases (Bošković and Tobutt 2001) (Figure 5). Allelic frequencies are significantly different ($\chi^2 = 84.17$; df = 13; $P<0.001$) with the absence of S_4 and S_5 in wild cherry being particularly notable. Should this observation persist after further determination of self-incompatibility genotypes in the wild cherry gene pool on which the Flemish breeding programme is based (i.e. selected accessions and certified seed stands), then the occurrence of S_4 and/or S_5 in populations of wild cherry may indicate introgression from sweet cherry cultivars. Besides the absence of S_{10} in sweet cherry cultivars, it should be noted that S_7 and S_{12} to S_{16} are rare. Likewise S_9 is rare in the wild cherry collection. Such information could be useful for checking that seed or seedlings offered for sale as wild are indeed of wild origin. It is thought that sweet cherry derives particularly from introductions from the Near East rather than from northern European populations. A fuller discussion of differences in allelic frequencies between sweet and wild

cherry must await data from a wider range of wild populations. Recently Békefi *et al.* (2003) and Wünsch and Hormaza (2004) have indicated that alleles matching S_{22} and S_{21} may be present in the cultivars 'Rita' and 'Taleguera Brillante' respectively.

Genotypes

Only six wild cherry accessions share a genotype with one of the 22 incompatibility groups defined in the 'Harmonization Table' drawn up by Tobutt *et al*. (2001). If a minimum of two accessions having the same *S*-genotype are needed to establish an incompatibility group, then 16 new groups could be defined. However, as the incompatibility group system has been developed primarily for sweet cherry cultivars it may be unhelpful to introduce many wild cherry accessions into the Harmonization Table. Seventeen combinations of *S*-alleles are represented by one accession only and thus can currently be regarded as unique genotypes and therefore as universal pollinators.

As the accessions constitute the components of a seed orchard, their cross-compatibility is of major importance. Out of the 2080 possible combinations of parents, 51 are cross-incompatible, 620 are semi-compatible and 1409 are fully compatible. Although the proportion of cross-incompatible combinations is low (2.45 %), knowledge of these is important as pollination in wild cherries occurs mainly between neighbouring trees.

The determination of incompatibility genotypes of the 65 wild cherry accessions will serve the study of gene flow in the seed orchard 'Mommedeel' as a test case. The main goal is the unravelling of the pollination pattern through paternity analysis of half-sib seedlings. Once assessed, gaps in the pollination pattern, if any, may be explained by the cross-incompatibility of neighbouring trees, by a major discrepancy in flowering period or the lethality of certain progenies. Paternity analyses are most often based on microsatellite

markers. However, by analogy with microsatellite loci, the power of discrimination PD (Kloosterman *et al.* 1993) of the *S*-locus can be computed: $PD = 1 - \Sigma(P_i)^2$ where, in this case, P_i represents the frequency of each of the incompatibility genotypes identified in the collection of 65 accessions (i = 1 to 38). The high PD-value (0.96) shows that for purposes of paternity analysis, the information provided by scoring microsatellite markers can be usefully complemented by genotyping the highly polymorphic *S*-locus with 19 alleles identified so far.

The productivity of seed orchards depends not only on the spatial arrangement of mother trees but also on the selection of their constituents. Genotyping of the *S*-locus with consensus primers allows the identification of triploid accessions. Spontaneous triploids have rarely been reported in wild cherry and can be exceptionally vigorous (Schelhorn 1947) which could favour their selection as plus trees. However, as they probably produce weak aneuploid seedlings they are unlikely to be useful for breeding purposes and therefore should not be included in seed orchards.

Forward look

The analytical approaches and allele characterization data presented here will be useful for studying the *S*-locus not only in seed orchards but in wild populations of cherry.

Other rosaceous tree species of economic importance for timber production (e.g. *Sorbus torminalis*) or of conservation interest (e.g. *Malus sylvestris*, *Pyrus pyraster*) are likely to have a ribonuclease-based gametophytic incompatibility system. The advent of PCR-primers for *S*-alleles in related fruit crops should facilitate the genotyping of individuals for studying the population genetics of the *S*-locus in seed orchards and wild populations of these other species.

Acknowledgements

Bart De Cuyper aknowledges receipt of the funding of a sabbatical leave from the Training, Personnel and Organisation Administration of the Ministry of the Flemish Community. He thanks East Malling Research for acting as a host location for this sabbatical leave and for providing the technical facilities necessary to carry out the research.

References

Békefi Zs, Tobutt KR, Sonneveld T (2003) Determination of (in)compatibility genotypes of Hungarian sweet cherry (*Prunus avium* L.) accessions by PCR based methods. International Journal of Horticultural Science, **9** (3-4), 37-42

Berger K (1963) Fertilitätsuntersuchungen an der Wildkirsche *Prunus avium* L. var. *avium*. Kulturpflanze, **11**, 358-375

Bošković R, Tobutt KR (1996) Correlation of stylar ribonuclease zymograms with incompatibility alleles is sweet cherry. Euphytica, **90**, 245-250

Bošković R, Tobutt KR (2001) Genotyping cherry cultivars assigned to incompatibility groups, by analysing stylar ribonucleases. Theoretical and Applied Genetics, **103**, 475-485

Bošković R, Russell K, Tobutt KR (1997) Inheritance of stylar ribonucleases in cherry progenies, and reassignment of incompatibility alleles to two incompatibility groups. Euphytica, **95**, 221-228

Campbell JM, Lawrence MJ (1981) The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. II. The number and frequency of *S*-alleles in a natural population (R106). Heredity, **46**, 81-90

Corpet F (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Research, **16**, 10881-10890

Crane MB, Lawrence WJC (1929) Genetical and cytological aspects of incompatibility and sterility in cultivated fruits. Journal of Pomology and Horticultural Science, **7**, 276-301

Goulson D, Hawson SA, Stout JC (1998) Foraging bumblebees avoid flowers already visited by conspecifics or by other bumblebee species. Animal Behaviour, **55**, 199-206

Heinrich B (1976) The foraging specializations of individual bumblebees. Ecological Monographs, **46**, 105-128

Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. Proceedings of the National Academy of Sciences of the United States of America, **89**, 10915-10919

Kloosterman AD, Budowle B, Daselaar P (1993) PCR-amplification and detection of the human D1S80 VNTR locus. Amplification conditions, population genetics and application in forensic analysis. International Journal of Legal Medicine, **105**, 257-264

Ma R-C, Oliveira MM (2001) Molecular cloning of the self-incompatibility genes *S1* and *S3* from almond (*Prunus dulcis* cv. Ferragnès). Sexual Plant Reproduction, **14**, 163-167

Matthews P, Dow KP (1969) Incompatibility groups: sweet cherry (*Prunus avium*). In: Knight RL(ed.), Abstract Bibliography of Fruit Breeding and Genetics to 1965, *Prunus*. Commonwealth Agricultural Bureaux, Farnham Royal, 540-544

O'Donnell S, Lawrence MJ (1984) The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. IV. The estimation of the number of alleles in a population. Heredity, **53**, 495-507

Schelhorn M, von (1947) Über eine triploide Vogelkirsche. Züchter, **6/8**, 232-235

Schütt P, Schuck HJ, Lang UM, Roloff A (1995) Enzyklopädie der Holzgewächse; Handbuch und Atlas der Dendrologie, 2. Ecomed, Landsberg am Lech.

Sonneveld T, Robbins TP, Bošković R, Tobutt KR (2001) Cloning of six cherry self-incompatibility alleles and development of allele-specific PCR detection. Theoretical and Applied Genetics, **102**, 1046-1055

Sonneveld T, Tobutt KR, Robbins TP (2003) Allele-specific PCR detection of sweet cherry self-incompatibility (S) alleles S_1 to S_{16} using consensus and allele-specific primers. Theoretical and Applied Genetics, **107**, 1059-1070

Sonneveld T, Robbins TP, Tobutt KR. Improved discrimination of self-incompatibility *S*-RNase alleles in cherry and high throughput genotyping by automated sizing of first intron PCR products. Molecular Ecology Notes (submitted)

Tobutt KR, Sonneveld T, Bošković R (2001) Cherry (in)compatibility genotypes – Harmonization of recent results from UK, Canada, Germany Japan and USA. Eucarpia Fruit Breeding Section Newsletter, **5**, 41-46

Tobutt KR, Sonneveld T, Bekefi Z, Bošković R Cherry (in)compatibility genotypes – An updated cultivar table. Acta Horticulturae. In press

Ushijima K, Sassa H, Tao R, Yamane H, Dandekar AM, Gradziel TM, Hirano H (1998) Cloning and characterization of cDNAs encoding *S*-RNases from almond (*Prunus dulcis*): primary structural features and sequence diversity of the *S*-RNases in Rosaceae. Molecular and General Genetics, **260**, 261-268

Wünsch A, Hormaza JI (2004) Cloning and characterization of genomic DNA sequences of four self-incompatibility alleles in sweet cherry (*Prunus avium* L.). Theoretical and Applied Genetics, **108**, 299-305

Author Information Box

B De Cuyper is a forest geneticist with special interest in selection and breeding of wild cherry. His research focuses on the establishment of clonal seed orchards, based on gene flow studies combined with assessment of self-incompatibility genotypes of mother trees. KR Tobutt works on breeding and genetics of fruit crops, timber trees and woody ornamentals and is interested in rosaceous incompatibility. T Sonneveld recently completed a PhD on self-incompatibility in sweet cherry in which she developed consensus and allele-specific primers for *S*-alleles. This work represents an application of techniques developed for fruit breeding to the improvement of timber trees.

Figures

Fig. 1 PCR band patterns after amplification with consensus primers for the second intron of the 19 accessions carrying one of the new alleles S_{17} to S_{22} (*S*-genotype in brackets). Bands are grouped in five size classes D, G, I, J and L (mean size (bp) in brackets). La = 1 Kb Plus ladder; C = distilled water as negative control

Fig. 2 The 13 size classes A to M of *S*-allele second intron PCR-products of 65 wild cherry accessions. Box and whiskers plots indicate the median, upper and lower quartile, range and outliers for each class. Between brackets on the x-axis and y-axis are shown, respectively, the number of bands assigned to each class and the size of the PCR-products for the known alleles S_1 to S_{16} (Sonneveld *et al*. 2003)

Fig. 3 Alignment of amino acid sequences of cherry self-incompatibility alleles S_1 to S_{22} deduced from the coding sequence between the conserved regions C2 and C5. Residues different from the consensus and gaps (dashes) are highlighted in black boxes, except for the gaps from position 117 onward. For the annotation of conserved and variable regions the convention of Ushijima *et al*. (1998) was followed. No attempt was made to align the sequences in the variable region with 'indels' just upstream of C5 (position 116 to 136)

Fig. 4 Comparison of the relative occurrence of *S*-alleles between two populations of wild cherry, Solre-Saint-Géry (ν) and Anor (□)

Fig. 5 Comparison of the relative occurrence of *S*-alleles between a collection of 65 wild cherry accessions (ν) and 67 sweet cherry cultivars (□)

Table 1 Sizes of the PCR-products after amplification with the consensus primers for the first and second intron for wild cherry accessions carrying new *S*-alleles. Based on the 95 % confidence interval of the log-linear regression, values for the second intron are estimates with a confidence level of ± 4 %.

Size class	Accession	Known allele	New allele			
			2nd intron gel score (bp)	class mean (bp)	1st intron fluorescent score (bp)	Label
D (part)	SSG5	S_6	797	788	396	S_{17}
	SSG6	S_6	795		396	
	SSG16†	S_{13}	772		396	
	SSG21	S_{13}	773		396	
	SSG26	S_7	778		396	
	AN43	S_3	805		396	
	AN44	S_3	798		396	
G	SSG18	S_2	937	935	342	S_{18}
	AN69†	S_3	934		342	
	AN102	S_{14}	936		342	
	AN103	S_{14}	932		342	
	AN10†	S_{14}	971	971	427	S_{19}
I (part)	AN22†	S_2	1724	1736	326	S_{20}
	AN23	S_2	1735		326	
	AN97	S_{14}	1748		326	
J	AN57*	S_7S_{12}	2014	2010	375	S_{21}
	AN104†	S_{14}	2006		375	
L (part)	SSG1†	S_{13}	2258	2264	422	S_{22}
	SSG35	S_6	2270		422	

* Triploid accession

† Accessions selected for the partial sequencing of the new alleles

Table 2 Determining the *S*-alleles in 65 wild cherry accessions based on estimated length of the second intron PCR-product and on the amplification with allele-specific primers

Accession	Genotyping with consensus primers 2nd intron				Amplification with allele-specific primers												Conclusion‡
	Allele 1		Allele 2														
	# bp	Candidates*	# bp	Candidates*	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_9	S_{10}	S_{12}	S_{13}	S_{14}	S_{16}	
SSG1	861	S_1 / S_{13}	2258	S_2	-	-									+			$S_{13}S_{22}$
SSG2	872	S_1 / S_{13}	2211	S_2	+	+									-			S_1S_2
SSG3	584	S_6	871	S_1 / S_{13}	+					+								S_1S_6
SSG4	740	S_{10} / S_{14}	882	S_1 / S_{13}	-								+		+			$S_{10}S_{13}$
SSG5	591	S_6	797	S_9						+		-						S_6S_{17}
SSG6	591	S_6	795	S_9						+		-						S_6S_{17}
SSG7	740	S_{10} / S_{14}	2239	S_2		+							+					S_2S_{10}
SSG8	738	S_{10} / S_{14}	902	S_3			+						+					S_3S_{10}
SSG9	591	S_6	2419	S_7						+	+							S_6S_7
SSG10	740	S_{10} / S_{14}	2240	S_2		+							+					S_2S_{10}
SSG11	588	S_6	2227	S_2		+				+								S_2S_6
SSG12	577	S_6	2214	S_2		+				+								S_2S_6
SSG13	860	S_1 / S_{13}	2225	S_2	+	+												S_1S_2
SSG14	861	S_1 / S_{13}	2388	S_7	-						+				+			S_7S_{13}
SSG15	570	S_6	716	S_{14} / S_{10}						+			+			-		S_6S_{10}
SSG16	772	S_9	857	S_1 / S_{13}	-							-			+			$S_{13}S_{17}$
SSG17	583	S_6	2210	S_2		+				+								S_2S_6
SSG18	937	S_3	2226	S_2		+	-											S_2S_{18}
SSG19	871	S_1 / S_{13}	2206	S_2	+	+												S_1S_2
SSG20	873	S_1 / S_{13}	2406	S_7							+				+			S_7S_{13}
SSG21	773	S_9	859	S_1 / S_{13}	-							-			+			$S_{13}S_{17}$
SSG22	716	S_{14} / S_{10}	2207	S_2		+										+		S_2S_{14}
SSG23	726	S_{10} / S_{14}	1430	S_{16}									+				+	$S_{10}S_{16}$
SSG24	579	S_6	726	S_{10} / S_{14}						+			+					S_6S_{10}
SSG25	887	S_3	2389	S_7			+				+							S_3S_7
SSG26	778	S_9	2388	S_7							+	-						S_7S_{17}
SSG27	1429	S_{16}	1742	S_{12}										+			+	$S_{12}S_{16}$
SSG28	871	S_1 / S_{13}	1426	S_{16}	+												+	S_1S_{16}
SSG29	726	S_{10} / S_{14}	1752	S_{12}									+	+				$S_{10}S_{12}$
SSG30	716	S_{14} / S_{10}	1433	S_{16}												+	+	$S_{14}S_{16}$
SSG31	1420	S_{16}	2193	S_2		+											+	S_2S_{16}
SSG32	878	S_1 / S_{13}	2400	S_7							+				+			S_7S_{13}
SSG33	583	S_6	864	S_1 / S_{13}	-					+					+			S_6S_{13}
SSG34	576	S_6	2216	S_2		+				+								S_2S_6
SSG35	576	S_6	2270	S_2		-				+								S_6S_{22}
SSG36	868	S_1 / S_{13}	2404	S_7	+						+							S_1S_7
SSG37	1437	S_{16}	2210	S_2		+											+	S_2S_{16}
SSG38	573	S_6	725	S_{10} / S_{14}						+			+					S_6S_{10}
SSG39	576	S_6	2404	S_7						+	+							S_6S_7

Table 2 Continued

Accession	Genotyping with consensus primers 2nd intron				Amplification with allele-specific primers												Conclusion‡
	Allele 1		Allele 2														
	# bp	Candidates*	# bp	Candidates*	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_9	S_{10}	S_{12}	S_{13}	S_{14}	S_{16}	
AN8	871	S_1 / S_{13}	1430	S_{16}	+												+	S_1S_{16}
AN9	715	S_{14} / S_{10}	891	S_3			+									+		S_3S_{14}
AN10	715	S_{14} / S_{10}	971													+		$S_{14}S_{19}$
AN11	2220	S_2	2413	S_7		+					+							S_2S_7
AN20	785	S_9	2215	S_2		+						+						S_2S_9
AN22	1724	S_{12}	2223	S_2		+								-				S_2S_{20}
AN23	1735	S_{12}	2206	S_2		+								-				S_2S_{20}
AN26	735	S_{10} / S_{14}	2224	S_2		+							+					S_2S_{10}
AN28	737	S_{10} / S_{14}	2209	S_2		+							+					S_2S_{10}
AN30	735	S_{10} / S_{14}	2256	S_2		+							+					S_2S_{10}
AN31	909	S_3	2392	S_7			+				+							S_3S_7
AN43	805	S_9	918	S_3			+					-						S_3S_{17}
AN44	798	S_9	916	S_3			+					-						S_3S_{17}
AN48	916	S_3	2430	S_7			+				+							S_3S_7
AN50	727	S_{10} / S_{14}	2260	S_2		+							-			+		S_2S_{14}
AN57†	1780 & 2415	S_{12} & S_7	2014								+			+				$S_7S_{12}S_{21}$
AN59	1773	S_{12}	2410	S_7							+			+				S_7S_{12}
AN60	710	S_{14} / S_{10}	887	S_3			+									+		S_3S_{14}
AN69	882	S_3	934				+											S_3S_{18}
AN76	1439	S_{16}	2380	S_7							+						+	S_7S_{16}
AN78	736	S_{10} / S_{14}	2218	S_2		+							+					S_2S_{10}
AN97	712	S_{14} / S_{10}	1748	S_{12}										-		+		$S_{14}S_{20}$
AN102	712	S_{14} / S_{10}	936	S_3			-									+		$S_{14}S_{18}$
AN103	716	S_{14} / S_{10}	932	S_3			-									+		$S_{14}S_{18}$
AN104	716	S_{14} / S_{10}	2006													+		$S_{14}S_{21}$
AN109	719	S_{14} / S_{10}	894	S_3			+									+		S_3S_{14}

* Decreasing order of likelihood

† Triploid accession

‡ Determination and labelling of new alleles S_{17} to S_{22} in accordance with Table 1

Figure 1

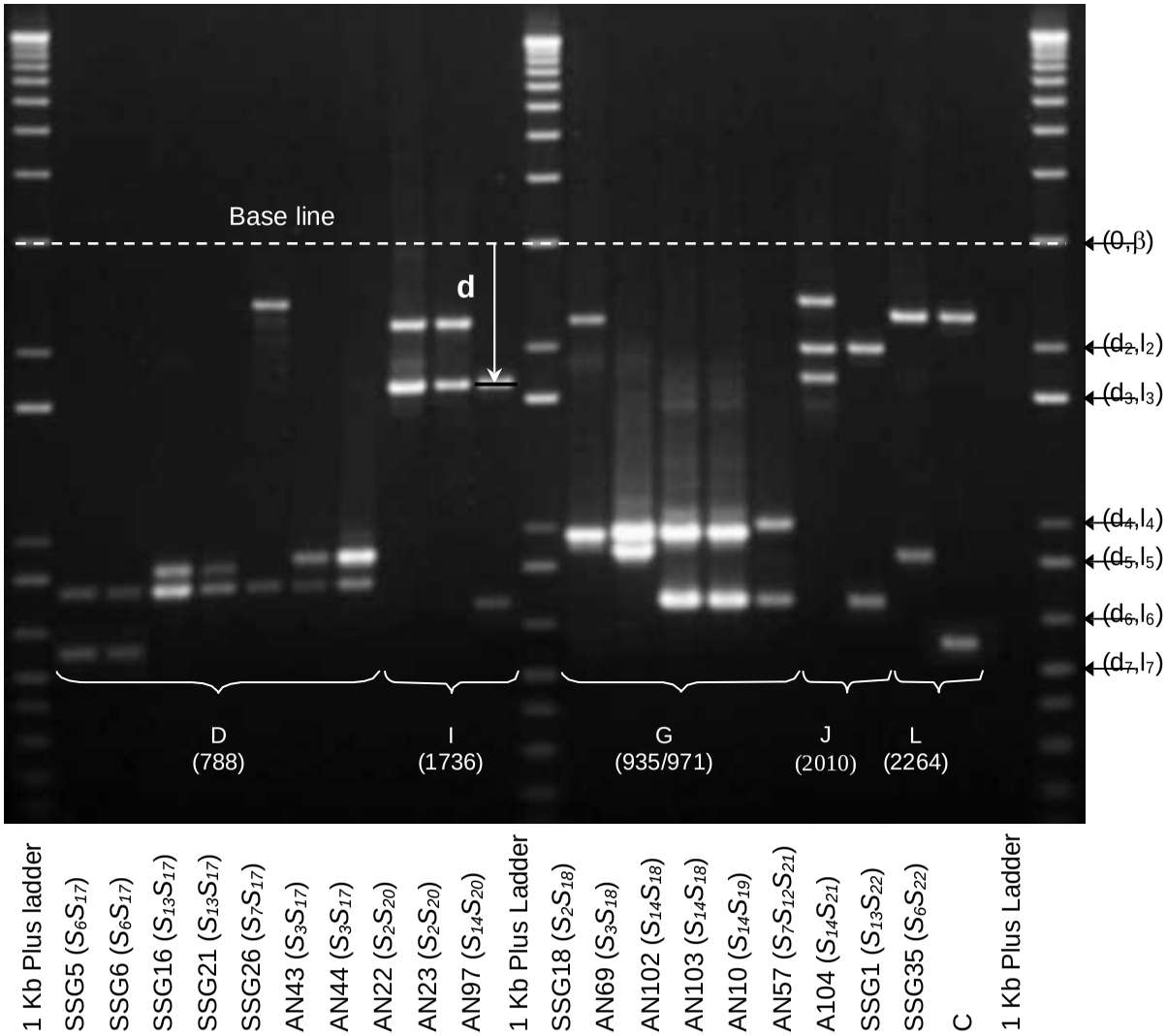

Base line
d
(0,β)
(d2,l2)
(d3,l3)
(d4,l4)
(d5,l5)
(d6,l6)
(d7,l7)
D
(788)
I
(1736)
G
(935/971)
J
(2010)
L
(2264)
1 Kb Plus ladder
SSG5 (S6S17)
SSG6 (S6S17)
SSG16 (S13S17)
SSG21 (S13S17)
SSG26 (S7S17)
AN43 (S3S17)
AN44 (S3S17)
AN22 (S2S20)
AN23 (S2S20)
AN97 (S14S20)
1 Kb Plus Ladder
SSG18 (S2S18)
AN69 (S3S18)
AN102 (S14S18)
AN103 (S14S18)
AN10 (S14S19)
AN57 (S7S12S21)
A104 (S14S21)
SSG1 (S13S22)
SSG35 (S6S22)
C
1 Kb Plus Ladder

Figure 2

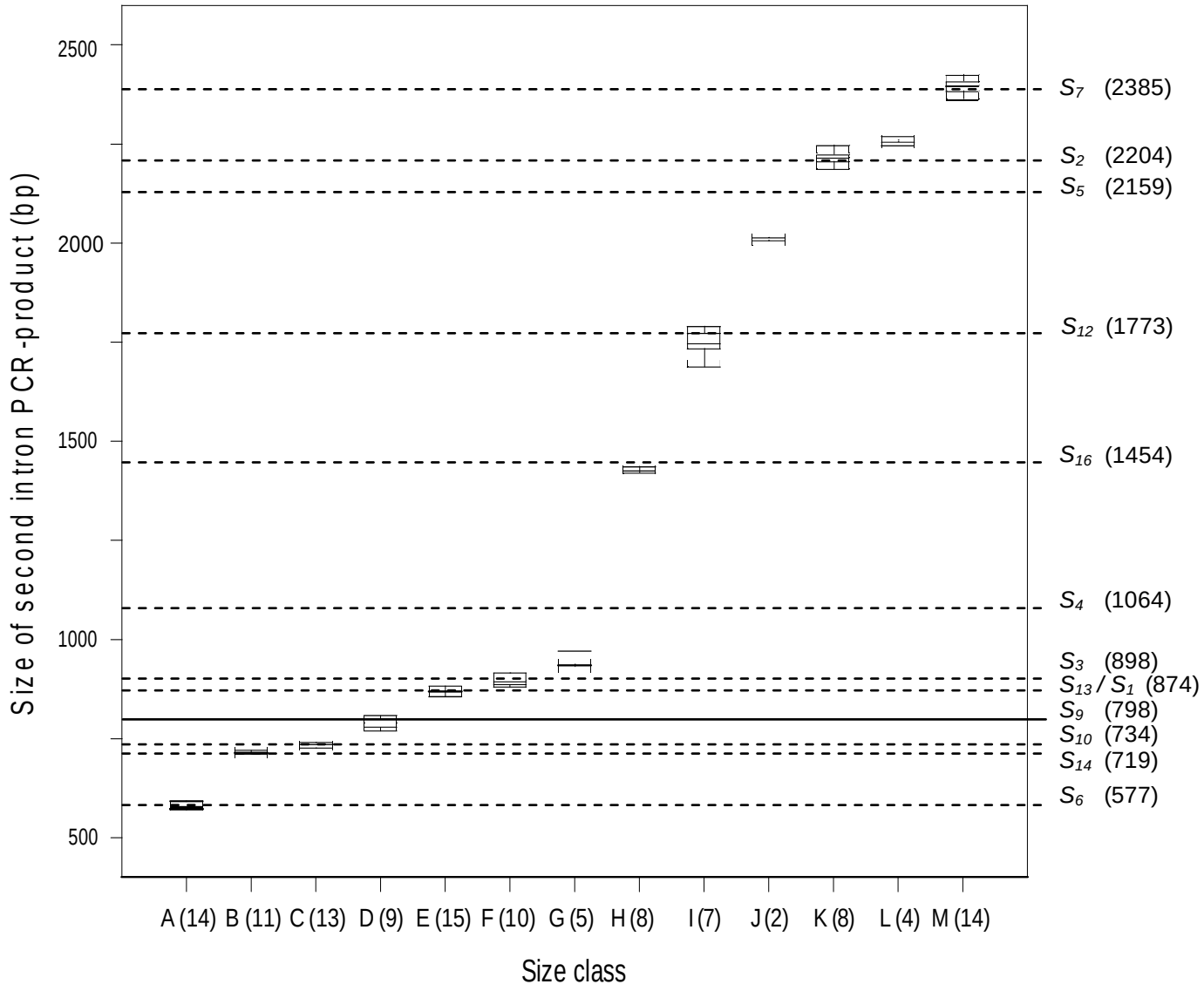

Size of second intron PCR-product (bp)
2500
2000
1500
1000
500
S7 (2385)
S2 (2204)
S5 (2159)
S12 (1773)
S16 (1454)
S4 (1064)
S3 (898)
S13 / S1 (874)
S9 (798)
S10 (734)
S14 (719)
S6 (577)
A (14) B (11) C (13) D (9) E (15) F (10) G (5) H (8) I (7) J (2) K (8) L (4) M (14)
Size class

Figure 3

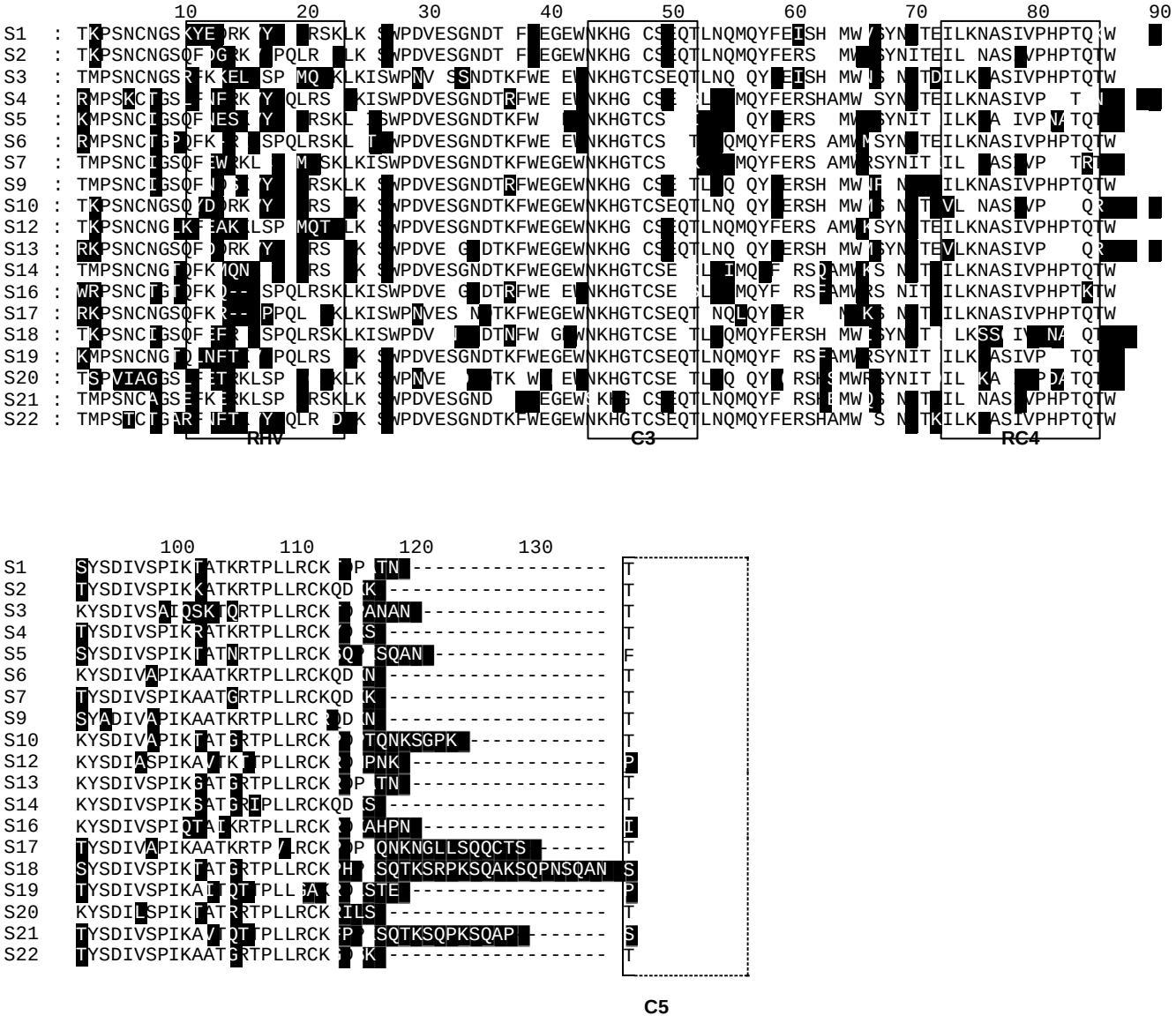


Figure 4

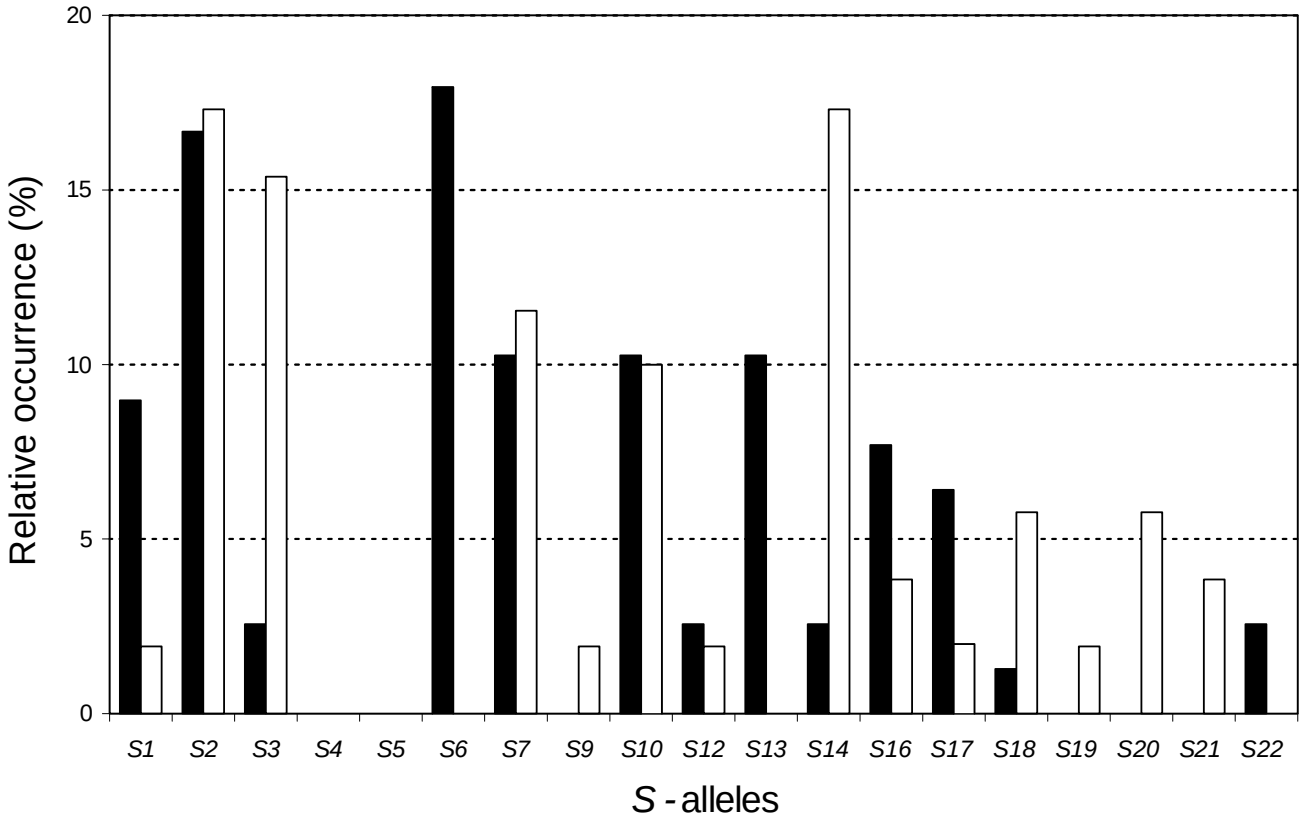
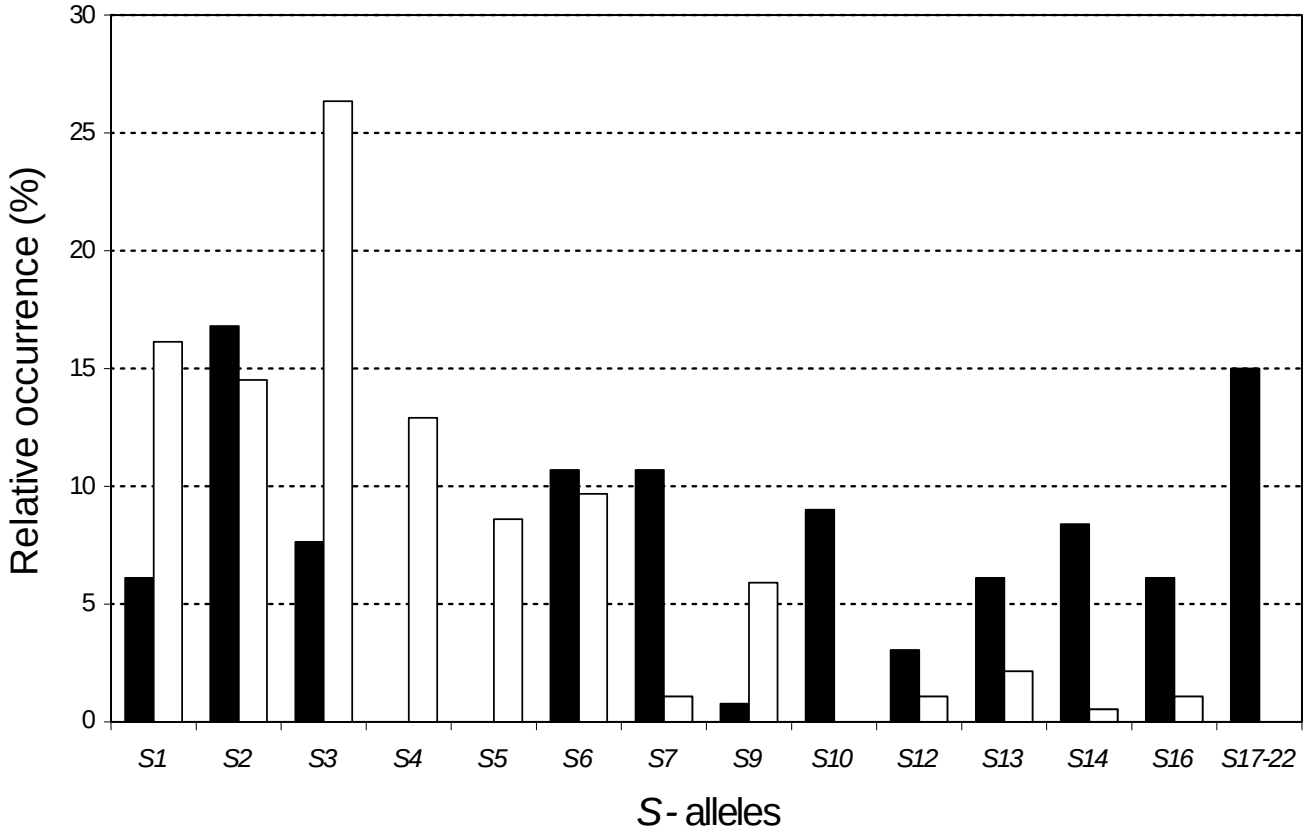

20
15
10
5
0
Relative occurrence (%)
S1
S2
S3
S4
S5
S6
S7
S9
S10
S12
S13
S14
S16
S17
S18
S19
S20
S21
S22
S - alleles

Figure 5


30
25
20
15
10
5
0
Relative occurrence (%)
S1
S2
S3
S4
S5
S6
S7
S9
S10
S12
S13
S14
S16
S17-22
S - alleles