# GENETIC MANIPULATION OF SELF-INCOMPATIBILITY IN DIPLOID POTATO SPECIES

**DANIEL KWADJO DZIDZIENYO, BSc, MSc.**

**Thesis submitted to the University of Nottingham**

**for the degree of Doctor of Philosophy**

**September 2012**

## ABSTRACT

Many of the wild and some cultivated species of potato are true diploids and are therefore more amenable for genetic studies than the majority of tetraploid cultivars. However, the use of these diploid *Solanum* species is complicated by almost universal self-incompatibility (SI). In *Solanum*, SI is gametophytic and pistil specificity is controlled by a polymorphic ribonuclease (S-RNase), as found in other members of the Solanaceae. The genetic engineering of self-compatible (SC) diploid potato lines would benefit potato breeding in general and allow inbred lines to be established for the first time. This would facilitate genetic analyses including that of complex traits such as drought resistance or yield. The aim of this thesis is to downregulate the expression of S-RNases in diploid potatoes using the RNAi technique and established procedures for *Agrobacterium*-mediated transformation. This approach to engineering self-compatibility has already been successfully demonstrated in SI *Petunia inflata* (Lee *et al.*, 1994) and other species of the Solanaceae.

To date just a handful of S-RNase sequences are available for potato species. The characterization of S-RNases in targeted diploid *Solanum* species was an initial requirement for our approach. To develop the tools, S-alleles have initially been characterized in both *Petunia inflata* and *P. hybrida cv* Mitchell both phenotypically (by pollination tests using a diallel cross) and/or genotypically (by RT-PCR). This approach was then transferred to three diploid potato species, specifically accessions of *Solanum stenotomum*, *Solanum phureja* and *Solanum okadae*. These

wild species are important sources of new traits studied by The James Hutton Institute (formally the Scottish Crop Research Institute).

The approach taken to amplify partial S-RNase sequences from pistil RNA was RT-PCR using a degenerate primer. PCR products were cloned using a TA vector (Invitrogen) and sequenced. For two alleles full length sequences were obtained by 5'RACE. Database searches with these sequences, revealed sixteen S-RNases several of which are novel. Phylogenetic analysis was carried out with the cloned S-RNases together with selected published S-RNase and S-like RNase sequences of solanaceous species. The S-RNases revealed extensive trans-generic evolution and are clearly distinct from and distantly related to S-like RNases. For two alleles ($S_{o1}$ and $S_{o2}$), S-RNase gene expression profiling was performed to check the developmental expression of the S-RNase gene, tissue-specific expression and also test whether these S-RNases (*e.g.* $S_{o1}$- and $S_{o2}$-RNases) are expressed at a similar level. Wide variation in S-RNase gene expression levels have been reported in the literature.

An RNAi construct has been designed to down-regulate two specific S-RNases in an $S_{o1}/S_{o2}$ heterozygote of *S. okadae*. To increase the chance of silencing, the RNAi construct has been designed to use a chimeric S-RNase gene involving the 5' end of the $S_{o1}$-RNase and the 3' end of the $S_{o2}$-RNase. The correct chimeric S-RNase construct ($S_{o1}/S_{o2}$-RNase) has now been identified and inserted into an RNAi vector (pHellsgate8) using Gateway® technology. This RNAi construct (pHG8-$S_{o1}/S_{o2}$) is now a valuable resource for use in S-RNase gene silencing in potato leading to

the development of self-compatible diploid potato lines and ultimately the development of the first inbred lines of *S. okadae*.

## ACKNOWLEDGEMENTS

I thank the almighty God for His care, mercies, favour and grace which has brought me this far. My profound gratitude goes to my supervisor Dr Tim Robbins whose mentoring, inspirations, tolerance, friendship, criticisms, comments, contributions and invaluable guidance this research was made possible. I also acknowledge with sincere appreciations, the support, encouragement and advice of my co-supervisor Dr Glenn Bryan of the James Hutton Institute (JHI, Dundee, Scotland) throughout this research. Special thanks also go to Dr Jennifer Stephens, Dr Mark Taylor, Dr Laurence Ducreux, Gail Wilde, Karen McLean and Dr Sanjeev Kumar Sharma all of JHI for their help and support throughout this study. I further express my gratitude to Dr Ian Taylor, Dr Paul Anthony, Mark Meacham, Rose Beacham and all the academic, administrative and technical staff at the Plant and Crop Sciences Division, School of Biosciences, University of Nottingham. I sincerely thank Dr Tineke Sonneveld, Dr Upendra Kumar Devisetty, Dr Rafal Czyzyk, Dr Jose Fernandez, Benjamin Taylor and all my colleagues and friends for their friendship, encouragement and assistance throughout my research work. My heartfelt appreciations also go to my parents and siblings for their continuous encouragement and support throughout my education to this level. A very special thank you also go to The James Hutton Institute and the Plant and Crop Sciences Division, University of Nottingham for the joint studentship, without which this research would not have been made possible. To all I say God bless you.

TABLE OF CONTENTS

*ABSTRACT* ..... *i*

*ACKNOWLEDGEMENTS* ..... *iv*

*LIST OF FIGURES* ..... *x*

*LIST OF TABLES* ..... *xiv*

*LIST OF ABBREVIATIONS* ..... *xv*

*CHAPTER 1: GENERAL INTRODUCTION* ..... *1*

**1.1 Fertilization in flowering plants** ..... **1**

**1.2 General overview of self-incompatibility** ..... **4**

**1.3 Homomorphic self-incompatibility systems** ..... **6**

1.3.1 Sporophytic self-incompatibility (SSI) ..... 8

1.3.2 Gametophytic self-incompatibility (GSI) ..... 15

1.3.2.1 S-RNase based (Solanaceae type) self-incompatibility system ..... 16

1.3.2.1.1 Identity and function of the pollen S-gene ..... 19

1.3.2.1.2 Modifier genes involved in S-RNase-based GSI system ..... 21

1.3.2.1.3 Proposed models for the S-RNase-based GSI system 23

1.3.2.2 GSI system in Papaveraceae ..... 33

**1.4 The origin of self-incompatibility systems** ..... **36**

1.4.1 Self-incompatibility evolution and maintenance ..... 36

1.4.2 Diversity of SI systems and polymorphism at the S-locus ..... 40

**1.5 Biotechnology and manipulation of self-incompatibility** ..... **41**

1.5.1 S-RNase regulation in transgenic plants ..... 41

**1.6 Aims of the project** ..... **43**

*CHAPTER 2: GENERAL MATERIALS AND METHODS* ..... *46*

**2.1 PCR GENOTYPING** ..... **46**

2.1.1 Genomic DNA extraction ..... 46

2.1.2 Primer design for allele specific PCR ..... 48

2.1.3 PCR/allele-specific PCR for S-genotyping ..... 49

2.1.4. Agarose gel electrophoresis ..... 49

**2.2 GENERAL METHODS FOR RT-PCR CLONING** ..... **50**

2.2.1 RNA extraction ..... 50

2.2.2 RNA quality control and quantification ..... 52

2.2.3 RT-PCR and primers ..... 52

2.2.4 Reverse transcription (RT)-PCR reaction: 3'RACE ..... 54

2.2.5 Reverse transcription (RT)-PCR reaction: 5'RACE ..... 55

**2.3 GENERAL METHODS FOR CLONING S-RNase PCR PRODUCTS ..... 59**

2.3.1 Ligation into pCR®2.1 vector ..... 59
2.3.2 Transformation into TOP10F' chemically competent cells..... 59
2.3.3 Colony screening by PCR ..... 60
2.3.4 Plasmid DNA extraction ..... 60
2.3.5 Glycerol stock preparation of plasmids ..... 61
2.3.6 Restriction enzyme digestion ..... 61
2.3.7. Sequencing of plasmid inserts ..... 61

**2.4. BIOINFORMATICS ..... 62**

2.4.1. Sequence data analysis tools ..... 62

***CHAPTER 3: SELF-(IN)COMPATIBILITY STUDIES IN PETUNIA.... 64***

**3.1 INTRODUCTION ..... 64**

3.1.1 The genus *Petunia* ..... 64
3.1.2 Self-incompatibility in *Petunia* ..... 65
3.1.3 Self-compatibility in *Petunia* ..... 66
3.1.3.1 The molecular basis of self-compatibility in *Petunia* ..... 67

**3.2 MATERIALS AND METHODS ..... 70**

3.2.1 Plant Materials ..... 70
3.2.2 S-allele phenotypic analysis ..... 71
3.2.2.1 Pollinations ..... 71
3.2.2.2 Self-incompatibility phenotype determination ..... 71

**3.3 RESULTS ..... 72**

3.3.1 *Petunia inflata* S-allele characterization ..... 72
3.3.1.1 Pollination and S-phenotype determination ..... 72
3.3.1.2 Cloning and sequencing of *P. inflata* S-alleles: 3'RACE-PCR ..... 73
3.3.2 S-allele characterization for N401 family ..... 77
3.3.2.1 Pollination and S-phenotype determination ..... 77
3.3.2.2 Cloning and sequencing of Mitchell $S_m$-RNase ..... 79
3.3.2.3 S-genotype determination of N401 using PCR technique . 83
3.3.2.4 Confirmation of the segregation of $S_m$-allele with self-compatibility ..... 84

**3.4 DISCUSSION ..... 86**

3.4.1 *Petunia inflata* S-allele characterization ..... 86
3.4.1.1 Pollination and S-phenotype determination for *Petunia inflata* ..... 86
3.4.1.2 Cloning and sequencing of *P. inflata* S-alleles ..... 87
3.4.2 Characterization of S-alleles in N401 family ..... 88
3.4.2.1 Pollination and S-genotype determination for N401 family89

3.4.2.2 Cloning and sequencing of Mitchell S-alleles ..... 89
3.4.2.3 The basis of self-compatibility in Mitchell ..... 91

***CHAPTER 4: SELF-INCOMPATIBLITY STUDIES IN POTATO..... 95***

**4.1 INTRODUCTION ..... 95**
4.1.1 Potato crop ..... 95
4.1.2 Potato breeding ..... 97
4.1.3 Self-incompatibility in potato ..... 100

**4.2 MATERIALS AND METHODS ..... 102**
4.2.1 Plant materials ..... 102
4.2.2 Pollinations ..... 103
4.2.3 S-RNase gene cloning and sequencing ..... 104
4.2.4 Sequence alignment, comparison and phylogenetic analysis of S-RNases ..... 104

**4.3 RESULTS ..... 106**
4.3.2 Confirming the SI status and the compatibility relationships of potato germplasm ..... 106
4.3.3 RT-PCR (3'RACE) of potato pistil RNA ..... 107
4.3.3.1 Sequencing and alignment of the 3'RACE products ..... 109
4.3.4 5'RACE of selected S-RNases ..... 119
4.3.4.1 Sequencing and alignment of the 5'RACE products ..... 120
4.3.5 Phylogenetic analysis of S-RNases ..... 123

**4.4 DISCUSSION ..... 127**
4.4.1 SI status confirmation and the compatibility relationships of the potato stocks ..... 127
4.4.2 S-allele characterization in selected diploid potato plants ..... 128
4.4.3 Solanaceous S-RNases exhibit allelic diversity and intraspecific-sequence polymorphism ..... 132
4.4.4 Phylogenetic analysis of S-RNases ..... 133
4.4.4.1 S-RNases are distinct from S-like RNases ..... 133
4.4.4.2 Phylogenetic analysis of solanaceous S-RNases ..... 134

***CHAPTER 5: S-RNase GENE EXPRESSION ANALYSIS..... 137***

**5.1 INTRODUCTION ..... 137**
5.1.1 S-RNases have unique role in pollen rejection ..... 137

**5.2 MATERIALS AND METHODS ..... 139**
5.2.1 Selection of reference genes for quantitative RT-PCR ..... 139
5.2.2 Semi-quantitative RT-PCR ..... 140
5.2.3 Tissue specificity and developmental expression of the S-RNase gene ..... 141
5.2.4 Real-time quantitative RT-PCR ..... 141

**5.3 RESULTS ..... 143**
5.3.1 Selection of reference genes for quantitative RT-PCR ..... 143

5.3.2 Semi-quantitative RT-PCR.....145
5.3.2.1 Tissue-specificity of the S-RNase gene.....145
5.3.2.2 Developmental expression of the S-RNase gene.....146
5.3.3 Quantitative RT-PCR .....148

**5.4 DISCUSSION.....152**
5.4.1 The choice of housekeeping gene is essential for gene expression studies .....152
5.4.2 S-RNase gene expression is confined to pistil tissues.....155
5.4.3 The S-RNase gene is developmentally expressed.....156
5.4.4 S-RNase gene expression can be variable under different genetic backgrounds .....158

***CHAPTER 6: S-RNase GENE SILENCING: RNAi CONSTRUCT DESIGN .....160***

**6.1 INTRODUCTION.....160**
6.1.1 RNAi in plants: a brief overview.....160
6.1.2 Gateway® cloning: a brief overview.....164

**6.2 MATERIALS AND METHODS.....168**
6.2.1 Dialysis membrane preparation and gel purification .....168
6.2.1.1 Dialysis membrane preparation.....168
6.2.1.2 Electroelution dialysis DNA gel purification and ethanol precipitation .....168
6.2.2 Primers for constructing the chimeric S-RNase gene and the RNAi construct.....170
6.2.3 Propagation/maintenance of gateway compatible vectors.....171
6.2.4 *att*B-PCR products purifications .....172
6.2.5 BP recombination reaction: creating an entry clone.....173
6.2.6 LR recombination reaction: creating an expression clone/RNAi construct .....174
6.2.7 Transformation of RNAi constructs into *Agrobacterium tumefaciens* .....174

**6.3 RESULTS.....175**
6.3.1 Construction of a chimeric S-RNase gene: $S_{o1}/S_{o2}$-RNase.....175
6.3.1.1 $S_{o1}$-vector preparation for constructing the chimeric S-RNase.....178
6.3.1.2 $S_{o2}$-insert preparation for constructing the chimeric S-RNase.....179
6.3.1.3 Ligation of $S_{o1}$-vector and $S_{o2}$-insert .....183
6.3.2 BP recombination reaction: creating an entry clone.....185
6.3.3 LR recombination reaction: creating an expression clone/RNAi construct .....190
6.3.4 Transformation of RNAi construct into *A. tumefaciens* .....195

**6.4 DISCUSSION.....196**

6.4.1 $S_{o1}/S_{o2}$-RNase chimeric gene .....196
6.4.2 RNAi construct design for $S_{o1}/S_{o2}$-RNase gene silencing .....198

***CHAPTER 7: GENERAL DISCUSSION*.....202**

**7.1 PHENOTYPIC CHARACTERIZATION OF S-ALLELES.....203**
7.1.1 Self-incompatibility and compatibility relationships .....203
7.1.2 The analysis of self-compatibility trait in Petunia Mitchell .....204

**7.2 MOLECULAR CHARACTERIZATION OF S-ALLELES .....206**
7.2.1 Degenerate primer approach to S-RNase cloning .....206
7.2.2 Sequence analysis of S-RNase genes.....207
7.2.3 Phylogenetic analysis of S-RNase genes.....209

**7.3 ANALYSIS OF THE S-RNase GENE EXPRESSION .....210**

**7.4 RNAi CONSTRUCT DESIGN FOR GENE SILENCING .....212**

**7.5 TRANSFORMATION OF THE RNAi CONSTRUCT *IN PLANTA* AND IMPLICATIONS FOR FUTURE WORK .....214**

**7.6 CONCLUDING REMARKS.....217**

***REFERENCES*.....219**

***APPENDICES*.....246**

## LIST OF FIGURES

**Figure 1.1** A schematic diagram of the process of pollination, fertilization and germination in angiosperms.....2

**Figure 1.2** Gametophytic self-incompatibility in angiosperms.....7

**Figure 1.3** Sporophytic self-incompatibility in angiosperms.....8

**Figure 1.4** Model for the mechanism of sporophytic self-incompatibility response in the Brassicaceae.....13

**Figure 1.5** Proposed S-RNase degradation model.....25

**Figure 1.6** Proposed modified S-RNase degradation model.....26

**Figure 1.7** Proposed S-RNase compartmentalization model.....28

**Figure 1.8** S-specific pollen rejection models for S-RNase based SI.....31

**Figure 1.9** Model for the mechanism of GSI response in *Papaver rhoeas*.....35

**Figure 1.10** Phylogenetic tree of T2-type RNase in plants.....37

**Figure 2.1** Typical Solanaceae S-RNase structure and 3'RACE strategy.....55

**Figure 2.2** An overview of the 5'RACE strategy.....58

**Figure 3.1** Colony PCR of a typical *P. inflata* pistil S-RNase using M13 universal primers.....74

**Figure 3.2** Alignment of the partial deduced amino acid sequence of *P. inflata* $S_d$-RNase with two closely matching S-RNases from *N. alata* ($S_7$-RNase and $S_{C10}$-RNase) and one from *S. peruvianum* ($S_{p16}$-RNase).....76

**Figure 3.3** Pedigree of N401 family.....77

**Figure 3.4** Colony PCR using M13 universal primers for 3'RACE clones of putative $S_m$-RNase.....80

**Figure 3.5** Colony PCR using M13 universal primers for 5'RACE clones of putative $S_m$-RNase.....81

**Figure 3.6** Alignment of the deduced amino acid sequence of $S_m$-RNase with three other most closely matching S-RNases, $S_X$, $S_o$ and $S_1$ from *P. hybrida*.....82

**Figure 3.7** $S_m$-genotyping using $S_m$-allele specific primers.....84

**Figure 4.1** RT-PCR of OKA 9 diploid potato pistils.....108

**Figure 4.2** Colony PCR of OKA 9 diploid potato pistil S-RNase using M13 universal primers.....109

**Figure 4.3** Alignment of the deduced amino acid sequence of the 16 novel S-RNases cloned from *S. okadae*, *S. phureja* and *S. stenotomum* with one published S-RNase, $S_{11}$ from *Solanum chacoense* (Acc. No: S69589.1).....117

**Figure 4.4** 5'RACE PCR of OKA 1 ($S_{o2}$-RNase).....119

**Figure 4.5** Colony PCR using M13 universal primers of OKA 1 ($S_{o2}$-RNase) 5'RACE.....120

**Figure 4.6** Alignment of the deduced amino acid sequence of the full gene sequence of $S_{o2}$-RNase (S.okaSo2) cloned from *S. okadae* and $S_{s2}$-RNase (S.steSs2) from S. *stenotomum* with three published full length potato S-RNases.....122

**Figure 4.7** Phylogenetic tree of S-RNases and S-like RNases.....124

**Figure 4.8** Phylogenetic tree of S-RNases from Solanaceae.....126

**Figure 5.1** PCR profiles for housekeeping gene evaluation on OKA 1...143

**Figure 5.2** PCR profiles for housekeeping gene evaluation on OKA 3...144

**Figure 5.3** Semi-quantitative RT-PCR profiles for OKA 3 tissues.....146

**Figure 5.4** The floral developmental stages from which pistils were harvested in OKA 1.....147

**Figure 5.5** Semi-quantitative RT-PCR for the developmental stage expression of $S_{o1}$- and $S_{o2}$-RNase in OKA 1 pistils.....148

**Figure 5.6** Relative expression of the S-RNase gene at different pistil developmental stages.....150

**Figure 5.7** QRT-PCR of the S-RNase gene accumulation in the stigma and style.....151

**Figure 5.8** Relative expression of the $S_{o1}$- and $S_{o2}$-RNase gene in different plant backgrounds.....152

**Figure 6.1** RNA silencing mechanism.....161

**Figure 6.2** A schematic of RNAi construct design using Gateway cloning.....167

**Figure 6.3** A schematic of the chimeric $S_{o1}/S_{o2}$-RNase gene construction.....177

**Figure 6.4** Double restriction enzyme digest and purification of $S_{o1}$-RNase.....179

**Figure 6.5** Amplification of $S_{o2}$-RNase using So2NdeI-F1/So2-R1 primer pairs.....180

**Figure 6.6** Colony PCR using M13 universal primers for pCR2.1-So2(NdeI) clones.....181

**Figure 6.7** Double restriction enzyme digests and gel purification of $S_{o2}$-RNase.....183

**Figure 6.8** Confirming the ligation reaction.....184

**Figure 6.9** Colony PCR of the putative chimeric $S_{o1}/S_{o2}$-RNase.....185

**Figure 6.10** attB-PCR .....186

**Figure 6.11** A schematic of the entry clone, pENTR207-So1/So2.....187

**Figure 6.12** Colony PCR of the entry clone.....188

**Figure 6.13** The sequence of the chimeric S-RNase gene in pENTR207-So1/So2.....189

**Figure 6.14** A schematic of the map of primers used for screening the putative pHG8-So1/So2 clones.....190

**Figure 6.15** Colony PCR of the putative RNAi clones.....192

**Figure 6.16** PCR to confirm the selected putative RNAi clones.....193

**Figure 6.17** Restriction enzyme digests to confirm the putative RNAi clones.....194

**Figure 6.18** Map of the designed RNAi construct pHG8-So1/So2.....195

## LIST OF TABLES

**Table 2.1** Primers for genotyping populations segregating for S-RNases....48

**Table 2.2** Sequence of the primers designed and the *Not*I primers used for S-RNase cloning in this study....53

**Table 3.1** Incompatibility groups derived from diallel crosses of *P. inflata*....73

**Table 3.2** Pollination data for N401 family....78

**Table 3.3** S-genotype and pollination data for N401 family....85

**Table 4.1** Diploid potato species and clone designations used for the study....103

**Table 4.2** Pollination data for *S. okadae*....107

**Table 4.3** Summary of the sixteen deduced S-genotype of *Solanum*...111

**Table 4.4** Results of blast searches (NCBI) with cloned S-RNases for *S. okadae*, *S. stenotomum* and *S. phureja*....112

**Table 4.5** S-RNases for potato retrieved from EBI Database....114

**Table 4.6** Percentage amino acid similarity of the sixteen cloned S-RNases....118

**Table 5.1** Housekeeping gene primers evaluated for gene expression studies....139

**Table 5.2** S-RNase allele specific primers (QRT-PCR primers) used for gene expression studies....140

**Table 6.1** Primers for creating chimeric S-RNase and confirming RNAi constructs....171

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ARC | Armadillo repeat containing protein |
| C1-C5 | Conserved domains (1 to 5) of solanaceous S-RNases |
| GSI | Gametophytic self-incompatibility |
| HVa & HVb | Hypervariable domains (a & b) of solanaceous S-RNases |
| PCR | Polymerase chain reaction |
| QRT-PCR | Quantitative real time-polymerase chain reaction |
| RT-PCR | Reverse transcription-polymerase chain reaction |
| RACE | Rapid amplification of cDNA ends |
| RISC | RNA-induced silencing complex |
| RNAi | RNA interference |
| SBP | S-RNase binding protein |
| SC | Self-compatibility |
| SCR | S-locus cysteine rich |
| SFB | S-haplotype-specific F-box |
| SI | Self-incompatibility |
| SLF | S-locus F-box |
| SLG | S-locus glycoprotein |
| SIi | S-locus inhibitor |
| SP11 | S-locus protein 11 |
| S-RNase | S-locus ribonuclease |
| SRK | S-locus receptor kinase |
| SSI | Sporophytic self-incompatibility |
| SSK | SLF-interacting SKP1-like |
| THL | Thioredoxin-h proteins |

# CHAPTER 1: GENERAL INTRODUCTION

## 1.1 Fertilization in flowering plants

Fertilization in plants is the outcome of a series of events that begins following successful pollination, as a result of the landing of a pollen grain on a receptive stigmatic surface and proceeds through a series of developmental stages. During a compatible pollination, the pollen grain hydrates causing a pollen tube to emerge from the grain which grows through the transmitting tissue of the style to reach the ovary to ultimately effect fertilization.

The growing pollen tube consists of a single large cell containing the vegetative nucleus and a generative nucleus which carries the sperm cells in its growing tip. Once the growing pollen tube gets to the ovule, the sperm cells are released and one of these sperm cells fuses with the haploid egg cell to form the embryo and the second sperm cell fuses with the central cells to form the endosperm which serves as the food storage organ for the developing embryo (Dodds *et al.*, 1996). The process of pollination and fertilization through to germination is summarised in Figure 1.1 (Hunter, 2009).

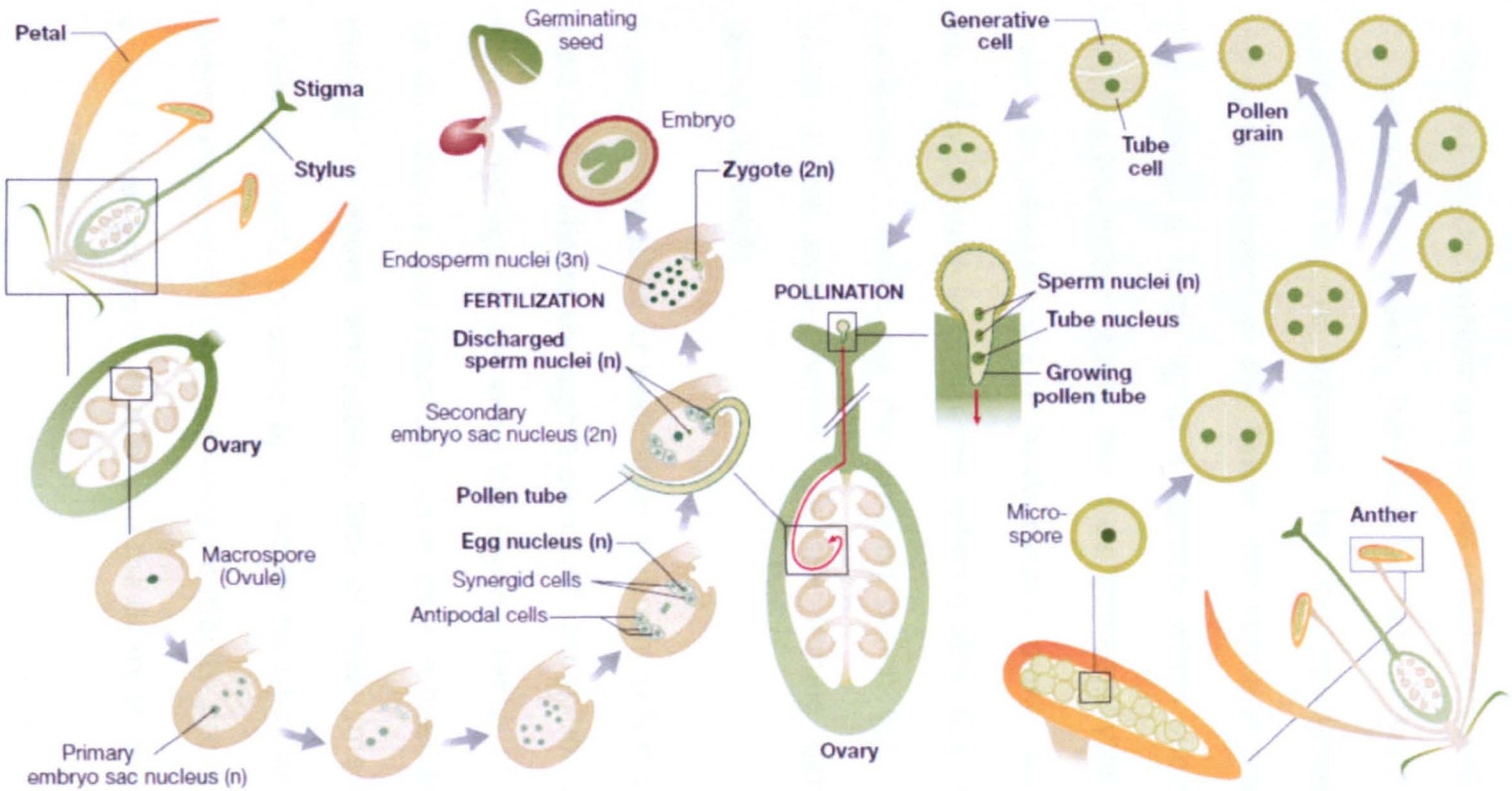



**Figure 1.1 A schematic diagram of the process of pollination, fertilization and germination in angiosperms** (Reprinted from Hunter (2009) by permission from Nature Publishing Group).

Flower formation in angiosperms, is one of the key evolutionary novelties distinguishing them (angiosperms) from all other groups of plants. The flower is the reproductive structure of plants and is made up of four distinct floral organ types which are arranged in four concentric whorls. Arranged on the outside are sepals which serve as a protective cover for the developing flower, followed by petals, serving as attractants for pollinators. This is followed by the specialized male and female reproductive parts of the flower: the stamens and carpels respectively. The stamen is made up of a filament which supports the anther which contains the pollen grains, the male gametophytes, in two pairs of pollen sacs. The carpels (one or more carpels makes up the pistil) consists of the stigma which receives the pollen grains, the style which serves as a transmission organ for the growing pollen tube and the ovary which contains the ovules which are the female gametophytes (Rea and Nasrallah, 2008).

During the process of pollination through to fertilization, the pollen and pistil communicate through a series of recognition systems that allow the pistil to distinguish between the diverse range of pollen types that land on the stigma and therefore allow only compatible pollen grains to go through to effect fertilization. One of these mechanisms employed by flowering plants to serve as a recognition barrier to self fertilization is self-incompatibility (SI). Self-incompatibility prevents genetically related or self pollen grains from reaching the ovary to effect fertilization.

## 1.2 General overview of self-incompatibility

Flowering plants exhibit a wide variation of floral display and design which includes hermaphroditism, monoecism and others that exhibit gender polymorphism with dioecism and gynodioecism being the commonest. The evolution of pollination and mating systems are associated with much of the functional basis of this remarkable floral diversity (Barrett, 1998). The majority of flowering plants, being hermaphroditic in nature, often have their reproductive organs located in close proximity. This feature enables pollen grains from an individual plant to have a high tendency of landing on the stigma of the same flower. One might assume that this feature will impose or promote self pollination on flowering plants leading to self fertilization, a feature which would otherwise have been generally deleterious for angiosperm evolution as a whole. However, flowering plants have evolved several strategies to avoid self fertilization, one of which is the mechanism of self-incompatibility (Gaude *et al.*, 2006).

In one definition of SI, the landing of a pollen grain from a self-incompatible plant onto a receptive stigmatic surface on the same plant, or a genetically related plant, results in a failure of pollen germination or the arrest or seizure of pollen tube growth (de Nettancourt, 1977). The mechanism of self-incompatibility enables the female reproductive organ of a flower, the pistil, to distinguish self pollen from non-self pollen, thereby allowing only the non-self pollen to effect fertilization (Kao and McCubbin, 1996). Self-incompatibility is the most widespread mechanism by which flowering plants prevent inbreeding and promote outcrossing,

thereby ensuring genetic variability within a plant population. The evolution of mechanisms to prevent inbreeding in flowering plants is partly responsible for their reproductive success, thereby making them one of the most successful terrestrial groups of plants (Silva and Goring, 2001).

In many self-incompatible species, it has been found that self-incompatibility is controlled developmentally and is not expressed in immature flowers but only in mature ones. Consequently, the pollination of the pistils of immature buds with mature pollen has been found to result in fertilization and the production of viable seeds. Immature bud pollination has regularly been used for the production of plants homozygous for S-alleles (Bernatzky *et al.*, 1988). This technique has made it practical to overcome self-incompatibility in members of the Solanaceae family leading to the production of homozygous stocks (*e.g.* Robbins *et al.*, 2000) but not feasible in some families *e.g.* Rosaceae.

Flowering plants have evolved several mechanisms of SI. Based on floral morphology, self-incompatibility systems have been classed into two categories; heteromorphic and homomorphic. Heteromorphic SI occurs in flowering plants that produce morphologically distinct flowers characterized by the relative positions of their reproductive organs; while plants with homomorphic SI produce morphologically identical flowers (Bernatzky *et al.*, 1988; Ebert *et al.*, 1989; Kao and McCubbin, 1996). The molecular genetics of the heteromorphic system of self-incompatibility still remains elusive although its biology and genetics have been described in some plant species (*e.g. Primula*). Heteromorphic

SI systems use a combination of floral morphological differences (heterostylous polymorphism) controlled genetically and diallelic sporophytically controlled SI system to pose barriers for self-pollination and subsequently self-fertilization (Barrett, 1998; McCubbin, 2008). Homomorphic systems, on the other hand have been studied extensively and two types have been identified so far; sporophytic self-incompatibility (SSI) and gametophytic self-incompatibility (GSI) (Ebert *et al.*, 1989).

The SI response studied in many plant species to date has been found to be controlled by a single multi-allelic locus, designated the S-locus in the exception of for instance the grasses which have two unlinked with multi-allelic loci (*i.e.* S- and Z-loci). The S-locus contains at least two highly polymorphic genes, called the S-determinant genes, which are expressed in the pollen and pistil (McCormick, 1998). However, recent studies by several authors have revealed that other loci are also involved in the SI response, with the S-locus playing the major role of self and non-self pollen recognition (Gaude *et al.*, 2006).

## 1.3 Homomorphic self-incompatibility systems

The mode of genetic control of the pollen phenotype in self-incompatible species has formed the basis for classifying homomorphic self-incompatibility in flowering plants. The pollen phenotype can be derived sporophytically or gametophytically. For the gametophytic self-incompatibility system (GSI), the pollen incompatibility phenotype is

determined by its own haploid genotype (Figure 1.2) while in the sporophytic self incompatibility system (SSI), the diploid genotype of the pollen producing plant determines the phenotype of the pollen (Figure 1.3) (Silva and Goring, 2001).

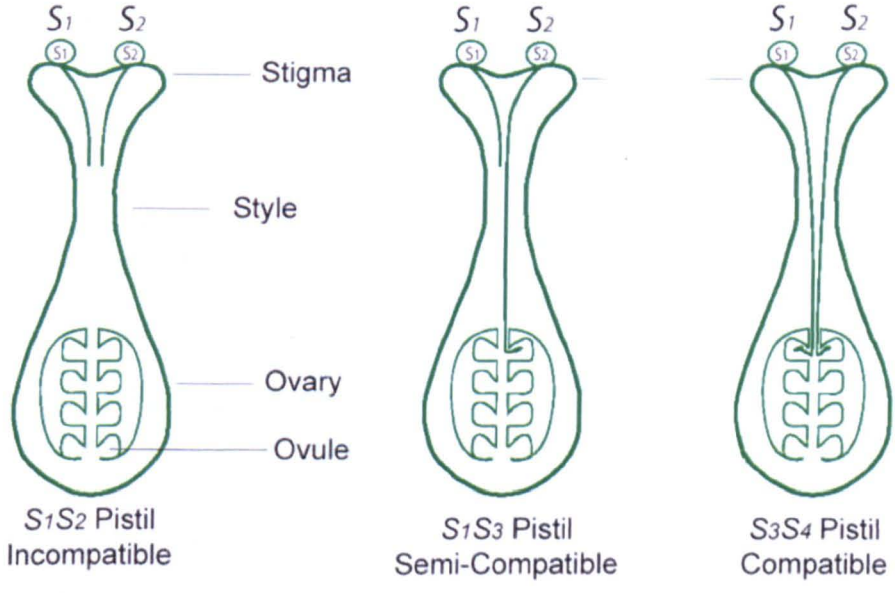



**Figure 1.2 Gametophytic self-incompatibility in angiosperms.** The growth of pollen tube is prevented when the SI determinant of the pollen matches one of the pistil SI determinants. The $S_1S_2$ alleles expressed in the pollen are self-incompatible with the $S_1S_2$ alleles expressed in the pistil (left pistil). On the other hand, the matching of only one of the pollen S alleles, thus $S_1$, with either of those of the pistil, results in half of the pollen grains, the $S_2$ pollen, to be compatible on the $S_1S_3$ pistil (centre pistil). Full compatibility however is observed when the $S_1S_2$ alleles expressed in the pollen are different from the $S_3S_4$ alleles that are expressed in the pistil (right pistil), thereby allowing the $S_1$ and $S_2$ pollen to germinate and penetrate the style to reach the ovary allowing fertilization to occur (Redrawn from Silva and Goring (2001)).

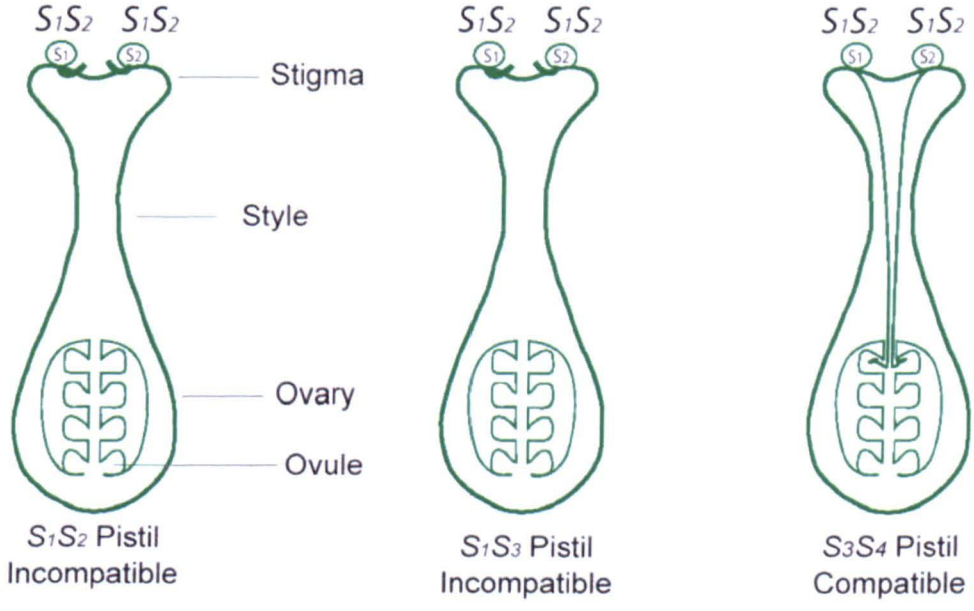



**Figure 1.3 Sporophytic self-incompatiblity in angiosperms.** The matching of one of the S alleles in the pollen producing parent with that of the pistil causes the arrest of pollen germination at the stigmatic surface. Pollen grains expressing the $S_1S_2$ alleles are inhibited on both pistils expressing the $S_1S_2$ allele (left pistil) and the $S_1S_3$ allele (centre pistil) as a result of the alleles that are matching. On the other hand, pollen grains expressing the $S_1S_2$ allele are fully compatible on pistils expressing the $S_3S_4$ allele (right pistil), hence allowing the occurence of fertilization (Redrawn from Silva and Goring (2001)).

### 1.3.1 Sporophytic self-incompatibility (SSI)

Sporophytic self-incompatibility has been studied extensively at the molecular level in the Brassicaceae (Nasrallah and Nasrallah, 1993; Hiscock and Tabah, 2003) although it has also been described in members of the Compositae and Convolvulaceae. The recognition and subsequent arrest of self-pollen in plants that exhibit SSI occurs very rapidly on the stigmatic surface (Nasrallah and Nasrallah, 1993; Silva and Goring, 2001). Unlike the GSI system which is known to operate under

two distinct mechanisms, only one system of operation has been characterized at the molecular level for the SSI system so far (Hiscock and Tabah, 2003). SSI in the Brassicaceae involves cell to cell interactions between the stigmatic papillae and the pollen grain or pollen tube. The self pollen recognition and rejection is marked by the inhibition of pollen hydration, germination or pollen tube penetration through the surface of the stigma (Nasrallah and Nasrallah, 1993; Dickinson, 1995).

In the Brassicaceae, both the male and female SI determinants, which are the pollen and pistil SI determinants respectively, have been identified and characterized. In addition there are a number of downstream genes which play roles in the SI response. The male determinant has been identified as SCR (S-locus cysteine rich protein) and the female determinant as SRK (S-locus receptor kinase) (Silva and Goring, 2001). Initially, the search for the female determinant of SI in the Brassicaceae identified stigma glycoproteins called S-locus glycoproteins (SLGs) which co-segregate with different S-haplotypes. SLG isolation and identification paved the way for the identification of the second S-locus gene, called S-locus receptor kinase (SRK) (Takayama and Isogai, 2005). SRK was found to share extensive sequence homology with SLG (Stein *et al.*, 1991).

Confirmation that the female S-determinant gene was SRK was obtained in a gain of function experiment using transgenic plants of *Brassica rapa* (Takasaki *et al.*, 2000). This study showed that only SRK determines the S-haplotype specificity of the stigma, with SLG playing an enhancing role to the activity of SRK in the SI recognition response. The precise role of

SRK was further confirmed, again in another gain of function experiment, using *Brassica napus* (Silva *et al.*, 2001). However, the enhancing role of SLG was not observed in this case. Varied conclusions have been drawn on the possible role of SLG from studies by several different researchers (Silva and Goring, 2001). There may be variability among different S-haplotypes in SLG requirement during SI response in the Brassicaceae (Silva and Goring, 2001; Takayama and Isogai, 2005).

The male determinant of SSI was identified through extensive sequence analysis of the S-locus region between the SRK and SLG genes of *Brassica rapa*. It was found to encode a small cysteine-rich pollen specific protein designated SP11 (S-locus protein 11) (Suzuki *et al.*, 1999) or SCR (S-locus cysteine rich) (Schopfer *et al.*, 1999). This gene (SP11/SCR) has been shown to encode the Brassicaceae pollen S determinant by functional analysis in transgenic plants (Schopfer *et al.*, 1999; Suzuki *et al.*, 1999; Takayama *et al.*, 2000).

The SI response in the Brassicaceae occurs as a result of S-haplotype specific ligand-receptor interaction between the SCR/SP11 and SRK. It is presumed that, the SCR/SP11 acts as pollen-borne ligand thereby activating the stigmatic SRK receptor by binding to its extracellular domain in an S-haplotype specific manner during a self-incompatible pollination. The activation of SRK then causes the initiation of a signalling cascade in the stigmatic papillae, thereby leading to the eventual rejection of the incompatible pollen. The activation of SRK may also be responsible for the activation of one or more other signalling pathways which may ultimately be responsible for the blocking of vital steps, such

as the release of water for pollen hydration and the release of stigmatic papillar cell wall-degrading enzymes which are required for the growth of the pollen tube (Silva and Goring, 2001).

In an attempt to elucidate, the downstream signalling pathway(s) involved in the cellular responses in the pistil leading to the rejection of self pollen, a yeast two-hybrid assay was used to screen for proteins that interact with the SRK kinase domain (Goring, 2000). These approaches led to the identification of several proteins and other signalling molecules including both negative and positive regulators of the signalling pathway (Silva and Goring, 2001). Through a yeast-two hybrid screen system using the kinase domain of SRK as a bait, THL1 and THL2 which are two stigmatic thioredoxin-h proteins were identified as binding partners with SRK. SRK and THL1/2 interactions were shown to be independent of phosphorylation and required a conserved cysteine residue at the SRK transmembrane domain (Bower *et al.*, 1996; Mazzurco *et al.*, 2001). THL1 and THL2 are thought to be acting as negative regulators of the kinase activity of SRK *in vitro* (Cabrillac *et al.*, 2001). The analysis of antisense THL1 and THL2 transgenic plants of *B. napus* (Haffani *et al.*, 2004) lend credence to this assumption. Results from the experiment of Haffani *et al.* (2004) showed that, the suppression of THL1/2 transcripts in stigmatic tissues resulted in a low level or weak constitutive rejection of the pollen grains of *B. napus* which are normally otherwise self compatible. THL1/2 may therefore be functioning to keep SRK in an inactive state by possibly interacting with SRK upstream of the SCR-SRK complex (Kaothien-Nakayama *et al.*, 2010).

ARC1 (Armadillo Repeat Containing protein) has also been identified to be interacting with the SRK kinase domain in a yeast-two hybrid screen. In contrast to THL1/2, ARC1 and SRK interaction was found to be dependent on phosphorylation. ARC1 specifically binds to the phosphorylated kinase domain of SRK leading to the phosphorylation of ARC1 by SRK (Gu *et al.*, 1998). ARC1 is a positive regulator of the Brassicaceae SI signalling pathway. When the ARC1 transcripts were suppressed in the pistils of self-incompatible *B. napus* plants, the resulting transgenic plants were found to have a partial breakdown in SI response (Stone *et al.*, 1999). The structural domains of ARC1 implicate it as an E3-ubiquitin ligase, which is an enzyme involved in the process of ubiquitination. ARC1 has been confirmed to function as an E3-ubiquitin ligase promoting stigmatic protein ubiquitination during self-incompatible response leading to the rejection of self pollen grains (Stone *et al.*, 2003). ARC1 could therefore be functioning downstream of the SCR-SRK complex by degrading papillar proteins during the SI response through ubiquitination and proteasome degradation (Stone *et al.*, 2003). Figure 1.4 shows a model for the SSI response in the Brassicaceae.

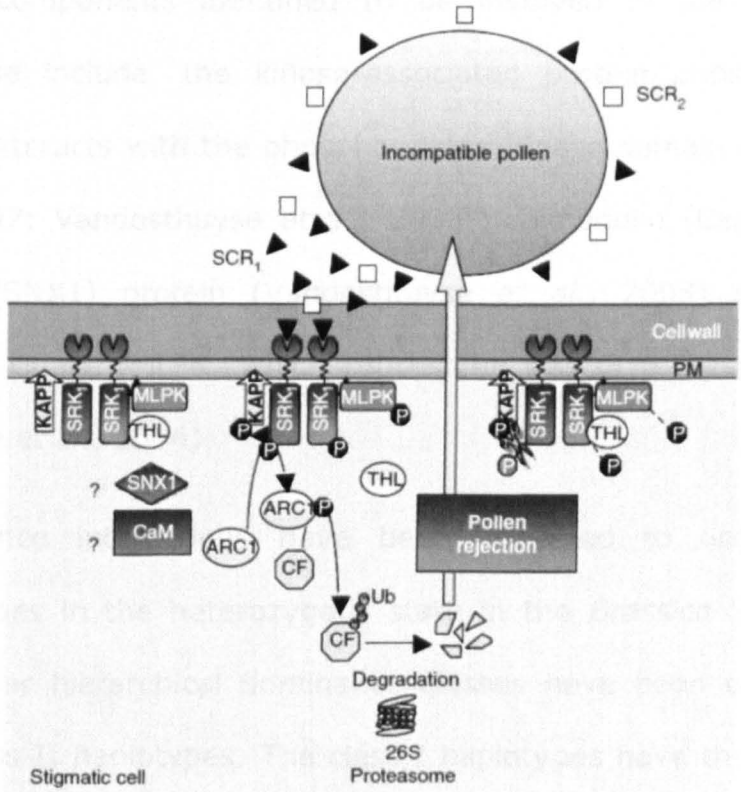



**Figure 1.4 Model for the mechanism of sporophytic self-incompatibility response in the Brassicaceae.** During self pollination reaction, SCR moves from the surface of pollen grains onto the stigmatic papillar plasma membrane. SCR peptides having the same S-allele specificity ($S_1$), corresponding to that of SRK interacts with extracellular domain of SRK. SRK in the basal state is inactivated by THL. However, the interaction between SRK1/SCR1 leads to the release of the negative regulation of THL thereby activating the kinase domain of SRK. ARC1, an E3-ubiquitin ligase is then phosphorylated by SRK. The phosphorylated ACR1 has been proposed to then recruit stigmatic compatibility factors (CFs) which promotes their polyubiquitination and degradation through the 26S proteasome. Following the degradation of these still elusive CFs, the growth of pollen tubes is arrested and the subsequent failure of self-fertilization occurs. MLPK is assumed to be involved in the signalling cascade leading to the degradation of the CFs. KAPP might be involved in regulating the signalling pathway negatively by dephosphorylating SRK. The role of CaM and SNX1 is yet to be ascertained but they might also be serving as negative regulators of the signalling pathway based on to their role in down-regulating receptor kinases in animals (Reprinted from Gaude *et al.* (2006) by permission from CAB International).

Other components identified to be involved in the Brassicaceae SI response include, the kinase-associated protein phosphatase (KAPP), which interacts with the phosphorylated kinase domain of SRK (Braun *et al.*, 1997; Vanoosthuyse *et al.*, 2003), calmodulin (CaM) and a sorting nexin (SNX1) protein (Vanoosthuyse *et al.*, 2003) and the M-locus protein kinase (MLPK) which causes the loss of SI response in *B. rapa* (Murase *et al.*, 2004).

Dominance interactions have been observed to occur between S-haplotypes in the heterozygous state in the *Brassica* SSI system. Two non-linear hierarchical dominance classes have been observed: class I and class II haplotypes. The class I haplotypes have the tendency to be dominant over the class II haplotypes, and conferring stronger SI phenotypes than the class II haplotypes. Seed production could be seen following self pollination in the class II types. Also the class II haplotypes are recessive to the class I haplotypes. Although co-dominance occurs among S-haplotypes in the stigma, the occurrence of dominant interactions in anther tissues may also be observed. These independent hierarchies of dominance relationships existing in the stigma and anther (pollen) may be an indication that allelic interactions operate using different mechanisms in the two organs (Gaude *et al.*, 2006).

In the pollen, dominant relationships are regulated at the mRNA level for SCR alleles. The pollen recessive allele expression is drastically down regulated in heterozygotes having a dominant allele. The inhibition of the expression of SCR recessive allele has been observed in both *Arabidopsis lyrata* and *Brassica* (Kusaba *et al.*, 2002; Shiba *et al.*, 2002) indicating

the occurrence of a common mechanism controlling allelic interaction in the pollen for both species (Gaude *et al.*, 2006). In the stigmas however, the expression level of SRK by itself was not able to give an explanation or account for the dominance relationships in *Brassica* and in *A. lyrata* (Hatakeyama *et al.*, 2001; Kusaba *et al.*, 2002). Therefore, the control of allelic interactions in the stigma may be operating with a different mechanism (Gaude *et al.*, 2006). Recent studies have implicated, a *trans*-acting small non-coding RNA (sRNA) termed *Smi* (SP11 methylation inducer) and situated in the region flanking the dominant SP11-allele to be responsible for the dominant/recessive relationship determination in *Brassica* by transcriptionally silencing the recessive SP11-allele (Tarutani *et al.*, 2010).

### 1.3.2 Gametophytic self-incompatibility (GSI)

Gametophytic self-incompatibility represents the most prevalent form of self-incompatibility existing in more than 60 flowering plant families. The Solanaceae, Rosaceae, Plantaginaceae (previously called Scrophulariaceae), Leguminoceae, Onagraceae, Papavaraceae and Poaceae are among the plant families exhibiting this form of self-incompatibility (Franklin *et al.*, 1995; Kao and McCubbin, 1996). The extensive study of GSI at the molecular level revealed that it operates by two different mechanisms to achieve self pollen recognition and rejection. One of these is the stylar ribonuclease (S-RNase) mechanism which has been initially identified and characterized in members of the Solanaceae, and later in the Rosaceae and Plantaginaceae. The other mechanism is

found in the Papavaraceae, in particular *Papaver rhoeas* (Franklin-Tong and Franklin, 2003). There are likely to be more distinct GSI systems as yet undiscovered (McClure and Franklin-Tong, 2006). The identification of the elusive pollen S-locus component has been one of the primary goals of research groups working on the two GSI systems (Franklin-Tong and Franklin, 2003).

#### *1.3.2.1 S-RNase based (Solanaceae type) self-incompatibility system*

The S-allele products in the styles of the Rosaceae, Plantaginaceae and Solanaceae were identified to be basic glycoproteins having ribonuclease activity responsible for blocking the growth of incompatible pollen tubes in the style (Franklin-Tong and Franklin, 2003). These S-locus encoded glycoproteins were first identified in the Solanaceae and found to be approximately 32kDa. These glycoproteins were found to be ribonucleases (RNases) and were later proven to have ribonuclease activity (Bredemeijer and Blaas, 1981; Anderson *et al.*, 1986; McClure *et al.*, 1989). It has been generally accepted that the S-RNase gene controls the female specificity by degrading ribosomal RNA (rRNA) or messenger RNA (mRNA) in self-incompatible tubes thereby inhibiting fertilization (McClure *et al.*, 1989; Franklin-Tong and Franklin, 2003).

The majority of wild species in the Solanaceae family including *Petunia, Nicotiana, Lycopersicon* and *Solanum* are self-incompatible although some self-compatible variants could be observed. *Petunia inflata,*

*Lycopersicon peruvianum*, *Nicotiana alata*, and *Solanum chacoense* have contributed enormously towards much of the molecular data available on the gametophytic system of self-incompatibility (McCubbin and Kao, 2000).

Molecular analyses of S-RNases in the Solanaceae has identified five conserved domains (C1, C2, C3, C4, and C5) and two hypervariable regions (HVa and HVb) located between the C2 and C3 domains (Ioerger *et al.*, 1991). However, in the Rosaceae S-RNases, the C4 domain found in the Solanaceae is absent. In addition only one hypervariable region (RHV) has been identified in the Rosaceae, unlike two such hypervariable regions (HVa and HVb) found in the Solanaceae (Takayama and Isogai, 2005).

The conserved domains of S-RNases were initially predicted to be responsible for the catalytic activity of S-RNases whilst the HV domains were predicted to be responsible for the allelic-specificity determination of S-RNases (Kaothien-Nakayama *et al.*, 2010). Using X-ray crystallographic analysis to deduce the three dimensional structure of solanaceous and rosaceous S-RNases, it was revealed that, the HV regions could be responsible for S-RNase specificity determination because they are exposed on the surface of S-RNases and therefore could possibly be interacting with the male determinant of the SI response (Ida *et al.*, 2001; Matsuura *et al.*, 2001).

However, the precise role of the HV regions in the S-RNase specificity determination was tested in a domain-swapping experiment (Matton *et*

*al.*, 1997). In their experiment, they use two very closely similar S-RNases, $S_{11}$- and $S_{13}$-RNase of *Solanum chacoense*. These two S-RNases differ by only ten amino acid residues, four of which are located in the HV region. The transformation of plants with $S_{11}$-RNase chimeric gene constructs having the four variable amino acid located in its HV region swapped with those four found in the $S_{13}$-RNase resulted in transgenic plants which have gained $S_{13}$-RNase specificity.

Although this study by Matton *et al.* (1997), showed that HVs are necessary and sufficient for controlling S-RNase specificity, it contradicts previous work that concluded that, HV regions are necessary but not sufficient for controlling S-RNase specificity (Kao and McCubbin, 1996). Transgenic plants transformed with a chimeric construct of $S_3$-RNase having the HV region of $S_1$-RNase, aside losing the ability to reject $S_3$ pollen, also failed to gain the ability to reject $S_1$ pollen, although the chimeric S-RNases were observed to have normal levels of ribonuclease activity (Kao and McCubbin, 1996). Regions outside the HV and the conserved regions might also play a role in the S-allelic specificity determination of S-RNases (Verica *et al.*, 1998).

The C2 and C3 domains in the Solanaceae S-RNases have been found to share a very high degree of sequence homology with RNase T2 and RNase Rh, which are fungal RNases. However, the S-RNases were found to be more similar to RNase T2 than the RNase Rh. Two conserved catalytic histidine residues are shared by both the S-RNases and the fungal RNases (Ioerger *et al.*, 1991).

The deduced amino acid sequences of the identified S-RNases in different solanaceous species shows a high level of divergence with amino acid sequence similarity ranging from as low as 38% to as high as 98% (McCubbin and Kao, 2000). The highest sequence diversity can be found in the two hypervariable regions (Ioerger *et al.*, 1991).

*1.3.2.1.1 Identity and function of the pollen S-gene*

Although the identity of the pistil S-gene has long been known in the S-RNase system, it was not until relatively recently that the elusive pollen component was identified. The identified pollen S-gene was initially doubted because of insufficient polymorphism shown between S-haplotypes to determine S-specificity (McClure and Franklin-Tong, 2006). The pollen S-determinant has now been identified as an F-box protein called *SLF* (S-locus F-box) or *SFB* (S-haplotype-specific F-Box) gene and was shown to be the best candidate for the pollen S-gene (Lai *et al.*, 2002; Entani *et al.*, 2003; Ushijima *et al.*, 2003). This was later proven indisputably to be the pollen S-gene by analysis of transgenic plants (Sijacic *et al.*, 2004) or mutants (Ushijima *et al.*, 2004; Sonneveld *et al.*, 2005).

SLF/SFBs are assumed to interact with S-RNases to elicit the SI response. However, in a compatible pollination, SLFs have been proposed to interact with S-RNases thereby targeting them for degradation in an attempt to protect pollen tubes from the cytotoxic effect of S-RNases. For instance, in a compatible pollination involving *Antirrhinum*, an

increase in stylar protein ubiquitination and a subsequent decrease in S-RNase was observed (Qiao *et al.*, 2004). However on the contrary, in a study involving *Nicotiana*, S-RNase amounts were found to be stable following a compatible pollination (Goldraij *et al.*, 2006).

Despite the fact that, SLF was identified to be the pollen-S gene, it puzzles researchers as to how a gene with such low sequence diversity will be able to identify the pistil S-gene which has much higher sequence diversity and differentiating self from non-self S-RNases (Newbigin *et al.*, 2008). The pollen S-gene could therefore possibly be a complex of genes. Consistent with this, in apple (*Malus domestica*) and also in Japanese pear (*Pyrus pyrifolia*) multiple F-box genes which could be putative pollen-S gene were identified and named SFBBs (S-locus F-box brothers) (Sassa *et al.*, 2007). Recently many more of such multiple and related SFBB genes were identified in apple and were shown to have linkage with the S-RNase gene, have S-haplotype specific polymorphism and are expressed specifically in the pollen (Minamikawa *et al.*, 2010). Contrary to the initial thought that the SLF/SFBs in each S-haplotype of plants in the Plantaginaceae, *Prunus* (subfamily Prunoideae of Rosaceae) and Solanaceae are single copy whilst those of the Maloideae of Rosaceae, apple and Japanese pear are multiple copies (Minamikawa *et al.*, 2010), and although SLF/SFB-like genes have been identified and linked to the S-locus of self-incompatibility species in the Solanaceae, Plantaginaceae and Rosaceae (Entani *et al.*, 2003; Zhou *et al.*, 2003; Wheeler and Newbigin, 2007; Sassa *et al.*, 2007; Minamikawa *et al.*, 2010), it has recently been shown that the pollen-S gene of *Petunia* (of the

Solanaceae) is made up of multiple functional SLF genes (as discussed later below) (Kubo *et al.*,2010).

*1.3.2.1.2 Modifier genes involved in S-RNase-based GSI system*

Although the functional relationship between S-RNases and SLFs still remains inconclusive, the elucidation of other components involved in eliciting the SI response in the S-RNase based SI system will provide the impetus to the understanding of these interactions (Kaothien-Nakayama *et al.*, 2010). These components, termed modifiers (modifier genes), were found to be unlinked to the S-locus and could be affecting either the pistil component or the pollen component of the SI response.

The two pistil modifier genes identified so far are the HT-B protein and the 120kDa protein. In a differential screen used to identify stylar genes that are expressed during a self-incompatible reaction in *Nicotiana*, a gene designated *HT-B* was identified. HT-B was identified to be a small asparagine-rich protein expressed in the late development of the style (McClure *et al.*, 1999). The down-regulation of HT-B protein using anti-sense and RNAi technique resulted in transgenic plants that lost the ability to reject self-pollen, indicating a possible role of *HT-B* in eliciting the self-incompatibility response (McClure *et al.*, 1999; O' Brien *et al.*, 2002). The 120kDa glycoprotein, an abundant protein was found in the extra cellular matrix (ECM) of the style where it is taken up by the growing pollen tubes (Lind *et al.*, 1994, 1996). Using an S-RNase affinity column, the 120kDa glycoprotein was found to directly interact with S-

RNases (Cruz-Garcia *et al.*, 2005). Both proteins accumulate with S-RNases in a vacuole-like compartment in compatible pollen tubes. The 120kDa protein can also be detected during the early stages of the SI response in incompatible pollen tubes (Goldraij *et al.*, 2006).

Although the precise role of both HT-B and the 120kDa glycoprotein in the SI response still remains uncertain, their participation in the SI response was confirmed using transgenic experiments. The antisense down-regulation of the HT-B protein in *Nicotiana* (McClure *et al.*, 1999) and *Solanum* (O'Brien *et al.*, 2002) produced transgenic plants that have lost the pistil SI function. Similar observations were made when the 120kDa protein in *Nicotiana* was down-regulated using RNAi (Hancock *et al.*, 2005). The role of HT-B protein and the 120kDa protein in the self-incompatibility response was further elucidated in an immunolocalisation experiment. In *Nicotiana*, HT-B protein degradation was found to occur in cross pollen tubes whilst the co-localisation of the 120kDa protein was found in the vacuolar compartments in which S-RNases exist (Goldraij *et al.*, 2006). Also the self-compatible plants arising from the *HT-B* antisense and the *120kDa* RNAi plants were observed to have resulted from prolonged S-RNase compartmentalizations. HT-B and 120kDa protein might therefore be playing a facilitating role in the events involving the compartmentalization of S-RNases within the growing pollen tubes (Goldraij *et al.*, 2006; Kaothien-Nakayama *et al.*, 2010).

The modifier genes identified on the pollen SI determinant side include the SSK1 (SLF-interacting SKP1-like1) and SBP1 (S-RNase Binding Protein 1). A yeast two hybrid system screening of *A. hispanicum* pollen

cDNA library using AhSLF-$S_2$ as bait, resulted in the identification of SSK1, which was found to be a homolog of SKP1 (Huang *et al.*, 2006). In a pull-down assay, the AhSSK1 protein was reported to be possibly serving as a link connecting AhSLF to CUL1 protein. CUL1 is a component of the SCF complex involved in ubiquitination. The SCF complex consists of four subunits namely; SKP1, CUL1, F-box protein and a RING-finger protein (Hua and Kao, 2006). Hence SLF and SSK1 are proposed to be involved in the SCF complex leading to the ubiquitination of S-RNases (Zhang and Xue, 2008).

SBP1 was identified during the screening of a yeast two-hybrid library from *Petunia* mature pollen (Sims and Ordanic, 2001). Similar approaches were used to identify the SBP1 homolog in *S. chacoense* (O' Brien *et al.*, 2004). Sequence analysis revealed that, all the identified SBP1 proteins contain RING-finger domains which could possibly be functioning as E3-ubiquitin ligase (Hua and Kao, 2006). SBP1 could therefore be involved in the ubiquitination and degradation of S-RNases. This possible role of SBP1 in S-RNase ubiquitination and degradation has been shown, where PiSBP1 was found to be interacting with the following; PiCUL1, a novel ubiquitin-conjugating enzyme and a novel E3-ligase complex (Sims, 2005; Hua and Kao, 2006).

*1.3.2.1.3 Proposed models for the S-RNase-based GSI system*

Different models have been proposed in an attempt to elucidate the interaction between the male determinant and the female determinants

of the SI response in the S-RNase based GSI system following the discovery that S-RNases are responsible for pollen tube inhibition. The first of these models initially proposed that S-haplotypes encode distinct S-RNases that are secreted into the extracellular matrix (ECM) of the style. These S-RNases are then released into pollen tubes where they exert their cytotoxic activity resulting in RNA degradation in the growing pollen tube in an incompatible pollination causing failure of the expression of pollen genes leading to the eventual arrest of pollen tube growth (Luu *et al.*, 2000; Franklin-Tong and Franklin, 2007).

However, following the discovery of the pollen determinant gene (SLF) of the SI reaction, to be a member of the F-box proteins (which are known to be involved in polyubiquitination and subsequent degradation by the 26S-proteasome) led to the proposal of a new model; S-RNase degradation model (Hua and Kao, 2006, 2008; Hua *et al.*, 2008). In this model, S-RNases were proposed to be taken up into the cytoplasm of the growing pollen tubes where they interact with SLF. During a compatible pollination, this S-RNase/SLF interaction will lead to ubiquitination and degradation of the non-self S-RNases by the 26S-proteasome thereby inhibiting their (S-RNases) cytotoxic activity in the growing pollen tubes. Conversely, in an incompatible pollination, the interaction will have no effect on the S-RNases leaving them to execute their cytotoxic activity on growing pollen tubes leading to their eventual arrest (Figure 1.5) (Franklin-Tong and Franklin, 2007).

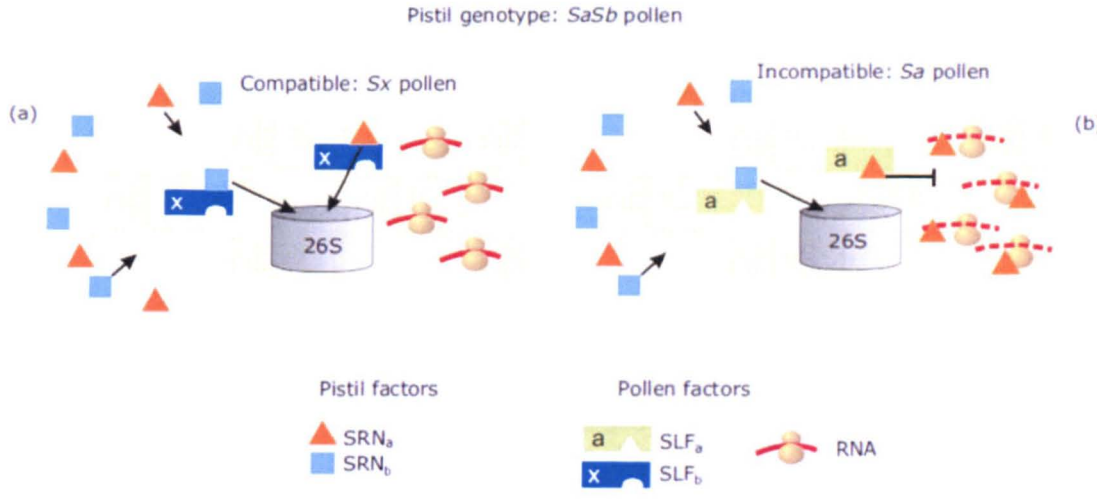



**Figure 1.5 Proposed S-RNase degradation model.** Compatible pollinations ($S_x$-, left) and incompatible pollinations ($S_a$-, right) are shown on the pistil of S-genotype, $S_aS_b$. S-RNases enter into the pollen tube cytoplasm from the extra cellular matrix of the style (arrows). (a) During a compatible pollination, non-self S-RNase and SLF interacts resulting in ubiquitination and degradation of the S-RNases by the 26S proteasome. This eliminates the cytotoxic activity of the S-RNases leading to the continued growth of pollen tubes. (b) During a self-incompatible pollination however, the interaction between self- S-RNase and SLF does not lead to the degradation of S-RNases. This leaves the S-RNases free to exert their cytotoxic activity on pollen tubes leading to their inhibition (Redrawn from Franklin-Tong and Franklin (2007)).

The observation of preferential and strong interaction of SLFs with non-self S-RNases rather than with its self S-RNases has lent further support to the predictions made by the S-RNase degradation model that non-self S-RNases are targeted by SLFs for ubiquitination and degradation (Hua and Kao, 2006; Hua *et al.*, 2008). In an attempt to identity the unique features of SLF that allow it to function in the self-incompatibility response, Hua *et al.* (2007) identified three SLF-specific functional

domains (FD) namely FD1, FD2 and FD3. This led to the modification of the S-RNase degradation model to reflect all the identified interacting functional domains (Figure 1.6) (Hua *et al.*, 2008).

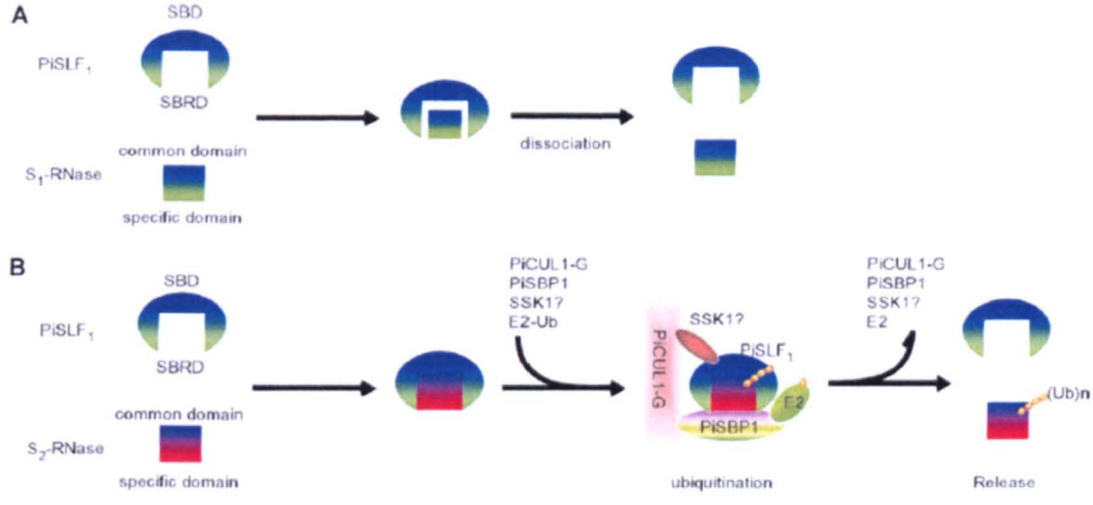



**Figure 1.6 Proposed modified S-RNase degradation model.** (a) During a self-incompatible reaction the interaction between PiSLF$_1$ and S$_1$-RNase results in the formation of an unstable complex. This is due to the weakening of the otherwise strong interaction between the SBD and the common domain of all S-RNases caused by interaction between the SBRD of PiSLF$_1$ and self S$_1$-RNase via its matching S-allele specificity domain. This unstable complex then dissociates thereby releasing the self S$_1$-RNase allowing it to exert its cytotoxic activity leading to the inhibition of pollen tube growth. (b) In a self-compatible reaction on the other hand, since the SBRD of PiSLF$_1$ and the S-allele specific domain of S$_2$-RNase (or any other non-self S-RNase) do not match, the SBD domain of PiSLF$_1$ will interact with S$_2$-RNase via the common domain of all S-RNases leading to the formation of a stable complex. This will cause further interaction with PiSBP1, PiCUL1-G and possibly SSK1 thereby forming an E3-like complex which mediates the ubiquitination and the subsequent degradation of the non self S$_2$-RNase by the 26S proteasome. This eliminates the cytotoxic activity of the S-RNase leading to the continued growth of pollen tubes (Reprinted from Hua *et al.* (2008) by permission from Oxford University Press).

Using an *in vitro* binding assay to examine the S-RNase binding properties of each of the functional domains, it was observed that FD2 is the primary functional region used for the SLF/S-RNase interaction with FD1 and FD3 acting negatively to regulate these interactions (*i.e.* between the FD2 domain of the SLF and the S-RNase) (Hua *et al.*, 2007). Based on these observations, it was proposed that the FD2 domain could be functioning as the S-RNase-binding-domain (SBD) which interacts with a common domain on all S-RNases. In addition, FD1 and FD3 are the S-RNase-binding-regulating-domains (SBRD) which function to weaken the strong FD2 interaction with S-RNase during self-interactions and also determine S-allele- specificity of SLFs (Hua *et al.*, 2008).

Recent identification of components (*e.g.* HT-B) other than the S-locus genes involved in the SI response, has led to the proposal of yet another new model by Goldraij *et al.* (2006), called the S-RNase compartmentalization or sequestration model (see Figure 1.7; Franklin-Tong and Franklin, 2007). This model is based on observations made using immunolocalization where S-RNases have been shown to be compartmentalized into vacuoles and are only released in a self-incompatibility reaction through the breakdown of the compartment in the presence of HT-B protein (Goldraij *et al.*, 2006, see section 1.3.2.1.2). The released S-RNase then exerts its cytotoxic activity on pollen tubes resulting in their subsequent arrest. Whereas in a self-compatible reaction, HT-B degradation could be observed and S-RNases remain sequestered in their vacuolar compartment, hence are not able to

cause the arrest of pollen tube growth (Goldraij *et al.*, 2006; Franklin-Tong and Franklin, 2007; Kaothien-Nakayama *et al.*, 2010).

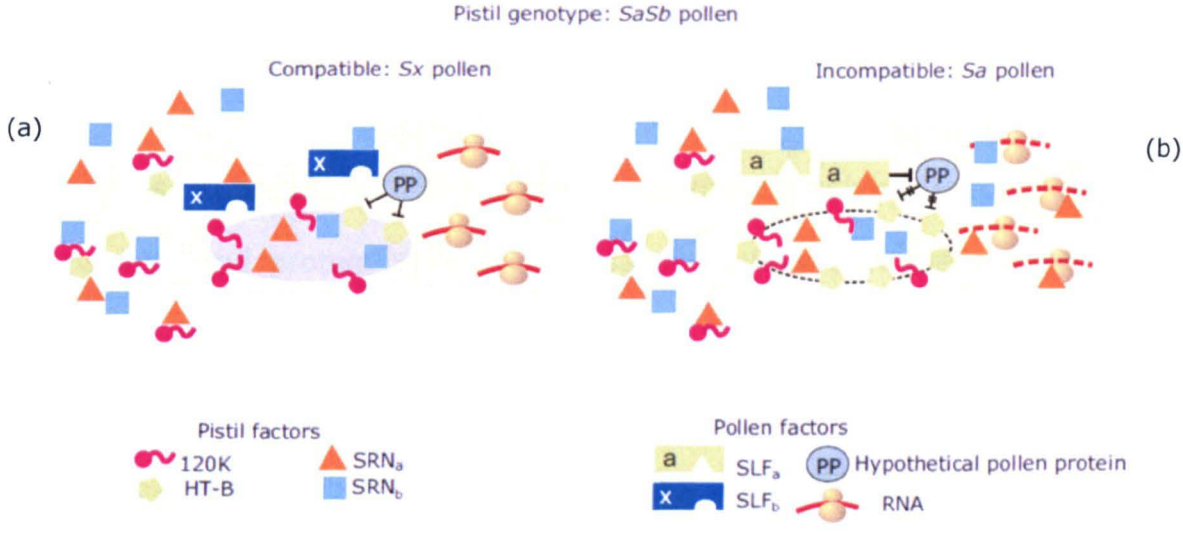



**Figure 1.7 Proposed S-RNase compartmentalization model.** Compatible pollinations ($S_x$-, left) and incompatible pollinations ($S_a$-, right) are shown on the pistil of S-genotype, $S_aS_b$. Through endocytosis, S-RNases, 120kDa and HT-B are taken up into pollen tubes and are sorted into vacuoles. (a) During a compatible interaction, HT-B degradation catalyzed by a hypothetical pollen protein (PP) occurs. S-RNases then remain sequestered in their compartments and are not free to exhibit their cytotoxic activities. However, how the S-RNase gains access to SLF still remains a mystery. (b) During an incompatible interaction, the vacuolar compartment containing S-RNases is degraded mediated by HT-B. This leads to the release of S-RNase into the cytoplasm of the growing pollen tubes. The S-RNase then degrades RNA in the pollen tube leading to their arrest (Redrawn from Franklin-Tong and Franklin (2007)).

SLFs are known to have much lower sequence diversity compared to their widely diverse S-RNases counterparts thereby raising an important question as to how a gene (SLF) showing such low allelic sequence diversity will be able to recognize such a large range of non-self RNases thereby allowing for the occurrence of cross-compatible pollinations (Kubo *et al.*, 2010). In an attempt to answer some of these puzzling questions, Kubo *et al.* (2010) in an experiment involving *Petunia*, used transgenic and protein interaction assays to show that the role of the pollen SI determinant is fulfilled by at least three types of divergent SLF products. Each of these three SLFs functions to recognize a subset of non-self S-RNases. The authors therefore concluded that the pollen determinant of SI is made up of multiple types of SLF products and proposed a new model for the S-RNase based GSI system called "collaborative non-self recognition" model. This model proposes that, each type of pollen-S product (SLF) (in *Petunia*) interacts with a subset of non-self S-RNases within an S-haplotype. Also the products of the multiple types of SLFs including those yet to be identified and characterized are needed for the recognition and detoxification of the entire suite of non-self S-RNases (Kubo *et al.*, 2010). Multiple *SLF/SFB* genes and single *S-RNase* gene systems have been found at the S-locus in other species in the Solanaceae, Plantaginaceae and Rosaceae family (Entani *et al.*, 2003; Zhou *et al.*, 2003; Wheeler and Newbigin, 2007; Sassa *et al.*, 2007; Minamikawa *et al.*, 2010).

The increased array of *SLF* genes constituting pollen-S would be of advantage because this could lead to an increasing number of potential

mating partners hence enabling the recognition and inactivation of more non-self S-RNases by pollen. On the contrary, however, an increase in *S-RNase* gene divergence could cause new S-RNases not to be detoxified by the existing array of *SLF* gene products (Kubo *et al.*, 2010). Although a lot is known about the S-RNase based GSI system, the precise interaction between S-RNase and SLF still remains elusive or inconclusive. The observations and implications of multiple SLF products in SI response by Kubo *et al.* (2010) in *Petunia*, has provided a new and fresh perspective and dimension on the molecular basis of S-RNase based GSI system in general. This would soon provide the thrust for a better understanding of the interaction between self/non-self S-RNases with SLFs thereby resulting in either a compatible or incompatible pollination.

Based on the observations of Kubo *et al.* (2010) and other findings, McClure *et al.* (2011) has revised the S-RNase degradation model and the S-RNase compartmentalization or sequestration model (described previously above and depicted in Figure 1.5) to include the involvement of the multiple SLF genes. In their revised model for the S-RNase degradation model as shown in Figure 1.8 (left) the collaboration of multiple SLFs (SLF1, red; SLF2, yellow; SLF3, black) enables the provision of resistance to $S_x$-RNase. A unique range of allelic *SLF* genes for instance $SLF1_x$ and $SLF1_y$ are expressed by different S-haplotypes. In such instances, the binding of a given S-RNase may or may not occur by individual SLF proteins. In a compatible pollination (in the model, Figure 1.8, top left) the binding of $S_x$-RNase occurs thereby preventing its degradation by only $SLF1_y$ and $SLF2_y$.

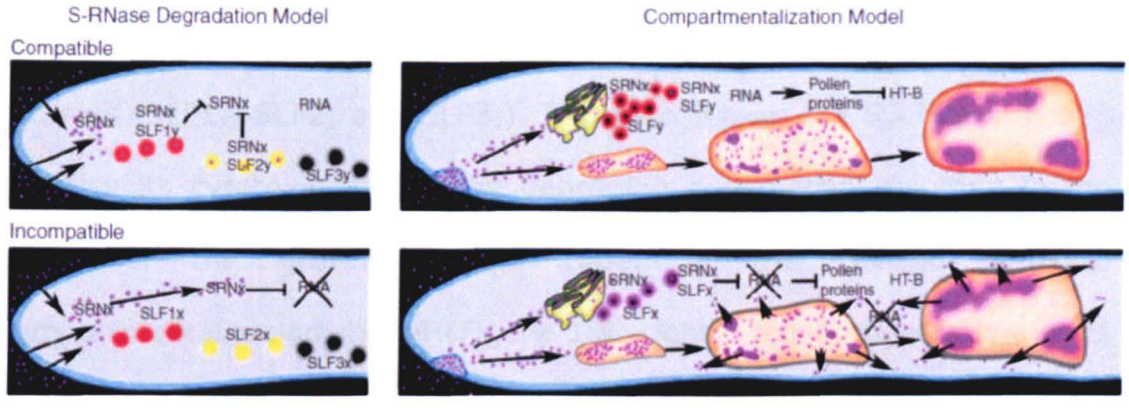



**Figure 1.8 S-specific pollen rejection models for S-RNase based SI.** Pollen tubes containing a single S-RNase ($SRN_x$, purple colour) are shown in the pistil ECM (although two S-RNases would be present in a typical S-heterozygote. Compatible pollination (top figure) is shown with $S_y$-pollen tube in a pistil which expresses the $S_x$-RNase and incompatible pollination (bottom figure) is shown with $S_x$-pollen tube. The S-RNase degradation model (shown on the left), implicates the involvement of multiple SLF proteins (*i.e.* SLF1, red; SLF2, yellow; SLF3, black, irrespective of whether they are obtained from $S_x$-haploype or $S_y$-haplotype) collaborate to cause the degradation of S-RNases. These models however, do not specify the ECM route to the pollen cytoplasm where the interaction of S-RNase and SLF occurs. For the S-RNase compartmentalization model (shown on the right), it is depicted that, S-RNases uptake by endocystosis and their progressive trafficking by default to larger vacuoles in the more matured regions of the growing pollen tube occurs. For the S-RNase and SLF interaction to occur the S-RNases must exit the endomembrane system; a single SLF ($SLF_x$, red; $SLF_y$, blue) is shown. The degradation of pollen RNA (RNA cross out) in incompatible pollen tubes is depicted by the two models, however this process does not occur in compatible pollen tube (RNA not crossed out). All S-RNase degradation models proposed the general degradation of S-RNases whiles S-RNase compartmentalization models on the contrary proposed the isolation of S-RNases from the cytoplasm (Reprinted from McClure *et al.* (2011) by permission from Oxford University Press).

On the other hand, for an incompatible pollination (Figure 1.8, bottom left), the binding of $S_x$-RNase does not occur by any of the depicted SLF proteins ($SLF1_x$, $SLF2_x$ and $SLF3_x$). This then leaves the $S_x$-RNase free to exercise its cytotoxic activity by degrading pollen RNA resulting in the arrest of the growing pollen tube. In the revised S-RNase compartmentalization model (Figure 1.8, right), the homeostatic effects of the expression of pollen genes in their relation to endomembrane system maintenance and HT-B protein elimination is highlighted. For a compatible pollination (Figure 1.8, top right), the degradation of HT-B protein occurs and it's assumed to be a result of the normal expression of pollen genes. Unlike the S-RNase degradation model, the interaction between SLF and S-RNase in the compartmentalization model does not result in the general S-RNase degradation; however they are rather protected and made mostly inaccessible. The compartmentalization model in general however fails to show the direct SLF/S-RNase interaction and the direct function of SLF-containing ubiquitin ligase complexes. However in this revised model for the sake of convenience, although the biochemical nature of these SLF/S-RNase interactions is not known, the model proposed an alternate complexes for both compatible and incompatible interactions, *i.e.* $SRN_x/SLF_y$ (top right) and $SRN_x/SLF_x$ (Figure 1.8, bottom right) respectively. In the incompatible pollination (Figure 1.8, bottom right), it is presumed that the SLF complex is activated by self S-RNases leading to the targeting and degradation of a pollen protein whose function is to degrade HT-B. The HT-B protein therefore remains available to release S-RNases from their containment thereby allowing them to exercise their cytotoxic action causing the

rejection of pollen tubes. Although the actual target of SLF-complex is not known in this model, the facts remain that during an incompatible pollination the stabilization of HT-B, the degradation of pollen RNA, the loss of integrity by the endomembrane system and the release of S-RNase could be observed. SI in this model therefore is viewed as a self-reinforcing mechanism (McClure *et al.*, 2011).

#### *1.3.2.2 GSI system in Papaveraceae*

The GSI mechanism in Papaveraceae (poppy) is quite different from the GSI system observed in the Solanaceae. In the poppy system, two tightly linked multi-allelic S-genes which code for a stigmatic S-protein and a pollen S-protein are responsible for the SI reaction. Unlike the Solanaceae system where the arrest of incompatible pollen occurs in the style, in poppy the arrest of pollen tubes occurs on the stigmatic surface (Silva and Goring, 2001). Until recently, only the female S-determinant gene had been identified in poppy and this female determinant has been found to induce a $Ca^{2+}$-dependent signalling pathway leading to the ultimate death of incompatible pollen (Takayama and Isogai, 2005). The stigmatic S-determinant was identified to be a small glycoprotein (~15kDa) that is secreted into the extracellular space (Foote *et al.*, 1994).

A highly polymorphic pollen expressed gene named *PrpS* (*Papaver rhoeas* pollen S determinant) which encodes a novel protein of ~20kDa has been identified as the male S-determinant (Wheeler *et al.*, 2009). Following

this discovery, the authors have proposed the renaming of the pistil S-gene from *S* to *PrsS* (*Papaver rhoeas* stigma S determinant). PrpS proteins exhibit similar level of polymorphism compared to PrsS proteins. Also PrpS proteins have two conserved domains with one of the domains overlapping with part of a predicted extracellular domain. The other domain is part of a hydrophobic region around the centre of the protein (Wheeler *et al.*, 2009, 2010).

Prior to the identification of the male determinant of the poppy SI response, the use of *in vitro* germination assays and recombinant forms of S-glycoproteins has led to the understanding of the signal transduction pathway occurring in pollen tubes after the initial SI recognition response (Gaude *et al.*, 2006). The induction of SI in poppy triggers a transient increase in cytosolic $(Ca^{2+})i$. This involves the influx of extracellular $Ca^{2+}$ at pollen tube shank. This results in the rapid (within 1-2 min) inhibition of pollen tube growth. $Ca^{2+}$ might be acting as a secondary messenger which triggers the signalling cascade leading to the irreversible pollen tube rejection. The loss of the apical gradient of $(Ca^{2+})i$, the reorganization of the pollen actin cytoskeleton and the extensive F-actin depolymerisation could be observed following the influx of $Ca^{2+}$. The influx of $Ca^{2+}$ from the shank and not the tip of pollen tubes indicate that, the pollen S receptors could be found in this region and could be involved directly in controlling the influx of $Ca^{2+}$ (Franklin-Tong *et al.*, 2002; Gaude *et al.*, 2006). A proposed model for the mechanism of SI response in poppy is shown in Figure 1.9.

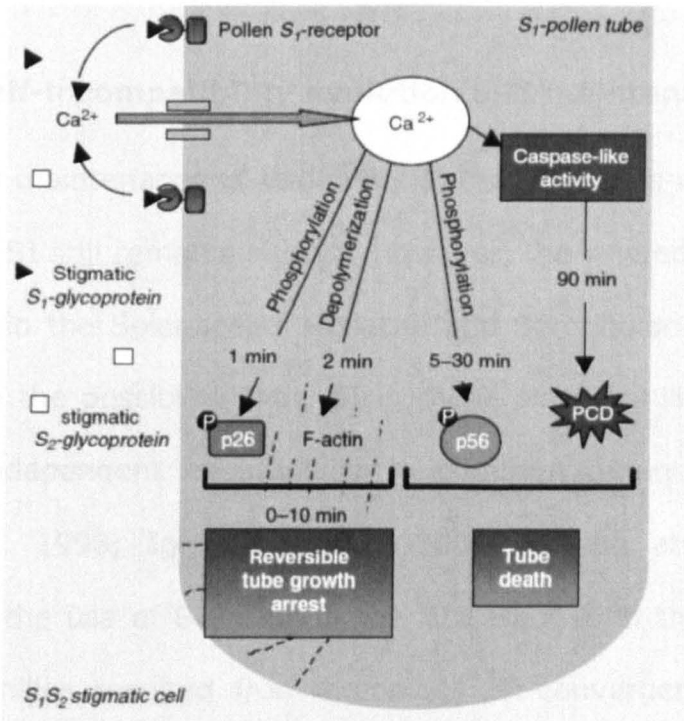



**Figure 1.9 Model for the mechanism of GSI response in *Papaver rhoeas*.** S-glycoproteins secreted by the stigmatic cells interact with pollen S-receptor in S-specific manner. The interaction between the specific stigmatic S-glycoprotein and the pollen S-receptor (here $S_1$) triggers an instant influx of $Ca^{2+}$ and an increase in cytosolic free $Ca^{2+}$ within the self-pollen tubes, thereby triggering an intracellular $Ca^{2+}$-dependent signalling cascade. This signalling cascade involves the phosphorylation of both a soluble inorganic pyrophosphatase protein named p26 and a mitogen-activated protein kinase (MAPK), named p56 protein, and also the depolymerisation of F-actin. 5 min after the induction of SI, p26 MAPK activity increases and peaks at 10 min, remaining higher for at least 30 min. p26 activation is proposed to trigger a programmed cell death (PCD) thereby ensuring an irreversible arrest of the growth of self-pollen tubes (Reprinted from Gaude *et al.* (2006) by permission from CAB International).

## 1.4 The origin of self-incompatibility systems

### 1.4.1 Self-incompatibility evolution and maintenance

Although maintenance of variability at the S-locus is well understood, the origin of SI still remains elusive. However, the relatedness of the pistil S-proteins in the Solanaceae, Rosacae and Scrophulariaceae to S-RNases, points to the possibility that, SI in these plant families might be due to either independent evolution or a common origin (Charlesworth and Awadalla, 1998; Igic and Kohn, 2001). In an attempt to elucidate whether the use of S-RNase in the SI response in the above mentioned plant families resulted from homology or convergence, Igic and Kohn (2001) estimated the evolutionary relationship among T2-type RNases (see legend of Figure 1.10 for details on T2-type RNases) using phylogenetic studies (Figure 1.10) and patterns of intron structure variations. Igic and Kohn (2001) concluded from their findings that S-RNase-based SI system is homologous in the Solanaceae, Rosacae and Scrophulariaceae and in core eudicots and is the ancestral form of SI system existing in the majority of dicots and was present before most of them diversified.

In several angiosperms, the gene family to which the S-locus belongs has been identified by detailed molecular studies (Charlesworth and Awadalla, 1998). As elucidated by these studies, achieving SI involves diverse systems (depending on the plant family) which do not have the same mode of action, since the identified determinant genes of SI have divergent structures. However, a common scheme describes all these

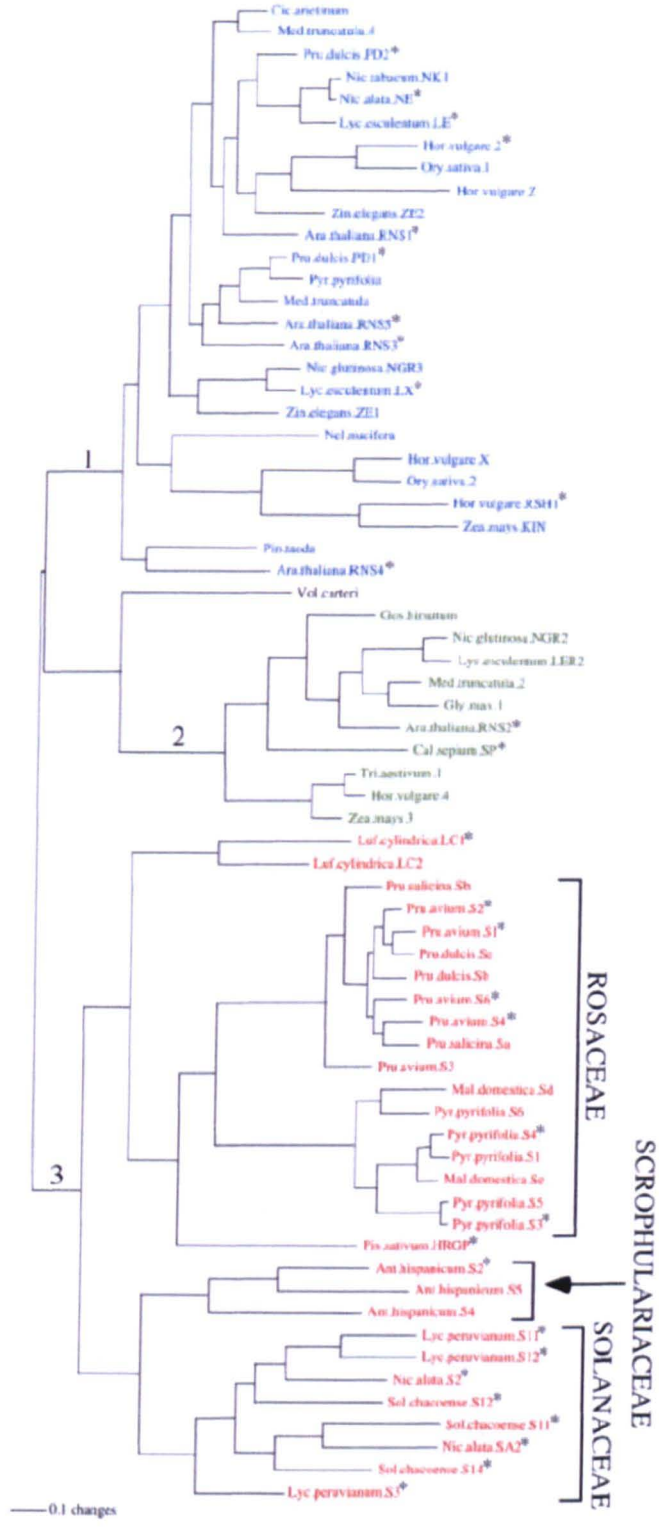



**Figure 1.10 Phylogenetic tree of T2-type RNase in plants.** All the S-RNases could be observed to form a single clade (clade 3) from the non-S-RNases (clades 1 and 2). T2-type RNases in plants have been grouped in three classes; Class I (clade 1) and II (clade 2) comprises non-S-RNases found in higher plants, i.e. they are not located at the S-locus and are not involved in SI but

may be involved in pathogen defence mechanisms or induced in response to phosphate starvation; Class III (clade 3) comprises S-RNases located at the S-locus and are involved in SI reaction. The monophyletic grouping of this Class III T2-type genes from the Solanaceae, Rosacae and Scrophulariaceae indicates homology of the S-RNase-based system in these plant families, i.e. they are present in a common ancestor of these families and passed down to multiple descendant taxa. *Antirrhinum*= Ant, *Arabidopsis*= Ara, *Calystegia*= Cal, *Cicer*= Cic, *Hordeum*= Hor, *Luffa*= Luf, *Lycopersicon*= Lyc, *Malus*= Mal, *Medicago*= Med, *Nelumbo*= Nel, *Nicotiana*= Nic, *Oryza*= Ory, *Pinus*= Pin, *Pisum*= Pis, *Prunus*= Pru, *Pyrus*= Pyr, *Solanum*= Sol, *Triticum*= Tri, *Volvox*= Vol, *Zea*= Zea, *Zinnia*= Zin. The asterisks (*) represents genes which have available information on the structure of their introns (Source: Igic and Kohn, 2001).

divergent SI systems, that of the S-locus comprising at least two tightly linked genes which are arranged in pairs with one functioning as the male determinant and the other the female determinant of SI (Takayama and Isogai, 2005).

The identification of the SI determinants first in the Brassicaceae and later on in the Solanacae, Rosacae and Scrophulariacae revealed that, all these plant families, with the exception of Brassicaceae, share the same identified type of SI determinants. Also, the identified female determinant gene in poppy bears no similarity to the equivalent genes identified in the other plant families, although they all share a unifying scheme of multi-allelic two gene recognition systems. Concluding from these findings, SI might have evolved independently and probably several times in angiosperms (Takayama and Isogai, 2005). Studies involving the analysis of cloned S-allele sequences have lent credence to this hypothesis that,

the self-incompatibility loci have evolved from independent origins several times in angiosperms (Charlesworth and Awadalla, 1998).

The comparison of S-alleles across species or genera reveals that, S-allele lineages have been maintained over tens of millions of years. The high degree of allelic polymorphism and the persistence of the ancient lineages of S-alleles have indicated a strong balancing selection as a result of the transmission advantage accruing to pollen expression of rare allele specificities (Uyenoyama, 2000). The extensive study of S-alleles for the different incompatibility types in plant families often reveals major differences and multiple amino acid differences between S-alleles. Also the identification of the conserved regions of S-alleles has paved the way for the design of primers for PCR-based analysis of S-alleles. The combination of PCR-based techniques coupled with restriction enzyme digests has made it possible to obtain specific bands for the characterization of S-alleles. This gave the impetus for sequencing thereby providing the thrust for studying allelic diversity among natural populations of plants. S-allele sequencing is important in order to move from counting alleles and focus on the quantitative description of the amino acid and base sequence diversity in S-alleles. The polymorphism existing in the S-alleles can then be subjected to molecular evolutionary analysis leading to the testing of recombination and selection events at sites of these loci (Charlesworth and Awadalla, 1998).

### 1.4.2 Diversity of SI systems and polymorphism at the S-locus

The ancestral state or the evolutionary relationships among various SI systems have been very difficult to establish as a result of the diverse state of SI systems and their widespread distribution in angiosperms. The use of molecular phylogenetics in recent years has enabled robust phylogenetic trees to be constructed leading to the identification of the most basal lineages of angiosperms (Allen and Hiscock, 2008).

Differences in the genetic control of SI, the numbers of loci controlling it, and the differences in the site and the timing of the incompatibility response defines the diversity of SI systems in angiosperms. These diverse SI mechanisms in angiosperms are as a result of the selection for more intense and effective mechanism for preventing self-pollen from landing on the stigma to effect self-fertilization thereby promoting outcrossing in bisexual flowers (Hiscock and McInnis, 2003; Allen and Hiscock, 2008).

As a result of diversifying selection at the self-incompatibility loci in angiosperms, there is an extreme level of polymorphism in terms of the numbers of alleles maintained, the divergence of alleles at the molecular level and the time to coalescence of these polymorphisms (Igic *et al.*, 2003). The extreme level of polymorphism in S-proteins indicates an unusual event of evolutionary forces operating at the S-locus (Ioerger *et al.*, 1990). The use of molecular population surveys and interspecific studies in the Solanaceae have revealed the abundance of S-alleles which are functionally distinct in natural populations. More than 50% difference in amino acid residues of S-alleles could be observed in plants of the

same species. A strong selection for diversification and creating polymorphism could explain this phenomenon (Ioerger *et al.*, 1990; Richman *et al.*, 1995, 1996). For instance, an evolutionary study involving S-alleles from the Solanaceae reveals that, S-alleles within a species share a lot of relatedness to S-alleles from different species rather than to those of the same species. This phenomenon termed trans-specific polymorphism might occur as a result of the passing down of S-alleles with modification from a common ancestor to multiple descendant taxa. The authors have concluded that, S-allele polymorphism in the Solanaceae predates the divergence of species (Ioerger *et al.*, 1990; Igic *et al.*, 2003).

## 1.5 Biotechnology and manipulation of self-incompatibility

### 1.5.1 S-RNase regulation in transgenic plants

Conventional breeding strategies have been widely used for the breeding and improvement of crop plants although with some associated limitations. However the birth of the era of biotechnology has opened up many possibilities thereby complementing the limitations of the conventional breeding strategies (Sharma *et al.*, 2005). The advent of recombinant-DNA technology as a result of advances made in the field of molecular biology has ushered in a new age of research into the analysis of the structure and functions of genomes. The technology has made it possible for researchers to exchange genetic materials between divergent species regardless of their phylogenetic standing (Sinsheimer, 1977).

Recombinant-DNA technology coupled with available plant genetic transformation techniques have allowed the transfer of genes from unrelated plants, microbes and animals into crop plants. However like any other new and emerging technology, recombinant-DNA technology is not without challenges or limitations (*e.g.* the availability of efficient transformation methods) (Sharma *et al.*, 2005). The significant advances made in genetic engineering over the past decades and its associated benefits makes this technology a promising tool which can be integrated into current breeding strategies thereby making it (genetic engeenering) one of the widespread available tools to modern plant breeding.

Genetic engineering has been used to successfully manipulate self-incompatibility in several plant species *e.g.*, in *Petunia* and *Nicotiana* in attempts to prove that S-RNases are responsible for S-allele-specific pollen recognition and rejection in the pistil. Lee *et al.* (1994) performed a transgenic loss-of-function experiment, by introducing an antisense S-RNase gene into transgenic SI *Petunia* plants. This caused the transgenic plants to lose the ability to reject pollen borne by the corresponding affected S-allele. The resulting phenotype is a self-compatible transgenic plant. In their experiment, they inhibited the expression of both $S_3$ and $S_2$ proteins in $S_2S_3$ *Petunia inflata* using an antisense $S_3$ gene, which resulted in the failure of the transgenic plants to reject $S_3$ and $S_2$ pollen. The antisense $S_3$ gene was introduced into the *Petunia inflata* plant under the control of the $S_3$ gene promoter in an *Agrobacterium*-mediated transformation.

Also in a gain of function experiment, the introduction of a transgene which encodes $S_3$ protein into $S_1S_2$ genotypes of *Petunia inflata* enables the transgenic plants to reject $S_3$ pollen (Lee *et al.*, 1994). Similar results were obtained in independent transgenic experiments involving *Nicotiana alata* (Murfett *et al.*, 1994). Self-compatibility has also been achieved in an otherwise self-incompatible apple using co-suppression gene silencing approach (Broothaerts *et al.*, 2004). The transformation of a co-suppression construct containing $S_3$-RNase under the influence of CaMV 35S promoter in their experiment resulted in the silencing of endogenous $S_3$-RNase creating a self-compatible apple tree with no adverse effect on the growth of the tree and the quality of fruits. Although engineering self-compatibility through the manipulation of S-RNase is feasible as shown from previous examples this has not yet been performed in potatoes. To the best of my knowledge this thesis is the first report of attempts to perform this in diploid potato species. Engineering self-compatibility will benefit potato breeding by allowing inbred lines to be established thereby enabling genetic analyses to be carried out more easily on the crop.

## 1.6 Aims of the project

This project aims to manipulate the expression of S-RNases in diploid potatoes using the RNAi technique, leading to the development of self-compatible diploid potato lines. This approach to engineering self-compatibility has already been successfully demonstrated in SI *Petunia* (Lee *et al.*, 1994). To date relatively few S-RNase sequences are

available for *Solanum* species compared to other members of the solanaceae. The characterization of S-RNases in *Solanum* species is therefore an initial requirement for the approach. This should generate novel S-allele sequence data in SI diploid potato lines.

The S-RNases present in some diploid potato species provided in collaboration with the James Hutton Institute, JHI (formally the Scottish Crop Research Institute, SCRI) will be characterized. Specifically accessions of *Solanum stenotomum*, *Solanum phureja* and *Solanum okadae* which are important sources of new traits studied by JHI. RT-PCR using degenerate primers will be used to amplify S-RNase sequences from pistil RNA. PCR products will be cloned and sequenced. Selected S-genotypes will be confirmed at the phenotypic level by pollination tests using a diallele cross. To develop familiarity with the techniques S-alleles present in both *Petunia inflata* and *Petunia hybrida cv* Mitchell were initially characterized. This approach was then transferred to the selected diploid potato species.

Following the S-RNase characterization, RNAi constructs will be designed to down-regulate specific *Solanum* S-RNases. As a preliminary study to the transformation work, S-RNase gene expression profiling will be performed to check the developmental stages of expression of the *Solanum* S-RNase genes and also check whether these S-RNases are expressed at a similar level. Wide variation in S-RNase gene expression levels have been reported in the literature.

The long term aim of this thesis is *in planta* transformation of potato with the designed RNAi silencing construct thereby creating SC diploid potatoes leading to the development of the first inbred lines of *S. okadae* (and possibly *S. phureja* and *S. stenotomum*). More general strategies for manipulating SI will be explored if time allows.

# CHAPTER 2: GENERAL MATERIALS AND METHODS

## 2.1 PCR GENOTYPING

### 2.1.1 Genomic DNA extraction

Two leaf discs weighing approximately 10-100 mg of young leaves were harvested into labelled 1.5 ml microcentrifuge tubes. The harvested leaves were then frozen immediately in liquid nitrogen ($N_2$). The tissues were ground to a fine powder with a micro-pestle (pre-chilled in liquid $N_2$) holding the tube in liquid $N_2$. The tubes containing the powder were left in the liquid $N_2$ until the addition of extraction buffer. DNA was harvested from the leaves using the GenElute Plant Genomic DNA Miniprep kit (Sigma-aldrich).

The procedure according to the manufacturer's recommendation with little modifications was as follows: the tubes were taken out of the liquid $N_2$ and 350 µl of Lysis A and 50 µl of Lysis B Solutions were added to it. The mixture was thoroughly mixed by vortexing and inverting (upon the addition of Lysis B solution, a cloudy precipitate will be formed). The mixture was then incubated at 65°C for 10 mins with occasional inversion to dissolve the resultant precipitate.

130 µl of Precipitation Solution was then added to the mixture and mixed completely by inverting the tubes and the samples placed on ice for 5 mins. The tubes were then centrifuged at maximum speed (13,000 rpm) for 5 mins to pellet cellular debris, proteins and polysaccharides. The supernatant was carefully pipetted onto a labelled GenElute filtration

column (blue insert with a 2 ml collection tube) and centrifuged at a maximum speed for 1 min. This was followed by the addition of 700 µl of Binding Solution directly to the flow-through liquid and mixing by inversion.

To prepare the GenElute Miniprep Binding Column (with a red o-ring) to maximize the binding of DNA to the membrane, 500 µl of Column Preparation Solution was added to each miniprep column and centrifuged at maximum speed for 1 min and the flow through liquid discarded.

700 µl of the resulting mixture from the addition of Binding Solution step was then carefully pipetted onto the prepared binding column and centrifuged at maximum speed for 1 min. The flow-through liquid was discarded and the collection tube retained. The column was returned to the collection tube and the remaining lysate pipetted onto the binding column and the centrifugation step above repeated. The flow-through liquid and the collection tube was then discarded and the binding column retained.

The binding column was then placed in a new 2 ml collection tube and 500 µl of Wash Solution added onto it. This was then centrifuged at maximum speed for 1 min and the flow-through liquid discarded but the collection tube retained. Another 500 µl of Wash Solution was added to the column and centrifuged at maximum speed for 3 min to dry the column (ensuring that the binding column does not touch the flow-through liquid).

The binding column was then transferred to a new 2 ml collection tube and 50 µl (instead of 100 µl recommended by the manufacturer) of pre-warmed (65°C) Elution Solution added onto it and centrifuged at maximum speed for 1 min. The elution step was repeated again and the eluate stored at -20°C.

### 2.1.2 Primer design for allele specific PCR

The genotyping primers were designed to show allele-specificity and amplify S-alleles from plants segregating for known S-RNases. The primers were designed to have optimal melting temperature of 65°C. The primers are shown in Table 2.1 below. The V13-F1 & V13-R1 and So-F2 & V26-R5 primer pairs were designed previously (Robbins *et al.*, 2000) whilst the SmR-F2 & SmR-R2 primer pair was designed for this project.

**Table 2.1** Primers for genotyping populations segregating for S-RNases

| Primer Name | Direction | Sequence (5'-3') | Tm °C |
|---|---|---|---|
| **SmR-F2** | Forward | CATGGATTATGTTGTTCAGATG | 59.3 |
| **SmR-R2** | Reverse | CACAGATTAAAAGATAAATCACAAAAG | 59.9 |
| **V13-F1** | Forward | GGACGAAGCTGATTGTAAGGG | 64.6 |
| **V13-R1** | Reverse | CGATTTTCATATATGGC | 52.8 |
| **So-F2** | Forward | GGTATTGCAAAATTAGGG | 54.5 |
| **V26-R5** | Reverse | ATGTTCTCTTCGAGTTCGCG | 64.4 |

The SmR-F2 & SmR-R2 primers amplify petunia plants segregating for the $S_m$-RNase, V13-F1 & V13-R1 primers amplify $S_v$-RNase, whilst So-F2 & V26-R5 primer pair amplifies $S_b$-RNase.

### 2.1.3 PCR/allele-specific PCR for S-genotyping

Populations segregating for known S-RNases were genotyped using allele-specific primers. Genomic PCR reactions (unless otherwise stated) were generally performed in a 25 µl reaction volume comprising of 1X PCR buffer, 3 mM $MgCl_2$, 0.2 mM dNTPs (Bioline, London, UK), 0.4 µM of forward and reverse primers each and 2 U of Taq DNA polymerase (Bioline, London, UK). PCR amplification was performed in a PTC-200 Thermal Cycler (MJ Research, Watertown MA, USA) under the following cycling conditions: an initial 3 min denaturation at 94°C, followed by 35 cycles of 30 sec at 94°C, 30 sec of annealing at a temperature depending on the Tm (melting temperature) of the primer, 1 min at 72°C and a final extension of 7 min at 72°C.

### 2.1.4. Agarose gel electrophoresis

PCR products were generally checked on 0.8-1.5% (w/v) agarose gel stained with 0.15 µg/ml ethidium bromide (EtBr), in 1X TBE buffer (Appendix A1). An aliquot of 15 µl of the amplified product (genomic DNA, plasmid DNA or cDNA) was mixed with 3 µl of 6X loading buffer and loaded into the wells of the agarose gel and run at a constant voltage of 80V for two hours. Depending on the expected amplicon size, 5 µl of either HyperLadder I or Hyperladder II (Bioline, London, UK) was also run alongside the PCR products to serve as a size marker. After the electrophoresis, the gels were visualized under ultra violet light and photographed using a digital imaging system (Syngene, Cambridge, UK).

## 2.2 GENERAL METHODS FOR RT-PCR CLONING

### 2.2.1 RNA extraction

Pistils from the individual plants were harvested in 1.5 ml microcentrifuge tubes. Fifteen to twenty pistils were harvested depending on the size of the tissue. The harvested pistils weighing approximately 10-100 mg were placed in a 1.5 ml labelled microcentrifuge tube and frozen immediately in liquid nitrogen ($N_2$). The tissues were ground to a fine powder with a micro-pestle (pre-chilled in liquid $N_2$) holding the tube in liquid $N_2$. The tubes containing the powder were left in the liquid $N_2$ until the addition of extraction buffer. RNA was harvested from the pistils using the RNeasy Mini kit (QIAGEN, Hilden, Germany).

The procedure according to the manufacturer's recommendation was as follows: the tubes were taken out of the liquid $N_2$ and 450 µl of Buffer RLT (containing Beta mercaptoethanol, ß-ME) was added in a fume cupboard and vortexed vigorously. The lysate was then transferred to a QIAshredder spin column placed in a 2 ml collection tube and centrifuged at maximum speed for 2 minutes (min). The supernatant of the flow-through was carefully transferred into a new microcentrifuge tube without disturbing the cell-debris pellet in the collection tube.

A 500 µl volume of ethanol (100%) was then added to the cleared lysate and mixed immediately by vortexing. The sample, including any precipitate that may have formed, was then transferred to an RNeasy spin column placed in a 2 ml collection tube and centrifuged for 15

seconds (sec) at maximum speed. The flow through was discarded and the collection tube reused in the next step.

The spin column membrane was washed by the addition of 700 µl of Buffer RW1 to the RNeasy spin column and centrifuged at maximum speed (13,000 rpm) for 15 sec. The RNeasy spin column was carefully removed and the flow-through discarded ensuring that the collection tube was completely emptied before being reused in the next step. The spin column membrane was further washed again by the addition of 500 µl of Buffer RPE to the RNeasy spin column and centrifuged at maximum speed for 15 sec. The RNeasy spin column was removed and the flow-through discarded. An additional 500 µl of Buffer RPE was added to the RNeasy spin column and centrifuged at maximum speed for 2 min to dry the spin membrane ensuring that no ethanol is carried over during RNA elution. The RNeasy spin column was placed in a new 2 ml collection tube and centrifuged at maximum speed for 1 min. This step is necessary to prevent any possible carryover of Buffer RPE.

The RNeasy spin column was placed in a new 1.5 ml microcentrifuge collection tube and 30 µl of RNase-free water added directly to the spin column membrane and centrifuged at maximum speed for 1 min to elute the RNA. The RNA quality was checked and quantified (see section 2.2.2) and stored at -80°C.

### 2.2.2 RNA quality control and quantification

After the RNA extraction, 3 µl of the extracted RNA was mixed with 3 µl of 6X loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, 40% sucrose) and electrophoresed on a 1.5% agarose gel in 1X TBE buffer (see section 2.1.4) to check for its quality. RNA quantification and purity check was carried out using the NanoDrop Spectrophotometer (ND-1000 Spectrophotometer, NanoDrop Technologies, Inc., Wilmington, USA).

### 2.2.3 RT-PCR and primers

Reverse Transcription (RT) reactions were carried out using an oligo-dT primer (NotI d(T)18) consisting of a 27bp anchor part with a *Not*I restriction site (included at the 5′ end) and 18 thymidine (T) nucleotides at the 3′ end. A degenerate primer (SolC2-F1.3) was designed and synthesized based on the C2 domain of Solanaceae S-RNases (Ioerger *et al.*, 1991). This was used to amplify partial S-RNase sequences from pistil RNA using the RT-PCR based 3′ RACE (Rapid Amplification of cDNA Ends) technique. The degenerate primer (SolC2-F1.3) was used together with another primer (NotI-anchor) which has the same recognition sequence as the anchor part of the NotI d(T)18 primer used in the first strand cDNA synthesis (see Figure 2.1).

Additional gene specific primers (GSPs) were designed from the partial sequences isolated from the 3′RACE (Rapid Amplification of cDNA Ends) cloning for the identified S-RNases and used for full length cDNA cloning

using the 5'RACE-PCR technique and a commercial kit (Invitrogen, Carlsbad CA, USA). The sequences of the designed primers and the *Not*I primers used for the amplifications are shown in Table 2.2.

**Table 2.2** Sequence of the primers designed and the *Not*I primers used for S-RNase cloning in this study

| Primer Name | Direction | Sequence (5'-3') | Tm °C |
|---|---|---|---|
| **SolC2.F1.3** | Forward | TTTACNRTNCATGGNCTNTGGCC | 64.8 |
| **SdR-GSP1** | Reverse | TGATTTTCTGAACGGTATATTGC | 61.6 |
| **SdR-GSP2** | Reverse | AAAGTATTGATCTTCATTGTAGCTACC | 60.7 |
| **SdR-GSP3** | Reverse | TCCATGCTTATTATACTCATGTCTC | 60.0 |
| **SmR-GSP1** | Reverse | TCTGAGAATAGTCAATAGATCAGTCC | 60.6 |
| **SmR-GSP2** | Reverse | TGTAGACATCTGAACAACATAATCC | 60.7 |
| **SmR-GSP3** | Reverse | TTGCTTGTGGTTAACTCAGG | 60.4 |
| **So2-GSP1** | Reverse | ATTATACCATGATTTCGGAGAGC | 61.9 |
| **So2-GSP2** | Reverse | AGATCGATACTACACGTTCCATG | 61.7 |
| **So2-GSP3** | Reverse | CAGTGATACTCCAGTTGTTTGC | 61.2 |
| **Ss2-GSP1** | Reverse | AGAAGTAATACCATTCTTTCCGAG | 60.4 |
| **Ss2-GSP2** | Reverse | AACACGTTCCATGCTTAATG | 60.1 |
| **Ss2-GSP3** | Reverse | CCAGATGTATTCCAGAGCTTC | 60.3 |
| **Not1-d (T)18** | Forward | AACTGGAAGAATTCGCGGCCGCAGGAA TTTTTTTTTTTTTTTTTT | 81.6 |
| **Not1-anchor primer** | Reverse | AACTGGAAGAATTCGCGG | 62.7 |

Degenerate terms: N=A, C, G or T, R=A or G. The degeneracy fold of the primer (SolC2.F1.3) is 512.

### 2.2.4 Reverse transcription (RT)-PCR reaction: 3'RACE

First strand cDNA synthesis was achieved using QIAGEN Omniscript® Reverse Transcriptase (QIAGEN). The reaction component comprised of 1X RT Buffer, 0.5 mM dNTPs, 1 μM *Not*I d(T)18 primers, 4 U of Omniscript Reverse Transcriptase and approximately 1-2 μg of denatured (at 65°C for 5 min) total RNA template in RNase-free water. The final RT reaction mixture was incubated at 37°C for 1 hr for cDNA synthesis. The RT reaction was set up without RNase inhibitor in some cases (although this was recommended by the manufacturer).

RT-PCR amplification was performed in a 25 μl total reaction mix comprising a 2.5 μl aliquot of RT reaction (cDNA reaction), 20 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTPs (Bioline, London, UK), 0.4 μM of SolC2-F1.3 (forward), 0.2 μM NotI-anchor (reverse) primers and 2 U of Taq DNA polymerase (Bioline, London, UK). Amplification was performed in a PTC-200 Thermal Cycler (MJ Research, Watertown MA, USA) under the following cycling conditions: an initial 3 min denaturation at 94°C, followed by 35 cycles of 30 sec at 94°C, 30 sec annealing at 55°C, 1 min at 72°C and a final extension of 5 min at 72°C. The 3'RACE strategy is depicted in Figure 2.1.

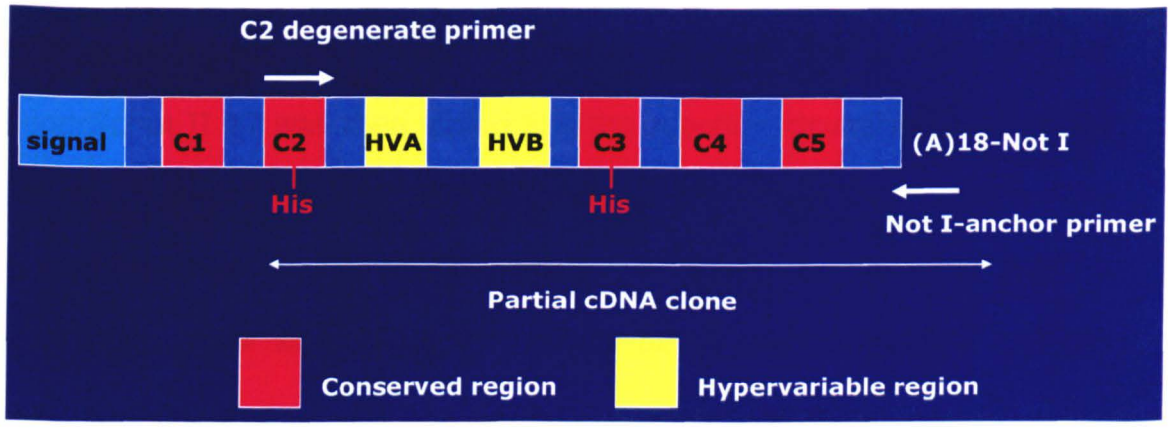



**Figure 2.1 Typical Solanaceae S-RNase structure and 3'RACE strategy.** C1-C5 are the conserved domains of the S-RNase gene and HVA & HVB are the two hypervariable. The two histidine (His) residues located in the C2 and C3 regions are responsible for the ribonuclease activity of the S-RNase gene.

### 2.2.5 Reverse transcription (RT)-PCR reaction: 5'RACE

First strand cDNA synthesis and 5'RACE-PCR was carried out using the 5'RACE System for Rapid Amplification of cDNA Ends, Version 2.0 kit (Invitrogen, Carlsbad CA, USA), according to the manufacturers' recommendations as described below.

The first strand cDNA was synthesized in a reaction comprising 2.5 pmoles of the first gene specific primer (GSP1), 2.5 µg of total RNA, and DEPC-treated water sufficient to make a 15.5 µl final volume. The reaction mixture was then incubated at 70°C for 10 min to denature the RNA. The sample was then chilled on ice for 1 min and the contents collected by brief centrifugation.

After the completion of the first strand cDNA synthesis, the following were added in the order given: a 1X PCR buffer, 2.5 mM $MgCl_2$, 1 mM

dNTP mix and 10 mM DTT. The reaction was mixed gently, collected by brief centrifugation and then incubated at 42°C for 1 min. 1 µl of SuperScript II RT was then added and the reaction incubated at 42 °C for 50 min followed by 70°C for 15 min to terminate the reaction. The reaction mixture was centrifuged briefly for 20 sec and placed at 37°C. 1 µl of RNase mix was added and mixed gently but thoroughly and the reaction incubated at 37°C for 30 min.

The synthesized cDNA was then separated from unincorporated dNTPs, GSP1 and proteins using S.N.A.P Column purification (Invitrogen) by adding 120 µl of binding solution (6M NaI) to the first strand reaction (cDNA). The resulting cDNA/NaI solution was then transferred to a S.N.A.P Column and centrifuged at 13,000 rpm for 20 sec. The cartridge insert was removed and the flow-through transferred to a new microcentrifuge tube (until the recovery of the cDNA was confirmed). The cartridge insert was then placed back into the tube and 0.4 ml of cold (4°C) 1X wash buffer added to the spin cartridge and centrifuged at 13,000 rpm for 20 sec. The flow-through was discarded and the wash step repeated for three additional times. The cartridge was then washed twice with 400 µl of cold (4°C) 70% ethanol and centrifugation at 13,000 rpm for 20 s. After discarding the final 70% ethanol wash from the tube, the spin cartridge was centrifuged at 13,000 rpm for 1 min. The spin cartridge insert was then placed in a new sample recovery tube and 50 µl of sterilized distilled water (preheated to 65°C) was added and centrifuged at 13,000 rpm for 20 sec to elute the cDNA.

After the purification of the cDNA, homopolymer tails were added to the 3'end of the cDNA using terminal deoxynucleotidyl transferase (TdT) to create Abridged Anchor Primer binding sites on the cDNA. The TdT tailing reaction consists of: 1X tailing buffer (10 mM Tris-HCl; pH 8.4, 25 mM KCl, 1.5 mM $MgCl_2$) 0.2 mM dCTP, 10 µl of the purified cDNA sample and 6.5 µl of DEPC-treated water. The reaction mixture was then incubated at 94°C for 3 min and chilled on ice for 1 min. 1 µl of TdT was added with gentle mixing and the reaction mixture incubated at 37°C for 10 min. The TdT was then heat inactivated at 65°C for 10 min. The content of the sample was then collected by brief centrifugation and placed on ice.

PCR amplification of the dC-tailed cDNA was performed in a 25 µl reaction mix comprising of a 2.5 µl aliquot of the dC-tailed cDNA reaction, 20 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.4 µM of Abridged Anchor Primer (AAP) (forward), 0.4 µM of nested GSP2 primer (reverse) and 2 U of Taq DNA polymerase (Bioline, London, UK). Amplification was performed in a PTC-200 Thermal Cycler (MJ Research, Watertown MA, USA) under the following cycling conditions: an initial 3 min denaturation at 94°C, followed by 35 cycles of 30 sec at 94°C, 30 sec annealing at 55°C, 1 min at 72°C and a final extension of 5 min at 72°C.

Following the PCR amplification of the dC-tailed cDNA, a 1 in 100 dilution of the primary PCR product was made and used in a second round of PCR amplification (nested PCR) in order to increase the specificity of the PCR products. Abridged Universal Anchor Primer (AUAP) (forward) and a nested GSP3 primer (reverse) was used to prime the synthesis of the

DNA. The PCR reaction mix and the PCR running conditions are the same as described above except that 0.2 µM of each primer was used in the PCR reaction mix. The PCR products were analysed as described in section 2.7. An overview of the 5'RACE strategy is shown in Figure 2.2.

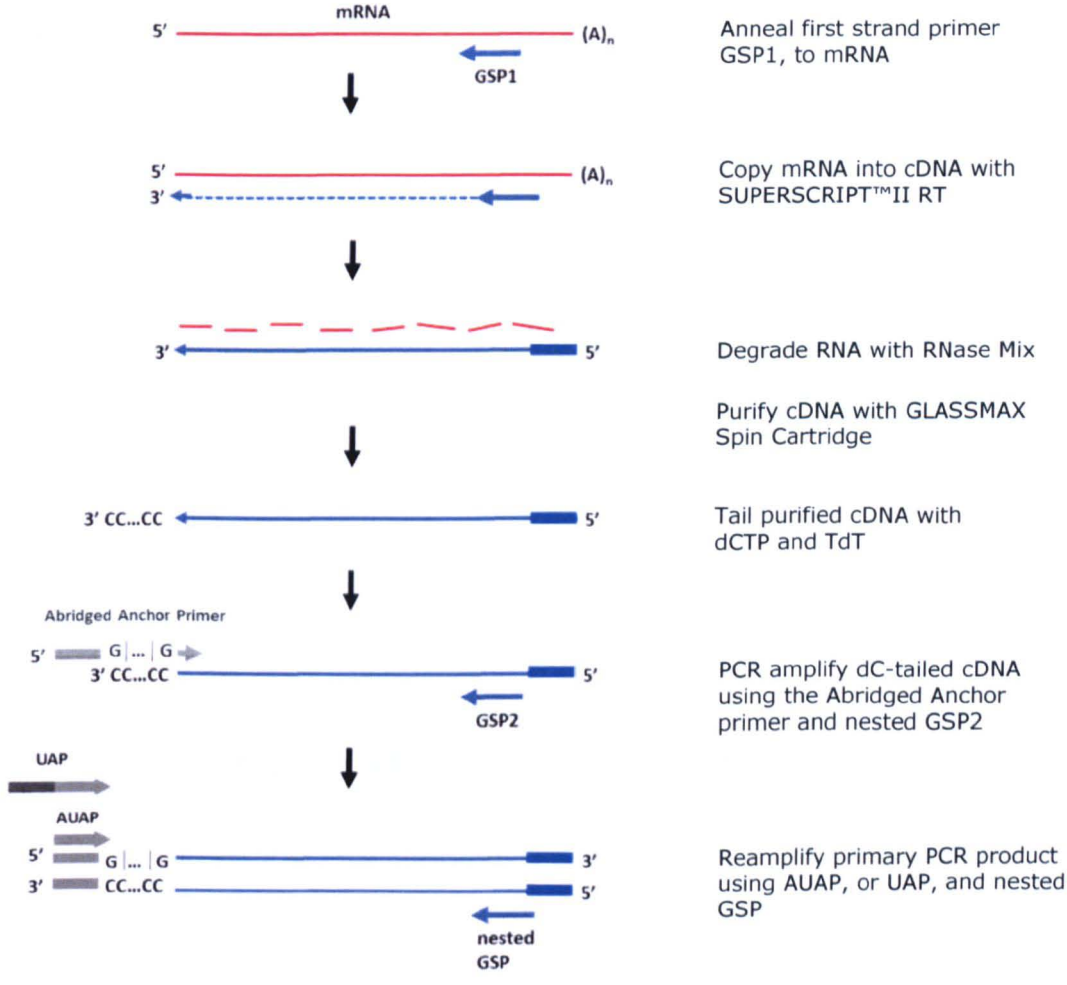



**Figure 2.2 An overview of the 5'RACE strategy.** A gene specific primer (GSP1) and SuperScript II is used to synthesize the first strand cDNA from total RNA. An RNase mix is used to degrade the original RNA template from the RNA:DNA heteroduplex formed. A S.N.A.P Column is used to purify the resulting cDNA from unincorporated dNTPs, GSP1 and proteins. TdT and dCTP is used to add homopolymer tails to the 3'end of the cDNA. This is followed by PCR amplification of the dC-tailed cDNA using AAP and GSP2. The primary PCR product is then re-amplified using AUAP and GSP3 (Redrawn from www.invitrogen.com).

## 2.3 GENERAL METHODS FOR CLONING S-RNase PCR PRODUCTS

RACE-PCR products were cloned using the TA cloning kit as described below (Invitrogen, Carlsbad CA, USA).

### 2.3.1 Ligation into pCR®2.1 vector

The vial containing pCR®2.1 vector was centrifuged to collect all the liquid in the bottom of the vial. The concentration of the fresh PCR (less than 1 day old) needed to ligate with 50 ng (20 fmoles) of pCR®2.1 vector was determined and the ligation reaction set up comprising 1.5 µl of fresh PCR products, 1 µl of 10X ligation buffer, 2 µl of 25 ng pCR®2.1 vector, 2.5 µl of sterile water and 1 µl of 4.0 Weiss units T4 DNA ligase. The ligation mixture was then incubated at 14°C overnight.

### 2.3.2 Transformation into TOP10F' chemically competent cells

The vials containing the ligation reactions was centrifuged briefly and placed on ice. One 50 µl vial of frozen "One Shot" competent cells (TOP10F') for each ligation was thawed on ice. 2 µl of each ligation reaction was pipetted directly into the vial containing TOP10F' competent cells and mixed by stirring gently with the pipette tip. The vials were incubated on ice for 30 min followed by heat shocking for exactly 30 s in the 42°C water bath (ensuring not to mix or shake the vials). The vials were then removed from the water bath and placed on ice.

250 µl of SOC medium (at room temperature) was then added to each tube. The vials were then placed on a shaking incubator and shaken horizontally at 225 rpm for 1 hour at 37°C. Two different volumes of 50 µl and 150 µl of each transformation vial was spread on labelled LB agar plates containing 40 mg/ml of X-Gal, 100 mM of IPTG (Appendix A1) and either 50 µg/ml of kanamycin or 100 µg/ml of ampicillin. The liquid was then allowed to be absorbed by the LB agar plates and then the plates were inverted and placed in a 37°C incubator for at least 18 hours. The plates were then shifted to 4°C for 2-3 hours to allow for full colour development by the X-gal assay.

### 2.3.3 Colony screening by PCR

Generally, 20 white colonies or more were picked with pipette tips onto a new master LB plate containing the appropriate antibiotics and the pipette tips placed into PCR reaction mixtures for colony PCR using universal M13 forward and reverse primers (or other vector specific recommended primers or S-allele specific primers).

### 2.3.4 Plasmid DNA extraction

White colonies with the expected insert size as revealed by the colony PCR were sampled and grown overnight on LB broth containing the appropriate antibiotics. Plasmid DNA was isolated using the QIAprep Spin Miniprep Kit (QIAGEN, Hilden, Germany) according to the manufacturers' recommendations.

### 2.3.5 Glycerol stock preparation of plasmids

Generally, 1.5 ml of cultured cells was added to 500 µl of 60% sterile glycerol stocks in screw capped tubes. The content of the tubes were then mixed by vigorous shaking and then flash frozen in liquid nitrogen and stored at -80 °C. The details of the plasmids were then logged into a glycerol stock record book for future reference.

### 2.3.6 Restriction enzyme digestion

Restriction enzyme digests were usually performed in 15 µl reactions volume containing 1X reaction buffer (supplied) 1X BSA, 400 ng of plasmid DNA and 0.25 units of restriction enzyme (New England Biolabs, NEB). The reactions were routinely incubated for 3hrs-overnight at the optimal temperature for particular restriction enzyme. Heat inactivations of the restriction enzymes were also performed based on the enzymes optimal inactivation temperature. Double digests were performed where necessary following the restriction enzyme provider's recommendations (NEB).

### 2.3.7. Sequencing of plasmid inserts

The sequence of the inserted DNA was determined using forward and reverse M13 universal primers (or other vector-based recommended primers) at either the QIAGEN Genomic Services/Sequencing Services (QIAGEN, Hilden, Germany) or The James Hutton Institute sequencing facility (JHI, Scotland, formally SCRI).

## 2.4. BIOINFORMATICS

### 2.4.1. Sequence data analysis tools

**DNAstar**

DNAstar software (DNAstar Inc. Madison WI, USA) is an all-inclusive tool from Lasergene (http://www.dnastar.com/) used for DNA and protein sequence data analysis. The cloned sequences were edited using this software to remove all vector sequences and also translate the DNA sequences into protein sequences where necessary.

**BLAST search and sequence retrieval**

Blast searches were performed using the National Center for Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov) in order to ascertain whether the sequences are published alleles or novel. The European Bioinformatics Institute (EBI) database (http://srs.ebi.ac.uk) was also used for published sequence retrieval.

**BioEdit**

Sequence alignments were performed using the BioEdit sequence alignment editor (www.mbio.ncsu.edu/BioEdit/bioedit).

**MEGA5**

Phylogenetic analysis was carried out using the Molecular Evolutionary Genetics Analysis (MEGA) software (http://www.megasoftware.net; Tamura *et al.*, 2011).

**Primer3 and Primer3Plus software**

Primers were designed for this project using either the Primer3 primer design tool (http://frodo.wi.mit.edu/) or Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) primer design tool.

# CHAPTER 3: SELF-(IN)COMPATIBILITY STUDIES IN PETUNIA

## 3.1 INTRODUCTION

### 3.1.1 The genus *Petunia*

*Petunia* is a genus in the Solanaceae family. Approximately 30 (sub)-species could be found in this genus. *Petunia* was first described in 1803 by Jussieu, based on materials collected by Commerson in Montevideo, Uruguay (Gerats and Vandenbussche, 2005; Stehmann *et al.,* 2009). It is one of the most important members of the Solanaceae that has ornamental value and is cultivated around the world (Stehmann *et al.,* 2009).

*Petunia* is one of the first bedding plants cultivated and has continued to be one of the preferred genera used for new variety development (Gerats and Vandenbussche, 2005). The most commonly cultivated garden petunia is *Petunia hybrida* and is known to have been derived from a cross between two *Petunia* species (*P. integrifolia* and *P. axillaris*) of South American origin. The temperate and sub-tropical regions of Argentina, Uruguay, Paraguay, Bolivia and Brazil constitute the geographical distribution of *Petunia* with its centre of diversity found in southern Brazil (Stehmann *et al.,* 2009).

### 3.1.2 Self-incompatibility in *Petunia*

The cultivated *Petunia* (*Petunia hybrida*) has contributed immensely towards the unravelling of knowledge in the field of self-incompatibility (Robbins *et al.*, 2000; Sims and Robbins, 2009) over the past decades. *Petunia* is among the earliest plant species used for self-incompatibility research because it thrives very well in glasshouse conditions and shows a high degree of self-incompatibility reaction specificity. Under optimal conditions this can result in a strong pollen tube growth inhibition and total failure of seed set. Also the self-incompatibility reaction is strictly located in the style during pollen tube growth half-way towards the ovary, making the site of the inhibition reaction accessible to experimental approaches. In addition *Petunia* has large homostylous flowers which aid the emasculation of large number of flowers by researchers before flower anthesis. In comparison with other self-incompatible plants, *Petunia* has large styles which give ample tissues for biochemical analysis. *Petunia* is relatively easy to clonally propagate by vegetative cuttings and has simple culture conditions in the nursery. It is also possible to maintain self-incompatible *Petunia* strains by bud pollination (Linskens, 1975) a technique which has been used to generate S-allele homozygous stocks.

*Petunia* exhibits the gametophytic self-incompatibility system like all other solanaceous species investigated to date. The essential features of self-sterility in *Petunia* were first described by Darwin (Darwin, 1891), however the concept that specificity of the pollen phenotype is controlled gametophytically was established later (Harland and Atteck, 1933). Other

very early research works in *Petunia* showed that, the self-incompatibility system was controlled by a single polymorphic locus, the S-locus and the recognition and the rejection of self-pollen is gametophytically determined by the expression of S-alleles in the pollen (Mather, 1943; Linskens, 1975; de Nettancourt, 1977).

### 3.1.3 Self-compatibility in *Petunia*

Although the majority of the wild species found in the Solanaceae family of which *Petunia* is a member, are self-incompatible, some self-compatible species could also be identified. For instance, the genus *Petunia* has 19 known natural taxa, and 4 out of these are self-compatible (Tsukamoto *et al.*, 1998; McCubbin and Kao, 2000). Also breeding has rendered most of the commercial cultivars in the Solanaceae, self-compatible. Thus the SI trait has been selected out during the domestication process of these plant species or at an early stage during the production of inbred lines which are homozygous for desirable traits. One of the self-compatible plants in the *Petunia* genus is the garden *Petunia* (*Petunia hybrida)* (McCubbin and Kao, 2000).

Most cultivars of *P. hybrida* are self-compatible but occasional SI varieties occur (Robbins *et al.*, 2000). One typical example of these self-compatible *Petunia* is *P. hybrida cv* Mitchell (Clark *et al.*, 2009), however the basis of its SC behaviour is currently unknown. Self-compatibility in *P. hybrida* has been studied by Ascher (1984), where the term pseudo-self compatibility (PSC) was coined to describe plants showing both

physiological and genetic breakdown of self-incompatibility following self-pollinations. Pseudo-self compatibility as defined by Ascher (1984) is the ability of plants which are otherwise self-incompatible to set variable levels of seeds following self-pollination or following crosses between individuals bearing the same S-alleles. The partial breakdown of self-incompatibility in plants exhibiting PSC has been distinguished from self-compatibility exhibited by plants which lack an active self-incompatibility. To measure PSC quantitatively, the SI behaviour was expressed as percent (%) PSC by determining the ratio of the number of seeds set in an incompatible cross to that produced by a cross which is fully compatible using the same individuals (Ascher, 1984; Sims and Robbins, 2009).

#### *3.1.3.1 The molecular basis of self-compatibility in Petunia*

In populations of self-incompatible species, the genetic mechanisms governing the self-compatible (fertile) variants found are just as complicated and diverse as their SI counterparts. Most of the factors that modulate a plant becoming self-compatible are caused by mutations. These mutations could be found occurring in the S-alleles themselves or in genes that are closely linked to S-genes and thereby altering their expression. Also the mutations could occur in genes that play key roles in the rejection pathways or genes that are unlinked to the S-locus but play roles in modulating the expression or the turnover of S-allele products. These unlinked genes could also affect how the SI system operates by

altering the pistil environment (Good-Avila *et al.*, 2008). These S-locus unlinked genes that modulate the expression of S-products are termed modifiers (see section 1.3.2.1).

In many plant species belonging to the family Asteraceae, Poaceae, Fabaceae, Rosaceae and Solanaceae, self-compatibility (self-fertility) is known to be associated with specific S-alleles. The conversion of these S-alleles into non-functional alleles could be as a result of mutation at the S-locus or in genes which are tightly linked to the S-locus (Levin, 1996; Good-Avila *et al.*, 2008). The recent understanding of the molecular biology of RNase-based GSI system has enabled researchers to investigate the cause of non-functional S-alleles in the Solanaceae and Rosaceae (Good-Avila *et al.*, 2008). For instance, self-compatibility based on a defective pollen SI component was found in a population of *P. axillaris* from Uruguay. The loss of function was found to be associated with the $S_{17}$ allele. Another population of the same *Petunia* species from Uruguay was found to be self-compatible as a result of the suppression of a specific S-RNase gene by the presence of an S-locus linked gene (Tsukamoto *et al.*, 2003a; Tsukamoto *et al.*, 2003b).

Self-compatibility has also been observed to occur specifically in the Solanaceae as a result of duplication of the S-locus in a phenomenon termed "competitive interaction" (de Nettancourt 1977, 1997; Stone, 2002). Competitive interaction is the breakdown of self-incompatibility in the pollen resulting in the production of self-compatible plants. GSI is well-known to breakdown (at the pollen SI determinant) in tetraploid plants when the S-locus of the diploid parent is heterozygous (Golz *et al.*,

1999; 2001). It was discovered in the mid-20th century that, naturally occurring or induced tetraploidy in several plant families resulted in loss of SI (Stone, 2002). In tetraploids, homozygous and heterozygous pollen production is expected to be at a ratio of 50:50. When tetraploid plants were pollinated with the homozygous pollen grains, they were observed to produce the same effect as normal haploid pollen grains by arresting pollen tube growth in styles. However, pollen grains with two different S-alleles (heteroallelic pollen) were not arrested leading to successful fertilization. Pollinations performed using diploid pollen on diploid styles containing the same two S-alleles, resulted in the production of seeds by only the heterozygous pollen tubes, thus showing that the self-compatibility behaviour observed was not due to tetraploidy in the styles. To confirm that the duplication of the S-locus affects only the pollen component of SI but not the style component, the pollination of tetraploid stigma with haploid pollen grains resulted in no seed set (Stone, 2002). Heterozygous styles of tetraploids will continue to have the ability to reject haploid pollen with matching S-alleles; however diploid heteroallelic pollen will not be rejected when they land on either diploid or tetraploid styles leading to self-compatibility (Sims and Robbins, 2009).

The timing or magnitude of expression of the S-locus genes, their turnover or the normal functioning of the pathways that eventually lead to pollen or pollen tube rejection could be potentially affected by many other genes (Tsukamoto *et al.*, 1999). Other modifiers unlinked to the S-locus are known to interfere with the strength of the SI response leading

to the occurrence of self-compatibility. Due to the complex nature of the SI response, the S-locus unlinked genes which are involved in the self rejection process could undergo mutations and depending on the nature of such mutations could possibly influence self-compatibility either qualitatively or quantitatively even in plant stocks which have their S-locus intact and functional (Tsukamoto *et al.*, 1999; Good-Avila *et al.*, 2008).

## 3.2 MATERIALS AND METHODS

### 3.2.1 Plant Materials

A variety of *Petunia* stocks were used including: *Petunia inflata* (MSc teaching materials supplied by Dr P. Anthony, University of Nottingham), *P. hybrida cv* Mitchell (supplied by Dr R. Koes, Free University of Amsterdam) and populations segregating for previously described functional alleles obtained from Mr U. Devisetty and Dr T. Sonneveld (University of Nottingham). All *Petunia* stocks were grown under controlled glasshouse conditions of 16hr photoperiod and 25°C day/18°C night temperatures in the School of Biosciences, Sutton Bonington Campus, University of Nottingham, UK).

The S-genotype and SI status of the *P. inflata* plants were unknown prior to this study. The N401 plant families were derived from crosses made between S-allele stocks maintained at the University of Nottingham (Dr T. Robbins) and *P. hybrida cv* Mitchell. 'Mitchell' is known to be self-compatible (Clark *et al.*, 2009), and although the basis of its SC

behaviour is unknown the plant is known to be homozygous for an uncharacterised S-allele. A total of 18 plants from the N401 family were used in the study. The pedigree of the N401 family is shown in Figure 3.3 in results section.

### 3.2.2 S-allele phenotypic analysis

#### *3.2.2.1 Pollinations*

Controlled self- and cross-pollinations were carried out using the diallel method after visual inspection of the stigma for the presence of exudates confirming its maturity. Self-pollination was performed using the pollen of an individual flower to pollinate the stigma of the same flower, except in some cases where pollination was performed using pollen from other flowers of the same plant because the stigma had matured before the anthers. For the controlled cross-pollinations, flowers were emasculated two days before anthesis by removing all the stamens to prevent self-pollination whilst leaving the stigma intact. Pollination was carried out by applying pollen to the stigma using sterilized forceps two to three days after emasculation. All pollinations were labelled with a string tag placed around the pedicel.

#### *3.2.2.2 Self-incompatibility phenotype determination*

At maturity (approximately four weeks after pollination) the number of seed set (capsules) produced per plant was scored to determine the

SI/SC phenotype of the plants. Plants were characterised as self-compatible (SC), if all the pollinations resulted in full seed set indicated by the presence of capsules. Alternatively plants were characterized as self-incompatible (SI), if all of the pollinations resulted in no seed set determined by the absence of a capsule. Plants setting intermediate pods or having variable levels of seed set were characterized as pseudo-self-compatible (PSC). To account for errors if a single pollination deviates from the majority of results then the pollinations were repeated to confirm the phenotypes.

## 3.3 RESULTS

### 3.3.1 *Petunia inflata* S-allele characterization

#### *3.3.1.1 Pollination and S-phenotype determination*

A total of 15 *P. inflata* plants growing under controlled glasshouse conditions were initially selfed to determine their SC/SI phenotype. Although the glasshouse was not insect proof, the possibility of insect pollination was reduced because pollinations were carried out during winter (between October and December). From the selfing results, 11 plants were found to be fully SC and were therefore eliminated from the study. A diallel cross design was then used to determine and characterize the possible S-phenotype of the remaining 4 SI plants (assigned plant IDs: P2, P5, P7 and P8). Five pollinations were performed per cross and the seed score of the various crosses recorded (see section 3.2.2.2).

It was observed that, crosses between P2 and P8 resulted in no seed set in either direction (Table 3.1). Similar results were obtained for crosses between P5 and P7. However, crosses of either P2 or P8 with P5 or P7 were found to be successful in either direction resulting in seed set. Based on these results, the four plants were grouped into two possible incompatibility groups (Group I and II) (Table 3.1). This grouping allowed the selection of plants from each group for S-allele cloning and sequencing.

**Table 3.1** Incompatibility groups derived from diallel crosses of *P. inflata*.

| | Pollen Donor | | | |
|---|---|---|---|---|
| | Incompatibility Group I | | Incompatibility Group II | |
| Plant ID | P2 | P8 | P5 | P7 |
| P2 | - | - | + | + |
| P8 | - | - | + | + |
| P5 | + | + | - | - |
| P7 | + | + | - | - |

A compatible cross is indicated by (+), and an incompatible cross is indicated by (-). P2, P5, P7 and P8 represent the individual *P. inflata* plant IDs.

### *3.3.1.2 Cloning and sequencing of P. inflata S-alleles: 3'RACE-PCR*

Following the characterization of the *P. inflata* plants into incompatibility groups through the diallel cross, one plant per group, *i.e.* P8 from Group I and P5 from Group II (Table 3.1) were selected. Pistil RNA was prepared as described in section 2.2.1 and their S-RNase alleles amplified

using the 3′ RACE-PCR technique and cloned using the TA cloning kit (sections 2.2.4; 2.3). To allow the selection of colonies having the expected insert size, the colonies growing following successful transformation were screened using M13 universal forward and reverse primers (Figure 3.1). Results from the screening of the colonies showed that, most of the colonies have the expected insert size and three such colonies were selected and plasmid DNA was extracted from them and sequenced. The sequencing results were edited using DNAstar (DNAstar, Inc. Madison WI, USA) to remove all vector sequences and the resultant putative S-RNase sequences were submitted to NCBI blast searches.

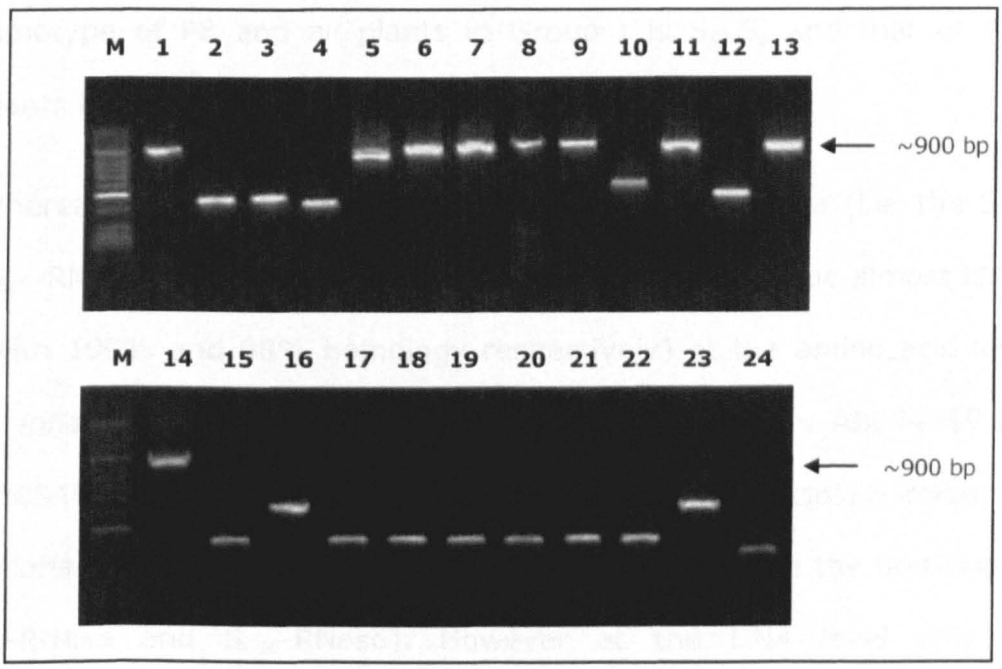



**Figure 3.1 Colony PCR of a typical *P. inflata* pistil S-RNase using M13 universal primers.** Lanes M represent Hyperladder II, lanes 1-23 represent possible transformed colonies, lane 24 represents empty vector. Colonies which are ~900 bp (*e.g.* colonies 1 & 5) seems to have the expected insert size and were therefore selected for sequencing.

Results from the NCBI database searches reveal that, these plants (in the two self-incompatibility groups) harbour three different S-alleles with one of them found to be common to the two incompatibility groups. One of the three cloned putative S-RNases was found to be a previously published S-RNase, *i.e.* $S_{3L}$, therefore reference will be made to this allele henceforth as $S_{3L}$. One of the remaining two was also found from database searches to be almost identical to a published S-allele called $S_{k1}$, but however showed some slight differences at the amino acid level and could be a putative new allele hence will be provisionally called $S_{k1'}$ in this report. The last remaining one as revealed by database searches was found to be novel and hence provisionally called $S_d$. The deduced S-genotype of P8 and all plants in Group I is $S_{k1'}S_d$ and that of P5 and plants in Group II is $S_{3L}S_d$.

Whereas the putative S-alleles cloned from SI *P. inflata* (*i.e.* the $S_{3L}$ and $S_{k1'}$-RNases) were found from the database search to be almost identical (with 100% and 98% homology respectively) at the amino acid level to *P. inflata* S-alleles (with database accession numbers AB094599.1 and AB094600.1 respectively), $S_d$ was found to be most closely related with *Nicotiana alata* S-RNases at the amino acid level (with the first two been $S_7$-RNase and $S_{C10}$-RNase). However at the DNA level only three significant hits could be found from the database search, both the $S_7$-RNase and $S_{C10}$-RNase from *Nicotiana* mentioned above and $S_{p16}$ from *Solanum peruvianum*. The $S_d$-RNase shares 68% homology with the $S_7$-RNase and 70% homology with $S_{C10}$-RNase at the amino acid level and 78% homology with the $S_{p16}$-RNase at the DNA level. Figure 3.2 shows

the partial amino acid sequence alignment of the $S_d$-RNase with the $S_7$-RNase, $S_{C10}$-RNase and $S_{p16}$-RNase. The two hypervariable regions (HVa and HVb) and three (C3, C4 and C5) of the five conserved domains found in all solanaceous S-RNases are highlighted.

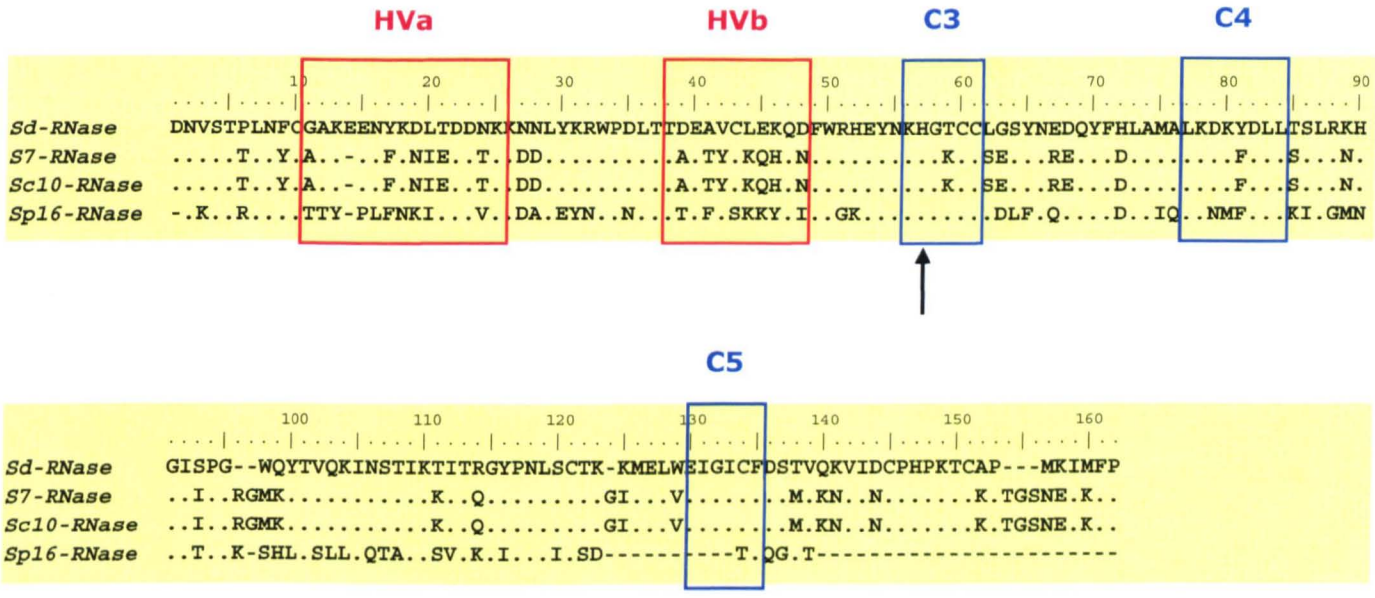



**Figure 3.2 Alignment of the partial deduced amino acid sequence of *P. inflata* $S_d$-RNase with two closely matching S-RNases from *N. alata* ($S_7$-RNase and $S_{C10}$-RNase) and one from *S. peruvianum* ($S_{p16}$-RNase).** The hypervariable regions and conserved regions are boxed. The dots in the alignment indicate amino acid identities among the four sequences with reference to the $S_d$-RNase. One of the conserved histidine residues (C3) involved in the ribonuclease activity of S-RNases is marked with an arrowhead. The database accession numbers for $S_7$-, $S_{C10}$- and $S_{p16}$-RNases are U13255.1, U45959.1 and HM357227.1 respectively. See Appendix A7 for the nucleotide sequence of $S_d$-RNase.

### 3.3.2 S-allele characterization for N401 family

#### *3.3.2.1 Pollination and S-phenotype determination*

*P. hybrida cv* Mitchell is a self-compatible plant (Clark *et al.*, 2009), and although the basis of its SC behaviour is unknown the plant is assumed to be homozygous for an uncharacterised S-allele provisionally called $S_m$. Self-pollinations were performed on all 18 members of the N401 family and their SI/SC phenotype scored in order to investigate and understand the basis of self-compatibility in *Petunia hybrida cv.* Mitchell. The segregation ratio observed was used to indicate whether the $S_m$-allele of *P. hybrida cv.* Mitchell was linked with self-compatibility. Figure 3.3 shows the pedigree of the N401 family used for this study.

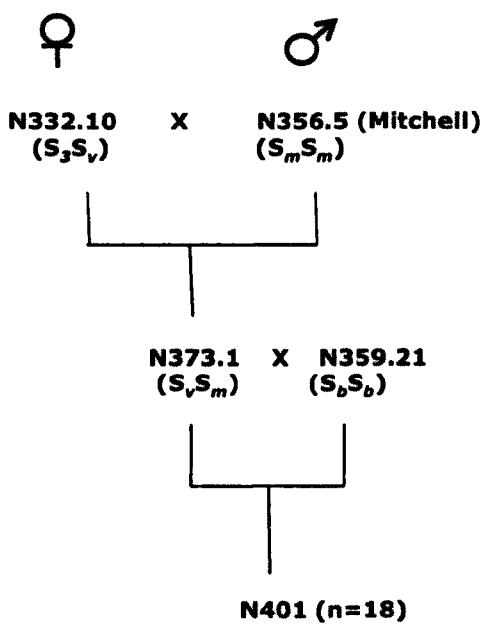



**Figure 3.3 Pedigree of N401 family.** N373.1 is an $F_1$ hybrid between SI line N332.10 and Mitchell stock N356.5. N401 is a cross between SC line N373.1 and SI line N359.21. S-genotypes are indicated in parentheses. The Mitchell line is assumed to be homozygous for an as yet unidentified allele provisionally called $S_m$. The N401 family comprises 18 plants. All crosses are shown using the convention of pistillate parent x staminate parent.

The N401 family was generated by crossing *P. hybrida cv.* Mitchell, N356.5 ($S_mS_m$) and SI plant N332.10 ($S_3S_v$). From the cross the resultant $F_1$ hybrid, N373.1 ($S_vS_m$) was found to be self-compatible. The $F_1$ hybrid was then test crossed to SI line N359.21 ($S_bS_b$) to generate the N401 population. The $S_b$-RNase was characterized and is known to be a functional allele (Wright, 2004; Czyzyk, 2010). A total of 18 plants were raised and tested for their self-compatibility and self-incompatibility phenotype. The results are summarised in Table 3.2.

**Table 3.2** Pollination data for N401 family

| Plant ID | Number of Pollinated Flowers | Number of Seed Set (Capsule) | SC/SI |
|---|---|---|---|
| *N401.1 | 10 | 3 | PSC |
| N401.4 | 5 | 5 | SC |
| N401.6 | 5 | 0 | SI |
| N401.7 | 5 | 0 | SI |
| N401.8 | 5 | 0 | SI |
| N401.9 | 5 | 5 | SC |
| N401.10 | 5 | 0 | SI |
| N401.11 | 5 | 4 | SC |
| N401.12 | 5 | 5 | SC |
| N401.14 | 5 | 5 | SC |
| N401.15 | 5 | 0 | SI |
| N401.16 | 5 | 5 | SC |
| N401.17 | 5 | 5 | SC |
| N401.19 | 5 | 5 | SC |
| N401.20 | 5 | 0 | SI |
| *N401.22 | 5 | 3 | PSC |
| *N401.23 | 5 | 4 | PSC |
| *N401.24 | 10 | 2 | PSC |

Plants exhibiting pseudo-self-compatibility are marked with an asterisk (*). These plants have pod or capsule sizes that are intermediate between SI and fully SC.

It was observed that the majority of the plants are either SC or SI. Moreover it was also observed that, some of the plants have variable levels of seed set, a phenomenon termed pseudo self-compatibility (PSC). From the whole family of 18 individuals, 8 plants (44.4%) were found to be SC, 6 (33.3%) were found to be SI and 4 (22.2%) were found to be exhibiting PSC.

### *3.3.2.2 Cloning and sequencing of Mitchell $S_m$-RNase*

The $S_m$-RNase was an uncharacterised allele prior to this study, although it was observed to have RNase activity (Ajiboye and Sonneveld, unpublished). Therefore an attempt was made to clone and sequence it. The 3'RACE technique was used to obtain the partial sequence of a pistil expressed S-RNase. Following successful transformation, resulting in large numbers of colonies growing on the plates, M13 universal primers were used to screen the colonies. The PCR results showed that, most of the colonies have the expected insert size. Figure 3.4 shows an example of the PCR from the screening of the colonies resulting from the 3'RACE cloning. Three such colonies having the expected insert size were selected and plasmid DNA was extracted from them and sequenced. The sequencing results were edited using DNAstar (DNAstar, Inc.) to remove all vector sequences. All the three resultant putative S-RNase sequences were observed to be identical and were submitted to NCBI blast searches.

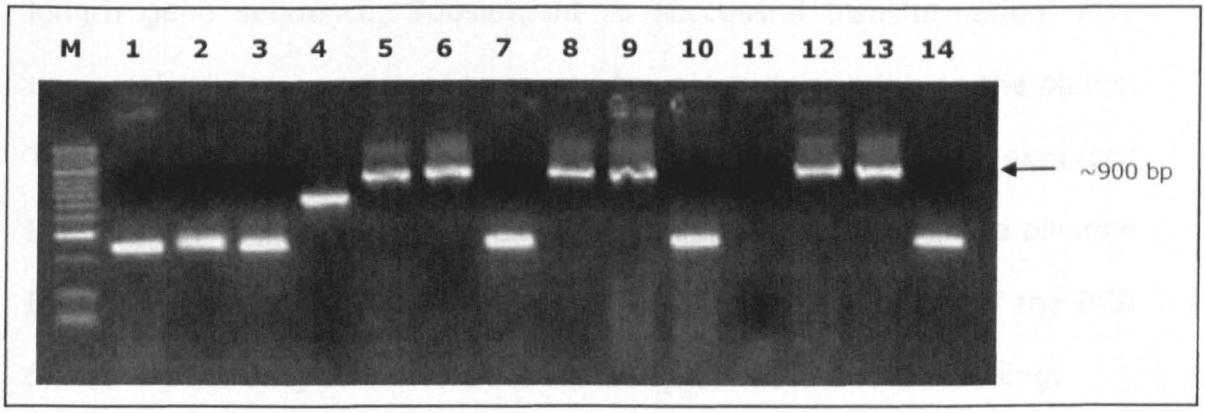



**Figure 3.4 Colony PCR using M13 universal primers for 3'RACE clones of putative $S_m$-RNase.** Lanes M represent Hyperladder II, lanes 1-14 represent possible transformed colonies. Colonies which are ~900 bp (*e.g.* colonies 5 & 6) seems to have the expected insert size and were therefore selected for sequencing.

Database searches with the partial sequence obtained showed that, the closest matches of $S_m$-allele were $S_x$-, $S_o$- and $S_1$-protein from *P. hybrida* and $S_{21}$-protein from *P. integrifolia* with database accession numbers M81686.1, FJ490180.1, U07362.1 and AF301180.1 respectively. However, the partial amino acid sequence alignment of $S_m$ together with the other closely matching S-alleles as revealed by BLAST search showed that $S_m$ is a novel S-allele (data not shown). Figure 3.6 shows that, the HV region is quite unique for this novel allele.

Following the confirmation that $S_m$-RNase is indeed a novel allele, an attempt was made to clone the full length sequence of this gene. Three antisense gene specific primers (SmR-GSP1, SmR-GSP2 and SmR-GSP3) (section 2.2.3; Table 2.2) were designed from the partial sequence obtained through the 3'RACE cloning and used in 5'RACE for the full

length gene sequence. Subsequent to successful transformation, M13 universal primers were used to screen the colonies growing on the plates. The PCR results showed that, most of the colonies have the expected insert size (~600 bp) and three such colonies were selected and plasmid DNA extracted and sequenced. Figure 3.5 shows an example of the PCR from the screening of the colonies resulting from the 5'RACE cloning.

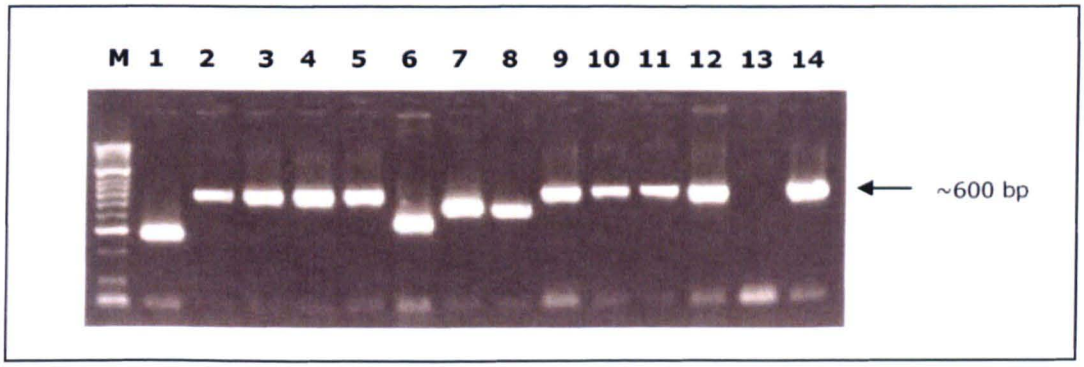



**Figure 3.5 Colony PCR using M13 universal primers for 5'RACE clones of putative $S_m$-RNase.** Lane M represents Hyperladder II, lanes 1-14 represent possible transformed colonies. Colonies which are ~600 bp (*e.g.* colonies 2 & 3) seem to have the expected insert size and were therefore selected for sequencing.

The sequencing results were edited using DNAstar (DNAstar, Inc.) to remove all vector sequences. The 5'RACE cloning was found to be successful following the complete generation of the full length gene sequence by the addition of the C1 and C2 domains to the 5' region of the $S_m$-RNase. The C1 and C2 domains were not part of the initial partial sequence alignment of the $S_m$-RNase. The amino acid sequence alignment of the full length of the $S_m$-RNase with its most closely

matching S-RNases from the database search is shown in Figure 3.6. The two hypervariable regions (HVa and HVb) and the five conserved domains (C1, C2, C3, C4 and C5) found in all solanaceous S-RNases are indicated. From the alignment, it could be observed that, the HV regions (HVa and HVb) of the $S_m$-RNase are quite distinct from $S_x$-, $S_o$- and $S_1$-RNases, which are the most similar alleles from the NCBI database.

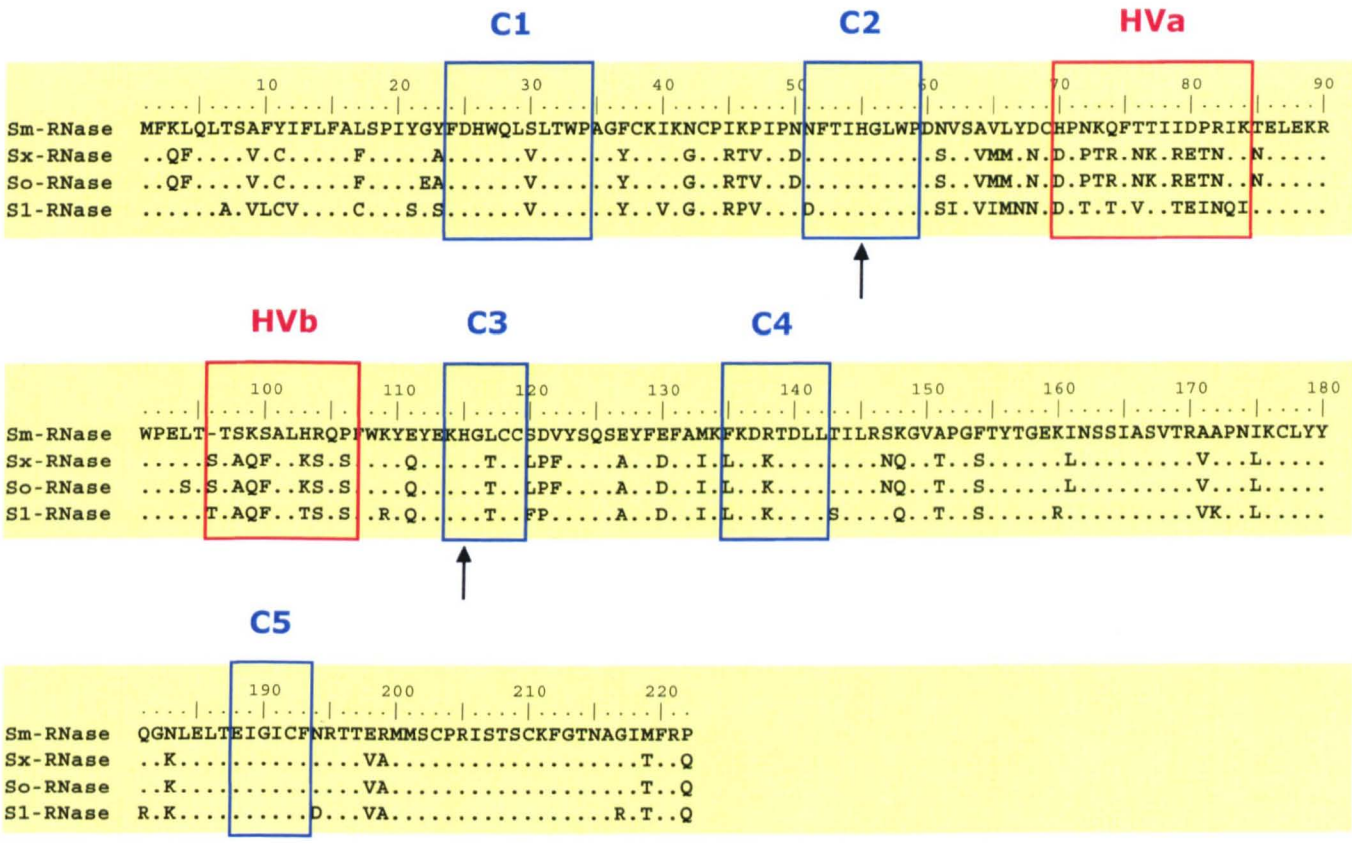



**Figure 3.6 Alignment of the deduced amino acid sequence of $S_m$-RNase with three other most closely matching S-RNases, $S_x$, $S_o$ and $S_1$ from *P. hybrida*.** The hypervariable regions and conserved regions are boxed. The dots in the alignment indicate amino acid identities among the four sequences with reference to the $S_m$-RNase. The two conserved histidine residues known to be involved in the ribonuclease activity of S-RNases are marked with an arrowhead. Database accession numbers for the $S_x$-, $S_o$- and $S_1$-RNases are M81686.1, FJ490180.1 and U07362.1 respectively. See Apendix A7 for the nucleotide sequence of $S_m$-RNase.

### *3.3.2.3 S-genotype determination of N401 using PCR technique*

Genomic DNA was extracted from plants in the N401 family and the segregation of S-alleles in the family was analysed using PCR with allele specific primers (section 2.1; Table 2.1). The family was initially genotyped using $S_v$ and $S_b$ allele specific primers. The $S_v$ and $S_b$ allele specific primers were already available (Robbins *et al.*, 2000; Wright, 2004) because they are previously characterized alleles unlike the $S_m$-allele which was uncharacterised at the beginning of this study, hence the need to design new primers in order to amplify it.

The partial sequence obtained through the sequencing of the $S_m$-allele (see section 2.2) enabled us to design allele specific primers for the $S_m$-allele specific genotyping as well. Results from the PCR showed that, 100% of the plants are positive for the $S_b$-allele, 55.6% are positive for $S_v$-allele and 44.4% are positive for $S_m$-allele (Table 3.3). An example of the PCR profiles for the $S_m$-allele specific primer is shown in Figure 3.7.

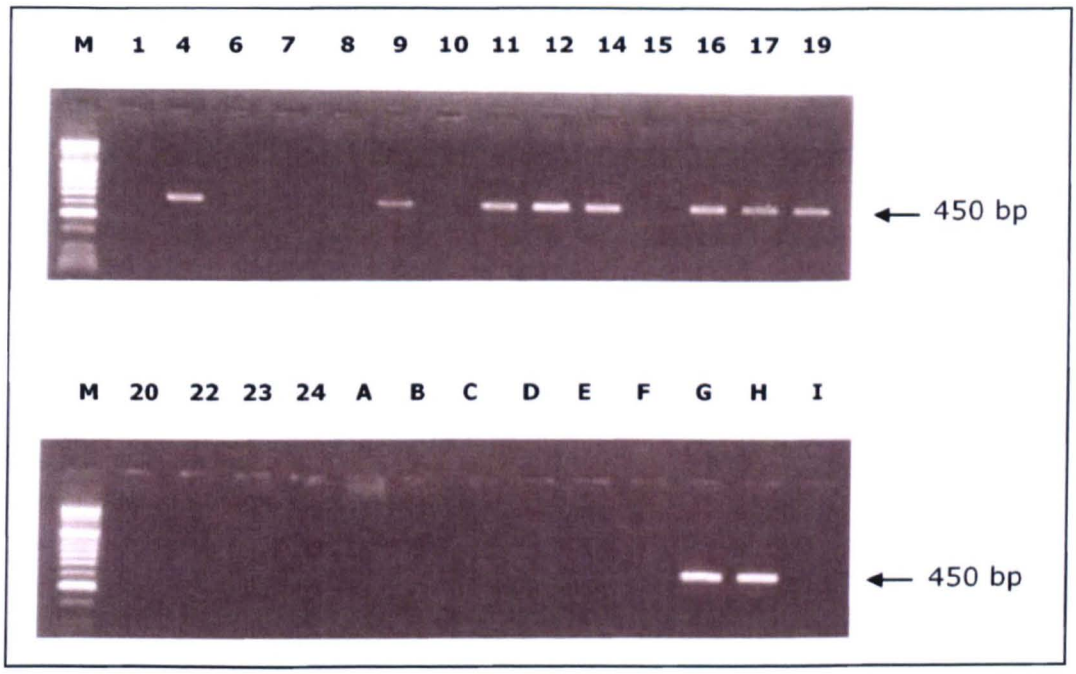



**Figure 3.7 $S_m$-genotyping using $S_m$-allele specific primers.** Lane M represents Hyperladder II, lanes labelled 1-24 represent the 18 plants in the N401 family, lanes A-F represent negative controls to test allele specificity (A and B indicate $S_bS_b$, C and D indicate $S_vS_v$, E and F indicate $S_3S_3$), G and H represent positive controls ($S_mS_m$) and I represents water control.

### *3.3.2.4 Confirmation of the segregation of $S_m$-allele with self-compatibility*

The N401 family pollination results (section 3.3.2.1) were compared with the S-genotyping results (section 3.3.2.3) in order to ascertain the possible linkage between the $S_m$-allele and self-compatibility (Table 3.3). From the pooled results, it was observed that all the plants that were fully self-compatible (SC) were also amplified by the $S_m$-allele specific primers indicating a possible linkage of the $S_m$-allele to self-compatibility. However, those plants showing variable levels of seed set (pseudo self-compatible plants) could not be amplified successfully by the $S_m$-primer.

From the results, the plants that are fully SC have a deduced S-genotype of $S_bS_m$ and those that are either SI or PSC have deduced S-genotype of $S_bS_v$. The segregation of modifier loci which is possibly unlinked to the S-locus could perhaps account for the PSC feature of these $S_bS_v$ plants which otherwise should have been SI.

**Table 3.3** S-genotype and pollination data for N401 family

| Plant ID | Sb | Sm | Sv | S-genotype | SI/SC Phenotype |
|---|---|---|---|---|---|
| *N401.1 | + | - | + | $S_bS_v$ | PSC |
| N401.4 | + | + | - | $S_bS_m$ | SC |
| N401.6 | + | - | + | $S_bS_v$ | SI |
| N401.7 | + | - | + | $S_bS_v$ | SI |
| N401.8 | + | - | + | $S_bS_v$ | SI |
| N401.9 | + | + | - | $S_bS_m$ | SC |
| N401.10 | + | - | + | $S_bS_v$ | SI |
| N401.11 | + | + | - | $S_bS_m$ | SC |
| N401.12 | + | + | - | $S_bS_m$ | SC |
| N401.14 | + | + | - | $S_bS_m$ | SC |
| N401.15 | + | - | + | $S_bS_v$ | SI |
| N401.16 | + | + | - | $S_bS_m$ | SC |
| N401.17 | + | + | - | $S_bS_m$ | SC |
| N401.19 | + | + | - | $S_bS_m$ | SC |
| N401.20 | + | - | + | $S_bS_v$ | SI |
| *N401.22 | + | - | + | $S_bS_v$ | PSC |
| *N401.23 | + | - | + | $S_bS_v$ | PSC |
| *N401.24 | + | - | + | $S_bS_v$ | PSC |

The presence of an unambiguous PCR product is indicated by (+), absence of a product is indicated by (-), plants exhibiting variable levels of seed set are indicated by (*), self-compatibility is indicated by (SC), self-incompatibility is indicated by (SI), (PSC) indicates pseudo self-compatibility as shown in Table 3.2.

## 3.4 DISCUSSION

### 3.4.1 *Petunia inflata* S-allele characterization

#### *3.4.1.1 Pollination and S-phenotype determination for Petunia inflata*

S-allele characterization can be carried out phenotypically by pollination tests and/or genotypically using PCR based techniques. The characterization of S-alleles using pollination tests can be very laborious and time consuming and may not reveal the identity of the S-alleles harboured by these plants (especially in previously un-genotyped plant populations). However, this method can help to identify plants that are cross and self-incompatible in a large plant population. This enables us to put the plant population into incompatibility groups thereby cutting down on the number of plants (with unknown S-alleles) whose identity have to be determined using molecular-based genotyping techniques.

A diallel cross design was used to characterize *P. inflata* plants of unknown S-genotypes growing under controlled glasshouse conditions in this study. Following the results obtained from the various crosses, those plants showing SC or PSC features were selected out and the remaining SI plants were later grouped into two incompatibility groups (Table 3.1). Plants belonging to the same group are self-incompatible and can only set seed with crosses with plants in the other group. This implies that, plants in the same group harbour the same S-alleles hence crosses among them are expected to result in the arrest of pollen tube growth leading to the eventual failure of fertilization occurring. A compatible

reaction will only occur resulting in successful pollination and the ultimate growth of pollen tube when the haplotypes expressed in the pollen and styles do not match.

### *3.4.1.2 Cloning and sequencing of P. inflata S-alleles*

In *Petunia*, cDNA sequences coding for S-proteins were first reported in *P. hybrida* (Clark *et al.*, 1990) and in *P. inflata* (Ai *et al.*, 1990). The isolated cDNA sequences in *Petunia* were found to share similarity with the first S-RNase sequence that was isolated from *Nicotiana* (Anderson *et al.*, 1986). Following the identification of the first S-protein sequence, a large number of S-RNase sequences have been isolated in other species of the Solanaceae. These S-RNases were found to share regular patterns of five interspersed highly conserved amino acid domains (C1-C5) and two highly variable domains (HVa and HVb) (Ioerger *et al.*, 1991).

The grouping of the four *P. inflata* plants into two incompatibility groups (Table 3.1) in this study enabled us to select one plant per group in order to clone and sequence their S-alleles. The sequencing results revealed that these *P. inflata* plants harbour three different S-alleles with one of them (provisionally called $S_d$-RNase) found to be novel. This allele shares homology with two *Nicotiana alata* S-alleles ($S_7$-RNase and $S_{C10}$-RNase) and one tomato ($S_{p16}$ from *Solanum peruvianum*) from the database rather than *Petunia* S-alleles. This observation is typical for Solanaceae where inter-specific similarity among S-RNases is a commonly observed feature (Ioerger *et al.*, 1990). One of the other two identified S-RNases

($S_{3L}$) was found to be already present in the NCBI database. The last identified allele (provisionally called $S_{k1'}$) could be a potential new allele, since it shares only 98% homology with $S_{k1}$ from database searches and not 100% homology unlike $S_{3L}$-RNase. A number of S-RNase sequences have been identified in natural populations of *Petunia*; however, this has not been exhaustively performed at population genetics level (Sims and Robbins, 2009). In *P. inflata*, quite a number of S-RNase sequences have been identified. For instance, in one study, 19 different S-haplotypes were identified in a population comprising 100 individuals (Wang *et al.*, 2001).

### 3.4.2 Characterization of S-alleles in N401 family

In other to ascertain whether the self-compatibility phenotype observed in *Petunia hybrida cv* Mitchell is linked to the S-locus, a series of pollinations were performed on plant populations derived from crosses between Mitchell and lines with a functional SI background. This plant population called the N401 family is made up of plants segregating for known S-alleles which include the $S_v$-, $S_b$- and $S_m$-RNases (see Figure 3.3). This was followed by the cloning of the $S_m$-allele (as discussed in section 3.4.2.2) and the design of allele specific primers for PCR genotyping (discussed below).

### *3.4.2.1 Pollination and S-genotype determination for N401 family*

Results from the pollination data showed that, the majority of the crosses resulted in plants that are either SC or SI. It was also observed that some of the plants set variable levels of seeds (PSC), a feature which could be result of the effect of a modifier. Overall, 44.4% were observed to be SC, 33.3% were SI and 22.2% were exhibiting PSC. Similar results were obtained from a previous study involving segregation of Mitchell $S_m$-allele in a distinct population by a previous student (Ajiboye, 2008). To confirm these observations, S-allele specific PCR genotyping was carried out on these plants. Results from the PCR genotyping using the allele specific primers designed showed that, all plants harbouring the Mitchell $S_m$-allele are indeed self-compatible and are of $S_bS_m$ genotype. It was also observed that, all the plants that were self-incompatible or exhibiting pseudo-self-compatibility were not amplified by the $S_m$-allele specific primer. These plants are of the $S_bS_v$ genotype as deduced from the PCR results. These observations therefore strongly linked the segregation of SC observed in this population to the $S_m$-allele of Mitchell.

### *3.4.2.2 Cloning and sequencing of Mitchell S-alleles*

Prior to this study and the sequencing of the $S_m$-allele, there had not been any mention of the Mitchell S-allele in the literature. However, the $S_m$-allele has been referenced by other authors recently (Puerta *et al.*, 2009a; Puerta *et al.*, 2009b) although the sequence is not yet available in the database. Therefore an attempt was made to clone the gene

sequence of the $S_m$-RNase from *P. hybrida cv.* Mitchell in this study. The partial cDNA cloning was initially achieved successfully and revealed a novel S-RNase sequence. This enabled us to successfully clone and sequence the full gene length of this novel allele. The full length gene sequence of this novel S-allele will be deposited in public databases after publication.

The observation from database searches that, the closest matches of $S_m$-RNase was alleles ($S_o$ and $S_x$) which were previously found to be associated with self-compatibility although they were reported to be functional S-RNases (see section 3.4.2.3) was very surprising. This observation however begs the question, that if they are functional S-RNases (although rendered inactive by unidentified modifiers and/or other compatibility factors) why do they show up together during database searches? And what is so unique about these S-RNases that they cluster together as closest matches in database searches? Although these S-RNases were observed to encode active S-RNases (see section 3.4.2.3), there is however the need to have a second and a critical look at their sequences. Thus a detailed sequence comparison of additional alleles has to be performed between the S-RNases conferring SC phenotype when present and those that confer SI phenotype. This could possibly provide further insight for these S-RNases ($S_m$, $S_o$ and $S_x$ and others) and why they cluster together. This might probably answer some of the questions or uncertainties surrounding their identity as active S-RNases but failing to confer SI phenotype.

In a study involving the $S_o$ and another allele $S_b$ from *P. hybrida*, it was found that, the structure and the expression pattern of both the S-RNase and *SLF* gene are identical for these two S-haplotypes, however, one confers self-compatibility ($S_o$) and the other self-incompatibility ($S_b$). These observations therefore raise similar questions as to why these two S-haplotypes confers different phenotypes (Wright, 2004; Tumusiime, 2006; Czyzyk, 2010).

### *3.4.2.3 The basis of self-compatibility in Mitchell*

Although *P. hybrida* has contributed a lot to the understanding of self-incompatibility, most cultivars are however, self-compatible (Robbins *et al.*, 2000). Several explanations could account for the self-compatibility behaviour observed in Mitchell. The SC phenotype of Mitchell could be a result of a defective stylar protein, loss of pollen-S function, the presence of S-locus linked or unlinked modifier genes. The above causes have been reported by several researchers to be responsible for the loss of SI in different plant families. For instance, a defective stylar component as a result of mutations in the pistil determinant of SI has been reported in *Solanum peruvianum* to be responsible for the loss of SI (Royo *et al.*, 1994). However, a defective stylar protein has been ruled out as a possible cause of the SC of Mitchell in an experiment using the Isoelectric Focusing (IEF) technique and ribonuclease activity staining (Ajiboye and Sonneveld, unpublished). Results from this study showed that the $S_m$-allele has RNase activity. Also the sequencing of the $S_m$-RNase in this

current study showed that it has all the structural features of a functional S-RNase.

This is not the first report of SC in *Petunia* that has been associated with the segregation of a particular S-allele. For instance, in cultivated *P. hybrida*, two S-alleles ($S_o$ and $S_x$) have also been found to co-segregate with functional S-alleles but do not always confer SI phenotype, although they encode active S-RNases. The presence of $S_o$ (Ai *et al.*, 1991; Harbord *et al.*, 2000; Robbins *et al.*, 2000) has been found to be associated with SC whilst $S_x$ (Ai *et al.*, 1991) when present could either render the plants SI or completely SC.

Although defective stylar protein has been ruled out as a possible cause of loss of SI in Mitchell, and although the $S_m$-allele appears to show all the structural features of functional S-RNases and has been shown to have ribonuclease activity, its expression profiling has not been conducted yet and it is worth mentioning that, a reduced expression of the S-RNase gene could also be the cause of SC in Mitchell. The presence of modifiers or compatibility factors rendering the $S_m$-allele of Mitchell non-functional could also account for the self-compatibility of Mitchell. A similar effect of modifiers was found involving the $S_{13}$-RNase of *P. axillaris*. The cause of SC in this case was observed not to be as a result of a deletion of the S-RNase ($S_{13}$) gene or mutations in its promoter but by the presence of a modifier locus causing the suppression of the $S_{13}$-RNases gene expression thereby resulting in SC (Tsukamoto *et al.*, 1999; 2003b). A similar phenomenon could be occurring with the $S_m$-RNase in this case and since full self-compatibility observed in the population

involved in this study was only linked with the $S_m$-allele (and not the other alleles involved $S_b$ and $S_v$), these unidentified modifiers segregating in the population causing the SC could possibly be allele-specific (thus $S_m$-allele linked). The SC nature of Mitchell could also be as a result of a defective pollen SI component. The loss of SI in *P. axillaris* was found to be associated with a defective pollen component (Tsukamoto *et al*, 2003a).

The observation, that some of the plants that were supposed to be self-incompatible with reference to the PCR data (Table 3.3) turned out to be setting variable levels of seed as seen in the pollination data (Table 3.2) was surprising, since all the alleles involved ($S_b$ and $S_v$) are known to be functional alleles. This phenomenon termed pseudo self-compatibility (Ascher, 1984) has been observed in some *Petunia hybrida* varieties (Clark *et al.*, 1990). This feature has also been reported in *Senecio* (Hiscock, 2000). From this current study, all the plants that seem to be exhibiting this characteristic could not be successfully amplified by the $S_m$-allele specific primer, thus their PSC did not arise from Mitchell because these plants don't harbour the $S_m$-allele. However, in a *Petunia* cross involving *P. hybrida* and *P. inflata,* some of the resulting progenies that carry the $S_x$-allele (an allele found to be associated with both SC and SI) were found to be exhibiting PSC (Ai *et al.*, 1990) and possibly indicating a link between the PSC feature observed and the $S_x$-allele. But this is not the case in this current study with the $S_m$-allele since none of the PSC plants carries the $S_m$-allele. Breakdown in SI is not always attributable to the pollen and pistil SI determinants. For instance, in the

study of PSC in *Nemesia*, the pollen or stylar part of SI determinants could not account for part of the system responsible for their SC feature observed (Robacker and Ascher, 1982). PSC in *P. hybrida* was initially thought to be as a result of its hybrid origin. However, experiments involving *P. integrifolia* showed that, PSC in cultivated *Petunia* did not arise as a result of artefact of the inter-specific hybridization that produced *P. hybrida* (Dana and Ascher, 1985). In this current study, the segregation of modifier loci causing weakening of self-incompatibility could possibly be responsible for the pseudo self-compatibility feature observed. The genetic mechanisms governing self-compatibility could be as complicated and diverse as those controlling self-incompatibility and needs the necessary attention for their elucidation. Current advances in the genetics and molecular biology of self-(in)compatibility research will soon pave the way for a better understanding of these vital processes involved in the developmental biology of flowering plants.

# CHAPTER 4: SELF-INCOMPATIBLITY STUDIES IN POTATO

## 4.1 INTRODUCTION

### 4.1.1 Potato crop

Potato, *Solanum tuberosum* L. is the world's fourth most important food crop after wheat, maize and rice, with an annual production of 18 million hectares (Hawkes, 1990; Lang, 2001). It possibly has the largest number of wild and cultivated relatives (approximately 220 wild and cultivated relatives) of any known food or feed crop (Camadro *et al.*, 2004). Potatoes can be seen in both temperate and tropical habitats and at heights up to 4000 meters above sea level (Carputo *et al.*, 2005). The species of potato belong to a very large and diverse genus, *Solanum,* which comprises about 2,300 species (Barroso *et al.*, 1986 as cited by Aversano *et al.*, 2007; Hawkes, 1994). Potato species are largely grouped into the section *Petota Durmont* which comprises of subsections *Estolonifera* and *Potatoe*. Species in the subsection *Estolonifera* are strictly diploids which lacks stolons and tubers. Subsection *Potatoe* species on the other hand have underground stolons and undergoe tuberization under varying environmental conditions. They show different ploidy levels ranging from diploid to hexaploid. However, the majority of them are diploids and the remaining range from triploids, tetraploids, pentaploids and hexaploids. They are grouped into 16 taxonomical series. Seven species of cultivated potatoes as well as their most closely related wild species are known in the series Tuberosa. They occur in polyploid

series with a monoploid number of 12 and range from diploid to pentaploids. However, because of the relative similarity of these potato species, they were not considered as distinct species, but instead they have been bulked into a single species *S. tuberosum* (Hawkes, 1994; Camadro *et al.*, 2004). A combined morphological data (Huaman and Spooner, 2002) and a recent molecular data (Spooner *et al.*, 2007), has however led to the proposal of the reclassification of cultivated potatoes into four species namely *S. tuberosum*, *S. ajanhuiri*, *S. juzepczukii* and *S. curtilobum* ( Spooner *et al.*, 2007).

The potato plant is ideal material for creating and maintaining variation as a result of its reproductive biology. This reproductive behaviour has made it adaptable to a wide range of environmental conditions and end-uses (Bradshaw *et al.*, 2006). The distribution of potato species spans from the southwestern USA via Mexico and Central America through to southern Chile and central Argentina along the Andes mountains towards Brazil, Paraguay and Uruguay in the East. However, the highest number of species per degree latitude could be located from northcentral Peru to central Bolivia and in the central highlands of Mexico (Hijmans and Spooner, 2001; Camadro *et al.*, 2004). The introduction of potatoes into Europe occurred in the 1570s, followed by their subsequent cultivation in many parts of the world (Hawkes and Francisco-Ortega, 1993; Hawkes, 1990; Pandey and Kaushik, 2003). *Solanum tuberosum* L. subsp. andigena Hawkes, the Andean form of tetraploid potatoes, was the first potato to be introduced into Europe. These potatoes were initially adapted to form tubers under 12 hour daylength of the Andes but not

under the long 16-18 hour day length that exists in Europe. However, in the late 18th and 19th centuries, these short day adapted potatoes were selected to suit the 16-18 hour daylength, thereby causing their large scale field cultivation in many parts of Europe (Hawkes, 1994).

The crop has excellent nutritional characteristics and high yield potential, thereby serving as a source of nutrition as a staple crop (Bradshaw *et al.*, 2006). Potato gives important nutrients to the human diet by serving as a non-fattening, nutritious and wholesome food. Also potato tubers contain significant amount of essential amino acids and vitamin C. In addition, they also serve as an important source of other essential vitamins (at least 12) and minerals (Carputo *et al.*, 2005). Sale of fresh potatoes provides an important source of income for many economies. Other uses of the crop include processing into fries and crisps or chips and other processed products. Potatoes can be used for the production of starch and alcohol. The crop is also used for feeding animals (Carputo *et al.*, 2005; Bradshaw *et al.*, 2006). Novel products such as designer starches can also be produced from potatoes (Davies, 1998) and they also hold the potential to be used as bioreactors for the production of biopharmaceuticals (Sonnewald *et al.*, 2003).

### 4.1.2 Potato breeding

The advent of scientific breeding techniques which is based on the sound knowledge of the inheritance of economically important traits has revolutionized plant breeding. This has resulted in rapid progress in the

breeding of many agronomically important crops. However, potato breeding seems to be relatively less successful and more inconsistent in comparison with other crops because of the narrow genetic base of the crop (potato) and its complex inheritance patterns due to the tetrasomic inheritance shown by potato cultivars (Bradshaw and Dale, 2005).

The progress in potato breeding can be judged by how well the available improved cultivars have met the demands and needs of both farmers and consumers (Bradshaw *et al.*, 2006) alike. Hence the efficiency of production and sustainability has been the main emphasis of many potato breeding schemes, with a focus on the development of improved cultivars having multiple resistance and improved quality traits (Hayes and Thill, 2002). For potatoes to meet the needs and demands of the changing world there is the need for a continual improvement of quality traits (Dale and Mackay, 1994). Quality traits are classed into two major categories depending on the market utilization specificities. The first quality category is classed "external quality". External quality traits are vital for fresh consumptions where they have a tendency of influencing the choice of consumers. These external quality traits comprise aspects relating to skin colour, the size and shape of tubers and eye depth. The second category is classed the "internal quality" and includes aspects relating to nutritional properties, culinary values, after cooking properties and processing qualities. Traits such as dry matter content, flavour, sugar and protein content, the quality of starch and the type and amount of glycoalkaloids defines internal quality traits (Carputo *et al.*, 2005).

The common cultivated potato, *Solanum tuberosum*, is a tetraploid (2n=4x=48) showing tetrasomic inheritance patterns. The possession of four sets of chromosomes, and the associated complex inheritance patterns, makes it difficult to carry out genetic studies on the crop at the tetraploid level (Ortiz and Peloquin, 1994). These inherent genetic and biological factors make breeding work on tetraploid potato very burdensome and complicated (Carputo *et al.*, 2005). However, many of the wild and some cultivated species of potato are true diploids (2n=2x=24), and are therefore more amenable for genetic studies and breeding (Tai, 1994; Ortiz and Peloquin, 1994). About 80% of *Solanum* species which are notable for potato breeding are diploids and diploidy is the naturally occurring state for these plants (Carputo and Barone, 2005). These wild diploid *Solanum* species are sources of valuable germplasm with rich resources and traits of agronomic importance that can be introgressed into the cultivated potato to help broaden the genetic base of the cultivated potato (Carputo *et al.*, 2005).

Tuber-bearing wild *Solanum* species provide an excellent and unique source of genetic diversity for breeding purposes due to their evolutionary diversity and also due to the comparative narrow genetic base of *Solanum tuberosum* (Aversano *et al.*, 2007). Also for breeding purposes it is worth noting that, diploid *Solanum* species undergoes normal chromosome pairing and crossing over (Carputo *et al.*, 2005) compared to their tetraploid counterparts. This may help in facilitating key genetic analysis in potato and also help in simplifying the potato breeding process. Breeding at the diploid level, amongst other

advantages, may help in shortening the time period required for the production of a new variety, help in the rapid elimination of deleterious recessive alleles and also aid the efficient introduction of desired agronomically important traits from diploid (2x) species (Ortiz and Peloquin, 1994).

### 4.1.3 Self-incompatibility in potato

Although many of the wild and some cultivated potatoes are true diploids genetic analysis is complicated by the poor fertility and almost universal self-incompatibility (SI) in diploid *Solanum* species. Most of the diploid tuber bearing *Solanum* species rely on outcrossing, due to gametophytic self-incompatibility which is controlled by a single multi-allelic S-locus (Pushkarnath 1942; Pandey 1962; Cipar *et al.*, 1964).

In the SI *Solanum* species, the products of the S-locus, are presumed to be responsible for causing the arrest of pollen tube growth (Hawkes, 1994), thereby preventing self-fertilization. The pistil specificity determinant of SI in *Solanum* is assumed to be controlled by a polymorphic ribonuclease (S-RNase) as found in other members of the Solanaceae (see General Introduction). In *Solanum*, S-RNase genes have been mapped on chromosome 1 of *S. tuberosum* and *S. chacoense* (Gebhardt *et al.*, 1991; Rivard *et al.*, 1996).

The SI existing in diploid tuber-bearing *Solanum* species is an impediment for developing inbred lines (Birhman and Hosaka, 2000). The production of self-compatible (SC) diploid potato lines would therefore

benefit breeding and allow inbred lines to be established. Self-compatible homozygous diploid potato lines, which are vigorous and exhibiting both male and female fertility and are free of deleterious recessive genes, would facilitate many genetic analyses (Jacobsen and Ramanna, 1994) including that of complex traits.

Some self-compatible variants have been reported among self-incompatible diploid potato species, however these cannot be selfed over several generations continuously (Phumichai *et al.*, 2005). The conversion of an SI species into a permanently SC species, could be as a result of mutations of either the female determinant (stylar) or the male determinant (pollen) of the SI response (Qin *et al.*, 2001) or other factors (modifier factors) unlinked to the S-locus (see section 1.3.2.1 on modifiers). These defects on female and male side and the influence of modifier factors responsible for the conversion of SI species to SC species have been reported by various authors in different plant species (*e.g.* see section 3.1.3.1 in Chapter 3).

In a study using a self-compatible variant of diploid potatoes (*S. chacoense*), it was revealed that a dominant gene designated *Sli* (S-locus inhibitor gene) which was unlinked to the S-locus has the ability of altering self-incompatible plants to become self-compatible (Hosaka and Hanneman, 1998a). This gene has been mapped to the most distal end of chromosome 12 and has been proposed to have a possible tight linkage with a recessive lethality gene (Hosaka and Hanneman, 1998a,b). The introductions of *Sli* gene into diploid potatoes can lead to the successful development of $S_2$ progenies (Birhman and Hosaka, 2000). *Sli* behaviour

has been interpreted as a single dominant factor displaying sporophytic inhibition of self-incompatibility (Hosaka and Hanneman, 1998a,b). Although the *Sli* gene is known to be acting in pollen, its precise role leading to SC still remains unclear. The effect of *Sli* inhibiting SI in practise could occur at several stages including the uptake of S-RNases, S-RNase and SLF interactions, or later events leading to the rejection of pollen. Even though the elucidation of the precise role of *Sli* in causing SC is currently an area which is under-studied by researchers trying to understand SI mechanisms, the clarification of the precise role of *Sli* could provide new perspectives on these (McClure *et al.*, 2011).

## 4.2 MATERIALS AND METHODS

### 4.2.1 Plant materials

Accessions of *Solanum okadae*, *Solanum stenotomum* and *Solanum phureja* (Table 4.1) are maintained at the James Hutton Institute, JHI (formally Scottish Crop Research Institute, SCRI) as part of the Commonwealth Potato Collection (CPC) and also duplicate clones were maintained at the University of Nottingham, UoN. Pistils were harvested at JHI and shipped on dry ice and some harvested from those maintained at the UoN. The plants were grown under controlled glasshouse conditions of 16hr photoperiods and 25°C day/18°C night temperatures during winter and natural day lengths during summer with supplementary lighting where necessary.

**Table 4.1** Diploid potato species and clone designations used for the study

| Potato Species | Plant ID |
|---|---|
| *S. okadae* | OKA 7129-1 |
| | OKA 7129-3 |
| | OKA 7129-5 |
| | OKA 7129-7 |
| | OKA 7129-9 |
| *S. stenotomum* | STN 4679 |
| | STN 4679-68 |
| | STN 4679-72 |
| | STN 4711-61 |
| | STN 4741 |
| | STN 4741-119 |
| | STN 4741-135 |
| | STN 4786-80 |
| *S. phureja* | DB 226-70 |
| | DB 337-37 |
| | DB 536-102 |

The OKA accessions will henceforth be referred to as OKA -1, -3, -5, -7 and -9 in this chapter. All others (STN and DB accessions) remain the same.

### 4.2.2 Pollinations

Controlled self- and cross-pollinations were carried out at both JHI and UoN in order to confirm the SI status of the potato stocks and also to determine their possible compatibility relationships. The anthers of potatoes are hollow tubes (anther cones) that open by small apical pores. In order to collect pollen from these tubular anthers, a buzzer was used to vibrate the anther cone and the pollen collected into a microcentrifuge tube. The pollen was then deposited onto the stigma using a paint brush after the stigma has reached maturity. Crosses were scored ~4 weeks after pollination as (self)incompatible if there were no berries (or no seed

set) formed and fully (self)compatible if berries (or seed set) could be seen following pollinations.

### 4.2.3 S-RNase gene cloning and sequencing

Pistil RNA was extracted as described in section 2.2.1. Partial S-RNase sequences were then amplified from the extracted pistil RNAs using the 3'RACE strategy outline in section 2.2.3 and 2.2.4. Full length sequences were then obtained where necessary using the 5'RACE technique as described in section 2.2.5. The resulting RACE PCR products were then cloned (TA cloning) and the confirmed colonies sequenced as outlined in section 2.3 to obtain the putative S-RNase sequences.

### 4.2.4 Sequence alignment, comparison and phylogenetic analysis of S-RNases

The alignment, comparison and phylogenetic analysis of the putative S-RNase sequences were performed using the sequence data analysis tools listed in section 2.4. DNAstar software (DNAstar Inc. Madison WI, USA) was used to deduce the amino acid sequence of the cloned putative S-RNases. The deduced amino acid sequences were then subjected to the NCBI database BLAST (http://blast.ncbi.nlm.nih.gov) searches and the accession numbers of the four best hits noted. The accession numbers were then entered into the EBI database (http://srs.ebi.ac.uk) for the retrieval of the amino acid sequence of these best hits. The accession numbers of the S-RNases are shown in the results section.

The cloned S-RNases were aligned together with one of the selected S-RNase sequences from the database to act as a reference gene. Alignments were performed by ClustalW method (Thompson *et al.*, 1994) using BioEdit (www.mbio.ncsu.edu/BioEdit/bioedit) and also MegAlign™ as implemented in DNAstar using the default settings and edited by hand.

The percentage similarities of the amino acid sequence of the cloned S-RNases was also determined by calculating the sequence distances using the Neighbour-Joining method (Saitou and Nei, 1987) on the basis of the alignment using ClustalW method as implemented in MegAlign™ (DNAstar).

Phylogenetic analysis of the cloned S-RNase sequences together with database (http://srs.ebi.ac.uk) retrieved selected solanaceous S-RNases and S-like RNases were performed using the Neighbour-Joining Tree method as implemented in MEGA5 software (Tamura *et al.*, 2011). The evolutionary distances used to infer the phylogenetic tree was calculated using a Poisson correction method (Zuckerkandl and Pauling, 1965) as implemented in MEGA5. To show how well the topology of the tree was supported, bootstrap analysis (Felsenstein, 1985) was performed using 1000 replicates. Three *Antirrhinum* S-RNase amino acid sequences (with database accession numbers X96464, X96465 and X96466) were also included in the phylogenetic tree as an out-group to root the tree and in other cases $T_2$-RNase of *Aspergillus oryzae* (accession number CAA43400.1) was used to root the phylogenetic tree. Positions which contain gaps and missing data were eliminated from the phylogenetic analysis.

## 4.3 RESULTS

### 4.3.2 Confirming the SI status and the compatibility relationships of potato germplasm

In order to confirm the SI status and the compatibility relationships among some of the potato genotypes used in this study, pollinations were performed following a diallel cross design. Pollinations were performed per cross and the seed/berry score of the various crosses recorded (see section 4.2.2). Results from the crosses on *S. okadae* accessions (Table 4.2) showed that, all self-pollinations resulted in no berry set (self-incompatible). Also it was observed that some of the *S. okadae* accessions may possibly be harbouring the same pair of S-alleles particularly OKA 1 and OKA 3 accessions. Crosses between these two parents consistently did not produce any berries in either direction thereby indicating that both parents possess the same pair of S-alleles. Crosses involving the other parents were observed to set seeds in either direction and could represent a compatible cross or semi-compatible cross. The full details of the crosses are in Appendix A2. The compatibility relationship of other potato plants used in this study (*S. stenotomum* and *S. phureja*) could not be checked using pollination tests due to the limited number of flowers available or lack of flowering time synchronisation.

**Table 4.2** Pollination data for *S. okadae*

| | Pollen Donor | | | | |
|---|---|---|---|---|---|
| **Plant ID** | **OKA 1** | **OKA 3** | **OKA 5** | **OKA 7** | **OKA 9** |
| **OKA 1** | * | * | 106 | 125 | 125 |
| **OKA 3** | * | * | 50 | 71 | 82 |
| **OKA 5** | 19 | 60 | * | 37 | 87 |
| **OKA 7** | 52 | 50 | 106 | * | 90 |
| **OKA 9** | 113 | 115 | 100 | 79 | * |

The figures in the table are the average number of seed set per berry (seeds/berry) for the various crosses. The asterisks (*) represents the failed (incompatible) crosses.

### 4.3.3 RT-PCR (3'RACE) of potato pistil RNA

A degenerate primer designed from the C2 domain based on the alignment of solanaceous S-RNases and used in combination with a NotI-anchor primer (see section 2.2.3) enabled the amplification of pistil S-RNases from *S. stenotomum*, *S. okadae* and *S. phureja*. RT-PCRs gave the expected amplicon size of ~850bp (Figure 4.1) for almost all accessions. The first round RT-PCRs always gave clean and specific PCR products and hence were used directly for cloning without the need for a nested PCR step before cloning.

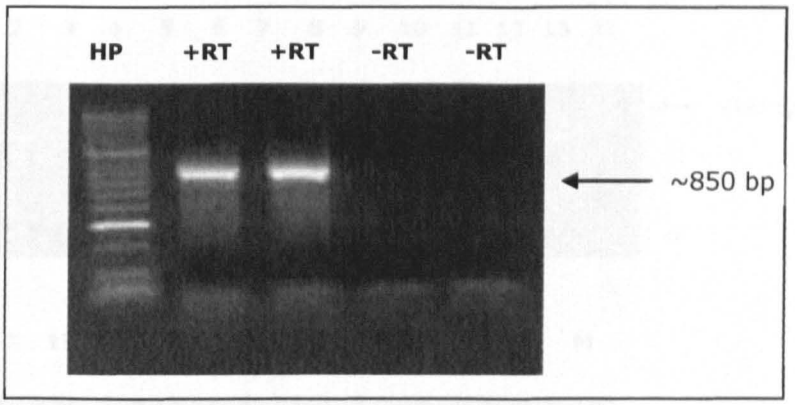



**Figure 4.1 RT-PCR of OKA 9 diploid potato pistils.** Lane HP represents Hyperladder II, +RT represents sample with RT (Reverse Transcriptase), -RT represents sample without RT.

RT-PCR products were cloned into the TA cloning vector (see section 2.3). The amount of the PCR product was estimated from the gel (*e.g.* ~30ng/ul for sample in Figure 4.1) and the optimal amount needed to ligate with the vector determined following the manufacturer's recommendations. In general, transformations were successful with large numbers of colonies growing on the plates. To enable selection for colonies having the right insert size, colony PCR using M13 universal forward and reverse primers was performed. Results from the colony PCR (Figure 4.2) showed that most of the colonies seem to have the expected insert size (~900 bp, corresponding to insert plus part of vector sequence). Three such colonies were then selected and plasmids extracted from them ready for sequencing of their inserts.

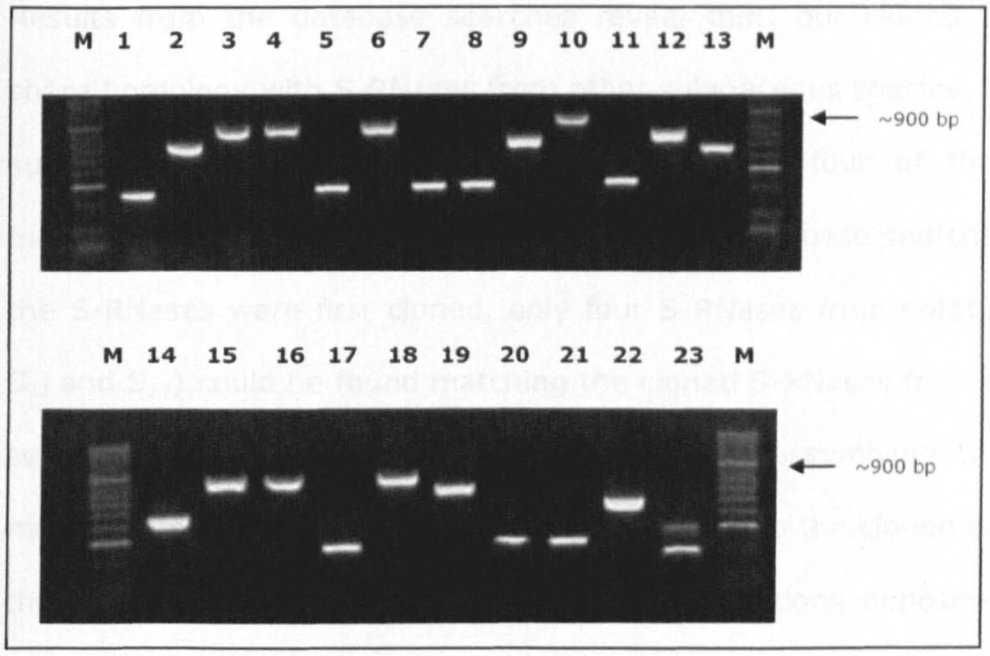



**Figure 4.2 Colony PCR of OKA 9 diploid potato pistil S-RNase using M13 universal primers.** Lanes M represent Hyperladder II, lanes 1-22 represent possible transformed colonies, lane 23 represents empty vector. Colonies which are ~900 bp (*e.g.* colonies 3 & 4) have the expected insert size and were therefore selected for sequencing.

### *4.3.3.1 Sequencing and alignment of the 3'RACE products*

A minimum of three different plasmids revealed through colony PCR to have the expected insert sizes were sequenced. The sequencing results produced a total of seventeen putative S-RNases from the accessions of *S. okadae*, *S. stenotomum* and *S. phureja* following database searches. These S-RNases were provisionally called $S_{o1}$ to $S_{o5}$ for the *S. okadae* S-RNases, $S_{s1}$ to $S_{s10}$ for the *S. stenotomum* S-RNases and $S_{p1}$ & $S_{p2}$ for the *S. phureja* ones as distributed among the accession as shown in Table 4.3. The S-RNase designation is based on the order in which they were identified in each of the species.

Results from the database searches reveal that, our cloned S-RNases share homology with S-RNases from other solanaceous species. Table 4.4 summarises all the 16 S-RNases together with four of the closest matches at the amino acid level from the NCBI database searches. When the S-RNases were first cloned, only four S-RNases from potato ($S_2$, $S_3$, $S_{11}$ and $S_{14}$) could be found matching the cloned S-RNases from database searches. However, during the later stages of assembling this thesis more sequences could be found as best matches to the cloned ones from the NCBI database. Some of these current additions deposited in the database include those matching the cloned S-RNases, $S_{s5}$, $S_{s8}$ and $S_{s9}/S_{p1}$ (see Table 4.4). Database mining for S-RNases published in potato so far has revealed only fourteen S-RNase sequences (Table 4.5).

**Table 4.3** Summary of the deduced S-genotype of *Solanum*

| SPECIES | PLANT ID | CLONED S-ALLELE | | | | DEDUCED S-GENOTYPE |
|---|---|---|---|---|---|---|
| | | Allele | Freq | Allele | Freq | |
| *Solanum okadae* | OKA 1 | $S_{o1}$ | 3 | $S_{o2}$ | 4 | $S_{o1}S_{o2}$ |
| *Solanum okadae* | OKA 3 | $S_{o1}$ | 2 | $S_{o2}$ | 3 | $S_{o1}S_{o2}$ |
| *Solanum okadae* | OKA 5 | $S_{o1}$ | 13 | $S_{o5}$ | 2 | $S_{o1}S_{o5}$ |
| *Solanum okadae* | OKA 7 | $S_{o3}$ | 10 | - | - | $S_{o3}S_{o}$? |
| *Solanum okadae* | OKA 9 | $S_{o1}$ | 7 | $S_{o4}$ | 2 | $S_{o1}S_{o4}$ |
| *Solanum stenotomum* | STN 4679 | $S_{s1}$ | 3 | $S_{s9}$ | 2 | $S_{s1}S_{s9}$ |
| *Solanum stenotomum* | STN 4679-72 | $S_{s5}$ | 4 | - | - | $S_{s5}S_{s}$? |
| *Solanum stenotomum* | STN 4711-61 | $S_{s2}$ | 3 | $S_{s3}$ | 2 | $S_{s2}S_{s3}$ |
| *Solanum stenotomum* | STN 4741 | $S_{s4}$ | 8 | $S_{s10}$ | 1 | $S_{s4}S_{s10}$ |
| *Solanum stenotomum* | STN 4741-135 | $S_{s8}$ | 4 | - | - | $S_{s8}S_{s}$? |
| *Solanum stenotomum* | STN 4786-80 | $S_{s6}$ | 3 | $S_{s7}$ | 4 | $S_{s6}S_{s7}$ |
| *Solanum phureja* | DB 226 | $S_{p1}$ | 4 | - | - | $S_{p1}S_{p}$? |
| *Solanum phureja* | DB 337 | $S_{p2}$ | 4 | - | - | $S_{p2}S_{p}$? |
| *Solanum phureja* | DB 536 | $S_{p2}$ | 4 | - | - | $S_{p2}S_{p}$? |

The deduced S-genotype and the distribution of the sixteen alleles cloned from the 14 accessions of the three *Solanum* species used for this study are shown. Most of the accessions contain the two alleles expected of a heterozygote whilst only one could be identified in others. Freq = cloned frequency of each allele.

**Table 4.4** Results of blast searches (NCBI) with cloned S-RNases for *S. okadae*, *S. stenotomum* and *S. phureja*

| Cloned S-RNase | Database Matches | | | | |
|---|---|---|---|---|---|
| | Matching allele | E-value | (%) Iden | Accession number | Species |
| $S_{o1}$ | $S_{11}$ | 7e-108 | 93 | AAB30528.1 | *S. chacoense** |
| | $Lpf_{SRN-1}$ | 3e-82 | 73 | BAC00940.1 | *S. neorickii#* |
| | $S_{24}$ | 7e-82 | 73 | BAC00932.1 | *S. peruvianum#* |
| | $S_{15}$ | 4e-80 | 72 | BAC00924.1 | *S. peruvianum#* |
| | | | | | |
| $S_{o2}$ | $S_3/S_8$ | 3e-94 | 78 | AAV69976.1 | *N. glauca* |
| | $S_{25}$ | 1e-77 | 69 | BAC00933.1 | *S. peruvianum#* |
| | $S_3$ | 2e-71 | 63 | CAA53666.1 | *S. peruvianum#* |
| | $S_5$ | 1e-70 | 62 | AAB26702.1 | *S. peruvianum#* |
| | | | | | |
| $S_{o3}$ | $S_2$ | 3e-63 | 70 | BAC00939.1 | *S. habrochaites#* |
| | $S_2$ | 5e-63 | 56 | CAA40216.1 | *S. chacoense** |
| | $S_{16}$ | 1e-61 | 58 | AAG40749.1 | *P. integrifolia* |
| | $S_{23}$ | 4e-61 | 74 | ABV46013.1 | *S. chilense#* |
| | | | | | |
| $S_{o4}$ | $S_{3L}$ | 2e-86 | 75 | BAE73274.1 | *P. inflata* |
| | $S_{12}$ | 5e-86 | 75 | AAG40746.1 | *P. integrifolia* |
| | $S_{19}$ | 2e-79 | 92 | ABV46023.1 | *S. chilense#* |
| | $S_{13}$ | 5e-79 | 69 | AAG40747.1 | *P. integrifolia* |
| | | | | | |
| $S_{o5}$ | $S_{15}$ | 5e-86 | 76 | BAC00924.1 | *S. peruvianum#* |
| | $S_1$ | 6e-86 | 76 | BAC00934.1 | *S. chilense#* |
| | $Lpf_{SRN-1}$ | 2e-85 | 76 | BAC00940.1 | *S. neorickii#* |
| | $S_{24}$ | 1e-84 | 75 | BAC00932.1 | *S. peruvianum#* |
| | | | | | |
| $S_{s1}$ | $S_{12}$ | 2e-84 | 76 | BAA04146.1 | *S. peruvianum#* |
| | $S_{11}$ | 4e-84 | 76 | AAA77039.1 | *S. peruvianum#* |
| | $S_6$ | 1e-83 | 75 | ADC52410.1 | *S. habrochaites#* |
| | $S_{22}$ | 2e-69 | 61 | BAC00930.1 | *S. peruvianum#* |
| | | | | | |
| $S_{s2}$ | $S_9$ | 2e-75 | 70 | AAV69974.1 | *N. glauca* |
| | $S_{10}$ | 1e-67 | 62 | AAG40744.1 | *P. integrifolia* |
| | $S_{20}$ | 4e-64 | 59 | AAG40752.1 | *P. integrifolia* |
| | $S_3$ | 4e-63 | 60 | CAA40217.1 | *S. chacoense** |
| | | | | | |
| $S_{s3}$ | $S_{12}$ | 1e-83 | 72 | AAG40746.1 | *P. integrifolia* |
| | $S_{3L}$ | 2e-82 | 73 | BAE73274.1 | *P. inflata* |
| | $S_{k1}$ | 2e-80 | 71 | BAE73275.1 | *P. inflata* |
| | $S_{15}$ | 2e-80 | 71 | AAK15437.1 | *P. axillaris* |
| | | | | | |
| $S_{s4}$ | $S_{p7}$ | 6e-84 | 92 | ADR51134.1 | *S. peruvianum#* |
| | $S_{17}$ | 4e-79 | 93 | ABV46028.1 | *S. chilense#* |
| | $S_{15}$ | 3e-68 | 76 | ABD16163.1 | *L. parishii* |
| | $S_{17}$ | 7e-62 | 57 | AAG40750.1 | *P. integrifolia* |

| | | | | | |
|---|---|---|---|---|---|
| $S_{s5}$ | $S_3$ | 3e-88 | 96 | ADO33170.1 | *S. stenotomum** |
| | $Lpf_{SRN-1}$ | 7e-85 | 75 | BAC00940.1 | *S. neorickii*# |
| | $S_{15}$ | 4e-84 | 73 | BAC00924.1 | *S. peruvianum*# |
| | $S_1$ | 5e-84 | 73 | BAC00934.1 | *S. chilense*# |
| | | | | | |
| $S_{s6}$ | $S_{14}$ | 4e-94 | 79 | AAF36980.1 | *S. chacoense** |
| | $S_6$ | 5e-94 | 79 | CAA81334.1 | *S. peruvianum*# |
| | $S_{15}$ | 3e-75 | 79 | AEQ63300.1 | *S. chilense*# |
| | $S_{16}$ | 8e-75 | 79 | AEQ63301.1 | *S. chilense*# |
| | | | | | |
| $S_{s7}$ | $S_{12}$ | 1e-85 | 77 | BAA04146.1 | *S. peruvianum*# |
| | $S_{11}$ | 4e-85 | 77 | AAA77039.1 | *S. peruvianum*# |
| | $S_6$ | 1e-84 | 76 | ADC52410.1 | *S. habrochaites*# |
| | $S_2$ | 1e-66 | 75 | BAC00935.1 | *S. chilense*# |
| | | | | | |
| $S_{s8}$ | $S_{36}$ | 3e-92 | 98 | ADO33171.1 | *S. phureja** |
| | S1 | 3e-89 | 78 | BAC00934.1 | *S. chilense*# |
| | S15 | 3e-89 | 78 | BAC00924.1 | *S. perivianum*# |
| | $Lpf_{SRN-1}$ | 2e-88 | 78 | BAC00940.1 | *S. neorickii*# |
| | | | | | |
| $S_{s9}$/ $S_{p1}$ | 60_A | 4e-91 | 99 | AEN02425.1 | *S. stenotomum** |
| | 47_D | 1e-90 | 99 | AEN02424.1 | *S. stenotomum** |
| | 60_E | 2e-90 | 99 | AEN02429.1 | *S. stenotomum** |
| | 60_B | 3e-90 | 99 | AEN02426.1 | *S. stenotomum** |
| | | | | | |
| $S_{s10}$ | $S_{12}$ | 4e-91 | 75 | AAG40746.1 | *P. integrifolia* |
| | $S_{3L}$ | 1e-83 | 74 | BAE73274.1 | *P. inflata* |
| | $S_{k1}$ | 2e-83 | 73 | BAE73275.1 | *P. inflata* |
| | $S_{15}$ | 3e-83 | 73 | AAK15437.1 | *P. axillaris* |
| | | | | | |
| $S_{s9}$/ $S_{p1}$ | 60_A | 4e-91 | 99 | AEN02425.1 | *S. stenotomum** |
| | 47_D | 1e-90 | 99 | AEN02424.1 | *S. stenotomum** |
| | 60_E | 2e-90 | 99 | AEN02429.1 | *S. stenotomum** |
| | 60_B | 3e-90 | 99 | AEN02426.1 | *S. stenotomum** |
| | | | | | |
| $S_{p2}$ | $S_{14}$ | 3e-116 | 99 | AAF36980.1 | *S. chacoense** |
| | $S_6$ | 7e-92 | 79 | CAA81334.1 | *S. peruvianum*# |
| | $S_{15}$ | 4e-81 | 84 | AEQ63300.1 | *S. chilense*# |
| | $S_{16}$ | 1e-80 | 85 | AEQ63301.1 | *S. chilense*# |

Database searches with the cloned S-RNases reveal that the alleles are similar to other functional S-RNases from the Solanaceae family. Potato S-RNases are marked with (*) to differentiate them from one of their close relatives, tomato, marked with (#), which were formerly *Lycopersicon* but reclassified recently as *Solanum*. All species with L=*Lycium*, N=*Nicotiana*, P=*Petunia*, S=*Solanum*.

**Table 4.5** S-RNases for potato retrieved from EBI Database

| Potato Plant | S-allele | Accession # | Reference |
|---|---|---|---|
| *S. chacoense* | $S_2$ | X56896.1 | Xu *et al.*, 1990 |
| *S. chacoense* | $S_3$ | X56897.1 | Xu *et al.*, 1990 |
| *S. chacoense* | $S_{11}$ | S69589.1 /L36464.1 | Seba-El-leil *et al.*, 1994 |
| *S. chacoense* | $S_{12}$ | AF176533.1/AF19 1732.1 | Qi *et al.*, 2001 |
| *S. chacoense* | $S_{13}$ | L36667.1 | Despres *et al.*, 1994 |
| *S. chacoense* | $S_{14}$ | AF232304 | O'Brien *et al.*, 2002 |
| *S. chacoense* | $S_{16}$ | DQ007316 | Marcellan *et al.*, 2006 |
| *S. stenotomum* | $S_3$ | HM446648 | Kear 2010, unpublished |
| *S. stenotomum* | 60_A | AEN02425.1 | Kear and Malinski, 2010, unpublished |
| *S. stenotomum* | 47_D | AEN02424.1 | Kear and Malinski, 2010, unpublished |
| *S. stenotomum* | 60_E | AEN02429.1 | Kear and Malinski, 2010, unpublished |
| *S. stenotomum* | 60_B | AEN02426.1 | Kear and Malinski, 2010, unpublished |
| *S. phureja* | $S_{36}$ | HM446649 | Kear 2010, unpublished |
| *S. tuberosum* | $S_2$ | X62727 | Kaufman *et al.*, 1991 |

Database searches for published S-RNases from potato reveal only 14 alleles most of which were from *S. chacoense*. Date of retrieval of the S-RNase sequences from the database: March, 2012.

Some of the plants from which we have isolated S-RNases harbour the expected two S-alleles whilst only one allele could be found in others notably *S. phureja* accessions (Table 4.3). Attempts were made to sequence the second allele of those in which only one could be found which proved unsuccessful for some but not all genotypes studied. For instance a second allele was identified for OKA 9 but a similar number of clones for OKA 7 was not successful. In order to identify the second allele of the accessions with just one allele, a negative screening approach was taken. Allele specific primers were designed and used to screen all

colonies revealed by M13 primers to have inserts with the expected sizes. Those that did not amplify with the allele specific primers were thought to be the putative second allele of the accessions screened and were sent for sequencing. However, after screening large numbers of colonies only a few of these putative second allele colonies could be identified. The previously cloned allele from which the primers were designed was present in almost all the colonies screened (data not shown), suggesting that the allele specific primers did not amplify the target allele in every case.

From the sequencing results, $S_{o1}$-RNase was found to be present in four of the *S. okadae* accessions, *i.e.* OKA 1, OKA 3, OKA 5 and OKA 9. $S_{o2}$-RNase was also found to be present in both OKA 1 and OKA 3. Also, $S_{s9}/S_{p1}$ was found to be present in both *S. stenotomum* and *S. phureja* with accession numbers STN 4679 and DB 226 respectively. This allele was initially cloned from *S. stenotomum* and named $S_{s9}$ and was later cloned from *S. phureja* and named $S_{p1}$ (see Table 4.3). However sequence comparison showed that, these two S-RNases shared exactly the same deduced amino acid sequence (data not shown) and hence was provisionally renamed $S_{s9}/S_{p1}$ to represent both alleles although they could represent two different functional alleles.

The alignment of the deduced partial amino acid sequence of the sixteen novel putative *Solanum* S-RNases is shown in Figure 4.3. Three of the conserved domains (C3-C5) and the two hypervariable domains (HVa and HVb) are highlighted and are part of the primary structural features of solananceous S-RNases as defined by Ioerger *et al.* (1991). One of the

two catalytic histidines (His) known to be involved in the ribonuclease activity of S-RNases is indicated in the alignment of the C3 region. Also six out of the eight conserved cysteine residues found in selected solanaceous S-RNases (Ioerger *et al.*, 1991) are also indicated. These cysteine residues are important for determining the tertiary structure of S-RNases (Ishimizu *et al.*, 1996). The remaining two were not shown because they are located on regions which are not part of this alignment (*i.e.* between C1 and C2). However one of the cysteine residues shown (in this alignment) was not perfectly conserved. Thus all sixteen cloned S-RNases contained all six cysteine residues expected with the exception of $S_{o2}$-*RNase* which lacks one such residue at the 3′ end (*i.e.* the last conserved cysteine located just after the C5 region). This particular cysteine residue however, is also missing in some solanaceous S-RNases published in the database.

The percentage amino acid sequence similarity among the sixteen S-RNases ranged from 32.9 to 94.5% (Table 4.6). This low level of sequence similarity observed here is consistent with the high level of polymorphism known to exist at the S-locus in the Solanaceae (Ioerger *et al.*, 1990). From the table, it could be observed that, amino acid similarity within species could be observed to be as low as 32.9% for the *S. okadae* S-RNases, 33.1% for the *S. stenotomum* S-RNases and 44.2% for the two *S. phureja* S-RNases. Also the isoelectric point (pI) value was calculated for the full and partial deduced amino acid sequences (data not shown) using DNAstar and the values revealed that all the S-RNases are basic proteins and having a pI value range of 8.6-9.6.

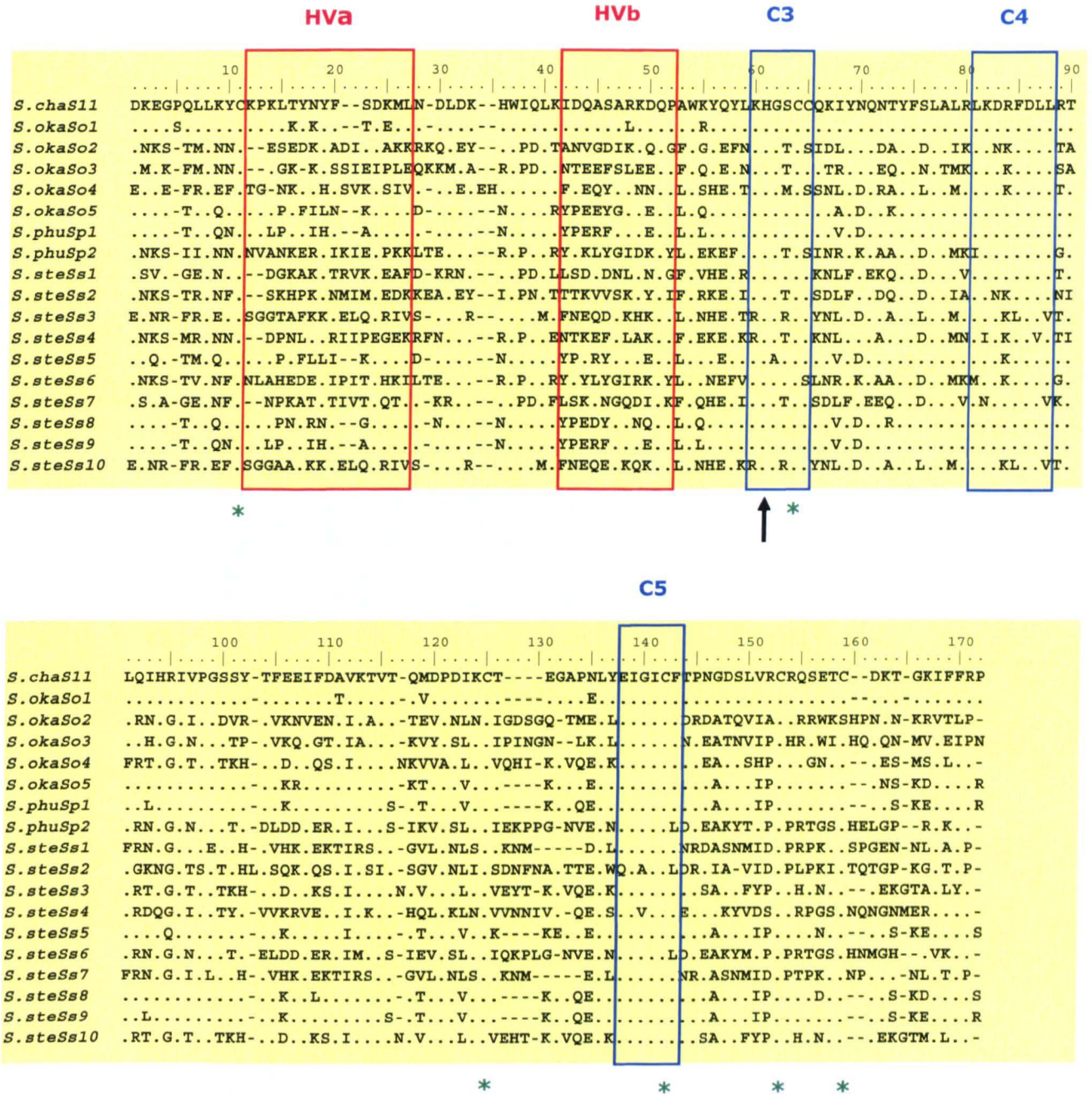



**Figure 4.3 Alignment of the deduced amino acid sequence of the 16 novel S-RNases cloned from *S. okadae, S. phureja* and *S. stenotomum* with one published S-RNase, $S_{11}$ from *Solanum chacoense* (Acc. No: S69589.1)**. Their hypervariable regions (HVa and HVb) and conserved regions (C3-C5) are boxed. The dots in the alignment indicate identities among the 16 sequences with reference to the $S_{11}$-RNase. Gaps in the alignment are indicated by (-). One of the conserved histidine residue involved in the ribonuclease activity of S-RNases is marked with an arrowhead. Six out of the eight conserved cysteine residues are marked with asterisks (*) under the alignment. See Appendix A7 for the nucleotide sequence of the cloned S-RNases.

**Table 4.6** Percentage amino acid similarity of the sixteen cloned S-RNases

| S-allele | $S_{o1}$ | $S_{o2}$ | $S_{o3}$ | $S_{o4}$ | $S_{o5}$ | $S_{s1}$ | $S_{s2}$ | $S_{s3}$ | $S_{s4}$ | $S_{s5}$ | $S_{s6}$ | $S_{s7}$ | $S_{s8}$ | $S_{s9}/S_{p1}$ | $S_{s10}$ | $S_{p2}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $S_{o1}$ | - | 32.9 | 42.2 | 52.3 | 73.9 | 43.7 | 33.5 | 50.6 | 41.9 | 72.0 | 38.3 | 41.1 | 75.8 | 76.4 | 51.3 | 39.6 |
| $S_{o2}$ | | - | 47.5 | 38.0 | 34.8 | 44.6 | 48.4 | 34.4 | 45.3 | 37.4 | 44.1 | 42.0 | 35.5 | 35.5 | 34.4 | 45.3 |
| $S_{o3}$ | | | - | 42.3 | 44.8 | 44.2 | 37.7 | 39.2 | 43.8 | 44.8 | 42.1 | 43.6 | 44.2 | 44.2 | 39.9 | 44.7 |
| $S_{o4}$ | | | | - | 53.5 | 40.3 | 38.0 | 70.8 | 40.5 | 54.8 | 44.6 | 38.3 | 56.8 | 56.1 | 72.7 | 44.6 |
| $S_{o5}$ | | | | | - | 39.7 | 33.5 | 50.6 | 39.4 | 84.7 | 43.5 | 40.4 | 86.0 | 86.6 | 50.0 | 44.2 |
| $S_{s1}$ | | | | | | - | 42.9 | 35.9 | 39.5 | 39.1 | 37.2 | 74.5 | 39.7 | 38.4 | 36.5 | 37.2 |
| $S_{s2}$ | | | | | | | - | 33.1 | 43.5 | 34.8 | 36.2 | 45.5 | 33.5 | 33.5 | 35.0 | 38.8 |
| $S_{s3}$ | | | | | | | | - | 40.6 | 53.8 | 38.4 | 34.0 | 53.2 | 52.6 | 94.5 | 40.9 |
| $S_{s4}$ | | | | | | | | | - | 40.6 | 43.8 | 40.1 | 41.9 | 41.3 | 42.5 | 48.8 |
| $S_{s5}$ | | | | | | | | | | - | 43.5 | 37.7 | 82.8 | 84.1 | 53.2 | 42.9 |
| $S_{s6}$ | | | | | | | | | | | - | 37.8 | 41.6 | 42.2 | 40.3 | 80.4 |
| $S_{s7}$ | | | | | | | | | | | | - | 39.7 | 38.4 | 36.5 | 39.1 |
| $S_{s8}$ | | | | | | | | | | | | | - | 87.9 | 53.8 | 42.9 |
| $S_{s9}/S_{p1}$ | | | | | | | | | | | | | | - | 53.8 | 44.2 |
| $S_{s10}$ | | | | | | | | | | | | | | | - | 42.1 |
| $S_{p2}$ | | | | | | | | | | | | | | | | - |

Species of origin for S-alleles are abbreviated as follows: So = *S. okadae*, Ss = *S. stenotomum*, Sp = *S. phureja*

### 4.3.4 5'RACE of selected S-RNases

Following the confirmation that the cloned sequences are indeed genuine S-RNases, an attempt was made to clone the full length sequence of selected alleles, *i.e.* $S_{o2}$-RNase from *S. okadae* and $S_{s2}$-RNase from *S. stenotomum*. Three antisense gene specific primers (GSPs) (Table 2.2) were designed from the partial sequence obtained through the 3'RACE cloning and used in 5'RACE (section 2.2.5) for the full length gene sequencing. The RT-PCR products gave the expected amplicon (Figure 4.4) and were cloned using the TA cloning vector.

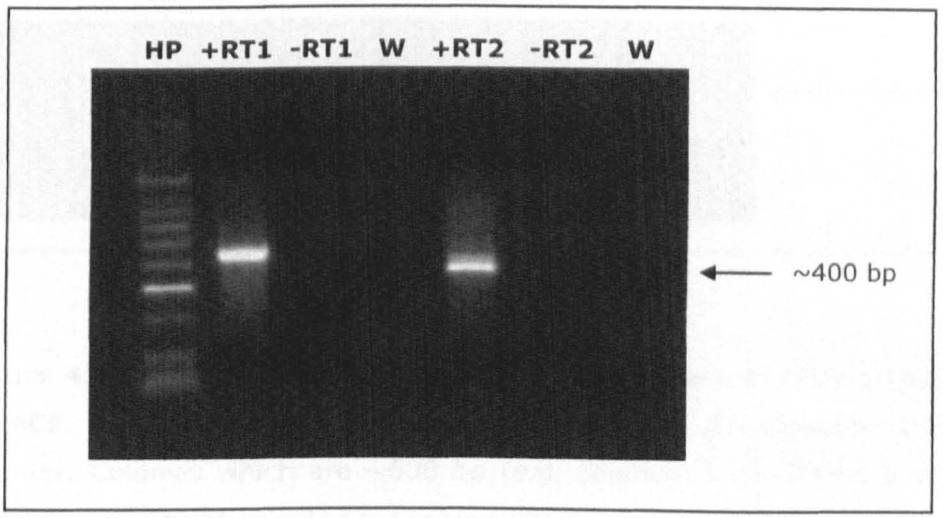



**Figure 4.4 5'RACE PCR of OKA 1 ($S_{o2}$-RNase).** Lane HP represents Hyperladder II, +RT1 represents sample with RT (Reverse Transcriptase) from 1st round PCR, -RT1 represents sample without RT from 1st round PCR, +RT2 represents sample with RT from 2nd round (nested) PCR, sample –RT2 represents sample without RT from second round PCR, W represents water control.

Following successful transformation, three of the colonies having the expected insert size as revealed by colony PCR (Figure 4.5) were selected for plasmid DNA extraction and sequencing of inserts.

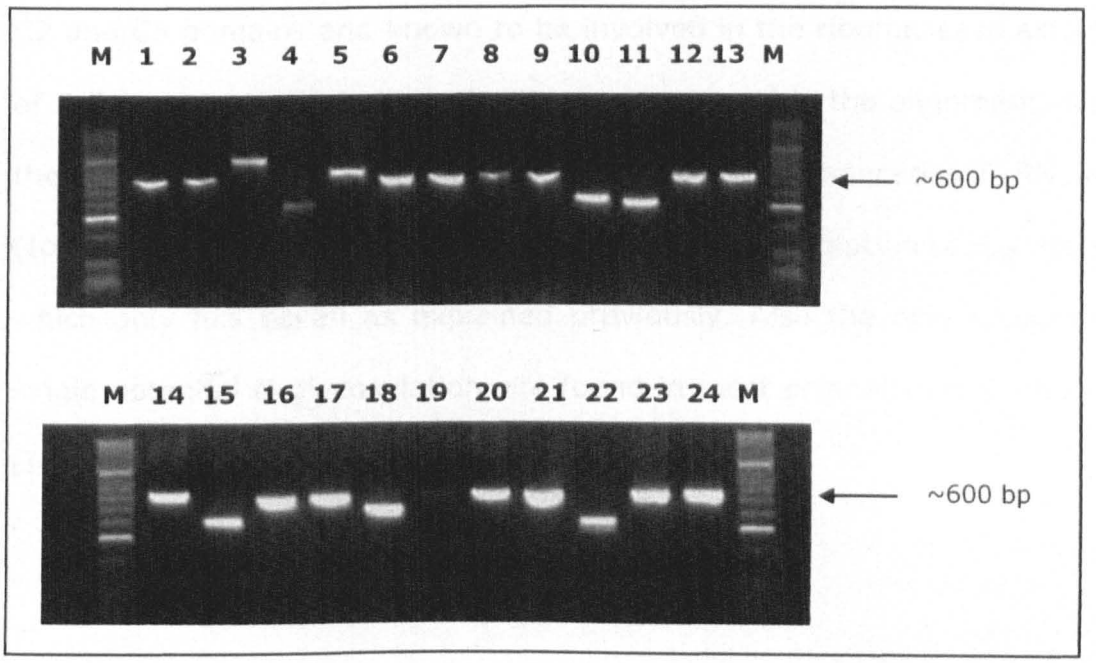



**Figure 4.5 Colony PCR using M13 universal primers of OKA 1 ($S_{o2}$-RNase) 5'RACE.** Lanes M represent Hyperladder II, lanes 1-24 represent transformed colonies. Colonies which are ~600 bp (*e.g.* colonies 1 & 2) have the expected insert size and were selected for sequencing.

### *4.3.4.1 Sequencing and alignment of the 5'RACE products*

The 5'RACE cloning has been found to be successful in terms of the addition of the C1 and C2 domains to the 5' region of the selected S-RNases. The C1 and C2 domains were not part of the initial partial sequence alignment of the cloned S-RNases. The amino acid sequence alignment of the full length of two of the cloned S-RNases ($S_{o2}$ from *S.*

*okadae* and $S_{s2}$ from *S. stenotomum*) with three of the published full length S-RNase sequences from the database is shown in Figure 4.6. The two hypervariable regions (HVa and HVb) and the five conserved domains (C1, C2, C3, C4 and C5) found in all solanaceous S-RNases are indicated. The two conserved histidine residues which are located on the C2 and C3 domains and known to be involved in the ribonuclease activity of S-RNases are also indicated with an arrow head in the alignment. Also the eight conserved cysteine residues found in solanaceous S-RNases (Ioerger *et al.*, 1991) are also indicated, with the exception of $S_{o2}$-RNase which only has seven as explained previously. Also the only conserved single potential N-glycosylation site found in most solanaceous S-RNases (Ioerger *et al.*, 1991) is also shown in the alignment.

```
                                C1                      C2                      HVa
                 10        20        30        40        50        60        70        80        90
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.okaSo2 MIKLQVLSALLILLFALAPVLGNFEYLQLVLQWPATFCHTRTCPVK--PIPNNFTIHGLWPDNK-STMLNNCES--EDKYADISDAKKRK
S.steSs2 .FR-.LA.VFVL..VT.S.GY.........T.........S.-.R---...........-..R..F.SK--HP..NM.M.ED.KE
S.chaS11 .F.SLLT.T.F.V..S.SSTY.D.DK.....T..PS...ANN.QR--I-V.K.........KEGPQL.KY.KP--KLT.NYF..-.MLN
S.chaS12 .F.-SLA.I.I...H-.S.GN.T..Q....FT..TA...KVN.-.R---...........-.RR..F.-K--ST..IKST.EG.KA
S.chaS14 .F.S.LT.VFFM.....S.IY.D..L.E..ST.....YAYG.SRRRP-...........-.II....NVANKER.IK.E.P..LT
                                         *    *         #                    *

                 HVb                C3                      C4
                 100       110       120       130       140       150       160       170       180
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.okaSo2 QLEYHWPDLTANVGDIKKQQGFWGYEFNKHGTCSIDLYNQDAYFDLAIKLKNKFDLLTALRNHGIIPGDVR-TVKNVENAIKAVTTEVPN
S.steSs2 A...I..N..TTKVVS..Y.I..RK.YI.....CS..F...Q......A........NI.GKN..TS.TSHL.SQKIQS...SI.SG...
S.chaS11 D.DK..IQ.KIDQASAR.D.PA.K.QYL...S.CQKI...NT..S..LR..DR....RT.QI.R.V..SSY-.FEEIFD.V.T..QMD.D
S.chaS12 Y...R..N..TTEVES..N.F..EK.YI.....CLP..D.N...K..VD..D.....NL.GK...R..TTHL.SQKIA...RTE.RGI..
S.chaS14 E.DKR..Q.RYDKLYGIDK.YL.EK..L.......NR.K.A......MNI.DR....GT......N..STY-DLDDI.R...T.SIK..S
                                  *

                 C5
                 190       200       210       220
         ....|....|....|....|....|....|....|....|....|....
S.okaSo2 LNCIGDS-GQTMELLEIGICFDRDATQVIACRRRWKSHPNKNKRVTLP-
S.steSs2 .I.SDNFNAG.T..WQ.A..L..N-IA..D.PLPKICTQTGP.GI.F.-
S.chaS11 IK.TEGA-PN---.Y......TPNGDSLVR..QSETCDKTG--KIFFRP
S.chaS12 IS.YD.F-QG.S.........PN-AI.KN.F.PKSCL.KGTTG..F.-
S.chaS14 .K..EKP-PGNV..N.....L.PE.KYTVP.P.TGSC.ELGP-.IKFR-
         *         *    *    *
```

**Figure 4.6 Alignment of the deduced amino acid sequence of the full cDNA sequence of $S_{o2}$-RNase (*S.okaSo2*) cloned from *S. okadae* and $S_{s2}$-RNase (*S.steSs2*) from *S. stenotomum* with three published full length potato S-RNases.** Hypervariable regions (HVa and HVb) and conserved regions (C1-C5) are boxed. The dots in the alignment indicate identites among the 8 sequences. The conserved histidine residues involved in the ribonuclease activity of S-RNases is marked with an arrowhead. The eight conserved cysteine residues are marked with asterisk (*) under the alignment. The only conserved potential N-glycosylation site found in solanaceous S-RNases is marked with hash symbol (#) under the alignment. Accession numbers for the three published *S. chacoense* sequences (*S.chaS11, S.chaS12, and S.chaS14*) are in Table 4.5.

### 4.3.5 Phylogenetic analysis of S-RNases

Phylogenetic trees were constructed using the Neighbour-Joining method as implemented in MEGA 5. Bootstrap values of the tree were calculated based on 1000 replicates in order to allow an estimate of how well the topology at a branch on the tree is supported.

Plant $T_2$-type RNases have been classified into three categories; class I, II and III. All S-RNase genes identified to date belong to the Class III type/group (Igic and Kohn, 2001; Nowak *et al.*, 2011) and type I & II groups are non-S-RNases and are generally referred to as S-like I and S-like II respectively. In an attempt to confirm that our cloned S-RNases belong to the class III RNase group and are genuine S-RNases and not S-like RNases (as evident from the blast search), a phylogenetic tree (Figure 4.7) was constructed based on an alignment using our cloned S-RNases and selected class I & II RNase members from the Solanaceae. The phylogenetic results showed that the S-RNases cluster differently from S-like RNases. The fungal RNase ($T_2$-RNase) which shares high homology with S-RNases (and from which the name $T_2$-type RNase was derived) was used as an out-group to root the phylogenetic tree. The accession numbers of the selected S-like RNases used is shown in Appendix A4.

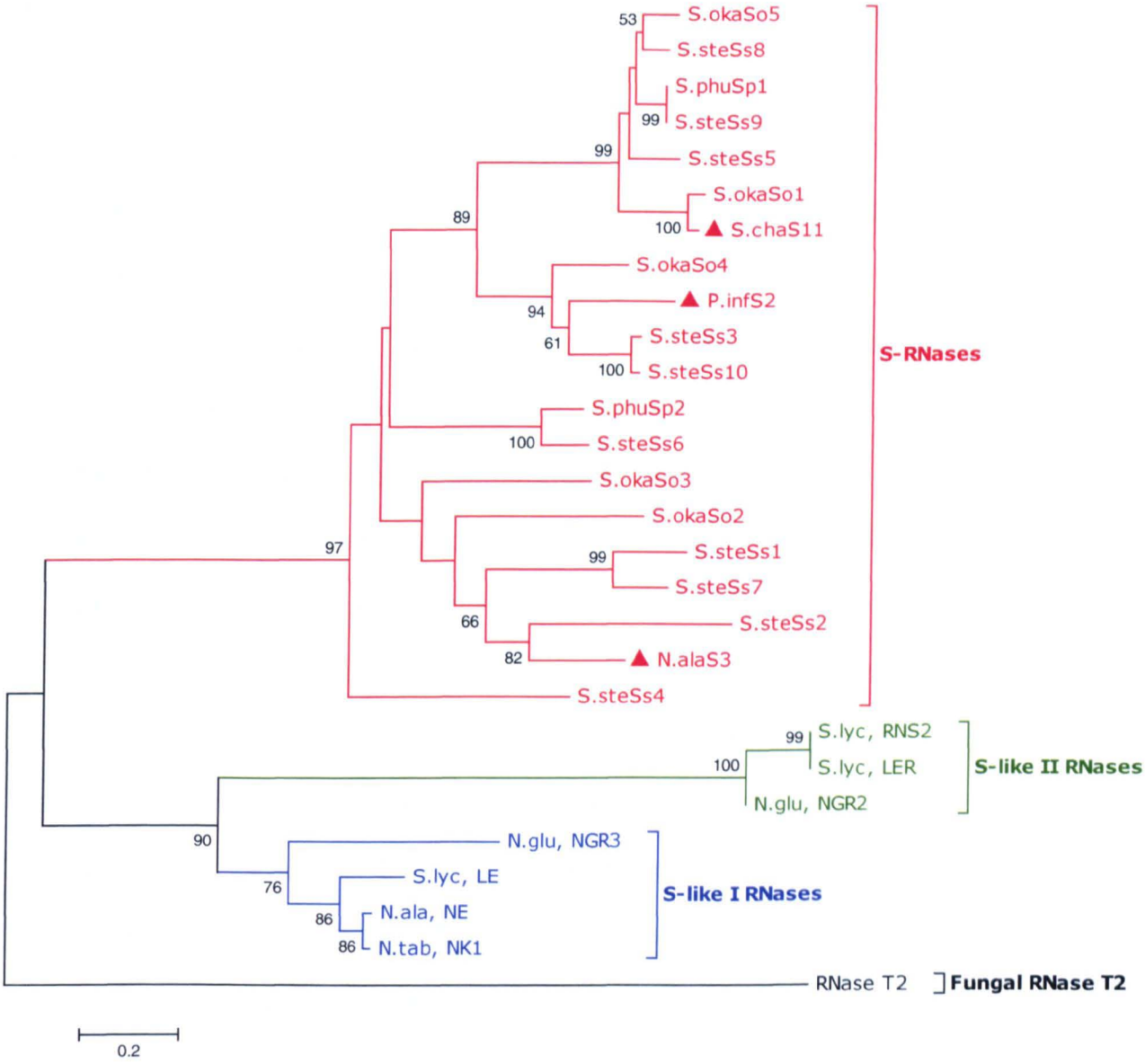



**Figure 4.7 Phylogenetic tree of S-RNases and S-like RNases.** A phylogenetic tree using the partial sequences of the cloned sixteen S-RNases and selected S-like RNases from Solanaceae was constructed. Three published S-RNases from *Petunia*, *Nicotiana* and *Solanum chaoense* (highlighted with a triangle) were included as controls. Fungal RNase $T_2$ of *Aspergillus oryzae* was included as an out-group to root the phylogenetic tree. Numbers are bootstrap values expressed as a percentage and only those exceeding 50% are shown. Bootstrap values were based on 1000 replicates. Phylogenetic tree was drawn using MEGA 5 software. N.ala= *Nicotiana alata,* N.glu= *Nicotiana glutinosa*, N.tab= *Nicotiana tabacum,* S.lyc= *Solanum lycopersicon*, S.oka= *Solanum okadae*, S.phu= *Solanum phureja,* S.ste= *Solanum stenotomum.*

The phylogenetic tree shown in Figure 4.8 was based on an alignment involving the partial deduced amino acid sequence of the sixteen cloned S-RNases and selected (based on the order of their identity) solanaceous S-RNases retrieved from the database (see Appendix A3). Three *Antirrhinum* S-RNase sequences are also included as an 'out-group' to root the phylogenetic tree. Results from the phylogenetic tree showed that, the cloned potato S-RNases are dispersed across many solanaceous lineages. Also interspecific similarities rather than intraspecific similarities are often observed among the S-RNases, *i.e.* the cloned S-RNases do not often cluster into species specific clades but rather formed clades with alleles (closest matches) from other members of the Solanaceae and are supported by high bootstrap values in most cases. Furthermore the S-RNases could be observed in some cases to be clustering with S-RNases from other genera of the Solanaceae rather than forming genus specific clades. The clustering of the S-RNases of one species/genus with members of other species/genus but not with the same species/genus is a commonly observed feature of solanaceous S-RNases (Ioerger *et al.*, 1990).

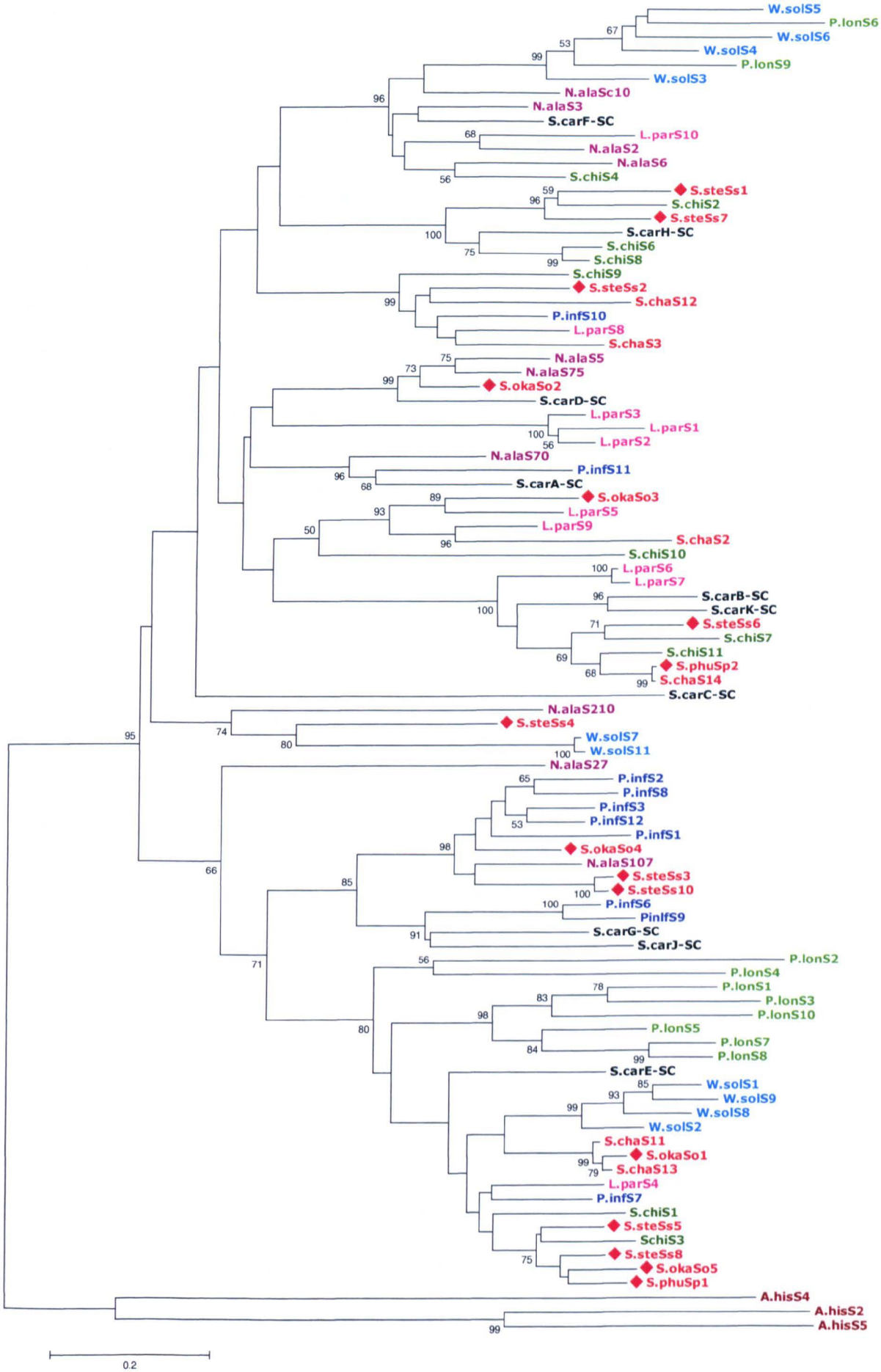

W.solS5
P.lonS6
W.solS6
W.solS4
P.lonS9
W.solS3
N.alaSc10
N.alaS3
S.carF-SC
L.parS10
N.alaS2
N.alaS6
S.chiS4
S.steSs1
S.chiS2
S.steSs7
S.carH-SC
S.chiS6
S.chiS8
S.chiS9
S.steSs2
S.chaS12
P.infS10
L.parS8
S.chaS3
N.alaS5
N.alaS75
S.okaSo2
S.carD-SC
L.parS3
L.parS1
L.parS2
N.alaS70
P.infS11
S.carA-SC
S.okaSo3
L.parS5
L.parS9
S.chaS2
S.chiS10
L.parS6
L.parS7
S.carB-SC
S.carK-SC
S.steSs6
S.chiS7
S.chiS11
S.phuSp2
S.chaS14
S.carC-SC
N.alaS210
S.steSs4
W.solS7
W.solS11
N.alaS27
P.infS2
P.infS8
P.infS3
P.infS12
P.infS1
S.okaSo4
N.alaS107
S.steSs3
S.steSs10
P.infS6
PinlfS9
S.carG-SC
S.carJ-SC
P.lonS2
P.lonS4
P.lonS1
P.lonS3
P.lonS10
P.lonS5
P.lonS7
P.lonS8
S.carE-SC
W.solS1
W.solS9
W.solS8
W.solS2
S.chaS11
S.okaSo1
S.chaS13
L.parS4
P.infS7
S.chiS1
S.steSs5
SchiS3
S.steSs8
S.okaSo5
S.phuSp1
A.hisS4
A.hisS2
A.hisS5
0.2

**Figure 4.8 Phylogenetic tree of S-RNases from Solanaceae.** Sixteen of the cloned S-RNases are indicated with red diamonds. Three *Antirrhinum* S-RNases are included as an out-group to root the phylogenetic tree. Other sequences are selected solanaceous S-RNases retrieved from databases. Numbers are bootstrap values expressed as a percentage and only those exceeding 50% are shown. Bootstrap values were based on 1000 replicates. Phylogenetic tree was drawn using MEGA 5 software. *Solanum* (potato) S-RNases are labelled in red, *Solanum* (tomato) S-RNases are in green, *Solanum carolinense* S-RNases are in black, *Petunia* S-RNases are in blue, *Lycium* in fuchsia, *Nicotiana* S-RNases in purple, *Witheringia* solanaceae S-RNases are in aqua, *Physalis* S-RNases are in lime and *Antirrhinum* S-RNases are in maroon. A.his= *Antirrhinum hispanicum,* L.par= *Lycium parishii*, N.glu= *Nicotiana glutinosa*, P.inf= *Petunia inflata*, P.lon = *Physalis longifolia*, S.car= *Solanum carolinense*, S.chac= *Solanum chacoense*, S.chil= *Solanum chilense*, S.oka= *Solanum okadae*, S.ste= *Solanum stenotomum,* W.sol = *Witheringia solanacae*. Partial sequences of the S-RNases were used for constructing the phylogenetic tree.

## 4.4 DISCUSSION

### 4.4.1 SI status confirmation and the compatibility relationships of the potato stocks

The SI mechanism in angiosperms enables the female reproductive organ of a flower the ability to recognize and distinguish between self-pollen from non-self-pollen and hence to allow only the non-self-pollen to go through to effect fertilization. In the GSI system, depending on the nature of the S-haplotypes expressed in both the pollen and the pistil, different compatibility relationships could be observed. Thus a compatible or semi-compatible reaction will occur when at least one of the haplotypes expressed in the pollen and pistils do not match resulting into successful pollination and the ultimate growth of pollen tubes. However

when both of the expressed S-haplotypes in the pollen and pistil matches it will result in an incompatible reaction leading to the arrest of the growing pollen tubes (see section 1.3).

The diallel cross design has established the compatibility relationship among the *S. okadae* genotypes studied. For instance, crosses between OKA 1 and OKA 3 did not yield any berries or seeds in either direction. This implies that both parents harbour the same pair of S-alleles hence the arrest of any growing pollen tube resulting into an incompatibility reaction. Crosses among all the other *S. okadae* genotypes resulted in the production of berries and seeds in either direction (Table 4.2). From these crosses, two plausible compatibility relationships could be inferred. The first is a compatible reaction (cross) where the parents involved harbour different pair of S-alleles and therefore could not lead to the arrest of the growing pollen tube. Alternatively the parents could differ at least by one S-allele in a reaction called semi-compatible cross where one of the alleles go through to effect fertilization whilst the growth of the pollen tube corresponding to the allele shared by both parents will be arrested. Also all of the self-pollinated *S. okadae* plants did not yield any berries/seeds thereby confirming that the accessions are indeed self-incompatible.

### 4.4.2 S-allele characterization in selected diploid potato plants

Relatively few S-RNase sequences are available in *Solanum* (potato) species compared to other members of the Solanaceae. For instance, as

of March 2012, the NCBI database mining for S-RNases revealed only fourteen S-RNases for potato (Table 4.5), most of which were cloned from *S. chacoense*. Several alleles of *S. stenotomum* have also been recently deposited in the database. There are some additional S-alleles from *S. tuberosum* mentioned in the literature (Kirch *et al.*, 1989; Kaufmann *et al.*, 1991) however not all of them are in public databases.

The use of the C2 domain degenerate primer and the strategy described has enabled the cloning of an additional sixteen novel putative S-RNases from accessions of three diploid potato plants; *S. okadae*, *S. stenotomum* and *S. phureja*. Plants exhibiting the S-RNase based GSI system are expected to be heterozygotes bearing two different S-alleles. However the observation that only one allele could be cloned from some of the accessions could be explained by the differential amplification of the S-RNases by the degenerate primer. Alternatively the unidentified S-RNases may have very low transcript levels relative to the other allele hence could not be detected or amplified. Alleles at the S-locus can show significant variation in their transcript levels or their stability (Roldan *et al.*, 2010).

Analysis of the partial deduced amino acid sequence obtained showed that, all the cloned sequences are genuine S-RNases and they have the primary structural features of solanaceous S-RNases as defined by Ioerger *et al.* (1991). In addition, the cloned partial S-RNase sequences contain one of the active histidine (His) residues located in the C3 region (the other is known to be located in the C2 region) which are involved in the ribonuclease activity of the S-RNase gene (see section 1.3.2.1). The

other histidine which is located in the C2 region is not part of the partial sequences obtained for most of the S-RNases cloned here, because the degenerate primer used for the cloning is from the C2 region hence was cut off from all partial sequences due to nucleotide sequence ambiguity in that region.

The cloned putative S-RNase sequences contain six out of the eight conserved cysteine residues found in functional S-RNases (with the exception of $S_{o2}$-RNase) which can form potential disulphide bonds (Ioerger *et al.*, 1991). The partial putative $S_{o2}$-RNase sequence was observed to contain only five of the expected six cysteine residues. This observation was not surprising since this particular cysteine residue is absent from some published functional solanaceous S-RNases (*e.g.* $S_{o2}$ matches in Table 4.4). The formation of disulphide bridges by the cysteine residues is considered vital for forming and stabilizing the tertiary structure of the proteins (Ishimizu *et al.*, 1996; Ida *et al.*, 2001). All of the cloned putative S-RNase proteins were predicted to have strong basic isoelectric point values (8.6-9.6) which is consistent with observations made with functional S-RNases involved in the self-incompatibility reaction, *i.e.* functional S-RNases are basic proteins having an isoelectric point (pI) value of >7.5 (Nowak *et al.*, 2011), and usually between ~8-10 (Roalson and McCubbin, 2003). All the identified putative S-RNases are novel and no cDNA sequences were available for them prior to this study.

An attempt to clone the full length of selected S-RNases using the partial sequences obtained through the initial 3'RACE cloning has been

successful for two alleles. The full length cloning of the $S_{o2}$-RNase (one of the S-RNases earmarked for silencing in our transgenic work) from *S. okadae* and $S_{s2}$-RNase from *S. stenotomum* has been successfully performed and the sequences confirmed. The full length sequences revealed the two conserved catalytic histidine residues involved in the ribonuclease activity of S-RNases and located in the C2 and C3 regions (Figure 4.6). Also the eight conserved cysteine residues could be observed in $S_{s2}$-RNase and only seven for $S_{o2}$-RNase which lacks one such residue as mentioned earlier.

Variable number of potential N-glycosylation sites can be identified in S-RNase sequences (*e.g.* Oxley *et al.*, 1998; Qin *et al.*, 2001). However the analysis of solanaceous S-RNase sequence has identified one single conserved potential N-glycosylation site found in the C2 conserved region (Ioerger *et al.*, 1991) and this was observed to be conserved in the full length sequence of the two cloned putative S-RNases. The presence of this conserved glycosylation site could possibly be responsible for modulating the S-RNase ribonuclease activity (Ioerger *et al.*, 1991). However studies have shown that, the removal of the glycan-side chains and glycosylation effect did not alter the enzymatic activity of the S-RNase gene both *in vitro* and *in vivo* (Broothaerts *et al.*, 1991; Karunanandaa *et al.*, 1994). The full length and the partial length sequence of the cloned putative S-RNases have typical primary structural features of functional solanaceous S-RNases and represent novel additions to the relatively few S-RNase sequences currently available in *Solanum*.

### 4.4.3 Solanaceous S-RNases exhibit allelic diversity and intraspecific-sequence polymorphism

An early report of the extreme level of polymorphism at the S-locus of a narrow endemic species, *Oenothera organensis* which was estimated to contain *ca.* 500 individuals (Emerson, 1939) generated interest in the understanding of the gametophytic S-locus of angiosperms by population geneticists (Igic *et al.*, 2003). The low level of amino acid sequence similarity observed in the partial sequences of the cloned putative S-RNases (32.9%) is consistent with earlier observations made in solanaceous S-RNases (Ioerger *et al.*, 1990; McCubbin and Kao, 2000). Ioerger *et al.* (1990) analysed the S-locus of three species of the Solanaceae and observed a low level of amino acid sequence similarity within species as low as 40%. This is consistent with this study where the amino acid sequence similarity within species was as low as 32.9 - 44.2%. Polymorphism at the S-locus and the high level of sequence divergence in solanaceous S-RNases could partly account for these observations.

Solanaceous S-allele polymorphism could lead to observations where S-alleles of one species or genus were found to be more closely related to alleles from other species or genus; a case of trans-specific or trans-generic evolution of these S-alleles (Ioerger *et al.*, 1990). Thus the S-RNases are exceptionally old and have been inherited from a common ancestor and passed down to multiple descendant taxa, *i.e.* the S-allele lineage origins predates their current species origin. Polymorphism at the S-locus in angiosperms is a result of diversifying selection, the age of S-

alleles and the absence of recombination at the S-locus which acts to preserve and maintain allelic variations at the S-locus (Ioerger *et al.*, 1991; Richman and Kohn, 2000; Igic *et al.*, 2003; section 1.4.2).

### 4.4.4 Phylogenetic analysis of S-RNases

The use of molecular phylogenetic analysis tools has enabled robust phylogenetic trees to be constructed leading to the identification of the most basal lineages of angiosperms. With the help of phylogenetic studies, S-genes have been characterised allowing conclusions to be drawn about the evolutionary history of their sequences (Allen and Hiscock, 2008). For instance, the use of phylogenetic tools has enabled Igic and Kohn (2001), to make predictions that GSI is ancestral to ~75% of eudicots and that RNase-based self-incompatibility of the GSI system was the ancestral state of self-incompatibility system existing in the majority of dicots.

#### *4.4.4.1 S-RNases are distinct from S-like RNases*

All the identified S-RNases known to be involved in the GSI reaction in plants belong to the $T_2$-type Class III gene subfamily (Igic and Kohn, 2001). Other non-S-RNases called S-like RNases have also been identified in many plant species. These S-like RNases share domain structure with S-RNases and appear in phylogenetic analyses to be related to S-RNases but have no role in self-incompatibility reaction. In an attempt to show that, the cloned *Solanum* putative S-RNases are

genuine S-RNases and not S-like RNases, a phylogenetic tree was constructed using an alignment of the cloned S-RNases and selected S-like RNases from Solanaceae (Figure 4.7). The observation from the phylogenetic analysis that the non-S-RNases (S-like I and S-like II-RNases) fall outside the S-RNase clade was not surprising. S-like RNases are known to be unlinked to the S-locus (and are not involved in GSI) and may be involved in pathogen defence mechanisms or induced in response to phosphate starvation (Kao and McCubbin, 1996; Dodds *et al.*, 1996; Hugot *et al.*, 2002), hence are expected to cluster differently from S-RNases.

Although the relationship between S-RNases and S-like RNases still remains equivocal, the S-like RNases may be the ancestral genes involved in defence against pathogen attack in the style that were recruited with modifications to function in the self-incompatibility reaction (Kao and McCubbin, 1996). All the putative S-RNases cloned from this study represents genuine and novel set of alleles in *Solanum* and are clearly distinct from and distantly related to S-like RNases.

#### *4.4.4.2 Phylogenetic analysis of solanaceous S-RNases*

The use of phylogenetic tools has enabled us to compare the cloned S-RNases with other alleles from the Solanaceae. The general observation that the S-RNases do not form species-specific clades (Figure 4.8) and that some interspecies clades (showing interspecific similarities) are higher than intraspecies specific ones is not surprising. Some accessions

however, could be observed to form intraspecific clades but the topologies at some of those branches are not supported by high bootstrap values (*e.g. P. inflata* alleles). Similar results were obtained by Ioerger *et al.* (1990), when they analysed the S-locus of three species of the Solanaceae and observed a higher interspecies similarity rather than intraspecies similarity. They concluded that, polymorphism at the S-locus predates the divergence of the species in the Solanaceae and this polymorphism has been maintained to the present time. This extreme level of polymorphism in S-proteins indicates an unusual event of evolutionary forces which are operating at the S-locus (Ioerger *et al.*, 1990).

Furthermore, from the phylogenetic tree (Figure 4.8), some trans-generic clades could be observed. For instance, S-RNases from *Solanum* (Potato, tomato), *S. carolinense, Lycium, Petunia* and *Nicotiana* could be observed to form extensive trans-generic clades indicating extensive diversifications of S-alleles in all these genera. However, a reduced or very limited trans-generic lineages could be observed in S-RNases from *Witheringia* and *Physalis, i.e.* they could be observed to form almost genus specific clades possibly indicating that very little diversification of S-alleles occurred in these two genera particularly in *Physalis*. Limited diversification of S-alleles in a genus can cause the clustering together of the alleles in a phylogenetic tree (Richman and Kohn, 2000). Similar observations were made by Richman *et al.* (1996) where they interpreted their observations that the loss of trans-generic lineages in *Physalis crassifolia* was an outcome of the effect of severe population bottleneck

imposed on this genus. Similar explanations could account for the reduced or very limited trans-generic lineages observed in *Witheringia* (Stone and Pierce, 2005). The observation of trans-generic evolution in solanaceous species is evidence for the inheritance of several S-alleles through events of multiple speciation (Richman *et al.*, 1996).

# CHAPTER 5: S-RNase GENE EXPRESSION ANALYSIS

## 5.1 INTRODUCTION

### 5.1.1 S-RNases have unique role in pollen rejection

A glycoprotein with RNase activity termed S-RNase is the pistil expressed S-gene product in plant families that exhibit the S-RNase based GSI system (Bredemeijer and Blaas, 1981; Anderson *et al.*, 1986; McClure *et al.*, 1989). Predictions that, genes encoding SI specificity in the pistil should exhibit high allelic sequence diversity formed the initial basis for the methods used for the S-RNase protein identification. These predictions were made when S-haplotypes defined by pollinations were observed to be co-segregating with polymorphic pistil proteins based on their molecular mass/and or isoelectric point differences (McCubbin and Kao, 2000). Subsequent to the cloning of the first S-RNase gene and the observation that high concentrations could be found in the stylar region where incompatible pollen tube inhibition occurs (Anderson *et al.*, 1986), *in situ* hybridization of the pistils of *Nicotiana* was used to confirm that, the expression of the gene was located in the transmitting tract of the style (Cornish *et al.*, 1987).

Following the production of S-RNases by the cells of the stylar transmitting tract and their subsequent secretion into the extracellular matrix of the style, the S-RNase cytotoxic effect is believed to cause the arrest of incompatible pollen tubes in the style by the degradation of RNA in the growing pollen tubes (McClure *et al.*, 1989). The secretion and

accumulation of S-RNases inside pollen tubes in an S-haplotype independent manner has been demonstrated (Luu *et al.*, 2000; Goldraij *et al.*, 2006), although the mode of S-RNases penetration and entering of pollen tubes still remains uncertain (Qin *et al.*, 2006). However, it is assumed that endocytosis through their inclusion into a membrane bound compartment or through the recognition of a conserved domain of the S-RNase by a receptor or a receptor complex (McClure, 2004; Kao and Tsukamoto, 2004) could be responsible for this. All the identified S-RNases have been observed to have two catalytic Histidine (His) residues, a similarity which is shared with RNase $T_2$ of *Aspergillus oryzae*. These two His residues were found to be important and necessary for the ribunuclease activity of S-RNases thereby initially linking RNase enzymatic activity in the self-incompatibility reaction (Kawata *et al.*, 1988; McClure *et al.*, 1989; Ioerger *et al.*, 1991), a notion which has been proven later unequivocally by several researchers in the field of SI (*e.g.* Royo *et al.*, 1994; Lee *et al*, 1994).

Through the use of developmental and transgenic approaches, it was observed that, low to moderate expression of the S-RNase gene in the styles of self-incompatible plants did not cause the arrest of self-pollen tubes, *i.e.* the ability of the S-RNase to reject self-pollen tubes is dependent on a threshold level (Clark *et al.*, 1990; Lee *et al.*, 1994; Murfett *et al.*, 1994, Zurek *et al.*, 1997). Although the minimum amount of S-RNases required for causing the arrest of incompatible pollen tubes has not been well determined empirically in SI studies, Qin *et al.* (2006) has developed a method for quantifying this threshold in single styles of

self-incompatible *Solanum chacoense* lines and has determined a definite threshold for self-pollen tube rejection in a particular genotype of *S. chacoense*. However, variations could occur in this S-RNase threshold required for pollen rejection depending on the pollen donors' genetic background. Qin *et al.* (2006) concluded that, SI phenotype determination is caused primarily by S-RNases.

## 5.2 MATERIALS AND METHODS

### 5.2.1 Selection of reference genes for quantitative RT-PCR

Six pairs of potato housekeeping gene primers (Nicot *et al.*, 2005) were ordered and tested on our potato stocks to select those that will be useful for the S-RNase gene expression work in this study. Primers coding for the gene *actin* (*Actin*), *β-tubulin* (*β-tub*), *cyclophilin* (*CyP*), *elongation factor 1-α* (*ef1-α*), *adenine phosphoribosyl transferase* (*aprt*) and *cytoplasmic ribosomal protein* (*L2*) were tested (Table 5.1).

**Table 5.1** Housekeeping gene primers evaluated for gene expression studies

| Primer Name | Direction | Sequence (5'-3') | Tm °C |
|---|---|---|---|
| **Actin-F** | Forward | GCTTCCCGATGGTCAAGTCA | 67.5 |
| **Actin-R** | Reverse | GGATTCCAGCTGCTTCCATTC | 66.9 |
| **β-tub-F** | Forward | ATGTTCAGGCGCAAGGCTT | 67.0 |
| **β-tub-R** | Reverse | TCTGCAACCGGGTCATTCAT | 67.7 |
| **CyP-FP** | Forward | CTCTTCGCCGATACCACTCC | 65.9 |
| **CyP-RP** | Reverse | TCACACGGTGGAAGGTTGAG | 66.3 |
| **ef1-αF** | Forward | ATTGGAAACGGATATGCTCCA | 65.2 |
| **ef1-αF** | Reverse | TCCTTACCTGAACGCCTGTCA | 66.8 |
| **aprt-F** | Forward | GAACCGGAGCAGGTGAAGAA | 66.5 |
| **aprt-R** | Reverse | GAAGCAATCCCAGCGATACG | 66.7 |
| **L2-F** | Forward | GGCGAAATGGGTCGTGTTAT | 65.7 |
| **L2-R** | Reverse | CATTTCTCTCGCCGAAATCG | 66.7 |

### 5.2.2 Semi-quantitative RT-PCR

Following the first strand cDNA synthesis (RT-reaction) as described in section 2.2.4., PCR reactions were set up with variable number of cycles (typically between 18 and 24 cycles were tested). The PCR reactions were performed using haplotype specific primers (Table 5.2) designed for this experiment. Semi-quantitative RT-PCR amplification was performed in a 25 µl total reaction mix comprising a 0.5 µl aliquot of RT reaction (cDNA reaction), 20 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTPs (Bioline, London, UK), 0.4 µM of forward primer, 0.4 µM of reverse primer and 2U of Taq DNA polymerase (Bioline, London, UK). Amplification was performed in a PTC-200 Thermal Cycler (MJ Research, Watertown MA, USA) under the following cycling conditions: an initial 3 min denaturation at 94°C, followed by between 18 and 24 cycles of 30 sec at 94°C, 30 sec annealing at 62°C, 20 sec at 72°C and a final extension of 5 min at 72°C.

**Table 5.2** S-RNase allele specific primers (QRT-PCR primers) used for gene expression studies

| Primer Name | Direction | Sequence (5'-3') | Tm °C |
|---|---|---|---|
| **qpcrSo1-F** | Forward | TGGATCCTGCTGTCAGAAAA | 63.4 |
| **qpcrSo1-R** | Reverse | GTTCCGGTGCTCCTTCAGTA | 64.2 |
| **qpcrSo2-F** | Forward | ATCACTGGCCTGATTTGACC | 63.9 |
| **qpcrSo2-R** | Reverse | GTTTGGCCAGAATCACCAAT | 63.6 |

The $S_{o1}$ and $S_{o2}$ primers were designed to be allele specific based on the alignment of the two gene sequences. See Appendix A5 for the position of the primers in the sequence of the two alleles.

### 5.2.3 Tissue specificity and developmental expression of the S-RNase gene

To test the tissue specific expression of the S-RNase gene, RNA was extracted using the RNeasy Mini kit (section 2.2.1) from individually harvested potato leaves, petals, anther cone, bud, stigma and styles. Furthermore to test for the developmental expression of the S-RNase gene, pistils were harvested from potato flowers at the following stages of floral development; namely the bud stage, *i.e.* immature bud, the balloon stage; *i.e.* before the flower reflexes, and the open stage; *i.e.* fully opened flower. The bud stage, the balloon stage and open stage will be referred to as Stage 1, Stage 2 and Stage 3 respectively in this report (see Figure 5.4 at results section for description of the stages). First strand cDNA (RT- reaction) was then synthesized from these materials (section 2.2.4) and semi-quantitative RT-PCR performed on them (section 5.2.1). Typically $S_{o1}$- and $S_{o2}$-RNase expressions were checked and semi-quantitative RT-PCR approach was used in an attempt to quantify the expression of the two genes present together in the tested plant materials. *Cyclophilin* (*CyP*) housekeeping gene primer (section 5.2.1) was selected and used as a control.

### 5.2.4 Real-time quantitative RT-PCR

Real-time quantitative RT-PCR (QRT-PCR) was performed on the reverse transcribed RNA samples (in section 5.2.3) in a real-time cycler. A master mix for each PCR run was performed using Quantitect SYBR® Green PCR

kit (QIAGEN, Germany). The reaction components comprised 1X Quantitect SYBR® Green PCR master mix (HotStarTaq® DNA polymerase, Quantitect SYBR® Green PCR buffer containing Tris-HCl, KCl, $(NH_4)_2SO_4$, and 5 mM $MgCl_2$ (pH 8.7), dNTP mix including dUTP of ultrapure quality, SYBR Green I and ROX™ fluorescent dyes), 0.5 µl of the RT-reaction and 0.22 µM of forward and reverse primers (Table 5.1 & 5.2). The QRT-PCR reaction run was performed in triplicates for each of the evaluated samples. Amplification was performed for 40 cycles in a LightCycler® 480 II instrument (Roche Diagnostics Ltd., Switzerland) under the following cycling conditions: an initial activation step for the HotStarTaq® DNA polymerase for 15 min at 95°C , followed by 15 sec denaturation at 95°C, 30 sec annealing at 62°C, 30 sec at 72°C for the PCR amplification. An additional data was acquired for the dissociation curve for 1 min at 95°C, 30 sec at 62°C and 30 sec at 95°C. Following data acquisition as $C_t$ values, a relative quantification method was used to determine the relative expression of the S-RNase gene in the tissues examined. *Cyclophilin* (*CyP*) housekeeping gene primer (section 5.2.1) was selected and used for normalising the expression levels. The dissociation curve data was used for melting curve analysis in order to check the specificity of the primers used.

## 5.3 RESULTS

### 5.3.1 Selection of reference genes for quantitative RT-PCR

The cDNA synthesized from the potato plant tissues to be used for the expression work was amplified using housekeeping gene primers from potato as previously published in order to choose those that will be useful in this current study. Results from the PCR showed that, depending on the type of tissues used, not all the evaluated housekeeping gene primers worked efficiently and reliably. For instance when the housekeeping gene primers were evaluated on tissues to be used for checking the developmental expression of the S-RNase gene, some of the primers produced multiple amplification (Figure 5.1 a & b), resulting in

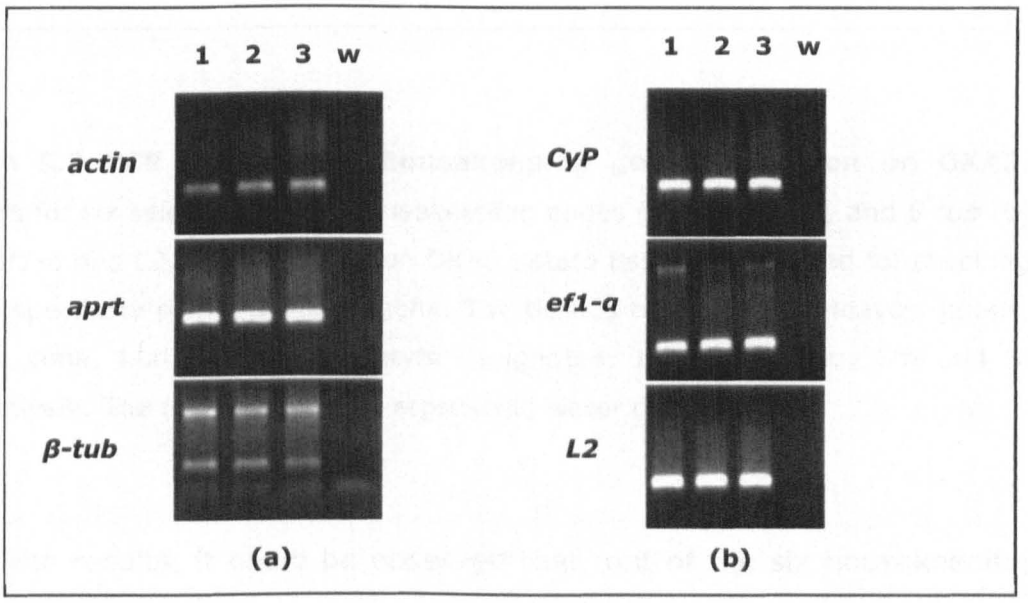



**Figure 5.1 PCR profiles for housekeeping gene evaluation on OKA1.** Primers for six selected potato housekeeping genes (a) *Actin, aprt* and *β-tub* (b) *CyP, ef1-a* and *L2,* were checked on OKA1 potato pistils at three different developmental stages namely bud stage (Stage 1), balloon stage (Stage 2) and open flower stage (Stage 3) designated: 1, 2 and 3 respectively. The (w) in the figure represents water control.

non-specific banding patterns and hence were discarded. Similar results were obtained for tissues used for checking the tissue specificity of the S-RNase gene expression (Figure 5.2 a & b).

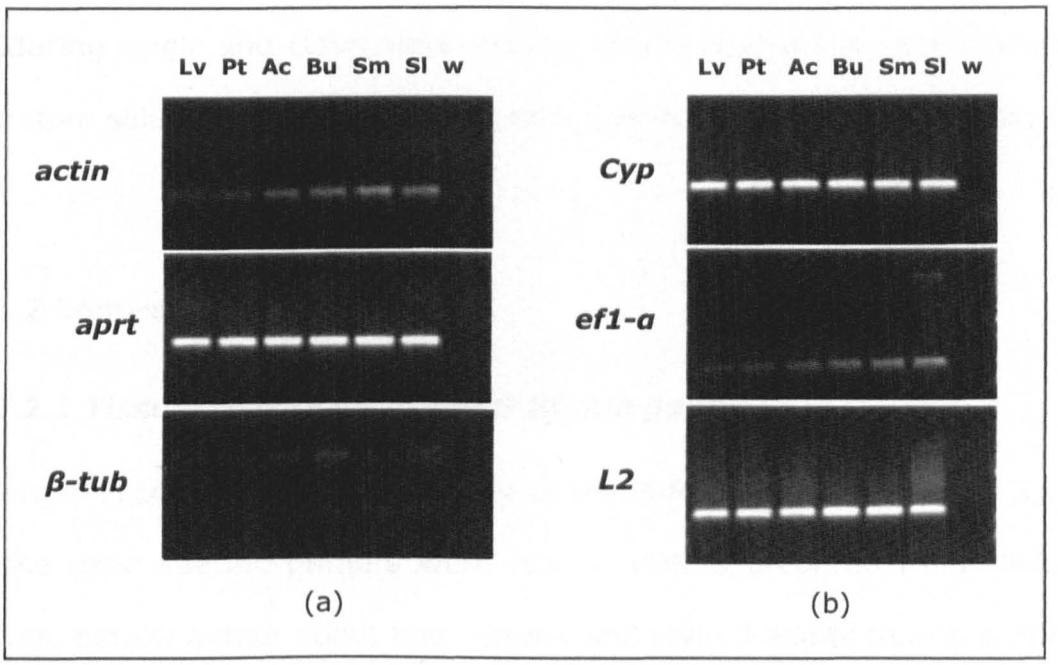



**Figure 5.2 PCR profiles for housekeeping gene evaluation on OKA3.** Primers for six selected potato housekeeping genes (a) *actin, aprt,* and *β-tub* (b) *CyP, ef1-a* and *L2*, were checked on OKA3 potato tissues to be used for checking tissue specificity of the S-RNase gene. The tissues tested include leaves, petals, anther cone, bud, stigma and style designated: Lv, Pt, Ac, Bu, Sm and Sl respectively. The (w) in the figure represents water control.

From the results, it could be observed that, out of the six housekeeping genes evaluated, *β-tub* and *ef1-a* primer pairs produced multiple bands in almost all the samples tested (Figure 5.1 a & b; Figure 5.2 a & b). However, primers coding for *Actin*, *CyP*, *aprt* and *L2* were found to produce single amplicons corresponding to the expected sizes in all cases (Figure 5.1 a & b; Figure 5.2 a & b). However not all of them gave very

clean amplicons *i.e.* some could be seen producing smears (*e.g. L2* primers). Also not all of them gave consistent banding patterns, thus they showed variation in their expression in all the tested tissues (*e.g. Actin* primers). The *aprt* and *CyP* primers appear to be consistent producing single and clean amplicons for all the tested tissues and were therefore selected to be the housekeeping genes of choice in this study.

### 5.3.2 Semi-quantitative RT-PCR

#### *5.3.2.1 Tissue-specificity of the S-RNase gene*

In order to test the tissue specificity of the S-RNase gene, $S_{o1}$- and $S_{o2}$-RNase gene specific primers were run on cDNAs prepared from OKA3 leaves, petals, anther cone, bud, stigma and style. Results from the PCR amplification (Figure 5.3) showed that the expression of the S-RNase gene is restricted to the floral tissues, specifically the stigma and style tissues, which is consistent with what is known of the S-RNase gene. A semi-quantitative RT-PCR approach with reduced number of cycles was used for the amplification in an attempt to quantify which of these floral tissues expressed the S-RNase gene at the highest level. From the results, it could be observed that the S-RNase gene is expressed slightly higher in the styles than the stigma. No expression (no amplification) could be detected in the leaves, petals, anther cone and bud tissues. Also, it could be observed from the results that, the expression of $S_{o2}$-RNase is slightly higher than the expression of $S_{o1}$-RNase in both the stigma and style of OKA3. The results from this experiment showed that

the expression of both alleles of the S-RNase gene in OKA3 is highly tissue specific.

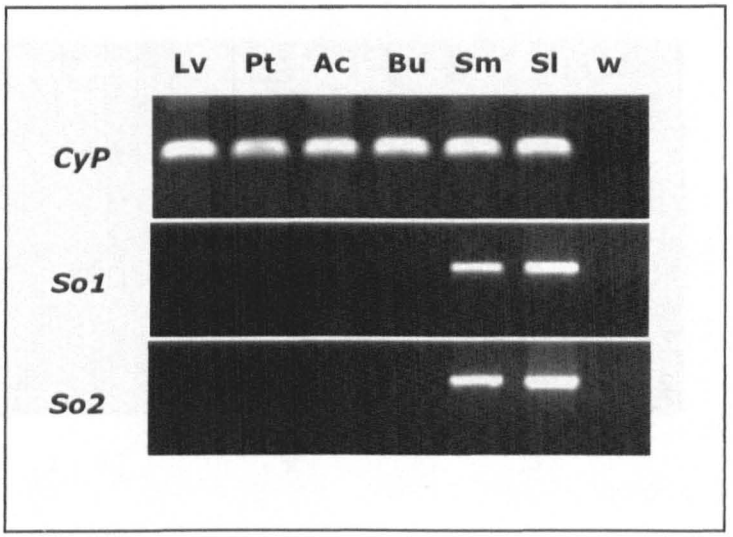
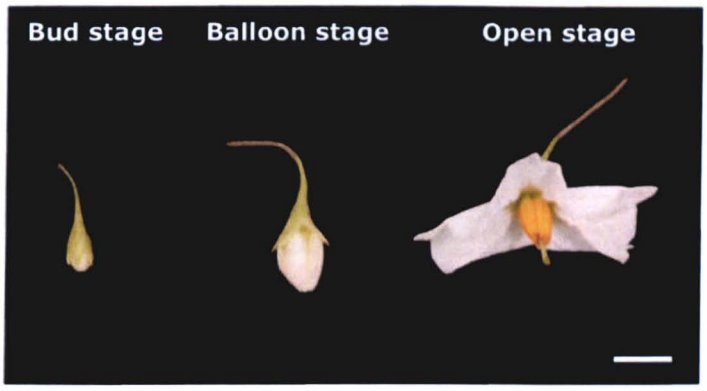



**Figure 5.3 Semi-quantitative RT-PCR profiles for OKA3 tissues.** $S_{o1}$- and $S_{o2}$-RNase gene specific primers were used in RT-PCR reactions with RNA from various OKA3 potato tissues to check for the tissue specificity of the S-RNase gene. The PCR also attempts to quantify the expression level of the S-RNase gene in a semi-quantitative RT-PCR with 20 cycles of amplification. The tissues used include leaves, petals, anther cone, bud, stigma and style designated Lv, Pt, Ac, Bu, Sm and Sl respectively. The (w) in the figure represents water control. *Cyp* housekeeping gene primers were used as a control to check the quality and quantity of the synthesized cDNAs from the tissues.

### *5.3.2.2 Developmental expression of the S-RNase gene*

S-RNase transcript accumulation is known to occur in a developmentally regulated manner in floral organs. To check the developmental expression of the $S_{o1}$- and $S_{o2}$- RNase gene in OKA1, S-RNase sequences were amplified in pistils harvested from floral tissues at different

developmental stages including the bud (Stage 1), balloon (Stage 2) and open flower (Stage 3) stages (Figure 5.4).




1 2 3

**Figure 5.4 The floral developmental stages from which pistils were harvested in OKA1.** (1) The bud stage refers to immature buds (2) the balloon stage refers to the stage just before the flower bud reflexes and (3) the open stage refers to a fully opened mature flower. Pistils were harvested from these three stages for both the semi-quantitative RT-PCR and QRT-PCR in order to test for the developmental expression of the S-RNase gene. Scale bar: 1 cm.

From the semi-quantitative RT-PCR results (Figure 5.5), it could be observed that, the tissues followed the expected developmental profile of the S-RNase gene expression. There were higher levels of expression found in the opened flower (Stage 3), followed by the balloon stage (Stage 2) and then the bud stage (Stage 1) where the lowest expression of the gene was observed. Also it appears the $S_{o1}$-RNase was expressed at a slightly low level compared to $S_{o2}$-RNases for all the three stages studied.

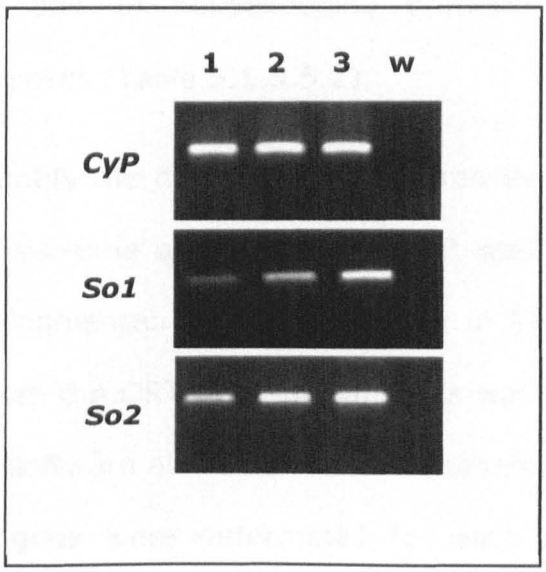



**Figure 5.5 Semi-quantitative RT-PCR for the developmental stage expression of $S_{o1}$- and $S_{o2}$-RNase in OKA1 pistils.** $S_{o1}$- and $S_{o2}$-RNase gene specific primers were used in RT-PCR reactions with RNA from various OKA1 pistils at different floral developmental stages. The semi-quantitative RT-PCR involved 20 cycles of amplification in an attempt to quantify the expression level of the $S_{o1}$- and $S_{o2}$-RNase in OKA1 at the different floral stages. The three stages used include the bud stage (Stage 1), balloon stage (Stage 2) and open flower stage (Stage 3) designated: 1, 2 and 3 respectively. The (w) in the figure represents water control. *Cyp* housekeeping gene primer was used as a control to check the quality of the synthesized cDNAs used for the PCR.

### 5.3.3 Quantitative RT-PCR

Although semi-quantitative RT-PCR approach could give an idea of the relative expression levels of a gene in a tissue, it cannot be used for accurate transcript quantification studies. In order to accurately quantify the expression of a particular gene, a quantitative RT-PCR approach is necessary. Real-time quantitative RT-PCR was performed using the LightCycler® 480 II instrument (Roche Diagnostics Ltd., Switzerland) using two pairs of gene specific primers amplifying the $S_{o1}$- and $S_{o2}$-

RNase and one pair of housekeeping primers amplifying *CyP* for normalisation purposes (Table 5.1 & 5.2).

To check and quantify the developmental expression of the $S_{o1}$- and $S_{o2}$-RNase in OKA1, real-time quantitative RT-PCR was performed on pistils at different developmental stages (described in Figure 5.4). Following data collection from the QRT-PCR run, analysis was performed using the LightCycler® 480 software attached to all instruments. Relative quantities of the S-RNase gene were determined for each of the tissues after normalisation to housekeeping gene (*CyP*) levels. The bud stage was used as a calibrator for the other stages and was given a value of 1. Results from the analysis showed that, there were significant differences in the expression of the S-RNase gene at different stages of pistil development in OKA1 (Figure 5.6). The highest expression level was detected in the fully opened flower stage (Stage 3), followed by the balloon stage (Stage 2) and then the bud stage (Stage 1) being the least for both $S_{o1}$- and $S_{o2}$-RNases. The results showed that, the S-RNase genes in OKA1 were developmentally expressed with the highest expression found in mature pistils.

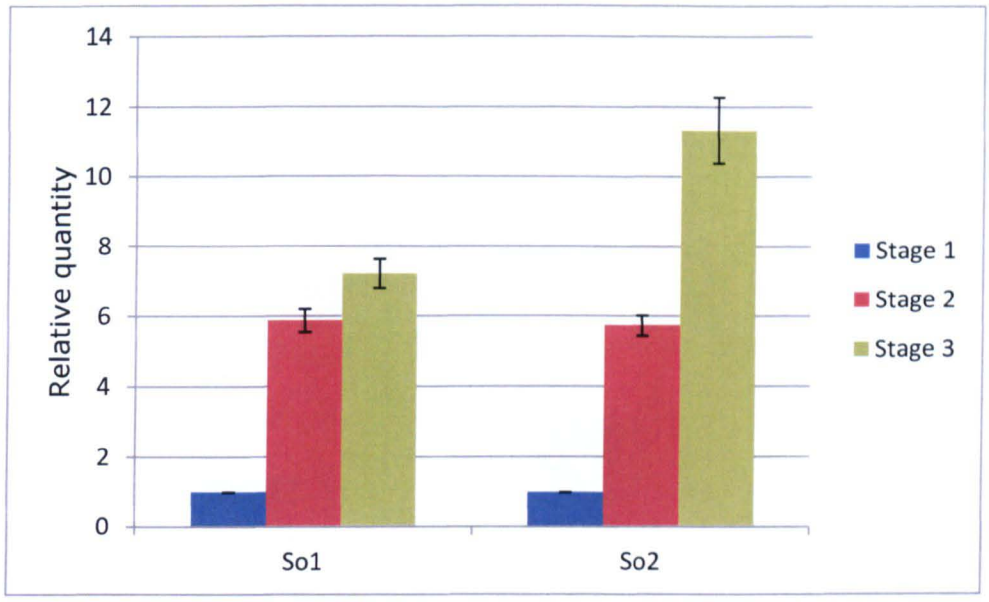



**Figure 5.6 Relative expression of the S-RNase gene at different pistil developmental stages.** Relative quantities of the $S_{o1}$- and $S_{o2}$-RNase were determined at different developmental stages of the pistils in OKA1. The stages used include the bud stage (Stage 1), balloon stage (Stage 2) and open flower stage (Stage 3). The bud stage (Stage 1) was used as a calibrator and was given a value of 1.

The relative expression of the $S_{o1}$- and $S_{o2}$-RNase gene was also determined in the stigma and style of OKA3, to evaluate where the S-RNase gene accumulates the most. Results from the QRT-PCR analysis (Figure 5.7) showed that, there were significant differences in the expression of $S_{o1}$- and $S_{o2}$-RNase in the stigma and styles of OKA3. The S-RNase gene was found to be expressed more in the style than in the stigma for both the $S_{o1}$- and $S_{o2}$-RNase.

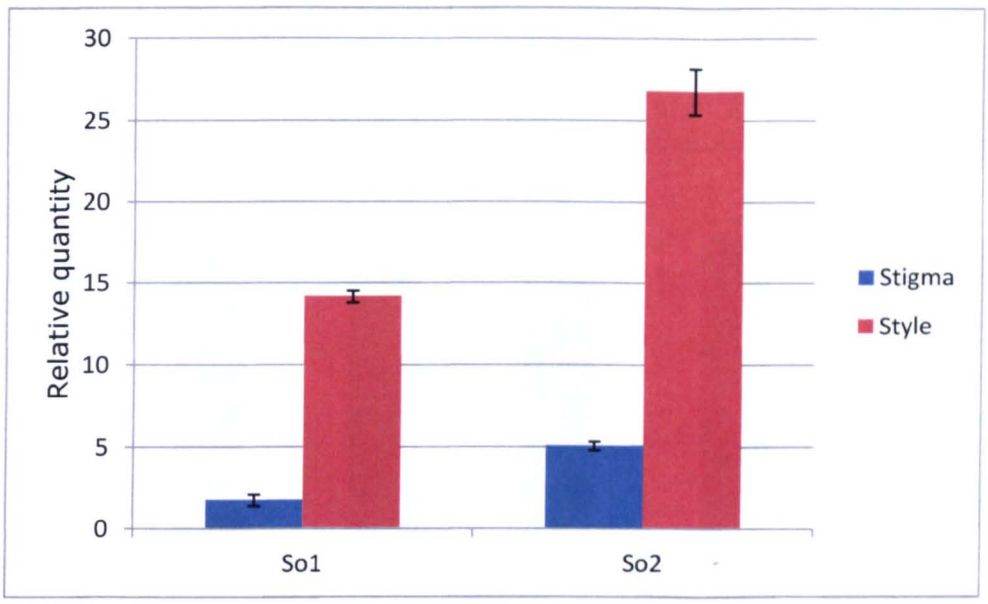



**Figure 5.7 QRT-PCR of the S-RNase gene accumulation in the stigma and style.** Relative expressions of the $S_{o1}$- and $S_{o2}$-RNase were determined in the stigma and styles of OKA3.

The expression of the same S-allele could differ in different genetic backgrounds. The genotyping of OKA1 and OKA3 in this current study revealed that, they both harbour the $S_{o1}$- and $S_{o2}$-RNase. Therefore an attempt was made to check the expression of the $S_{o1}$- and $S_{o2}$-RNase in the pistils of these two plants (*i.e.* both OKA1 and OKA3). Results from the QRT-PCR analysis (Figure 5.8) showed that, there are some slight differences in the expression of the $S_{o1}$- and $S_{o2}$-RNase gene in OKA1 and OKA3. However, the differences in the expression that was observed are not highly significant.

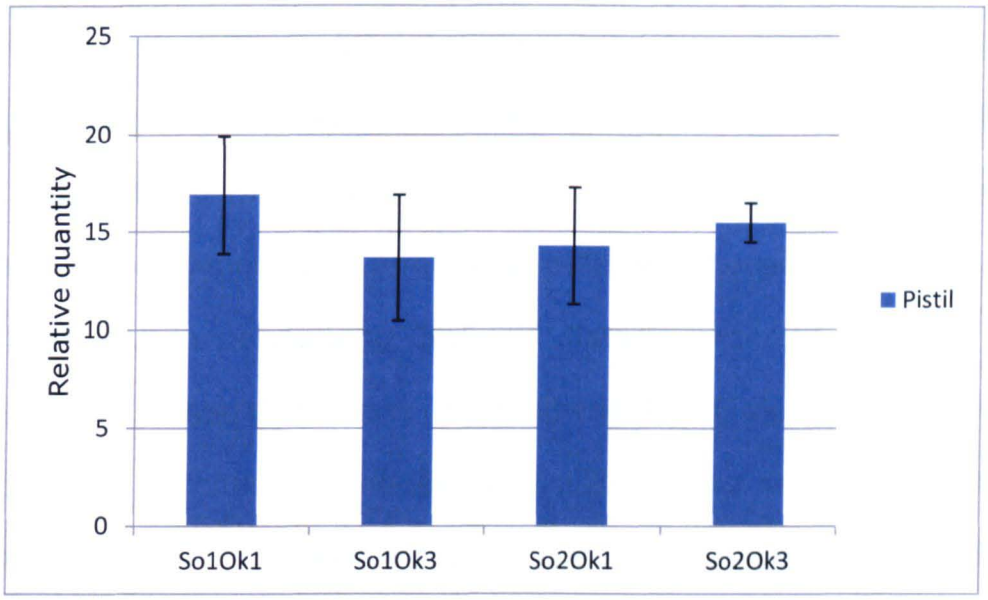



**Figure 5.8 Relative expression of the $S_{o1}$- and $S_{o2}$-RNase gene in different plant backgrounds.** Relative quantities of the $S_{o1}$- and $S_{o2}$-RNase were determined in the pistils of both OKA1 and OKA3. So1Ok1 represent $S_{o1}$-RNase in OKA1, So1Ok3 represent OKA3 $S_{o1}$-RNase, So2Ok1 represent $S_{o2}$-RNase in OKA1 and the $S_{o2}$-RNase in OKA3 is represented by So2Ok3.

## 5.4 DISCUSSION

### 5.4.1 The choice of housekeeping gene is essential for gene expression studies

Analytical techniques which are very sensitive, accurate and are highly reproducible are vital for the measurement of specific mRNA transcripts (Nicot *et al.*, 2005). The use of QRT-PCR for gene expression analysis provide lots of advantages which include high sensitivity and specificity, good reproducibility giving a wide dynamic range of gene quantification and its currently the most sensitive method for detecting low abundance

mRNA. However, accurate data normalisations coupled with other variables (*e.g.* PCR conditions) are essential for getting consistent gene quantification results using PCR based techniques (Bustin, 2000; Mascia *et al.*, 2010) for instance in a real-time quantitative RT-PCR. Internal control genes generally referred to us reference/housekeeping genes are usually used in normalising QRT-PCR data in order to avoid bias and these genes should normally be independent of the conditions of the experiment (Schmittgen and Zakrajsek, 2000). But several studies have shown that most housekeeping genes used for normalising QRT-PCR data can show variations with the conditions of the experiments thereby leading to erroneous results (Thellin *et al.*, 1999; Warrington *et al.*, 2000; Sturzenbaum and Kille, 2001; Radonic *et al.*, 2004). For these reasons the suitability of particular housekeeping gene should be determined empirically for each experiment before use.

In order to evaluate the variability of selected internal controls and chose the appropriate one for our expression studies, RNA transcript levels from potato were determined for six housekeeping genes. Results from the analysis (Figures 5.1 and 5.2) showed that not all the housekeeping genes could be used as internal controls in this experiment, an observation which is consistent with a different experiment using these same primer sets (Nicot *et al.*, 2005). In this current study, some of the housekeeping gene primers could be observed to be giving non-specific amplicons and others showed variability in their expression in the tested tissues hence could not be considered for use in this gene expression study.

The housekeeping gene of choice in any gene expression experiment should show very little variation compared to the gene of interest which could vary during the experimental period (Dean *et al.*, 2002). However, this was not the case with some of the housekeeping genes evaluated in this study. For instance, the *actin* primer tested was observed to show variability in its expression and hence was deemed unsuitable for use as an internal control in this study although *actin* has been one of the internal controls commonly used for expression studies (Ruan and Lai, 2007). Similar results were obtained by Nicot *et al.* (2005) in a QRT-PCR experiment using this same *actin* primer set where they found variable expression results for different treatments. Also genes coding for *β-tub* and *ef1-α* could be observed to show non-specificity by producing multiple bands. A similar observation could be made for the *L2* primer, which also showed some level of non-specificity by producing bands with smears which could also account for transcript levels when used for normalisation thereby giving erroneous results. Following the results obtained from the housekeeping gene evaluation in this study, two of the tested housekeeping genes (*aprt* and *CyP*) were found to be the only housekeeping genes meeting all criteria for gene expression studies. They were observed to show minimal variation and high specificity thereby making them a reliable housekeeping gene for normalisation in this study.

### 5.4.2 S-RNase gene expression is confined to pistil tissues

Exclusive pistil specific expression which is localized in the upper part of the style (where the arrest of incompatible pollen tubes occurs) was one of the characteristics expected (and observed) for the female determinant of the SI response (Anderson *et al.*, 1986; 1989; Cornish *et al.*, 1987). Unlike S-RNases whose expression is limited to floral tissues and shows high levels of tissue specificity (Anderson *et al.*, 1986; 1989), other RNases, termed S-like RNases, have been identified which are not related to the GSI system. These S-like RNases usually lack polymorphism and are not related to the S-locus. Nevertheless, these S-like RNases share similar domain structure and appear in phylogenetic studies to be related to S-RNases. However, they (S-like RNases) show very little tissue specificity in their expression and also their cellular and biochemical functions are not well defined (MacIntosh *et al.*, 2010; Nowak *et al.*, 2011).

In an attempt to show that, the cloned S-RNase from our potato stocks are tissue-specific and are expressed exclusively in the pistil, a semi-quantitative RT-PCR was performed on various tissues (which includes leaves, petals, anther cone, buds, stigma and style) of OKA1 using allele specific primers designed from the cloned $S_{o1}$- and $S_{o2}$-RNases (Figure 5.3). Results from the analysis showed that the cloned S-RNases showed high tissue specificity by showing expression in only the stigma and style with the highest expression in the styles. A quantitative real-time-PCR was also performed to check where the S-RNase gene accumulates the most in the pistils of our potato stock (OKA3) (Figure 5.7). Results from

the analysis showed that the S-RNase gene is more highly expressed in the style than in the stigma which is consistent with what is known of the S-RNase gene. The S-RNase gene is known to cause the inhibition of incompatible pollen tube in the S-RNase-based GSI system when the pollen tubes is approximately half-way through the style, hence one will expect the S-RNase gene to be highly expressed in that part of the pistil, thus in the style. In mature pistils, S-RNases are abundant in extremely high concentrations and about 10-50mg/ml of S-RNases could be found in the extracellular matrix of the transmitting tract in species of the Solanaceae (Jahnen *et al.*, 1989; Broothaerts *et al.*, 1990). This thesis shows that, the $S_{o1}$- and $S_{o2}$-RNase genes are genuine S-RNases and are not S-like RNases and are highly tissue specific with their expression confined to the pistil tissues. This is consistent with the phylogenetic placement of these cloned S-RNases with S-like RNases (section 4.3.3.3), where S-like RNases could be observed clustering together differently from the S-RNases.

### 5.4.3 The S-RNase gene is developmentally expressed

One of the characteristics of the S-RNase gene is that, it is developmentally expressed. Pistils of immature buds have very low levels of expression which increases as the flower develops to a point where its full expression corresponds with the pistils acquiring a SI phenotype at maturity (McCubbin and Kao, 2000). Attempts to profile the developmental expression of both $S_{o1}$-and $S_{o2}$-RNase in the pistils of OKA1 in this current study (Figures 5.5 and 5.6) was consistent with

what is known of the S-RNase gene. Analysis of both semi-quantitative and quantitative real-time PCR showed that both $S_{o1}$- and $S_{o2}$-RNase are expressed at very low levels in immature buds, increasing prior to flower opening (balloon stage) and reaching its highest expression levels in fully opened flowers (matured pistils), when SI is known to attain its full functionality.

S-RNases in mature pistils could account for about 1-10% of the total proteins found in the pistil (Roalson and McCubbin, 2003). This variable nature of expression of the S-RNase gene has been exploited by researchers enabling them to self-pollinate incompatible plants during the bud stage of pistil development. Through immature bud pollinations, plants homozygous for S-alleles have been obtained (Bernatzky *et al.*, 1988), a technique which has made it practical for overcoming self-incompatibility in members of the Solanaceae family which lead to the production of homozygous stocks (*e.g.* Robbins *et al.*, 2000). Self-compatibility could be attained in the bud self-pollination of otherwise self-incompatible plants because the very low level of S-RNase gene expressed at the bud stage during flower development is unable to cause the rejection of self-pollen, hence allowing self-pollen tubes to go through to eventually effect fertilization.

### 5.4.4 S-RNase gene expression can be variable under different genetic backgrounds

Attempts to quantify the expression of the $S_{o1}$- and $S_{o2}$-RNase gene in two different backgrounds, *i.e.* OKA1 ($S_{o1}S_{o2}$) and OKA3 ($S_{o1}S_{o2}$) using QRT-PCR (Figure 5.8) showed that, although there appear to be some slight changes in the expression of both $S_{o1}$- and $S_{o2}$-RNase in the different backgrounds, the difference was not that significant. It is not surprising that some slight differences in expression could be observed in the expression of an allele under different genetic backgrounds. Genotype-specific difference in the expression of the same S-RNase gene under different backgrounds is commonly observed in the Solanaceae although it has not been given much consideration. For instance, Qin *et al.* (2006) analysed the $S_{11}$-RNase in the styles of different genotypes of *Solanum chacoense* and observed insignificant differences in $S_{11}$-RNase expression. Also Qin *et al.* (2001) analysed the $S_{12}$-RNase in the styles of *S. chacoense* and found significant genotype-specific difference in its expression in different plant lines. Similar results were observed in transgene variation (although less surprising) in species of *Nicotiana* where the expression of the $S_{A2}$-RNase construct showed differences under different genetic backgrounds (Murfett and McClure, 1998). Based on all these findings, the S-RNase gene could show some level of variations in its expression in different plant genotypes. However these variations might not have any effect on the activity of the S-RNase gene because the resulting change might be very minimal and therefore might

not have any significant effect on the minimal threshold of S-RNases required quantitatively for incompatible pollen tube rejection.

# CHAPTER 6: S-RNase GENE SILENCING: RNAi CONSTRUCT DESIGN

## 6.1 INTRODUCTION

### 6.1.1 RNAi in plants: a brief overview

RNAs were initially thought to have only two main functions in cells, *i.e.* single stranded RNAs called messenger RNAs (mRNA) which are known to play important roles in the expression of genes by conveying information between DNA and protein. Also ribosomal RNAs (rRNA) and transfer RNAs (tRNA) are known to play important roles during the synthesis of proteins. The discovery that RNA molecules can regulate gene expression is perhaps the most important advance in molecular biology over recent decades (Novina and Sharp, 2004). The mechanisms of RNA silencing (depicted in Figure 6.1) were initially recognized as innate antiviral strategies used by organisms to protect them from RNA viruses or a mechanism to prevent transposable elements from integrating into their genome (Waterhouse *et al.*, 2001; Meister and Tuschl, 2004). Thus the introduction of a naturally occurring, viral RNA (which is usually dsRNA during replication), self-complementary, single-stranded 'hairpin' RNA (hpRNA) or dsRNA into cells results into its degradation, producing dsRNA fragments (siRNAs) of ~21 nt. The siRNAs are then introduced into a nuclease-complex referred to as RNAi silencing complex (RISC), which causes the degradation of cognate mRNAs that have complementary sequences to the single stranded siRNAs that are linked with the RISC complex (McManus and Sharp, 2002; Waterhouse and Helliwell, 2003).

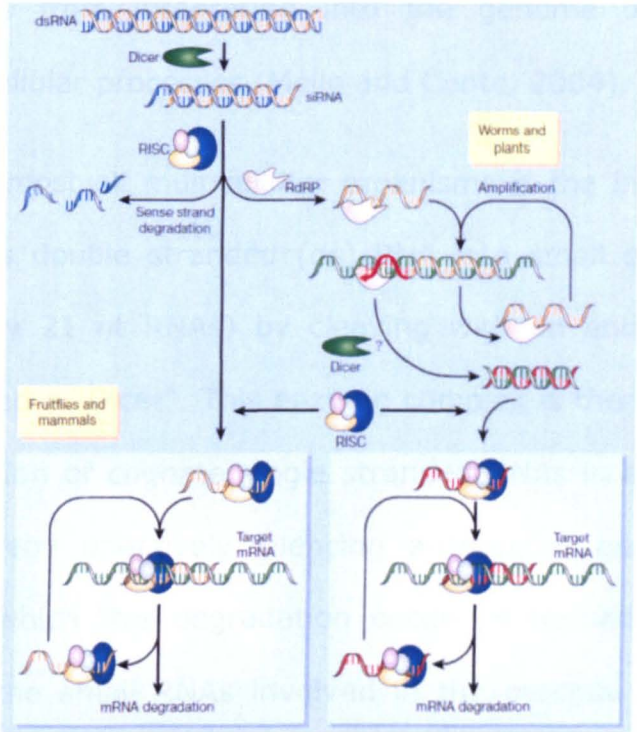



**Figure 6.1 RNA silencing mechanism**. A dsRNA (produced from a transgene, an intruding viral particle or a rogue genetic material) introduced into a cell results in the triggering of the RNAi mechanism. This double-stranded RNA is then cleaved by an enzyme complex called "Dicer" into siRNA. The siRNA is then introduced into the RISC complex where the sense strand (blue) of the siRNA is degraded whilst the antisense strand (yellow) is used for targeting genes for silencing. Depending on the organism involved, this silencing-inducing strand has one of several fates. For instance, in fruitflies and mammals, it is directly incorporated into the RISC complex for targeting complementary mRNA (green) for degradation. On the contrary, in plants and worms, the antisense strand *i.e.* silencing-inducing strand first undergoes amplification initiated by an RNA-dependent RNA polymerase (RdRP) by pairing with a complementary mRNA strand (green) to produce a new dsRNA. This dsRNA is then cleaved by "Dicer" into a new siRNA (red) which is then introduced into the RISC complex leading to the eventual degradation of the target mRNA (Reprinted from Novina and Sharp (2004) by permission from Nature Publishing Group).

The objective of the degradation or the silencing of the foreign invading nucleic-acids (*e.g.* from viruses) by most organisms is to prevent these

nucleic acids from integrating into the genome of the host or to destabilize cellular processes (Mello and Conte, 2004).

Present in almost all multicellular organisms is the inherent mechanism that converts double stranded (ds) RNA into small oligonucleotides (of approximately 21 nt RNAs) by cleaving with an endonuclease enzyme complex called a "Dicer". This enzyme complex is then used for directing the degradation of cognate single stranded RNAs in a sequence-specific manner thereby effectively silencing a targeted gene sequence. The process by which this degradation occurs is termed RNA interference (RNAi) and the small RNAs involved in the process are referred to as short interfering (si) RNAs (Watson *et al.*, 2005).

The RNAi term was coined when the phenomenon was discovered first in the nematode *Caenorhabditis elegans* (Fire *et al.*, 1998; Hannon, 2002). RNAi, a post transcriptional gene silencing event is a conserved mechanism amongst various organisms, which include plants and animals. RNAi is highly specific and has high efficacy and hence is widely used as an efficient tool for the analysis of the function of genes. The practical use of RNAi in the genetic enhancement of crop plants is projected and has been established in some crop plants; however its wider applicability still demands further feasibility studies (Kusaba, 2004). Before the discovery of RNAi in plants, RNAi-related phenomena had been previously observed in plants. One typical example is co-suppression which involves the silencing of a gene mediated by a sense transgene (Napoli *et al.*, 1990). Co-suppression was found to be either a transcriptional gene silencing (TGS) or a post-transcriptional gene

silencing (PTGS) process. In co-suppression, the expression of both the transgene and the endogenous homologous genes are suppressed. Coat protein mediated protection (CPMP), where a virus sense coat protein encoding transgene introduced into a plant confers resistance to that virus on the plant, is another example of an RNAi-related phenomenon in plants. There is genetic and biochemical evidence that, RNAi and RNAi-related phenomena have mutual mechanistic similarities. Also the biological pathway underlying dsRNA-induced gene silencing exists in most eukaryotes (Hannon, 2002; Kusaba, 2004).

RNAi in plants is achieved through a transgene producing a hairpin RNA (hpRNA) with a dsRNA region. Antisense mediated gene silencing was the conventional method of analysing gene function in plants and although it is an RNAi-related phenomenon, the use of hpRNA-induced RNAi has been found to give much greater efficiency than antisense mediated gene silencing (Chuang *et al.*, 2000; Serio *et al.*, 2001; Waterhouse and Helliwell, 2003). An hpRNA-producing vector has the target genes cloned in inverted repeats which are spaced by unrelated sequences and are under the control of a strong promoter, for instance the 35S CaMV promoter in the case of dicots and maize ubiquitin 1 promoter in the case of monocots (Kusaba, 2004). Introns are most often used as the spacers for these RNAi constructs, and were found to increase the efficiency of silencing, although the mechanism still remains elusive (Helliwell and Waterhouse, 2003).

In RNAi, different delivery systems have been used to introduce the gene silencing-inducing RNAs (dsRNA or hpRNA) into plants. These delivery

methods have both their advantages and disadvantages. The methods that have been used so far include the use of transgenes that encode hpRNAs or viral RNAs to stably transform plants. The silencing-inducing RNA can also be delivered into plants transiently through bombardment with beads which are coated with nucleic acids (microprojectile bombardment). Also plant cells can be infiltrated using transgene-carrying *Agrobacterium tumefaciens* (agroinfiltration) or by using a virus to infect plants either in isolation or in combination with a satellite-virus (Virus-induced gene silencing, VIGS) (Waterhouse and Helliwell, 2003).

Gene silencing, or the down-regulation of specific endogenous gene expression, has received some attention in potato research. This has been done mainly for functional analysis of cloned genes but also to obtain potatoes with improved traits. Most of the gene silencing in potato has been achieved using the antisense RNA technique (Wolters and Visser, 2000). Genes involved in starch biosynthesis or sucrose metabolism have been effectively silenced in potato (Kuipers *et al.*, 1997). One gene that has been comprehensively researched in terms of gene silencing is the gene that codes for the enzyme granule-bound starch synthase I (GBSSI) (Wolters and Visser, 2000).

### 6.1.2 Gateway® cloning: a brief overview

The drive for the elucidation of gene functions has necessitated the need for high-throughput directional gene cloning techniques that are rapid, flexible, efficient, simple to use and are compatible with any piece of DNA

fragment to be cloned (Fu *et al.*, 2008) thereby replacing the traditional DNA ligase-based restriction enzyme cloning systems. One of the universal cloning technologies developed to facilitate the cloning process is the Gateway® cloning technology (Hartley *et al.*, 2000; Invitrogen, Carlsbad CA, USA). Gateway cloning system takes advantage of the site-specific recombination system of bacteriophage λ (that allows the integration and excision of phage λ in and out of the chromosome of bacteria); thereby enabling the shuttling of gene sequences between plasmids with compatible recombination sites. Bacteriophage λ uses site-specific recombination to integrate into the genome of its host during lysogeny and prophage excision. The phage λ integrates at specific attachment (*att*) sites in the genome of the bacteria host. The integration is a sequential exchange process which depends on proteins encoded by both the bacteria and the phage (Fu *et al.*, 2008).

The Gateway® cloning system offers a simple method for transferring DNA sequences from one vector to another without affecting the reading frame and orientation of the inserted DNA sequence. Two recombination reactions constitute the Gateway® cloning system; the BP recombination reaction and the LR recombination reaction. Both the BP and LR reactions involve two recombination sites that are used for cloning and transfer of genes from one vector to the other; thus recombination event between the *att*B and *att*P sites in the BP reaction and *att*L and *att*R sites in the LR reaction. The BP recombination reaction involves site-specific recombination of a PCR product (or cloned gene) flanked-*att*B sites (*att*B1 and *att*B2) with *att*P sites (*att*P1 and *att*P2) contained in a

gateway-compatible vector called a donor vector. The BP reaction is mediated by a mixture of enzymes containing proteins called Int (Integrase protein) and IHF (Integration Host Factor). The product of the BP reaction is called an entry clone flanked by new recombination sites *i.e. att*L1 and *att*L2 (Invitrogen, Carlsbad CA, USA; Fu *et al.*, 2008).

The LR reaction involves the transfer of the insert in the entry vector into an expression/destination vector. During this process, the *att*L sites contained in the entry clone recombines with the *att*R recombination sites (*att*R1 and *att*R2) contained in the expression/destination vector to create an expression clone which has new *att*B recombination sites (*att*B1 and *att*B2). The LR reaction is mediated by a mixture of enzymes containing proteins called Int (Integrase protein), IHF (Integration Host Factor protein) and Xis (Excisionase protein) (Invitrogen, Carlsbad CA, USA; Fu *et al.*, 2008). Gateway cloning has become the cloning method of choice for many researchers and the development of transformation vectors compatible for the technique is on the increase for different functional assays (Karimi *et al.*, 2007). For instance, different gateway compatible destination vectors have been designed to express hairpin RNA (hpRNA) for gene silencing by RNAi in plants. These vectors carry two independent gateway cassettes separated by an intron spacer (*att*R1-*ccd*B-*att*R2-intron-*att*R2-*ccd*B-*att*R1) which allows for the recombination of entry clones (*att*L1-*target gene*-*att*L2) initially created through BP reaction to create a hpRNA expression clone/RNAi construct (*att*B1-*target gene*-*att*B2-intron-*att*B2-*target gene*-*att*B1) in a single LR reaction (Karimi *et al.*, 2007). The *ccd*B cassette present in both donor

and destination vectors acts as a negative selective marker gene which enables the selection against vectors that have not undergone recombination and transformation. An example of the RNAi construct design with gateway cloning procedure is depicted in Figure 6.2 using two gateway compatible vectors; pDONR201 vector (in BP reaction) and pHellsgate vector (in LR reaction) (Helliwell and Waterhouse, 2003).

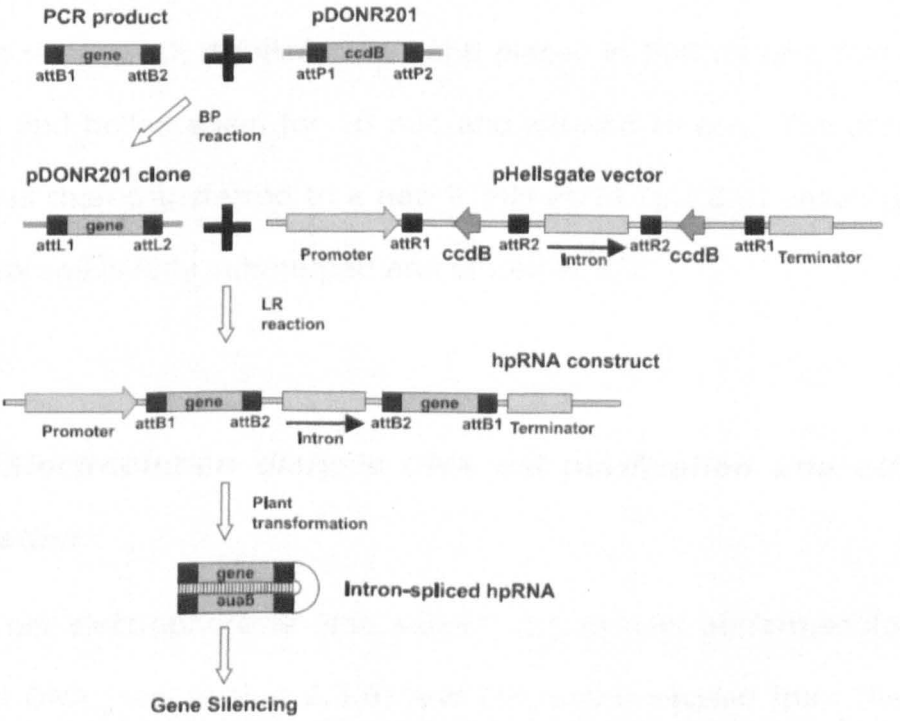



**Figure 6.2 A schematic of RNAi construct design using Gateway cloning.** An outline of RNAi construct design for gene silencing is shown. An *att*B-attached primer was used to amplify the gene of interest to create an *att*B-flanked PCR product. A recombination event between the *att*B1/*att*B2 and *att*P1/*att*P2 sites of the PCR product and a donor vector (pDONR201) respectively in a BP reaction creates an entry clone with new *att*L-sites. The created entry clone with *att*L1/*att*L2 sites then recombines with a destination vector (pHellsgate) with *att*R1/*att*R2 sites in an LR reaction to create an expression clone (RNAi construct) with a newly formed *att*B sites. The expression of the RNAi construct in plants results into the production of an intron-spliced hairpin RNA (hpRNA) which triggers silencing (Reprinted from Helliwell and Waterhouse (2003) by permission from Elsevier).

## 6.2 MATERIALS AND METHODS

### 6.2.1 Dialysis membrane preparation and gel purification

#### *6.2.1.1 Dialysis membrane preparation*

The dialysis tubing (19 mm width, Medicell International Ltd., London) was cut to required sizes (~4-6 cm) and boiled in 500 ml of solution containing 2% $NaHCO_2$ and 1 mM EDTA (pH 8.0) for 10 min. The tubing was then rinsed with distilled water and placed in 500 ml of 1 mM EDTA (pH 8.0) and boiled again for 10 min and allowed to cool. The prepared tubing was then transferred to a new 1 mM EDTA (pH 8.0) ensuring that the membrane is fully submerged and stored at 4°C.

#### *6.2.1.2 Electroelution dialysis DNA gel purification and ethanol precipitation*

Agarose gel electrophoresis (see section 2.1.4) was performed for the restricted DNA (see section 2.3.6) and the bands excised from the gels using a sterile razor blade (ensuring minimal exposure of the gel to UV light) and placed in a 1.5 ml microcentrifuge tube. Approximately three gel lanes were excised per sample ensuring that the size of the gel was trimmed to fit into the smallest size dialysis tubing.

The prepared hydrated dialysis membrane (section 6.2.1.1) was washed both inside and outside with 0.5X TBE and one of the ends of the bags clipped closed with a dialysis bag clip. 2 ml of 0.5X TBE buffer was then placed into the membrane and the excised gel pieces were then placed

into the tubing. The liquid (0.5X TBE) in the membrane was squeezed up to the top of the tubing while dangling the gel strip in an attempt to remove excess liquid and ensuring all air bubbles are taken out leaving behind about half of the TBE in the tubing. The other end of the tubing was then clipped closed using a new dialysis bag clip. The dialysis membrane was then placed in an electrophoresis tank containing 0.5X TBE and run at 100 V for ~1.5 hr ensuring that the membrane is oriented parallel to the electrodes and perpendicular to the electric field. After running the DNA fragments out of the gel strip, the electric field was reversed for ~30 sec (to allow the DNA to move off the walls of the tubing into solution) and the DNA in solution checked under UV lamp.

One of the clips on the tubing bag was then removed and buffer containing the DNA was mixed gently with a P1000 pipette and transferred into a 1.5 ml microcentrifuge tube. An equal volume of butanol was then added to concentrate the DNA and mixed briefly followed by centrifugation at maximum speed (13,000 rpm) for 3 min. The organic phase formed (top layer) was then carefully removed with the pipette tip leaving the aqueous phase (bottom layer) in the microcentrifuge tube. This step was repeated for ~4-6 times until ~50-60 µl of DNA was retained in the aqueous phase in the microcentrifuge tube. The aqueous phase containing the concentrated DNA was then pipetted into a new microcentrifuge tube. This was then ethanol precipitated by adding 0.1X volume of 3M sodium acetate (NaAc, pH 5.2) and 2.5X volume of absolute (100%) ethanol. The mixture was then kept at -20°C for 30 min or generally overnight. The samples were then centrifuged at

maximum speed for 15 min and the supernatant carefully removed. This was followed by the addition of 250 µl of 70% (v/v) ethanol and centrifugation at maximum speed for 15 min to wash the DNA. The supernatant was slowly removed and the pellet was air dried at room temperature for ~30 min. The samples were re-suspended in 15-50 µl (depending on the expected DNA concentration) of distilled water or TE buffer.

### 6.2.2 Primers for constructing the chimeric S-RNase gene and the RNAi construct

The primers used for constructing the chimeric gene and RNAi construct in this study are listed in Table 6.1. The vector specific primers used in this study were all kindly provided by the donors of the various vectors as acknowledged later in this chapter. All the other primers (gene primers) were specifically designed either using the "Primer3" software (section 2.4) or manually by hand. The primer pairs were designed to have their melting temperature close together (at ~60°C), and should not to be complementary at both ends, have their GC content within 40-60% and also be devoid of secondary structures.

**Table 6.1** Primers for creating chimeric S-RNase and confirming RNAi constructs

| Primer Name | Direction | Sequence (5'-3') | Tm °C |
|---|---|---|---|
| **So2NdeI-F1**[1] | Forward | GGGAATTCCATATGACAATGCTTAAT AACTGCGAGTC | 61.1 |
| **So2-R1**[1] | Reverse | ATCTCGGTCGAAACATATTCC | 61.7 |
| **So1So2-FP**[2] | Forward | TGAAGTACTGCAAGCCGAAA | 63.5 |
| **So1So2-RP**[2] | Reverse | GCCTTAATGGCATTCTCGAC | 63.4 |
| **attB1-So1**[3] | Forward | GGGGACAAGTTTGTACAAAAAAGCA GGCTTAGGATAAGGAGGGATCACA | 57.3 |
| **attB2-So2**[3] | Reverse | GGGGACCACTTTGTACAAGAAAGCT GGGTAATCTCGGTCGAAACATAT | 54.2 |
| **pDONR207-FP**[4] | Forward | TCGCGTTAACGCTAGCATGGATCTC | 72.4 |
| **pDONR207-RP**[4] | Reverse | GTAACATCAGAG ATT TTGAGACAC | 58.3 |
| **Ph8xhof**[5] | Forward | GAAAGGCTATCATTCAAGATCTCTCT GCCG | 72.9 |
| **Ph8xhor**[5] | Reverse | GCTAGTATATCATCTTACATGTTCGA TCAAATTCATTA | 68.5 |
| **Ph8xbaf**[5] | Forward | CATGATAGATCATGTCATTGTGTTAT CATTGATC | 69.5 |
| **Ph8xbar**[5] | Reverse | CATGCGATCATAGGCGTCTCGCATAT CTC | 76.0 |
| **HELLINT2**[5] | Forward | GAATAAACAAGGTAACATGATAGATC | 57.8 |
| **OCSrev**[5] | Reverse | CATAATTCTCGGGGCAGCAAGTCGG TTAC | 75.3 |

1: Primers used for adding *Nde*I site to $S_{o2}$-RNase. The *Nde*I site in the forward primer is underlined

2: Primers used checking the chimeric S-RNase gene and also confirming the entry clones and RNAi construct

3: Primers used for generating the *att*B-flanked PCR products

4: Primers used for checking the entry clones

5: Primers used for confirming the RNAi construct

### 6.2.3 Propagation/maintenance of gateway compatible vectors

Donor vector (pDONR207, see Appendix A6.2) kindly donated by Dr J Stephens (James Hutton Institute, JHI) was propagated in Library Efficiency® DB3.1™ competent cells (Invitrogen, Carlsbad CA, USA).

DB3.1™ competent cells are resistant to the lethal gene *ccd*B effects and can support the propagation of gateway compatible vectors. Following the transformation (section 2.3.2), the cells were plated on LB agar plates containing 25 µg/ml of gentamycin and 15 µg/ml of chloramphenicol to maintain the integrity of the vector. Destination vector (pHellsgate8, see Appendix A6.3) kindly donated by Dr M. Taylor/Dr L. Ducreux (JHI) was also propagated in Library Efficiency® DB3.1™ competent cells. The transformed colonies were grown on LB plates containing 100 µg/ml of spectinomycin. Both the pDONR207 and pHellsgate8 colonies were screened using vector specific primers (Table 6.1) and confirmed through restriction enzyme digestion and plasmid extracted from the confirmed colonies and glycerol stocks made and kept for storage and future use (section 2.3.5).

### 6.2.4 *att*B-PCR products purifications

Following PCR amplification (section 2.1.3) at an annealing temperature of 50°C in a 50 µl reaction mix with the *att*B site-flanked primers (Table 6.1) and the confirmation that, the expected PCR product (chimeric S-RNase gene) was amplified, the *att*B-PCR product was then subsequently purified to remove all *att*B primers and primer dimers. The procedure for purifying *att*B-PCR products provided by Invitrogen (Invitrogen, Carlsbad CA, USA) was as follows; 150 µl of TE buffer (pH 8.0) was added to the 50 µl *att*B-PCR products to dilute it and then precipitated using 100 µl of 30% PEG 8000/30mM $MgCl_2$ solution and vortex briefly followed by centrifugation at 13,000 rpm for 15 min to pellet the DNA. The

supernatant was removed and the pelleted DNA resuspended in 50 µl of TE buffer (pH 8.0) and the quality and quantity of the purified *att*B-PCR products checked on agarose gel.

### 6.2.5 BP recombination reaction: creating an entry clone

Entry clone, pENTR207-$S_{o1}$/$S_{o2}$ was created by incubating *att*B-flanked PCR products with *att*P containing donor vector, pDONR207 in a BP recombination reaction using BP Clonase™ enzyme mix (Invitrogen, Carlsbad CA, USA). The reaction component following the enzyme provider's recommendation comprised; 56.6 ng (100 fmol) of *att*B-PCR product, 300 ng of pDONR207, 4 µl of 5X BP clonase buffer, 4 µl of BP Clonase™ enzyme mix and the final reaction volume adjusted to 20 µl with 1.4 µl of TE buffer (pH 8.0). The reaction set-up was then incubated at 25°C overnight. 2 µl of 2 µg/µl Proteinase K solution was then added to terminate the reaction and the mixture incubated again at 37°C for 10 min. The reaction was then kept at -20°C (not more than 1 week) until ready for transformation into competent cells. Following the BP recombination reaction, an aliquot of the reaction mix (2-8 µl) was then used to transform Library Efficiency® DH5α™ Chemically competent *E. coli* cells (Invitrogen, Carlsbad CA, USA). The cells were then plated on LB agar plates containing 25 µg/ml of gentamycin and incubated overnight at 37°C (see section 2.3.2 for transformation procedure).

### 6.2.6 LR recombination reaction: creating an expression clone/RNAi construct

The expression clone/RNAi construct (pHellsgate8-$S_{o1}/S_{o2}$ or pHG8-$S_{o1}/S_{o2}$) was created in an LR recombination reaction where the *att*L-containing entry clone (pENTR207-$S_{o1}/S_{o2}$) recombines with *att*R-containing destination vector (pHellsgate8). The recombination reaction comprised; 50 ng of pENTR207-$S_{o1}/S_{o2}$, 100 ng of pHellsgate8 vector, 2 µl of LR Clonase™ II enzyme mix (Invitrogen, Carlsbad CA, USA) and the final reaction volume adjusted to 10 µl with 6.7 µl of TE buffer (pH 8.0). The reaction set-up was then incubated at 25°C overnight. 1 µl of 2 µg/µl Proteinase K solution was then added to terminate the reaction and the mixture incubated again at 37°C for 10 min. An aliquot of the LR reaction mix (2-8 µl) was then used to transform Library Efficiency® DH5α™ Chemically competent *E. coli* cells (Invitrogen, Carlsbad CA, USA). The cells were then plated on LB agar plates containing 100 µg/ml of spectinomycin and the plates incubated overnight at 37°C (see section 2.3.2 for transformation procedure).

### 6.2.7 Transformation of RNAi constructs into *Agrobacterium tumefaciens*

The RNAi construct carrying the gene of interest was electroporated into *Agrobacterium tumefaciens* strain LBA4404. A vial of 100 µl LBA4404 competent cells was thawed on ice for each transformation reaction. 300 ng of the RNAi constructs were added to the competent cells and mixed

gently with a pipette tip and incubated on ice for 2 min. The cells were then transferred to a pre-chilled electroporation cuvette (pre-sterilized, 1 mm gap, Molecular BioProducts, Inc., San Diego CA, USA) and electroporated for ~1 sec using a Bio-Rad GenePulser with the settings 2.5 kv, 25 μFD, 400 Ω. The cells were recovered by the addition of 250-500 μl of LB medium to the cuvette and then mixed well and the cells transferred to a 1.5 ml microcentrifuge tube. This was followed by incubating the cells on a 28°C shaking incubator for ~3 hrs. The cells were then plated on LB agar plates containing the appropriate antibiotics (100 μg/ml of spectinomycin and 50 μg/ml of rifampicin) and incubated at 28°C for up to 48-72 hrs. The resulting colonies growing on the LB plates were screened using colony PCR with both pHellsgate8 vector specific primers and $S_{o1}/S_{o2}$ gene specific primers (Table 6.1) in different combinations (see section 6.3.3) and the clones with the putative RNAi constructs selected and glycerol stocks made from them (section 2.3.5) ready for plant transformation.

## 6.3 RESULTS

### 6.3.1 Construction of a chimeric S-RNase gene: $S_{o1}/S_{o2}$-RNase

In order to increase the chances of silencing in a heterozygote ($S_{o1}S_{o2}$), a hybrid of the $S_{o1}$ and $S_{o2}$ gene found in OKA1 and OKA3 was created ($S_{o1}/S_{o2}$). Both alleles were initially cloned into the pCR2.1 vector (see section 2.3); hence an attempt was made to join them together to create a hybrid gene. A virtual restriction enzyme digest initially performed on

the $S_{o1}$ and $S_{o2}$ sequences enabled the identification of a common restriction enzyme that cuts only once in the gene of interest ($S_{o1}$ and $S_{o2}$) but not in the vector backbone. The virtual digest performed on the vector (pCR2.1) helped to identify an enzyme that cut once in the pCR2.1 vector and not the insert. The *Nde*I and *Kpn*I pair of restriction enzymes was selected as the preferred enzymes of choice following the virtual restriction enzyme digestion for the generation of the chimeric S-RNase gene. These pair of enzymes (*Nde*I and *Kpn*I) was selected because the insert was observed through sequencing results to be in the opposite orientation (3′-5′) in the vector. *Nde*I was selected for cutting in the inserts (although $S_{o2}$ lacks an *Nde*I site, it was attached through PCR). *Kpn*I is located in the polylinker of the vector and upstream of the gene (see Appendix A6.1). The selected pair of restriction enzymes was used in a double restriction enzyme digest reaction. A schematic of the $S_{o1}$-RNase and $S_{o2}$-RNase double digests and the regions of interest excised for constructing the chimeric S-RNase gene are shown in Figure 6.3.

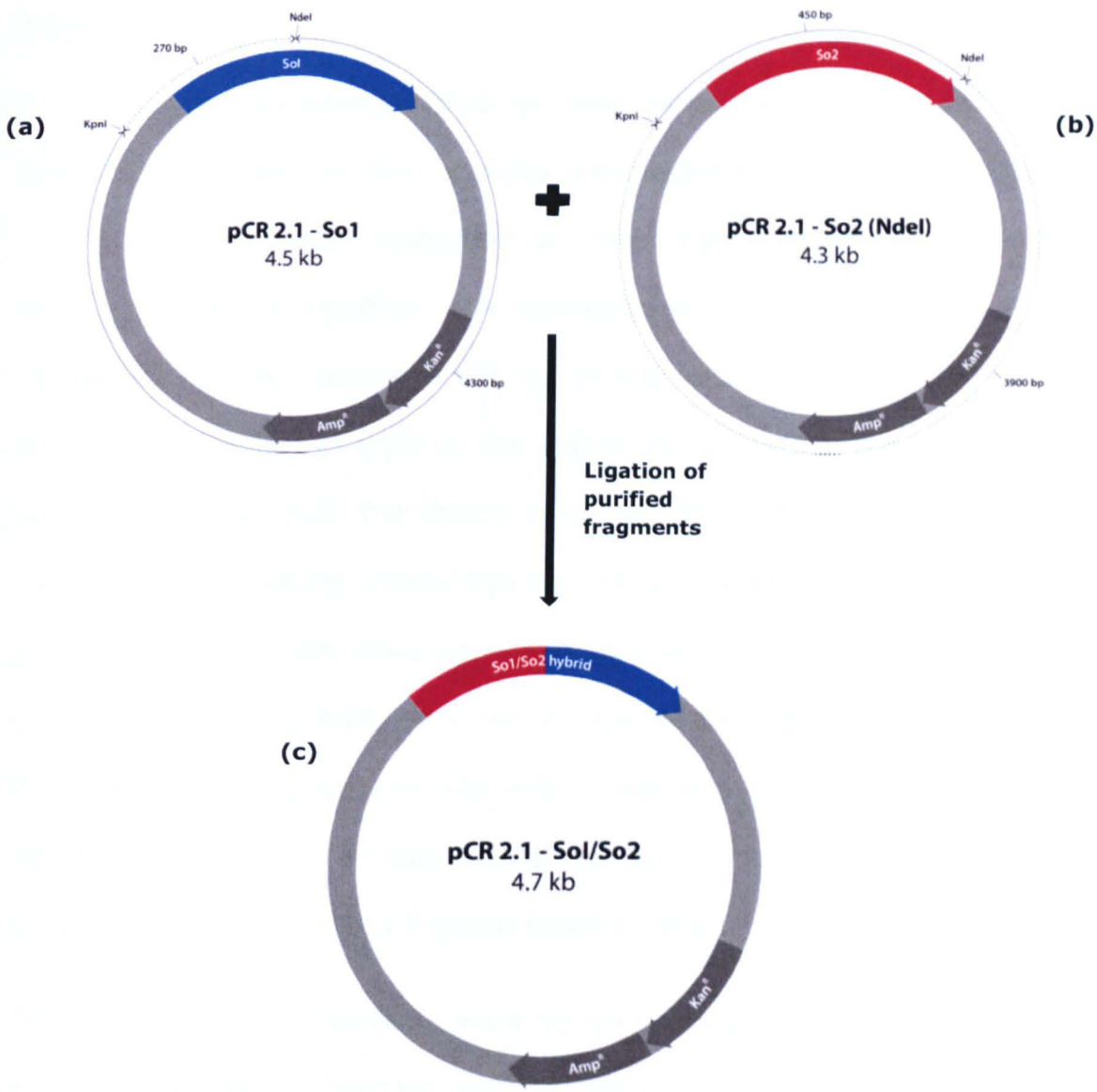



**Figure 6.3 A schematic of the chimeric $S_{o1}/S_{o2}$-RNase gene construction.** (a) A double restriction enzyme digest was performed on pCR2.1-So1 clones using *Nde*I restriction enzyme (located on the $S_{o1}$-RNase gene) and *Kpn*I restriction enzyme (located in the polylinker of the pCR2.1 vector). The region of interest (~4.3kb) was excised and purified. (b) A double restriction enzyme digest was performed on pCR2.1-So2(NdeI) clone using *Nde*I restriction enzyme (attached on $S_{o2}$-RNase gene through PCR) and *Kpn*I restriction enzyme (located in the polylinker of the pCR2.1 vector). The region of interest (~450bp) was excised and purified. (c) The purified $S_{o1}$ and $S_{o2}$ fragments were ligated to form the chimeric $S_{o1}/S_{o2}$-RNase clone (pCR2.1-So1/So2). Figures are not drawn to scale.

#### *6.3.1.1 $S_{o1}$-vector preparation for constructing the chimeric S-RNase*

With the prior knowledge (through sequencing results) that the $S_{o1}$-RNase (insert) was in the opposite orientation (3'-5') in the pCR2.1 vector, *Nde*I and *Kpn*I restriction enzymes were used for the double restriction enzyme digestion. The restriction enzyme *Nde*I was found to cut once in the $S_{o1}$ sequence but not in the vector backbone and *Kpn*I was also found to cut once in the vector backbone but not in the $S_{o1}$ sequence as expected. The double restriction enzyme digestion reaction was set up following manufacturers recommendations (see section 2.3.6). Since *Nde*I and *Kpn*I have incompatible buffers, in order to do a double digest using both enzymes a high fidelity form of *Kpn*I called *Kpn*I-HF which is compatible with *Nde*I buffer was used (*Kpn*I-HF will be referred to henceforth in this chapter as *Kpn*I). All restriction enzymes used were supplied by New England Biolabs (NEB).

The restricted DNA fragments were electrophoresed (see section 2.1.4) and the fragment of interest corresponding to part of the insert ($S_{o1}$ gene) and the vector backbone (see Figure 6.3) was excised and purified using the electroelution dialysis DNA purification method (see section 6.2.1). Since the restriction enzyme digestion was performed with two different restriction enzymes with non-complementary sticky ends there was no need for dephosphorylating the purified fragments. The digest gave the expected results (Figure 6.4a) with the region of interest having an approximate size of ~4.3kb. In order check the quality and quantity of the recovered DNA, an aliquot of the purified double digestion product

was run on an agarose gel (Figure 6.4b). An appreciable high quality DNA was found to be recovered from the gel with an estimated concentration of ~45 ng/µl. The purified product will henceforth be referred to as $S_{o1}$-vector in this chapter.

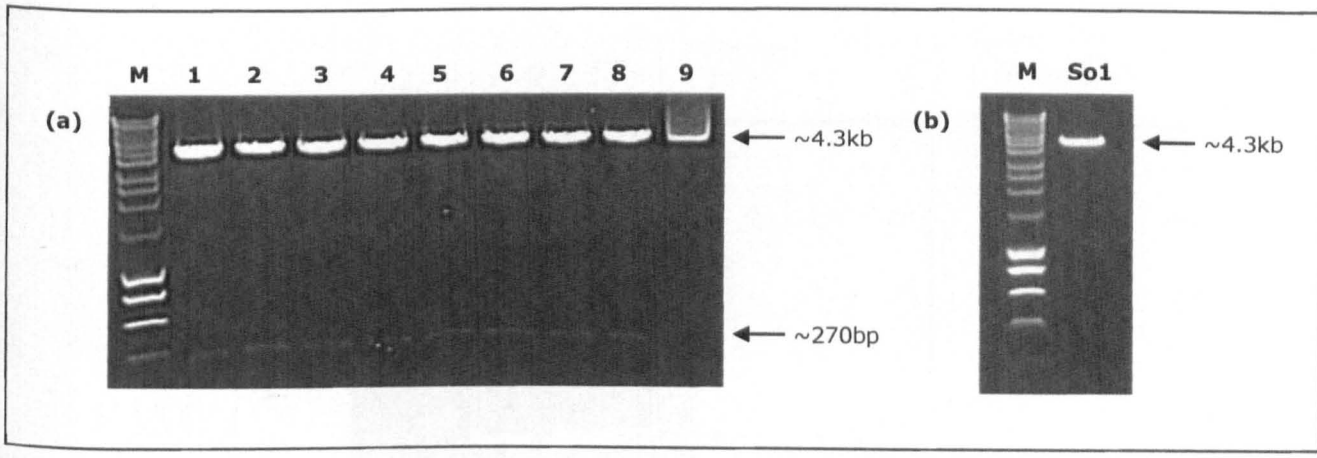



**Figure 6.4 Double restriction enzyme digest and purification of $S_{o1}$-RNase.** (a) Double restriction enzyme digests of pCR2.1-So1 using *Nde*I and *Kpn*I. (b) Purified $S_{o1}$-vector. The fragment of interest (~4.3kb) was excised from the gel and purified to create the $S_{o1}$-vector. The concentration of the $S_{o1}$-vector was estimated from the gel to be ~ 45 ng/µl. Lanes M represent Hyperladder I, lanes 1-8 represent digested pCR2.1-So1 clones, lane 9 represent empty pCR2.1 vector, lane So1 represent the purified DNA ($S_{o1}$-vector).

### *6.3.1.2 $S_{o2}$-insert preparation for constructing the chimeric S-RNase*

The initial virtual digest performed on $S_{o2}$-RNase showed that it lacks the selected *Nde*I site and hence a new site was introduced through PCR in order to facilitate the creation of the chimeric S-RNase gene. An $S_{o2}$-allele specific primer pair was designed with *Nde*I restriction site attached to

the 5′ end of the forward primer (Table 6.1). The amplification (section 2.1.3) of the $S_{o2}$-RNase from the pCR2.1-So2 clone with the allele specific primer was found to be successful yielding the expected amplicon size of ~390bp (Figure 6.5) following agarose gel electrophoresis with an aliquot of the PCR product.

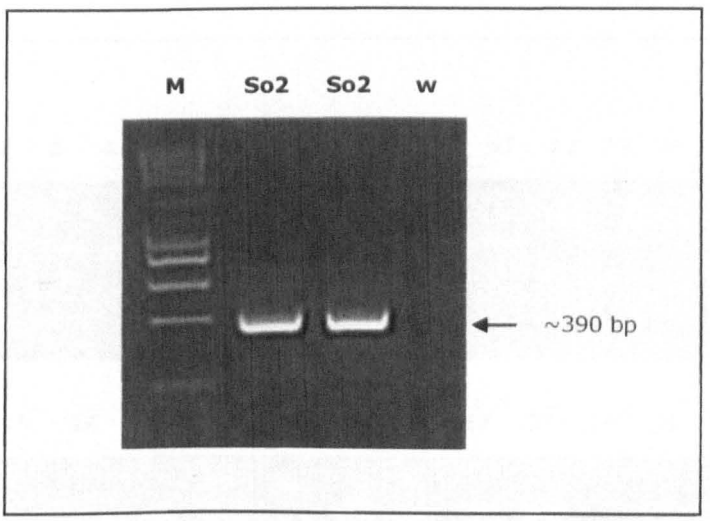



**Figure 6.5 Amplification of $S_{o2}$-RNase using So2NdeI-F1/So2-R1 primer pairs.** An $S_{o2}$-allele specific primer pair was used to add an *Nde*I restriction site to $S_{o2}$-RNase cloned in pCR2.1 (pCR2.1-So2). Lanes M represent Hyperladder II, lane So2 represents the amplified $S_{o2}$-RNase with the allele specific restriction primer, w represent water control. The concentration of the PCR product was estimated to be ~ 30 ng/µl.

The remaining PCR products were cloned back into pCR2.1 vector (see section 2.3) and the resulting colonies were initially screened through colony PCR (section 2.3.3) with M13 universal primers and the selected clones containing the expected inserts later confirmed with $S_{o2}$-allele specific primers (Table 6.1). Results from the colony screening showed

that, the cloning was successful with most of the colonies amplified with the M13 primers giving the expected amplicon size of ~600bp (Figure 6.6) and the $S_{o2}$-allele specific primers yielding the expected amplicon size of ~400bp for most of the colonies (data not shown). Plasmid DNA was then extracted (section 2.3.4) from the selected pCR2.1-So2 clones for sequencing of the inserts and further analysis.

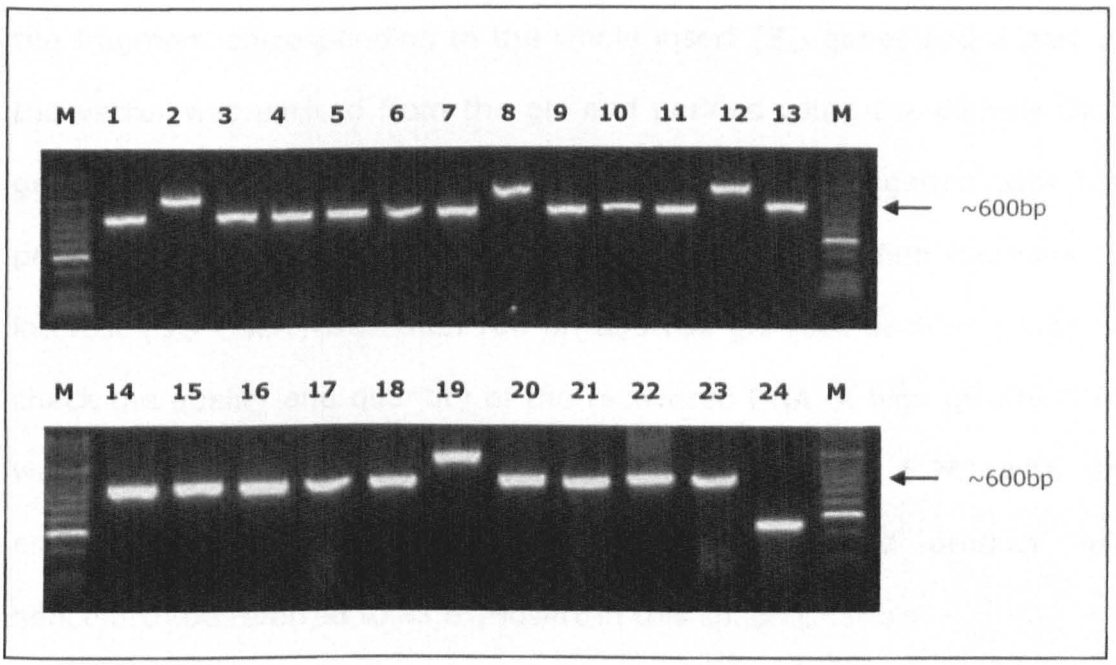



**Figure 6.6 Colony PCR using M13 universal primers for pCR2.1-So2(NdeI) clones.** Lanes M represent Hyperladder II, lanes 1-23 represent possible transformed colonies, lane 24 represents empty vector. Colonies which are ~600 bp (*e.g.* colonies 1 & 3) seems to have the expected insert size and were therefore selected for further screening.

To confirm that, the clones have acquired the *Nde*I restriction site, a double restriction enzyme digest (see section 2.3.6) was performed with *Nde*I and *Kpn*I on the extracted plasmids of the selected pCR2.1-So2

clones. The double digest results with *Nde*I and *Kpn*I gave the expected fragments and showed that, some of the $S_{o2}$ clones have acquired the *Nde*I restriction site. This was further confirmed by sequencing of the inserts in the selected clones. Figure 6.7a shows an example of one such clone restricted with *Nde*I and *Kpn*I. A selected clone with the *Nde*I site attached (pCR2.1-So2(NdeI)) was then used to set-up multiple digests for gel extraction (Figure 6.7b). The fragment of interest (~450bp) thus the fragment corresponding to the whole insert ($S_{o2}$ gene) and a part of the vector was excised from the gel and purified using the dialysis DNA gel purification method (see section 6.2.1) ready for ligation with the previously purified $S_{o1}$-vector (section 6.3.1.1). The purified fragment of interest ($S_{o2}$-insert) was then run on agarose gel (see section 2.1.4) to check the quality and quantity of the recovered DNA. A high quality DNA was observed to be recovered from the gel (Figure 6.7c) with an estimated concentration of ~15 ng/µl. The purified product will henceforth be referred to as $S_{o2}$-insert in this chapter.

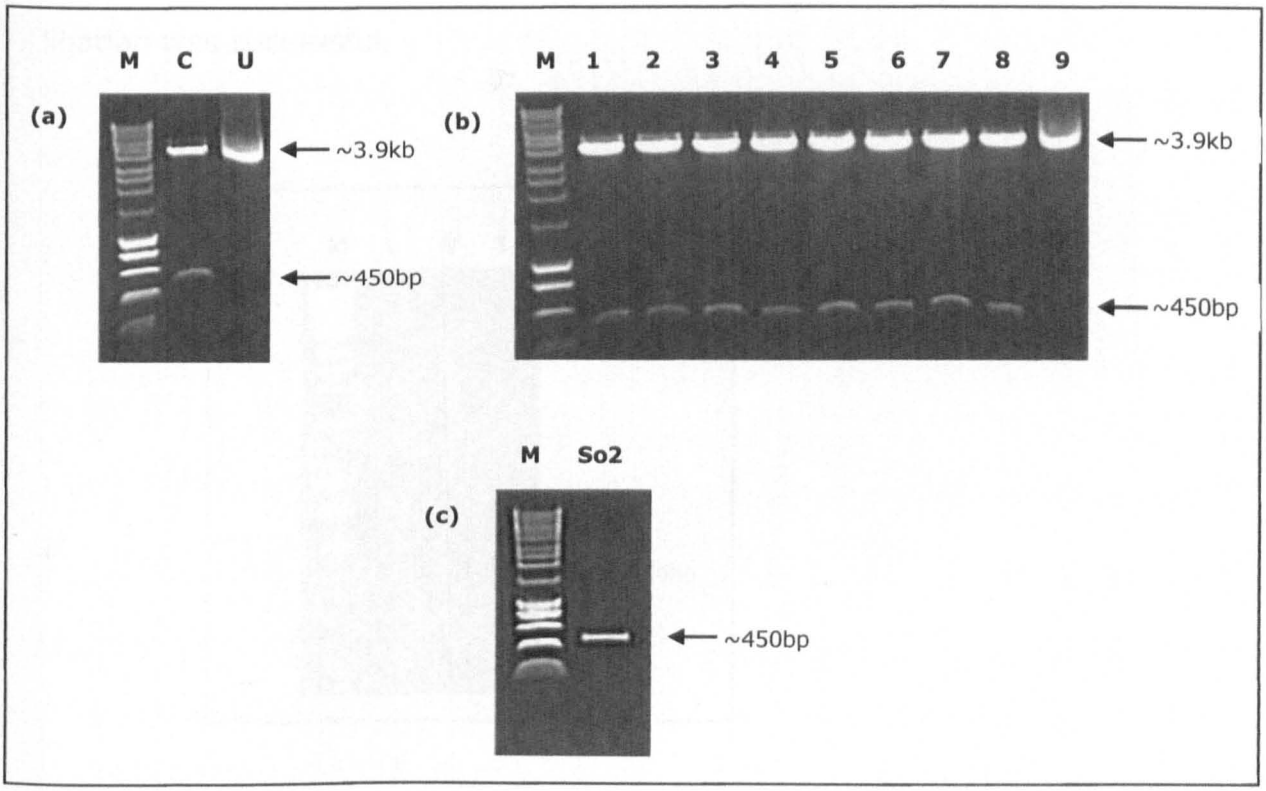



**Figure 6.7 Double restriction enzyme digests and gel purification of $S_{o2}$-RNase.** (a) Confirming that $S_{o2}$-RNase has acquired the *Nde*I site. Double digest was performed on one of the selected clones in Figure 6.6 using *Nde*I and *Kpn*I. (b) Duplicate digest reactions using the confirmed clone in lane C for dialysis DNA gel purification. (c) Purified $S_{o2}$-insert. The concentration of the $S_{o2}$-insert was estimated from the gel to be ~ 15 ng/µl. Lanes M represent Hyperladder I, lanes C and 1-8 represent digested $S_{o1}$-RNase clone in pCR2.1 vector, lane U and 9 represent empty pCR2.1 vector, lane So2 represent the purified DNA ($S_{o2}$-insert).

### *6.3.1.3 Ligation of $S_{o1}$-vector and $S_{o2}$-insert*

In order to create the chimeric $S_{o1}/S_{o2}$-RNase gene, 100 ng of the purified $S_{o1}$-vector and 15 ng of the purified $S_{o2}$-insert fragments were ligated (section 2.3.1) and an aliquot of the ligation reaction checked on agarose gel along with the $S_{o1}$-vector and $S_{o2}$-insert (Figure 6.8). The results

reveal the expected band shift in the ligation reaction indicating that, the ligation was successful.

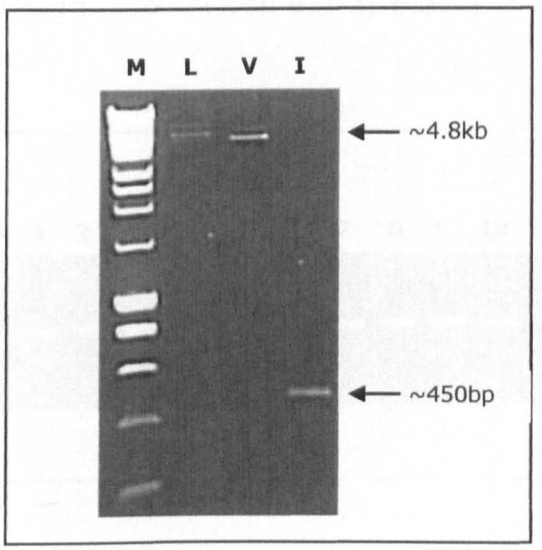



**Figure 6.8 Confirming the ligation reaction.** A 2 µl aliquot of the ligation reaction was electrophoresed alongside $S_{o1}$-vector and $S_{o2}$-insert on agarose gel to confirm that the ligation reaction was successful. The ligation reaction produced the expected band shift with a size of ~4.8kb compared with the $S_{o1}$-vector (~4.3kb). Lane M represents Hyperladder I, lane L represents the ligation reaction, lane V represents $S_{o1}$-vector, lane I represents $S_{o2}$-insert.

The ligation reaction (2 µl) was then transformed into TOP10F chemically competent cells (section 2.3.2). The transformation was found to be successful as large number of colonies grew on the plate. The resulting colonies were screened by colony PCR using M13 universal primers (section 2.3.3). Results from the colony screening showed that, almost all the clones have the expected amplicon size (~1kb) and could be the putative chimeric $S_{o1}/S_{o2}$-RNase gene. An example of the PCR results is shown in Figure 6.9. Plasmid DNA was extracted (section 2.3.4) from

these putative clones and their inserts confirmed through sequencing. The sequencing results showed all the selected clones harbour the chimeric $S_{o1}/S_{o2}$-RNase and one such clone was used for the RNAi construct design (see section 6.3.2; 6.3.3).

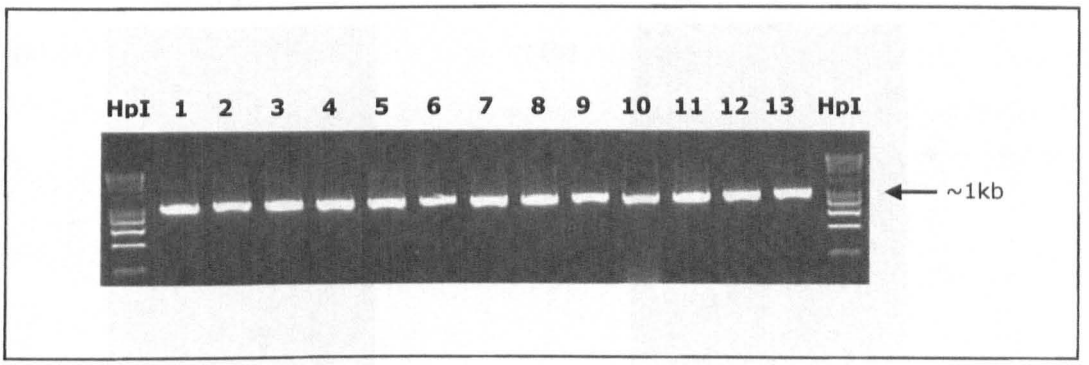



**Figure 6.9 Colony PCR of the putative chimeric $S_{o1}/S_{o2}$-RNase.** M13 universal primers were used to amplify the putative chimeric gene. Plasmid DNA was then extracted from the clones and sent for sequencing the inserts. Lane HpI represents Hyperladder I, lanes 1-13 represent putative chimeric colonies.

### 6.3.2 BP recombination reaction: creating an entry clone

In order to clone the chimeric $S_{o1}/S_{o2}$-RNase gene (pCR2.1-So1/So2) (see section 6.3.1) into a gateway compatible donor vector to create an entry clone, primers spanning the whole gene were designed with *att*B1 and *att*B2 sites attached to the 5'end of the forward and reverse primers respectively (Table 6.1). PCR performed (section 2.1.3) with these primers was found to be successful yielding the expected amplicon size of ~900bp (Figure 6.10a) and showing the right quantity and quality in agarose gel electrophoresis with an aliquot of the PCR product. The remaining *att*B-PCR product was purified (section 6.2.4) to remove all

*att*B primers and primer dimers. An aliquot of the resulting purified *att*B-PCR product was then checked again on agarose gel (Figure 6.10b) and its concentration estimated (~6 ng/µl).

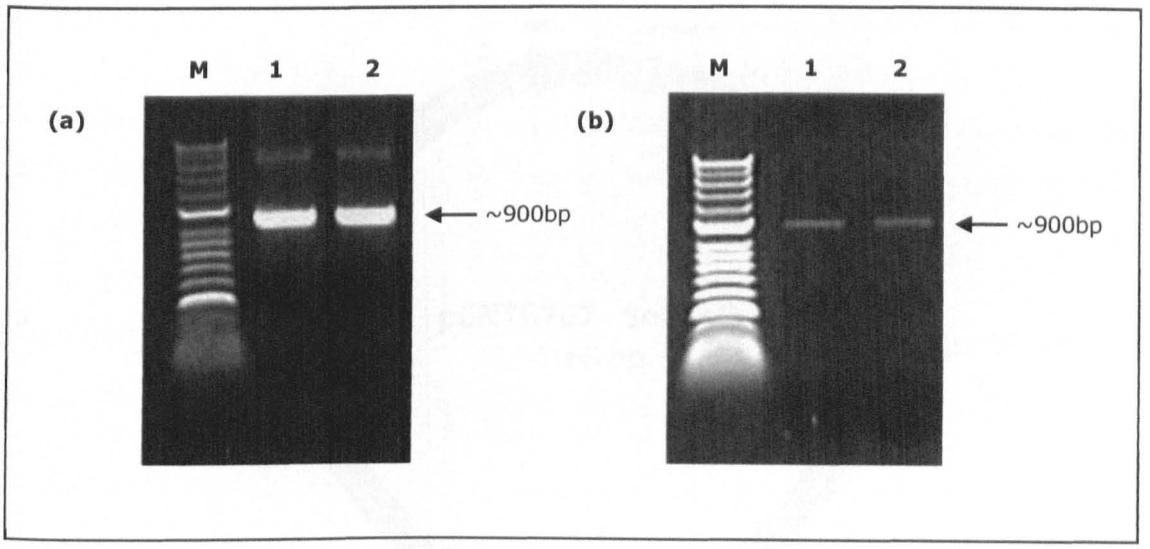



**Figure 6.10 *att*B-PCR.** PCR was performed using primers designed with *att*B site attached to amplify the whole chimeric gene ($S_{o1}/S_{o2}$-RNase). PCR purification was performed to remove *att*B-primers and primer dimers. (a) Non-purified *att*B-PCR profile. (b) Purified *att*B-PCR profile. The concentration of the purified *att*B-PCR product was estimated to be ~6 ng/µl. Lane M represents Hyperladder II, lanes 1 & 2 in both figures represent the *att*B-PCR products.

The purified *att*B-PCR product was then incubated with *att*P-containing donor vector, pDONR207 in a BP recombination reaction and the reaction products were transformed into DH5α competent cells. The transformation was found to be successful as large numbers of colonies grew on the plates. The colonies were then screened using both the pDONR207 vector specific primers and $S_{o1}/S_{o2}$ gene specific primers (Table 6.1) in different combinations and putative entry clones

(pENTR207-$S_{o1}$/$S_{o2}$) selected. A schematic of the entry clone, pENTR207-$S_{o1}$/$S_{o2}$ is shown in Figure 6.11.

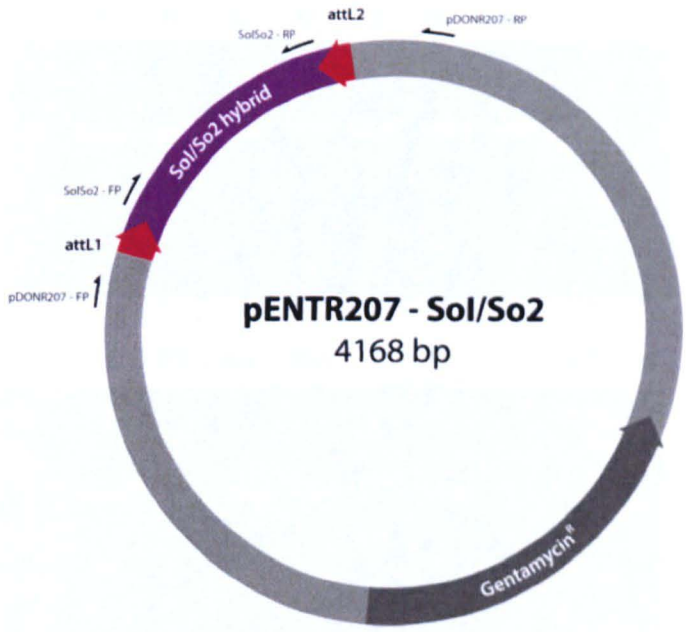



**Figure 6.11 A schematic of the entry clone, pENTR207-$S_{o1}$/$S_{o2}$.** The vector primers (pDONR207-FP and pDONR207-RP) used for initially screening the colonies growing on the plate and the allele specific primers (So1So2-FP and So1So2-RP) used for confirming the selected putative entry clones are shown. The vector primers were used for sequencing the insert to confirm their identity.

Results from the colony PCR showed that, almost all the colonies screened seem to yield the expected amplicon sizes for the various primer combinations. Examples of the colony PCR profiles are shown in Figure 6.12 (a, b & c).

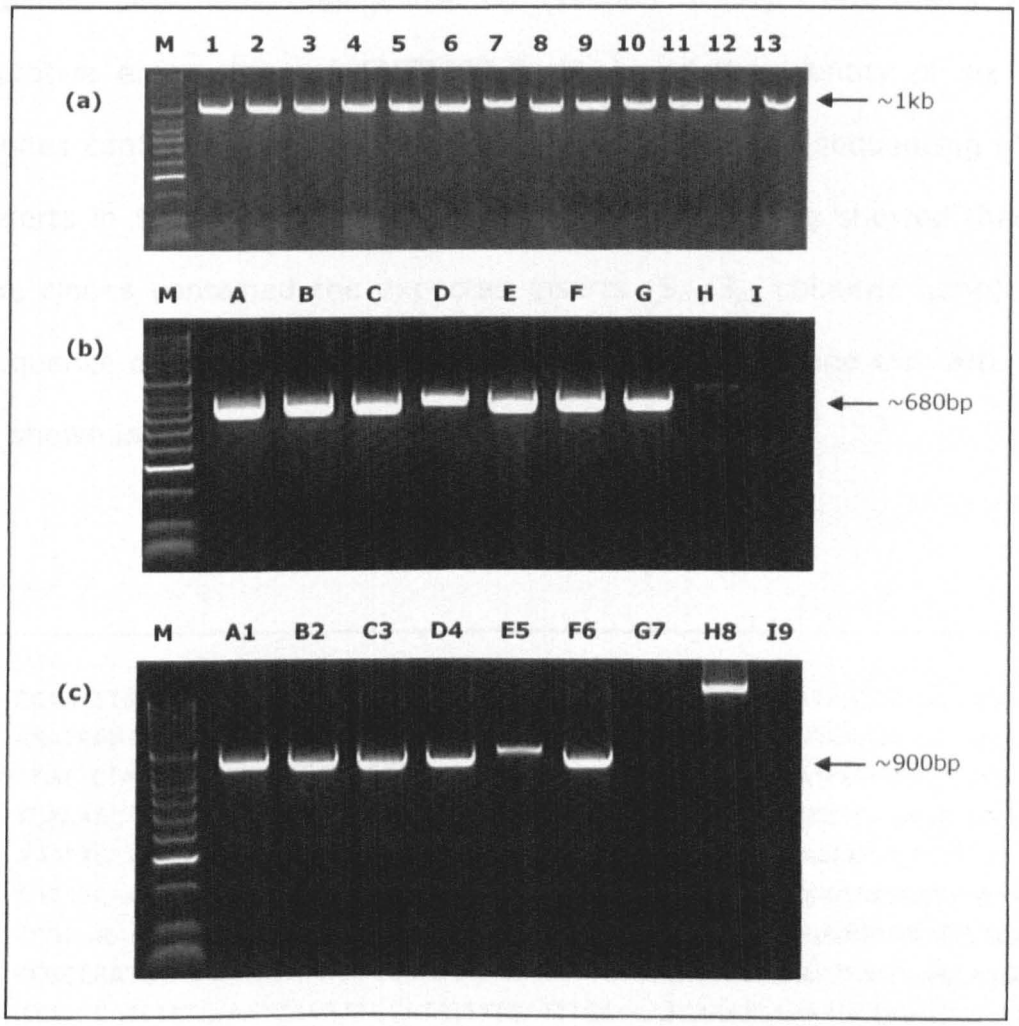



**Figure 6.12 Colony PCR of the entry clone.** The putative entry clones (pENTR207-$S_{o1}$/$S_{o2}$) were screened using both vector and gene specific primers and their combinations. (a) Initial screening of the colonies using vector specific primers (pDONR207-FP and pDONR207-RP). (b) Confirming some selected colonies in (a) using gene specific primers (So1So2-FP and So1So2-RP). (c) Further confirmation of the selected colonies using a combination of gene specific and vector specific primers (So1So2-FP and pDONR207-RP). All the PCR products gave the expected amplicon size of ~1027bp (a), ~680bp (b) and ~900bp (c). Plasmid DNA was then extracted from the clones and sent for sequencing. Lane M represents Hyperladder II, lanes 1-10, A-F & A1-F7 represent the putative entry clones, lanes G & G7 represent the chimeric $S_{o1}$/$S_{o2}$-RNase in pCR2.1, lanes H & H8 represents pDONR207 empty vector and I & I9 represents water control.

Plasmid DNA was then extracted (section 2.3.4) from the selected putative entry clones (pENTR207-$S_{o1}$/$S_{o2}$) and the identity of six such clones confirmed via restriction enzyme digestion and sequencing of the inserts in the clones. The results from the sequencing showed that, all the clones contained the expected inserts ($S_{o1}$/$S_{o2}$ chimeric gene). The sequence of the chimeric S-RNase gene in the entry clone with *att*L sites is shown in Figure 6.13.

**5'-** CCAACTTGTACAAAAAAGCAGGCTGGATAAGGAGGGATCACAGCTGTTGAAGTACTGCAAGCCGAAACTTAA
ATATAAATATTTCACTGATGAGATGCTCAATGATCTTGACAAACACTGGATTCAATTGAAGATTGATCAAGCTT
CTGCTCTAAAGGACCAACCAGCATGGAGATATCAATATCTAAAGCATGGATCCTGCTGTCAGAAAATCTACAA
TCAAAACACGTATTTTAGTTTAGCCTTGCGCTTAAAAGATAGGTTTGATCTTCTGAGAACTCTCCAAATACATA
GAATTGTTCCTGGATCAAGTTATACATTTGAAGAAATCTTTGACACCGTCAAAACAGTTACTCAGGTGGATCCT
GATATCAAGTGTACTGAAGGAGCACCGGAACTATATGAGATAGGCATATGACAATGCTTAATAACTGCGAGT
CCGAAGACAAGTATGCAGATATCTCGGATGCCAAGAAGCGCAAACAACTGGAGTATCACTGGCCTGATTTGA
CCGCCAATGTAGGTGATATTAAAAAACAACAAGGTTTCTGGGGATATGAATTTAATAAGCATGGAACGTGTA
GTATCGATCTCTATAATCAAGATGCATATTTTGATTTGGCCATCAAATTAAAAAACAAGTTTGATCTTTTGACG
GCTCTCCGAAATCATGGTATAATTCCCGGAGATGTTAGAACGGTAAAAAATGTCGAGAATGCCATTAAGGCTG
TGACTACAGAAGTTCCTAACCTCAACTGCATTGGTGATTCTGGCCAAACGATGGAATTATTGGAGATAGGAAT
ATGTTTCGACCGAGATACCCAGCTTTCTTGTACAAAGTTGG**-3'**

**Figure 6.13 The sequence of the chimeric S-RNase gene in pENTR207-So1/So2**. Parts of the *att*L1 and *att*L2 recombination sites are indicated in black and purple respectively, the green sequence represents the $S_{o1}$-RNase and the blue sequence represents the $S_{o2}$-RNase. The *Nde*I restriction enzyme used for linking the two S-RNase sequences together to form the hybrid gene is indicated in red.

### 6.3.3 LR recombination reaction: creating an expression clone/RNAi construct

The confirmed entry clones, pENTR207-$S_{o1}$/$S_{o2}$ were incubated with the *att*R-containing destination vector, pHellsgate8 in an LR recombination reaction and LR reaction transformed into DH5α competent cells. The transformation reaction was then plated on LB agar plates containing the appropriate antibiotics (100 µg/ml of spectinomycin). The transformation was found to be successful as large numbers of colonies grew on the plates. The colonies were then screened using both the vector specific primers and gene specific primers (Table 6.1) in different combinations and the putative RNAi clones (pHG8-$S_{o1}$/$S_{o2}$) selected. Figure 6.14 shows the map of the primers used for selecting the expression/RNAi clones.

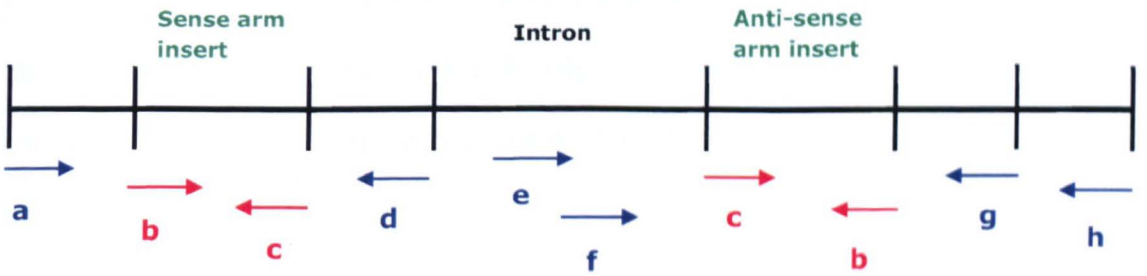



**Figure 6.14 A schematic of the map of primers used for screening the putative pHG8-$S_{o1}$/$S_{o2}$ clones.** A combination of both vector and gene specific primers were used to screen the colonies growing on the plate. The gene specific primers are shown in red and the vector primers are in blue and their names are abbreviated with letters (a-h). a= Ph8xof, b= So1So2-FP, c= So1So2-RP, d= Ph8xor, e= HELLINT2, f= Ph8xbaf, g= Ph8xbar, h= OCSrev (see Table 6.1).

The colony PCR results showed that, some of the colonies screened (~25 colonies) seem to yield the expected amplicon sizes for the various primer combinations. An example of the initial colony PCR performed to select colonies for further confirmation is shown in Figure 6.15 (a & b). Plasmid DNA was extracted (section 2.3.4) from the selected putative RNAi clones and further confirmation performed on them using various primer combinations (Figure 6.16 a, b & c) and restriction enzyme digestion (Figure 6.17 a & b) to check the orientation of the inserts and also ensure that the intron hasn't flipped in the construct. The results showed that, most of the selected clones contained the expected inserts and in the right orientations with intact intron directions in the vector. The PCR confirmation initially showed that, all the selected clones are the putative RNAi clones. However restriction enzyme digest results revealed that not all the clones observed in the PCR are the putative RNAi clones, *i.e.* only ~50% of the clones were found to be the expected putative RNAi clones. This was further confirmed through sequencing of the inserts in the constructs. The map of the created RNAi construct (pHG8-$S_{o1}$/$S_{o2}$) is shown in Figure 6.18.

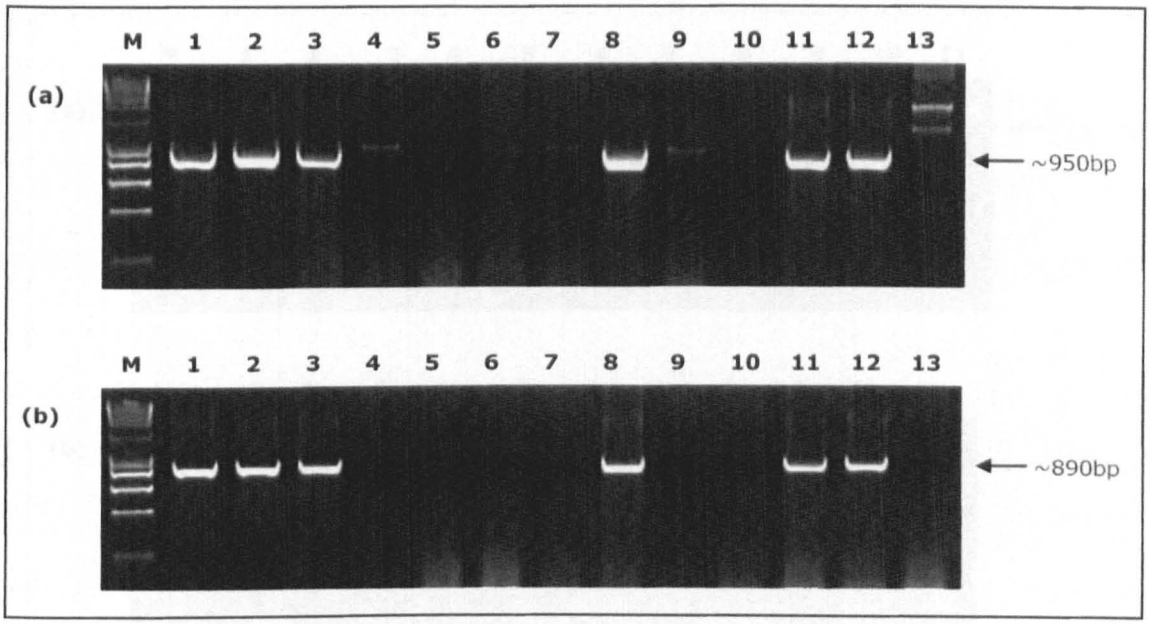



**Figure 6.15 Colony PCR of the putative RNAi clones.** The putative RNAi clones (pHG8-$S_{o1}$/$S_{o2}$) were initially screened using a combination of both vector and gene specific primers. (a) Initial screening of the sense arm of the putative RNAi clones using vector specific primer Ph8xhof and gene specific primer So1So2-RP. (b) Initial screening of the anti-sense arm of the putative RNAi clones using vector specific primer Ph8xbaf and gene specific primer So1So2-FP. The PCR products that amplified gave the expected amplicon of ~950bp (a) and ~890bp (b). Lane M represents Hyperladder I, lanes 1-12 represent the putative RNAi clones (lanes 1, 2, 3, 8, 11 and 12 gave the expected results and were selected), lane 13 represents the empty pHellsgate8 vector.

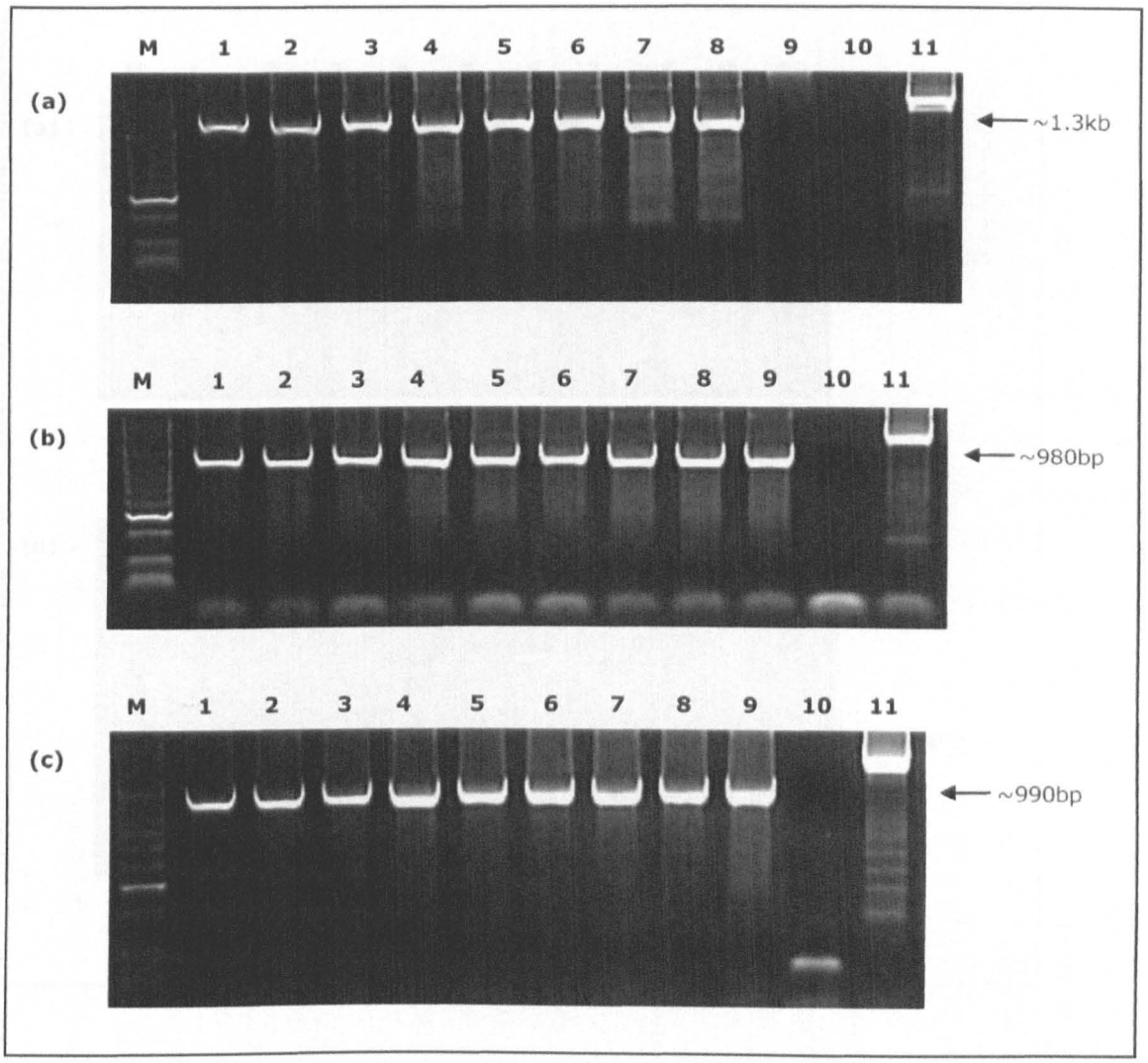



**Figure 6.16 PCR to confirm the selected putative RNAi clones.** The putative RNAi clones (pHG8-$S_{o1}$/$S_{o2}$) were confirmed using a combination of both vector and gene specific primers. (a) Confirming the sense arm of the putative RNAi clones using vector specific primers, Ph8xhof and Ph8xhor. (b) Confirming the anti-sense arm of the putative RNAi clones using vector specific primers, Ph8xbaf and Ph8xbar. (c) A forward primer (HELLINT2) designed from the intron region in the vector and another vector primer (Ph8xbar) was used to confirm that the intron separating the sense arm of the insert from the anti-sense arm hasn't flipped. The PCR products that amplified gave the expected amplicons of ~1337kb (a), ~980bp (b) and ~997bp (c). Lane M represents Hyperladder II, lanes 1-9 represent the putative RNAi clones, lane 10 represents water control, lane 11 represents the empty pHellsgate8 vector.

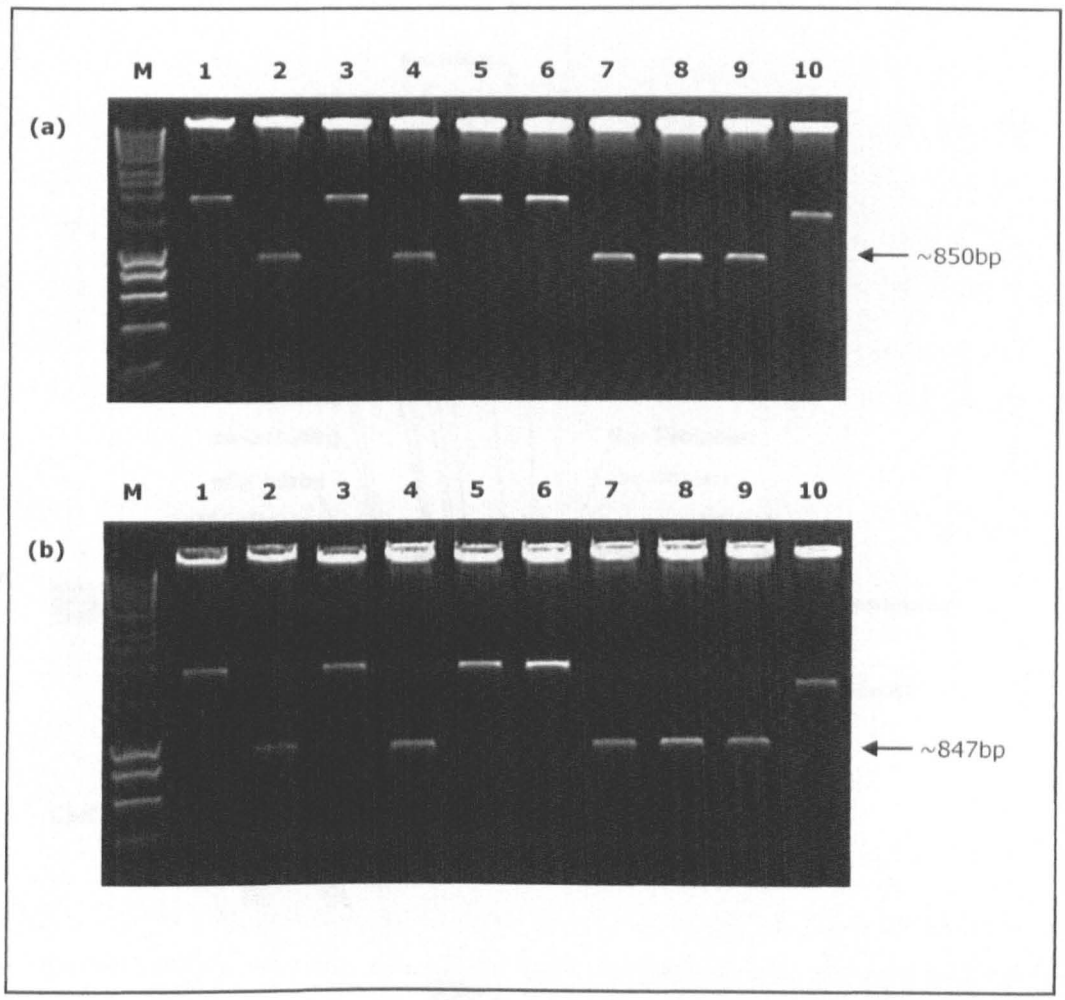



**Figure 6.17 Restriction enzyme digests to confirm the putative RNAi clones.** (a) Confirming the sense arm insert. The selected putative RNAi clones were digested with *Xho*I which directly flanks both the 5′ and 3′ ends of the insert in the sense arm of the construct. The putative RNAi clones gave the expected insert size of ~850bp following the digest. (b) Confirming the anti-sense arm insert. The selected putative RNAi clones were digested with *Xba*I which directly flanks both the 5′ and 3′ ends of the insert in the anti-sense arm of the construct. The putative RNAi clones gave the expected insert size of ~847bp following the digest. Based on the two (a & b) digest results, clones in lanes 2, 4, 7, 8 & 9 were selected as the putative RNAi clones (pHG8-$S_{o1}$/$S_{o2}$) and the inserts in the constructs confirmed through sequencing. See Figure 6.18 for the two enzyme positions in the RNAi construct. Lane M represents Hyperladder I, lanes 1-9 represent the putative RNAi clones, lane 10 represents the empty pHellsgate8 vector.

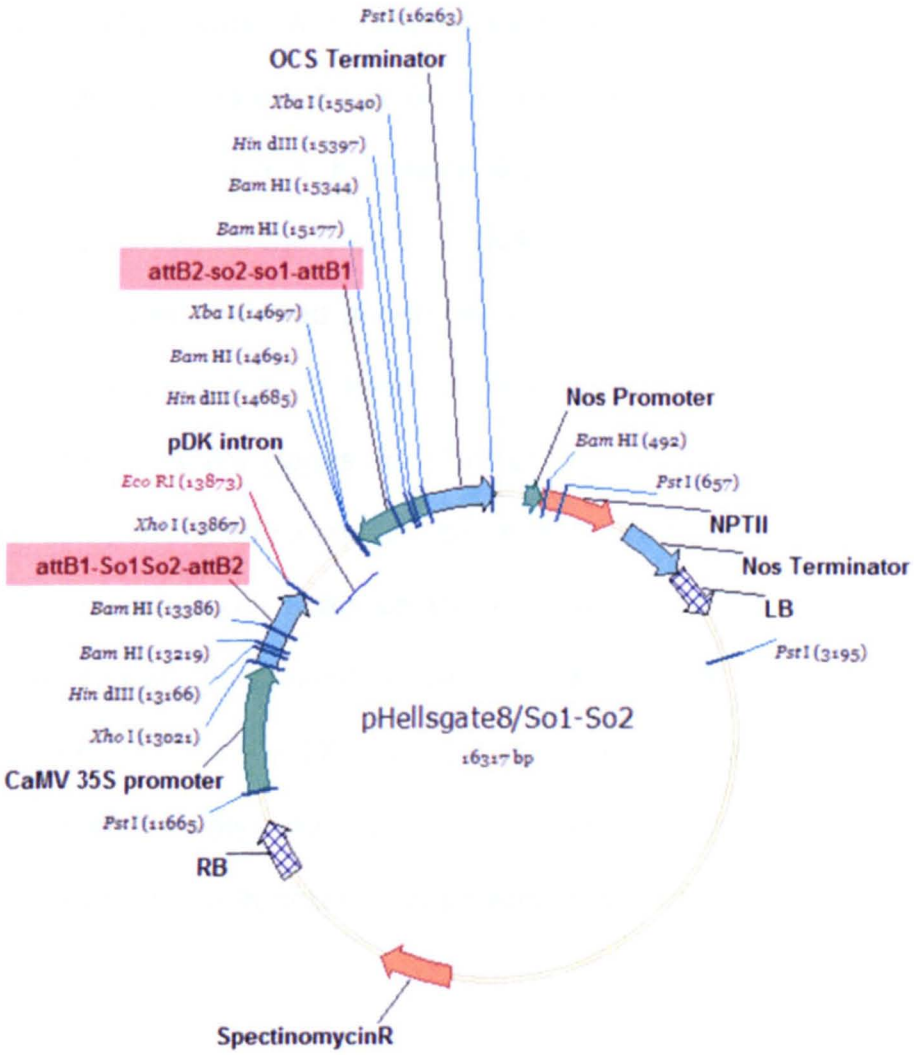



**Figure 6.18 Map of the designed RNAi construct pHG8-$S_{o1}/S_{o2}$.** The S-RNase sequences in the construct are labelled as attB1-So1So2-attB2 and attB2-So2So1-attB1 in the sense and antisense arm of the construct respectively.

### 6.3.4 Transformation of RNAi construct into *A. tumefaciens*

Three of the confirmed RNAi constructs (designated as pHG8-So1/So2.1, pHG8-So1/So2.2 and pHG8-So1/So2.3) carrying the gene of interest were transformed into *Agrobacterium tumefaciens* strain LBA4404 electro-competent cells using electroporation (see section 6.2.7). The

transformation was found to be successful as large number of colonies grew on the plates. The colonies were screened using the same primer combinations used in section 6.3.3 to confirm the RNAi clones (see Figure 6.14; Figure 6.15 a & b; Figure 6.16 a, b & c). Most of the colonies screened contained the expected RNAi clones (data not shown) consistent with what was observed previously in section 6.3.3 (see Figure 6.15 a & b; Figure 6.16 a, b & c). Plasmid DNA was then extracted (section 2.3.4) from the selected clones for further analysis. The restriction enzyme digestion performed to confirm the RNAi clones previously (Figure 6.17 a & b) was carried out again on these clones to confirm their identity. The results (data not shown) were consistent with what was previously observed (in Figure 6.17 a & b). The confirmed clones were then grown in LB broth and glycerol stocks made (section 2.3.5) from them and are now ready for use in potato plant transformation.

## 6.4 DISCUSSION

### 6.4.1 $S_{o1}/S_{o2}$-RNase chimeric gene

One of the main objectives of this study was to silence specific S-RNase genes in selected self-incompatible diploid potato species leading to the development of self-compatible diploid stocks. The alleles earmarked for silencing in this study are the $S_{o1}$- and $S_{o2}$-RNases found in some accessions of *S. okadae*. Since these two alleles seem to be the most predominant in the *S. okadae* accessions we have genotyped, the advantage of using a construct which can silence both alleles in a

heterozygote ($S_{o1}/S_{o2}$) would be more beneficial than just targeting one allele. This is because, it can offer the possibility of using the designed construct on *S. okadae* accessions other than the $S_{o1}/S_{o2}$ heterozygous ones, *i.e.* other accessions harbouring either of the alleles ($S_{o1}$ or $S_{o2}$). The simpler way to achieve this was by combining portions of $S_{o1}$-RNase and $S_{o2}$-RNase to create a chimeric gene ($S_{o1}/S_{o2}$-RNase). This is a novel strategy and has not been attempted prior to this study.

The hybrid $S_{o1}/S_{o2}$-RNase gene was created using conventional cloning approaches with selected restriction enzymes. The fragment length of the created chimeric gene is ~795bp (*i.e.* 393bp of $S_{o1}$-RNase and 402bp of $S_{o2}$-RNase). The length of a gene fragment is an important determinant for effective gene silencing in RNAi. The recommended length of gene fragments which can cause successful and maximum effective gene silencing in plants is 300-800bp (although gene fragments of ~1kb have also been used successfully). Shorter gene fragments however are less successful in achieving effective gene silencing (Helliwell and Waterhouse, 2005). It is worth pointing out that, since the regions used for creating the chimeric S-RNase gene include some conserved domains (C1-C5), it is possible that this construct could potentially be used for silencing other S-alleles, since the mechanism of silencing in RNAi is based on nucleic acid sequence recognition (see section 6.1.1). For instance an antisense construct designed by Lee *et al.* (1994) to suppress the expression of $S_3$-RNase in a self-incompatible $S_2S_3$ *Petunia inflata* plant, was observed not only to cause the failure of the transgenic plants to reject $S_3$ pollen but also its $S_2$ pollen producing a fully compatible plant

(see section 1.6.1). The sequence homology shared between these two alleles could account for these observations. Similar effects could be plausibly observed for the chimeric S-RNase gene used for designing the RNAi construct in this study.

### 6.4.2 RNAi construct design for $S_{o1}/S_{o2}$-RNase gene silencing

S-alleles have been successfully manipulated in SI species in attempts to prove that S-proteins are responsible for S-allele-specific pollen recognition and rejection in the pistil (*e.g.* Lee *et al.*, 1994; Murfett *et al.*, 1994). The chimeric S-RNase gene ($S_{o1}/S_{o2}$) constructed was used to create the RNAi silencing construct using the Gateway® cloning system. It has been shown that, when double stranded RNA is expressed in plants, it triggers the RNAi pathway which eventually leads to homologous mRNA degradation (Waterhouse *et al.*, 1998; Baulcombe, 2004). The cloning of part of a target gene in inverted orientation relative to each other downstream of a strong promoter will help trigger this RNAi response in plants. This was successfully achieved with the created chimeric $S_{o1}/S_{o2}$-RNase gene using two gateway compatible vectors; firstly a donor vector (pDONR207) to help with creating an entry clone and secondly a destination vector (pHellsgate8) to create the RNAi construct or expression clone.

Gateway® cloning has become the cloning system of choice by taking advantage of the λ-site specific recombination system of bacteriophage λ without the need for restriction enzyme digest and DNA ligase-mediated

cloning as in conventional systems. Two recombination reactions constitute the gateway system; the first is the BP recombination reaction between *att*B and *att*P sites to create an entry clone. The recombination of *att*B containing $S_{o1}/S_{o2}$ chimeric gene with *att*P containing donor vector, pDONR207 has enabled the creation of an *att*L containing entry clone (pENTR207-$S_{o1}/S_{o2}$). The created entry clone was then entered into the second gateway recombination reaction called the LR reaction.

In LR recombination reaction, the *att*L containing entry clone recombines with an *att*R containing destination vector to create an *att*B containing expression clone. Gateway cloning system depends on appropriate gateway-compatible plant destination vectors, many of which can replicate both in *E. coli* and in *A. tumefaciens*. In addition, these vectors contain the left and right border sequences to allow for *Agrobacterium*-mediated transfer of T-DNA (Earley *et al.*, 2006). For the RNAi construct design in this experiment, the pHellsgate vector series (pHellsgate8) was chosen. The pHellsgate vector series have been designed (Wesley *et al.*, 2001; Helliwell and Waterhouse, 2003) to carry target genes in both forward and inverted orientations spaced by an intron to help trigger the RNAi response (Smith *et al.*, 2000; Wesley *et al.*, 2001; Helliwell and Waterhouse, 2003). Beside other advantages, pHellsgate vector series have suitable restriction enzymes that can be used for confirming the inserts (gene of interest) and also check for the orientation of the spacer intron (Helliwell and Waterhouse, 2003).

The recombination of the created entry clone, pENTR207-$S_{o1}/S_{o2}$ with pHellsgate8 destination vector successfully generated an expression

clone/RNAi construct (pHG8-$S_{o1}$/$S_{o2}$). In selecting for the putative RNAi clones following the recombination reactions, they (clones) were initially screened using PCR based approaches. However although the PCR's enabled us identity potential putative RNAi clones from the background clones, it was observed later that, results from the PCRs alone were not conclusive enough for clone confirmation. For instance, when the putative clones identified by PCR were digested with restriction enzymes, some of them did not have the right inserts and were discarded. Therefore, it is worth mentioning that, although PCRs are useful for clone identification, other confirmation techniques (*e.g.* restriction enzyme digest and sequencing) are necessary to confirm the identified putative clones especially those that will be used for further downstream applications.

The successful delivery of RNAi constructs into plants is dependent on the availability of appropriate and efficient delivery systems. Different delivery systems have been used to introduce gene silencing constructs into plants in RNAi, one of which is the *Agrobacterium*-mediated transformation techniques. *Agrobacterium*-mediated transformation is a well-established technique in potato. However it is worth mentioning that, although potato transformation is a fairly straightforward approach once the protocol is established (compared with other plant species), it is highly genotype dependent (Wheeler *et al.*, 1985; Badr *et al.*, 2008), *i.e.* there are no universal protocols for transforming potatoes and has to be optimized and established empirically for each potato genotype before use. In order to introduce our RNAi construct (pHG8-$S_{o1}$/$S_{o1}$) into potato plants, the construct carrying the transgene was introduced into

*Agrobacterium tumefaciens* strain LBA4404. The resulting colonies were screened and confirmed ready for use for potato plant (OKA1/OKA3 and other accessions harbouring either one of $S_{o1}$/$S_{o2}$-alleles) transformation. Hopefully when this is performed, it will generate transgenic plants with reduced expression of the targeted S-RNase genes which can be self-pollinated to produce self-compatible diploid potato plants leading to the development of the first inbred lines of *S. okadae*.

# CHAPTER 7: GENERAL DISCUSSION

Angiosperms have devised various strategies to avoid inbreeding and promote outcrossing thereby maintaining genetic diversity among species of populations. One such widespread strategy is the mechanism of self-incompatibility. In a self-incompatibility mechanism, a pollen grain recognized as self is arrested and prevented from reaching the ovary to effect fertilization whilst non self-pollen grains grow successfully to reach the ovary to effect fertilization. Different forms of self-incompatibility exist in flowering plants; however the most prevalent form is the gametophytic self-incompatibility system (GSI). The Solanaceae, Rosaceae, Plantaginaceae, Leguminoceae, Onagraceae, Papavaraceae and Poaceae are among the plant families found to be exhibiting the GSI system. Gametophytic self-incompatibility however has been observed to operate by two different mechanisms of recognizing and rejecting self-pollen tubes. One of which is found only in the Papavaraceae so far and the other commonly referred to as the S-RNase-based GSI system has been studied extensively in the Solanaceae, Rosaceae and Plantaginaceae. This project ultimately aims to characterize the S-RNases in diploid potato species specifically accessions of *Solanum okadae*, *Solanum phureja* and *Solanum stenotomum* and subsequently manipulate the expression of these S-RNases using the RNAi technique.

## 7.1 PHENOTYPIC CHARACTERIZATION OF S-ALLELES

### 7.1.1 Self-incompatibility and compatibility relationships

The characterization of S-alleles can be carried out phenotypically by pollination tests. Although this method can be very laborious and may not reveal the identity of the S-alleles harboured by these plants, especially in previously un-genotyped plant populations, pollination tests can identify plants that are cross and self-incompatible in a large plant population. Where the identity of the S-alleles involved are necessary, then pollination tests could be followed up with molecular based genotyping techniques.

For instance the use of pollination tests in a diallel cross design has enabled the characterization of compatibility relationships in a population of *Petunia inflata* of previously unknown S-genotypes (section 3.4.1.1). This technique was useful in checking the SI status of the plant population (the SC lines were discarded) and also allocating the remaining SI lines into incompatibility groups. This approach reduced the number of plants with unknown S-alleles whose identity were determined using molecular based genotyping techniques. In general, two incompatibility groups were derived for the SI plants in this *Petunia inflata* population. The common observation was that, plants belonging to the same group are self-incompatible and only set seed with crosses with plants in the other incompatibility group. The explanation for this is that, plants found in the same incompatibility group harbour the same pair of S-alleles hence crosses among them are expected to result in self-pollen tube growth arrest leading to the eventual failure of fertilization. A

compatible reaction and successful fertilization however occurs in the inter-group crosses because the S-haplotypes expressed in the pollen and styles do not match hence resulting in successful pollination and the fertilization.

A similar approach was used to predict the compatibility relationship among potato stocks of *Solanum okadae* and also confirm their SI status (section 4.3.2). With the help of pollination tests in a diallel cross design, plants belonging to the same incompatibility groups were determined enabling us to identify those that are cross-incompatible (bearing the same S-allele) and those that are cross-compatible (*i.e.* differing by at least one S-allele). Also through the pollination test, the self-incompatibility status of the *S. okadae* accessions was confirmed where all the self-pollinated plants were observed to lack berries/seed set thereby confirming that the accessions are indeed self-incompatible.

### 7.1.2 The analysis of self-compatibility trait in Petunia Mitchell

*Petunia hybrida cv* Mitchell is known to be self-compatible and although the basis of its SC behaviour is unknown the plant is assumed to be homozygous for an uncharacterised S-allele (provisionally called $S_m$-RNase). The use of pollination tests (section 3.4.2.1) has enabled the S-genotype relationship determination in plant populations derived from crosses between Mitchell and lines with a functional SI background. In addition the pollination test has enabled the ascertainment of whether the self-compatibility phenotype observed in *Petunia hybrida cv* Mitchell

is linked to the S-locus. Through the pollination tests, it was possible to observe the segregation of SC and SI traits in the plant population segregating for the known S-alleles. Results from the pollination data revealed that there is a possibility of the linkage of the S-allele of Mitchell to the observed self-compatibility trait in the plant population. This observation was confirmed with PCR S-genotyping with allele-specific primers (see section 3.3.2.4) on the plant population. The observation from the analysis showed that, all the SC plants harbour the $S_m$-allele but those that were SI did not; an observation which strongly implicates the $S_m$-allele of Mitchell to the segregating self-compatibility trait in the plant population studied.

This observation raised the obvious question why Mitchell is self-compatible and what could be the genetic basis of this? Self-compatibility could arise out of an otherwise self-incompatible plant species as a result of defective stylar protein, loss of pollen-S function, the presence of S-locus linked or unlinked modifier genes (Royo *et al.*, 1994; Tsukamoto *et al.*, 1999; 2003a; 2003b). Any of these effects could be a potential cause of the SC feature of *Petunia hybrida cv* Mitchell. However, a defective stylar protein is unlikely to be the cause of SC in Mitchell since it has been shown to have ribonuclease activity (Ajiboye and Sonneveld, unpublished) and also the sequencing of the $S_m$-RNase (see section 3.3.2.2) showed that it exhibited all the primary structural features of a functional S-RNase. It is nevertheless, worth mentioning that a defective stylar protein ($S_m$-RNase) could still account for the failure of Mitchell to reject self-pollen. This is because its ($S_m$-RNase) expression profiling has

not been conducted yet hence a reduced expression of the $S_m$-RNase gene far below its rejection threshold could plausibly account for the cause of self-compatibility feature of *Petunia hybrida cv* Mitchell. Another conceivable cause of the self-compatibility feature of Mitchell (resulting from the pistil SI determinant) is the presence of S-locus linked modifiers or compatibility factors rendering the $S_m$-RNase of Mitchell non-functional.

## 7.2 MOLECULAR CHARACTERIZATION OF S-ALLELES

### 7.2.1 Degenerate primer approach to S-RNase cloning

S-allele cloning and characterization can be carried out genotypically by using PCR based approaches with degenerate oligonucleotide primers. Over recent years, there has been a surge in the availability of gene sequences for many plant species as a result of the advances in the use of molecular approaches thereby making it possible to design degenerate primers from related species for use in cloning genes from other members of the gene family. The use of degenerate oligonucleotides has facilitated the cloning and identification of genes in many plant species. A degenerate oligonucleotide primer designed and synthesized (in this current study) based on the conserved C2 domain of solanaceous S-RNases and used in RT-PCR has enabled the cloning of S-RNase sequences from *Petunia* (chapter 3 of this thesis) and potato (chapter 4 of this thesis) species. To test the universality of this degenerate primer for cloning S-alleles in other solanaceous species, the primer was tried on

tomato species also and has been observed to have successfully amplified S-RNases whose sequences were cloned and sequenced to confirm their identity (data not reported in this thesis). This degenerate oligonucleotide primer (SolC2-F1.3) therefore represents a valuable primer which can be used for potentially cloning S-RNases from any solanaceous species as evident in this study.

### 7.2.2 Sequence analysis of S-RNase genes

With the help of the degenerate primer in an RT-PCR approach, four putative S-RNase sequences have been reported here from *Petunia* species out of which one is a full length gene sequence and sixteen from potato species out of which two are full length gene sequences. S-RNase genes involved in the S-RNase-based GSI system have been observed to share some primary structural features. The analysis of the sequences of the putative S-RNase genes cloned with the help of the degenerate primer in this study showed that all the cloned putative S-RNase genes possess the primary structural features of solanaceous S-RNases as defined by Ioerger *et al.* (1991). The analysis of solanaceous S-RNases revealed five conserved domains and two hypervariable domains (Ioerger *et al.*, 1991), a feature which is distinctive from rosaceous S-RNases which lack one such conserved domain and one hypervariable domain (Ishimizu *et al.*, 1998). The conserved region of the S-RNase gene is responsible for the catalytic activity of the gene whilst the hypervariable domains could plausibly be responsible for allele specificity determination (see section 1.3.2.1). Also the cloned putative S-RNases were observed

to have strong basic isoelectric point (pI) values, which conform to the pI value of functional S-RNases. It is however worth pointing out that, although functional assays have not been performed for the cloned S-RNases, their sequence characterization coupled with pollination tests showed that they are definitely genuine S-RNases and function in the SI reaction and no cDNA sequence was available for them prior to this study.

Subsequent to the identification of the first S-protein sequence from *Nicotiana* (Anderson *et al.*, 1986), a large number of S-RNase sequences have been isolated from other species of the Solanaceae. However, unlike *Petunia* species and some other members of the Solanaceae, relatively few S-RNase gene sequences are available for potato species. The relatively large number of putative S-RNases identified from the relatively small number of potato genotypes from this current study implies that, there is possibly a high S-RNase gene variability and diversity in potato. However, this S-gene variability (in potato) has not been well exploited and characterized by researchers studying the field of GSI compared to other plant species particularly in the Solanaceae. The S-RNase gene reported here therefore represents a unique addition to the rather few available potato S-RNase gene sequences.

### 7.2.3 Phylogenetic analysis of S-RNase genes

Three classes of $T_2$-type RNases have been identified (Igic and Kohn, 2001). RNases involved in the GSI system belong to the class III $T_2$-type RNases. The phylogenetic analysis of the cloned putative *Solanum* S-RNases with selected S-like RNases (class I & II of $T_2$-type RNases), where the cloned S-RNases were observed to fall outside the S-like RNases clade confirmed that, the cloned S-RNases are clearly distinct and distantly related to the S-like RNases and represents genuine S-RNases. Although class I & II $T_2$-type RNases are unrelated to the S-locus and play no role in self-incompatibility reaction, they are known to be structurally related to functional S-RNases (belonging to the class III $T_2$-type RNases).

The S-locus was suggested to be under the influence of unusual evolutionary forces which tends to maintain an exceptionally high level of polymorphism of the S-locus genes. The distribution and polymorphism of the S-locus alleles within plant populations are governed by high frequency-dependent selection (Wright, 1939; Ioerger *et al.*, 1990). In the Solanaceae, it has been well established that S-allele polymorphism predates the divergence of species. The phylogenetic tree in this study shows that interspecies specific clades are prevalent over intraspecies specific ones lending credence to the ancient polymorphism at the solanaceous S-locus. This observation was supported when the amino-acid sequence comparison of the cloned putative potato S-RNases (section 4.3.3.1) showed that, sequence similarity within a species could be as low as 32.9%, a finding similar to the 40% observed by Ioerger *et*

*al.* (1990) in other solanaceous species. Also the trans-generic clades observed in the phylogeny of solanaceous S-alleles in this study is consistent with the trans-generic evolution found in solanaceous species, an observation which provides evidence that, their S-lineages are inherited through multiple events of speciation (Richman *et al.*, 1996). Thus S-RNase allelic variants are extremely old and have been inherited from a common ancestor and passed down to multiple descendant taxa; *i.e.* the S-lineage origin predates the origin of the current species in which they reside.

## 7.3 ANALYSIS OF THE S-RNase GENE EXPRESSION

S-locus ribonuclease genes termed S-RNases that function in the self-incompatibility reaction are known to be expressed exclusively in the pistil tissues (Anderson *et al.*, 1986; 1989; Cornish *et al.*, 1987). Other RNases unrelated to the S-locus termed S-like RNases, have also been identified and observed to share similar domain structure with S-RNases but lack the polymorphism observed in S-RNases, although they (S-like RNases) appear to be related to S-RNases in phylogenetic studies. However, they show very little tissue specificity (*i.e.* are ubiquitous) in their expression unlike S-RNases (MacIntosh *et al.*, 2010; Nowak *et al.*, 2011). In the S-RNase-based GSI system, incompatible pollen tube inhibition is caused by the S-RNase gene when the pollen tube is approximately half-way through the style, the site of the SI reaction. The analysis of the tissue specific expression pattern of the $S_{o1}$- and $S_{o2}$-

RNases was consistent with what is known of the S-RNase gene. Both alleles showed high tissue (pistil) specificity with the highest expression found in the style. This lends further support to the observation that, the cloned S-RNases are genuine ribonucleases involved in the SI reaction and not S-like RNases involved in defence against pathogens or released in response to phosphate starvation.

The S-RNase gene expression increases during flower development, where its full expression coincides with the flower attaining maturity. The pistil is known to acquire its full SI phenotype once the flower has reached its fully matured stage (Clark *et al.*, 1990; McCubbin and Kao, 2000). Pistils at early bud stage development accumulate very little S-RNase, below the threshold for rejecting self-pollen tubes, thereby allowing successful fertilization resulting in seed set. The developmental profiling of the S-RNase gene in this study conforms to this observation where the $S_{o1}$- and $S_{o2}$-RNase expression was observed to be very low in immature buds, increasing prior to flower opening and reaching its highest expression levels in fully opened matured flowers. This feature of the S-RNase gene has been exploited for overcoming SI through immature bud self-pollination (Bernatzky *et al.*, 1988) leading to the production of S-allele homozygous stocks in the Solanaceae (*e.g.* Robbins *et al.*, 2000).

The comparison of expression for two alleles of the S-RNase gene ($S_{o1}$- and $S_{o2}$-RNase) in this study showed that there were no significant differences in their transcript levels. Significant differences or no differences can be observed for the same S-RNase genes under different

genetic backgrounds (Qin *et al.*, 2001; 2006). However, it is worth mentioning that, in such cases where differences are found in the transcript levels of the same S-RNase expressed under different genotypes, the variation might not have any significant effect on the minimal threshold of S-RNases required quantitatively for incompatible pollen tube rejection. The ability of S-RNase genes in self-pollen tube rejection is known to be dependent on a threshold level of the S-RNase transcript (see section 5.1.1).

## 7.4 RNAi CONSTRUCT DESIGN FOR GENE SILENCING

Sequence-specific gene silencing has received much attention in recent years basically because of its fascinating biology and its usefulness as a powerful experimental tool. One such gene silencing mechanism that has received much study and understanding is the phenomenon of RNA interference (RNAi, see Chapter 6 of this thesis). The discovery that the mechanism of RNAi, aside its other roles (*e.g.* in response to pathogenic and parasitic nucleic acid invasion), also functions in basic cellular events (*e.g.* gene regulation and heterochromatin formation) has generated a great deal of interest in this field by researchers partly leading to its rapid widespread recognition. This discovery is perhaps the most important advance made in molecular biology over recent decades (Denli and Hannon, 2003; Novina and Sharp, 2004). The use of RNAi has enabled the silencing of genes in both plants and animals thereby providing the thrust for the rapid discovery and confirmation of gene functions. RNAi

has become the gene silencing method of choice compared with antisense-mediated gene silencing and co-suppression (see section 6.1.1) partly because it (RNAi) offers high efficiency and stability and also has the ability to suppress the expression of transgenes in multi-gene families. Silencing in RNAi occurs using a transgene which produces a hairpin RNA (hpRNA) with a dsRNA region triggering gene silencing (Waterhouse and Helliwell, 2003; Kusaba, 2004).

The silencing of S-RNases in selected diploid potato species leading to the development of inbred lines was one of the main objectives of this study. A combination of fragments of two alleles ($S_{o1}$-RNase and $S_{o2}$-RNase) was performed to create a chimeric gene prior to the RNAi construct design. This was in order to increase the chances of silencing both alleles in a heterozygote and also offer a wide usage of the designed construct in silencing S-alleles in other accessions besides the targeted accessions. Since the regions used for creating the chimeric S-RNase gene include some conserved domains and since the mechanism of silencing in RNAi phenomenon is based on homology-dependent sequence recognition and silencing, it is possible that the chimeric gene could be used for silencing other S-alleles mediated by the conserved S-RNase sequences in the silencing construct. This could potentially make the designed RNAi construct a 'universal' S-RNase silencing construct in potato when tested.

The cloning of the chimeric S-RNase gene in both sense and inverted repeats in pHellsgate8 vector to create an RNAi silencing construct was achieved via an LR recombination reaction using LR Clonase™ II enzyme mix. The pHellsgate8 RNAi vector carries the target gene ($S_{o1}$/$S_{o2}$-RNase)

in both forward and inverted orientations spaced by an intron to help trigger the RNAi response. The cloning of sense and antisense genes separated by an intron spacer driven by the same promoter leads to the production of hairpin RNA (hpRNA) which triggers gene silencing (Smith *et al.*, 2000). Although the use of introns as spacers for RNAi constructs was observed to increase the efficiency of silencing, the mechanism still remains elusive (Smith *et al.*, 2000; Helliwell and Waterhouse, 2003). Intron splicing from the RNAi construct could help enhance the efficiency of silencing by possibly helping in the alignment of the complementary arms of the hairpin in an environment which promotes the hybridization of RNA thereby encouraging duplex formation. On the other hand, intron splicing could lead to an increase in hpRNA either by aiding or impeding the passage of the hairpin from the nucleus or by forming a loop which is less nuclease-sensitive (Smith *et al.*, 2000). The RNAi construct (pHG8-$S_{o1}/S_{o1}$) designed in this study is now a valuable resource for use in S-RNase gene silencing in potato leading to the development of self-compatible diploid potato lines.

## 7.5 TRANSFORMATION OF THE RNAi CONSTRUCT *IN PLANTA* AND IMPLICATIONS FOR FUTURE WORK

An efficient and suitable means of construct delivery *in planta* is a prerequisite for any successful transgenic work and *Agrobacterium*-mediated transformation is a well-established and commonly used technique in potato although it requires the need for optimization for

each potato genotype before use. Thus published regeneration methods for potato have been observed to be cultivar- or genotype-dependent. This is because of the special media requirement needs and composition of every potato genotype (Wheeler *et al.*, 1985; Badr *et al.*, 2008). To allow for transformation *in planta*, the RNAi construct (pHG8-$S_{o1}/S_{o1}$) carrying the transgene of interest was successfully introduced into *Agrobacterium tumefaciens*. Over the years, there has been an improvement in plant transformation vectors and techniques for increased efficiency offering stability of transgene expression in crop plants. The transformation of plants using *Agrobacterium Ti*-plasmid-based vectors remains to be one of the best transformation systems partly because of its simplicity and the accurate transgene integration besides other advantages (Veluthambi *et al.*, 2003). This method has been one of the most commonly and widely used for the transformation of potatoes.

Although the RNAi constructs designed (and transformed into *Agrobacterium tumefaciens*) in this study has not been tested *in planta* yet due to time constraints, it is expected that the constructs will serve as a valuable resource for use in S-RNase gene silencing in potato leading to the development of self-compatible diploid potato lines. When *in planta* transformation of the constructs is finally achieved, the reduced expression of the S-RNases of primary transformants will be checked using allele-specific primers and quantitative RT-PCR. The baseline expression level for S-RNases prior to transformation has been established (see Chapter 5). Controlled pollinations will be used to

determine the SI/SC phenotype in transgenic lines showing reduced expression of the S-RNase gene. Seeds from the transgenic lines exhibiting self-compatibility will then be collected and sown to initiate the development of the first inbred lines of *S. okadae* (and other selected diploid *Solanum* species, *e.g. S. stenotomum* and *S. phureja*). The availability of the self-compatible (homozygous) diploid potato lines (exhibiting desired characteristics) would therefore play key roles in facilitating many genetic analyses including that of complex traits.

It is worth commenting at this point that, the silencing of S-RNases is not the only possibly way of causing an otherwise self-incompatible plant to become self-compatible. One such idea which could be potentially explored in future projects for creating self-compatible diploid potatoes is to silence the *HT-B* modifier gene. *HT-B* was identified to be a small asparagine-rich protein that was observed to play a crucial role in eliciting the self-incompatibility response (see section 1.3.2.1) although unlinked to the S-locus. S-RNases were observed to be compartmentalized in a vacuole and remains sequestered in a self-compatible reaction and requires HT-B protein to breakdown the vacuoles in a self-incompatible reaction to release the S-RNase to exert its cytotoxic effect. HT-B protein degradation was observed to occur in compatible reactions whilst its accumulation was observed in the vacuolar compartments containing S-RNases in an incompatible reaction (Goldraij *et al.*, 2006). Kondo *et al.* (2002) in a study observed that the cause of the loss of self-incompatibility in SC tomato was as a result of a reduced or lack of expression of *HT-B* transcript. All the *HT-B* genes identified so

far show high levels of sequence conservation and therefore could be exploited for creating 'universal' silencing constructs leading to self-compatibility engineering. RNAi constructs therefore designed and targeted towards *HT-B* could prevent the breaking down of the vacuolar compartment containing S-RNases hence sequestering them (S-RNases) thereby rendering them unavailable and inactive allowing the growing pollen tube to reach the ovary to effect fertilization. Also now that the pollen-S determinant gene (*SLF*) has been identified for the S-RNase-based GSI system and as more information becomes available; another option or strategy which could be potentially explored for manipulating self-incompatibility leading to the creation of self-compatible plants could be based on manipulating the pollen-S determinants (*SLF's*) function.

## 7.6 CONCLUDING REMARKS

The use of degenerate oligonucleotide primer in RT-PCR and the use of allele specific primers have enabled the identification and characterization of S-alleles in *Petunia* and potato species in this investigation leading to the development of an RNAi construct for gene silencing in potato. The identified S-RNases in this study showed the characteristic distribution of conserved and variable sequences of solanaceous S-RNases. Also the identified S-RNases showed a general pattern of allelic variance (polymorphisms) and also form clades in phylogenetic studies with functional alleles. The identified alleles represent unique and useful additions for studying diversity and phylogeny of S-alleles particularly in

*Solanum* (potato). For instance allele specific primers could be designed from the sixteen putative potato S-RNases identified in this investigation and used for S-allele genotyping and studying of the diversity of S-RNases in the large collection of potato DNA-bank at the James Hutton Institute (JHI). Although diversity of S-alleles has been studied in other solanaceous species, to the best of our knowledge this is yet to be conducted comprehensively in potato. Performing this, could therefore give further idea of the diversity of S-alleles in potato species thereby contributing significantly to the existing knowledge of the diversity of S-RNases in the Solanaceae family. Despite the enormous progress made so far in unravelling knowledge in one of the novelties of angiosperms success and evolution, *i.e.* the self-incompatibility system; the elucidation of more information regarding this system continues to be one of the principal aspirations of many researchers in this field. These collective efforts promise to continue to shed more light on the understanding of the intricacies of this fascinating but complex system.

## REFERENCES


Ai Y, Kron E, Kao TH (1991) S-alleles are retained and expressed in a self-compatible cultivar of *Petunia hybrida*. *Molecular and General Genetics* **230**: 353-358.

Ai Y, Singh A, Coleman CE, Ioerger TR, Kheyr-Pour A, Kao TH (1990) Self-incompatibility in *Petunia inflata*: isolation and characterisation of cDNAs encoding three S-allele associated proteins. *Sexual Plant Reproduction* **3**: 130-138.

Ajiboye J (2008) Investigating the basis of self-compatibility in *Petunia hybrida* 'Mitchell'. MSc Thesis at University of Nottingham.

Allen AM, Hiscock SJ (2008) Evolution and phylogeny of self-incompatibility systems in angiosperms. In: Self-incompatibility in Flowering Plants: Evolution, Diversity and Mechanisms (Franklin-Tong VE, Ed). Springer-Verlag, Berlin Heidelberg, pp. 73-101.

Anderson MA, Cornish EC, Mau S-L, Williams EG, Hoggart R, Atkinson A, Bonig I, Grego B, Simpson R, Roche PJ, Haley JD, Penschow JD, Niall HD, Tregear GW, Coughlan JP, Crawford RJ, Clarke AE (1986) Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. *Nature* **321:** 38-44.

Anderson MA, McFadden GI, Bernatzky R, Atkinson A, Orpin T, Dedman H, Tregear G, Fernley R, Clarke AE (1989) Sequence variability of three alleles of the self-incompatibility gene of *Nicotiana alata*. *Plant Cell* **1**: 483-491.

Ascher PD (1984) Self-incompatibility. In: *Petunia*: Monographs on Theoretical and Applied Genetics: Volume 9 (Sink KC, Ed). Springer-Verlag, Berlin, pp. 92-109.

Aversano R, Ercolano MR, Frusciante L, Monti L, Bradeen JM, Cristinzio G, Zoina A, Greco N, Vitale S, Carputo D (2007) Resistance traits and AFLP characterization of diploid primitive tuber-bearing potatoes. *Genetic Resources and Crop Evolution* **54**: 1797-1806.

Badr E, Mabrouk Y, Rakha F, Ghazy AH (2008) *Agrobacterium tumefaciens*-mediated transformation of potato and analysis of genomic instability by RAPD. *Research Journal of Agriculture and Biological Sciences* **4**: 16-25.

Barrett SCH (1998) The evolution of mating strategies in flowering plants. *Trends in Plant Science* **3**: 335-341.

Barroso GM, Peixoto AL, Costa CG, Ichaso CLF, Guimaraes EF, Lima HC (1986) Sistematica de Angiospermas do Brasil, vol 3. Editora da Universidade Federal de Vicosa, Vicosa, Brazil.

Baulcombe D (2004) RNA silencing in plants. *Nature* **431**: 356-363.

Bernatzky R, Anderson MA, Clarke AE (1988) Molecular genetics of self-incompatibility in flowering plants. *Developmental Genetics* **9**: 1-12.

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

Bower MS, Matias DD, Fernandes-Carvalho E, Mazzurco M, Gu T, Rothstein SJ, Goring DR (1996) Two members of the thioredoxin-h family interact with the kinase domain of a *Brassica* S locus receptor kinase. *Plant Cell* **8:** 1641-1650.

Bradshaw JE, Bryan GJ, Ramsay G (2006) Genetic resources (including wild and cultivated *Solanum* species) and progress in their utilisation in potato breeding. *Potato Research* **49**: 49-65.

Bradshaw JE, Dale MFB (2005) Potato breeding in the 21st century. In: Aspects of Applied Biology 76, Production and Protection of Sugar Beet and Potatoes (Champion G, Dale MFB, Jaggard K, Parker WE, Pickup J, Stevens M, Eds). Association of Applied Biologists, Warwick, United Kingdom, pp. 109-115.

Braun DM, Stone JM, Walker JC (1997) Interaction of the maize and *Arabidopsis* kinase interaction domains with a subset of receptor-like protein kinases: implications for transmembrane signaling in plants. *Plant Journal* **12:** 83-95.

Bredemeijer GM, Blass J (1981) S-specific proteins in styles of self incompatible *Nicotiana alata. Theoretical and Applied Genetics* **59**: 185-190.

Broothaerts W, Keulemans J, Van Nerum I (2004) Self-fertile apple resulting from S-RNase gene silencing. *Plant Cell Reports* **22**: 497-501.

Broothaerts W, Vanvinckenroye P, Decock B, Van Damme J, Vendrig J (1991) *Petunia hybrida* S-proteins: ribonuclease activity and the role of their glycan side chains in self-incompatibility. *Sexual Plant Reproduction* **4**: 258-266.

Broothaerts WJ, van Laere A, Witters R, Preaux G, Decock B, van Damme J, Vendrig JC (1990) Purification and N-terminal sequencing of style glycoproteins associated with self-incompatibility in *Petunia hybrida*. *Plant Molecular Biology* **14**: 93-102.

Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology* **25**: 169-193.

Cabrillac D, Cock JM, Dumas C, Gaude T (2001) The S-locus receptor kinase is inhibited by thioredoxins and activated by pollen coat proteins. *Nature* **410:** 220-223.

Camadro EL, Carputo D, Peloquin SJ (2004) Substitutes for genome differentiation in tuber-bearing *Solanum*: interspecific pollen-pistil incompatibility, nuclear-cytoplasmic male sterility, and endosperm. *Theoretical and Applied Genetics* **109**: 1369-1376.

Carputo D, Aversano R, Frusciante L (2005) Breeding potato for quality traits. *Acta Horticulturae* **684**: 55-64.

Carputo D, Barone A (2005) Ploidy level manipulations in potato through sexual hybridisation. *Annals of Applied Biology* **146**: 71-79.

Charlesworth D, Awadalla P (1998) Flowering plant self-incompatibility: the molecular population genetics of *Brassica* S-locus. *Heredity* **81:** 1-9.

Chuang CF, Meyerowitz EM (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 4985-4990.

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

Clark DG, Pichersky E, Verdonk J, Dudareva N, Haring M, Klahre U, Schuurink (2009) Benzoids dominate the fragrance of *Petunia* flowers. In: *Petunia*: Evolutionary, Developmental and Physiological Genetics (Gerats T, Strommer J, Eds). Springer Science+Business Media, Spring Street, New York, USA, pp. 51-69.

Clark KR, Okuley JJ, Collins PD, Sims TL (1990) Sequence variability and developmental expression of S-alleles in self-incompatible and pseudo-self- compatible *Petunia*. *The Plant Cell* **2**: 815-826.

Cornish EC, Pettitt JM, Bonig I, Clarke AE (1987) Developmentally controlled expression of a gene associated with self-incompatibility in *Nicotiana alata*. *Nature* **326**: 99-102.

Cruz-Garcia F, Hancock CN, Kim D, McClure B (2005) Stylar glycoproteins bind to S-RNase *in vitro*. *The Plant Journal* **42:** 295-304.

Czyzyk R (2010) Role of the *SLF* gene in self-compatibility in *Petunia hybrida*. PhD thesis at University of Nottingham.

Dale MFB, Mackay GR (1994) Inheritance of table and processing quality. In: Potato Genetics (Bradshaw JE, Mackay GR, Eds). CAB International, Wallington, Oxford, United Kingdom, pp. 285-315.

Dana MN, Ascher PD (1985) Pseudo-self compatibility (PSC) in *Petunia integrifolia*. *Journal of Heredity* **76**: 468-470.

Darwin C (1891) The effects of cross and self fertilisation in the vegetable kingdom. John Murray, 3rd Ed, London, pp. 188-189.

Davies HV (1998) Prospects for manipulating carbohydrate metabolism in potato tuber. *Aspects of Applied Biology* **52**: 245-254.

de Nettancourt D (1977) Incompatibility in angiosperms. Springer, Berlin Heidelberg New York.

de Nettancourt D (1997) Incompatibility in angiosperms. *Sexual Plant Reproduction* **10**: 185-199.

Dean JD, Goodwin PH, Hsiang T (2002) Comparison of relative RT-PCR and northern blot analyses to measure expression of β-1,3-glucanase in *Nicotiana benthamiana* infected with *Colletotrichum destructivum*. *Plant Molecular Biology Reporter* **20**: 347-356.

Denli AM, Hannon GJ (2003) RNAi: an ever-growing puzzle. *Trends in Biochemical Sciences* **28**: 196-201.

Despres C, Saba-El-Leil MK, Rivard SR, Morse D, Cappadocia M (1994) Molecular cloning of two *Solanum chacoense* S-alleles and a hypothesis concerning their evolution. *Sexual Plant Reproduction* **7**: 169-176.

Dickinson H (1995) Dry stigmas, water and self-incompatibility in *Brassica*. *Sexual Plant Reproduction* **8:** 1-10.

Dodds PN, Clarke AE, Newbigin E (1996) A molecular perspective on pollination in flowering plants. *Cell* **85**: 141-144.

Dodds PN, Clarke AE, Newbigin E (1996) Molecular characterisation of an S-like RNase of *Nicotiana alata* that is induced by phosphate starvation. *Plant Molecular Biology* **31**: 227-238.

Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, Pikaard CS (2006) Gateway-compatible vectors for plant functional genomics and proteomics. *The Plant Journal* **45**: 616-629.

Ebert PR, Anderson MA, Bernatzky R, Altschuler M, Clarke AE (1989) Genetic polymorphism of self-incompatibility in plants. *Cell* **56**: 255-262.

Emerson S (1939) A preliminary survey of the *Oenothera organensis* population. *Genetics* **24**: 524-537.

Entani T, Iwano M, Shiba H, Che F-S, Isogai A, Takayama S (2003) Comparative analysis of the self-incompatibility (*S*-) locus region of *Prunus mume*: identification of a pollen-expressed F-box gene with allelic diversity. *Genes to Cells* **8**: 203-213.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783-791.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**: 806-811.

Foote HC, Ride JP, Franklin-Tong VE, Walker EA, Lawrence MJ, Franklin FC (1994) Cloning and expression of a distinctive class of self-incompatibility (S) gene from *Papaver rhoeas* L. *Proceedings of the National Academy of Sciences of the United States of America* **91:** 2265-2269.

Franklin FCH, Lawrence MJ, Franklin-Tong VE (1995) Cell and molecular biology of self-incompatibility in flowering plants. *International Review of Cytology* **158**: 1-64.

Franklin-Tong V, Franklin C (2007) Self-incompatibility. *Encyclopedia of Life Sciences*. Advanced article, doi: 10.1002/9780470015902.a0002041.pub2.

Franklin-Tong VE, Franklin FCH (2003) Gametophytic self-incompatibility inhibits pollen tube growth using different mechanisms. *Trends in Plant Science* **12**: 598-605.

Franklin-Tong VE, Holdaway-Clarke TL, Straatman KR, Kunkel JG, Hepler PK (2002) Involvement of extracellular calcium influx in the self-incompatibility response of *Papaver rhoeas*. *The Plant Journal* **29:** 333-345.

Fu C, Wehr DR, Edwards J, Hauge B (2008) Rapid one-step recombinational cloning. *Nucleic Acids Research* **36**: e54.

Gaude T, Fobis-Loisy I, Miege C (2006) Control of fertilization by self-incompatibility mechanisms. In: The Molecular Biology and Biotechnology of Flowering (Jordan BR, Ed) CAB International, Wallington, Oxford, United Kingdom, pp. 269-297.

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

Gerats T, Vandenbussche M (2005) A model system for comparative research: *Petunia. Trends in Plant Science* **10**: 251-256.

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

Golz JF, Oh YH, Su V, Kusaba M, Newbigin E (2001) Genetic analysis of *Nicotiana* pollen-part mutants is consistent with the presence of an S-ribonuclease inhibitor at the S-locus. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 15372-15376.

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

Good-Avila SV, Mena-Ali JI, Stephenson AG (2008) Genetic and environmental causes and evolutionary consequences of variations in self-fertility in self-incompatible species. In: Self-incompatibility in Flowering Plants: Evolution, Diversity and Mechanisms (Franklin-Tong VE, Ed). Springer-Verlag Berlin Heidelberg, pp. 33-51.

Goring DR (2000) The search for components of the self-incompatibility signalling pathway(s) in *Brassica napus*. *Annals of Botany* **85:** 171-179.

Gu T, Mazzurco M, Sulaman W, Matias DD, Goring DR (1998) Binding of an arm repeat protein to the kinase domain of the S-locus receptor kinase. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 382-387.

Haffani YZ, Gaude T, Cock JM, Goring DR (2004) Antisense suppression of thioredoxin-h mRNA in *Brassica napus* cv. Westar pistils causes a low level constitutive pollen rejection response. *Plant Molecular Biology* **55:** 619-630.

Hancock CN, Kent L, McClure BA (2005) The stylar 120kDa glycoprotein is required for S-specific pollen rejection in *Nicotiana*. *The Plant Journal* **43:** 716-723.

Hannon GJ (2002) RNA interference. *Nature* **418**: 244-251.

Harbord RM, Napoli CA, Robbins TP (2000) Segregation of T-DNA markers linked to the self-incompatibility (S) locus in *Petunia hybrida*. *Genetics* **154**: 1323-1333.

Harland SC, Atteck OS (1933) The inheritance of self-fertility in *Petunia violacea*. *Genetica* **15**: 89-102.

Hartley JL, Temple GF, Brasch MA (2000) DNA cloning using *in vitro* site-specific recombination. *Genome Research* **10**: 1788-1795.

Hatakeyama K, Takasaki T, Suzuki G, Nishio T, Watanabe M, Isogai A, Hinata K (2001) The S receptor kinase gene determines dominance relationships in stigma expression of self-incompatibility in *Brassica*. *Plant Journal* **26:** 69-76.

Hawkes JG (1990) The potato, evolution, biodiversity and genetic resources. Belhaven Press, London.

Hawkes JG (1994) Origins of cultivated potatoes and species relationships. In: Potato Genetics (Bradshaw JE, Mackay GR, Eds). CAB International, Wallington, Oxford, United Kingdom, pp. 3-42.

Hawkes JG, Francisco-Ortega J (1993) The early history of the potato in Europe. *Euphytica* **70**: 1-7.

Hayes RJ, Thill CA (2002) Co-current introgression of economically important traits in a potato-breeding program. *American Journal of Potato Research* **79**: 173-181.

Helliwell C, Waterhouse P (2003) Constructs and methods of high-throughput gene silencing in plants. *Methods* **30**: 289-295.

Helliwell C, Waterhouse P (2005) Constructs and methods for hairpin RNA-mediated gene silencing in plants. *Methods in Enzymology* **392**: 24-35.

Hijmans RJ, Spooner DM (2001) Geographic distribution of wild potato species. *American Journal of Botany* **88**: 2101-2112.

Hiscock SJ (2000) Genetic control of self-incompatibility in *Senecio squalius* L. (Asteraceae): a successful colonizing species. *Heredity* **85**: 10-19.

Hiscock SJ, McInnis SM (2003) The diversity of self-incompatibility systems in flowering plants. *Plant Biology* **5:** 23-32.

Hiscock SJ, Tabah DA (2003) The different mechanisms of sporophytic self-incompatibility. *Philosophical Transactions of the Royal Society of London. Series* B **358**: 1037-1045.

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

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

Hua Z, Fields A, Kao TH (2008) Biochemical models for S-RNase-based self-incompatibility. *Molecular Plant* **1**: 575-585.

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

Hua Z, Kao TH (2008) Identification of major lysine residues of $S_3$-RNase of *Petunia inflata* involved in ubiquitin-26S proteasome-mediated degradation *in vitro*. *The Plant Journal* **54**: 1094-1104.

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

Huaman Z, Spooner DM (2002) *American Journal of Botany* **89**: 947-965.

Huang J, Zhao L, Yang Q, Xue Y (2006) AhSSK1 a novel SKP1-like protein that interacts with the S-locus F-box protein SLF. *Plant Journal* **46:** 780-793.

Hugot K, Ponchet M, Marais A, Ricci P, Galiana E (2002) A tobacco S-like RNase inhibits hyphal elongation of plant pathogens. *Molecular Plant - Microbe Interactions* **15**: 243-250.

Hunter P (2009) Me, myself and I: the genetics and molecular biology behind self-incompatibility and the avoidance of inbreeding in plants. *European Molecular Biology Organization* **10:** 1297-1300.

Ida K, Norioka S, Yamamoto M, Kumasaka T, Yamashita E, Newbigin E, Clarke AE, Sakiyama F, Sato M (2001) The 1.55 A resolution structure of *Nicotiana alata* $S_{F11}$-RNase associated with gametophytic self-incompatibility. *Journal of Molecular Biology* **314:** 103-112.

Igic B, Bohs L, Kohn JR (2003) Historical inferences from the self-incompatibility locus. *New Phytologist* **161:** 97-105.

Igic B, Kohn JR (2001) Evolutionary relationships among self-incompatibility RNases. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 13167-13171.

Ioerger TR, Clark AG, Kao TH (1990) Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proceedings of the National Academy of Sciences of the United States of America* **87**: 9732-9735.

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

Ishimizu T, Norioka S, Kanai M, Clarke AE, Sakiyama F (1996) Location of cysteine and cystine residues in S-ribonucleases associated with gametophytic self-incompatibility. *European Journal of Biochemistry* **242**: 627-635.

Ishimizu T, Shinkawa T, Sakiyama F, Norioka S (1998) Primary structural features of rosaceous S-RNases associated with gametophytic self-incompatibility. *Plant Molecular Biology* **37**: 931-941.

Jacobsen E, Ramanna MS (1994) Production of monohaploids of *Solanum tuberosum* L. and their use in genetics, molecular biology and breeding. In: Potato Genetics (Bradshaw JE, Mackay GR, Eds). CAB International, Wallington, Oxford, United Kingdom, pp. 155-170.

Jahnen W, Lush WM, Clarke AE (1989) Inhibition of *in vitro* pollen tube growth by isolated S-glycoproteins in *Nicotiana alata*. *Plant Cell* **1**: 501-510.

Kao TH, McCubbin AG (1996) How flowering plants discriminate between self and non-self pollen to prevent inbreeding. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 12059-12065.

Kao TH, Tsukamoto T (2004) The molecular and genetic bases of S-RNase-based self-incompatibility. *Plant Cell* **16**: S72-S83.

Kaothien-Nakayama P, Isogai A, Takayama S (2010) Self-incompatibility systems in flowering plants. In: Plant Developmental Biology-Biotechnological Perspectives: Volume 1 (Pua EC, Davey MR, Eds) Springer-Verlag, Berlin Heidelberg, pp. 459-485.

Karimi M, Depicker A, Hilson P (2007) Recombinational cloning with plant Gateway vectors. *Plant Physiology* **145**: 1144-1154.

Karunanandaa B, Huang S, Kao T (1994) Carbohydrate moiety of the *Petunia inflata* $S_3$ protein is not required for self-incompatibility interactions between pollen and pistil. *The Plant Cell* **6**: 1933-1940.

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

Kawata Y, Sakiyama F, Tamaoki H (1988) Amino-acid sequence of ribonuclease $T_2$ from *Aspergillus oryzae*. *European Journal of Biochemistry* **176**: 683-697.

Kear PJ (2010) How did flowering plants learn to avoid blind date mistakes? Self-incompatibility in plants and comparisons with non-self rejection in the immune response. Unpublished.

Kear PJ, Malinski LT (2010) Identification of S-RNases in self and interspecific crosses: can S-RNase identity be used to predict crossability in potato? Unpublished.

Kirch HH, Uhrig H, Lottspeich F, Salamini F, Thompson RD (1989) Characterization of proteins with self-incompatibility in *Solanum tuberosum*. *Theoretical and Applied Genetics* **78**: 581-588.

Kondo K, Yamamoto M, Itahashi R, Sato T, Egashira H, Hattori T, Kowyama Y (2002) Insights into the evolution of self-compatibility in *Lycopersicon* from a study of stylar factors. *The Plant Journal* **30**: 143-153.

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

Kuipers AGJ, Jacobsen E, Visser RGF (1997) Applications of antisense technology in plants. In: Antisense Technology: A Practical Approach (Lichtenstein C, Nellen W, Eds). Oxford University Press, Oxford, pp. 191-219.

Kusaba M (2004) RNA interference in crop plants. *Current Opinion in Biotechnology* **15**: 139-143.

Kusaba M, Tung CW, Nasrallah ME, Nasrallah JB (2002) Monoallelic expression and dominance interactions in anthers of self-incompatible *Arabidopsis lyrata*. *Plant Physiology* **128:** 17-20.

Lai Z, Ma W, Han B, Liang L, Zhang Y, Hong G, Xue Y (2002) An F-box gene linked to the self-incompatibility (S) locus of *Antirrhinum* is expressed specifically in pollen and tapetum. *Plant Molecular Biology* **50**: 29-42.

Lang J (2001) Notes of a potato watcher. Texas A and M University Press, College Station, Texas.

Lee HS, Huang S, Kao TH (1994) S proteins control rejection of incompatible pollen in *Petunia inflata*. *Nature* **367**: 560-563.

Levin DA (1996) The evolutionary significance of pseudo self-fertility. *The American Naturalist* **148**: 321-332.

Lind JL, Bacic A, Clarke AE, Anderson MA (1994) A style-specific hydroxyproline-rich glycoprotein with properties of both extensins and arabinogalactan proteins. *Plant Journal* **6:** 491-502.

Lind JL, Bonig I, Clarke AE, Anderson MA (1996) A style-specific 120kDa glycoprotein enters pollen tubes of *Nicotiana alata in vivo*. *Sexual Plant Reproduction* **9:** 75-86.

Linskens HF (1975) Incompatibility in *Petunia*. *Proceedings of the Royal Society of London. Series B* **188**: 299-311.

Luu DT, Qin X, Morse D, Cappadocia M (2000) S-RNase uptake by compatible pollen tubes in gametophytic self-incompatibility. *Nature* **407**: 649-651.

MacIntosh GC, Hillwig MS, Meyer A, Flagel L (2010) RNase $T_2$ genes from rice and the evolution of secretory ribonucleases in plants. *Molecular Genetics and Genomics* **283**: 381-396.

Marcellan ON, Acevedo A, Camadro EL (2006) $S_{16}$, a novel S-RNase allele in the diploid species *Solanum chacoense*. *Genome* **49**: 1052-1054.

Mascia T, Santovito E, Gallitelli D, Cillo F (2010) Evaluation of reference genes for quantitative reverse-transcription polymerase chain reaction normalization in infected tomato plants. *Molecular Plant Pathology* **11**: 805-816.

Mather K (1943) Specific differences in *Petunia.* I. Incompatibility. *Journal of Genetics* **45**: 215-235.

Matsuura T, Sakai H, Unno M, Ida K, Sato M, Sakiyama F, Norioka S (2001) Crystal structure at 1.5-A resolution of *Pyrus pyrifolia* pistil ribonuclease responsible for gametophytic self-incompatibility. *Journal of Biological Chemistry* **276:** 45261-45269.

Matton DP, Maes O, Laublin G, Xike Q, Bertrand C, Morse D, Cappadocia M (1997) Hypervariable domains of self-incompatibility RNases mediate allele-specific pollen recognition. *Plant Cell* **2:** 1757-1766.

Mazzurco M, Sulaman W, Elina H, Cock JM, Goring DR (2001) Further analysis of the interactions between the *Brassica* S receptor kinase and three interacting proteins (ARC1, THL1 and THL2) in the yeast two-hybrid system. *Plant Molecular Biology* **45:** 365-376.

McClure B, Cruz-Garcia F, Romero C (2011) Compatibility and incompatibility in S-RNase-based systems. *Annals of Botany* **108**: 647-658.

McClure B, Mou B, Canevascini S, Bernatzky R (1999) A small asparagine-rich protein required for S-allele-specific pollen rejection in *Nicotiana*. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 13548-13553.

McClure BA (2004) S-RNase and SLF determine S-haplotype-specific pollen recognition and rejection. *Plant Cell* **16**: 2840-2847.

McClure BA, Franklin-Tong V (2006) Gametophytic self-incompatibility: understanding the cellular mechanisms involved in "self" pollen tube inhibition. *Planta* **224**: 233-245.

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

McCormick S (1998) Self-incompatibility and other pollen-pistil interactions. *Current Opinion in Plant Biology* **1**: 18-25.

McCubbin A (2008) Heteromorphic self-incompatibility in *Primula*: twenty-first century tools promise to unravel a classic nineteenth century model system. In: Self-incompatibility in Flowering Plants: Evolution, Diversity and Mechanisms (Franklin-Tong VE, Ed). Springer-Verlag, Berlin Heidelberg, pp. 289-308.

McCubbin AG, Kao TH (2000) Molecular recognition and response in pollen and pistil interactions. *Annual Review of Cell and Developmental Biology* **16**: 333-364.

McManus MT, Sharp PA (2002) Gene silencing in mammals by small interfering RNAs. *Nature Reviews Genetics* **3**: 737-747.

Meister G, Tuschl T (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**: 343-349.

Mello CC, Conte D (2004) Revealing the world of RNA interference. *Nature* **431**: 338-342.

Minamikawa M, Kakui H, Wang S, Kotoda N, Kikuchi S, Koba T, Sassa H (2010) Apple S-locus region represents a large cluster of related, polymorphic and pollen-specific F-box genes. *Plant Molecular Biology* **74**: 143-154.

Murase K, Shiba H, Iwano M, Che FS, Watanabe M, Isogai A, Takayama S (2004) A membrane anchored protein kinase involved in *Brassica* self-incompatibility signaling. *Science* **303:** 1516-1519.

Murfett J, Atherton TL, Mou B, Gasser CS, McClure BA (1994) S-RNase expressed in transgenic *Nicotiana* causes *S*-allele-specific pollen rejection. *Nature* **367**: 563-566.

Murfett J, McClure BA (1998) Expressing foreign genes in the pistil: a comparison of S-RNase constructs in different *Nicotiana* backgrounds. *Plant Molecular Biology* **37**: 561-569.

Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible co-suppression of homologous genes *in trans*. *Plant Cell* **2**: 279-289.

Nasrallah JB, Nasrallah ME (1993) Pollen-stigma signalling in the sporophytic self-incompatibility response. *Plant Cell* **5:** 1325-1335.

Newbigin E, Paape T, Kohn JR (2008) RNase-based self-incompatibility: puzzled by pollen S. *The Plant Cell* **20**: 2286-2292.

Nicot N, Hausman JF, Hoffmann L, Evers D (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany* **56**: 2907-2914.

Novina CD, Sharp PA (2004) The RNAi revolution. *Nature* **430**: 161-164.

Nowak MD, Davis AP, Anthony F, Yoder AD (2011) Expression and trans-specific polymorphism of self-incompatibility RNases in *Coffea* (Rubiaceae). *PLoS ONE* **6**: e21019. doi:10.1371/journal.pone.0021019.

O'Brien M, Kapfer C, Major G, Laurin M, Bertrand C, Kondo K, Kowyama Y, Matton DP (2002) Molecular analysis of the stylar-expressed *Solanum chacoense* small asparagine-rich protein family related to the HT modifier of gametophytic self-incompatibility in *Nicotiana*. *Plant Journal* **32:** 985-996.

O'Brien M, Major G, Chanth SC, Matton DP (2004) Isolation of S-RNase binding proteins from *Solanum chacoense*: identification of an SBP1 (RING finger protein) orthologue. *Sexual Plant Reproduction* **17:** 81-87.

Ortiz R, Peloquin SJ (1994) Use of 24-chromosome potatoes (diploids and dihaploids) for genetical analysis. In: Potato Genetics (Bradshaw JE,

Mackay GR, Eds). CAB International, Wallington, Oxford, United Kingdom, pp. 133-154.

Oxley D, Munro SL, Craik DJ, Bacic A (1998) Structure and distribution of N-glycans on the $S_7$-allele stylar self-incompatibility ribonuclease of *Nicotiana alata*. *Journal of Biochemistry* **123**: 978-983.

Pandey KK (1962) Interspecific incompatibility in *Solanum* species. *American Journal of Botany* **49**: 874-882.

Pandey SK, Kaushik SK (2003) Origin, evolution, history and spread of potato. In: The Potato Production and Utilization in Sub-tropics (Khurana SMP, Minhas JS, Pandey SK, Eds) Mehta Publishers, New Delhi, pp. 15-24.

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

Puerta AR, Takagawa M, Sassa H (2009b) Pistil non-S-specific self-incompatibility factor *HT-B* is tightly linked to the *HT*-like gene *HTL* in *Petunia*. *Breeding Science* **59**: 99-101.

Puerta AR, Ushijima K, Koba T, Sassa H (2009a) Identification and functional analysis of pistil self-incompatibility factor *HT-B* of *Petunia*. *Journal of Experimental Botany* **60**: 1309-1318.

Pushkarnath (1942) Studies on sterility in potatoes. 1. The genetics of self- and cross-incompatibilities. *Indian Journal of Genetics and Plant Breeding* **2**: 11-36.

Qiao H, Wang H, Zhao L, Zhou J, Huang J, Zhang Y, Xue Y (2004) The F-box protein AhSLF-S2 physically interacts with S-RNases that may be inhibited by the ubiquitin/26S proteasome pathway of protein degradation during compatible pollination in *Antirrhinum*. *Plant Cell* **16:** 582-595.

Qin X, Liu B, Soulard J, Morse D, Cappadocia M (2006) Style-by-style analysis of two sporadic self-compatible *Solanum chacoense* lines supports a primary role for S-RNases in determining pollen rejection thresholds. *Journal of Experimental Botany* **57**: 2001-2013.

Qin X, Luu D, Yang Q, Maes O, Matton D, Morse D, Cappadocia M (2001) Genotype-dependent differences in $S_{12}$-RNase expression lead to sporadic self-compatibility in *Solanum chacoense*. *Plant Molecular Biology* **45:** 295-305.

Radonic A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A (2004) Guideline to reference gene selection for quantitative real-time PCR. *Biochemical and Biophysical Research Communications* **313**: 856-862.

Rea AC, Nasrallah JB (2008) Self-incompatibility systems: barriers to self-fertilization in flowering plants. *International Journal of Developmental Biology* **52:** 627-636.

Richman AD, Kao TH, Schaeffer SW, Uyenoyama MK (1995) S-allele sequence diversity in natural populations of *Solanum carolinense* (Horsenettle). *Heredity* **75:** 405-415.

Richman AD, Kohn JR (2000) Evolutionary genetics of self-incompatibility in the Solanaceae. *Plant Molecular Biology* **42**: 169-179.

Richman AD, Uyenoyama MK, Kohn JR (1996) Allelic diversity and gene genealogy at the self-incompatibility locus in the Solanaceae. *Science* **273:** 1212-1216.

Rivard SR, Cappadocia M, Landry BS (1996) A comparison of RFLP maps based on anther culture derived, selfed, and hybrid progenies of *Solanum chacoense*. *Genome* **39**: 611-621.

Roalson EH, McCubbin AG (2003) S-RNases and sexual incompatibility: structure, functions, and evolutionary perspectives. *Molecular Phylogenetics and Evolution* **29**: 490-506.

Robacker CD, Ascher PD (1982) Discriminating styles (DS) and pollen-mediated pseudo self-incompatibility (PMPSC) in *Nemesia strumosa* benth. *Theoretical and Applied Genetics* **61**: 289-296.

Robbins TP, Harbord RM, Sonneveld T, Clarke K (2000) The molecular genetics of self-incompatibility in *Petunia hybrida*. *Annals of Botany* **85**: 105-112.

Roldan JA, Quiroga R, Goldraij A (2010) Molecular and genetic characterization of novel S-RNases from a natural population of *Nicotiana alata*. *Plant Cell Reports* **29**: 735-746.

Royo J, Kunz C, Kowyama Y, Anderson M, Clarke AE, Newbigin E (1994) Loss of a histidine residue at the active site of S-locus ribonuclease is associated with self-compatibility in *Lycopersicon peruvianum*. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 6511-6514.

Ruan W, Lai M (2007) Actin, a reliable marker of internal control? *Clinica Chimica Acta* **385**: 1-5.

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

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**: 406-425.

Sassa H, Kakui H, Miyamoto M, Suzuki Y, Hanada T, Ushijima K, Kusaba M, Hirano H, Koba T (2007) S locus F-Box brothers: multiple and pollen-specific F-Box genes with S haplotype-specific polymorphisms in apple and Japanese pear. *Genetics* **175**: 1869-1881.

Schmittgen T, Zakrajsek BA (2000) Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. *Journal of Biochemical and Biophysical Methods* **46**: 69-81.

Schopfer CR, Nasrallah ME, Nasrallah JB (1999) The male determinant of self-incompatibility in *Brassica*. *Science* **266**: 1697-1700.

Serio FD, Schob H, Iglesias A, Tarina C, Bouldoires E, Meins FJ (2001) Sense- and antisense-mediated gene silencing in tobacco is inhibited by the same viral suppressors and is associated with accumulation of small RNAs. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 6506-6510.

Sharma KK, Bhatnagar-Mathur P, Thorpe TA (2005) Genetic transformation technology: status and problems. *In Vitro Cellular and Developmental Biology-Plant* **41**: 102-112.

Shiba H, Iwano M, Entani T, Ishimoto K, Shimosato H, Che F-S, Satta Y, Ito A, Takada Y, Watanabe M, Isogai A, Takayama S (2002) The dominance of alleles controlling self incompatibility in *Brassica* pollen is regulated at the RNA level. *Plant Cell* **14:** 491-504.

Sijacic P, Wang X, Skirpan AL, Wang Y, Dowd PE, McCubbin AG, Huang S, Kao TH (2004) Identification of the pollen determinant of S-RNase-mediated self-incompatibility. *Nature* **429**: 302-305.

Silva NF, Goring DR (2001) Mechanisms of self-incompatibility in flowering plants. *Cellular and Molecular Life Science* **58**: 1988-2007.

Silva NF, Stone SL, Christie LN, Sulaman W, Nazarian KAP, Burnett LA, Arnoldo MA, Rothstein SJ, Goring DR (2001) Expression of the *S* receptor kinase in self-compatible *Brassica napus* cv. Westar leads to the allele-specific rejection of self-incompatible *Brassica napus* pollen. *Molecular Genetics and Genomics* **265**: 552-59.

Sims TL (2005) Pollen recognition and rejection in different self-incompatibility systems. *Recent Research Development Plant Molecular Biology* **2:** 31-62.

Sims TL, Ordanic M (2001) Identification of a S-ribonuclease-binding protein in *Petunia hybrida*. *Plant Molecular Biology* **47:** 771-783.

Sims TL, Robbins TP (2009) Gametophytic self-incompatibility in *Petunia*. In: *Petunia*: Evolutionary, Developmental and Physiological Genetics (Gerats T, Strommer J, Eds). Springer Science+Business Media, Spring Street, New York, USA, pp. 85-106.

Sinsheimer RL (1977) Recombinant DNA. *Annual Review of Biochemistry* **46**: 415-38.

Smith NA, Singh SP, Wang MB, Stoutjesdijk PA, Green AG, Waterhouse PM (2000) Total silencing by intron-spliced hairpin RNAs. *Nature* **407**: 319-320.

Sonneveld T, Tobutt KR, Vaughan SP, Robbins TP (2005) Loss of pollen-*S* function in two self-compatible selections of *Prunus avium* is associated with deletion/mutation of an *S* haplotype-specfic F-box gene. *Plant Cell* **17**: 37-51.

Sonnewald U, Hajirezaei MR, Biemelt S, Mueller M (2003) Designer tubers for production of novel compounds. In: Congress Proceedings, BCPC International Congress: Crop Science and Technology (Bracknell UK, Ed). Glasgow, UK, pp. 123-132.

Spooner DM, Nunez J, Trujillo G, Herrera MR, Guzman F, Ghislain M (2007) Extensive simple sequence repeat genotyping of potato landraces supports a major re-evaluation of their gene pool structure and classification. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 14694-14699.

Stehmann JR, Lorenz-Lemke AP, Freitas LB, Semir J (2009) The genus *Petunia*. In: *Petunia*: Evolutionary, Developmental and Physiological Genetics (Gerats T, Strommer J, Eds). Springer Science+Business Media, Spring Street, New York, USA, pp. 1-28.

Stein JC, Howlett B, Boyes DC, Nasrallah ME, Nasrallah JB (1991) Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleracea*. *Proceedings of the National Academy of Sciences of the United States of America* **88**: 8816-8820.

Stone JL (2002) Molecular mechanisms underlying the breakdown of gametophytic self-incompatibility. *The Quarterly Review of Biology* **77**: 17-32.

Stone JL, Pierce SE (2005) Rapid recent radiation of S-RNase lineages in *Witheringia solanacea* (Solanaceae). *Heredity* **94**: 547-555.

Stone SL, Anderson EM, Mullen RT, Goring DR (2003) ARC1 is an E3 ubiquitin ligase and promotes the ubiquitination of proteins during the rejection of self-incompatible *Brassica* pollen. *Plant Cell* **15:** 885-898.

Stone SL, Arnoldo M, Goring DR (1999) A breakdown of *Brassica* self-incompatibility in ARC1 antisense transgenic plants. *Science* **286:** 1729-1731.

Sturzenbaum SR, Kille P (2001) Control genes in quantitative molecular biological techniques: the variability of invariance. *Comparative Biochemistry and Physiology Part B* **130**: 281-289.

Suzuki G, Kai N, Hirose T, Fukui K, Nishio T, Takayama S, Isogai A, Watanabe M, Hinata K (1999) Genomic organization of the *S* locus: identification and characterization of genes in *SLG/SRK* region of $S_9$ haplotype of *Brassica campestris* (syn. *rapa*). *Genetics* **153**: 391-400.

Tai GCC (1994) Use of 2n gametes. In: Potato Genetics (Bradshaw JE, Mackay GR, Eds). CAB International, Wallington, Oxford, United Kingdom, pp. 109-132.

Takasaki T, Hatakeyama K, Suzuki G, Watanabe M, Isogai A, Hinata K (2000) The *S* receptor kinase determines self-incompatibility in *Brassica* stigma. *Nature* **403**: 913-16.

Takayama S, Isogai A (2005) Self-incompatibility in flowering plants. *Annual Review of Plant Biology* **56**: 467-489.

Takayama S, Shiba H, Iwano M, Shimosato H, Che F-S, Kai N, Watanabe M, Suzuki G, Hinata K, Isogai A (2000) The pollen determinant of self

incompatibility in *Brassica campestris*. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 1920-1925.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* (submitted).

Tarutani Y, Shiba H, Iwano M, Kakizaki T, Suzuki G, Watanabe M, Isogai A, Takayama S (2010) Trans-acting small RNA determines dominance relationships in *Brassica* self-incompatibility. *Nature* **466**: 983-987.

Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, Grisar T, Igout A, Heinen E (1999) Housekeeping genes as internal standards: use and limits. *Journal of Biotechnology* **75**: 291-295.

Thompson J, Higgins D, Gibson T, Thompson J D, Higgins DG, Gibson T J (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673-4680.

Tsukamoto T, Ando T, Kokobun H, Watanabe H, Masada M, Zhu X, Marchesi E, Kao TH (1999) Breakdown of self-incompatibility in a natural population of *Petunia axillaris* (Solanaceae) in Uruguay containing both self-incompatible and self-compatible plants. *Sexual Plant Reproduction* **12:** 6-13.

Tsukamoto T, Ando T, Kokubun H, Watanabe H, Sato T, Masada M, Marchesi E, Kao TH (2003b) Breakdown of self-incompatibility in a natural population of *Petunia axillaris* caused by a modifier locus that suppresses the expression of an S-RNase gene. *Sexual Plant Reproduction* **15**: 255-263.

Tsukamoto T, Ando T, Kokubun H, Watanabe H, Tanaka R, Hashimoto G, Marchesi E, Kao TH (1998) Differentiation in the status of self-incompatibility among all natural taxa of *Petunia* (Solanaceae). *Acta Phytotaxonomica et Geobotanica*. **49**: 115-33.

Tsukamoto T, Ando T, Takahashi K, Omori T, Watanabe H, Kokubun H, Marchesi E, Kao TH (2003a) Breakdown of self-incompatibility in a natural population of *Petunia axillaris* caused by loss of pollen function. *Plant Physiology* **131**: 1903-1912.

Tumusiime S (2006) Self-incompatibility in *Petunia hybrida* and *Coffea canephora*. PhD thesis at University of Nottingham.

Ushijima K, Sassa H, Dandekar AM, Gradziel TM, Tao R, Hirano H (2003) Structural and transcriptional analysis of the self-incompatibility locus of almond: identification of a pollen-expressed F-Box gene with haplotype-specific polymorphism. *The Plant Cell* **15**: 771-781.

Ushijima K, Yamane H, Watari A, Kakehi E, Ikeda K, Hauck NR, Iezzoni AF, Tao R (2004) The *S* haplotype-specific F-box protein gene, *SFB*, is defective in self-compatible haplotypes of *Prunus avium* and *P. mume*. *Plant Journal* **39**: 573-586.

Uyenoyama MK (2000) Evolutionary dynamics of self-incompatibility alleles in *Brassica*. *Genetics* **156:** 351-359.

Vanoosthuyse V, Tichtinsky G, Dumas C, Gaude T, Cock JM (2003) Interaction of calmodulin, a sorting nexin and kinase-associated protein phosphatase with the *Brassica oleracea* S locus receptor kinase. *Plant Physiology* **133:** 919-929.

Veluthambi K, Gupta AK, Sharma A (2003) The current status of plant transformation technologies. *Current Science* **84**: 368-380.

Verica JA, McCubbin AG, Kao TH (1998) Are the hypervariable regions of S-RNases sufficient for allele-specific recognition of pollen? *Plant Cell* **10:** 314-317.

Wang X, Hughes AL, Tsukamoto T, Ando T, Kao TH (2001) Evidence that intragenic recombination contributes to allelic diversity of the S-RNase gene at the self-incompatibility (S) locus in *Petunia inflata*. *Plant Physiology* **125**: 1012-1022.

Warrington JA, Nair A, Mahadevappa M, Tsyganskaya M (2000) Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. *Physiological Genomics* **2**: 143-147.

Waterhouse PM, Graham MW, Wang MB (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 13959-13964.

Waterhouse PM, Helliwell CA (2003) Exploring plant genomes by RNA-induced gene silencing. *Nature Reviews Genetics* **4**: 29-38.

Waterhouse PM, Wang MB, Lough T (2001) Gene silencing as an adaptive defence against viruses. *Nature* **411**: 834-842.

Watson JM, Fusaro AF, Wang M, Waterhouse PM (2005) RNA silencing platforms in plants. *Federation of European Biochemical Societies Letters* **579**: 5982-5987.

Wesley SV, Helliwell CA, Smith NA (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant Journal* **27**: 581-590.

Wheeler D, Newbigin E (2007) Expression of 10 S-Class SLF-like genes in *Nicotiana alata* pollen and its implications for understanding the pollen factor of the S locus. *Genetics* **177**: 2171-2180.

Wheeler MJ, de Graaf BHJ, Hadjiosif N, Perry RM, Poulter NS, Osman K, Vatovec S, Harper A, Franklin FCH, Franklin-Tong VE (2009) Identification of the pollen self-incompatibility determinant in *Papaver rhoeas*. *Nature* **459**: 992-995.

Wheeler MJ, Vatovec S, Franklin-Tong VE (2010) The pollen S-determinant in *Papaver*: comparisons with known plant receptors and protein ligand partners. *Journal of Experimental Botany* **61:** 2015-2025.

Wheeler VA, Evans NE, Fougler D, Webb KJ, Karp A, Franklin J, Bright SW (1985) Shoot formation explant cultures of fourteen potato cultivars and studies of the cytology and morphology of regenerated plants. *Annals of Botany* **55**: 321-329.

Wolters A-MA, Visser RGF (2000) Gene silencing in potato: allelic differences and effect of ploidy. *Plant Molecular Biology* **43**: 377-386.

Wright PD (2004) The molecular basis of self-compatibility in *Petunia hybrida*. PhD thesis at University of Nottingham.

Wright S (1939) The distribution of self-sterility alleles in populations. *Genetics* **24**: 538-552.

Xu B, Mu J, Nevins DL, Grun P, Kao TH (1990) Cloning and sequencing of cDNAs encoding two self-incompatibility associated proteins in *Solanum chacoense*. *Molecular and General Genetics* **224**: 341-346.

Zhang Y, Xue Y (2008) Molecular biology of S-RNase-based self-incompatibility. In: Self-incompatibility in Flowering Plants: Evolution, Diversity and Mechanisms (Franklin-Tong VE, Ed). Springer-Verlag, Berlin Heidelberg, pp. 193-215.

Zhou J, Wang F, Ma W, Zhang Y, Han B, Xue Y (2003) Structural and transcriptional analysis of S-locus F-box genes in *Antirrhinum*. *Sexual Plant Reproduction* **16**: 165-177.

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Evolving Genes and Proteins (Bryson V, Vogel HJ, Eds). Academic Press, New York, pp. 97-166.

Zurek DM, Mou B, Beecher B, McClure B (1997) Exchanging the sequence domains between S-RNases from *Nicotiana alata* disrupts pollen recognition. *The Plant Journal* **11**: 797-808.

# APPENDICES

**Appendix A1** Media recipes

| **10X TBE buffer** | |
|---|---|
| Tris base | 108 g |
| Boric acid | 55 g |
| EDTA (0.5 M, pH 8.0) | 40 ml |
| Make up to 1 L with $dH_2O$; dissolve well and store at room temperature | |

| **1X TE buffer** | |
|---|---|
| Tris-HCl (1 M, pH 8.0) | 1 ml |
| EDTA (0.5 M, pH 8.0) | 200 µl |
| Make up to 100 ml with $dH_2O$ | |
| Autoclave and store at room temperature | |

| **EDTA (0.5 M)** | |
|---|---|
| $Na_2$-EDTA | 186.12 g |
| $dH_2O$ | ~500 ml |
| Adjust pH to 8.0 with NaOH | |
| Make up to 1 L with $dH_2O$ | |
| Autoclave to sterilize and store at room temperature | |

| **LB (Luria-Bertani) medium** | |
|---|---|
| Bacto-Tryptone | 10 g |
| Bacto-yeast extract | 5 g |
| NaCl | 10 g |
| Agar (if for solid medium) | 15 g |
| Make up to 1 L with $dH_2O$ | |
| Adjust pH to 7.0 and autoclave to sterilize. Cool to ~50°C and add appropriate antibiotics | |

**Appendix A1** (continuation)

| **X-Gal (40 mg/ml)** | |
|---|---|
| X-Gal | 400 mg |
| Dimethylformamide, DMF | 10 ml |
| Dissolve well and store at -20°C in a brown bottle to protect from light | |

| **IPTG (100 mM)** | |
|---|---|
| IPTG | 238 mg |
| $dH_2O$ | 10 ml |
| Dissolve well and filter-sterilize. Store in 1 ml aliquots at -20°C | |

**Appendix A2** Pollination data for *Solanum okadae*

| Female Parent | (X) | Male Parent | Berries | Seeds | Average Seed/Berry |
|---|---|---|---|---|---|
| OKA 1 | X | OKA 1 | 0 | 0 | 0 |
| OKA 1 | X | OKA 3 | 0 | 0 | 0 |
| OKA 1 | X | OKA 5 | 9 | 950 | 106 |
| OKA 1 | X | OKA 7 | 8 | 1000 | 125 |
| OKA 1 | X | OKA 9 | 12 | 1500 | 125 |
| OKA 3 | X | OKA 1 | 0 | 0 | 0 |
| OKA 3 | X | OKA 3 | 0 | 0 | 0 |
| OKA 3 | X | OKA 5 | 14 | 700 | 50 |
| OKA 3 | X | OKA 7 | 17 | 1200 | 71 |
| OKA 3 | X | OKA 9 | 14 | 1150 | 82 |
| OKA 5 | X | OKA 1 | 9 | 172 | 19 |
| OKA 5 | X | OKA 3 | 10 | 600 | 60 |
| OKA 5 | X | OKA 5 | 0 | 0 | 0 |
| OKA 5 | X | OKA 7 | 11 | 412 | 37 |
| OKA 5 | X | OKA 9 | 7 | 611 | 87 |
| OKA 7 | X | OKA 1 | 7 | 365 | 52 |
| OKA 7 | X | OKA 3 | 4 | 200 | 50 |
| OKA 7 | X | OKA 5 | 8 | 850 | 106 |
| OKA 7 | X | OKA 7 | 0 | 0 | 0 |
| OKA 7 | X | OKA 9 | 10 | 900 | 90 |
| OKA 9 | X | OKA 1 | 15 | 1700 | 113 |
| OKA 9 | X | OKA 3 | 20 | 2300 | 115 |
| OKA 9 | X | OKA 5 | 12 | 1200 | 100 |
| OKA 9 | X | OKA 7 | 17 | 1350 | 79 |
| OKA 9 | X | OKA 9 | 0 | 0 | 0 |

**Appendix A3** Solanaceous S-RNases used for the phylogenetic studies

| S-RNase | Accession # | Species |
|---|---|---|
| S1 | DQ367853 | *Lycium parishii* |
| S2 | DQ367854 | *Lycium parishii* |
| S3 | DQ367855 | *Lycium parishii* |
| S4 | DQ367856 | *Lycium parishii* |
| S5 | DQ367857 | *Lycium parishii* |
| S5 | DQ367858 | *Lycium parishii* |
| S7 | DQ367859 | *Lycium parishii* |
| S8 | DQ367860 | *Lycium parishii* |
| S9 | DQ367861 | *Lycium parishii* |
| S10 | DQ367862 | *Lycium parishii* |
| | | |
| S2 | X03803 | *Nicotiana alata* |
| S3 | U66427 | *Nicotiana alata* |
| S5 | GQ375151 | *Nicotiana alata* |
| S6 | U08861 | *Nicotiana alata* |
| Sc10 | U45959 | *Nicotiana alata* |
| S27 | GQ375153 | *Nicotiana alata* |
| S70 | GQ375150 | *Nicotiana alata* |
| S75 | GQ375152 | *Nicotiana alata* |
| S107 | GQ375154 | *Nicotiana alata* |
| S210 | GQ375155 | *Nicotiana alata* |
| | | |
| S1 | M67990 | *Petunia inflata* |
| S2 | AY136628 | *Petunia inflata* |
| S3 | M67991 | *Petunia inflata* |
| S6 | AF301167 | *Petunia inflata* |
| S7 | AF301168 | *Petunia inflata* |
| S8 | AF301169 | *Petunia inflata* |
| S9 | AF301170 | *Petunia inflata* |
| S10 | AF301171 | *Petunia inflata* |
| S11 | AF301172 | *Petunia inflata* |
| S12 | AF301173 | *Petunia inflata* |
| | | |
| S1 | AF281180 | *Physalis longefolia* |
| S2 | AF281181 | *Physalis longefolia* |
| S3 | AF281182 | *Physalis longefolia* |
| S4 | AF281183 | *Physalis longefolia* |
| S5 | AF374420 | *Physalis longefolia* |
| S6 | AF281184 | *Physalis longefolia* |
| S7 | AF281185 | *Physalis longefolia* |
| S8 | AF281186 | *Physalis longefolia* |
| S9 | AF281187 | *Physalis longefolia* |
| S10 | AF374421 | *Physalis longefolia* |

**Appendix A3** (continuation)

| S-RNase | Accession # | Species |
|---|---|---|
| A-SC | L40539 | *Solanum carolinense* |
| B-SC | L40540 | *Solanum carolinense* |
| C-SC | L40541 | *Solanum carolinense* |
| D-SC | L40542 | *Solanum carolinense* |
| E-SC | L40543 | *Solanum carolinense* |
| F-SC | L40544 | *Solanum carolinense* |
| G-SC | L40545 | *Solanum carolinense* |
| H-SC | L40546 | *Solanum carolinense* |
| J-SC | L40547 | *Solanum carolinense* |
| K-SC | L40548 | *Solanum carolinense* |
| | | |
| S1 | EF680106 | *Solanum chilense* |
| S2 | EF680089 | *Solanum chilense* |
| S3 | EF680103 | *Solanum chilense* |
| S4 | EF680094 | *Solanum chilense* |
| S6 | EF680086 | *Solanum chilense* |
| S7 | EF680109 | *Solanum chilense* |
| S8 | EF680085 | *Solanum chilense* |
| S9 | EF680093 | *Solanum chilense* |
| S10 | EF680088 | *Solanum chilense* |
| S11 | EF680110 | *Solanum chilense* |
| | | |
| S1 | AY454099 | *Witheringia solanaceae* |
| S2 | AY454100 | *Witheringia solanaceae* |
| S3 | AY454115 | *Witheringia solanaceae* |
| S4 | AY454102 | *Witheringia solanaceae* |
| S5 | AY454103 | *Witheringia solanaceae* |
| S6 | AY454104 | *Witheringia solanaceae* |
| S7 | AY454105 | *Witheringia solanaceae* |
| S8 | AY454106 | *Witheringia solanaceae* |
| S9 | AY454107 | *Witheringia solanaceae* |
| S11 | AY454109 | *Witheringia solanaceae* |

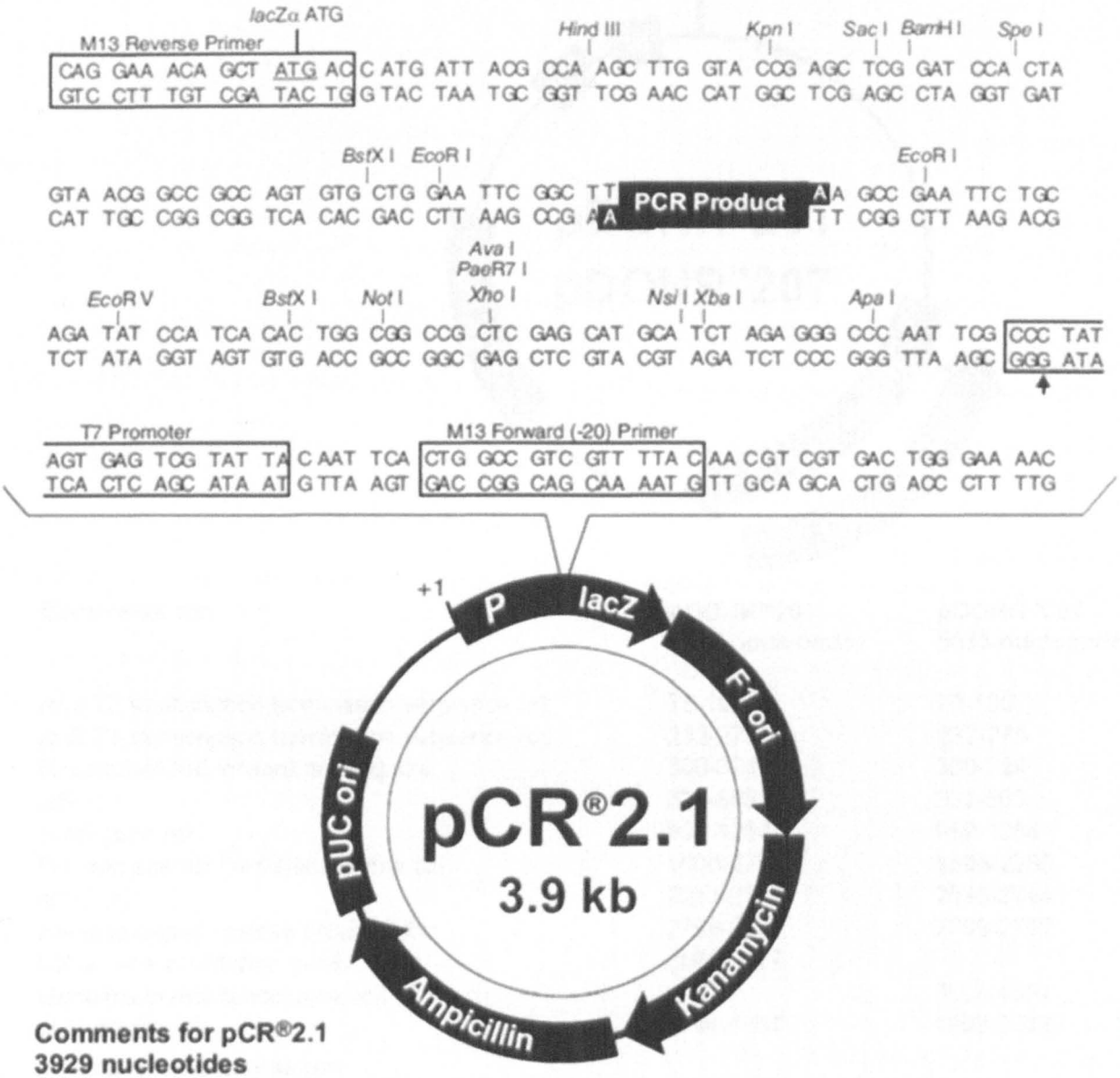
**Appendix A4** Solanaceous S-like RNases used for the phylogenetic studies

| S-like RNase | Class | Accession # | Species |
|---|---|---|---|
| NE | I | U13256.1 | *Nicotiana alata* |
| NGR3 | I | AB032257 | *Nicotiana glutinosa* |
| NK1 | I | AB034638 | *Nicotiana tabacum* |
| LE | I | X79337 | *Solanum lycopersicon* |
| RNS2 | II | AK324819.1 | *Solanum lycopersicon* |
| NGR2 | II | AB032256 | *Nicotiana glutinosa* |
| LER | II | CAL64053 | *Solanum lycopersicon* |

**Appendix A5** Alignment of partial DNA sequences of $S_{o1}$- and $S_{o2}$-RNases showing the QRT-PCR primer positions (see Table 5.2). $S_{o1}$-RNase forward (qpcrSo1-F) and reverse (qpcrSo1-R) primers are indicated in blue and light blue respectively. $S_{o2}$-RNase forward (qpcrSo2-F) and reverse (qpcrSo2-R) primers are indicated in green and light green respectively. The alignment highlights the allele specificity of the two primers by showing the nucleotide differences at the positions used for the primer design.

```
                   10        20        30        40        50        60        70        80
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
SO1-RNase  GGATAAGGAGGGATCACAGCTGTTGAAGTACTGCAAGCCGAAACTTAAATATAAATATTTCACTGATG---AGATGCTCA
SO2-RNase  ......CA..A.--....A-..C.T..TA.....G..T.CG..GAC..G...GC.G..A..T.G....CCA...A..G..

                   90       100       110       120       130       140       150       160
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
SO1-RNase  ATGATCTTGACAAACACTGGATTCAATTGAAGATTGATCAAGCTTCTGCTCTAAAGGACCAACCAGCATGGAGATATCAA
SO2-RNase  .AC.A..G..GT.T......CC.G.T....CCGCCA..GT..G.GA.AT.AA...AC.A...GGTTTC...G.....G..

                  170       180       190       200       210       220       230       240
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
SO1-RNase  TATCTAAAGCATGGATCCTGCTGTCAGAAAATCTACAATCAAAACACGTATTTTAGTTTAGCCTTGCGCTTAAAAGATAG
SO2-RNase  .T.AAT.........A.G..TA..ATCG.TC....T......G.TG.A......GA...G...A.CAAA......A.C.A

                  250       260       270       280       290       300       310       320
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
SO1-RNase  GTTTGATCTTCTGAGAACTCTCCAAATACATAGAATTGTTCCTGGATCAAGTTATACATTTGAAGAAATCTTTGACACCG
SO2-RNase  ..........T...CGG......G..AT...G.T..AA....C...GATGT.AGA..GG.AA..A.TG..GAGA.TG..A

                  330       340       350       360       370       380       390       400
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
SO1-RNase  TCAAAACAGTTACTCAGGTGGATCCTGATATCAAGTGTACTGAAGG---------AGCACCGGAACTATATGAGATAGGC
SO2-RNase  .T..GG.T..G...ACA.AA.T....A.CC....C..C.T..GT.ATTCTGGCCA.A GAT....T...TG........A

                  410       420       430       440
           ....|....|....|....|....|....|....|....|...
SO1-RNase  ATATGTTTTACCCCAAATGGAGATAGTCTAGTTCGTTGTCGTC
SO2-RNase  ........CGA..G.G...C.AC.CAGG.GA..GC........
```

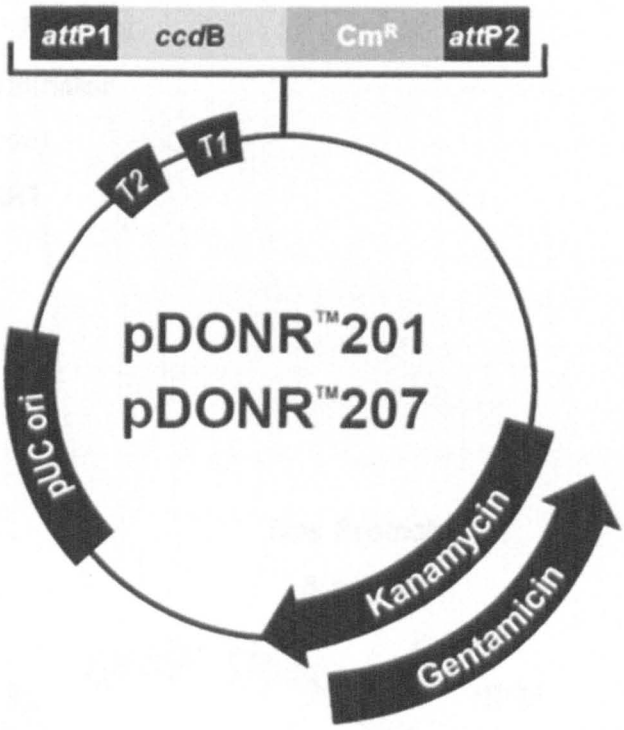
## **A6.1** pCR2.1 vector




**Comments for pCR®2.1**
**3929 nucleotides**

*LacZα* gene: bases 1-545
M13 Reverse priming site: bases 205-221
T7 promoter: bases 362-381
M13 (-20) Forward priming site: bases 389-404
f1 origin: bases 546-983
Kanamycin resistance ORF: bases 1317-2111
Ampicillin resistance ORF: bases 2129-2989
pUC origin: bases 3134-3807

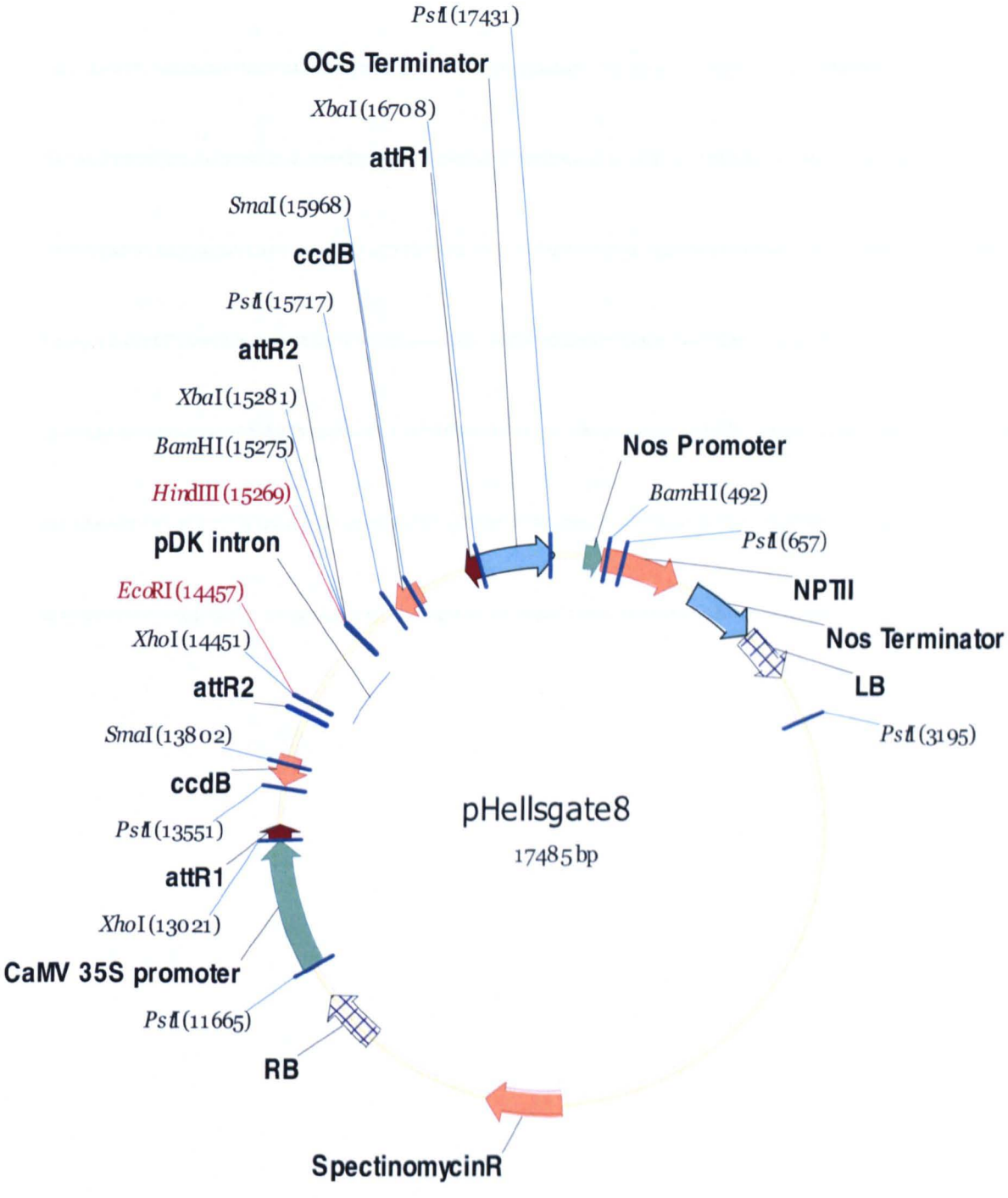
**A6.2** pDONR207 vector




| Comments for: | pDONR™201<br>4470 nucleotides | pDONR™207<br>5585 nucleotides |
|---|---|---|
| *rrn*B T2 transcription termination sequence (c): | 73-100 | 73-100 |
| rrnB T1 transcription termination sequence (c): | 232-275 | 232-275 |
| Recommended forward priming site: | 300-324 | 300-324 |
| *att*P1: | 332-563 | 332-563 |
| *ccd*B gene (c): | 959-1264 | 959-1264 |
| Chloramphenicol resistance gene (c): | 1606-2265 | 1606-2265 |
| *att*P2 (c): | 2513-2744 | 2513-2744 |
| Recommended reverse priming site: | 2769-2792 | 2769-2792 |
| Kanamycin resistance gene: | 2868-3677 | --- |
| Gentamicin resistance gene (c): | --- | 3528-4061 |
| pUC origin: | 3794-4467 | 4909-5582 |

(c) = complementary strand

**A6.3** pHellsgate8 vector



**Appendix A7** cDNA sequence for the cloned S-RNases

## **A7.1** $S_d$-RNase from *P. inflata*: partial sequence

```
        10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGATAACGTCAGTACACCGCTGAATTTTTGTGGTGCCAAAGAAGAAAACTACAAAGACTTAACGGATGATAACAAAAAGAATAATTTGTA

       100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TAAACGTTGGCCTGACTTGACCACCGATGAAGCTGTATGTTTGGAAAAGCAAGATTTCTGGAGACATGAGTATAATAAGCATGGAACGTG

       190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TTGTTTAGGTAGCTACAATGAAGATCAATACTTTCATTTAGCCATGGCCCTAAAAGACAAGTATGATCTTCTAACATCTTTGAGAAAGCA

       280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGGAATTAGTCCTGGCTGGCAATATACCGTTCAGAAAATCAATAGCACCATCAAGACAATAACTCGAGGGTATCCTAACCTCTCGTGCAC

       370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TAAGAAAATGGAACTATGGGAGATAGGAATATGTTTCGACTCGACGGTACAAAAGGTGATCGATTGTCCTCATCCTAAGACATGCGCGCC

       460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GATGAAAATTATGTTTCCATAAATAAAATTTCATTTTCTCTCTTATGTAAACCGCTACGATAATCAGTACAATCAAATAAAATTTTATTT

       550       560       570       580       590       600       610       620
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GAAAATTTGATATCAAATTTATAAAAACTACAATTGAATTTAAACATTCTATTCTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

## **A7.2** $S_m$-RNase from *P. hybrida* Mitchell: complete sequence

```
        10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AAAATCAAAATGTTTTTTTCGAACTTTAGAATGTTCAAGTTACAGCTCACATCAGCTTTCTACATTTTCCTTTTTGCTCTTTCCCCCATT

       100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TATGGGTATTTCGACCACTGGCAACTCTCTTTAACATGGCCTGCAGGTTTTTGCAAAATTAAGAATTGCCCGATAAAACCGATTCCGAAC

       190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AACTTTACCATTCATGGACTTTGGCCAGATAACGTCTCCGCTGTGTTGTATGACTGCCATCCCAATAAACAGTTTACAACGATCATAGAT

       280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CCAAGAATTAAAACCGAACTGGAGAAGCGCTGGCCTGAGTTAACCACAAGCAAATCTGCTTTGCACCGTCAACCTTTCTGGAAATATGAA

       370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TATGAGAAGCATGGATTATGTTGTTCAGATGTCTACAGTCAATCAGAATATTTTGAATTTGCTATGAAATTTAAAGACAGGACTGATCTA

       460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TTGACTATTCTCAGAAGTAAGGGAGTTGCTCCGGGATTTACTTATACTGGAGAAAAAATCAACAGTTCCATCGCTTCGGTAACTCGAGCG

       550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GCACCTAACATCAAGTGCCTTTATTATCAAGGAAATTTGGAATTAACTGAGATAGGAATATGTTTTAACCGAACAACAGAACGTATGATG

       640       650       660       670       680       690       700       710       720
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TCGTGTCCTCGGATTAGTACGTCATGCAAATTTGGGACTAATGCGGGGATTATGTTTCGACCGTGAAAAACTACGGAAAGTAGCGCCATC

       730       740       750       760       770       780       790       800       810
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GTCATAATAATATATTGTGGTTATAAAATCATTACGATATAGGATCCATAATATACTTAACTTTTGTGATTTATCTTTTAATCTGTGCGT

       820       830
....|....|....|....|....|....
ATTTTTAGTTCAAAAAAAAAAAAAAAAAA
```

**A7.3** $S_{o1}$-RNase from *S. okadae*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGATAAGGAGGGATCACAGCTGTTGAAGTACTGCAAGCCGAAACTTAAATATAAATATTTCACTGATGAGATGCTCAATGATCTTGACAA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ACACTGGATTCAATTGAAGATTGATCAAGCTTCTGCTCTAAAGGACCAACCAGCATGGAGATATCAATATCTAAAGCATGGATCCTGCTG

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TCAGAAAATCTACAATCAAAACACGTATTTTAGTTTAGCCTTGCGCTTAAAAGATAGGTTTGATCTTCTGAGAACTCTCCAAATACATAG

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AATTGTTCCTGGATCAAGTTATACATTTGAAGAAATCTTTGACACCGTCAAAACAGTTACTCAGGTGGATCCTGATATCAAGTGTACTGA

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGGAGCACCGGAACTATATGAGATAGGCATATGTTTTACCCCAAATGGAGATAGTCTAGTTCGTTGTCGTCAAAGTGAAACATGCGACAA

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AACAGGGAAAATATTTTTTCGTCCCTGAACAACTTCTATTCCTTCATATATTATTCAAGGTAAGGTAAGTGCTACAAAATACAAAGCCAA

        550       560       570       580       590       600       610
....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CATAATAATCTAAACCATCAATAATATAAGTAGTCTGAATGACCACGTAGGAAACCCTGTTGAAAAAAAA
```

## **A7.4** $S_{o2}$-RNase from *S. okadae*: complete sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ACTATTCGAACTACGGAATGATTAAGCTACAGGTCTTATCAGCTCTCTTAATCTTGCTTTTTGCACTTGCTCCCGTTTTAGGAAATTTCG

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGTACCTTCAACTCGTTTTACAATGGCCAGCAACTTTTTGTCACACTAGAACTTGCCCAGTTAAGCCAATTCCAAACAACTTTACAATTC

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ACGGGCTTTGGCCGGATAACAAGAGCACAATGCTTAATAACTGCGAGTCCGAAGACAAGTATGCAGATATCTCGGATGCCAAGAAGCGCA

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AACAACTGGAGTATCACTGGCCTGATTTGACCGCCAATGTAGGTGATATTAAAAAACAACAAGGTTTCTGGGGATATGAATTTAATAAGC

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ATGGAACGTGTAGTATCGATCTCTATAATCAAGATGCATATTTTGATTTGGCCATCAAATTAAAAAACAAGTTTGATCTTTTGACGGCTC

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TCCGAAATCATGGTATAATTCCCGGAGATGTTAGAACGGTAAAAAATGTCGAGAATGCCATTAAGGCTGTGACTACAGAAGTTCCTAACC

        550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TCAACTGCATTGGTGATTCTGGCCAAACGATGGAATTATTGGAGATAGGAATATGTTTCGACCGAGATGCAACTCAGGTGATTGCTTGTC

        640       650       660       670       680       690       700       710       720
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GTCGACGTTGGAAAAGCCACCCTAACAAAAATAAGAGGGTTACTCTTCCATAGTGAATAGCTTCCGTTTTCTTTCCCTTTTTTTCTATGC

        730       740       750       760       770       780       790       800       810
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TTTTTTCTGAGTACAGTAAATGAAATACGTCTGATAATATAATGACCTACCAGACAAAATAATCAGTATGCCAACATTGTAAATTGTTTA

        820       830       840       850       860       870       880       890
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TCAGTTTATGATAATCAATATTATTTTGAAATGAGTATTGAAAACACCCTAAACTTAATGAGAAAAAAAAAAAAAAAAAA
```

**A7.5** $S_{o3}$-RNase from *S. okadae*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGACATGGAAAAATTTATGTTGAATAATTGCAAGGGTAAAAAGTATTCTAGTATAGAGATTCCCCTTGAACAAAAAAAAATGGACGCTCG

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTGGCCAGACTTGAAAAACACAGAAGAATTCAGCCTGGAAGAACAACCGTTCTGGCAATATGAATACAATAAGCACGGGACGTGTTGTCA

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GACACGCTACAATCAAGAACAATATTTTAATCTAACCATGAAATTAAAAGACAAGTTTGATCTTTTGTCAGCTCTCCAACATCATGGAAT

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TAATCCTGGATCAACTCCTACGGTCAAGCAAATCGGCACTGCCATTGCGACAGTTACTAAAGTATATCCAAGCCTAAAGTGCATTCCTAT

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CAACGGAAATCTTAAACTGTTGGAGATAGGCATATGTTTCAACCCAGAGGCAACAAATGTAATTCCTTGTCATCGATCTTGGATATGTCA

        460       470       480       490       500       510       520
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....
CCAGGATCAGAATATGGTGATTGAAATTCCAAATTAAGTTTTTTCTTTCCTTAAAAAAAAAAAAAAAAAAAAA
```

**A7.6** $S_{o4}$-RNase from *S. okadae*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CGAAAAAGAGGAGTTTCGATTGGAGTTCTGCACTGGCAATAAGTATAATCATTTCAGTGTAAAAGATAGTATAGTCAATGATCTGGAGAA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGAACACCATTGGATTCAATTGAAGTTCGATGAACAATACGCCAGGAATAATCAACCTCTCTGGAGCCATGAGTACACAAAGCATGGAAT

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GTGCTCTTCAAATCTTTACGATCAGAGAGCATATTTTTTATTAGCCATGCGCTTAAAAGATAAGTTTGATCTTTTGACAACTTTCAGAAC

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TCATGGAATTACTCCCGGAACAAAGCATACATTTGATGAAATCCAAAGTGCCATCAAGACAGTTACTAATAAAGTAGTTGCTGATCTCAA

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GTGTGTCCAACATATCAAAGGAGTACAGGAACTAAAAGAGATAGGCATATGTTTTACCCCAGAGGCAGATAGCTCTCATCCATGTCGTCA

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGGTAACACATGCGATGAAAGTATGAGTATTTTGTTTCGATGAATAGTCAAGATTATTGGGCAGAAAATATAATGATTCACCTAACTATG

        550       560       570       580       590       600       610
....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
GAACCAGTCGGTCTAAACCTATTGTCTGGTAACGAATATCATTATTGAATTTACAAAAAAAAAAAAAAAAAAAAA
```

## **A7.7** $S_{o5}$-RNase from *S. okadae*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGATAAAGAGGGAACACTTCTGCAATACTGCAAGCCAAAACCTACTTTTATATTAAACAAGGATAAGATGCTTGATGATCTTGACAAGAA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTGGATTCAGTTAAGGTATCCAGAAGAGTATGGTCGAAAGGAACAACCTTTATGGCAATATCAATATCTAAAGCATGGATCCTGTTGTCA

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GAAAGCCTACGATCAAAACAAGTATTTTAGTCTAGCTTTGCGCTTAAAAGACAGGTTTGATCTTCTGAGAACTCTCCAAATACATCGAAT

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGTTCCTGGATCAAGTTATACATTTAAAAGGATATTTGATGCCGTCAAGACAGTTACTAAAACAGATCCTGACGTCAAGTGTACAAAAGG

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGCACCGGAGCTATATGAGATAGGCATATGTTTCACCCCAAATGCAGATAGTCTGATTCCTTGTCGTCAAAGTGAAACATGCGACAACTC

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GAAAGACATCTTTTTTCGAAGATGAACAATTCCATTCGTTTTTATTTTATATATGTTAAGTGTTATGAAGTACAAAACTTACAGAATAAA

        550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TTGTTATGAAATACAAAACCTACAGTATAATCTAAACTATCAAGATTATAAGCACTTTGAATGACCAGGTTGGAAACTTTTCAGAAGAAA

        640       650       660
....|....|....|....|....|....|....|....
ATATCTTATTAATTGATGAGAAAAAAAAAAAAAAAAAAA
```

## **A7.8** $S_{p1}$-RNase from *S. phureja*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGATAAAGAGGGAACACTGCTGCAGAACTGCAAGCCATTACCTACGTATATACATTTCGCGGATAAGATGCTCAATGATCTTGACAAAAA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTGGATTCAATTGAAGTATCCAGAACGTTTTGCTCGAAAGGAACAACCTTTATGGCTATATCAATATCTAAAGCATGGATCCTGTTGTCA

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GAAAGTTTACGATCAAAACACGTATTTTAGTCTAGCTTTGCGCTTAAAAGACAGGTTTGATCTTCTGAGAACTCTCCAATTACATCGAAT

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGTTCCTGGATCAAGTTATACATTTAAAGAAATCTTTGATGCCGTCAAGACAGTTAGTCAAACAGATCCTGACGTCAAGTGTACAAAAGG

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGCACAGGAACTATATGAGATAGGCATATGTTTCACCCCAAATGCAGATAGTCTGATTCCTTGTCGTCAAAGTGAAACATGTGACAAATC

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GAAAGAAATCTTTTTTCGTAGATGAACAATTCCATTCGTTTTTATTTTATATATGTTAAGTGTTATGAAGTACAAAACTTACAGAATAAA

        550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GTGTTATGAAATGCAGAACCTACATGCAGAACCTACAGTATAATCTAAACTATTAAGATTATAAGCACTTTGATGACCAGGTTGGAAAAA

        640
....|....|....|.
AAAAAAAAAAAAAAAA
```

**A7.9** $S_{p2}$-RNase from *S. phureja*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGATAACAAGTCCATTATACTGAATAACTGCAATGTTGCCAATAAAGAACGTTACATTAAGATCGAGGATCCCAAAAAACTAACTGAGCT

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGACAAACGCTGGCCTCAACTCAGATACGACAAATTGTATGGCATAGATAAACAATATCTCTGGGAAAAAGAATTCCTAAAACATGGAAC

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTGTAGTATAAATCGCTACAAACAAGCAGCATACTTTGATTTAGCCATGAAAATAAAAGACAGGTTTGATCTATTGGGAACTCTCAGAAA

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TCATGGAATTAATCCTGGTTCAACTTATGATCTTGATGATATCGAACGTGCTATAAAGACAGTTTCTATAAAGGTTCCTAGCCTTAAGTG

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CATAGAAAAGCCACCTGGAAATGTGGAACTTAATGAGATTGGTATATGCTTAGACCCAGAAGCGAAATATACGGTTCCCTGTCCACGAAC

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGGGTCATGCCATGAATTGGGACCTAGGATAAAGTTCCGATGATAAATAAACTCTTCATTTCATTTCCTATTTTGTTGTGATAAGTAAAT

        550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TAAATGCAAACTCTTCAGGCGAGAAGCCAAATGTTACTCTAGCAGGTCACACTGTTTAATATTAGGAGTAAATTTCTTGTCAACGCTTGA

        640       650       660       670       680       690
....|....|....|....|....|....|....|....|....|....|....|....|...
ATAGCATTACTCAACTAAATGAGATTTTAATATGTTTACATCTAAAAAAAAAAAAAAAAAAAA
```

**A7.10** $S_{s1}$-RNase from *S. stenotomum*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGATAGCGTGGGTGGAGAACTGAATTACTGTGATGGCAAAGCTAAGTATACTAGGGTCAAGGATGAAGCATTTGATAAGAGGAATAAGCA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTGGCCTGACTTGTTACTAAGCGATGCTGATAATCTGAAAAACCAAGGTTTCTGGGTACATGAATACAGAAAGCATGGATCGTGTTGTAA

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AAATCTCTTCAATGAAAAACAATACTTTGATTTAGCATTGGTTTTAAAAGACAGGTTTGATCTTCTGACGACTTTCAGAAATCACGGAAT

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGTTCCTGAATCATCTCATACTGTTCATAAAATTGAAAAAACTATTAGGTCAGTTACTGGGGTTCTTCCTAATCTCTCTTGCACTAAAAA

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TATGGATCTTTTGGAGATAGGGATATGTTTCAACCGAGACGCAAGTAATATGATTGATTGTCCAAGGCCTAAGACGTGCAGCCCTGGCGA

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AAATAATCTGATTGCATTTCCATGATGATGATTATCTTTTTGTATCTTTCTTCATCCCCCTCCCCTTGCTGTAATTTGCTTAATAAATAC

        550       560       570       580       590
....|....|....|....|....|....|....|....|....|..
AACAAATGAAAAACCTATATAATCAATCCCATCAAGAAATAAAAAAAAAAAAAAAAA
```

**A7.11** $S_{s2}$-RNase from *S. stenotomum*: complete sequence

```
        10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TTGTTTGAACTGCAGAATGTTTAGACAACTCGCGTCAGTTTTCGTCCTTCTACTTGTTACTCTTTCTCCGGGTTATGGGAATTTCGAGTA

       100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTTGCAACTTGTTTTAACGTGGCCGGCAACTTTTTGCCACACGAGAAGCTGCGTCAGAATTCCAAATAACTTCACAATTCACGGGCTTTG

       190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GCCGGATAACAAGAGCACGCGGCTGAATTTCTGCAGTAAACATCCTAAGTATAATATGATCATGGATGAAGATAAGAAAGAAGCTCTGGA

       280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ATACATCTGGCCTAACTTGACAACAACTAAAGTTGTTTCAAAAAAGTATCAAATTTTTTGGAGGAAAGAATACATTAAGCATGGAACGTG

       370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TTGTTCAGATCTCTTTAATCAAGATCAGTATTTTGATTTAGCCATTGCGTTAAAAAACAAGTTTGATCTTTTGAATATTCTCGGAAAGAA

       460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGGTATTACTTCTGGAACAAGTCATCTTACCTCTCAAAAAATCCAAAGTGCTATCAAGTCCATTACTAGCGGGGTTCCAAATCTCATCTG

       550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTCTGATAATTTTAATGCTGGAACAACGGAACTATGGCAGATAGCCATATGTTTGGACAGAAACATCGCTGTGATAGATTGTCCTCTACC

       640       650       660       670       680       690       700       710       720
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AAAGATATGCACTCAAACAGGACCTAAGGGGATCACCTTTCCCTGATGAATAACTTCCAATTTCTTTCAGCTTCCCTTCTTTATGAAAAG

       730       740       750       760       770       780       790       800       810
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TATAATTTAAAGAGAACTATTCTGATAATCAAGAACTTTAATCATAGGAATTTTCAAGGACGATTACATTGTAGATGCGCCGTAAATCTC

       820       830       840       850
....|....|....|....|....|....|....|....|.
TTATGAATTAATGAAACTCATTAGTGAAAAAAAAAAAAAAAAAAA
```

## **A7.12** $S_{s3}$-RNase from *S. stenotomum*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CGAGAAGAATCGGTTTCGTCTGGAGTACTGCTCCGGCGGTACCGCGTTTAAGAAATTTGAATTACAAGATCGTATAGTCAGTGATCTGGA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TCGCCATTGGATTCAAATGAAGTTCAACGAACAAGATGCTAAACATAAACAACCTCTCTGGAACCACGAATACACAAGGCATGGAAGGTG

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TTGTTACAATCTCTACGATCAGAACGCATATTTTTTACTAGCCATGCGCTTAAAAGATAAATTAGATCTTGTAACAACTCTCAGAACTCA

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGGAATTACCCCAGGGACAAAGCATACATTTGATGAAATCAAAAGTGCTATTAAGACCGTTACTAATCAAGTAGATCCTGATCTCAAGTG

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CGTCGAGTATACGAAAGGAGTACAGGAATTAAAAGAGATAGGCATATGTTTTACCCCCTCAGCCGATAGTTTTTATCCATGTCGTCATAG

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TAATACATGCGATGAAAAGGGTACGGCGATTTTATATAGATAATGAATGACTTGCCATTACCAACTTTTCTGTTGGTTTATGGTTAGATA

        550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AACAAGATGCGTAACTACCCATTGAAGTAAATTCAATGTCAATTTACGAACATCATTAATAAATTTATATTCTCTAAAAAAAAAAAAAAA

....|..
AAAAAAA
```

## **A7.13** $S_{s4}$-RNase from *S. stenotomum*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGATAACAAATCCATGCGTCTGAATAACTGTGATCCCAATCTTAACTATCGGATCATCCCGGAGGGTGAGAAACGCTTCAATCTTGACAA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GCGCTGGCCTCAGTTAGAAAATACCAAAGAATTTGCTTTGGCAAAGCAACCATTCTGGGAAAAGGAATATAAAAGGCATGGAACGTGTTG

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CAAAAATCTGTATAACCAAGCAACATACTTTGATTTGGCCATGAATTTAATAGACAAGTTTGATGTTTTGACAATTCTCAGAGATCAGGG

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AATCATCCCTGGAACGTACTATGTCGTTAAAAGAGTAGAAGATGCCATCAAGAAAGTTACCCATCAACTTCCTAAGCTCAACTGCGTAGT

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CAATAATATTGTAGGACAGGAACTAAGTGAGATAGTCATATGTTTCGAACCTAATGGAAAGTATGTGGATAGTTGTCGTCGACCTGGGTC

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ATGCAACCAAAATGGAAATATGGAAAGGATTTTTTTTTCGATGATAACTTTCCGATGCTTTATTTATGCGGCTTTAATAAAATGCCAAAAC

        550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTCTGGAATAAACAAGTGCATCACGCAACCGAGTGTGTATCAATATTAAATGTGTATTGTCACATTGCCTGTTCATCGTCAGAGATTAAC

        640       650       660       670       680       690       700
....|....|....|....|....|....|....|....|....|....|....|....|....|....|...
TAGTATCCACGGATGGTTGATAATAAAATTGCATCATTAGCTTTTTTTGAAAAAAAAAAAAAAAAAAAAAAAA
```

## **A7.14** $S_{s5}$-RNase from *S. stenotomum*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGATAAACAGGGAACAATGCTGCAATACTGCAAGCCAAAACCTACGTTTCTATTGATCAAGGATAAGATGCTTGATGATCTTGACAAAAA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTGGATTCAGTTGAAGTATCCACAACGTTATGCTCGAAAGGAACAACCTTTATGGAAATATGAATATCTAAAGCATGCATCCTGTTGTCA

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GAAAGTCTACGATCAAAACACATATTTTAGTCTAGCTTTGCGCTTAAAAGACAAGTTTGATCTTCTGAGAACTCTCCAAATACATCAAAT

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGTTCCTGGATCAAGTTATACATTTAAGGAAATCTTTGATGCCATCAAGACAGTTACTCAAACAGATCCTGACGTCAAGTGTAAAAAAGA

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGCACCAGAACTATATGAGATAGGCATATGTTTCACCCCAAATGCAGACAGTCTGATTCCTTGTCGTCAAAGTAATACATGCGACAAATC

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AAAAGAAATCTTTTTTCGAAGCTGAACAACTTCCATTCCTTTTTATTTTATTTATGGTATGTGTTATGAATTACAAAGCCTACAGAATAA

        550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGTGTTATGAGAAACAAAACCTACGGTATAATCTAAACTATCAAGATTATAAGCATTCTGAATGACCAGGTTGGAAACCTTTCAGAACAG

        640       650
....|....|....|....|....|...
AACATCTTATTAAAAAAAAAAAAAAAAA
```

## **A7.15** $S_{s6}$-RNase from *S. stenotomum*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGATAACAAGTCCACAGTACTGAATTTCTGCAATTTAGCCCATGAAGATGAGTACATTCCGATCACGGATCACAAGATATTAACTGAGCT

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGACAAACGCTGGCCTCAACTCAGATACGATTATTTGTATGGCATACGTAAACAATATCTCTGGAAAAATGAATTCGTAAAACATGGAAG

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTGTAGTTTAAATCGCTACAAACAAGCAGCATACTTTGATTTAGCCATGAAGATGAAAGACAAGTTTGATTTATTGGGAACTCTCAGAAA

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TCATGGAATTAATCCTGGTTCAACTTATGAACTTGATGATATCGAACGTGCTATAATGACAGTTTCTATAGAGGTTCCTAGCCTCAAGTG

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CATACAAAAGCCACTTGGAAATGTGGAACTTAATGAGATTGGTATATGTCTAGACCCAGAAGCGAAATATATGGTTCCCTGTCCACGAAC

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGGGTCATGCCATAATATGGGACATAAAGTAAAGTTCCGATGATGAATAAACTCTTTAATTCACTTTCTTATTTTGTTATGATAAGTAAA

        550       560
....|....|....|....|....|..
TTTAAATGCAAAAAAAAAAAAAAAAAA
```

## **A7.16** $S_{s7}$-RNase from *S. stenotomum*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGATAGCGAGGCTGGAGAGCTGAATTTCTGTAATCCCAAAGCTACGTATACTATTGTCACGGATCAAACATTAAATAAGAGGGACAAGCA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTGGCCTGACTTGTTCCTAAGCAAAGCTAATGGTCAGGATATACAAAAATTCTGGCAACATGAATACATAAAGCATGGAACATGTTGTTC

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGATCTCTTCAATGAAGAACAATACTTTGATTTAGCGTTGGTCTTAAATGACAGGTTTGATCTTGTGAAGACTTTCAGAAATCACGGAAT

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TATTCCTCTATCTTCTCATACTGTTCATAAAATTGAAAAAACTATCAGGTCAGTTACTGGGGTTCTTCCTAATCTGTCTTGCACTAAAAA

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TATGGAACTTTTGGAGATAGGAATATGTTTCAACCGAAACGCAAGTAATATGATTGATTGTCCAACACCAAAGACGTGCAACCCTGACAA

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AACTAATTTGATTACGTTTCCATGATGATTAACTTTTTTTTATCTTTCTTCATTTTCCCCTTGCTGTATGATGAAAAAAATACAACAAAT

        550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AAGTTAATATTGAGCAACATTTTGTTGAACCAGCAGAAGTAAAATTATTGCTTATGGAAAATACTGCCCCTCTTTGTTCCAAAAAAAAAA

....|..
AAAAAAA
```

## **A7.17** $S_{s8}$-RNase from *S. stenotomum*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGATAAAGAGGGAACGCTGCTGCAGTACTGCAAGCCAAAACCTAATTATAGAAATTTCGGGGATAAGATGCTCAATAATCTTGACAAAAA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTGGATTCAATTGAAGTATCCAGAAGATTATGCTCGAAATCAACAACCTTTATGGCAATATCAATATCTAAAGCATGGATCCTGTTGTCA

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GAAAGTTTATGATCAAAACAGGTATTTTAGTCTAGCTTTGCGCTTAAAAGACAGGTTTGATCTTCTGAGGACTCTCCAAATACATCGAAT

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGTTCCTGGATCAAGTTATACATTTAAAGAAATCCTTGATGCGGTCAAGACAGTTACTCAAACAGATCCTGACGTCAAGTGTACAAAAGG

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGCACAGGAACTATATGAGATAGGCATATGTTTCACCCCAAATGCAGATAGTCTGATTCCTTGTCGTCAAAGTGATACATGCGACAAATC

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GAAAGACATATTTTTCCGTAGTTGAACGACTTTCATTGCTTTATATTTTATTTATGGCAAGTGTAATAAAGTACAAAACCTACGGAACAA

        550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGTGTTATGAAGTACAAAACCTACAGTATAATCTAAACTATCAAGATTATAAGCATTTTGAATGACCAGGTTGGAAACCTTTCTGAACAA

        640
....|....|....|....
AAAAAAAAAAAAAAAAAAA
```

## **A7.18** $S_{s9}$-RNase from *S. stenotomum*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GGATAAAGAGGGAACACTGCTGCAGAACTGCAAGCCATTACCTACGTATATACATTTCGCGGATAAGATGCTCAATGATCTTGACAAAAA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CTGGATTCAATTGAAGTATCCAGAACGTTTTGCTCGAAAGGAACAACCTTTATGGCTATATCAATATCTAAAGCATGGATCCTGTTGTCA

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GAAAGTTTACGATCAAAACACGTATTTTAGTCTAGCTTTGCGCTTAAAAGACAGGTTTGATCTTCTGAGAACTCTCCAATTACATCGAAT

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGTTCCTGGATCAAGTTATACATTTAAAGAAATCTTTGATGCCGTCAAGACAGTTAGTCAAACAGATCCTGACGTCAAGTGTACAAAAGG

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGCACAGGAACTATATGAGATAGGCATATGTTTCACCCCAAATGCAGATAGTCTGATTCCTTGTCGTCAAAGTGAAACATGTGACAAATC

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GAAAGAAATCTTTTTTCGTAGATGAACAATTCCATTCGTTTTTATTTTATATATGTTAAGTGTTATGAAGTACAAAACTTACAGAATAAA

        550       560       570       580
....|....|....|....|....|....|....|....|
GTGTTATGAAATGCAGAACCTAAAAAAAAAAAAAAAAAAA
```

## **A7.19** $S_{s10}$-RNase from *S. stenotomum*: partial sequence

```
         10        20        30        40        50        60        70        80        90
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CGAGAAGAATCGGTTTCGTCTGGAGTTCTGCTCCGGCGGTGCCGCGTATAAGAAATTTGAATTACAAGATCGTATAGTCAGTGATCTGGA

        100       110       120       130       140       150       160       170       180
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TCGCCATTGGATTCAAATGAAGTTCAACGAACAAGAGGCTAAACAAAAACAACCTCTCTGGAACCACGAATACAAAAGGCATGGAAGGTG

        190       200       210       220       230       240       250       260       270
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TTGTTACAATCTCTACGATCAGAACGCATATTTTTTACTGGCCATGCGCTTAAAAGATAAATTAGATCTTGTAACAACTCTCAGAACTCA

        280       290       300       310       320       330       340       350       360
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TGGAATTACCCCAGGGACAAAGCATACATTTGATGAAATCAAAAGTGCTATTAAGACCGTTACTAATCAAGTAGATCCTGATCTCAAGTG

        370       380       390       400       410       420       430       440       450
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CGTCGAGCATACAAAAGGAGTGCAGGAATTAAAAGAGATAGGCATATGTTTTACCCCCTCAGCCGATAGTTTTTATCCATGTCGTCATAG

        460       470       480       490       500       510       520       530       540
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TAATACATGCGATGAAAAGGGTACGATGATTTTATTTAGATAATGAATGACTTTCCATTGCCAACTTTTCTTTTTGTTTATGGTTAGATA

        550       560       570       580       590       600       610       620       630
....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AACAAGATGCGTAACTACCCATTGAAGTAAATTCAATGTCAATTTACGAACATCATTAATAAATTTATATTCTCTACTTTAGTCCACTAA

        640       650
....|....|....|....|....|
AAAAAAAAAAAAAAAAAAAAAAAAA
```