Promotors: Prof dr. ir. E. Van Bockstaele
Department of Plant Production, UGhent
Dr. I. Roldán-Ruiz
Department of Plant Genetics and Breeding, CLO

Dean: Prof. Van Langenhove

Rector: Prof. Dr. A. De Leenheer

Ir. Inge Van Daele

Identification of genes related to self-incompatibility in ryegrass (*Lolium perenne* L.)

Thesis submitted in fulfillment of the requirements
for the degree of Doctor (PhD) in Applied Biological Sciences

Identificatie van genen betrokken bij zelfincompatibilteit in raaigras (*Lolium perenne* L.)

Cover illustration: right: a compatible, full-grown pollen tube and left: an incompatible, mal-formed pollen tube

Van Daele Inge, “Identification of genes related to self-incompatibility in ryegrass (*Lolium perenne* L.)”, PhD thesis, Ghent university

ISBN-number: 90-5989-093-0

Voorwoord

Na mijn studies, zeven jaar geleden, kon ik als IWT doctoraatsbursaal beginnen op het departement voor Plantengenetica en -veredeling. Vier jaar fundamenteel onderzoek was onvoldoende om resultaten te behalen in een complexe materie waarover nog weinig onderzoek bestaat. Gelukkig kreeg ik de mogelijkheid van prof. dr. ir. Erik Van Bockstaele om te blijven werken aan hetzelfde onderwerp. Dit werk is er dan ook gekomen met de steun van heel wat mensen die ik uitdrukkelijk wil bedanken voor hun financiële steun, begeleiding en labo-werk.

Een woordje van dank aan mijn beide promotoren die ervoor gezorgd hebben dat ik mijn doctoraat tot een goed einde kon brengen. Bedankt prof. dr. ir. Erik Van Bockstaele voor de kans die jij mij bood om op het Departement voor Plantengenetica en -veredeling de eerste stappen te zetten in het onderzoek. Dr. Isabel Roldán-Ruiz, bedankt voor de steun en de vele leerrijke discussies die we gevoerd hebben. Je bleef altijd met een enorme interesse dit werk nalezen, sturen en verbeteren.

De honderden gels en de vele data gegenereerd tijdens dit onderzoek kwamen er dankzij de medewerking van Ariane, Carina, Cindy, Katrien, Lien, Nancy en Sabine. Bedankt allemaal voor de hulp, het enthousiasme en dat jullie er altijd waren als ik terug eens afkwam dat er nog iets ‘dringends’ moest gebeuren.

Bedankt ook aan de collega’s van de groep van dr. ir. Johan Van Huylenbroeck en dr. ir. Joost Baert voor het onderhoud van de planten. Ik weet het, het waren zeker geen gemakkelijke planten, het ene jaar waren ze geliefd door de konijnen en het andere jaar deden de sproeistoffen te goed hun werk. Maar jullie bleven ze onderhouden met de nodige zorg en bedrijvigheid zodat de planten terug weelderig bloeiden.

Bedankt dr. ir. Jan De Riek voor het oplossen van al mijn computerproblemen. Het waren er best veel en het was altijd dringend. Bedankt ook aan al de andere collega’s in het labo en op mijn bureau voor de gezellige en plezante momenten. Die afleiding kon ik best eens gebruiken.

Ook een woordje van dank aan Els Coart, Veerle Lamote en Hilde Muylle, mijn vroegere bureaugenoten, voor de hulp bij het schrijven van mijn doctoraat. Als ervaren gedoctoreerden waren jullie opmerkingen heel welkom.

Ook buiten het DvP zijn er een aantal personen die een belangrijk aandeel hadden bij het realiseren van dit werk. Bedankt Cindy Martens (VIB) voor de hulp bij de verwerking van de sequenties, dr. Paul Van Hummelen voor de tijd die je vrijmaakte om mij wegwijs te maken in de verwerking van de microarray data en de GRASP partners. Marie-Claire bedankt voor het kopiëren en inbinden van mijn doctoraat.

Graag wil ik de juryleden bedanken voor hun interesse in dit werk, de nuttige opmerkingen en de opbouwende kritiek. Bedankt prof. dr. Wout Boerjan, dr. Wim Broothaerts, dr. Marc De Loose, dr. Oene Dolstra, prof. dr. Danny Geelen, prof. dr. Lieve Gheysen, prof. dr. ir. Dirk Reheul, dr. Danny Thorogood en prof. dr. ir. Georges Hofman.

Tenslotte een woordje van dank aan mijn familie en vrienden. Bedankt voor jullie steun en vertrouwen. Een dikke zoen speciaal voor Frank en mijn twee kleine kabouters, Lien en Michiel.

List of abbreviations

µl	microlitre
A	Mean of (log_2(R) * log_2(G))
aa	Amino acids
ABC	ATP binding cassette transport
ABI	Applied Biosystems
ABP	Actin-binding protein
AFLP	Amplified Length Polymorphism
AGP	Arabinogalactan protein
AP	Adapter primer
AQBC	Adaptive Quality-Based Clustering
ARC	Arm Repeat Containing protein
ATP	Adenosine triphosphate
AUN	Agricultural University of Norway
BC	Back-cross
BCD	Prefix of RFLP probes isolated from barley
BEN	Belgian EMBnet Node
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
Ca^{2+}	Calcium
cDNA	Copy DNA
CDO	Prefix of RFLP probes isolated form oat
CDPK	Calcium Dependant Protein Kinase
CLO	Centrum voor Landbouwkundig Onderzoek
CLO	Centrum voor Landbouwkundig Onderzoek
cM	Centimorgan
CP	Cross-pollinated
CP1h	Cross-pollinated pistils 1 hour after pollination
CP24h	Cross-pollinated pistils 24 hours after pollination
CP4h	Cross-pollinated pistils 4 hours after pollination
CP8h	Cross-pollinated pistils 8 hours after pollination
CRK1	Cytokinin-regulated protein
CSU	Prefix of RFLP probes isolated from maize
CTAB	hexadecyl trimethylammonium bromide
CV	Coefficient of variation
DD	Differential display
DEVD	Amino acid sequence: Asp-Glu-Val-Asp
DH	Double haploid
DIAS	Danish Institute of Agricultural Sciences
DNA	Deoxyribonucleic acid
dNTP	Desoxyribonucleosidetriphosphate
DTT	Dithiothreitol
DVP	Departement voor Planten veredeling en genetica
DvP	Departement voor Plantengenetica en -veredeling
ECM	Extracellular Matrix
EDTA	Ethylenediaminetetraacetic acid
ESA	European Seed Association
EST	Expressed Sequence Tag
EtBr	Ethidium bromide

G	Background corrected, normalized green fluorescence of the dye Cy3
GADPH	Glyceraldehydes 3-phosphate dehydrogenase
GLP	Germin like proteins
GRASP	Development of ryegrass allele-specific makers for sustainable grassland improvement
GTP	Guaninetriphosphate
h	hour
HSP	Heat Shock Proteins
HV	Hyper-variable
IGER	Institute of Grassland and Environmental Research
ILGI	International *lolium* Genome Initiative
INRA	Institut National de la Recherche Agronomique
IP_3	Inositol 1,4,5-triphosphate
K_3PO_4	Potassium phosphate tribasic
KCl	Kalium chloride
LG	Linkage group
LIA	Lithuanian Institute of Agriculture
Lp	*Lolium Perenne*
M	$Log_2(R/G)$
MAPK	Mitogen-activated protein kinase
$MgCL_2$	Magnesium chloride
MIPS	Munich Information Centre fir Protein Sequences
MQ	MilliQ water
mRNA	Messenger RNA
NADP	Nicotinamide Adenine Dinucleotide Phosphate
NCBI	National Centre of Biotechnology Information
ng	Nanogram
NP	Non-pollinated pistils
PCD	Programmed Cell Death
PCR	Polymerase Chain Reaction
PELP	Pistil-specific Extensin-Like Protein
ppm	Parts per million
PRI	Plant Research International
PSR	Prefix of RFLP probes isolated from wheat
QTL	Quantitative Trait Loci
R	Background corrected, normalized red fluorescence of the dye Cy5
RACE	Rapid Amplification of cDNA ends
RAPD	Random Amplified Polymorphic DNA
RAP-PCR	Random Arbitrarily Primed PCR
REC	Recombination frequency
Rfo	Raffinose oligosaccharides
RNase	Ribonuclease
rRNA	Ribosomal RNA
RT-PCR	Reverse Transcriptase PCR
SAGE	Serial Analysis of Gene Expression
SBP	S-proteins binding protein
SCR	S-locus Cysteine Rich
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SFB or SLF	S-locus F-box gene

SI	Self-incompatibility
SLG	S-locus Glycoprotein
SNP	Single Nucleotide Polymorphism
SP	Self-pollinated
SP1h	Self-pollinated pistils 1 hour after pollination
SP24	Self-pollinated pistils 24 hours after pollination
SP4h	Self-pollinated pistils 4 hours after pollination
SP8h	Self-pollinated pistils 8 hours after pollination
SPH	S-protein homologues
SRK	S-locus Receptor Kinase
SSC	Sodium citrate buffer
SSH	Suppression Substractive Hybridisation
SSR	Simple Seguence Repeat
STS	Sequence Tagged Sites
TDF	Transcript Derived Fragment
TE	Tris EDTA
THL	Thioredoxin-h-like protein
TIGR	The Institute for Genomic Research
Tris-HCL	Tris (hydroxymethyl)aminomethane hydrochloride
TTS	Transmitting-Tissue Specific
U	units
YVAD	Amino acid sequence: Tyr-Val-Ala-Asp

Summary

Self-incompatibility (SI) is a mechanism that prevents self-pollination and inbreeding in many flowering plant species. SI in *Lolium perenne* is controlled by two multi-allelic loci, S and Z. The working of this SI-mechanism has important consequences for *L. perenne* breeding as it prevents the efficient production of inbred lines and hybrids. The aim of this study was to identify genes involved in the SI-response in *L. perenne*. Therefore, different approaches were applied. We investigated first whether homologues of genes involved in the SI-response of other plant families and grass species (Solanaceae, Brassicaceae and *Phalaris coerulescens*) are also present in *L. perenne*. Based on the results, we could conclude that no S-RNase based system (active in members of the Solanaceae) is present in ryegrass. This result was not surprising as very little features with regards to the SI-response are common for grasses and species displaying a S-RNase based system. On the other hand, a *L. perenne* homologue of SRK, the key component at the female side in members of the Brassicaceae, was identified. However, this gene was not only expressed in pistils but also in leaves indicating that it is not the recognition factor of the SI-response in ryegrass. This conclusion was also supported by its map position, as it mapped on LG4 and not on LG1 or LG2, where S and Z are located. Finally, we amplified in pistils and leaves of *L. perenne* the thioredoxin homologue known to be involved in the SI-response of *Phalaris coerulescens*, also a member of the Poaceae. The *L. perenne* thioredoxin homologue mapped in the neighbourhood of S (at 3cM), indicating that it is not S. Whether this thioredoxin gene is involved in the SI-response in ryegrass needs to be investigated. However, the involvement of thioredoxins in SI has been demonstrated in *Brassica*.

Another objective of this study was to identify genes involved in the signalling cascade triggered by a SI-response in *L. perenne*. Therefore we used two differential expression techniques: cDNA-AFLP and the cDNA-microarray. The genome-wide expression analysis carried out by cDNA-AFLP allowed us to identify 479 TDFs (Transcript Derived Fragments) as putatively related to SI in *L. perenne*. All the expression profiles were quantified using AFLP Quantar Pro and clustered into groups of co-regulated genes. A subset of these genes was selected to study further. The selected genes clustered into 11 functional categories. Several of them corresponded to genes involved in general cellular processes but homology was also been found to proteins known to be involved in fertilization and SI-processes in

other plant families such as ubiquitin-related and calcium-related proteins. The GRASP microarray was used to confirm the expression profiles of the TDFs selected by the cDNA-AFLP and to identify other genes related to SI. This microarray contains the TDFs selected by cDNA-AFLP and cDNA clones of genes involved in different biological processes. Using this technique, we were able to identify 371 differentially expressed genes putatively involved in the SI-response. Surprisingly the results from both techniques were contradictory. Confirmation of a selection of genes by Real-time RT-PCR revealed consistency between the results of Real-time RT-PCR and those of the cDNA-AFLP but not with those of the microarray. Further investigation is needed to clarify this point.

As a complementary approach, the genome location of the genes identified as putatively involved in SI was determined. Two mapping populations were used for this purpose: the ILGI population and the CLO-DvP population. Some of the mapped genes clustered around S and Z on LG1 and LG2, respectively. Several other markers mapped across the different linkage groups with a higher concentration on LG3 and LG4. Also a strong association was identified between LG1 and LG3 indicating that some genes of LG3 interact with LG1 during SI.

Finally, combining all these data, a selection of 21 genes putatively involved in SI in *L. perenne* was made. Confirmation of their involvement is the subject for further research.

Samenvatting

Zelfincompatibiliteit (SI) is een mechanisme dat zelfbestuiving en inteelt bij bloeiende planten verhindert. De SI-respons in *Lolium perenne* wordt gecontroleerd door twee multi-allelische loci, S en Z. Het SI-systeem in raaigras bepaalt mee de veredelingsstrategieën aangezien een verplichte zelfbestuiving slechts een zeer beperkte zaadoogst oplevert wat het vormen van lijnen en hybriden bemoeilijkt. Toch zou de ontwikkeling van hybriden, waarbij het heterosis effect kan geëxploiteerd worden, heel wat voordelen kunnen bieden. Dit onderzoek had als doel genen op te sporen die rechtstreeks of onrechtstreeks betrokken zijn bij de regulatie van zelfincompatibiliteit bij *L. perenne*. Daarvoor werden verschillende strategieën gevolgd. Eerst gingen we na of er gelijkenissen waren tussen het SI-mechanisme actief in raaigras en reeds gekende mechanismen uit andere plantenfamilies en grassoorten (Solanaceae, Brassicaceae en *Phalaris coerulescens*). Daaruit konden we besluiten dat geen RNase gebaseerd systeem (bestudeerd in de Solanaceae) actief is in de raaigrassen. Dit resultaat was te verwachten aangezien het SI-mechanisme actief in de raaigrassen weinig gelijkenissen vertoont met een RNAse gebaseerd SI-mechanisme. Wel, werd in *Lolium perenne* een homoloog van het SRK gen geamplificeerd. Het SRK gen is de vrouwelijke herkenningscomponent van het SI-mechanisme in soorten van de familie van de Brassicaceae. Aangezien dit gen niet alleen tot expressie kwam in de stampers maar ook in de bladeren, kan aangenomen worden dat het niet de herkenningsfactor is van het SI-systeem in raaigras. Dit resultaat werd ook nog bevestigd door de kaart positie van het gen, namelijk op LG4 en niet op LG1 en LG2, welke S en Z bevatten. In bladeren en stampers van *L. perenne* werd het thioredoxine homoloog geamplificeerd. Thioredoxine is betrokken in het SI-mechanisme van *Phalaris coerulescens*, ook behorend tot de familie van de Poaceae. Het thioredoxine gen karteert 3cM van S, wat aangeeft dat het niet S zelf is maar vermoedelijk wel betrokken is in de SI-respons. Op welke manier dit gen betrokken is moet nog verder onderzocht worden.

Aangezien het SI-mechanisme een ingewikkeld proces is waarbij verschillende componenten betrokken zijn, werden expressiestudies uitgevoerd met behulp van de cDNA-AFLP en microarray techniek om genen op te sporen die een interessant differentieel expressieprofiel vertonen in zelfbestoven, kruisbestoven en niet bestoven stampers. Met behulp van de cDNA-AFLP techniek werden 479 fragmenten met een

interessant expressieprofiel geïdentificeerd. De bekomen data werd geanalyseerd met behulp van de AFLP QuantarPro software en de verschillende expressieprofielen werden vervolgens geclusterd in groepen van co-gereguleerde genen. Uit die clusters werd een subset van genen geselecteerd voor verder onderzoek. Die fragmenten werden gesequeneerd en werden geclusterd in 11 functionele categorieën. Er werden genen geselecteerd die homologieën vertoonden met genen betrokken in algemeen biologische processen maar ook met genen die betrokken zijn in bestuivingsprocessen en SI-mechanismen, zoals ubiquitine gerelateerde en calcium gerelateerde proteinen. Om de resultaten bekomen met cDNA-AFLP te bevestigen en om nieuwe genen betrokken in SI te identificeren werd een GRASP microarray experiment uitgevoerd. De microarray bevatte de fragmenten geselecteerd bij de cDNA-AFLP analyse en cDNA clones van genen betrokken bij verschillende biologische processen. Met behulp van die techniek werden 371 differentieel geëxpresseerde genen, die mogelijks betrokken zijn in het SI mechanisme van raaigras geïdentificeerd. De resultaten bekomen met beide technieken vertoonden tegenovergestelde expressieprofielen. Om na te gaan welk expressieprofiel correct was, werd voor een bepaald aantal genen een Real-time experiment uitgevoerd. Daaruit bleek dat de expressie profielen bekomen met de cDNA-AFLP gelijkaardig waren aan deze van de Real-time maar tegenovergesteld aan deze van de microarray. Verder onderzoek is nodig om na te gaan wat er precies gebeurd is en om dit probleem op te lossen.

Met behulp van de homologie gebaseerde strategie en de expressie studies werden een aantal genen geselecteerd waarvan de genoom locatie werd bepaald. Daarvoor werden twee mapping populaties gebruikt: de ILGI en de CLO-DvP populatie. Een aantal van die fragmenten karteerden rond S en Z op LG1 en LG2. Verschillende andere merkers karteerden op de andere linkage groepen, met een hogere concentratie op LG3 en LG4.

Uiteindelijk kon uit al die data een selectie gemaakt worden van 21 interessante genen die mogelijks een belangrijke rol spelen in de SI-respons van *L. perenne*. Bevestiging hiervan en de manier waarop ze de SI-respons beïnvloeden is onderwerp voor verder onderzoek.

List of abbreviations

Summary

Samenvatting

Contents

Part A

General introduction

Chapter 1

Background and objectives

1.1 Ryegrasses

Of the 129 million hectares of farmland in Europe, over one third is permanent grassland and for the greater part, ryegrass cultivars (*Lolium* spp.) are used (http://europa.eu.int/comm/agriculture/envir/report/en/concl_en/report.htm). *Lolium* spp. is an economically important crop, and 66% of the 219,000 tonnes grass seed that were produced in Europe in 2004 were of *Lolium* spp. (European Seed Association (ESA) and Semzabel).

The *Lolium* spp. belong to the family of the Poaceae (syn. Gramineae), subfamily Pooideae and tribe Poeae. The *Lolium* genus consists of eight species. From an economic point of view, *Lolium perenne* (perennial ryegrass) (Figure 1.1) and *Lolium multiflorum* (Italian ryegrass) are the two most important species of the genus. *L. perenne* and *L. multiflorum* are cool-season bunchgrasses native to Europe, but they are widely spread throughout the world, including North and South America, New Zealand and Australia. They are important forage species used as pasture and hay crop and they are best adapted to cold, moist climates where winter kill is not a problem (Hannaway et al. 1999 a&b).

Figure 1.1 *Lolium perenne*

Perennial ryegrass is important in forage/livestock production systems. It has high yield potential, is winter hardy and persistent. Perennial ryegrass is primarily grown for pasture and hay but also for turf. Because of its high quality, forage-type perennial ryegrass is used primarily for lactating dairy cows and for all classes of livestock, especially those with high nutrient requirements such as young, growing animals. The use of *L. perenne* as turf has increased in recent years with selection of more dense-growing, persistent and darker-green types, with improved disease resistance. *L. multiflorum* has lower stress resistance and lower persistence but has a higher yield potential. It is an important short-duration grass, used when fast cover or quick feed is required (Hannaway et al. 1999a).

1.2 Ryegrass breeding

L. perenne and *L. multiflorum* are outcrossing, wind-pollinated species that can be inter-crossed resulting in a fertile hybrid (*L. x hybridum* or *L. x boucheanum*). Selfing is largely prevented by the action of a gametophytic self-incompatibility system controlled by two multi-allelic loci, S and Z, which are located in different chromosomes, and thus segregate independently. The presence of an SI-mechanism in these species prevents the efficient production of inbred lines and hybrids. As a consequence, the breeding of ryegrasses involves recurrent selection of multiple parental clones. Selected clones are polycrossed to generate synthetic varieties (Reheul and Ghesquiere 1996, Van Bockstaele 1998, Forster et al. 2001). Traditional forage grass breeding programs have succeeded in improving a wide range of valuable traits. However, the production of hybrids would make it possible to exploit heterosis effects. The most important criteria for selection of forage types are rust resistance, resistance to leaf spot and *Xanthomonas*, better stress tolerance, persistence, digestibility, fast regrowth, palatability, nutritional quality, N-efficiency and seed production (Van Bockstaele 1998-1999). For turf varieties the most important selection criteria are disease resistance, tolerance to high temperatures, wear tolerance, recovering ability, darker-green colour, better mowing characteristics, drought resistance, fineness of leaves and slow growth (Grasgids 2000).

1.3 Biotechnology in ryegrass breeding

Classical breeding approaches have been the basis of forage improvement over the past half century. More recently, molecular tools have been developed which facilitate and speed up breeding activities

(Spangenberg et al. 2001). Trait improvement of turf grass and forage grass through genetic engineering has potential to be important to the industry and for the environment. Beneficial traits of grasses such as fungal resistance, drought and stress tolerance that will reduce water usage, insect and pest resistance, phyto-remediation of soils, horticultural qualities such as aluminium tolerance, stay-green appearance, and forage quality traits such as digestibility, nutritional quality and palatability can be improved through genetic engineering. Genetic modification of these traits should enhance economics, animal health and the environment. At present, reproducible and efficient genetic transformation and regeneration protocols are available for *Lolium* spp. (Spangenberg et al. 2001, Altpeter et al. 2003). However, the target modification of biochemical pathways for forage crop improvement requires knowledge of the pathways themselves at the enzymatic and underlying genetic levels. So in a first step it is necessary to identify the genes involved in the different pathways. It is in this area that current research activities concentrate.

However, cultivated *L. perenne* and *L. multiflorum* genotypes cross readily with wild and feral ryegrasses and occasionally with *Festuca* species (Giddings et al. 1997a). The main problem of introducing genetically modified wind-pollinated grasses in Europe is the fact that the transgenes would spread from transgenic plants to related crops or wild species which grow here. In this situation the transgenic construct would be introduced into new genetic backgrounds where its stability and expression are uncertain and pleiotropic effects unpredictable (Kareiva 1993). Giddings et al. (1997b) developed models to describe pollen dispersal from a *L. perenne* source to traps arranged in concentric circles around it. They found that transgenic pollen can land on plants several kilometres away. Given the potential for long-distance dispersal the use of guard rows around transgenic crops to “mop” up pollen seems unlikely to be entirely effective at preventing pollen dispersal away from the crop (Giddings 2000). Male sterility in transgenic grasses could provide a good tool to prevent gene flow from genetically modified plants and should facilitate the application of genetic engineering in producing environmentally responsible grasses with enhanced traits (Luo et al. 2003).

Another important molecular tool that can be implemented in ryegrass breeding programs is DNA marker technology. In the first place, DNA markers can help to determine genetic distances between gene pools (Roldán-Ruiz et al. 2000), optimise the construction of new breeding

populations and identify the best parents for hybrid or synthetic varieties (Humphreys 2000). DNA technologies also allow more targeted approaches to the selection and introgression of valuable genes from a range of genetic resources while retaining the integrity of valuable genetic backgrounds. Finally, genetic mapping and Quantitative Trait Loci (QTLs) analysis allow the location of genes associated with valuable but difficult to measure traits, such as components of forage quality. Once identified and tagged, the 'positive' allelic variants of these genes are relatively easy to monitor during recurrent selection and introgression (Humphreys et al. 2003). Molecular genetic marker systems have been developed for implementation in the breeding of ryegrasses. As a result, several linkage maps for *Lolium* spp. are now available at different institutes and based on populations which segregate for different traits (Hayward et al. 1994 and 1998, Bert et al. 1999, Jones et al. 2002 a and b, Armstead et al. 2002 and Muylle 2005; see Chapter 6 for an overview). The constructed maps are based on conserved Restriction Fragment Length Polymorphisms (RFLPs), isozyme markers and PCR-based marker techniques such as Random Amplified Polymorphic DNAs (RAPDs), Amplified Fragment Length Polymorphisms (AFLPs), Simple Sequence Repeats (SSRs) and Sequence Tagged Sites (STS) markers. The next generation of molecular markers for plants, and also for *Lolium* spp. is based on the development of primers from characterised expressed sequences and target loci with gene information (Faville et al. 2004). These "functional" markers have clear advantages in comparison to "anonymous" markers. As primer binding sites are designed in coding DNA which has a higher degree of sequence conservation, such EST markers are very useful for aligning linkage maps and comparing QTLs. Genetic mapping with ESTs thus enables a more rapid transfer of linkage information between related species (Cato et al. 2001). Functional markers which are related to specific traits such as self-incompatibility, flowering, seed production, disease resistance, leaf morphology, forage quality and cold resistance are being developed in the context of the EU project GRASP (see Chapter 7).

1.4 Relevance of self-incompatibility in ryegrasses and other members of the Poaceae

Self-incompatibility (SI), or the inability of a fertile hermaphrodite seed-plant to produce zygotes after self-pollination (de Nettancourt 1977) is a very important trait in *Lolium perenne* breeding, as although it prevents the efficient production of inbred lines and hybrids, it ensures the maintenance of heterotic gene combinations (a detailed description of this mechanism in

members of the Poaceae and of other plant families is given in Chapter 2). SI has intrigued breeders, geneticists and molecular biologists for decades. As a consequence, self-incompatibility has been studied in different plant families, including the Poaceae. The SI-mechanism in ryegrass has been thoroughly described in terms of classical genetics and from a cytological point of view (Lundqvist 1954 & 1961).

The grass family (Poaceae) includes economically and ecologically important species, including cereals, sugar cane and many feed and forage crops (Baumann et al. 2000). SI is widely distributed among the Poaceae and present in at least 16 genera (de Nettancourt 1977, Connor 1979). An overview of some economically relevant species is given in Table 1.1. Usually, self-incompatible and self-compatible species are found in the same genus, the frequency of self-incompatibility being higher in perennial than in annual species (Beddows 1931). The existence of a distinct self-incompatibility system in the grasses has been observed in many species. Lundqvist (1954) first reported that self-incompatibility of *Secale cereale* was under gametophytic control of two unlinked multi-allelic loci, S and Z. Two years later, Hayman (1956) confirmed the S-Z system in *Phalaris*. Up to now, the S-Z system has been found in many other species but the genetic studies are confined so far to the tribes of Aveneae, Poeae and Triticeae belonging to the Pooideae subfamily (Baumann et al. 2000).

Table 1.1 Overview of self-compatible and self-incompatible species, belonging to the family of the Poaceae (Li et al. 1997).

Subfamily	Tribe	Species	Self-incompatible
Bambusoideae	Oryzae	Oryza sativa	No
Panicoideae	Andropogoneae	Zea mays	No
Pooideae	Poeae	Festuca pratensis	Yes
		Lolium rigidum	Yes
		Lolium perenne	Yes
		Lolium multiflorum	Yes
		Lolium temulentum	No
	Aveneae	*Phalaris coerulescens*	Yes
	Triticaceae	Triticum aestivum	No
		Hordeum bulbosum	Yes
		Hordeum vulgare	No
		Secale cereale	Yes

As the SI-mechanisms are conserved among members of the same plant family (Nasrallah 2005 and references therein) the results obtained for one given species of the Poaceae, would also contribute to a better understanding of the SI-response in other species of this family. This knowledge can be used in two directions. On the one hand, self-incompatible species could be made self-compatible, making it possible to construct inbred lines and hybrids. On the other hand, self-compatible crop plants, can be made self-incompatible by the introduction of the genes controlling self-incompatibility. This would facilitate the production of hybrid seed and, at the same time, reduce labour costs.

1.5 Objectives and general outline of this thesis

The molecular mechanisms responsible for the self-incompatibility response in ryegrass and other species of the Poaceae has had little attention in the past, with the remarkable exceptions of some studies in *Phalaris coerulescens* and *Secale cereale* (rye). Molecular research in *P. coerulescens* resulted in the identification of a thioredoxin gene, which was initially thought to represent the S-gene, but further studies demonstrated that the thioredoxin gene was located 2 cM from S (Langridge et al. 1999). Recently, two DNA-markers derived from expressed sequences have been found to cosegregate with S in *P. coerulescens* and another DNA-marker maps 0.9 cM from Z (Bian et al. 2004). Also recent work in rye has resulted in the identification of one STS marker (TC116908) which cosegregates with Z (Hackauf and Wehling 2005). The gene from which this marker was derived was strongly expressed in rye pistils but not in leaves. However, the implication of any of these genes in the SI-response has not been demonstrated yet.

It is beyond any doubt that a better understanding of the recognition and rejection processes between stigma and pollen would be of relevance for the development of improved breeding strategies for commercially important crops such as *Lolium*. **The main objective of this thesis is to perform a detailed study of the molecular basis of the SI-mechanism in *L. perenne*.** As mentioned above, the results obtained will also contribute to a better understanding of the SI-response in other species of the Poaceae.

Among the different strategies that can be followed to identify genes or gene products controlling specific processes and mechanisms in plants, two complementary approaches were selected to be used in this thesis. In the

first place, a homology-based strategy was applied to determine whether the kind of genes involved in the SI-response in other plant families play also a role in the SI-response in *L. perenne* and to eventually isolate these gene homologues in the ryegrass genome. The results of this approach are presented in Chapter 3. In the second place, a genome-wide gene expression analysis was carried out (Chapter 4). To this purpose, genes which are involved in the signalling cascade which follows a SI-reaction were identified using the cDNA-AFLP technique. Relevant cDNA-AFLP fragments were sequenced and classified into functional clusters (Chapter 5). Thereafter, the map position of genes identified by the two strategies was determined in the ILGI reference population (i.e. population with characterised SI-genotypes, Thorogood et al. 2002) and the DvP-CLO reference population (Muylle et al. 2005). The results of the mapping work are presented in Chapter 6. Finally, a cDNA microarray experiment was carried out to eventually confirm the expression profiles obtained using the cDNA-AFLP technique and to check whether other genes present in the *L. perenne* microarray used are involved in the SI-response (Chapter 7). Chapter 8 summarises the general conclusions of the study and outlines future perspectives for research.

Before embarking in the description of the methodology applied and the results obtained, a summary is presented of our present knowledge of fertilization and SI-mechanisms and factors in different plant families (Chapter 2).

Chapter 2

Reproduction system and barriers to fertilization

2.1 Pollination and fertilization

At anther dehiscence, the pollen grains of most flowering plants are binucleate, containing a vegetative cell and a generative cell and are highly dehydrated. The second mitotic division of the generative cell to produce two sperm cells occurs after germination on the style. In some plant families, like the Poaceae, this mitotic division takes place before the pollen is shed from the anthers. When the pollen grain is released from the anther it can be carried by wind, insects or other agents to the stigma of the same or other plants. If the stigma is a compatible one, the pollen grain adheres to the papillae cells of the stigma, hydrates and germinates. The result is a pollen tube that transports the sperm cells to the embryo sac.

Pollination can be divided into characteristic stages that include: 1) pollen adhesion, 2) pollen hydration and germination, 3) pollen tube growth through the style and guidance to the ovule micropyle and 4) delivery of the sperm to the embryo sac. The time required for these events to take place varies considerably between species: from only a few hours (e.g. *Brassica*) to some months (e.g. pines) (Wilhelmi and Preuss 1997). The whole process, from pollen delivery, to fertilization is highly regulated on both the male and the female sides and the concerted work of numerous factors is required to ensure successful fertilization, as illustrated below.

2.1.1 Pollen adhesion

Stigmas are of two broad types: wet and dry, depending on the amount of secreted matrices that the pollen encounters on their surface. The carbohydrate and lipid-rich matrix on the surface of wet stigmas, such as those found in the Solanaceae, promotes the adhesion and hydration of most pollen species. In contrast, control of pollen acceptance by adhesion represents a critical step in the whole fertilization process in species with a dry stigma such as members of the Brassicaceae (Zinkl and Preuss 2000). In species with a dry stigma, the pollen coat is essential for the recognition and adhesion of the pollen to the stigmatic surface. The pollen coat is an extracellular wall of long-chain lipids and proteins which are recognized by the stigma (Elleman et al. 1992, Zinkl and Preuss 2000, Heizmann et al. 2000, Mayfield et al. 2001). This coating facilitates communication in plants with dry stigmas, providing a function similar to the lipid-rich exudate on the surface of wet stigmas. Most likely, each species has its own set of male and female adhesion factors and identification of these components will reveal

how diversity among adhesion molecules determines the choice of mating partners.

2.1.2 Pollen hydration and germination

The water content of the pollen grain approximates 20% when it lands on the stigma (Dumas and Gaude 1982). After adhesive contacts have been established, the pollen grain rehydrates by acquiring water from the stigma exudates or from the stigma papillary cells. In plants with wet stigmas the lipids are supplied by the hydrophobic stigma coating (the exudates) (Johnson and Preuss, 2002). Hydration of the pollen grain is aided by the presence of the pollen coat in species with a dry stigma, which in *Arabidopsis* contains predominantly lipases and oleosins (Mayfield et al. 2001). The pollen coat lipids probably play a role in establishing a conduit that allows transport of water from the stigma. In *Arabidopsis cer* mutants, carrying deficiencies in pollen coat lipids hydration defects have been described (Hülskamp et al. 1995). It has also been found that the loss of an oleosin protein (a lipid binding protein), the most abundant pollen coat protein, from the coat in *Arabidopsis* results in a three-fold delay in pollen hydration (Wolter-Arts et al. 1998). The delay is due to a failure of pollen to interact with the stigma rather than failure to absorb water. Similarly, mutations that eliminate the lipid-rich exudates of wet *Nicotiana* stigmas result in female sterility, and addition of lipids to the surface of these stigmas restores fertility (Wolter-Arts et al. 1998). The most likely function of lipids in the hydration process is to form a water-tight seal between pollen and stigma, facilitating the rapid transport of water through channels in the stigma and pollen membranes.

However, lipids are not the sole factors important in pollen hydration. Other factors which control the water flow from the stigma to the pollen grain play also an important role in pollen hydration. For example in *Brassica*, aquaporin-like proteins in the stigma, named MIP-MOD, act as water-transporting molecular channels to accomplish this (Dixit et al. 2001).

Following hydration, reorganisation of the pollen grain cytoplasm determines the polarity of pollen tube extension, which, in many species, emerges through the pore adjacent to the stigma surface. Although the nature of the signals that direct polarized growth is not understood yet, there is evidence that in species with a dry stigma cell polarity is organised in

response to a gradient of water established by pollen coat lipids (Wolters-Arts et al. 1998).

2.1.3 Pollen tube growth

Pollen tubes grow between the papillar cells and migrate into the pistil at rates that depend on the species and the genotype studied but that can approach 1 cm/hour in maize (Barnabas and Fridvalczky 1984). The pollen tube grows through an extracellular matrix that consists of lipids, proteins and carbohydrates. It is believed that these components provide the signals that are critical for directing and supporting pollen tube growth.

The nature of the pollen tube growth is totally different from that of other plant cells, in which growth takes place over the entire surface of the cell. In contrast, pollen tube growth is restricted to the tip region. Vesicles and cell wall components are transported to the growth tip via highly active “cytoplasmic streaming” and are incorporated into a zone of elongation in the apical domain of the pollen tube tip. During growth the volume of the tube remains relatively constant as “plugs” of callose are periodically deposited in the cytoplasm (Figure 2.1).

A growing pollen tube consists of three different parts (Figure 2.1): zone 1) the older non-growing part with callose plugs in the cytoplasm, zone 2) the cytoplasm with callose deposition in the cell wall and carrying the vegetative nucleus, the generative cell and many organelles and zone 3) the growing tip without callose deposition (Franklin-Tong 1999).

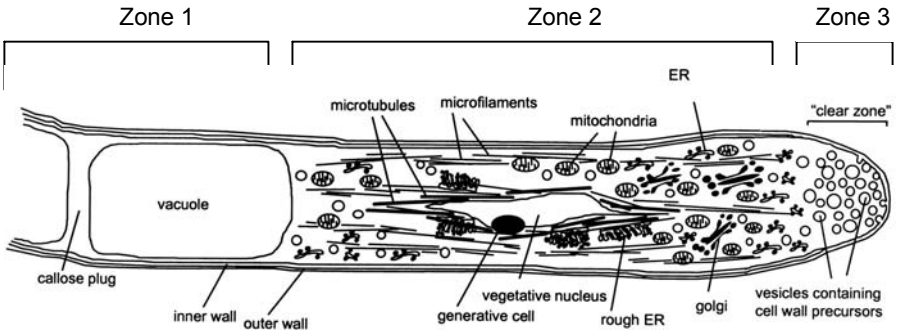


Figure 2.1 A diagrammatic representation of the basic structure of a pollen tube (Franklin-Tong 1999).

It is now clear that a tip-focused calcium ion concentration [Ca^{2+}] gradient, increasing towards the tip, and an intact actin cytoskeleton are crucial for the polarized tip growth and the movement of organelles in growing pollen tubes (Franklin-Tong 1999). This gradient is absent in non-growing pollen tubes and disruption of the gradient inhibits pollen tube growth. The cytoskeleton plays a key role in the delivery of vesicles at the growing tip. It consists of microtubules (polymers of tubulin), actin filaments and myosin. The microtubules form the rigid support on which the network of actin filaments rests. A dense network of actin microfilaments is found in the subapical region (zone 2) of the tube, parallel to the pollen tube axis whereas the tip of the tube contains an unstructured mass of relatively few shorter actin filaments (Cai et al. 1997).

Plants have probably a variety of ways to accept compatible pollen and guide the sperm cells to the ovule. To date there is no single system that is thoroughly understood, but studies in different plant species reveal that the pistil plays an active role in pollen tube guidance and that some mechanisms involve calcium, proteins and glycoproteins (Lord 2003). Calcium has long been proposed as a chemotrophic factor in pollination. An apical [Ca^{2+}] gradient is essential for polarized pollen tube growth. This [Ca^{2+}] gradient is likely to be maintained by both the influx of Ca^{2+} through Ca^{2+} channels located at the tip of the pollen tube and by Ca^{2+} release from other sources, such as the endoplasmatic reticulum (Franklin-Tong 1999). The signals involved in the Ca^{2+} release remain, for the most part, unidentified but there is evidence that a functional phosphoinositide pathway operates in growing pollen tubes of *Papaver rhoeas* (Franklin-Tong 1999). The flow of Ca^{2+} from intracellular stores can be the result of stimulation by IP_3 (inositol 1,4,5-triphosphate), which could, in turn, be triggered by extracellular signals (Cai et al. 2000). The role of this Ca^{2+} gradient in pollen tube growth has not been completely unravelled yet. A current model proposes that Ca^{2+}, through a series of intermediates, promotes the fusion of secretory vesicles in specific areas of the tip membrane (Malho and Trewavas 1996). Ca^{2+} is probably also able to modulate the organization of actin filaments in the apical/subapical regions of the pollen tube and can therefore participate in establishing the polar organization of organelles. However, the molecular mechanism at the basis of this finding is unclear yet (Cai et al. 2000).

In general, pollen tubes grow within the style through the stylar transmitting tissue (a nutrient-rich matrix secreted by the female cells, also called Extracellular Matrix, ECM). The pollen tube traffic in the style is

controlled by the female tissue through a matrix-driven mechanism that guides the tubes to the micropyle. The ECM is comprised of a complex mixture of proteins and varies depending on the species studied. Arabinogalactan proteins (AGPs), proline-rich glycoproteins and extensin-like proteins are abundant in the styles of many species. AGPs are highly glycosylated, hydroxyproline-rich glycoproteins implicated in various aspects of plant growth and development. Two classes of AGPs have been isolated in *N. tabacum*, the transmitting–tissue specific (TTS) proteins (Cheung et al. 1995, Wu et al. 1995) and the class III pistil-specific PELP proteins (PELPIII) (de Graaf et al. 2003). The TTS proteins adhere to the cell wall and to the tip of the growing pollen tube. A portion of the TTS are deglycosylated by pollen tubes and the sugars are incorporated into the pollen tube wall. Within the transmitting tissue TTS proteins display a gradient of increasing glycosylation from the stigmatic end to the ovarian end of the style, coincident with the direction of pollen tube growth. The fact that pollen tubes incorporate stylar molecules during passage down the style has led to the proposal that the stylar ECM is a source of nutrients for the fast-growing pollen tube but the exact function of TTS in pollen tube guidance is still unknown (Cheung et al. 1995, Wu et al. 1995, Wheeler et al. 2001). The PELPIII proteins are able to penetrate the pollen and accumulate in the style transmitting tissue from the early stage of development until pollination.

All these findings support the existence of a contact-stimulated guidance system that facilitates pollen tube growth through the pistil. The exact mechanism is still unknown.

2.1.4 Fertilization

After a pollen tube arrives at the micropyle, it breaks and releases two sperm cells, one that fertilizes the egg cell and a second one that fertilizes the central cell. The identification of the guidance signals that are important for this process is difficult because it occurs very rapidly. It has been suggested that chemo-attraction is important. Attractants might come from a synergid cell in the female gametophyte that, in many cases, begins to degenerate just before the pollen tube arrives. Actin filaments from this degenerating cell may participate in propelling the sperm cells to their targets (Wilhelmi and Preuss 1997, Figure 2.2).

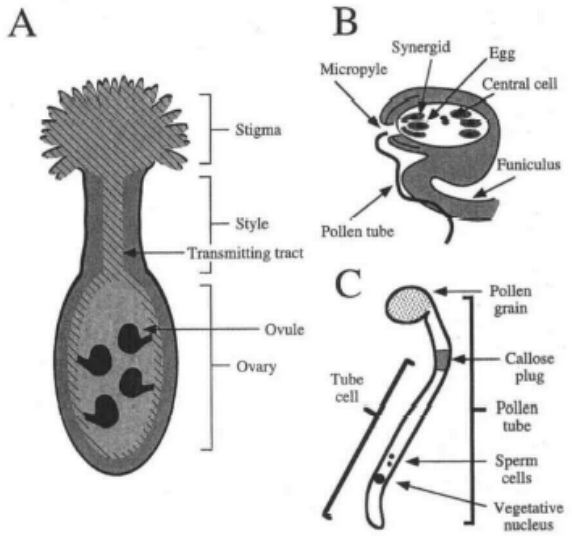


Figure 2.2 Reproductive structures in flowering plants. A. The pistil contains a stigma, style and basally located ovary that encases the ovules. Pollen tubes travel from the stigma to the ovary through the transmitting tract. B. The ovule contains the female gametophyte composed of an egg, a 2n central cell, two synergids that lay next to the egg and three antipodal cells that are opposite to the micropyle. Pollen tubes travel up the funiculus and enter the ovule through the micropyle. There they penetrate the synergid cell and deliver two sperm cells that eventually fuse with the egg and central cell, forming the zygote and 3n endosperm. C. Pollen grains, the male gametophyte, carry the two sperm and the vegetative nucleus to the female gametophyte by forming a pollen tube. Plugs of callose separate older parts of the tube cell at the growing tip (Wilhelmi and Preuss 1997).

2.2 Self-incompatibility

2.2.1 Introduction

The interactions described above, between pollen and female sporophytic tissues, are restricted to the events occurring in compatible pollinations. However, the angiosperms have developed several mechanisms to discriminate between desirable and undesirable pollen. These barriers are the result of incompatibility and incongruity. The incongruity model proposed by Hogenboom (1973) suggests that interspecific incompatibility is the result of barriers determined by evolutionary, physiological and/or morphological divergence between species. It occurs in interspecific crosses as a result of lack of genetic

congruity between the parent plants, necessary to complete pre- and post-fertilization processes and can be seen as an passive process. In contrast, self-incompatibility is mainly a pre-fertilization barrier.

In many interspecific crosses, inhibition of pollen tube growth operates in one direction, whereas the reciprocal cross is successful. This unilateral inhibition has been called 'unilateral incompatibility' (de Nettancourt 1977).

Intraspecific incompatibility, generally referred to as SI, is the focus of this PhD thesis. It can be defined as the inability of a fertile hermaphrodite seed-plant to produce zygotes after self-pollination (de Nettancourt 1977). Self-incompatibility prevents inbreeding in flowering plants and promotes outcrossing (de Nettancourt 1977). The basis for this mechanism is a cell-cell recognition between pollen and stigma. SI is not due to physical barriers in the style, but the stigma can distinguish between self (related) or foreign (unrelated) pollen. Only non self pollen of the same species is able to realize fertilization. Own pollen, self-incompatible pollen, is rejected at some point in the pollination process. Pollen rejection due to SI may happen at hydration, germination or during pollen tube growth through the style, depending on the species considered. This already indicates the existence of a variety of mechanisms in different SI-systems.

The SI-systems can be divided into two major classes, gametophytic and sporophytic, based on whether the pollen behaviour is determined by its own haploid genotype or by the genotype of the pollen producing mother plant (Figure 2.3). In gametophytic SI, the pollen phenotype is specified by its own haploid S-genotype. This mechanism has been well studied in members of the Solanaceae, Rosaceae and Papaveraceae. In sporophytic SI, the pollen phenotype is determined by the diploid S-genotype of the pollen producing plant. This indicates that the SI-gene products are carried by the pollen coat, which is derived from the tapetum and has therefore maternal tissue. The best studied sporophytic SI-system is that of the Brassicaceae.

In most species the SI-mechanism is controlled by a single multi-allelic S-locus, which contains genes encoding at least a stylar component and a pollen component. As described later, each of the SI-specificity-determining S-loci of the three single-locus SI-systems examined to date is a complex of at least two genes, and in all cases all S-locus variants, classically termed S-alleles, are more accurately referred to as S-haplotypes (Nasrallah, 2005). Some species represent an exception to this rule. The recognition

mechanism in the stigmatic gametophytic SI-system of the grasses is controlled by two multi-allelic loci, S and Z (Lundqvist 1954). In some other species, like *Beta vulgaris* and *Ranunculus acris* (Lundqvist et al. 1973) even up to four polyallelic genes participate in the determination of the gametophytic SI-system. These genes are unlinked and complementary in their effects. While the S-locus (S and Z loci in Poaceae) encodes the determinants of specificity in SI, in most species additional factors are also required. For example, in Brassicaceae and Solanaceae, factors that are not linked to the S-locus are required to give an SI-response. These factors are sometimes referred to as modifier genes because they influence SI (i.e. modify the response). Modifiers can be placed in three groups based on their interaction with specificity determinants (McClure et al. 2000). Group 1 factors directly affect the expression of genes that determine specificity. The factors belonging to group 2 interact either genetically or biochemically with the specificity determinants and are required for pollen rejection but have no role in pollination. Group 3 factors include genes that function in pollen rejection and in other pollen-pistil interactions as well. Different SI-systems evolved independently and are based on highly different mechanisms of recognition and rejection. In the following paragraphs a summary is given of our present knowledge of SI-systems in different plant families.

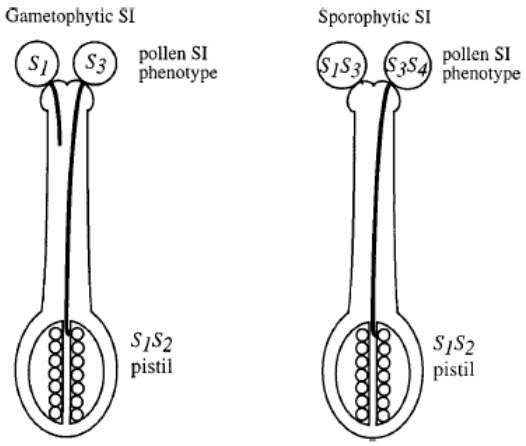


Figure 2.3 Illustration of the genetic basis of gametophytic and sporophytic self-incompatiblity. For the gametophytic type, the SI-phenotype of pollen is determined by the S-haplotype of its haploid genome, thus each pollen grain carries the product of one S-haplotype. For the sporophytic type, the SI-phenotype of pollen is determined by the S-genotype of its diploid parent, thus each pollen grain carries the products of two S-haplotypes. In both types, matching of one S-haplotype between pollen and pistil results in rejection of pollen (McCubbin and Kao 2000).

2.2.2 Sporophytic systems

In plants in which SI is sporophytically controlled, the SI-response is initiated when a stigmatic papillar cell receives a signal from the pollen coat (Stephenson et al. 1997). The growth of any incompatible pollen that succeeds in penetrating the stigmatic surface halts, following the deposition of callose produced in the stigmatic papillar cells. The mechanism of sporophytic self-incompatibility has been extensively studied in the Brassicaceae. In this family, the SI-mechanism is controlled by a complex, single multi-allelic S-locus. A number of genes have been shown to be physically linked to the S-locus: the S-locus Glycoprotein (SLG) (Nasrallah et al. 1985 and 1987), the S-locus Receptor Kinase (SRK) (Nasrallah et al. 1994, Stein et al. 1991) and S-locus Cysteine Rich protein (SCR, also called SP-11) (Schopfer et al. 1999). It has been demonstrated that SLG and SRK are highly expressed in stigmatic tissues and that SCR (SP11) is the male component of the SI-recognition system (Takasaki et al. 2000, Takayama et al. 2000, Shiba et al. 2001).

SRK is a membrane-associated protein with an extracellular, receptor domain, a transmembrane domain and a cytoplasmic serine/threonine kinase domain. The majority of sequence variation between SRK alleles lies within the receptor domain, where regions of hyper-variability are predicted to be responsible for S-specificity. Transgenic gain-of-function experiments have shown that a functional SRK alone determines S-specificity in the stigma (Takasaki et al. 2000).

SLG is an abundant glycoprotein secreted into the papillar cell wall matrix, similar in nucleotide sequence to the extracellular domain of SRK, for a given haplotype. SLG is highly polymorphic between different S-haplotypes but is not required for pollen recognition. SLG seems to enhance the SI-phenotype by stabilising SRK rather than being absolutely required for pollen rejection (Dixit et al. 2000, Takasaki et al. 2000). This was confirmed by Suzuki et al. (2003), who found that the S_{32}, S_{33} and S_{36} haplotypes from *Brassica rapa*, which lack SLG in their genome, display a normal SI-phenotype. However, as the S-haplotypes lacking SLG are a minority in *Brassica*, SLG must have some function in the SI-system. Nucleotide sequencing reveals that SRK and SLG genes can be classified into two groups that correlate with phenotypic dominance/recessivity relationships. Class-I alleles exhibit a strong SI-phenotype and are generally dominant or codominant, whereas class-II consists of three recessive (S_2, S_5 and S_{15})

alleles which exhibit a weaker SI-phenotype. Alleles within each class exhibit approximately between 90% and 70% homology with members of the other classes.

SCR (SP11) is a small (6-8 kDA), basic, cysteine-rich protein mainly expressed in anther tapetum cells and which accumulates in the pollen coat during pollen maturation (Kachroo et al. 2001). This is consistent with the sporophytic control of the SI-response in *Brassica*. The S-receptor domain of SRK binds the pollen ligand (SCR) in an allelic-specific manner such that the interaction only occurs when both SCR and SRK are from the same haplotype (Kachroo et al. 2001, Takayama et al. 2001). The structure of SP11 is similar to that of plant defensins except for loop 1. The sequences in loop 1 are more variable in length and therefore designated as the hyper-variable (HV) region that could serve as a specific binding site for the stigma receptor in an allelic-specific way (Mishima et al. 2003).

The signalling process downstream of SCR-SRK recognition is yet poorly understood. Till now three proteins putatively involved in the signal cascade have been isolated: THL1, THL2 and ARC1 (Arm Repeat Containing 1). THL1 and THL2 are thioredoxin-h-like proteins that can bind to the cytosolic domain of SRK in a specific and reversible manner, preventing spontaneous activation of the self-incompatibility signalling pathway (SRK autophosphorylation). Therefore, these thioredoxins act as inhibitors of SRK in the absence of a SCR pollen ligand (Cabrillac et al. 2001). The ARC1 component can interact with the kinase domain of SRK in a phosphorylation dependent manner. Suppression of ARC1 expression by an antisense transgene results in partial breakdown of SI, confirming that ARC1 has a positive effect of SI-signalling (Stone et al. 1999).

A model for the SI-response in *Brassica* has been postulated by Hiscock and Mclennis (2003, Figure 2.4). When a self-pollen lands on the stigma, SCR/SP11 penetrates the papilla cell wall and binds to the SRK/SLG receptor complex on the papilla cell membrane. This ligand-receptor interaction and consequent receptor activation occurs in an S-haplotype-specific manner. The binding induces autophosporylation of SRK, which triggers the signalling cascade that results in the rejection of self-pollen. Autophosphorylation of SRK induced by SP11 presumably causes subsequent phosphorylation of intracellular substrates of SRK, like ARC1. This appears to be the first step in an intracellular signalling cascade within

the papillar cell. The next steps of the cascade and how they finally result in the inhibition of pollen tube growth still remain unclear.

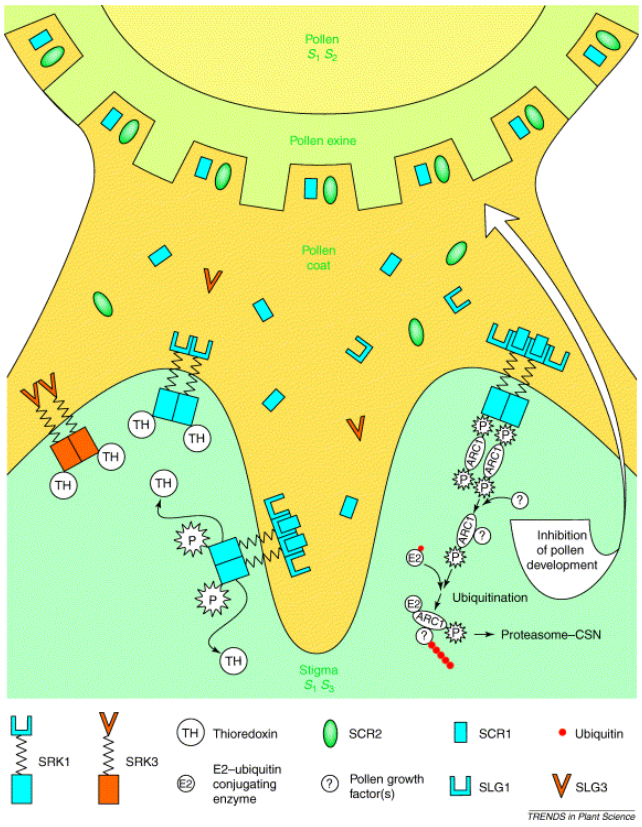


Figure 2.4 Model for the mechanism of S-haplotype-specific pollen recognition in *Brassica*. Proteins encoded by the same S-haplotype are given the same form. The pollen SCR_1 (/$SP11_1$) protein is recognised by, and binds to the extracellular receptor domain of SRK_1, presumably inducing dimerization between SRK_1 proteins and transphosphorylation on serines and threonines in their kinase domains. Activation of the SRK complex, which includes SLG in most S-haplotype-specific interactions, initiates an intracellular signalling cascade within the stigmatic papilla cell that culminates in localized rejection of pollen. SRK and SCR/SP11 alone are required for S-haplotype-specific pollen recognition and rejection, but the strength of the SI-response is enhanced by the presence of SLG in S-haplotypes that possess this protein. Signalling downstream of the activated SRK complex has yet to be characterized. However ARC1, an arm-repeat protein of yet unknown function, binds to the kinase domain of SRK in a phosphorylation-dependent manner and is essential for SI (Hiscock and Mclnnis 2003).

2.2.3 Gametophytic systems

In contrast to *Brassica* SI, the SI-response in the Solanaceae and poppy families targets the invading self-pollen tubes for destruction, and the pathway that leads to the pollen inhibition is triggered within the pollen tube, rather than within the cells of the pistil. Two distinct mechanisms to recognize and destroy self-pollen tubes are used (Nasrallah 2005), which are described in more detail in the following sections.

2.2.3.1 S-RNase based SI

This system is typical of species with bi-nucleate pollen and wet stigmas that receive exudates rich in the arabinogalactans necessary for pollen-tube growth through the style. It is typically characterised by an S-ribonuclease gene and SI involves the degradation of pollen RNA. In the families Solanaceae, Rosaceae and Scropulariaceae*,* which possess an S-**RNase**-based SI-system, both compatible and incompatible pollen grains germinate and begin to grow through the transmitting tract of the style. The growth of an incompatible pollen tube is arrested in the upper third segment of the style. The S-RNases are localised in the stigma and in the transmitting tract of the upper third segment of the style, sites that coincide with the zone of pollen tube inhibition in species displaying a S-RNase based SI-system. Gain-of-function and loss-of-function studies in *Petunia inflata* (Lee et al. 1994) and in *Nicotiana alata* (Murfett et al. 1994) revealed that catalytically active S-RNases in the pistil are necessary and sufficient for recognising and rejecting incompatible pollen. Thus, it is currently accepted that S-RNases are allele-specific cytotoxins.

S-RNases are glycoproteins of ca. 30 kDa, with one or more N-linked glycan chains. A large number of cDNAs of S-RNase proteins from several Solanaceae species have been isolated and sequenced to date. Sequence comparison has revealed a common structure, with five highly conserved regions (C1 to C5) and two highly variable regions (HVa and HVb) (Ioerger et al. 1991, Tsai et al. 1992). C2 and C3 show sequence similarity with the corresponding regions of fungal RNases (RNase T2 and RNase Rh). The experiments of Kuranandaa (1994) in *Petunia inflata* demonstrated that the S-haplotype specificity determinant of S-RNases resides in the backbone itself and not in the glycan side chains. The two hypervariable regions (Hva and HVb) seem to play a key role in S-haplotype specificity but, it remains

possible that amino acids outside HVa and HVb are also involved in S-haplotype specificity (Verica et al. 1998).

The characteristics of the pollen S-protein in species displaying an S-RNase based SI-system remains unclear. It is probably a gene other than the S-RNase, but genetically linked to it. Several candidates have been isolated and characterised to date. The strongest pollen candidate at present is an F-box gene, named **SLF or SFB** (S-Locus F-box gene). SFB was first identified in *Anthirrhinum hispanicum* (of the Scrophulariaceae) as a pollen-expressed, S-linked gene on a BAC clone carrying also the S_2-allele (Lai et al. 2002). The AhSLF F-box gene meets several of the key characteristics expected from the product of the pollen S-gene: (i) it is located at the S-locus; (ii) it exhibits S-haplotype-specific polymorphism; (iii) it is specifically expressed in the pollen. Also Yamane et al. (2003) identified in *Prunus* an F-box protein gene, SFB, with S-haplotype-specific polymorphism and Sijacic et al. (2004) in *Petunia inflata*, namely PiSLF. Sonneveld et al. (2005) identified two self-compatible pollen-part mutants in *Prunus* associated with a loss of function of the haplotype-specific SFB genes. These findings provide additional support that the SFB gene is the pollen S-gene in *Prunus* and other species with S-RNase based SI.

In general, the function of F-box proteins is to promote ubiquitin-mediated protein degradation together with ubiquitin-activating enzymes (E1s) and ubiquitin-conjugating enzymes (E2s). They are thus likely to interact with the 'non-self' S-RNases and send them down a degradation pathway, thereby preventing arrest of compatible pollen tube growth. There must be some mechanism to distinguish between 'non-self' and 'self' S-RNases, whereby in an incompatible situation, the 'self' S-RNases would not be targeted to the ubiquitination pathway and so be left to degrade incompatible pollen RNA and thereby affect the arrest of pollen tube. Recently, it has been shown that AhSLF-S_2 (one of the allelic forms of the S-locus F-box protein of *Anthirrhium hispanicum)*, is a likely component of the Skp1/Cullin or CDC53/F-box (SCF) complex. This complex is composed of Skp1, Cullin1, Rbx1 and an F-box protein and is involved in ubiquitin-mediated protein degradation by the 26S proteosome. Interestingly, AhSLF-S_2 interacts with both self and non-self S-RNases but appears to mediate degradation of only non-self S-RNases (Qiao et al. 2004). However, it is still unclear how the non ubiquitination of self-RNases is achieved. Compared to mammalian systems, 'self' S-RNases are probably protected from ubiquitination and

subsequent degradation, by binding to their cognate F-box protein as a pseudosubstrate.

In *Petunia hybrida* the pollen protein PhSBP1 seems to be involved in ubiquitin-mediated protein degradation and has also been found to interact with S-RNases (in a Yeast two-hybrid system). PhSBP1, contains a RING-HC finger domain at the C-terminus, found in many proteins that function as E3 ubiquitin ligases (Sims and Ordanic 2001).

Also in *Brassica* the stigma protein ARC1 (see above) has been demonstrated to possess E3-ubiquitin ligase activity. Thus, it seems increasingly likely that ubiquitination, and presumably protein degradation, is a key component of both sporophytic and gametophytic SI-systems (Franklin-Tong and Franklin, 2003).

Other factors putatively involved in the S-RNase mediated SI-response have been identified to date. For example, Dowd et al. (2000) identified 13 pollen expressed genes, which lie within 1 cM of the S-RNase gene in *Petunia inflata*. McClure et al. (1999) isolated a cDNA, designated **HT**, which is present in the SI species *N. alata* but not in *N. plumbaginifolia*, a closely related but self-compatible species. HT encodes an asparagine-rich protein of 101 residues. HT-transcript accumulation lags slightly behind the S-RNase transcript, but both reach their maximum levels at anthesis. Transgenic experiments revealed that HT is essential in the SI-response, as plants with reduced levels of HT-protein are defective in S-allele-specific pollen rejection. This effect is not mediated by a reduction in S-RNase expression, and HT has been classified as a group 2 factor following McClure et al. (2000). Although HT-proteins are implicated in S-allele-specific pollen rejection, their exact function remains unknown.

There exist two models, the receptor and the inhibitor model, for S-haplotype specific inhibition of pollen tubes in S-RNase based SI-systems (Figure 2.5). Both propose that S-RNases act as intracellular toxins by degrading pollen tube RNA, although they suggest very different roles for the pollen S-product. According to the first model, the pollen component is a receptor that allows the extracellular S-RNases to enter the pollen tube in an allelic-specific manner. The S-RNase then degrades pollen rRNA in the incompatible pollen tube, leading to the arrest of pollen tube growth (Figure 2.5 (left), model I). In the inhibitor model, all S-RNases enter the pollen tube via an aspecific transporter. The pollen S-protein acts then as an RNase

inhibitor that recognizes and inhibits all S-RNases through an interaction with a low-affinity binding site. When the S-RNase is allelic, binding to a high-affinity, haplotype-specific site occurs, and this somehow prevents the interaction with the low affinity site, leaving the S-RNase functionally active and thus able to inhibit pollen tube growth (Figure 2.5 (left), model II). The inhibitor model is favoured because it is consistent with the localization of S-RNases within both compatible and incompatible pollen tubes. A variation on the inhibitor model was proposed by Luu et al. (2001). According to these authors the pollen S-allele products only contain the S-allele-specificity domain and a general RNase inhibitor is responsible for the inhibition of S-RNases. The general RNase inhibitor produced by the pollen tube would bind and inactivate all non-self S-RNases. In the case of a self S-RNase, binding would be prevented because the binding between the matching S-allele specific domains of a pollen S-allele product and its cognate S-RNase (which is not a RNase inhibitor) is more thermodynamically favourable. A further refinement of the inhibitor model suggests that the pollen S-proteins function as a tetramer. This enables them to account for the behaviour of heteroallelic pollen (i.e. pollen carrying two S-alleles) produced by tetraploid plants or plants with a radiation-induced duplication of the S-locus, so called pollen-part mutants (Figure 2.5 (right)).

2.2.3.2 Self-incompatibility in Papaver

In *Papaver rhoeas* a gametophytic SI-system controlled by one multi-allelic locus is active. However, study at molecular and biochemical level has revealed that the stigmatic S-gene and the mechanisms involved in pollen inhibition are completely different from those of the Solanaceae*,* where the SI-response involves the action of cytotoxic RNases (see 2.2.3.1). This is related to differences in the physiology of the SI-system, since, in contrast to the Solanaceae, inhibition of incompatible pollen occurs on the stigmatic surface and arrest of pollen tube growth is rapid, occurring within minutes. The SI-response in the Papaveraceae is mediated by a complex signalling cascade and does not involve degradation of RNA (Franklin-Tong et al. 1991, Rudd and Franklin-Tong 2003).

The *Papaver* stigmatic S-proteins are small (15 kDa) secreted proteins, some of which are N-glycosylated. The S-proteins are highly polymorphic, sharing between 51.3 and 63.7% amino acid sequence identity. The *Papaver* S-proteins share some homology to identified members of a large gene family in *Arabidopsis*, called the S-protein homologues (**SPH**). The

function of the SPH-genes, which are expressed at certain developmental stages in *Arabidopsis*, a self-compatible species, is still unclear, but they seem to be involved in some aspects of cell-cell signalling, possibly during development (Ride et al. 1999). To identify the amino acid residues in the stigmatic S-proteins that are responsible for the allele-specific recognition in *P. rhoeas*, Jordan et al. (2000) constructed a range of mutants of the S_1-protein using site-directed mutagenesis. This study revealed that both variable and conserved amino acid residues in the hydrophilic surface loops 6 and 2 play a role in the recognition and inhibition of incompatible pollen in the pollen-pistil SI-reaction.

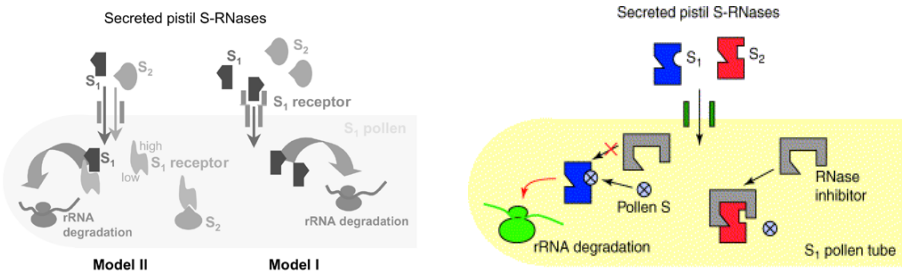


Figure 2.5 Models for the mechanism of pollen inhibition in the Solanaceae, Rosaceae and Scrophulariaceae. The products of the female S-gene, the pistil S-RNases are secreted into the transmitting tissue of the style. Pollen tubes (of S_1-genotype here) growing through the style encounter S-RNases. In the case of pollen that carries an S-allele corresponding to either of the alleles present in the style, inhibition occurs. (Left) Two models have been proposed for the inhibition mechanism. Model I: the S-RNAse (S_1 here) enters the S_1 pollen tube via an allele-specific receptor. The S-RNase then degrades the rRNA within the incompatible pollen tube and arrest the pollen tube growth. In Model II, all S-RNases enter the pollen tube via an unspecific transporter. On entry, they encounter the pollen S-receptor, which has two ligand-binding sites. There is a low-affinity site that binds and inhibits the S-RNases in an allele-independent manner (shown as an interaction with S_2-RNase), and a high-affinity site, which is allele specific. Binding to this site prevents binding of the S-RNase (S_1 here) to the low affinity, hence the protein remains active and degrades the pollen rRNA (Wheeler et al. 2001). (Right), modified inhibitor model. Secreted S-RNases, S_1 and S_2, are taken up, probably as part of a protein complex by a growing pollen tube. The S_2-RNase is recognized and bound by a RNase inhibitor that inactivates the S-RNase. By contrast, the S_1-RNase is bound at an allele-specific binding site by the S_1 pollen protein, which is possibly a tetramer. This interaction is thermodynamically favoured over the interaction with the RNase inhibitor. Hence the S_1-RNase remains active and degrades the pollen tube rRNA, resulting in the arrest of the pollen tube growth (Franklin-Tong and Franklin 2003).

Stigmatic SPH-proteins interact with the pollen S-gene product, which is believed to be a plasma membrane receptor. One candidate, an S-protein binding protein (**SBP**), expressed in mature pollen was identified by Hearn et al. (1996). SBP is a 70-120 kDa glycoprotein that is localized in the pollen plasma membrane and that binds S-proteins. Although this interaction appears to be crucial to the SI-reaction, the binding to the S-proteins is not allele specific. The lack of allelic discrimination by SBP suggests that it may function as an accessory receptor in the SI-reaction, modulating the interaction of stigmatic S-proteins with the pollen allele-specific S-receptor. However, till now no direct evidence exists that SBP and the S-receptor are separate entities (Jordan et al. 1999).

Despite this lack of knowledge about the female and male recognition factors, much is known about how the arrest of the pollen tube growth occurs after an incompatible reaction in *Papaver*. Inhibition of incompatible pollen is mediated by the activation of a calcium-based signal transduction pathway in the pollen. An increase in cytosolic free Ca^{2+} ($[Ca^{2+}]_i$) is the initial step in the signalling cascade, leading to loss of the apical high $[Ca^{2+}]_i$ gradient that is typical of growing pollen tube tips. This loss of high apical $[Ca^{2+}]_i$ is likely to play a part in the initial inhibition of pollen tube growth. It has been speculated that vesicle fusion, which occurs specifically at the tip and requires high $[Ca^{2+}]_i$ might be rapidly inhibited in the SI-response, and thus be responsible for pollen tube inhibition. The rapidity of the Ca^{2+} response is consistent with the idea that the SI-reaction is a receptor-mediated response, where the binding of the pistil S-protein to the receptor acts as a signal molecule that triggers increases in $[Ca^{2+}]_i$. The increase in $[Ca^{2+}]_i$ precedes a complex series of events that ultimately lead to death of the incompatible pollen (Franklin-Tong et al. 1993, 1995).

The increase of $[Ca^{2+}]_i$ is rapidly followed by the phosphorylation of at least two proteins, **P26** and **P68** (Rudd et al. 1996, 1997). Phosphorylation of P26 occurs within 90 sec, and is mediated by a Ca^{2+}-dependent serine/threonine protein kinase (CDPK). P26 is a protein with pyrophosphatase activity. Soluble inorganic pyrophosphatases often drive cellular biosynthetic reactions and play an important role in generating both ATP and the biopolymers required for making new membranes and cell walls. Under conditions of raised $[Ca^{2+}]_i$, when P26 is phosphorylated, its pyrophosphatase activity is reduced (Rudd and Franklin-Tong 2003). The reduction of pyrophosphatase activity of P26 during the SI-response could result in the depletion of biopolymers, such as long-chain carbohydrates and

proteins, which contribute to pollen tube membranes and cell walls. This would finally result in the inhibition of the tip growth. However this hypothesis needs to be tested further. Phosphorylation of P68, probably a protein kinase, occurs at a later stage in the incompatibility response and is Ca^{2+}-independent. Therefore, the SI-signalling cascade mediating the SI-response is complex, having both a Ca^{2+}-dependent and Ca^{2+}-independent phase.

Another pollen component that displays altered activity during the SI-response is the putative mitogen-activated protein kinase (MAPK) **P56** (Rudd et al. 2003). MAPK phosphorylation cascades are highly conserved signal transduction mechanisms. In higher plants, MAPKs have been implicated in a wide range of stress responses. P56 activation occurs downstream of Ca^{2+} increases. Activation peaks several minutes after initial inhibition of pollen tube growth and remains active for at least 30 min. P56 might therefore be involved in signalling to the irreversible inhibition of pollen tube growth (Thomas et al. 2003).

It has been demonstrated that the actin cytoskeleton is essential for apical growth in pollen tubes (see above). Snowmann et al. (2000, 2002) demonstrated that SI stimulates a dramatic and rapid rearrangement of F-actin that is sustained for at least 1 hour. This demonstrates that alterations to the cytoskeleton continue even after inhibition of tip growth. Actin depolymerization was also achieved in pollen by treatments that increase cytosolic Ca^{2+} artificially, indicating that actin alterations are on the Ca^{2+}-mediated signalling cascade. The dramatic and rapid reorganization and depolymerization of actin during the SI-response raise the question of how these alterations are mediated. Actin-binding proteins (ABPs) are thought to be responsible for modulating changes in actin organisation and dynamics, but which ABPs are involved in SI remains to be identified (Staiger and Franklin-Tong 2003, Thomas et al. 2003). It is possible that cytoskeleton destabilizing occurs during the SI-response of all plant families but it has only been studied in the family of the Papaveraceae. Whether the cytoskeleton destabilization is the cause or the consequence of the SI-reaction may depend on the family studied.

Actin depolymerization is probably linked to events that trigger apoptosis, as was found in some animal cell types. Similarly, programmed cell death (PCD) might be triggered by SI in incompatible pollen. Data that are consistent with the idea of PCD come from observations of S-specific

nuclear DNA fragmentation in incompatible pollen tubes (Jordan et al. 2000). As artificial increase of $[Ca^{2+}]_i$ in pollen tubes results in DNA fragmentation, Ca^{2+}-signalling might be involved in the induction of PCD in plant cells. Although DNA fragmentation in animal cells involves caspase-3 (cysteine-aspartate proteases, which have the capacity to cleave several cellular proteins) activity, no caspase homologues have been identified in *Arabidopsis* to date. This means that either these very highly conserved genes are not present in plants or that analogous genes may operate in plants. Experiments with Caspase inhibitors, such as DEVD (a caspase-3 inhibitor which is commonly used to block PCD in animal cells) and YVAD (a caspase-1 inhibitor, a potential negative control), clearly indicate that DEVDase/caspase-like activity is triggered in an S-specific manner in incompatible pollen and that this mediates DNA fragmentation (Thomas et al. 2003, Thomas and Franklin-Tong 2004). However, further studies to explore the evidence for a PCD signalling cascade and to provide firm links between this and the sustained depolymerization of F-actin are required (Staiger and Franklin-Tong 2003).

A model for the SI-response in *Papaver* has been proposed by Franklin-Tong et al. (2002) (Figure 2.6). The interaction between a pollen S-receptor, a single entity or a receptor complex with S-protein binding proteins, and an identical stigmatic S-protein results in an immediate increase of cytosolic free calcium $[Ca^{2+}]_i$ within the pollen. Three events occur very soon (probably within 1 min) after the initial SI-signal. The first event is loss of tip-focused $[Ca^{2+}]_i$ gradient. Since this is thought to be required for vesicle fusion, it could be responsible for immediate arrest of pollen tube growth. The second event is the rapid phosphorylation of P26 via CDPK and P68 in a Ca^{2+}-independent manner. As P26 has homology to an inorganic pyrophospatase, it is highly likely to contribute to an arrest of biosynthesis of cell membrane and wall material essential for growth. The third event is the dramatic reorganization and depolymerization of F-actin. Other events are activation of the putative MAPK P56 and triggering of a programmed cell death (PCD)-signalling cascade, eventually resulting in the fragmentation of nuclear DNA.

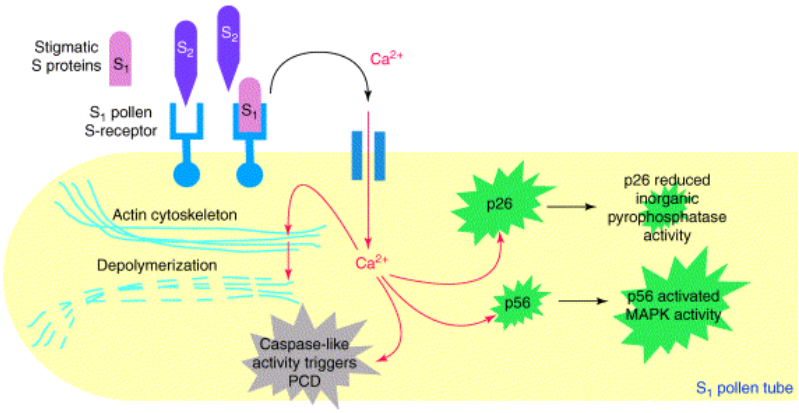


Figure 2.6 Current model of signals and targets involved in the SI-response in *P. rhoeas*. Induction of an incompatible SI-signalling cascade in pollen that interacts with a matching stigmatic S-protein triggers several rapid events, most of which are likely to contribute to the inhibition of pollen tube growth. These include: increases in $[Ca^{2+}]_i$ in the 'shank' region and loss of the apical high $[Ca^{2+}]_i$ gradient; alterations to the actin cytoskeleton; the phosphorylation of soluble pollen proteins including p26, with a postulated resultant reduction in pyrophosphatase activity. Most of these early events are likely to contribute to the rapid arrest of incompatible pollen tube growth. Later, further changes in the actin cytoskeleton are detected and a putative MAPK is activated; their roles are not known, but MAPKs are known to be involved in stress-induced responses. DNA fragmentation and other more recent data suggest that a PCD signalling pathway is triggered, suggesting that later events may 'commit' incompatible pollen to die. Thus, it appears that there are several 'fail-safe' mechanisms involved in pollen tube inhibition (adapted from Franklin-Tong et al. 2002).

2.2.3.3 Self-incompatibility in the Poaceae

Grass pollen is tri-cellular and short-lived. The stigma is dry and composed of two stylodia covered by secondary stigma branches made of up to five ranks of cells, the papillate tips of which support the pollen-receiving surfaces. After a compatible pollen grain lands on a receptive stigma surface it attaches to the 'sticky' receptive zone, hydrates and lifts the operculum at the germination aperture. The tube tip emerges and penetrates the stigma. The pollen tube grows through the intercellular spaces of the stigma branch and the transmitting tract in the stylodium and in the upper ovary wall. It enters the ovarian cavity and fertilization takes place.

The self-incompatibility reaction is rapid and the arrest of the pollen tube growth occurs almost immediately (in certain species as rapidly as 30 s after emergence of the tube tip) upon contact with the stigma (Wehling et al.

1994a). However, the time and site of the reaction response are very variable and change with the genotype of the individual plant (Heslop-Harrison 1982).

As mentioned above, the SI in grasses is controlled by two multi-allelic loci S and Z. Each locus can contain up to 40 different alleles (Fearon et al. 1994). Fertilization is prevented when the S and Z alleles present in the pollen are matched in the style. Although S and Z determine the recognition process between pollen and stigma, it is assumed that, similarly to what is found in other plant families, many other genetic factors are involved in the cascade of events that is triggered by an incompatible reaction in the pistil.

Due to the complementary interaction of two genes, S and Z, SI in the grasses has features distinct from those of single locus systems. First of all, the degree of compatibility (i.e. the percentage of compatible pollen) between two plants can be either: 0% compatibility (fully incompatible: no pollen tube has been observed), 50% compatible (half of the pollen tubes grow through the style), 75% compatible and 100% compatible (all the pollen grains germinate and produce a full grown pollen tube) (Table 2.1 and Figure 2.7). Second, differences are encountered in reciprocal crosses. For example, after fertilization of a plant with genotype $S_{11}Z_{12}$, with $S_{12}Z_{13}$ as pollen donor, 75% of the pollen tubes will be compatible. In the reciprocal cross, only 50% of the pollen germinate and result in pollen tube growth. This occurs when one parent is homozygous of one locus and there is one allele difference between parents. This should not happen in one locus systems as homozygosity at the S-locus do not happen, theoretically.

Perhaps because the rejection system (bi-factorial, multi-allelic, complementary and apparently restricted to monocots) lacks the simplicity that has been attributed to the mechanisms operating in the Solanaceae and Brassicaceae, relatively few laboratories are contributing to the understanding of SI in grasses (de Nettancourt 2001).

The study of the complete breakdown of SI has revealed the existence of mutants with interesting characteristics. Lundqvist (1954) identified mutations at or near the S and Z loci in rye (*Secale cereale*). All mutants were pollen-part mutants, meaning that the pollen grains had lost their SI-specificity whereas the functionality of the pistils of the same plant remained intact. Thorogood and Hayward (1991) analysed a self-compatible mutant of *Lolium perenne*, but in this case the gene causing self-fertility was not

located to S or Z, suggesting that a third locus was involved. The presence of a third locus was confirmed by a study of self-fertile mutants in *Phalaris coerulescens* (Hayman and Richter 1992). The mutation at the third locus (T) showed close linkage to a leaf peroxidase isozyme and had an effect only on the pollen and not on the pistil of the mutant plant. Similarly, Voylokov et al. (1993) identified a third locus conferring self-compatibility in *Secale cereale*.

Table 2.1 Possible incompatibility relations in ryegrass where the SI-response is controlled by two multi-allelic loci: (-) incompatible, (½) half-incompatible, (¾) ¾-compatible and (+) full compatible.

Pollinators		Female		
		Plant 1 $S_1S_2Z_1Z_2$	Plant 2 $S_1S_2Z_1Z_3$	Plant 3 $S_1S_1Z_1Z_3$
Plant 1 ($S_1S_2Z_1Z_2$)	S_1Z_1	-	-	-
	S_1Z_2	-	+	+
	S_2Z_1	-	-	+
	S_2Z_2	-	+	+
Incompatibility relation		-	½	¾
Plant 2 ($S_1S_2Z_1Z_3$)	S_1Z_1	-	-	-
	S_1Z_3	+	-	-
	S_2Z_1	-	-	+
	S_2Z_3	+	-	+
Incompatibility relation		½	-	½

Although none of the gene products of S and Z have been isolated to date, their chromosome location has been determined in rye (Lewis et al. 1980 and Gertz and Wricke 1989) and in ryegrass (Cornish et al. 1979, Thorogood 1991 and Thorogood et al. 2002). The S-gene is located on linkage group 1 (numbering of linkage groups follows the consensus map of the Triticeae) and is known to be linked to PGI-2, an isozyme phosphoglycoisomerase. The Z-locus is located on linkage group 2 (Fuong et al. 1993). In *Phalaris coerulescens*, high resolution maps of the S- and Z-regions have been generated recently using RFLP probes from wheat,

barley, oat and *Phalaris*. The S-locus was localized to the centromeric region of chromosome 1 and two markers (Xbm2 and Xbcd762) were found to cosegregate with S. The Z-locus, on the other hand, was localized to the long arm end of chromosome 2 and one marker, Xbcd266 mapped 0.9 cM from Z (Bian et al. 2004). Recent work in rye also allowed the identification of one STS marker (TC116908) which cosegregates with Z (Hackauf and Wehling 2005). The gene from which this marker was derived was strongly expressed in rye pistils but not in leaves. A more detailed description of this mapping work will be given in Chapter 6. Knowledge about the map-position of S and Z can help to develop strategies for the isolation of these genes using, for example, map-based cloning.

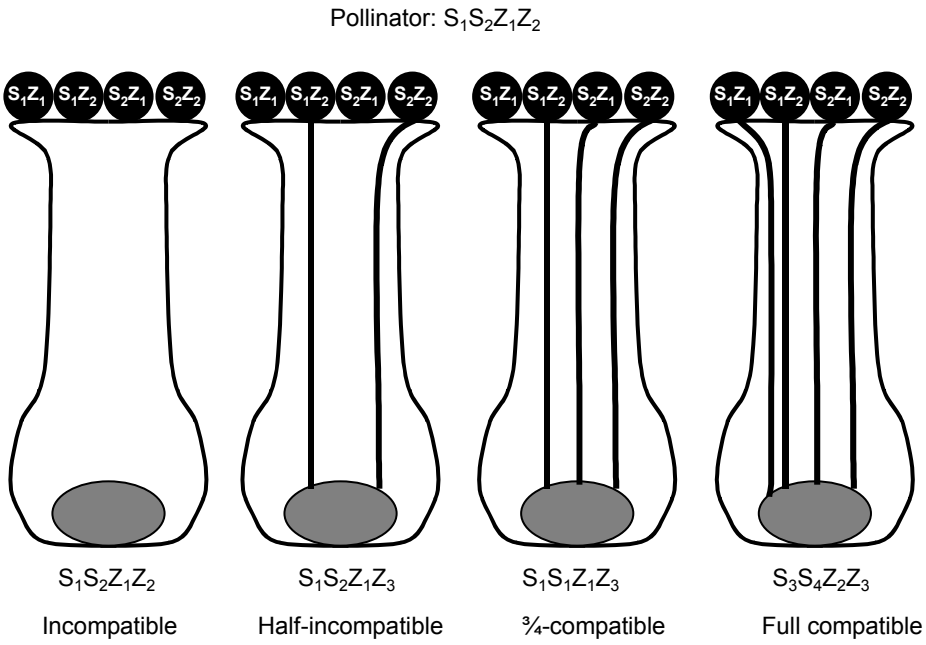


Figure 2.7 Possible incompatibility relations in ryegrass where the SI-response is controlled by two multi-allelic loci. Incompatible: none of the pollen grains effect fertilization, half-incompatible: 50% of the pollen grains effect fertilization, ¾-compatible: 75% of the pollen grains effect fertilization and full compatible: all the pollen grains effect fertilization.

Very little is yet known about the putative function or genomic organisation of other factors involved in the SI-response in the Poaceae. In the past *Secale cereale* and *Phalaris coerulescens* were used as 'models' to

analyse the SI-mechanism in the Poaceae. Different strategies have been followed in order to identify genes involved in the SI-response in these species. Speculating on the involvement of molecules similar to those controlling the SI-response in *Brassica*, the S-Locus Glycoprotein (SLG) and S-Receptor Kinase (SRK), a 280 bp genomic fragment was amplified in *Secale* with primers derived from the conserved *Brassica* SLG-13 sequence (Wehling et al. 1994b). Polymorphisms in this fragment correlated with the SI-genotypes of 46 rye plants tested. Wehling et al. (1994a) also demonstrated that protein phosporylation and Ca^{2+} gradients are involved in the SI-response in rye.

In *P. coerulescens*, Li et al. (1994) used a differential screening technique to identify the S-gene. A putative pollen specific cDNA clone, the *Bm2*-probe, was identified and used to identify an RFLP which cosegregated with the S-genotype of over 120 plants. The *Bm2* gene was strongly expressed in mature pollen but not in any other tissue tested. It consists of a highly conserved catalytic domain (the C terminus) and a variable, potentially allelic domain (the N terminus). All these characteristics strongly suggested that *Bm2*, coding for a protein with thioredoxin activity, represented the S-gene of *P. coerulescens*. However further research demonstrated that this thioredoxin gene does not represent the S-gene, but is located 2 cM apart from the S-gene in *P. coerulescens*. The S-thioredoxin gene plays probably a role in the SI-response but is not involved in the recognition process (Langridge et al. 1999).

More recently, four pollen-expressed genes potentially involved in pollen-stigma interaction have been isolated in *P. coerulescens* (Baumann et al. 2000). All four clones were predominantly or exclusively expressed in mature pollen and showed homology to protein kinases. However, mapping experiments demonstrated that none of the clones represents either the S- or Z-gene. The role of these protein kinases has to be studied further.

Wehling et al. (1994a) proposed a mechanism for the SI-reaction in rye that could be applicable to all SI members of the Poaceae (Figure 2.8). On the landing of an incompatible grain, the S- and Z-gene products formed in the stigma would associate with their self S- and Z-homologues of the pollen grain to trigger the phosphorylation of pollen proteins and an influx of Ca^{2+} (acting as a second messenger). According to this model, the growth of any pollen tube is inhibited only if the S- and Z-alleles of the pollen grain are activated by the S- and Z-components of the style. The signal-transduction

mechanism leading to this inhibition would induce a cascade of events mediated by a Ca^{2+} influx and subsequent increases in cytosolic calcium. S-proteins are probably not transferred across the plasma membrane during pollen germination and, as in *Papaver*, the incompatibility reaction occurs within the germinating pollen grain (Li et al. 1994). However, if we take into account our current lack of knowledge about the genetic factors involved in the SI-response in the Poaceae, this model is rather speculative.

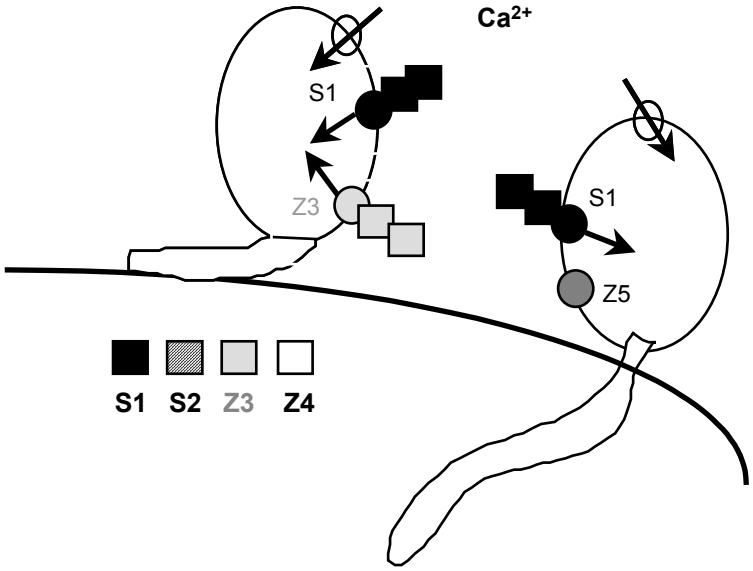


Figure 2.8 A model proposed for the SI-response in rye. Stigma SI-products (ligands) interact with their specific pollen receptors by mediation of Ca^{2+} as second messenger (Wehling et al. 1994a).

Part B

Results

Chapter 3

Identification of SI-related genes by homology-based approach

3.1 Introduction

As summarized in Chapter 2, in species of several plant families the SI-response has been studied in detail and some of the genes involved in SI have been identified. However, at the time this thesis was initiated, little was known about the genes involved in self-incompatibility in the grasses and no clues were available about the kind of genes which could be involved in the recognition and signaling processes in members of the grass family (with the exception of the works of Wehling et al. 1994b and Li et al. 1994, see further). Therefore, we analysed whether the kind of genes responsible for the recognition and signaling processes in other plant families played a role in the SI-response of *L. perenne*. Furthermore, we identified *L. perenne* homologues of the thioredoxin gene which had previously been identified in *Phalaris coerulescens* by Lee et al. (1994). This kind of homology-based approach was applied previously for example by Kowyama et al. (1995) to identify PCR fragments in *Ipomoea trifida* (Convulvaceae, sporophytic SI-system) which showed homology to *Brassica* SLGs and SRKs. One SRK-like gene was exclusively expressed in flower tissue but segregation studies revealed that it is not linked to the S-locus. These data strongly suggest that *Ipomoea* do not use the *Brassica* SRK/SCR system of SI (Kowyama et al. 1996).

As the grasses display a gametophytic SI-system, homologues of gene families known to be involved in SI in gametophytic systems were searched for in this study. To date only two forms of gametophytic SI-response have been elucidated in detail: the S-RNase based SI, active in members of the Solanaceae and of the Rosaceae, and the *Papaver rhoeas* SI-system. We therefore investigated whether RNase genes are also active in the SI-response in *L. perenne*. Although molecular and biochemical analyses have revealed that the stigmatic S-gene and the mechanism involved in pollen inhibition in *P. rhoeas* differ dramatically from that found in the Solanaceae; no SI-related genes have been identified in this species (see Chapter 2 for a review).

However, in species which show an S-RNase based SI-response, pollen germinates and begins to grow through the transmitting tract of the style. The growth of the incompatible pollen tube is arrested when it reaches about one-third of its way through the style. These features contrast with the typical SI-reaction in the grasses, where the inhibition reaction occurs immediately on the stigma surface. If we consider that the characteristics of the SI-

reaction in grasses (rapid arrest of the pollen tube on the stigma surface) are similar to the characteristics typical of sporophytic systems such as the one described in *Brassica,* it is possible that similar molecules are involved in the recognition and rejection reaction although the genetic control of SI is different. Therefore, we also investigated whether SLG- and SRK-homologues are involved in the SI-response of *L. perenne*. A similar approach was applied by Wehling et al. (1994a) in rye. They used primers derived from the conserved regions of the SLG-13 gene from *Brassica* in PCR reactions with rye DNA as template. A genomic 280 bp fragment was identified which displayed a polymorphism that co-segregated with different SI-genotypes. It was concluded, that the 280 bp fragment represents part of the S-gene in rye (Wehling et al. 1994a). In view of the partial homology found between the *Brassica* allele SLG13 and the putative S-locus in rye (or a gene closely linked to it), the authors concluded that a close relationship exists between gametophytic and sporophytic SI. Unfortunately, no further investigations were published on this putative S-gene.

As described in Chapter 2, in *P. coerulescens* a thioredoxin gene located 2 cM away from the S-locus seems to be involved in the SI-response (Li et al. 1995), although it does not represent the S-gene (Langridge et al. 1999). Thioredoxins are small, heat-stable molecules with catalytic activities that depend on several factors, such as the sequence beyond the protein core, the assembly of the protein and post-translational modifications (Li et al. 1995). Thioredoxins act in many biological reactions as hydrogen donors, substrates for reduction enzymes or receptors, and as subunits of viral DNA polymerases. It is also known now that the thioredoxins THL1 and THL2 are involved in the SI-response of *Brassica* (Cabrillac et al. 2001). We therefore tried to isolate the homologues in *L. perenne* and then determine its full-length sequence.

3.2 Objectives

The aims of the experiments described in this chapter were (i) to investigate whether homologues of genes known to be involved in SI in the Solanaceae, Brassicaceae and in *P. coerulescens* are also present in *L. perenne* and (ii) to determine whether these genes are involved in the SI-mechanism in ryegrass. Therefore PCR and RT-PCR reactions were performed with primers derived from conserved regions of SI-related genes S-RNases (Solanaceae and Rosaceae), SRK and SLGs (Brassicaceae) and S-thioredoxin (Poaceae).

3.3 Material and methods

3.3.1 Plant material

Young leaf material from four *Lolium perenne* genotypes (S_{22} 6/16, S_{22} 6/59, S_{22} 6/4 and G_2 15/39) was harvested. These plants were chosen from the collection of the Department of Plant Genetics and Breeding (CLO). Leaf material was stored at –80°C until required for DNA or RNA extraction.

In addition, non-pollinated and self-pollinated pistils of genotype S_{22} 6/16 were isolated for RNA extraction and RT-PCR analysis. Non-pollinated mature pistils were harvested from emasculated inflorescences to ensure that no cross-pollination had occurred. For the isolation of self-pollinated pistils, pollinations were carried out on excised inflorescences. Two plants of the same genotype (S_{22} 6/16 clones) were used as crossing partners (one as female and one as male partner). First, spikelets of the plant chosen as female were emasculated under a binocular to avoid unwanted self-pollination. Thereafter, the emasculated inflorescences were enclosed in cellophane bags till the pistils were mature and receptive for pollination (when the stylodia came out of the spikelet). To collect pollen, the inflorescences of the pollinator plants were enclosed in cellophane bags. The pollinator plants shared pollen around midday. The release of the pollen was stimulated by shaking pollinator inflorescences above the receptive pistils. It is important that no clumps of pollen, but only free-flowing, viable pollen was shed. The pollinated inflorescences were thereafter kept at room temperature for 24h. Pistils were harvested 1h and 24h after pollination. At 1h after pollination the SI-response was still active; after 24h, on the other hand, the mis-formed pollen tubes were clearly visible and the pistil started to degrade as no fertilization happened.

Before harvesting the pistils, the ‘fertilization’ was checked. To this purpose, stylodia of self-pollinated inflorescences were removed from the ovaries with a scalpel on microscope slides. Stigmas were placed under a coverslip in a drop of decolourised aniline blue (0.2% water-soluble aniline blue in 2% K_3PO_4). Pollen tube growth was then observed microscopically under UV fluorescent light. Aniline blue interacts with the callose layer of the pollen tube wall so that pollen tubes become visible (Martin, 1959). If the expected result was observed (malformed pollen tubes typical of self-incompatible reactions) the pistils were harvested. Samples were preserved at –80°C until RNA-extraction.

3.3.2 DNA and cDNA preparations

Genomic DNA from the four genotypes described above was extracted using a CTAB-based method (Weising et al. 1991). DNA-concentrations were estimated and standardized against known concentrations of λ-DNA on 1.5 % agarose gels. PCR reactions were performed on 30-50 ng of genomic DNA.

For the extraction of total RNA from leaves, the RNeasy plant mini kit (Qiagen) was used, according to the manufacturer's instructions. For the extraction of total RNA from pistils, the Tri-Reagent kit (Sigma) was used. The Tri-Reagent kit is based on the RNA isolation method of Chomczynski and Sacchi (1987). cDNA was synthesized from 1-5 µg of total RNA with Superscript II RNase H^- Reverse transcriptase (Invitrogen) in 20 µl of reaction buffer (50 mM Tris-HCL pH 8.3, 75 mM KCl and 3 mM $MgCl_2$) containing 20 mM DTT, 1 mM of each dNTP and 500 ng $Oligo(dT)_{12-18}$ primer. In each reaction 200 U of Superscript II and 20 U of Superase inhibitor (AMBION) were used. Two µl (around 50 ng cDNA) of the cDNA synthesis mix were used for RT-PCR reactions.

3.3.3 PCR and RT-PCR conditions

Table 3.1 describes all the primers used. To amplify S-RNase gene analogues in *L. perenne*, the degenerated primer pairs S-RNase/AS-RNase and S-appel/AS-RNase were combined. S-RNase was derived from the second conserved region of Solanaceae and Rosaceae S-RNases. AS-RNase was derived from the third conserved region of the Solanaceae and Rosaceae S-RNases (Figure 3.1). Primer S-apple was derived from the N-terminus of the S_2-RNase of apple (Figure 3.1). This was necessary as between the N-terminal sequences of known S-RNases from Solanaceae and apple (*Malus* spp.) little homology is present. The development of the primers was based on the amino acid sequence alignment information obtained by Broothaerts et al. (1995).

To amplify SRK and SLG analogues in *L. perenne*, we developed two pairs of degenerated primers: S-SRK2/AS-SRK2 and S-SLG2/AS-SLG2. S-SRK2 and AS-SRK2 were derived from the conserved amino acid region of SRK-9 from *Brassica napus* (accession number: D88193) and a receptor kinase domain of *Arabidopsis thaliana* (accession number: M80238) using the program BLAST of the Belgian EMBnet Node (BEN). The primers were

located on the third and the fifth exon of the *Brassica napus* sequence, respectively (Figure 3.1). S-SLG2 and AS-SLG2 were derived from the conserved region of SLG sequences and SRK sequences from *Brassica napus* and are located on the first exon of *Brassica napus SLG* (Figure 3.1).

Table 3.1 Description of the PCR primers used for the amplification of SI-gene analogues in *L. perenne*.

Primer code	Primer sequence (5'=>3')	Remarks
S-RNase	CNRTNCAYGGNYTNTGGCC	second conserved region of the S-RNases from the Solanaceae and Rosaceae (Broothaerts et al. 1995)
AS-RNase	YANCCRCANRYNCCRTGYTT	third conserved region of the S-RNases from the Solanaceae and Rosaceae (Broothaerts et al. 1995)
S-apple	TAYTTYCARTTYACNCARCARTA	N-terminus of the S_2-RNase from apple (Broothaerts et al. 1995)
S-SRK2	GGNTTYGGNATYGTNTAYGAR	Third exon of the *Brassica napus* sequence (Sakamoto et al. 1998)
AS-SRK2	RATNCKNGCCATNCCRAARTC	Fifth exon of the *Brassica napus* sequence (Sakamoto et al. 1998)
S-SLG2	GARATGAARCTNGGNTRBGAY	First exon of the *Brassica napus* sequence (Sakamoto et al. 1998)
AS-SLG2	CNAGNCKRATNTRRTCYTGRTG	First exon of the *Brassica napus* sequence (Sakamoto et al. 1998)
XUP	AAGCAAACAAGGATGGGAAAA	catalatic, thioredoxin domain of the S-thioredoxin of *Phalaris coerulescens* (Li et al. 1997)
XLP	GCGGAAAAGACACGGAAACTG	catalatic, thioredoxin domain of the S-thioredoxin of *Phalaris coerulescens* (Li et al. 1997)
LUP	CTACGACCAGGAGAGGAGACC	allelic, variable region of the S-thioredoxin of *Phalaris coerulescens* (Li et al. 1997)
LLP	AGGGAGACGAAGACGATGAGC	allelic, variable region of the S-thioredoxin of *Phalaris coerulescens* (Li et al. 1997)

In order to isolate the S-thioredoxin gene analogue in *L. perenne*, the non-degenerated primer pairs XUP/XLP and LUP/LLP published by Li et al. (1997), derived from the thioredoxin gene of *P. coerulescens* were used. XUP/XLP amplified a fragment within the catalytic domain. LUP/LLP

amplified a fragment within the, initially postulated, “allelic, variable region” of the S-thioredoxin of *P. coerulescens* (Li et al. 1997).

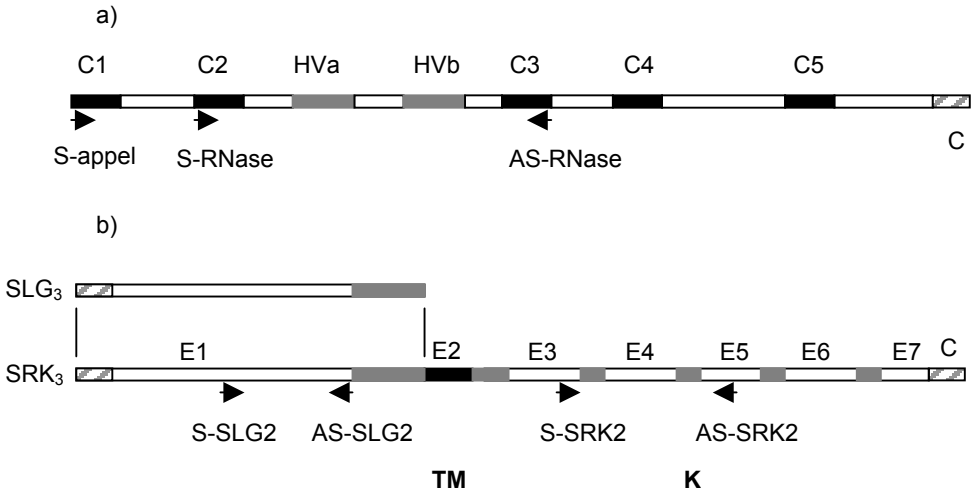


Figure 3.1 Overview of the primers used to amplify putative SI-related genes in *L. perenne*. (a) Scheme of the primers used for the amplification of the S-RNase homologues. C1-C5 refer to the five conserved regions and HVa-HVb to the hypervariable regions identified in S-proteins of Solanaceae (Broothaerts et al. 1995). (b) Schematic representation of the SLG_3 and SRK_3 proteins and the localization of the primers used to amplify SRK and SLG homologues in ryegrass. E1-E7 refer to the seven exons. The introns are indicated as grey boxes. S: S-domain; TM: transmembrane domain; K: kinase domain and C: C-terminal domain (Sakamoto et al. 1998 and Delorme et al. 1995).

The PCR reactions in which degenerated primers were used, were performed in a total volume of 50 µl with 50 ng genomic DNA or single stranded cDNA, 5 µl of 10 x PCR buffer, 125 pmol of each primer, 0.5 µl of dNTP mixture (10 nmol each), 50 nmol $MgCl_2$, 0.2 µl of Ampli *Taq* Gold DNA polymerase (5 u/µl) and MQ water. The PCR conditions were: 10 min at 96°C, 40 cycles of 1 min 96°C, 1.5 min 50°C and 1.5 min 72°C and 10 min at 72°C with a thermal cycler (9700, Perkin-Elmer).

The PCR reactions in which non-degenerated primers were used were performed by mixing 200 ng genomic DNA or single stranded cDNA with 5 µl 10 x PCR buffer, 20 pmol of each primer, 0.5 µl dNTPs (10 pmol each), 0.5 µl *Taq* DNA polymerase (5u/µl) and MQ water to make a final volume of 50 µl. The PCR conditions were: 30 sec of 94°C, 35 cycles of 30 sec of 94°C, 1

min of 58°C and 1 min of 72°C and final 10 min at 72°C on a thermal cycler (PE 9700).

The PCR products were loaded on 1.5 % agarose gels. λ-Pst I standard was used to estimate PCR product sizes. The gels were stained using EtBr and photographed. The PCR products were excised from the gel, purified using the QIAquick gel extraction purification kit (Qiagen) and sequenced. For each primer combination, positive controls for the respective species were considered (results not shown).

3.3.4 Cloning and sequencing

In a first step, PCR products were directly sequenced using the BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems), based on the dideoxy-chain termination method of Sanger et al. (1977). Sequencing reactions were carried out in a total volume of 20 µl with 50 ng of PCR product, 4 µl of Big Dyes version 1.1 (Applied Biosystems), 3.2 pmol of the corresponding PCR primer and 2 µl 5 x sequenase buffer.

The PCR products which could not be sequenced directly, were excised from the gel, purified using the QIAquick gel extraction purification kit (Qiagen), reamplified using the same primer set as used in the specific PCR reaction and cloned in the TOPO vector using the TA-TOPO cloning kit (Invitrogen). For each PCR product 8 colonies were picked, suspended in 200 µl MQ water and boiled for 5 min. In a next step, PCR reactions were carried out on 5 µl of the suspension using the M13 forward and reverse primers. PCR reactions were checked on 1.5 % agarose gels. Five colonies of each PCR product were used for sequencing.

All sequencing reactions were loaded on an automatic DNA sequencer (ABI Prism 377, Applied Biosystems™). Forward and reverse sequences were assembled using Autoassembler™ (Applied Biosystems™) to build a consensus sequence.

3.3.5 RACE

To obtain the full length cDNA of the thioredoxin gene, the marathon cDNA amplification kit from Clontech was used. Adaptor primers were combined with the sequence-specific primers XUP and XLP, following the manufacturer's instructions. The PCR products were sequenced and

sequence information was used to design a new primer pair (SI3/SI4) using the PrimerExpress software (Applied Biosystems), which allowed us to amplify the full length cDNA. This primer pair was used in PCR and RT-PCR reactions using 50 ng of DNA or cDNA, 5 µl 10 X PCR buffer, 20 pmol of each primer, 0.5 µl dNTPs (10 pmol each), 0.5 µl *Taq* DNA polymerase (5u/µl) and MQ water to make a final volume of 50 µl. The PCR conditions were: 30 sec of 94°C, 35 cycli of 30 sec of 94°C, 1 min of 58°C and 1 min of 72°C and final 10 min at 72°C on a thermal cycler (PE 9700).

3.4 Results and discussion

3.4.1 Detection of S-RNase homologues in Lolium perenne

The primer pair S-RNase/AS-RNase was used to amplify fragments at DNA and cDNA level. cDNA from leaves, non-pollinated and self-pollinated pistils, isolated at two different time points after pollination (1h and 24h) were used. At DNA level, fragments of 530 bp and 400 bp were amplified (Figure 3.2). On cDNA of leaves and self-pollinated pistils fragments of 400 bp and 250 bp were amplified (Figure 3.2**)**. In non-pollinated pistils no amplification was obtained after several repeats (results not shown).

With the primer pair S-apple/AS-RNase no amplification was obtained neither on DNA or cDNA. In apple, used as positive control, a fragment of 500 bp was amplified (Figure 3.3). This is not surprising, as the S-apple primer was designed in a highly polymorphic region.

All the PCR products shown in Figure 3.2 were excised from the gel, purified and sequenced. All the 400 bp DNA and cDNA fragments displayed identical sequences and all the 530 bp fragments were identical. From the 250 bp fragments, several DNA sequences were obtained after cloning. No homology among the fragments of 530 bp, 400 bp, and 250 bp was found. Database search in the protein database of the 'National Centre of Biotechnology Information' (NCBI) by the program Blastx (nucleotide sequence translated against protein databank) revealed no homology to known S-RNases or RNases. The fragments of 530 bp revealed 47% homology in 107 amino acids (aa) to a putative nucleotide-binding leucine-rich-repeat protein 1 of *Oryza* (Accession number: Q69L52), a protein involved in defense to pathogens. The fragments of 400 bp revealed 40% homology in 94 aa to an aminotransferase-like protein of *Oryza* (Accession number: Q6K3Y9). The sequences obtained for the 250 bp fragments did not match any known gene of the database.

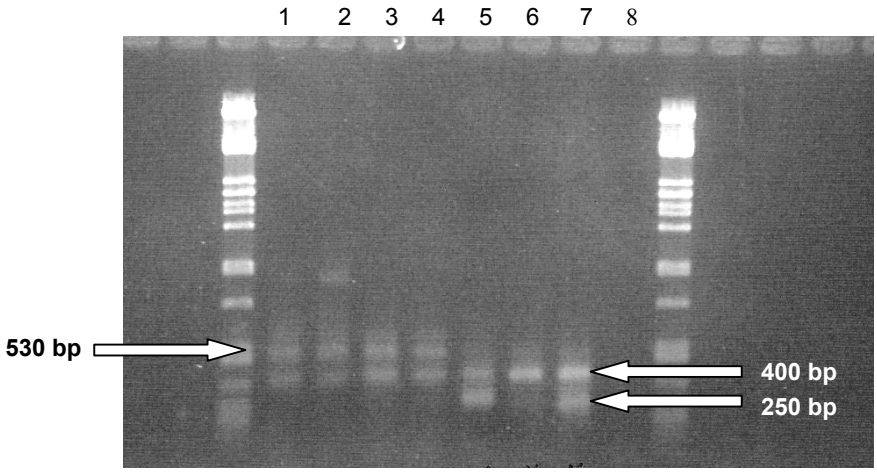


Figure 3.2 A detail of the amplification pattern obtained by PCR and RT-PCR reactions with S-RNase/AS-RNase. (1-4) DNA of the genotypes S_{22} 6/16, S_{22} 6/59, S_{22} 6/4 and G_2 15/39, respectively, (5) leaf cDNA, (6) cDNA of self-pollinated pistils 24 h after pollination, (7) cDNA of self-pollinated pistils 1 h after pollination and (8) water as negative control.

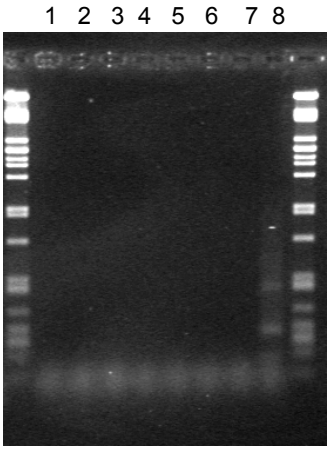


Figure 3.3 Amplification patterns obtained by PCR reactions with S-appel/AS-RNase. (1) water, as negative control (2-3) DNA of genotypes S_{22} 6/16 and S_{22} 6/59, respectively, (4) leaf cDNA, (5) cDNA of non-pollinated pistils, (6) cDNA of self-pollinated pistils 24 h after pollination, (7) cDNA of self-pollinated pistils 1 h after pollination and (8) apple DNA as positive control.

Given the lack of homology between the sequences identified in *L. perenne* and known S-RNases, there is no evidence that an S-RNase based SI-system is active in ryegrass. Except for the fact that also in *L. perenne* SI is under gametophytic genetic control, no common features with the SI-system of the Solanaceae are found. This is not surprising, if we take into consideration that the model which is currently very acceptable for the mode of action of S-RNases (Newbigin and Uyenoyama 2005), requires the incorporation of these cytotoxic proteins into the growing pollen tube (probably non-selectively). This requires that the pollen germinates and the pollen tube migrates through the transmitting tract of the style before the inhibition reaction takes place. As mentioned in Chapter 2, the inhibition reaction in *L. perenne* takes place on the stigma surface, a feature which is more consistent with a receptor-based SI-model as the one recently accepted for *Brassica*.

Whether the sequences identified in this study correspond to elements of the SI-response in ryegrass remains to be investigated. However, given the fact that they seem to be expressed in both reproductive and non-reproductive tissues, they are probably not related to SI. As no homology was found to known S-RNases we did not analyse these DNA fragments further.

3.4.2 Detection of SRK and SLG homologues in Lolium perenne

The primer pair S-SRK2/AS-SRK2 amplified two fragments of 500 bp and 200 bp respectively on genomic DNA (Figure 3.4). RT-PCR reactions were performed on cDNA of leaves, non-pollinated and self-pollinated pistils. In leaves and self-pollinated pistils a fragment of 400 bp was amplified (Figure 3.4). In the non-pollinated pistils, on the other hand, a weak band of 200 bp was amplified (Figure 3.4).The primer pair S-SLG2/AS-SLG2 amplified on DNA samples one fragment of 400 bp and on cDNA of leaves, one fragment of 300 bp (Figure 3.4). No amplification was obtained on cDNA of pistils (Figure 3.4).

All the amplified fragments were excised from the gel, reamplified, cloned and sequenced. Database search using the program FASTA (nucleotide against nucleotide) or BLASTx (nucleotide translated to protein) revealed homologies to different kind of genes (Table 3.2).

Table 3.2 Overview of the homologies found by database search for the fragments amplified by the primer pairs S-SRK2/AS-SRK2 and S-SLG2/AS-SLG2.

Origin	Size (bp)	Primer pair	% homology	Homology
DNA	500	S-SRK2/AS-SRK2	67% in 239 bp	*Brassica oleracea* mRNA for receptor kinase (EM_PL: BOSFR2)
			67% in 238 bp	*Brassica rapa* SRK22 mRNA for S-locus receptor kinase (EM_PL: AB054061)
DNA	200	S-SRK2/AS-SRK2		No sequence information available
cDNA	400	S-SRK2/AS-SRK2	83% in 149 aa	S-locus receptor-like kinase RLK14 mRNA of *Oryza sativa* (Q8LLI4)
			68% in 149 aa	Putative serine/threonine kinase of *Arabidopsis thaliana* (Q9SVK2)
DNA	400	S-SLG2/AS-SLG2		No homology to known genes of the database
cDNA	300	S-SLG2/AS-SLG2	80% in 89 aa	Putative ATP synthase gamma chain 1 of *Oryza* (Q84NW1)

None of the fragments amplified by the S-SLG2/AS-SLG2 primer pair displayed homology to known SLG genes. Based on these results, we concluded that no SLG gene homologues are present in the ryegrasses and no further experiments were performed with these PCR products. This result is not surprising as even in members of the Brassicaceae the precise function of the SLG protein in the SI-response is unclear, it probably just enhances the SI-response by stabilizing SRK (Chapter 2).

It was not possible to obtain the sequence information of the 200 bp product (Figure 3.4) despite several sequencing attempts. The 500 bp fragment amplified on DNA level and the 400 bp amplified on cDNA level by the primer pair S-SRK2/AS-SRK2 revealed homology to receptor kinases. Whether any of these PCR products correspond to S-receptor kinase genes remains to be tested. However, the 500 bp product was not amplified on cDNA samples and the 400 bp fragment was amplified in both reproductive and non-reproductive tissues (Figure 3.4). The 400 bp fragment amplified with primer pair S-SRK2/AS-SRK2 was converted into a DNA-marker suitable for mapping in segregating populations. This allowed us to check whether it co-segregates with the S- or Z-locus in ryegrass (see Chapter 6).

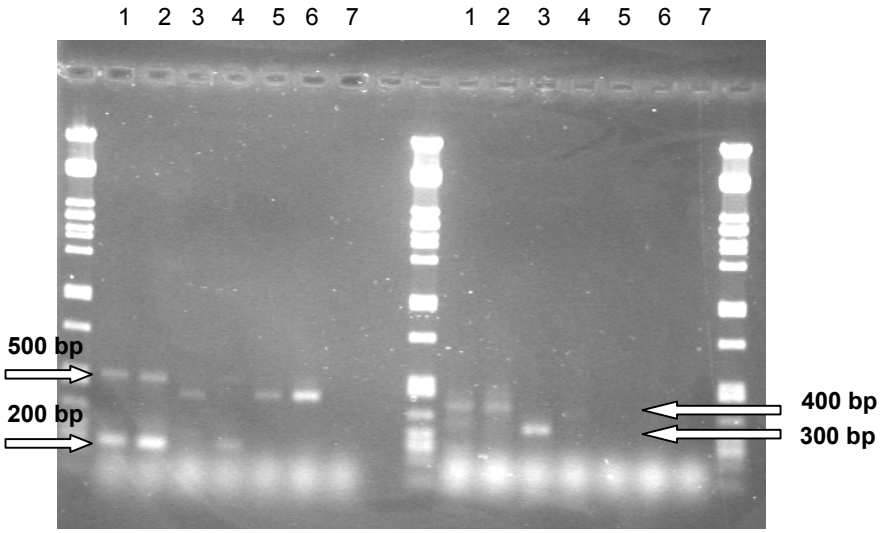


Figure 3.4 Amplification pattern obtained after PCR and RT-PCR reactions with the primer pairs S-SRK2/AS-ASRK2 (left) and S-SLG2/AS-SLG2 (right). (1-2) DNA of the *L. perenne* reference plants (S_{22} 6/16, S_{22} 6/59); (3) cDNA of leaves, (4) non-pollinated pistils, (5-6) self-pollinated pistils after two different time points of pollination (24 h and 1 h, respectively) and (7), water as negative control.

3.4.3 Detection of sequence homologues to the S-thioredoxin gene of Phalaris coerulescens (Bm2) in Lolium perenne

Two primer pairs XUP/XLP and LUP/LLP were used to identify S-thioredoxin homologues in *L. perenne*. The primers XUP/XLP flank the conserved thioredoxin domain and the primer pair LUP/LLP flanks the putative variable allelic region defined by Li et al. (1997). As expected (Li et al. 1997), the XUP-XLP primer pair amplified one fragment of 380 bp in leaf cDNA (Figure 3.5). Database search of this fragment, using the program FASTA, revealed homology of about 90% in 305 bp with the S_2-gene of *P. coerulescens*. The LUP-LLP primers amplified two fragments: one of the expected size (150 bp) and one a little bit longer (200 bp) (Figure 3.5). For the 200 bp fragment no sequence information was obtained. The 150 bp fragment displayed 82% homology in 120 bp to the S_2-gene of *P. coerulescens*. Amplification of this thioredoxin gene in leaves contrasts with what was found by Li et al. (1994), there the *Bm2* gene was highly expressed in mature pollen of *Phalaris* and not in other tissue types. Expression studies of the *Bm2* gene in other grass species such as *S.*

cereale, *H. bulbosum* and *L. perenne* indicated that *Bm2* mRNA was barely detectable in the pollen of these grass species compared to the expression level in *Phalaris* (Li et al. 1997). As the expression level of this gene differs from species to species *Bm2* could not be the S-gene. However, this gene is located 2 cM away from the S-gene in *P. coerulescens* and it might be involved in the SI-reaction (Langridge et al. 1999). Therefore, in a next step we characterized the full length cDNA of this gene.

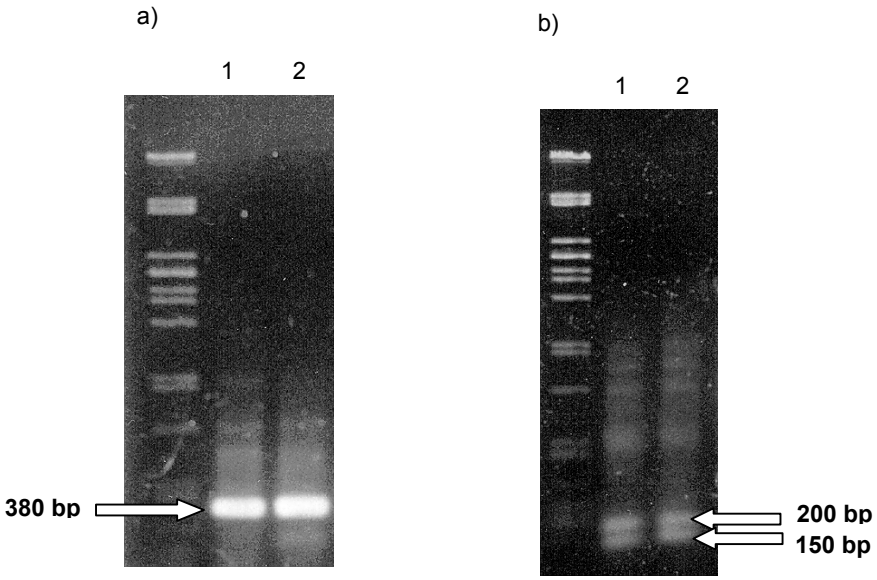


Figure 3.5 Amplification patterns. a) The XUP/XLP primerpair amplified on cDNA of leaves of the reference plants ((1) S_{22} 6/4 and (2) S_{22} 6/16) a fragment of 380 bp and b) the primers LUP/LLP amplified on cDNA two fragments, one of 200 bp and an other of 150 bp.

To obtain the full length cDNA of the thioredoxin gene 5' and 3' Rapid Amplification of cDNA Ends (RACE) reactions were performed. For the 5'-RACE reaction the AP1 (adaptor primer) / XLP (gene-specific) primer pair was used. The 3'-RACE reactions were performed with the XUP/AP1 primer pair (Figure 3.6). From the obtained sequences a new primer pair was developed, SI3 (ATCAATCGAACCACTCGCTGCG) and SI4 (GGCGGGAAAGACACAGACACTGT). A cDNA fragment of 800 bp and a DNA fragment of 2700 bp were amplified (Figure 3.7). Sequence information revealed that this fragment starts from the second exon of the thioredoxin gene of *Phalaris* and ends at the 3' non-translated region (Figure 3.8), as expected. It was confirmed that this gene is expressed in different tissue

types of ryegrass (Figure 3.7). In all samples a fragment of the same length, 800 bp was amplified. Database search of the sequences of these PCR products revealed a homology of 85% in 705 bp overlap with the thioredoxin-like protein (pTrx) of *P. coerulescens* (EM_PL: AF 159389) (Figure 3.9), confirming that this thioredoxin gene is expressed not only in pollen but also in pistils and non-reproductive tissues of *L. perenne*.

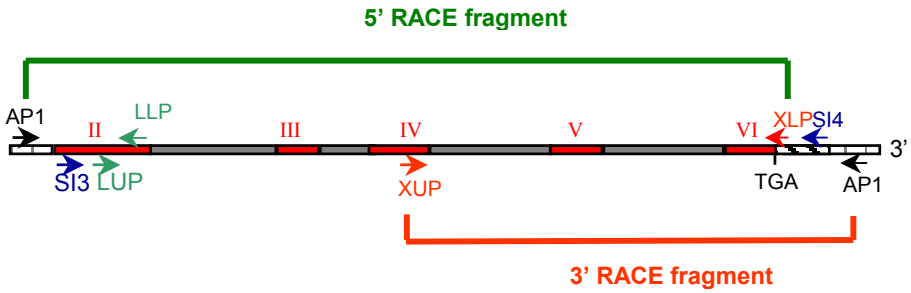


Figure 3.6 Scheme of the primers used in the RACE reactions. I-VI: exons, grey blocks: introns.

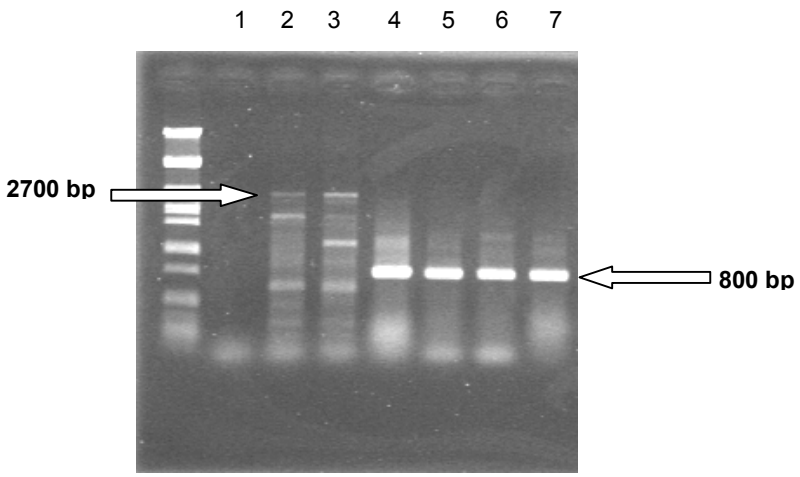


Figure 3.7 Amplification pattern of PCR reactions with the primer pair SI3/SI4 on (1) water, (2-3) DNA of reference plant (S_{22} 6/16, S_{22} 6/59), (4) cDNA of leaves, (5) non-pollinated pistils and (6-7) self-pollinated pistils 1 h and 24 h after pollination, respectively.

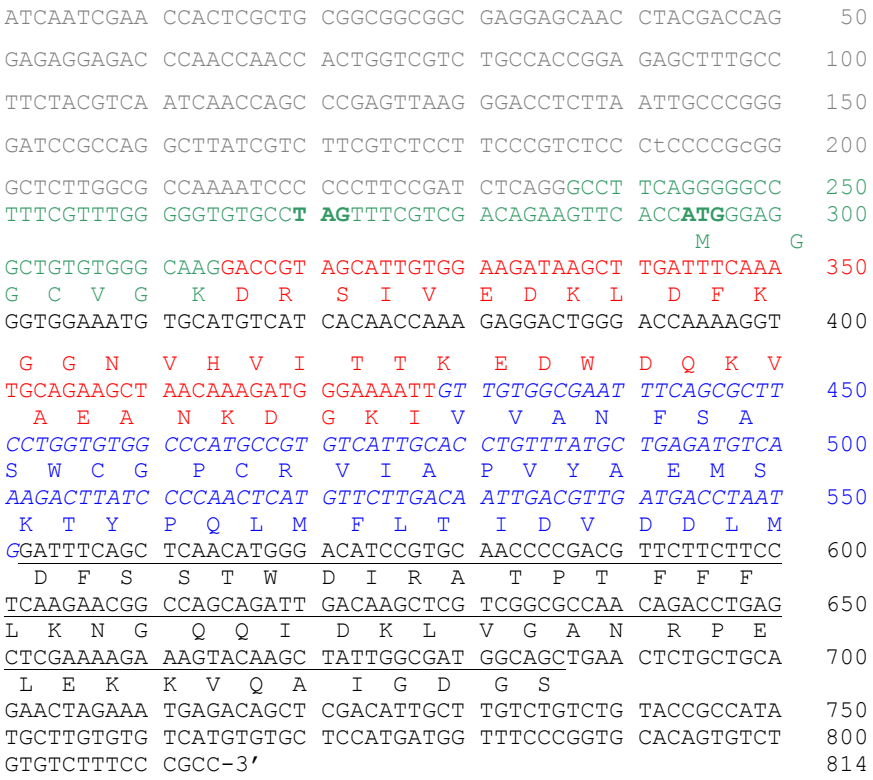


Figure 3.8 Scheme of the pistil cDNA sequence of the thioredoxin gene, amplified in *Lolium perenne* with the primer pair SI3/SI4, grey: exon 2, green: exon 3, red: exon 4, bleu: exon 5, underlined: exon 6. In exon 3: TAG: stop codon and ATG: start codon.

We employed different strategies (Invert PCR and Anchored PCR) to amplify the first exon of this thioredoxin gene in ryegrass but without success. Further work with *Bm2* homologues from three self-incompatible grasses (*L. perenne, H. bulbosum and S. cereale*) and two other alleles from *Phalaris* showed that their open reading frames (ORFs) were less than half the size from that predicted for *Bm2* by Li et al. in 1994 (Baumann et al. 2000). Comparing the new sequence data with the previously published putative *Phalaris* S_1-sequence, a discrepancy in the region of the first intron-second exon boundry was found. The cDNA sequence at the 5' end of the genes isolated by Baumann et al. (2000) identified regions of homology, and terminating within the first intron of the *Phalaris* S_1-sequence reported by Li et al. (1994). The newly isolated sequences translate into thioredoxin-like protein only and do not contain a so called, "allelic" domain. Comparing the

structure of the *Bm2* homologues from the various grasses revealed, in several cases, conserved positions of upstream stop codons, making clear that the thioredoxin-like proteins do not constitute a component of larger proteins (Baumann et al. 2000). The *Bm*2 homologue in this study, contains a stop codon in exon 3 and an ATG start codon directly before the thioredoxin domain (Figure 3.9). Similar results were reached for *S. cereale* at the Institute of Agricultural Crops in Germany (Bernd Hackauf, personal communication). Whether this gene contributes to the self-incompatibility reaction remains to be investigated. In a first step, the genome position of this thioredoxin gene in the ryegrass genome will be determined (see Chapter 6).

3.5 Conclusions

To unravel the molecular mechanism of self-incompatibility in ryegrass, we checked whether genes (or gene families) which are involved in the SI-response in Solanaceae and Brassicaceae could be involved in the SI-response in ryegrass. The DNA sequences amplified in *L. perenne* with primers derived from the S-RNases showed no homology to gene sequences of S-RNases or any other kind of RNases. This indicates that probably no RNase based SI-system is present in grasses. This is not unexpected, as in *L. perenne* the arrest of pollen tube growth takes place on the dry stigma surface. In contrast, in the species displaying an S-RNase based SI-system arrest of pollen tube growth takes place in the style and the SI-recognition process occurs within the growing pollen tube. Probably little similarity can be found between the Solanaceae and the Poaceae SI-systems beyond the nature of the genetic control of the SI-system which is gametophytic in both cases.

In our experiment, the primer pair derived from the SRK gene amplified two fragments of 500 bp and 200 bp at DNA level and one fragment of 400 bp on cDNA from leaves and self-pollinated pistils. These sequences displayed homology to receptor kinases. However, the 400 bp PCR-product was amplified in cDNA of all tissue types analysed, including leaves. Possibly, the sequence amplified does not correspond to a gene involved in the SI-response, but it represents a kinase gene with unknown biological function. However, at present the involvement of this gene in the SI-response cannot be ruled out. It could correspond to a gene located near the S- or Z-locus, but not the S-locus itself. Similarly, in rye (Poaceae) a 280 bp fragment was identified in genomic DNA with primers derived from SLG

sequences. This fragment displayed a polymorphism that co-segregated with the S-alleles of 46 plants tested (Wehling et al. 1994a), what indicated that it was either the S-gene or that it was located in the neighborhood of the S-locus. However, as far as we know, the position of this gene in the rye genome has never been determined. In ryegrass, the EST marker derived from the PCR product of 400 bp amplified with primer pair S-SRK2/AS-SRK2 mapped on LG4 in the ILGI population (see Chapter 6).

Finally, we screened the *L. perenne* genome for homologues of the thioredoxin gene, putatively involved in the SI-response of *P. coerulescens*. A *L. perenne* homologue was amplified in cDNA not only from leaves but also from non- and self-pollinated pistils. This contrasts with the observations of Li et al. (1994), who found that the *Bm2* gene was only expressed in mature pollen and not in other *P. coerulescens* tissues. It is now clear that this gene does not represent the S-locus, but whether it is involved in the SI-response is still unclear. More recent data show that *Bm2* maps 2 cM from the S-locus in *P. coerulescens* (Langridge et al. 1999). Therefore, we determined the map position of the thioredoxin sequence isolated in *L. perenne*, to determine if this gene is linked to the S-locus, also in ryegrass. The thioredoxin gene maps on LG1, 3 cM away from the S-gene in the ILGI population (see Chapter 6).

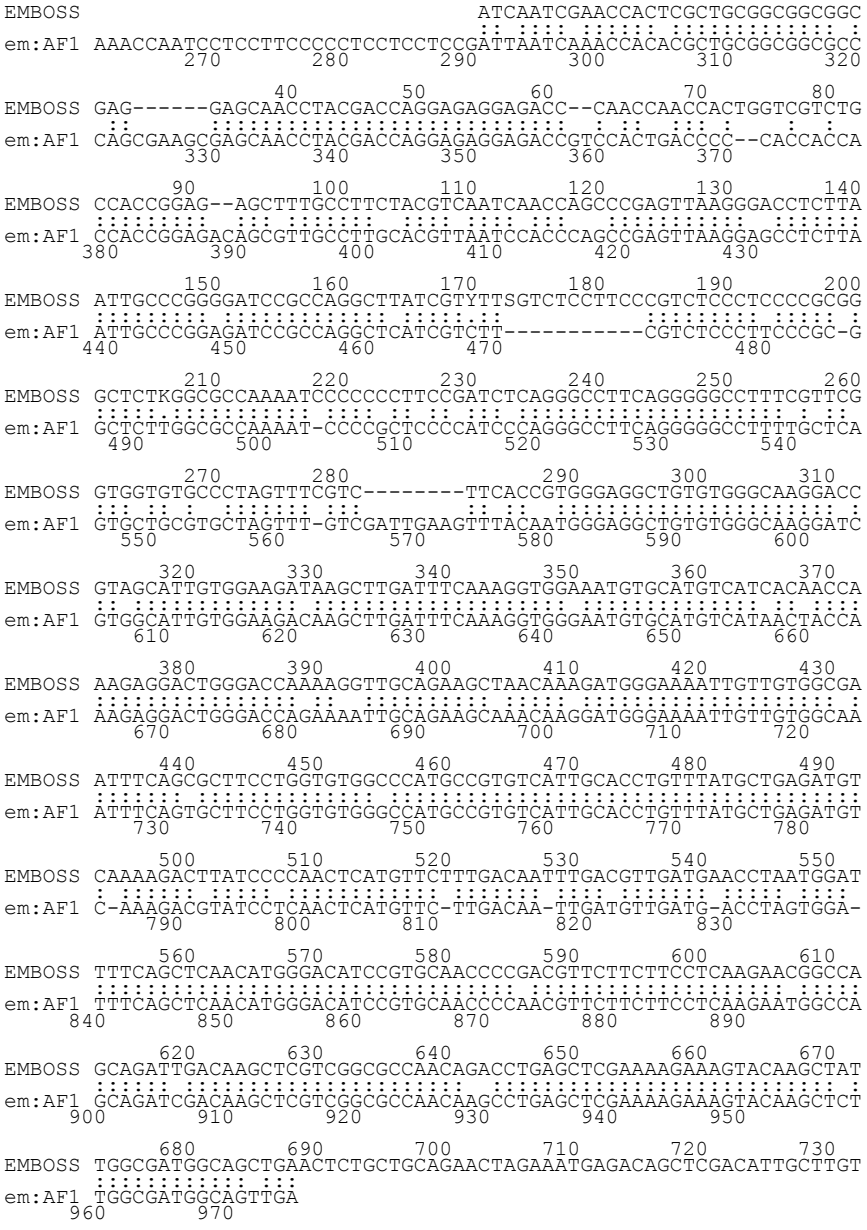


Figure 3.9 Sequence homology of 85% in 705 bp overlap between the 800 bp pistil cDNA fragment of ryegrass and the mRNA of the thioredoxin-like protein of *Phalaris coerulescens* (em_pl: AF 159389).

Chapter 4

Identification of genes involved in the SI-response using the cDNA-AFLP technique

4.1 Introduction

All SI-systems can be thought as consisting of two phases: a recognition phase, determined by the proteins encoded in the SI-specificity-determining locus (or loci), and a response phase, triggered by the initial recognition event and orchestrated by genes that reside outside the SI-specificity loci (Nasrallah 2005). Also in the grass family, although S and Z determine the recognition process between pollen and stigma, it is assumed that, similarly to what has been found in other plant families (see chapter 2), many other genes are involved in the cascade of events that is triggered by an incompatible reaction. A study of patterns of gene expression in the pistil can thus help us to gain insight into these processes and to identify the genes which are involved in the signalling cascade which follows a SI-reaction.

In general terms, a study of differential gene expression involves the comparison of mRNA populations between two samples. A variety of molecular techniques are available, including hybridisation-based, sequencing-based and PCR-based methods. Some techniques concentrate on the analysis of specific genes or gene families, for which probes or DNA-sequence information are available, such as Northern blotting (hybridisation-based) or RT-PCR (illustrated in Chapter 3). Other techniques allow a global analysis of gene expression. In this latter case expression changes of numerous genes are monitored simultaneously, and in some cases novel genes can be identified. Northern blotting remains the method of choice for confirming differential mRNA expression identified by screening techniques such as cDNA microarrays or cDNA-AFLP but nowadays Real-Time RT-PCR is also becoming widely used for this purpose. An overview of many of the methods available can be found in Lorkowski and Cullen (2002). This introduction will concentrates on techniques which allow **genome-wide transcript profiling**.

The first attempt to measure global levels of gene expression was based on large-scale Expressed Sequence Tag (EST) sequencing. This approach is however, too laborious and costly and alternative approaches were developed. Technologies for the study of differential gene expression can be subdivided into four main types: (i) microarray hybridisation, (ii) counting of sequence tags or signatures from cDNA fragments, (iii) subtractive hybridisation and (iv) gel-based analysis of cDNA tags. Each of these techniques will be briefly presented in the following paragraphs, with special attention for their advantages and disadvantages.

4.2.1 Microarray hybridisation

Both cDNA and oligonucleotide microarrays have proven powerful and are extensively used for genome-wide expression analysis in a wide range of organisms, including plants (Breyne and Zabeau 2001). The two main approaches involve either *in situ* synthesis of oligonucleotides ('oligonucleotide microarrays') or deposition of pre-synthesised DNA fragments ('cDNA microarrays') on solid surfaces. The strength of these techniques lies in the parallel nature of the analysis, which allows analysing simultaneously up to tens of thousands of genes. For some organisms such as yeast (Wodicka et al. 1997), a number of bacteria (Laub et al. 2000) and plants such as *Arabidopsis*, microarrays comprising complete gene sets are available. For a few other organisms, available microarrays contain a subset of the genes (see chapter 7 for the description of a *L. perenne* microarray).

The most important advantage of microarray-based technology is that large data sets from different experiments can be combined and analysed together in a single database (Breyne and Zabeau 2001). However, the current microarray technology has limitations. The most important one is the fact that for many organisms only a fraction of the genes is available and can be surveyed. Another important restriction is the difficulty to distinguish among different transcripts from genes of the same gene family. Techniques such as SAGE or cDNA-AFLP (see further) overcome this difficulty.

4.2.2 Counting of sequence tags

One of the best known techniques within this category is SAGE (Serial Analysis of Gene Expression). SAGE is a high-throughput technique to gain a qualitative and quantitative measure of gene expression without having prior sequence information (Velculescu et al. 1995). SAGE is based on the isolation of unique DNA sequence tags, derived from a defined position in individual mRNAs, which are concatenated serially into long DNA molecules for sequencing. In a first step, RNA is extracted from the tissue of interest and is transcribed to cDNA. Upon cleavage of the cDNA with an anchoring restriction enzyme the 3'-ends are collected and linkers are attached. The tagging restriction enzyme recognises its cleavage signal in the linker and cleaves downstream, resulting in the release of a small cDNA tag. These linker tags are ligated into linker ditags, which are amplified by PCR. Then the ditags are released by incubation with the anchoring restriction enzyme and ligated into concatemers. The concatemers are cloned into a plasmid

and sequenced to obtain information on gene expression. By counting the number of tags from each cDNA, one obtains an accurate measure of the absolute number of transcripts present in the mRNA sample. One important characteristic of SAGE is that independent data sets can be compiled in a single database allowing the comparative analysis of data from different experiments.

The principal advantage of SAGE is that it provides an absolute measure of gene expression instead of measuring relative expression levels. This technique is useful for the study of both rare and abundant transcripts and for the discovery of novel genes (Velculescu et al. 1995). However, as it relies on the knowledge of the genomic sequence of the species of interest and on the availability of SAGE databases to identify genes and gene expression profiles, it can only be used in organisms for which extensive sequence information data is available in databases. SAGE has been widely used in humans, especially for cancer research (Strausberg 2001), but not much in other organisms. This might change in the future, once the complete genome sequence of more plant species become available.

4.2.3 Subtractive hybridisation

For this technique, different modifications exist (reviewed by Lorkowsky and Cullen, 2002) but the principle is similar; the so-called 'tester' transcriptome (cDNA), which is the sample of interest, is hybridised with an excess amount (cDNA) of control genome, the so-called 'driver', in which the target sequence is absent (Gray and Collins 2000) and followed by separation of double-stranded nucleic acid hybrids (cDNAs in common) from single-stranded cDNAs (differentially expressed mRNAs). The resulting single-stranded target or subtracted cDNA is either used as labelled probe to screen genomic or cDNA libraries or for the construction of a subtracted cDNA library. To characterise the differentially expressed transcripts, sequencing is required. This technique has the advantage of requiring no specific biochemical information relating to the genes to be investigated. It can therefore be used to identify not yet characterised genes in the organism under investigation. The main disadvantages are the relatively large amount of plant material needed and the fact that it identifies abundant mRNAs more easily than rare mRNAs. In addition, the procedure is very laborious if several samples need to be compared since only two stages or conditions can be compared at the same time.

4.2.4 Gel-based analysis of cDNA tags

The first PCR-based techniques which were developed for whole-genome transcript profiling (RAP-PCR (Welsh et al. 1992) and differential display DD/RT-PCR (Liang and Pardee 1992) made use of arbitrary primers. The biggest advantages of these techniques over hybridisation-based techniques is that several developmental stages or tissue types can be compared at the same time and they allow the systematic screening of nearly all transcripts in a given biological system using small quantities of start material. However, the techniques based on the use of arbitrary primers have important problems with reproducibility (Bachem et al. 1996). Such problems arise primarily from the use of random oligonucleotides for PCR priming and from the relatively low annealing temperatures necessary to produce amplification products. To overcome these problems another method was developed: RNA fingerprinting based on AFLP, namely cDNA-AFLP (Bachem et al. 1996). Along with other PCR-based RNA fingerprinting methods, cDNA-AFLP has the great advantage of not requiring prior sequence information and it affords the possibility of screening large numbers of controls, simultaneously with the treatments used for the experiments. Due to the highly stringent conditions used during PCR, the rate of false-positives produced by mis-priming is low.

The original cDNA-AFLP method consists of four steps: 1) synthesis of cDNA using a poly-dT oligonucleotide, 2) production of primary template by restriction digest with two restriction enzymes (a rare cutter and a frequent cutter) and ligation of anchors to their termini, 3) pre-amplification with primers corresponding to anchors from the secondary template and 4) selective restriction fragment amplification with primers extended with one or more specific nucleotides. The final fingerprint is produced by radioactive labelling of one of the primers, allowing visualization of the amplification products. The original cDNA-AFLP protocol (Bachem et al. 1996) was modified by Breyne and Zabeau (2001) to permit a better transcript profiling of differentially expressed genes. In the modified protocol the number of tags which are amplified per transcript was reduced to one. The most frequently cutting restriction enzymes have several recognition sites per cDNA and as a result, several restriction fragments are produced from one single transcript after restriction digestion. By selecting the 3' end restriction fragment (Figure 4.1) only one AFLP tag is obtained per transcript, significantly reducing the total number of tags to be screened. The appropriate restriction enzymes for *A. thaliana* were selected by Breyne et

al. (2003) based on an *in silico* analysis of full-length cDNAs. They checked several tetra-, penta- and hexa-cutters and concluded that the enzymes with 4-base recognition sites provided the highest cDNA coverage, although the tags obtained were relatively short and the 3' end site occurred closer to the poly(A) tail than with 5- or 6-base cutters. On the other hand, most of the hexa-cutters generated more informative tags, but less than half of all the cDNAs are covered (Breyne et al. 2003). Among the different restriction enzymes tested, the enzyme *Bsf*YI was found to be the most appropriate in *A. thaliana* and was chosen in combination with *Mse*I for cDNA-AFLP analysis. The optimal number of selective nucleotides to screen and visualize the majority of both abundantly and weakly expressed genes was also determined *in silico* (Breyne et al. 2003). Of course, this kind of *in silico* approach is not possible in species for which no complete annotated sequence data are available. In this case the optimal conditions for a cDNA-AFLP experiment are determined experimentally.

Since the publication of the methodology, cDNA-AFLP has proven to be a robust and reliable method for monitoring the differential expression of genes in a wide range of biological systems. For example, Bachem et al. (2001) identified a gene which plays a role in the tuber life cycle and plant development in potato using cDNA-AFLP. In *Poa pratensis*, 179 cDNA-AFLP sequences were identified that differed qualitatively and quantitatively between apomictic and sexual genotypes (Albertini et al. 2004). In rice, 34 aluminium-regulated genes have been identified by cDNA-AFLP (Mao et al. 2004). cDNA-AFLP was also used to isolate genes involved in plant-pathogen interactions in *Chenopodium amaranticolor* (Cooper 2001), cassava (Kemp et al. 2005), sugar beet (Samuelian et al. 2004), tomato (Durrant et al. 2000) and barley (Eckey et al. 2004). In chicory, genes involved in root development have been identified using cDNA-AFLP (Goupil 2003). In all these applications, changes of gene expression were visualised under a wide range of biotic and abiotic conditions. In all cases, process-specific alterations of gene expression were identified that led to the isolation of genes directly or indirectly involved in the process under investigation. This demonstrates the utility of cDNA-AFLP for the study of biological questions involving induction or repression of gene expression. In a further advancement of the technique, cDNA-AFLP has also been used for the visualization of genetic variation within mapping populations and mapping the transcribed regions of the genome, the transcriptome. For example, Brugmans et al. (2002) used cDNA-AFLP to study transcripts expressed in segregating populations of *A. thaliana* and potato (*Solanum tuberosum*).

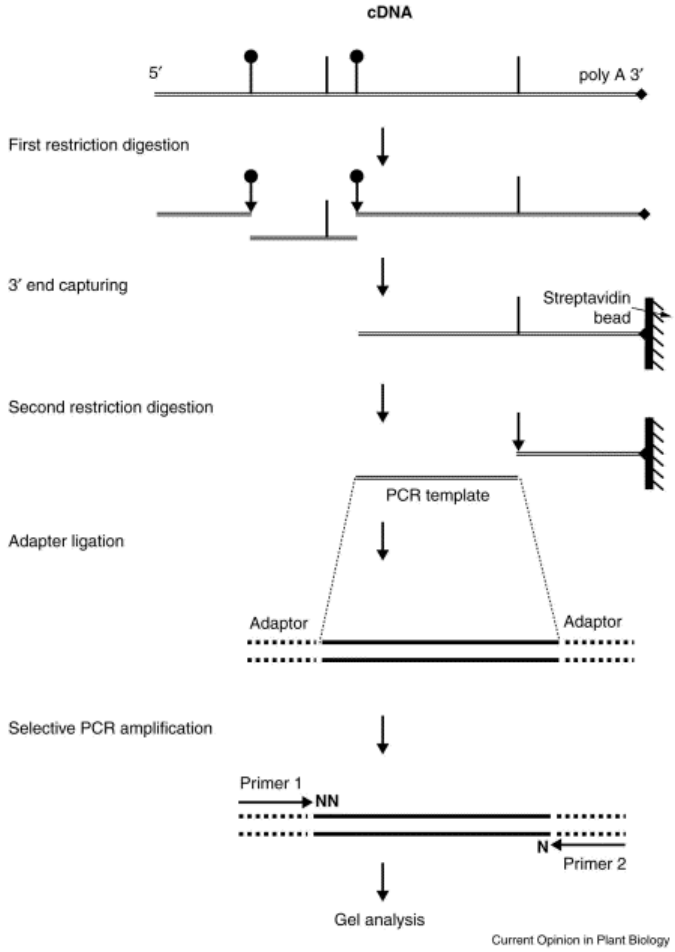


Fig 4.1 Outline of the cDNA-AFLP procedure. Messenger RNA is converted into double stranded cDNA using a biotinylated oligo-dT primer. After digestion with a first restriction enzyme, the 3' termini of the cDNA are recovered by binding to streptavidin beads. Following digestion with a second enzyme, restriction fragments are released that serve as PCR templates. After ligation of site-specific adaptors, primers that match the adapter sequences and that carry selective nucleotides at their 3' end are used to amplify subsets of the transcript fragments. Following amplification, fragments of between 100 and 1000 bases are separated and displayed on polyacrylamide gels (Breyne and Zabeau 2001).

However, as with all analyses involving DNA retrieval from denaturing polyacrylamide gels, in cDNA-AFLP experiments false positives are generated by co-amplification of unrelated sequences that may either underlie the targeted band or co-migrate very closely with the chosen

fragment (Lorkowki and Cullen 2002). Unfortunately, there is no method of getting around this problem.

In general terms, we can conclude that cDNA-AFLP is a very sensitive system, able to detect gene expression of rare messages and especially useful when tiny amounts of start material are available. The flexibility of cDNA-AFLP, in terms of choice of enzyme combinations, anchor design and primer extensions, allows the system to be readily adapted to a wide variety of biological problems.

4.2 Objectives

The objective of the experiment presented in this chapter was to carry out a cDNA-AFLP based genome-wide transcript profiling of *L. perenne* pistils. Expression profiles of non-pollinated, self-pollinated and cross-pollinated pistils at several time points after pollination were compared together with leaf material and pollen. This allowed us to identify genes which are differentially expressed in non-pollinated, self-pollinated and cross-pollinated pistils and which are putatively involved in the signalisation cascade triggered by an SI-reaction. The cDNA-AFLP expression profiles were analysed with the software AFLP-QuantarPro^{TM} (Keygene), which allowed us to identify clusters of transcripts with specific expression patterns during compatible and incompatible reactions.

4.3 Materials and methods

4.3.1 Plant material

Four *L. perenne* genotypes were used for controlled pollinations and RNA extraction:

- Genotypes LpM and LpN were chosen at random from the DvP collection.
- Genotypes LpT_3 and LpT_6 were two full-sibs derived from a cross between two plants selected at random from the DvP collection ($S_{22}6/16$ x $S_{22}6/56$).

In the first place we checked whether reciprocal pollinations between LpM and LpN and pollinations between LpT_3 and LpT_6 resulted in fully compatible reactions, using *in vitro* pollination tests. We used a slightly adapted version of the method described by Cornish (1979). Spikelets were removed from the inflorescences of the mother plants and whole pistils

(mature ovaries with stylodia attached) were collected from basal florets of spikelets just before anthesis. The mature ovaries were placed on petri-dishes plated with agar (2.5% agar, 25% sucrose and 25 ppm Boric acid) which provided a semi-solid medium on which the ovaries could be supported so that the stigmas were free from the medium and ready for good pollen reception. Two ovaries from the mother plant were placed on a single petri-dish. Just before anthesis, the pollinator plants were placed in a growth chamber at 23°C with a light level of about 300 µmol/m^2/s and low humidity, important for the viability of the pollen. To collect pollen, the inflorescences of the pollinators were enclosed in cellophane bags. The pollinator plants shared pollen around midday. The pollen collected in the bags was released by shaking over the pistils, ensuring that only free-flowing, non-clumped, viable pollen was used. Anthers in the bag were removed prior to pollination to ensure smooth pollen flow onto the plates. The pollinations were checked using the UV fluorescence microscopy method described in section 3.3.1. All tests performed resulted in fully compatible reactions, meaning that all the germinated pollen grains produced full-grown pollen tubes. These pairs of genotypes are thus suitable for the isolation of pistils after a compatible pollination.

4.3.2 Isolation of self-pollinated and cross-pollinated pistils

For the isolation of self-pollinated and cross-pollinated pistils, the *in planta* pollination strategy described in section 3.3.1 was followed. LpM and LpT_6 were used as female parents and were therefore emasculated. LpN and LpT_3 were used as pollinators to fertilize LpM and LpT_6 respectively.

Before pistils were harvested, we always checked that the cross-pollinations were successful and that the self-pollinations resulted in malformed pollen tubes, as described in section 3.3.1. If the expected result was observed, pistils (around fifty) were harvested at four different time points after pollination: 1, 4, 8 and 24 hours. We also harvested non-pollinated pistils, pollen and leaf material of the female plants and pollen of the pollinator plants. This resulted in a total of 12 samples to compare per pair (Table 4.1). These materials were kept at –80°C till RNA extraction.

cDNA-AFLP primer combinations 1-105 were carried out on plant materials collected from the genotypes LpM and LpN (LpM as female plant and LpN as pollen donor). cDNA-AFLP primer combinations 106-256 were

carried out on materials collected from LpT_6 and LpT_3 (LpT_6 as female plant and LpT_3 as pollen donor).

Table 4.1 Samples collected per cross (either LpM x LpN or LpT_6 x LpT_3) for RNA extractions and cDNA-AFLP analysis.

Genotype	Pistils									Pollen	Leaf
	NP	SP (1h)	SP (4h)	SP (8h)	SP (24h)	CP (1h)	CP (4h)	CP (8h)	CP (24h)		
Female plant	X	X	X	X	X	X	X	X	X	X	X
Pollen donor										X	

NP: Non-pollinated; SP: Self-pollinated; CP: Cross-pollinated

4.3.3 RNA isolation, cDNA synthesis and template preparation

Total RNA was extracted from 20 to 50 mg of fresh plant material using the 'guanidinium thiocyanate method' described by Chomczynski and Sacchi (1987). Double stranded cDNA was synthesised from total RNA (between 25-50 µg) using a biotinalyted oligo$(dT)_{25}$ primer according to standard protocols (Sambrook et al. 1989). The resulting double stranded cDNA was chloroform/isoamylalcohol extracted, ethanol precipitated and taken up in a final volume of 20 µl of water. The restriction enzymes used were *Bst*YI (rare cutter) and *Mse*I (frequent cutter) and the digestion was performed in two separate steps (Breyne and Zabeau 2001). After the digestion with the rare cutter, the 3' end fragments were collected using streptavidine magnetic beads, while the other fragments were washed away. After digestion with the second enzyme, the restriction fragments released from the beads were collected and used as template in subsequent AFLP steps. Adaptors were ligated to the ends of the restricted fragments and 25 cycles of pre-amplification (30 sec at 94°C, 1 min at 56°C and 1 min at 72°C) were performed on 5 µl of undiluted 'adaptor ligation mix' with primers without selective nucleotides and complementary to the anchors, following the standard procedure of AFLP reactions described by Vos et al. (1995). The products of the preamplification were checked on 1.5% agarose gels (a smear between 100 and 1000 bp is expected after successful preamplification). The template was diluted 10-fold in TE buffer. The

following adaptors and primers were used: *Bst*YI adaptor: 5'-CTCGTAGACTGCGTAGT-3' and 5'-GATCACTACGCAGTCTAC-3'; *Bst*YI primer: 5'-GATCACTACGCAGTCTAC(N_{1-2})-3'; *Mse*I adaptor: 5'-GACGATGAGTCCTGAG-3' and 5'-TACTCAGGACTCAT-3'; *Mse*I primer: 5'-ACGATGAGTCCTGAGTAA(N_{1-2})-3', where N represents the selective nucleotides. For the PCR reactions all 16 possible primers of two base extensions were tested, rendering a total of 16^2 (256) primer combinations.

4.3.4 Radioactive cDNA-AFLP reactions

Radioactive labelling of the *Bst*YI primer was carried out as described by Vos et al. (1995). Thermocycling was carried out with 35 cycles, including a 12 cycle touchdown (20 sec 94°C, 30 sec at 65°C, annealing temperature was reduced from 65°C to 56°C in 0.7°C steps for 12 cycles, 55 sec at 72°C), maintained at 56°C for 23 cycles and finally 15 min at 72°C. Samples were then boiled after the addition of loading dye and 50% formamide and separated on 50 cm, 5% polyacrylamide gels. All gels were run at standard conditions with a 10-bp ladder (Sequamark, Genetic Research) as length marker. Gels were dried on Whatman 3M paper using a slab gel dryer. Labelled DNA fragments were visualized by autoradiography. Gels and films were positionally marked prior to development. The two-nucleotide selective extensions yielded between 50 and 70 labelled fragments per lane.

4.3.5 Quantitative measurements of the expression profiles and data analysis

In a second step the gel autoradiograms were scanned and, after visual inspection and selection of a number of differentially expressed tags, the gel images were analysed quantitatively with AFLP-QuantarPro™ image analysis software (Keygene, Wageningen, The Netherlands). This software was designed for accurate lane definition, fragment detection and quantification of band intensities. All the interesting, differentially expressed AFLP fragments were scored and individual band intensities were measured. These data were used to determine the quantitative expression profile of each transcript, after correction and standardization. The raw data were first corrected for differences in overall lane intensities, which may arise due to loading errors or differences in the efficiency of PCR amplification with a given primer combination for one or more samples. A correction factor was estimated for each primer combination using the intensity of bands that remained constant across the different samples

analysed. For each primer combination, 8 to 10 invariable bands were selected and the intensity values of these bands were summed per lane. Each summed value was divided by the maximum value to obtain the correction factor for each primer combination. The raw values generated by QuantarPro for each differentially expressed transcript were divided by this correction factor.

Subsequently, each individual gene expression profile was variance-normalised by standard statistical approaches as used for microarray-derived data (Tavazoie et al. 1999). For each transcript, the mean expression value across the different samples was subtracted, and the value obtained was then divided by the standard deviation of the transcript across the different lanes. A coefficient of variation (CV) was calculated by dividing the standard deviation by the mean. This CV value was used to establish a cut-off value: expression profiles with a CV below 0.25 were considered to be constitutively expressed (Breyne et al. 2003).

Cluster analysis of the expression data was carried out using two complementary methods. First, the TIGR MultiExperiment Viewer software (version 2.2), which is freely available on the TIGR website (http://www.tm4.org/mev.html), was used for hierarchical, average linkage clustering (Saeed et al. 2003). The resulting clustering shows a hierarchical tree, in which each line represents a gene. The expression values are presented with color codes: red stands for high expression values and green for low expression values. Black is considered to be the mean. Missing data are represented in grey. The statistical support of the tree nodes was calculated using the program Support Trees (TIGR MultiExperiment Viewer software version 2.2), based on resampling strategies. For this, new trees were made by Jacknifing (resampling and leaving one gene out) the genes. The resampling was done on the genes by a user-specified number of iterations (100). The branches of the resulting tree were colour-coded to denote the percentage of times that a given node was supported over the resampling trials.

The second method used was the Quality-Based Clustering, performed with a recently developed software program which is accessible for use on http://www.esat.kuleuven.ac.be/~thijs/Work/Clustering.html (De Smet et al. 2002). This program is similar to K-means clustering, except that the number of clusters do not need to be defined in advance by the user and the expression profiles that do not fit in any cluster at user-specified threshold

values for specific 'quality' parameters are not assigned to any cluster. As output, genes are grouped into different clusters and are represented in graphs. The minimum number of tags in a cluster was set to 5 in this experiment.

4.4 Results and discussion

4.4.1 Cross-compatibility relationships and isolation of plant material

In a first step, the compatibility relations between M x M, M x N, T_6 x T_6 and T_6 x T_3 were determined using the UV fluorescence microscopy technique described by Martin (1959) and in section 3.1.1. With this UV fluorescence microscopy method it is possible to distinguish viable, fully-grown pollen tubes (a compatible reaction) and short, deformed pollen tubes (an incompatible reaction) (Figure 4.2). In incompatible reactions the growth of the pollen tube is arrested at or near the stigma surface and it becomes occluded with callose that fluoresces brightly under UV light. Compatible pollen grains, on the other hand, are able to germinate successfully and become rapidly empty into the pollen tube. These pollen grains stain weakly because there is no callose deposition (Figure 4.2). The same staining method was used to check whether the pollinations were successful, before harvesting the (self- and cross-pollinated) pistils to be used for cDNA-AFLP analysis. The crosses M x N and T_6 x T_3 were fully compatible and the selfings M x M and T_6 x T_6 were fully incompatible.

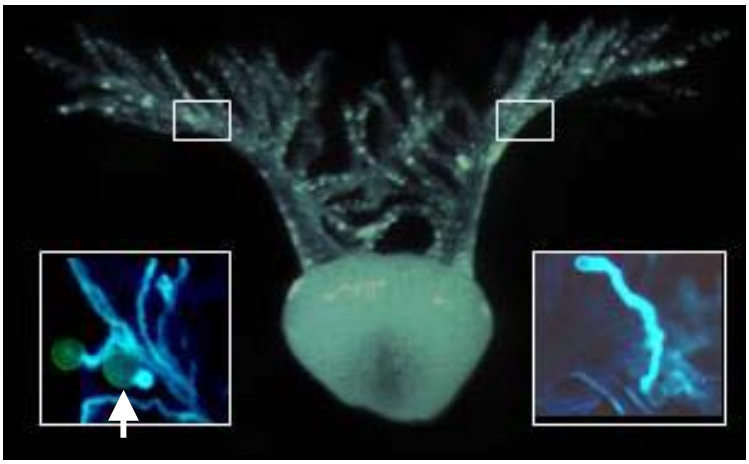


Figure 4.2 A mature pistil of ryegrass with the ovary, the two styles and the stigmas (25X). Left: an incompatible, short, deformed pollen tube (indicated by an arrow, 400X). Right: a compatible, full-grown pollen tube (400X).

4.4.2 Quantitative analysis of gene expression

We used the cDNA-AFLP technique to perform a quantitative analysis of genes putatively involved in the SI-response and in the fertilization process in *L. perenne*. An example of the generated cDNA-AFLP fingerprints is shown in Figure 4.3. The cDNA-AFLP fragments ranged in length from 60 to 600 bp and for each primer combination, 50 to 70 bands were visualized on a gel.

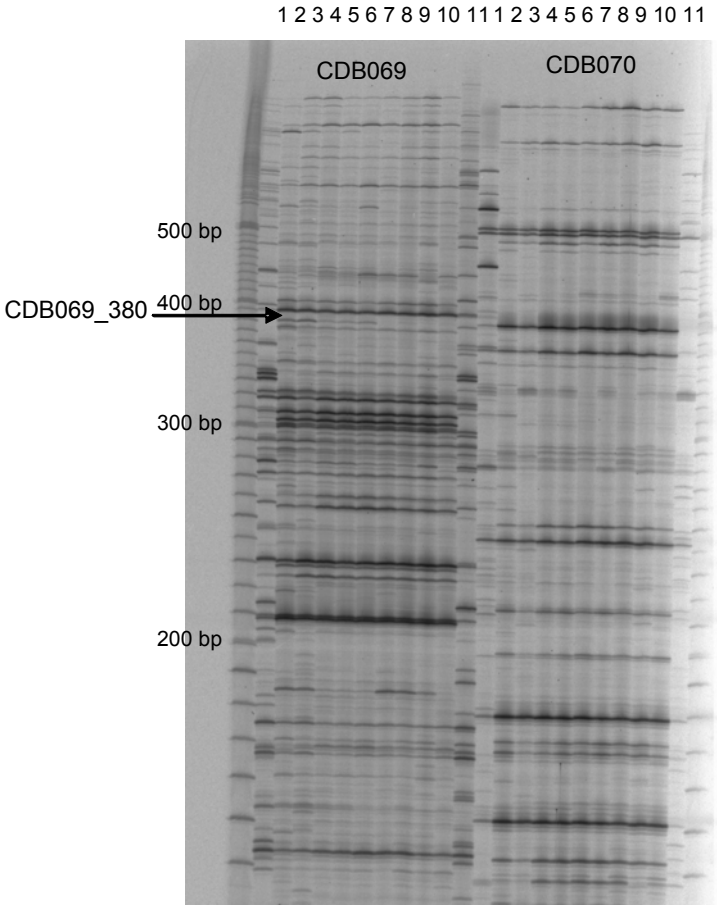


Figure 4.3 cDNA-AFLP gel with the primer combinations *BstY*I + AC and *Mse*I + CA (CDB069) and *BstY*I + AC and *Mse*I + CC (CDB070). 1: T_6 pollen; 2: T_6 non-pollinated pistils; 3 -6: T_6 x T_6 self-pollinated pistils after 1, 4, 8 and 24 hours of pollination; 7-10: T_6 x T_3 cross-pollinated pistils after 1, 4, 8 and 24 hours of pollination. 11: T_3 pollen. CDB069_380 represents a gene that is down regulated in the self-pollinated pistils and absent in the cross-pollinated pistils.

Per primer combination, the cDNA-AFLP fingerprints obtained were compared using AFLP QuantarPro^{TM} (Keygene, Wageningen, The Netherlands) to identify differentially expressed transcripts. Although differentially expressed and constant bands can be discriminated by visual scoring, automated analysis with AFLP QuantarPro^{TM} analysis is more sensitive and reliable and generates quantitative expression data (Breyne et al. 2003, De Paepe et al. 2004). In a first step all the transcripts in the size range 100-600 bp which displayed differential expression patterns, according to the visual inspection on the computer screen, were selected for processing. cDNA-AFLP bands which were amplified also in leaf samples were not selected, as these genes most probably have general biological functions and are not specifically involved in the SI-response. The raw band intensity data were first corrected for differences in total lane intensity. To calculate the correction factors, the band intensity values of the pollen samples were not taken into account as not enough constitutively expressed fragments were found between pollen and the rest of the samples (pistils). The corrected values were then standardized as described in material and methods (4.3.5). After normalisation of the data 7% of the tags that had been initially selected were excluded from the data set, as their coefficient of variation was smaller than 0.25 and were therefore considered to be constitutively expressed (Figure 4.4).

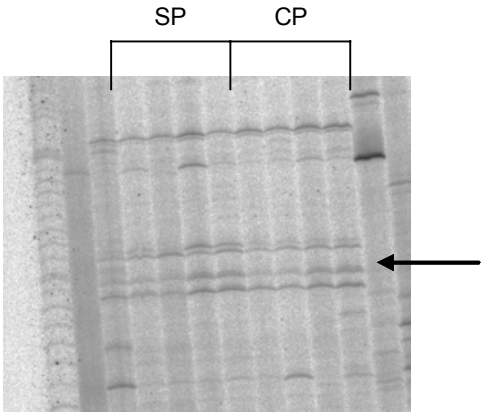


Figure 4.4 Fragment CDB237_390 initially thought to be differentially expressed. After normalisation this fragment was excluded from the data set as its coefficient of variation was <0.25.

Several patterns of differential expression were found among the remaining transcripts. However, only the expression profiles shown in Figure 4.5 and Table 4.2 represent genes putatively involved in SI as they are either down or up regulated in the self-pollinated pistils and absent or constitutively expressed in the cross- and non-pollinated pistils. Genes displaying other (differential) expression profiles are most probably not involved in the self-incompatibility response, as they are either up regulated or down regulated following a similar pattern in self- and cross-pollinated pistils. These genes could be involved in general growth, ripening or degradation processes of the pistils.

Of the around 18,000 transcripts that were screened in this experiment, 479 displayed a differential expression pattern after correction and standardization, and of these 151 displayed one of the expression patterns shown in Figure 4.5. The number of differentially expressed genes identified using the cDNA-AFLP technique depends strongly on the experiment that is performed. However, our results are in line with data obtained in previous experiments. For example, Durrant et al. (2000) identified around 290 differentially expressed transcripts during an Avr9- and Cf9-mediated defence response in tobacco cell lines. Zheng et al. (2004), on the other hand, identified only 12 differentially expressed fragments that control rice blast resistance.

4.4.3 Clustering of selected transcripts according to their expression patterns

The band intensities of the 479 transcripts which displayed differential expression patterns were subjected to data analysis using two alternative methods: the Adaptive Quality-Based Clustering (AQBC) method from De Smet et al. (2002) and the average linkage clustering method implemented in the TIGR MultiExperiment Viewer software (Saeed et al. 2003). Data on self-pollinated pistils 8h after pollination (SP8h) were not included in the analyses as they contained too many missing data and distorted the calculations.

Table 4.2 Overview of the types of expression profiles identified in the cDNA-AFLP analysis. Genes displaying one of these expression profiles are putatively involved in the SI-response. The ordering of the samples represents the way in which they were loaded on gel. SP: Self-pollinated, CP: Cross-pollinated, own: material collected on the plant used as female partner, foreign: material collected from the pollinator plants.

Expression profile	Leaf	Pollen	Pistils									Pollen
	own	own	Non-pollinated	SP 1h	SP 4h	SP 8h	SP 24h	CP 1h	CP 4h	CP 8h	CP 24h	foreign
1												
2												
3												
4												
5												
6												

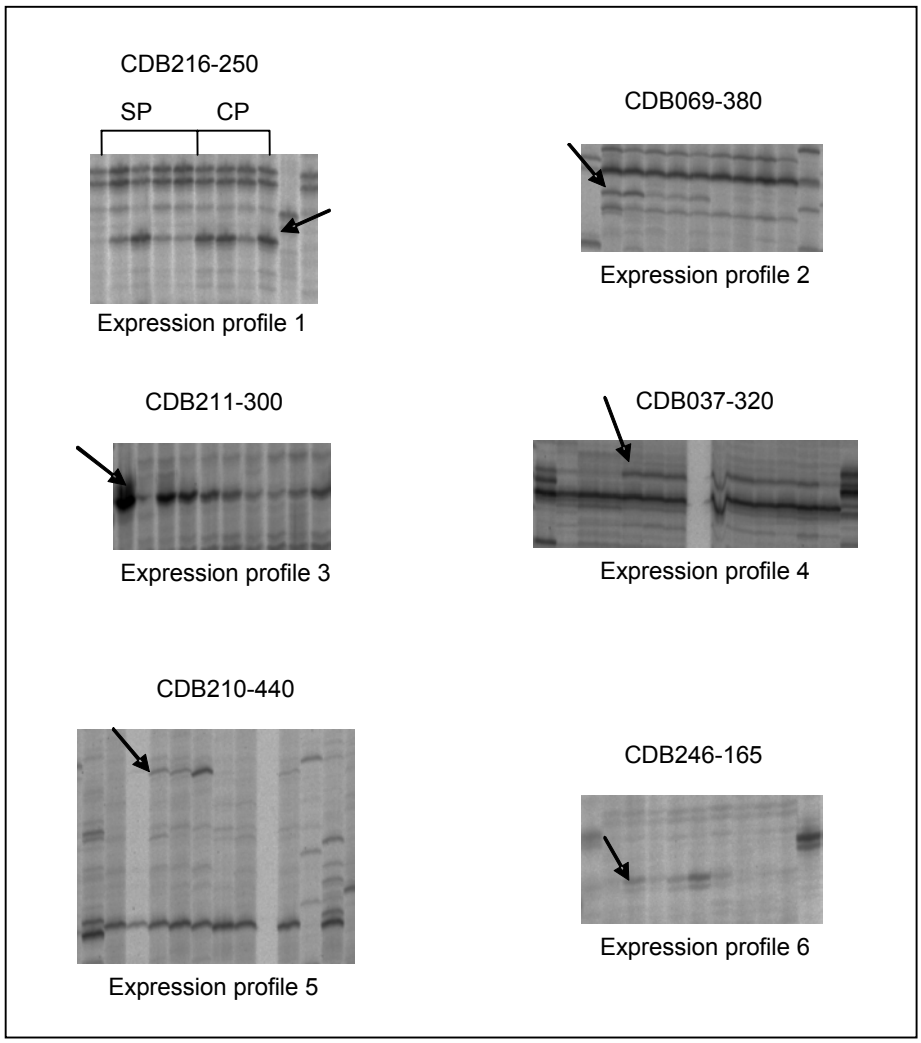


Figure 4.5 Examples of different expression profiles.

4.4.3.1 AQBC

The main advantage of the AQBC method above other clustering methods is that there is no need for predefinition of the number of clusters and expression profiles that do not fit in any cluster, according to predefined quality parameters, are rejected. This results in the formation of clusters of

significantly co-expressed genes. Two parameters, the minimal number of genes needed to form a cluster and the significance level (S), which has to be between 0.5 and 1 need to be defined by the user. In this study we set the minimum number of genes to 5. This implies that if a cluster is joined by less than five genes, it is not accepted as a 'real cluster' by the analysis. The default value for the significance level S (0.95) guarantees that a gene that has been assigned to a cluster, has a probability of 95% to belong to this cluster (this means that the probability of being a false positive is 5%). Different values of S were tested: 0.9, 0.93 and 0.95. The results obtained using different values of S were evaluated and compared. For a value of 0.9, 41% of the genes were clustered into 10 different groups. This value of S was not stringent enough, as genes with clearly different expression profiles were assigned to the same group. For a value of S=0.93 and S=0.95 AQBC identified 11 clusters comprising 36% of the genes and 10 clusters comprising 30% of the genes, respectively (Figure 4.6). Clusters 1 to 7 from the analysis with S=0.93 were also formed in the analysis S=0.95 but they contained less transcripts. No clear correspondence was found between the clusters 8-11 formed at S=0.93 and clusters 8-10 formed at S=0.95. We based the interpretation of the results on the results obtained for a value of S=0.93:

- Cluster 1 comprises transcripts that are down regulated during self-pollination (SI-reaction) but have also a low expression level in the cross-pollinated pistils (compatible reaction) 1 and 24 hours after pollination.
- Clusters 2 and 3 are probably the most relevant ones, as they contain transcripts with expression profiles fitting what is expected for genes involved in the SI-response. Both clusters contain transcripts which are down regulated in the self-pollinated pistils and absent in the cross-pollinated pistils. The main difference between the two clusters is the expression level in the non-pollinated pistils.
- Cluster 4 is probably less important with respect to SI as it contains transcripts which are constitutively expressed in self-pollinated pistils and which display differential expression in the cross-pollinated pistils. These can be genes involved in (successful) pollination and fertilization.
- Clusters 5 and 6 are formed by genes with a similar behaviour, down regulation in the self-pollinated pistils and up regulation in the cross pollinated pistils 24 hours after pollination.

- Clusters 7 and 8 are very similar and comprise genes with changing expression patterns which are difficult to explain as the genes are up, down and then again up regulated in the self-pollinated pistils.
- Clusters 9 and 10 comprise genes displaying increasing expression levels in both self-pollinated and cross-pollinated pistils. These clusters probably contain the genes which are involved in more general processes, such as degradation of the pistils at the end of the pollination.
- Cluster 11 groups together the genes with an up regulated expression profile in the self-pollinated pistils and a down regulated profile in the cross-pollinated pistils.

If we compare these results to the differential expression patterns which were defined in Figure 4.5, clear agreements are apparent (Table 4.3) and some clusters identified by the AQBC analysis match with the expression profiles shown in the figure. These groups of genes display expression profiles consistent with involvement in the SI-response in ryegrass.

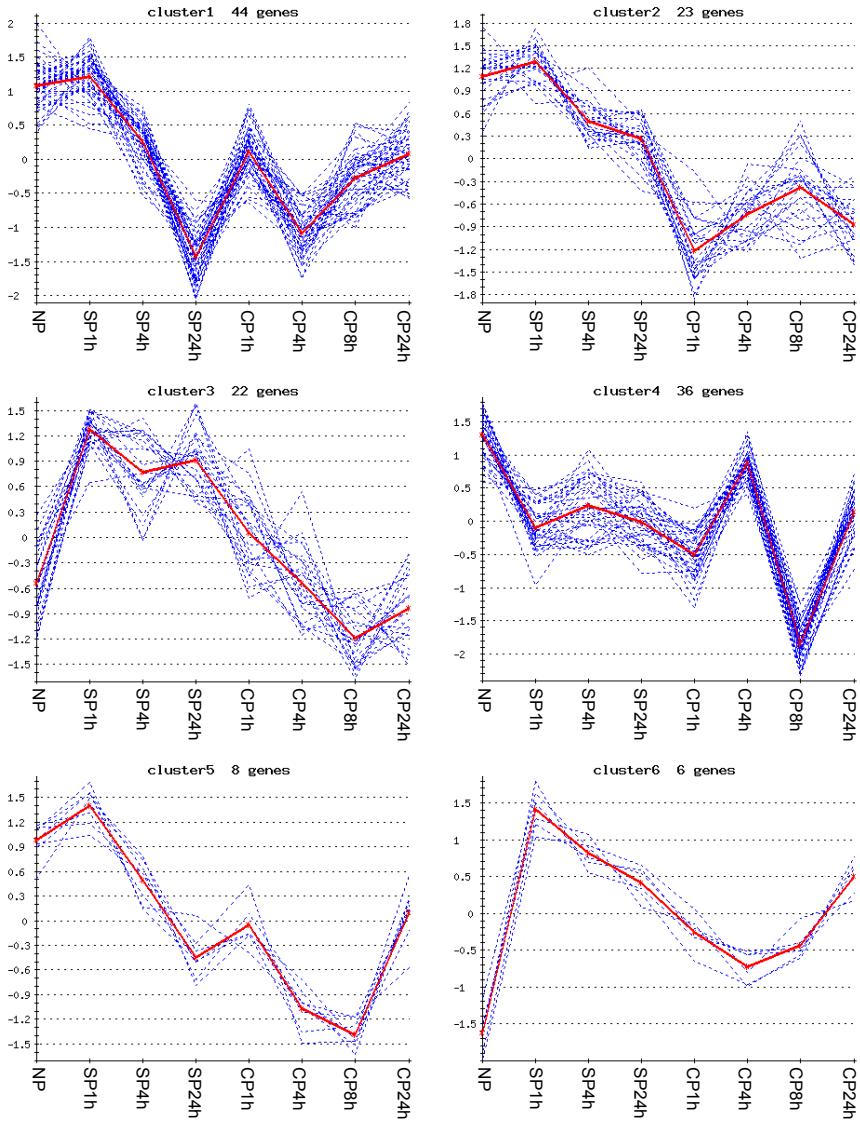
Table 4.3 Match between the expression patterns shown in Figure 4.5 and the results of the AQBC analysis.

Figure 4.5	AQBC analysis
Expression profile 1	Cluster 1
Expression profile 2	Cluster 2
Expression profile 3	Cluster 6
Expression profile 5	Cluster 11
Expression profile 6	Cluster 8

4.4.3.2 Hierarchical clustering

The results of AQBC were further compared to results obtained using hierarchical clustering. Average linkage hierarchical clustering classified the 479 genes into approximately ten groups of co-expressed genes (Figure 4.7). Some of the clusters identified by AQBC correspond to the groups identified by the average linkage clustering (the correspondence between the clusters identified by hierarchical techniques and the AQBC has also been represented in Figure 4.7). For the sake of clarity, we name ‘clusters’ the groupings identified by AQBC and ‘groups’ the groupings identified by average linkage. Group 2a corresponds to clusters 3 (black), 7 (pink) and

part of cluster 2 (green) from the AQBC analysis. Group 2b corresponds to clusters 3 and 6. The transcripts of cluster 2 are divided into two groups by the hierarchical clustering. All these transcripts display similar behaviour, strongly expressed in the self-pollinated pistils and not present in the cross-pollinated pistils. These transcripts are sometimes present and sometimes absent in the non-pollinated pistils. This is probably the most important group concerning SI.



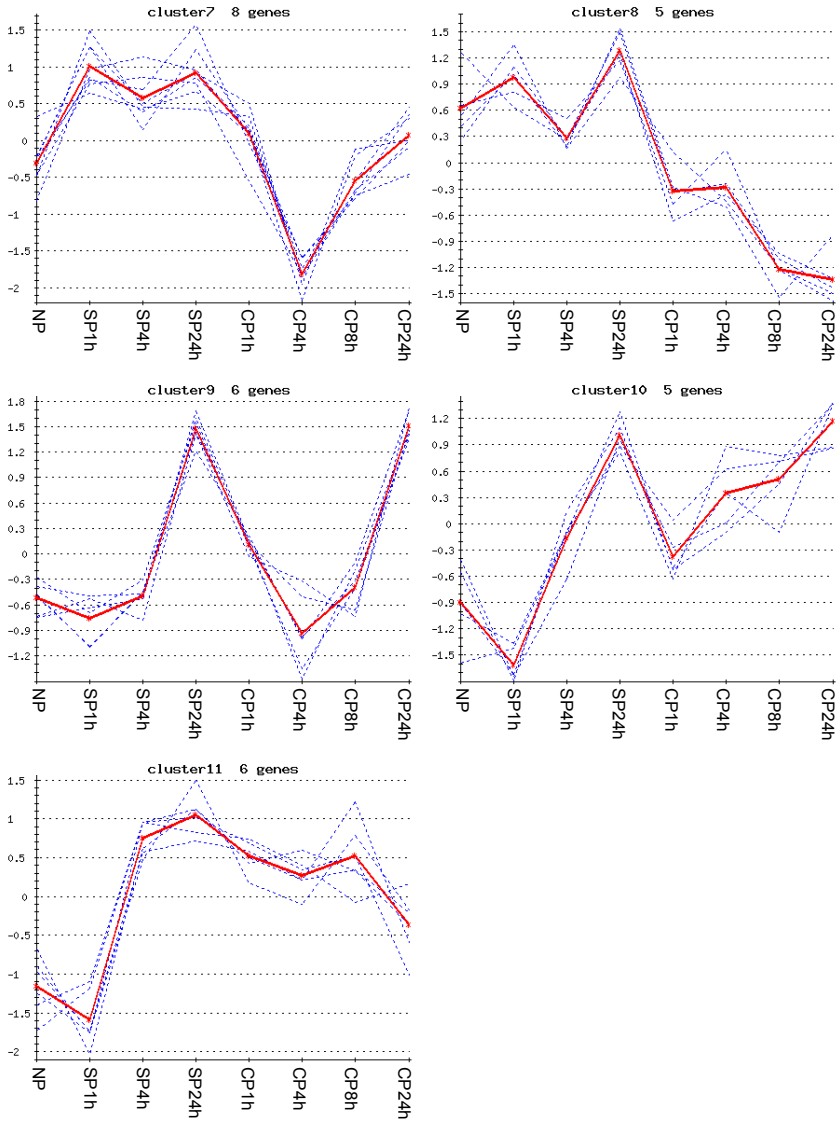


Figure 4.6 Clusters identified using the adaptive quality-based clustering (AQBC) method from De Smet et al. (2002). X-axis: samples, Y-axis: normalized intensity values.

In group 4, which is the largest, genes with a strong expression in the non-pollinated pistils and self-pollinated pistils after 1h of pollination and a down regulated in the other conditions cluster together. This group contains

all the genes of cluster 1 (red), which is the most consistent cluster of the AQBC analysis, and genes of cluster 5. This group seems to contain genes that are switched off during a self-compatible reaction, but which are strongly expressed before pollination or after a self-incompatible reaction.

Group 5 corresponds to cluster 4 (brown). These genes are probably not relevant for the SI- response as they are not expressed in the self-pollinated pistils. Groups 7-10 do not correspond to any of the AQBC clusters or correspond to clusters of less relevance, as explained in the previous section.

To evaluate the consistency of the clustering obtained by average linkage, support trees were constructed by Jacknifing the genes. The between-group support values were in general, very small. However, within-group support values of up to 70% were obtained (Figure 4.7). Hierarchical clustering formed consistent, but small groups.

4.4.3.3 Comparison of clustering techniques

The AQBC method is a user-friendly and fast approach which renders clusters containing only tightly related expression profiles. In comparison to previously proposed methods, such as K-means clustering, it does not require the predefinition of the number of clusters, and the 'quality' measure used is intuitive and meaningful. AQBC only assigned 36% of the genes analysed to clusters of co-expressed genes. The number of genes clustered in this experiment is lower compared to the results obtained by De Smet et al. (2002), in a microarray experiment in *Saccharomyces cerevisiae* 58% of the genes clustered in 42 clusters using the AQBC method. In the study of Breyne et al. (2002), 86% of all the tags were clustered in 21 distinct clusters, consisting of co-expressed genes and Vandenabeele et al. (2004) clustered 80% of the genes into 31 different clusters in *Arabidopsis*. This method provides, in a fast way, the different types of expression profiles found in the experiment. The hierarchical clustering method, by forcing each expression profile into a cluster, provides groupings which are less stringent than the clusters formed by AQBC and only groups of low hierarchical levels (containing few genes) are consistent. However, comparable groups of co-regulated genes were obtained with both techniques and we can conclude that the application of the two methods in combination is useful as an independent confirmation of the grouping of expression profiles.

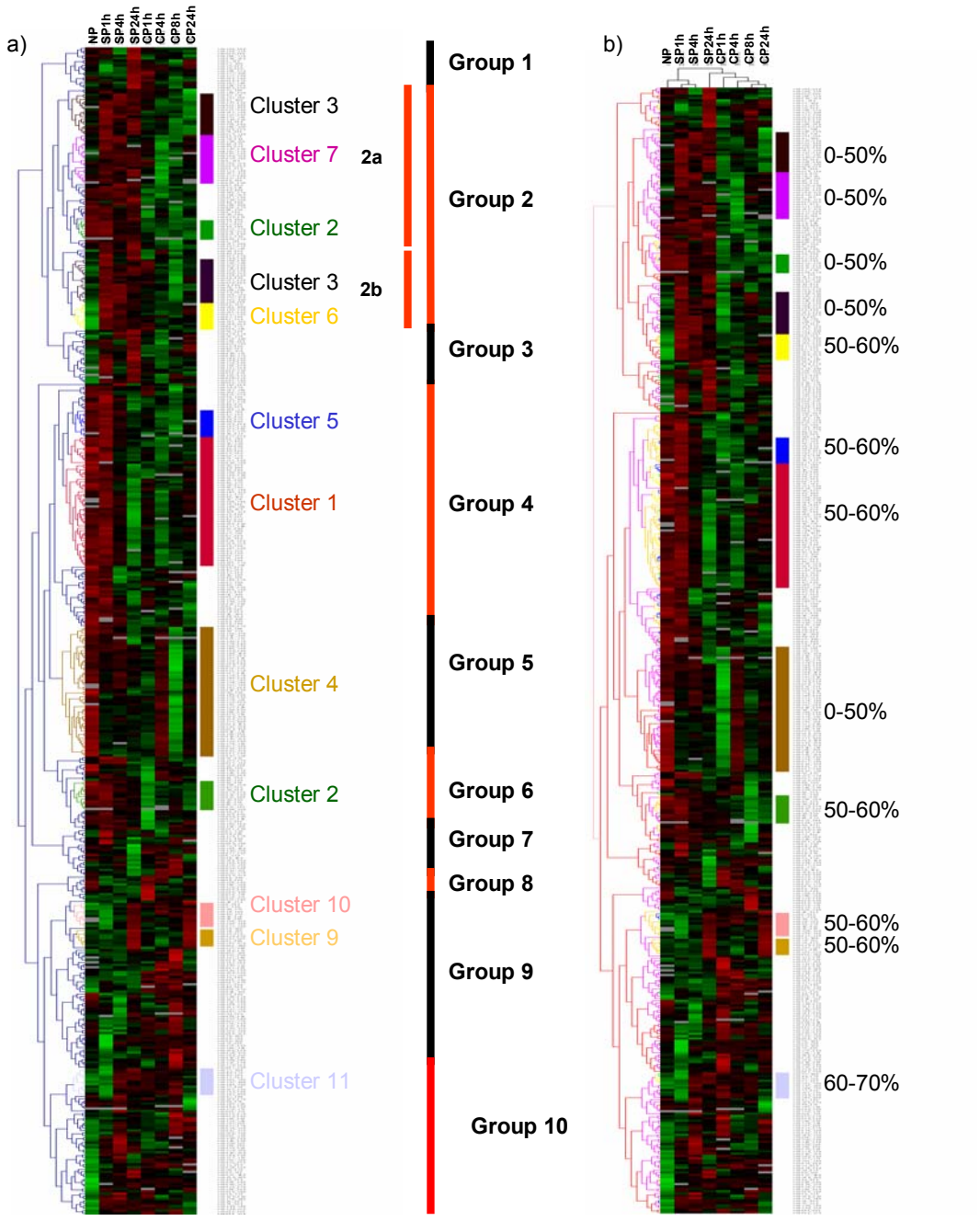


Figure 4.7 (a) Hierarchical clustering of the cDNA-AFLP fragments. **(b)** Jacknifing and Hierarchical clustering of the cDNA-AFLP fragments. The expression values are represented with color code: red=high expression values, green=low expression values and black is considered to be the mean.

4.5 Conclusions

The aim of this experiment was to identify genes involved in the SI-response in ryegrass. Using cDNA-AFLP and without any prior assumption about the kind of genes which could be induced or repressed during SI, we were able to analyze transcriptional changes in *L. perenne* pistils and to identify genes putatively involved in SI. According to our results, transcriptional changes take place during fertilization in *L. perenne*. Of the around 18,000 transcripts that were screened, 479 displayed a differential expression pattern.

Clusters 1, 2, 3, 5, 6 and 7 contain genes displaying interesting expression profiles concerning the self-incompatibility response: down regulation in the self-pollinated pistils and not expressed in the cross-pollinated pistils. Cluster 4 (equivalent to group 5 for the hierarchical clustering) is not an interesting group as the genes are constitutively expressed in the self-pollinated pistils and are down and up regulated in the cross-pollinated pistils. Genes from clusters 9 and 10 (group 8) fulfil putatively more general biological process in the pistil. The clusters found probably correspond to groups of co-regulated genes, as demonstrated in previous experiments (Breyne et al. 2002 and De Paepe et al. 2004).

Our results clearly illustrate that cDNA-AFLP constitutes a powerful technique to screen for genes if no prior sequence information is available. Due to the stringent conditions used, the rate of false positives produced by mis-priming is probably low. However, false-positives are generated by co-amplifying unrelated sequences that may either underlie the targeted band or co-migrate very closely with the chosen fragment and by differences created in the original RNA populations by nonstandardized extraction procedures. Therefore, a detailed verification is required to confirm the expression patterns using Real-time RT-PCR, Northern Blotting or cDNA microarray hybridisations. In our case Northern Blotting was not an option given the big quantities of plant material which are needed. Therefore, in a next step we checked the expression patterns observed in this experiment for a number of the selected genes in a microarray experiment (see chapter 7).

Although differential and constant TDFs can be discriminated by visual scoring, the use of software for image analysis allowed us to correct for lane to lane differences and to normalize the data. In this way, it was possible to

use an objective methodology for the selection of differential transcripts. Furthermore, the quantitative data generated could be used for cluster analysis and for the identification of groups of genes with similar expression profiles. A similar approach was applied recently for transcriptome analysis during cell division in plants (Breyne et al. 2001, 2002 and 2003) and to get insights into the early response to ethylene in *Arabidopsis* (De Paepe et al. 2004). Quantitative analysis of the expression cDNA-AFLP profiles is a very good alternative to microarrays and recent report demonstrate a good correlation between cDNA-AFLP and microarray results in *Saccharomyces cerevisiae* and *Arabidopsis* (Reijans et al. (2003) and De Paepe et al. (2004) respectively).

With regards the data analysis, the Adaptive Quality-Based Clustering (AQBC) method from De Smet et al. (2002) resulted in the assignment of 36% of the genes to 11 clusters.

Chapter 5

Sequencing of cDNA-AFLP sequence tags

5.1 Introduction

As shown in Chapter 4, cDNA-AFLP is an efficient method for the isolation of differentially expressed genes between self- and cross-pollinated pistils with no prior sequence information of genes involved in the SI-response. In our experiment, a total of 479 differentially expressed transcript derived fragments (TDFs) were identified. Using statistical software, the TDFs were grouped into clusters of genes displaying similar expression profiles. However, before any further step can be undertaken, it was necessary to gather sequence information of the selected TDFs. This will allow us to classify them according to their molecular functions. Among the annotated TDFs displaying homology to well-characterized gene classes, a connection to SI can then eventually be made. This strategy has also been used in several functional genomics studies to get insight into different biological processes, such as defence responses in *Arabidopsis* (Schenk et al. 2000), the identification of genes regulated by water-deficit stress in *Arabidopsis thaliana* (Bray 2004), to get insights into the early response to ethylene in *Arabidopsis* (De Paepe et al. 2004) and to identify genes differentially expressed during the hypersensitive response of cassava (Kemp et al. 2005).

In addition, sequence information can, in a latter step, be used to develop specific tests for confirmation of the differential expression patterns obtained for different tissues. Sequence information can also be used to develop polymorphic DNA-markers suitable for linkage analysis in segregating populations and to analyse the genomic organisation of the factors involved in SI (this approach is illustrated in Chapter 6).

5.2 Objectives

To get insight into the putative functions of genes identified in Chapter 4, 180 cDNA-AFLP fragments were selected for further analysis and characterisation. These cDNA-AFLP bands were excised from the gels, reamplified and sequenced. The obtained sequence information was compared to known sequences of publicly available databases of the 'National Centre of Biotechnology Information' (NCBI). This allowed us to classify the identified TDFs into functional categories.

5.3 Materials and methods

5.3.1 Reamplification and sequencing of cDNA-AFLP fragments

Bands corresponding to differentially expressed transcripts were excised from the gels and eluted in 200 µl of MQ water. A small volume (5 µl) of this elution was used for PCR amplification using the same conditions as for the pre-amplification step of the cDNA-AFLP procedure, but with the primers used in the corresponding selective amplification step (see materials and methods in Chapter 4). After control of the reamplification products on 1.5% agarose gels, PCR products were used for cloning.

The reamplified fragments were cloned in the TOPO-vector with the TOPO TA cloning kit® (Invitrogen), according to the manufacturer's instructions. For each cDNA-AFLP fragment at least 8 colonies were picked and dissolved in MQ water. In a next step direct colony PCR reactions were carried out with the universal M13 forward and reverse primers. For each cDNA-AFLP fragment, the PCR products obtained from 5 colonies were sequenced with the M13 forward or reverse primer using an automatic sequencer (Biosystems Prism™ 377). Fragments longer than 500 bp were sequenced in both directions. Cloning, sequencing reactions and sequencing runs were carried out as described in materials and methods of Chapter 3.

5.3.2 Homology searching and gene annotation

Primer sequences were removed from the DNA sequence files and reverse and forward sequences were assembled using the Autoassembler™ software (Applied Biosystems). DNA sequence information obtained for each transcript was compared with the nucleotide and protein sequences of the publicly available databases of NCBI using the BLAST algorithm (Altschul et al. 1997). This work was done in collaboration with Cindy Martens from the Department of Plant Systems Biology (UGhent and VIB). Two databases, a nucleotide sequences database containing 366.014 different sequences (Table 5.1) and a locally held and updated protein database, containing all the protein sequences of NCBI were used. The nucleotide sequences database contained genes from cereals and cDNAs from rice and *Arabidopsis*. In a first step, nucleotide sequences of the TDFs were compared to the protein database using the BLASTx algorithm. If no hits were found, the sequences were compared to the nucleotide sequences database, using the BLASTn algorithm. When a good BLASTn hit was

found, the hit was then compared to the protein database using the BLASTx algorithm.

The sequences were then classified into functional categories based on the MIPS classification system (Munich Information Centre for Protein Sequences, http://mips.gsf.de/projects/plants).

Table 5.1 Overview of the sequences used to construct a nucleotide sequences database

Species	Reference
cDNAs	
Oryza sativa	http://www.tigr.org/tdb/e2k1/osa1/data_download.shtml
Arabidopsis thaliana	ftp://ftp.tigr.org/pub/data/a_thaliana/ath1/SEQUENCES/ATH1.cdna
Cereal Gene Indices	
Zea mays	ftp://ftp.tigr.org/pub/data/tgi/Zea_mays/ZMGI.release_15.zip
Triticum aestivum	ftp://ftp.tigr.org/pub/data/tgi/Triticum_aestivum/TAGI.release_10.zip
Hordeum vulgare	ftp://ftp.tigr.org/pub/data/tgi/Hordeum_vulgare/HVGI.release_9.zip
Secale cereale	ftp://ftp.tigr.org/pub/data/tgi/Secale_cereale/RYEGI.release_3.zip
Sorghum bicolor	ftp://ftp.tigr.org/pub/data/tgi/Sorghum_bicolor/SBGI.release_8.zip

5.4 Results and discussion

5.4.1 Sequencing of differentially expressed transcripts

The results presented in Chapter 4 were used to select a subset of 180 differentially expressed cDNA-AFLP fragments. The selected TDFs were distributed over the different clusters identified by the AQBC method, with a higher concentration of those TDFs coming from the most relevant clusters concerning the self-incompatibility (Chapter 4). As only 39% of the TDFs identified in Chapter 4 were assigned to a cluster by the AQBC method, some of the TDFs which remained unassigned in the AQBC analysis but which displayed interesting expression profiles concerning SI were also selected for further analysis. These TDFs were excised from the gels, reamplified, cloned and sequenced. For 6.7% (12/180) of the selected TDFs, no good reamplification was obtained. Therefore, a set of 168 TDFs was used for cloning and sequencing.

Of all the cloned TDFs, 57 corresponded to two to four different sequences, while for 111 TDFs all clones rendered identical sequences. In conclusion, the patterns of differential expression observed on cDNA-AFLP

gels corresponded in most cases to single transcripts (66% or 111/168 TDFs). These results are in the line with what was found in the study of Cooper et al. (2001), namely 53% of the recovered sequences were unique. In the study of Mao et al. (2004), on the other hand, 97% of the selected differentially expressed fragments displayed unique genes. However, comparisons with literature data are difficult as most authors do not provide any indication of the number of unique genes found in similar cDNA-AFLP studies.

When different sequences were found for different clones of a given TDF, all were included in further analyses, as it was not easy for us to determine which one represented the 'original' differentially expressed gene. Sequences derived from the same cDNA-AFLP fragment where designated by the name of the cDNA-AFLP band followed a capital letter, which indicates the identity of the clone. For example CDB202_280A, and CDB202_280B correspond to sequences obtained for different clones of the same cDNA-AFLP band (a band of 280 bp amplified with primer pair 202).

Finally, a total of 259 sequences were compared with those present in the protein database of NCBI running the program 'BLASTx'. For 18% of the 259 sequences no significant hit against known proteins was found. The sequences which displayed homology to proteins with known function or to hypothetical proteins were classified into functional categories, using the MIPS classification system. However, 28% of the 259 tags remained unclassified as they showed homology to hypothetical proteins or to proteins for which no gene ontology information was available (Figure 5.1). The annotated genes were assigned to 11 functional categories (Figure 5.1).

5.4.2 Overview of the different functional groups

In the first place, we checked whether transcripts which were clustered together according to their expression patterns (Chapter 4) were assigned to the same functional group. No clear correlation between functional groups and clusters was found and clusters containing genes with similar expression profiles, consist of genes from different functional groups. In the next section the different functional groups are discussed.

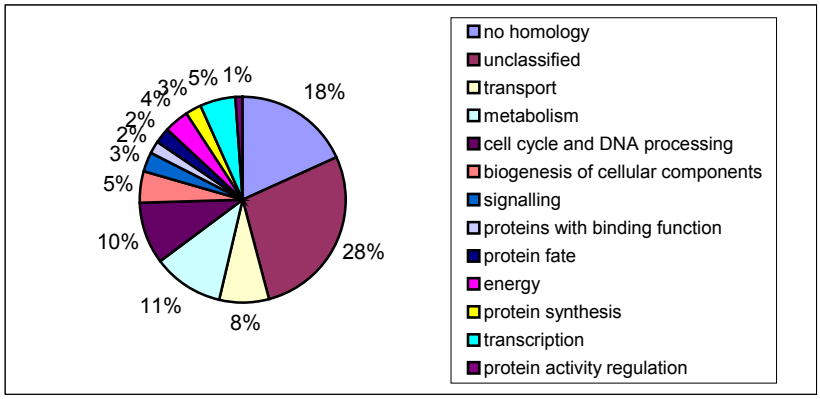


Figure 5.1 Functional classification of the sequenced tags based on the MIPS classification system.

5.4.2.1 Transport

Eight percent of the tags showed homology to proteins involved in the active transport of molecules (metabolites, proteins, etc) and ions across cell membranes or inside the cell (Table 5.2). These transport mechanisms are involved in numerous biological functions in living organisms. Active transport is a process whereby the cell uses both transport proteins and metabolic energy to transport substances across the membrane against a concentration gradient. Transport proteins involved in active transport include antiporters, symporters and the proteins of the ABC (ATP Binding Cassette) transport system (http://www.cat.cc.md.us/courses/bio141/lecguide/unit1/prostruct/cm.html#atp). The ABC transport system is an active transport system, widespread in Archae, Eubacteria and Eukaryotes (Higgins 1992). In all cases, it shows a common global organization, with three types of molecular components. Typically, it consists of two integral membrane proteins (permeases) composed of six transmembrane segments, two peripheral membrane proteins that bind and hydrolyse ATP and a periplasmatic (or lipoprotein) substrate-binding protein. This transport system is involved in the transport of certain sugars and amino acids.

Transport is an important function for pollen tube growth which consists of the incorporation of new cell membrane and cell wall material in a small area at the tip. This requires tip-directed targeted secretion, endocytosis and rapid organelle movement through the cytoplasm (Hepler et al. 2001). Which

proteins are involved in this transport is still unknown, but some of the genes identified in this study constitute interesting candidates.

Table 5.2 Overview of the cDNA-AFLP fragments showing homology to transport proteins.

Tag name	AQBC cluster	Most significant homology	Acession no.	E value
CDB005_350		OSJNBa0076N16.21 protein, *Oryza*	Q7XUE4	3E-41
CDB011_255B		P-type ATPase (Fragment), *Hordeum*	Q94IN2	5E-31
CDB013_285		Amino acid transporter-like, *Oryza*	Q6K4N7	1E-16
CDB021_290Y		Putative signal recognition particle receptor, *Oryza*	Q6Z245	1E-07
CDB031_290D		Putative ABC transport system ATP-binding protein	Q6NFI8	1
CDB049_280		Organic cation transporter (solute carrier family 22 member 2, isoform a), human	O15244	12
CDB051_370C		Organic cation transporter (Solute carrier family 22 member 2, isoform a), human	O15244	6.3
CDB069_380D	2	ABC transporter, ATP-binding protein, MsbA family, bacteria	Q5HR53	2.1
CDB098_220A		ATPase 2	Q7RNG4	4.3
CDB099_210G	7	Putative transmembrane transport protein, bacteria	Q9RL01	9.9
CDB100_340A		Oligopeptide ABC transporter permease, bacteria	Q5WDD5	0.26
CDB100_340B		Oligopeptide ABC transporter permease, bacteria	Q5WDD5	0.26
CDB100_340D		Oligopeptide ABC transporter permease, bacteria	Q5WDD6	0.26
CDB105_350		Putative ATFP7, *Oryza*	Q69NJ9	1E-24
CDB187_305		Putative cellular retinaldehyde-binding/triple function, *Oryza*	Q5TKJ2	1E-05
CDB197_255D		Triose phosphate translocator, *Triticum*	Q9ATY2	4E-34
CDB202_270A		Proton/sodium-glutamate symport protein	Q5V627	2.4

Also the actin cytoskeleton and transport of Ca^{2+} are important in pollen tube growth (see Chapter 2). It has been demonstrated in *Papaver* that loss of the Ca^{2+} gradient in the pollen tube tip plays a role in the initial inhibition of pollen tube growth (Franklin-Tong et al. 1993 and 1995). CDB011_255B and CDB098_220A display homology to P-type ATPases, which constitute a

superfamily of cation transport enzymes. P-type ATPases can be divided into 4 major groups: (1) Ca^{2+}-transporting ATPases, (2) Na^{+}/K^{+}-transporting ATPases, (3) plasma membrane H^{+}-transporting ATPases and (4) the bacterial P-type ATPases (http://www.ebi.ac.uk/interpro/IEntry?ac=IPR001757).

It is also clear that for most cDNA-AFLP fragments, the sequences of different clones display homology to genes with an identical function, e.g. CDB100_340A, B and D display all homology to an ABC transporter permease.

5.4.2.2 Metabolism

Eleven percent of the tags displayed homology to genes coding for enzymes involved in the basic metabolism of the cell (Table 5.3). Some cDNA-AFLP fragments represented genes encoding enzymes involved in the synthesis and breakdown of carbohydrates, lipids, amino acids and other household molecules like pyrimidines, purines etc. Tag CDB003_370 displayed homology to the *Zea mays* endoxylanase gene. In maize the endoxylanase protein is synthesized in the tapetum and exclusively present in the pollen coat. Xylanases are involved in the hydrolysis of the cell wall, which is important in certain aspects of sexual reproduction, such as the initial penetration of the pollen tube into the stigma (Yih Bih et al. 1999).

Some other tags showed homology to invertase, galactinol synthase and allene oxide cyclase. Allene oxide cyclase is a key component of the jasmonic acid biosynthesis. Jasmonates are important intracellular regulators mediating diverse developmental processes, such as seed germination, flower development, leaf abscission and senescence (Creelman and Mullet 1995). In addition, jasmonate induces plant defence responses against a group of pathogens and mechanical wounding (Creelman et al. 1992). Invertase activity can also be activated by wounding (Shanker et al. 1995). To a certain extent, pollen tube growth through the style is similar to the invasion caused by pathogens and could activate similar metabolic routes. Alternatively, expression of these genes could be caused by the mechanical wounding of the pistils during manipulation in this experiment.

We also found homology to galactinol synthase, which catalyses the production of rafinose oligosaccharides (RFOs), a group of soluble gactosyl-

sucrose carbohydrates playing a role in plant development. RFO accumulation has also been associated with environmental stress conditions, including cold, heat and dehydratation (Zhao et al. 2003).

Homology has also been found to glycosyltransferases, which are enzymes that catalyze the transfer of glycosyl (sugar) residues to an acceptor, both during degradation (cosubstrate is water or inorganic phosphate) and during biosynthesis of polysaccharides, glycoproteins and glycolipids.

Table 5.3 Overview of the tags showing homology to proteins involved in metabolism.

Tag name	AQBC cluster	Most significant homology	Acession no.	E value
Carbohydrate mechanism				
CDB003_150A		2-oxoglutarate dehydrogenase, E1 subunit (EC 1.2.4.2), *Arabidopsis*	Q9ZRQ2	5E-13
CDB003_370		Specific endoxylanase genes, Maize	Q9ZTB8	
CDB010_275		Putative chitinase, bacteria	Q9L115	1.1
CDB030_400		Cell wall invertase (EC 3.2.1.26) (Fragment), *Saccharum*	Q7XZS4	9E-28
CDB039_350		Glucose-6-phosphate dehydrogenase (Fragment), *Triticum*	Q9XJ53	1E-55
CDB052_360		Starch branching enzyme lib, *Hordeum*	Q9ZTB6	
CDB059_410A		TDP-desosamine glycosyltransferase, bacteria	Q9F826	0.14
CDB059_410D		TDP-desosamine glycosyltransferase, bacteria	Q9F826	0.23
CDB069_185		Galactinol synthase 1 (EC 2.4.1.123), Maize	Q5DVS7	1E-10
CDB070_250		Beta-1,6-glucanase precursor, *Trichoderma*	Q8TG97	9E-16
CDB071_185		Galactinol synthase 1 (EC 2.4.1.123), Maize	Q5DVS7	1E-10
CDB085_190		Glucosyltransferase-like protein, *Arabidopsis*	Q9SCP6	10
CDB187_410	1	TDP-desosamine glycosyltransferase, bacteria	Q9F826	0.18
CDB202_270D		Putative enolase; 31277-33713 (Putative enolase), *Arabidopsis*	Q9C9C4	1E-17
CDB232_500H	3	Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD- dependent formate dehydrogenase) (FDH), *Hordeum*	Q9ZRI8	2E-54
CDB246_150		Malate synthase, glyoxysomal (EC 2.3.3.9), Maize	P49081	2E-07

Table 5.3 *Continued*

Tag name	AQBC cluster	Most significant homology	Acession no.	E value
Amino acid, glycoprotein, lipid and phospholipid metabolism				
CDB018_205		Myotubularin related protein, putative, Arabidopsis	Q9LPP2	1E-11
CDB058_285B		Putative phosphatidate cytidylyltransferase domain-containing protein, Arabidopsis	Q6Z2M3	2E-34
CDB059_410A		TDP-desosamine glycosyltransferase, bacteria	Q9F826	0.14
CDB059_410D		TDP-desosamine glycosyltransferase, bacteria	Q9F826	0.23
CDB085_190		Glucosyltransferase-like protein, Arabidopsis	Q9SCP6	10
CDB096_280		Long-chain-fatty-acid CoA ligase, bacteria	O51162	5
CDB187_410F	1	TDP-desosamine glycosyltransferase, bacteria	Q9F826	0.18
CDB210_440		Serine carboxypeptidase iii precursor, Hordeum	P21529	
CDB219_240A		Methionine synthase protein, Arabidopsis	Q8W0Q7	2E-31
Nucleotide metabolism				
CDB048_285		Putative uracil phosphoribosyltransferase, Arabidopsis	Q9M336	2E-06
CDB216_135B		Hypothetical protein CBG23391, nematode	Q60LX5	18
CDB202_280D		Putative dihydroorotate dehydrogenase (Fragment), Oryza	Q8S3J6	1E-06
CDB214_290E	9	Putative dihydroorotate dehydrogenase, Oryza	Q8S3J6	
Others				
CDB117_300		Allene oxide cyclase, Oryza	Q8L6H4	3E-21

5.4.2.3 Cell cycle and DNA processing

Ten percent of the TDFs showed homology to genes that encode proteins involved in DNA replication (Gag-pol proteins, replication proteins and a endonuclease/exonuclease/phosphatase family protein) and in DNA recombination (putative polyproteins and BARE-1 polyprotein) (Table 5.4).

Table 5.4 Overview of the cDNA-AFLP fragments showing homology to proteins involved in the cell cycle and DNA processing.

Tag name	AQBC cluster	Most significant homology	Acession no.	E value
CDB008_240D		Nuclear protein-like, *Arabidopsis*	Q9FMF9	2E-07
CDB017_800D	1	Gag-pol polyprotein (Fragment)	Q5D7H0	5E-83
CDB017_800G	1	Hypothetical protein OSJNBa0053F13.6 (Putative polyprotein), *Oryza*	Q6ATA5	e-111
CDB017_800O	1	Hypothetical protein OSJNBa0053F13.6 (Putative polyprotein), *Oryza*	Q6ATA5	e-113
CDB017_800Q	1	Hypothetical protein OSJNBa0053F13.6 (Putative polyprotein), *Oryza*	Q6ATA5	e-109
CDB017_800R	1	Hypothetical protein OSJNBa0053F13.6 (Putative polyprotein), *Oryza*	Q6ATA5	e-109
CDB021_225P	2	Gag-pol polyprotein, Maize	Q8L7J3	3E-23
CDB022_230O		Replication protein A 70b, *Oryza*	Q7Y069	2E-08
CDB028_230		RNA helicase, putative	Q9U0H4	9.6
CDB028_345I		Putative helicase, *Oryza*	Q5W673	2E-08
CDB031_370	1	Putative gag-pol, Maize	Q7XBD9	1E-20
CDB053_285	9	putative nuclease, *Hordeum*	Q9ARD4	3E-32
CDB071_490A	1	Putative gag-pol protein, *Oryza*	Q7XEF2	4E-68
CDB071_490E		Putative gag-pol protein, *Oryza*	Q7XEF2	9E-69
CDB079_300		Putative polyprotein, *Oryza*	Q688H9	5E-10
CDB085_800		Prohibitin, Maize	Q9M586	5E-92
CDB093_280	1	Replication protein, bacteria	Q9F3Q9	0.62
CDB093_410	3	Putative phosphoesterase, *Oryza*	Q84JU8	
CDB101_440	3	Putative CRK1 protein, *Oryza*	Q6YU34	2E-14
CDB129_330	1	Gag-pol, Maize	Q8W1D1	2E-40
CDB139_230E		Putative gag-pol, Maize	Q7XBD9	6E-04
CDB216_135C		Putative endonuclease/exonuclease/phosphatase family protein, *Oryza*	Q75LU1	0.003
CDB246_410A	4	BARE-1 polyprotein (Fragment), *Hordeum*	Q9S9A8	2E-23
CDB246_410C	4	BARE-1 polyprotein (Fragment), *Hordeum*	Q9S9A8	4E-23
CDB246_410F	4	BARE-1 polyprotein (Fragment), *Hordeum*	Q9S9A8	7E-24

CDB085_800 showed homology to a prohibitin of maize. Prohibitins constitute a group of highly conserved proteins in eukaryotic cells playing a role in important cellular processes, such as cellular signaling,

transcriptional control, senescence, aging, cell death (Fusaro et al. 2003) and the regulation of mitochondrial activities (Tatsuta et al. 2005). However it remains largely elusive how prohibitins function at the molecular level and whether a similar of action underlies these apparently divergent roles (Tatsuta et al. 2005). This represents an interesting finding if we take into account that Programmed Cell Death (PCD) may be triggered by SI in incompatible pollen tubes of *Papaver* (Jordan et al. 2000). However further studies to explore the evidence for a PCD signaling cascade are required.

CDB101_440 displayed homology to a CRK1 (Cytokinin-Regulated kinase) of rice. CRK1 proteins are members of the Ca^{2+}-dependent protein kinases and were primarily found in tobacco, regulated by very low concentrations of cytokinin (Schäfer et al. 2002). It contains a highly conserved intracellular Serine/Threonin kinase domain and Schäfer et al. (2002) hypothised that CRK1 is involved in an early step of hormone signaling. Till now, nothing was known about the putative involvement of types of kinases in the SI-response.

5.4.2.4 Biogenesis of cellular components

Four homologues to basic components of the cystoskeleton, namely actin, and one homologue to an actin binding protein were identified (Table 5.5). This finding is not unexpected, when we take into account that actin is essential for apical growth of pollen tubes and SI stimulates a dramatic and fast rearrangement of the actin in *Papaver.* This rearrangement is mediated by actin-binding proteins (Snowmann et al. 2000, 2002). Which type of actin-binding proteins are involved in SI remains still unclear. Large actin depolymerization in incompatible pollen is probably linked to events that trigger apoptosis, as found in some animal cell types. (see Chapter 2 for further details). Although up to now the involvement of actin depolymerization has only been described in *Papaver*, these data suggest that it might be a more general phenomenon, also found in other SI-systems.

Homology to proteins involved in cell wall binding and catabolism was also found. Germins and Germin-Like Proteins (GLPs) are a family of developmentally regulated glycoproteins expressed during germination which may play a role in altering the properties of cell walls during germinative growth. It has been shown that wheat and barley germins act as oxalate oxidases, generating H_2O_2 from the oxidative breakdown of oxalate thereby playing a significant role in plant development and defense.

GLPs exhibit sequence and structural similarity with cereal germins but mostly lack oxalate oxidase activity. Germins and GLPs are thought to play a significant role during zygotic and somatic embryogenesis (wheat and barley, respectively), salt stress (barley), pathogen elicitation (wheat and barley), heavy metal stress etc. (Patnaik and Khurana 2001).

Table 5.5 Overview of the tags showing homology to proteins involved in the biogenesis of cellular components.

Tag name	AQBC cluster	Most significant homology	Acession no.	E value
CDB015_200	1	Actin, *Triticum Aestivum*	Q947D0	
CDB022_230I		Actin 2, *Oryza*	P17298	0.06
CDB022_230L		Actin 2, *Oryza*	P17298	0.06
CDB023_245		Putative pectin methylesterase, Oryza	Q7F0H7	0.004
CDB036_315		Germin B, *Hordeum*	Q9FYY4	0.092
CDB040_380		Actin, *Triticum*	Q947D0	
CDB071_370		Putative proline-rich cell wall protein, *Oryza*	Q8LMV5	0.01
CDB084_370	11	Putative Erwinia induced protein 1, *Oryza*	Q69T51	8E-51
CDB101_330B	9	Major sperm protein, nematode	Q18471	1
CDB101_330G	9	Drought-induced hydrophobic protein (Low temperature-induced low molecular weight integral membrane protein LTI6b), *Oryza*	Q6Q7C9	0.005
CDB117_305D		Putative alpha subunit of F-actin capping protein, *Oryza*	Q8H358	2E-09
CDB228_240		Germin D, *Hordeum*	Q9FYY2	0.1

5.4.2.5 Signalling

Considering the different processes triggered by the signalling cascade in the plant, like SI, senescence, programmed cell death, biotic stresses, dormancy and germination, it is not surprising that 3% of the TDFs displayed homology to genes involved in signalling such as kinases, receptors and GTP-binding proteins (Table 5.6). Different signal transduction mechanisms have been described in plants, such as calcium-based signalling, G-protein-mediated signalling and signalling involving inositol phospholipids (Clark et al. 2001). In the calcium-based signalling mechanism calcium (acting as a diffusible second messenger to the initial stimuli) and protein kinases play important roles. External signals such as hormones, light stress, pathogenesis lead to transient increases in calcium concentrations within the

cell. These increases lead to calcium binding by regulatory proteins, which turn the signal into a biological response. The G protein-mediated signalling involves GTP-binding proteins in coupling with many plasma membrane receptors with intracellular signalling pathways. These signalling routes are important in several physiological processes, including the regulation of pollen germination and pollen tube growth (Clark et al. 2001). The tags listed in Table 5.6 represent thus putative candidates of kinases and binding proteins involved in the signalling cascade which follows a SI-reaction in *L. perenne*.

Table 5.6 Overview of the tags showing homology to proteins involved in signalling.

Tag name	AQBC cluster	Most significant homology	Acession no.	E value
CDB016_320A		Putative GTP-binding protein, *Oryza*	Q6K2P6	6E-17
CDB016_320B		Putative GTP-binding protein, *Oryza*	Q6K2P6	6E-17
CDB022_230N		Olfactory receptor 13 (Fragment), Metazoa	Q68JG8	28
CDB062_270		S-locus-like receptor protein kinase (Fragment), *Prunus*	Q6DU53	9.1
CDB214_290A		Protein kinase, putative	Q6BGL9	3.2
CDB216_250		Protein kinase, putative	Q6BGL9	0.49
CDB222_140E		G protein-coupled receptor PGR3 (Fragment), human	Q86SP2	70

5.4.2.6 Proteins with binding activity

Three TDFs showed homology to RNA-binding proteins which are regulatory proteins of major importance for the control of cellular development (Table 5.7). Two TDFs showed homology to a calcium-binding protein, namely caleosin (CDB044_330A) and a calmodulin-binding protein (CDB141_160), respectively (Table 5.7**)**. Caleosins contain calcium-binding domains and have an oleosin-like association with lipid bodies. Caleosins are present at relatively low levels and are mainly bound to microsomal membrane fractions at the early stages of seed development. As the seed mature, overall levels of caleosins increase dramatically and are associated almost exclusively with storage lipid bodies. The calcium binding domain is related to the calcium EF-hand motif (http://www.ebi.ac.uk/interpro/IEntry?ac=IPR007736). Lipid bodies are found in seeds, pollen, and vegetative organs of plants to serve as energy stores during a forthcoming period of active metabolism. Caleosins may be

involved in processes such as membrane and lipid-body fusion (Murphy et al. 2000) and can consequently be involved in the pollen tube growth.

Calmodulin is a calcium-binding protein that can bind and regulate a multitude of different protein targets thereby affecting many different cellular functions such as metabolism, apoptosis, intracellular movement and signal transduction (Clark et al. 2001).

Table 5.7 Overview of the tags showing homology to proteins with binding activity.

Tag name	AQBC cluster	Most significant homology	Acession no.	E value
CDB015_230		Putative RNA-binding protein, *Oryza*	Q8H642	2E-24
CDB044_330A		Calcium-binding protein 1, caleosin, *Hordeum*	Q94B09	5E-21
CDB044_330B		Glycine-rich RNA-binding protein RGP-1c (Fragment), *Nicotiana*	Q6RY61	7E-13
CDB141_160		Putative calmodulin-binding protein, *Oryza*	Q6YU05	3E-14
CDB219_240B		U2 snRNP auxiliary factor, small subunit, *Oryza*	Q6AUG0	4.3

5.4.2.7 Protein fate, synthesis and activity regulation

An interesting aspect is the homology that has been found for five TDFs to ubiquitin-like proteins (Table 5.8). Ubiquitin-mediated proteolysis has emerged as being fundamentally important in many aspects of development such as hormone signalling, light perception and circadian rhythm as well as in plant defence signalling (Xu et al. 2002). Protein degradation mediated by ubiquitin-dependent proteolysis is an important aspect in the destruction of cell cycle regulators. Recently it is also been shown that ubiquitination processes, resulting in protein degradation, are probably involved in the SI-response in the Solanaceae (Sijacic et al. 2004) (see Chapter 2). Also in *Brassica*, it has been demonstrated that ARC_1, a stigma protein, possess E3-ubiquitin ligase activity (Franklin-Tong and Franklin 2003). Recently, Hackauf and Wehling (2005) identified also a STS marker in rye which co-segregated with Z and which showed homology to ubiquitin-specific protease genes. The involvement of this gene into the SI-response has to be confirmed. However, three TDFs (CDB022_310B, CDB022_310C and CDB211_300) showed a down regulated expression profile in the self-pollinated pistils and were not clearly expressed during cross-pollination. Their maximum expression level was achieved 1h after self-pollination.

Therefore, these fragments are good candidates to be involved in an ubiquitin-dependent proteolysis during the SI-response (Figure 5.2). The other two TDFs (CDB043_200D and CDB043_200F) are up regulated in both self- and cross-pollinated pistils, with an increasing expression up to 24h after pollination, and could be involved in degradation processes of the pistil at the end of the pollination (Figure 5.2). Therefore, the TDFs CDB022_310 and CDB211_300 represent promising targets for further investigation with respect to their relation to the SI-system of the grasses.

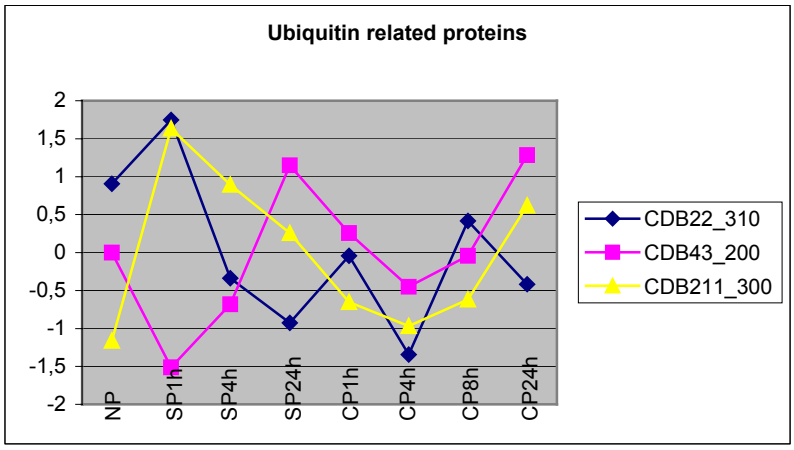


Figure 5. 2 Expression profiles of the TDFs, displaying homology to ubiquitin-related proteins which are probably involved in the protein degradation (see chapter 4).

It is has been demonstrated that the SI-response in the gametophytic SI-system of *Papaver* triggers programmed cell death (PCD) in the pollen, with the PCD response being dependent on a caspase-like activity (Thomas and Franklin-Tong 2004). In *Brassica*, the sporophytic SI-mechanism involves the ubiquitination of proteins via ARC1, an E3 ubiquitin ligase (Stone et al. 2003). In *Anthirrinum*, which together with members of the Solanaceae and Rosaceae represents a third genetic mechanism of SI involving S-RNases, evidence has accumulated that it is in compatible rather than in incompatible pollinations that ubiquitinated S-RNases are recruited for destruction via the 26S proteosome system. The degradation of the non-self S-RNases is mediated by the binding to a SCF (composed of Skp1, Cullin1, Rbx1 and an F-box protein) complex and an additional component that would protect the

self-S-RNases from ubiquitination in self-pollinations and, consequently, keeps it ready for inhibiting pollen tube growth (Qiao et al. 2004).

Table 5.8 Overview of the tagss showing homology to proteins involved in degradation, synthesis and activity regulation.

Tag name	AQBC cluster	Most significant homology	Acession no.	E value
Protein fate				
CDB022_310B	1	Ubiquitin protein ligase Praja1 (EC 6.3.2.-), mousse	O55176	0.024
CDB022_310C	1	Hypothetical protein AT4g15840, *Arabidopsis*	O23435	0.27
CDB043_200D		Ubiquitin-like protein (At5g57860), *Arabidopsis*	Q8L8S0	0.003
CDB043_200F		Hypothetical protein B2J23.190, *Neurospora*	Q9HFI3	3.4
CDB099_210H	7	Putative SKD1 protein, *Oryza*	Q6ETH5	7.6
CDB211_300	6	F6I1.6 protein, *Arabidopsis*	Q9FZ52	8.8
Protein synthesis				
CDB010_350A		Putative plastid ribosomal protein L34, *Oryza*	Q5VQX3	2E-15
CDB010_350C		Putative plastid ribosomal protein L34, *Oryza*	Q5VQX3	2E-15
CDB021_225C	2	Eukaryotic peptide chain release factor subunit 1-3 (ERF1-3), *Oryza*	Q75I90	7E-24
CDB021_290A		IDI2, *Hordeum*	Q9AYT7	
CDB028_275A		Putative hUPF2, *Oryza*	Q6H470	2E-14
CDB081_320	2	Putative translation initiation factor, *Oryza*	Q5ZBT6	9E-21
CDB099_210A	7	Putative 60S ribosomal L28 protein, *Oryza*	Q5TKP3	1E-08
Protein activity regulation				
CDB024_230E		Heat shock protein 101, *Arabidopsis*	P42730	
CDB038_450		Iron receptor-like protein, bacteria	Q6UVM2	0.034
CDB044_330D		Heat shock protein 101, *Arabidopsis*	P42730	

Finally, a very small group of differentially expressed fragments (only 1%) showed homology to regulatory proteins such as Heat Shock Proteins (HSPs). HSPs are synthesised in response to high temperature stress, but are also expressed at optimal growth temperatures and seem to play essential roles in normal growth, acting as molecular chaperones to assist the folding, assembly and transport of other proteins (Parsell and Lindquist 1993).

5.4.2.8 Energy

Seven TDFs showed homology to proteins involved in electron transport (Table 5.9). The electron transport is a coupled series of oxidation/reduction reactions during which ATP is generated by energy transfer as electrons move from high reducing state to lower reducing state. This is a general process in active living tissues.

Table 5.9 Overview of the genes showing homology to proteins involved in energy processes.

Tag name	AQBC cluster	Most significant homology	Acession no.	E value
CDB005_310B		Photosystem II reaction center H protein (Photosystem II 10 kDa phosphoprotein) (PSII-H), *Triticum*	P69555	1E-31
CDB016_225H		NADH dehydrogenase, *Geraea*	Q32270	13
CDB027_300		OSJNBb0004A17.1 protein, oxidase, *Oryza*	Q7XPI8	1E-39
CDB039_500		Putative cytochrome P450, *Oryza*	Q84JR8	1E-10
CDB062_265	2	NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3), vertebrates	Q6L7I8	21
CDB152_260F	6	Putative mandelonitrile lyase, *Oryza*	Q6Z290	5E-23
CDB154_460		Cytochrome P450, *Triticum*	Q6T485	2E-55

5.4.2.9 Transcription

Five percent of the TDFs fell in this functional group. Most of them were derived from genes encoding transcription factors of different well-known families in plants, such as MADS, Zn-finger, AP3 and auxin-response (Table 5.10). Transcription factors regulate the transcription, can be selectively activated or desactivated by other proteins and may play wider roles in different aspects of plant development (http://en.wikipedia.org/wiki/Transcription_factor). Which transcription factors are involved in the SI-response is still not known. The floral homeotic protein, apetala3 is a putative transcription factor involved in the genetic control of flower development. It is required for normal flower development of petals and stamens (Jack et al. 1992). Some others TDFs show homology to transcriptional regulators active in bacteria (Table 5.10).

Table 5.10 Overview of the genes showing homology to proteins involved in transcription.

Tag name	AQBC cluster	Most significant homology	Acession no.	E value
CDB022_310A	1	MADS box transcription factor AP3 (Fragment)	Q9LL99	0.35
CDB028_275B		Putative small nuclear ribonucleoprotein polypeptide G, *Oryza*	Q75K32	2E-28
CDB035_230E		Regulatory protein, bacteria	Q6STM0	3.3
CDB077_230A		Transcriptional regulator, MerR family, bacteria	Q87TU9	4.3
CDB101_600B	11	Putative ATP-dependent RNA helicase DB10, *Oryza*	Q5JKF3	2E-31
CDB104_390E		Putative zinc finger protein, *Oryza*	Q6L4C8	2E-13
CDB168_260A		Putative auxin-responsive factor (ARF1), *Oryza*	Q6YVY0	4E-17
CDB186_350A		MADS-box protein 7, *Hordeum*	Q9LEI1	1E-38
CDB202_270B		Putative splicing factor, *Oryza*	Q5ZCD9	3E-20
CDB202_280B		Putative splicing factor, *Oryza*	Q5ZCD9	4E-22
CDB214_275C		Floral homeotic protein apetala3, *Arabidopsis*	P35632	
CDB224_410B	4	OSJNBa0020I02.5 protein, *Oryza*	Q7XS75	2E-08
CDB230_195D	4	Transcriptional activator (b1), Maize		
CDB232_390A		Putative MADS box protein ZMM17, *Oryza*	Q6H711	2E-11

5.5 Conclusions

Till now very little is known about the identity and role of the genes involved in the SI-response in members of the grass family. A set of TDFs representing genes putatively involved in SI, and spread over the different groups defined in Chapter 4, were selected for cloning, sequencing and further characterisation. Of all the cloned TDFs, only 57 cDNA-AFLP bands correspond to more than one sequence.

Finally, a dataset of 259 sequences was obtained to compare with the nucleotide and protein database of NCBI. For 82% of the sequences a significant hit against genes with known function was found. The TDFs showing homology to annotated (putative) genes were classified into functional categories, using the MIPS classification system. In most cases, the different clones of the same fragment were more or less identical in sequence. Twenty-eight percent of the 259 tags remained unclassified as

they showed homology to hypothetical proteins or to proteins for which no gene ontology information was available. Finally TDFs were assigned to eleven functional groups. In many cases homologies were found to components of general cellular functions. In other cases, interesting homologies to proteins known to be involved in fertilization and SI were found. The results obtained allow us to make a selection of TDFs representing *L. perenne* genes which are putatively involved in SI and which should get priority in future research:

- CDB011_255B and CDB098_220A display homology to P-type ATPases, a superfamily of cation transport enzymes, including Ca^{2+}-transporting proteins. It is well known that transport of calcium is important for pollen tube growth and in signalling cascades. These signalling cascades are important in different biological processes, including the SI-response.

- CDB015_200, CDB022_230I, CDB022_230L, CDB040_380 and CDB117_305D constitute also a relevant group, as they display homology to genes coding for cellular components such as actin and actin binding proteins. Actin plays a very important role in pollen tube growth and SI triggers a dramatic and fast rearrangement of the actin in *Papaver*, leading to the inhibition of the pollen tube growth.

- CBD003_370 probably represents the *L. perenne* homologue of the endoxylanase protein previously isolated in maize pollen which is specifically expressed in tapetum tissue. Also relevant are TDFs CDB016_320A and CDB016_320B, which display homology to genes encoding GTP–binding proteins.

- CDB117_300 probably represents an allene oxide cyclase gene. As allene oxide cyclases are important in jasmonic acid synthesis, this result suggests that some of the processes activated in plants during fertilization and SI could be common to those triggered by pathogen attack. Although rather speculative at this stage, this hypothesis is very interesting if we take into account that the growth of the pollen tube through the transmitting tract of the style is a process with clear similarities to the growth of fungal pathogens in plants.

▪ CDB022_310, CDB043_200D and CDB211_300 show homology to ubiquitin-related proteins, involved in protein degradation. As discussed in Chapter 2 recent data demonstrate that the SI-response in several gametophytic and sporophytic systems involves ubiquitin-related protein degradation. Furthermore, a STS derived from an ubiquitin-specific protease gene seems to co-segregate with the Z locus in rye (Hackauf and Wehling 2005). Given the 'universal' importance of ubiquitin in SI, these TDFs represent very interesting candidates for further research in *L. perenne.*

Chapter 6

Development and mapping of EST markers derived from SI-related genes in two L. perenne populations

6.1 Introduction

Different DNA-marker systems are available for linkage map construction in plants. These include techniques based on restriction cleavage and hybridisation with labelled probes such as Restriction Fragment Length polymorphism (RFLP), techniques based on PCR such as Simple Sequence Repeats (SSR) and techniques which combine both approaches, such as Amplified Fragment Length Polymorphism (AFLP) and Cleaved Amplified Polymorphic Sequences (CAPS). Great potential for fine genetic mapping is offered by Single-Nucleotide Polymorphism (SNP) markers. SNPs are the result of single-base substitutions, the most abundant form of DNA polymorphisms in most organisms. They offer the opportunity to uncover allelic variation directly within expressed sequences of candidate genes (Snowdon and Friedt 2004). However, SNPs cannot be resolved by conventional gel electrophoresis and a variety of novel techniques and technologies have been developed or adapted for SNP detection (Henikoff and Comai 2003, Grattapaglia 2004).

6.1.1 Marker techniques available in L. perenne

In the case of *L. perenne*, RAPD (Huff et al. 1997), AFLP (Roldán-Ruiz et al. 2000, Roldán-Ruiz et al. 2001), DNA-derived SSR markers (Kubik et al. 1999 and 2001, Jones et al. 2001, Jensen et al. 2005), EST-derived SSR markers (Faville et al. 2004, Warnke et al. 2004, Jensen et al. 2005), RFLPs (Jones et al. 2002a, Inoue et al. 2004), some of which correspond to cDNA sequences from other *Poaceae* species and STS (Site Tagged Sequences) markers (Lem and Lallemand 2003) have been used to date, for different purposes. SNP techniques have not been applied in *Lolium* yet and appropriate tools and methodologies still need to be developed for efficient SNP detection and verification in this highly polymorphic species. SNP markers are currently being developed and validated in *L. perenne* in the context of the EU-project GRASP (Lübberstedt et al. 2003), in which DvP participates as partner. SNP markers are being identified in genes tightly associated with different traits related to quality, sustainability, disease resistance, fertility and flowering.

The DNA marker techniques applied to the segregating populations used in the mapping work described in this chapter are shortly introduced in the following paragraphs.

The most widely used co-dominant marker technique, not only in *L. perenne* but also in many other crop species has been RFLP. In addition to the ability to detect homologous DNA sequences in genomes, RFLP probes also detect heterologous sequences in the genomes of related species. As a consequence, RFLPs are very useful for comparative mapping and provide an opportunity to address biological questions about the evolution of plant genomes (Yu et al. 1996, Jones et al. 2002a, Sim et al 2005). However, RFLPs are not practical for routine mapping. The biggest disadvantage of the RFLP technique is the high amount of DNA required (about 50-200 µg of DNA per digest per individual, depending on the total size of the genome). Furthermore, the RFLP protocol, involving southern blotting, is time-consuming, labor intensive and not suitable for automation.

With the advent of the polymerase chain reaction (PCR; Mullis and Faloona 1987), the possibility was open to greatly increase the marker density in existing genetic maps. One of the most broadly used PCR-based DNA marker technique is AFLP. The AFLP protocol is able to generate a high number of markers starting from small amounts of DNA. AFLP can detect single nucleotide changes and amplifies in *Lolium* spp. around 95 DNA fragments in a single experiment using 6 selective nucleotides, in the absence of prior sequence knowledge (Roldán-Ruiz et al. 2000). In mapping studies, AFLP is commonly used to generate in an efficient way a dense backbone on which other markers (mostly co-dominant) are placed. The major disadvantage of the technique is that AFLP markers are dominant which means that the heterozygote cannot be distinguished from one of the homozygotes. This has implications for linkage mapping, as the markers are not fully informative which regarding the genetic composition of the plants. However, available linkage mapping software packages such as Joinmap 3.0 are able to cope with this aspect (Jansen and van Ooijen 2003, Maliepaard et al. 1997, see further).

SSR markers are highly reproducible, genetically co-dominant and multiallelic. They are therefore suitable for framework mapping as they can be used as anchor markers in alignment of maps obtained in different mapping populations. An important advantage of SSR markers is that they can be screened on automatic DNA sequencers, allowing high throughput genotyping and map construction. The main disadvantage of SSR markers is that the primers designed to amplify a SSR locus are often species-specific and cannot be easily used for comparative genetics. This is due to the high variability in the sequences surrounding the SSR sites (Peakall et

al. 1998). Recently, a set of 20 primer pairs have been published which allow the amplification of SSR markers located in gene sequences in *L. perenne* (EST-SSR; Faville et al. 2004). In recent years *L. perenne* SSR markers have also become publicly available (for an overview see Jensen et al. 2005). The main advantage of EST-SSR markers is that they allow the detection of polymorphisms in functional sequences. Although SSRs derived from cDNAs may show lower levels of intraspecific polymorphism than those derived from non-coding genomic regions, they provide the potential advantage of close linkage to significant gene variants (Faville et al. 2004). An important characteristic of EST-SSR markers is that they are, in general terms, more conserved across species than “anonymous” SSR markers. This is due to the fact that the primer binding sites are designed in coding DNA which has a higher degree of sequence conservation than non-coding DNA.

Another possibility to screen for polymorphisms in expressed sequences is the use of EST markers. EST markers are a special type of STS markers, which screen for polymorphisms in gene sequences. These “functional” markers have advantages in comparison to “anonymous” markers. In the first place, and similarly to what has been explained above for EST-SSRs, they are more conserved across species because the primer binding sites are designed in coding DNA. Therefore, such EST markers are very useful for aligning linkage maps and comparing QTL positions across mapping populations (Taylor et al. 2001). Genetic mapping with ESTs enables thus a more rapid transfer of linkage information between related species (Cato et al. 2001). Furthermore, as functional markers are derived from functional motifs they are in complete linkage with the target locus allele (Andersen and Lübberstedt, 2003).

6.1.2 Mapping populations and linkage map construction

DNA markers have been used in the development of detailed linkage maps of several crop species. A genetic linkage map graphically represents the arrangement of numerous loci, including morphological, isozyme as well as DNA markers along the chromosome. The distance between these loci is expressed in centimorgans (cM) which represent the recombination rates between loci (1 cM = 1% recombination). There is no linear relationship between the recombination distance and the physical distance (expressed in base pairs) because the rate of recombination varies along the length of the

chromosome. For example, regions near the centromeres show suppressed recombination which is reflected by the clustering of markers (Kumar 1999).

There are four well-defined steps in the construction of a genetic linkage map: 1) construction of a segregating population, 2) development of markers and identification of polymorphic ones, 3) scoring of the polymorphic markers in the parents and the progeny, 4) statistical analysis to determine the relationship among the DNA markers and their order along the chromosome. The order of the markers along the chromosomes is derived from the recombination frequencies and their standard errors for all pair-wise comparisons between loci. Distance units on the map are calculated using mapping functions. The construction of linkage maps is facilitated by the use of computer programs such as Linkage, Mapmaker or Joinmap (Kumar, 1999, Jansen and Van Ooijen, 2003).

In plant species genetic maps are usually constructed using dedicated segregating populations. Important factors that have to be considered in the design of a segregating population for linkage mapping purposes are the choice of the parents and the most appropriate type of population. The choice of parents to use depends on the amount of DNA polymorphisms in the species, since in absence of DNA polymorphism, segregation analysis and mapping are impossible. The parents should be genetically divergent enough to exhibit sufficient polymorphisms and at the same time should not be so far apart as to cause sterility in the progeny. Several kinds of mapping populations can be constructed, with each population structure having unique strengths and weaknesses. Most genetic mapping populations in plants have been derived from the F_1 of a cross between largely homozygous parents (inbred lines or doubled haploids). One possibility is to cross the F_1 hybrid (heterozygote plant) with one of the parents (homozygote plant), to produce a backcross (BC) population. Another possibility is to produce a F_2 population by selfing the F_1 hybrid. Doubled haploids can also be made by regenerating plants from single pollen grains of the F_1 hybrid and inducing chromosome doubling. These different types of populations vary in the time and effort they take to construct and the amount of information provided by each individual. In a backcross population, the alleles just from one parent segregate, in contrast to an F_2 population which segregates for the alleles of both homozygous parents. As a consequence, an F_2 individual harbors two times more information than a backcross individual. Because of this and also because F_2 populations harbor all possible combinations of parental alleles (AA, Aa, aa), the F_2 population has

been the most widely used in genetic linkage studies in plants (Paterson 1996).

The construction of F_2 and BC populations in outbreeding crop species such as ryegrasses is not straightforward. In many cases inbreeding by self-pollination to produce homozygous lines, if at all possible, would lead to reduced plant vigour. In this case a population derived from a cross between two heterozygous parents can be used for linkage analysis (this kind of population is usually referred to as 'cross-pollinator' population or 'CP'). Appropriate approaches for the estimation of recombination frequencies in CP populations, the so-called two-way pseudo-testcross approach, have been developed (Grattapaglia and Sederoff 1994, Maliepaard et al. 1998). In this approach, one linkage map is first produced for each parent. In a second step, the two parental linkage maps are integrated into a consensus map using codominant markers as genetic bridges. The major difference between a cross derived from inbred lines (BC or F_2) and a cross between two heterozygous parents is that in crosses from the inbred lines the number of alleles per locus is always two and in crosses with two outbred parents the number of alleles per polymorphic locus may be two, three or four (Jansen and Van Ooijen 2003).

6.1.3 Published linkage maps in L. perenne

Several linkage maps for *Lolium* spp. have been published or are being developed at different institutes. A non-exhaustive overview of the published *L. perenne* maps is given in Table 6.1. For example Jones et al. (2002b) published a reference linkage map of perennial ryegrass, based on a segregating population (p150/112), derived from a cross between a double haploid and a heterozygous *L. perenne* plant and is well known as the ILGI map (International *Lolium* Genome Initiative). The ILGI map contains 109 heterologous anchor RFLP probes from wheat, barley, oat and rice, allowing comparative relationships to be investigated between perennial ryegrass and other Poaceae. Interestingly, the S- and Z-loci have been mapped in this population by Thorogood et al. (2002).

Table 6.1 Published genetic maps of *Lolium* spp. The structure of the segregating population, the marker types and map length are given (DH=double haploid).

Reference	Plant species	Population type	Number of plants	Marker types	Map length (cM)
Hayward et al. (1998)	Interspecific	One-way pseudo-testcross DH (*L. perenne*) x F_1 hybrid (of *L. perenne* x *L. multiflorum*)	89	RFLP, RAPD, isozyme	692
Bert et al. (1999) Jones et al. 2002 (a&b)	***Lolium perenne***	One-way pseudo-testcross DH x F_1 (ILGI-population)	95	AFLP, EST, SSR, RFLP, isozyme	930
Armstead et al. (2002)	***Lolium perenne***	F_2 population (Selfing of F_1 hybrid derived from a cross between two distinct inbred lines)	180	RFLP, AFLP, STS, SSR	515
Alm et al. (2003)	*Festuca pratensis*	Two-way pseudo-testcross	138	RFLP, AFLP, SSR, isozyme	659
Muylle et al. (2005)	***Lolium perenne***	Two-way pseudo-testcross (CLO-DvP population)	252	RFLP, AFLP, SSR, STS	833
Faville et al. (2004)	***Lolium perenne***	Two-way pseudo-testcross	157	RFLP, EST, SSR	757
Maiko et al. (2004)	***Lolium multiflorum***	Two-way pseudo-testcross	82	RFLP, AFLP, SSR	1244
Warnke et al. (2004)	Interspecific	Two annual ryegrass plants were crossed with two perennial ryegrass plants. From the resultant F1 population, two plants were chosen and crossed to develop a three-generation population	91	AFLP, RAPD, RFLP, SSR, isozyme	537 712
Jensen et al. (2005)	*Lolium perenne*	Consensus map of four different populations		AFLP, RFLP, STS, SSR	772

Armstead et al. (2002) published a linkage map of perennial ryegrass based on an F_2 mapping population. This mapping population was produced for the genetic analysis of water soluble carbohydrate accumulation. Muylle et al. (2005) published a genetic linkage map of *L. perenne* generated on the basis of the two-way pseudo-testcross procedure. This population was generated for the genetic analysis of crown rust resistance.

Almost all published linkage maps have been aligned with each other on the basis of public available RFLP, SSR and STS markers (Jones et al. 2002, Lem and Lallemand 2003, Jensen et al. 2005).

6.1.4 Genomic regions known to be involved in SI in the Poaceae

S and *Z* have been mapped in the ILGI population by Thorogood et al. (2002). For this purpose, 139 individuals were phenotyped using in vitro pollination tests. The incompatibility genotype of the female DH parent was designated $S_{11}Z_{11}$. As the male parent was unrelated, and knowing that S and Z are highly polyallelic, it was assumed that it was heterozygous at both loci and that it carried other alleles than the female parent. Therefore, the male parent was assigned genotype $S_{23}Z_{23}$. According to the two-locus model of SI, the F_1 family derived from this cross segregates into a maximum of four possible incompatibility genotypes ($S_{12}Z_{12}$, $S_{12}Z_{13}$, $S_{13}Z_{12}$ and $S_{13}Z_{13}$). Inter-pollinations of these F_1 genotypes results in three types of incompatibility relations: fully incompatible, half compatible and 75% compatible (See Table 2.1, Chapter 2). No fully compatible reactions can be obtained as all F_1 plants produce 25% S_1Z_1 gametes that never effect successful pollinations. In this way it was possible to assign an incompatibility genotype to 139 F_1 plants of the mapping family. This information was then used to locate the S-locus on LG1 and Z-locus on LG2 of the *L. perenne* map (Thorogood et al. 2002). Thorogood et al. (2002) identified also a strong 'interaction' between LG1 and LG3, associated with significantly distorted segregation ratios for the majority of the markers on LG3. They concluded that the association of the S-locus alleles with markers on LG3 indicates that there are genes on LG3 that interact with the S-locus (or a gene tightly linked to it).

Significant and consistent distorted marker segregation ratios have also been used to identify markers associated with SC mutations of incompatibility loci. Such distortions have been identified in *P. coerulescens* and rye. In *P. coerulescens* a third locus (T) was identified (Hayman and

Richter 1992) and in rye a self-compatibility locus (S5) was detected on LG5 (Fuong et al, 1993 and Voylokov et al. 1993). Also Thorogood et al. (2004) identified in a *L. perenne* mapping population large regions of marker segregation distortion on LG5 and LG7 (see also Armstead et al. 2004). The LG5 region of marker distortion found in *L. perenne* is particularly attractive as a location for an SC-locus, as in rye the S5-locus is also located in LG5.

In *P. coerulescens* high-resolution maps of the S- and Z-regions were generated from populations with distorted segregation for incompatibility related genes and applying the method developed by Leach (1988) to calculate the recombination frequency between a marker and S (or Z) using partially compatible populations. They used RFLP probes from wheat, barley, oat and *Phalaris*, which mapped on chromosomes 1 and 2 of the Triticeae. The S-locus is located in the centromeric region of chromosome 1 and Z might be located on the long arm of *Phalaris* chromosome 2. The marker Xbm2 (Bian 2001) cosegregated with S in *Phalaris* and the corresponding rice BAC clone is useful for identifying more markers linked to S and candidate S-genes (Bian et al. 2004).

Recently, Hackauf and Wehling (2005) identified 12 STS-markers corresponding to gene orthologues from rice genes and flanking Z in rye. The closest flanking markers mapped at distances of 0.5 cM (marker TC89057) and 1 cM (marker TC101821) from Z. Marker TC116908 cosegregated with Z in a testcross.

6.2 Objectives

Two loci, S and Z, control the recognition process between pistil and pollen in *L. perenne*. Some of the genes involved in the signalling cascade resulting in a SI reaction will be located at either S or Z, but other genes may be dispersed all over the genome or clustered in certain genomic regions. The function or genomic organisation of other factors involved in the SI-response in the Poaceae is still unknown. In the Brassicaceae and Solanaceae it is well known that S-loci are located in chromosome positions where recombination is suppressed and which contain several genes interspersed with repetitive sequences and non-functional ORFs. These regions not contain only the genes which code the components which determine SI-specificity, but also other genes involved in SI and fertilization (Fukai et al. 2003, Wang et al. 2004). In the previous chapters genes putatively involved in SI were identified using cDNA-AFLP and a homology-

based strategy. Knowing the genome location of these other factors can help in the future to exploit knowledge in model organisms such as rice to elucidate the molecular mechanism of SI in members of the grass family. This chapter describes the development and mapping of polymorphic DNA-markers derived from (i) the cDNA-AFLP fragments displaying differential expression and discussed in Chapters 4 and 5, (ii) a thioredoxin gene, presented in Chapter 3 (iii) a kinase gene (Chapter 3) and (iv) PCR products amplified in *L. perenne* using six of the primer pairs of Hackauf and Wehling (2005).

The mapping work was carried out in two mapping populations. The main objective of mapping of polymorphisms associated with expressed sequence tags representing genes putatively involved in SI was to get insight into the genomic localisation of the factors involved in the SI-response in ryegrass.

6.3 Material and methods

6.3.1 Plant material

Two perennial ryegrass mapping populations were screened for polymorphisms and used for genetic mapping of the candidate genes:

1) The *L. perenne* mapping family P150/112 (ILGI) was used in collaboration with the Institute of Grassland and Environmental Research, Wales (IGER). This family is the core mapping family for the International *Lolium* Genome Initiative (ILGI). The population was derived from a cross between an anther culture-derived double haploid plant (female) and an unrelated heterozygous plant donated by Dr Peter Wilkins' forage breeding programme (Jones *et al.* 2002a). This population has been extensively genotyped and phenotyped (Jones et al. 2002a, Bert et al. 1999, Cogan et al. 2005, Thorogood et al. 2002). The core linkage map consists of 240 loci (RFLP, STS, AFLPs, SSRs and isozymes) and spans 811 cM. S and Z have been mapped in this population by Thorogood et al. (2002). We used the heterozygous parent of the population (the double haploid parent is no longer available) and 86 progeny plants from this population for linkage mapping.

2) The CLO-DvP reference population TC1 x SB2 was derived from a cross between two heterozygous *L. perenne* plants. A dense genetic map, generated from 252 individuals and that spans 833 cM was available by the start of this study. The map consisted of 230

loci (RFLPs, AFLPs, SSRs and STSs) and was constructed using the two-way pseudo-testcross approach (Muylle et al. 2005). We used the two parent plants and 100 progeny plants for linkage mapping in this study.

6.3.2 DNA-extraction and PCR conditions

Genomic DNA was extracted using the Qiagen DNA midi purification kit. PCR reactions were performed in an 20 µl reaction volume, containing 30 ng of genomic DNA, 6.5 µl dNTPs (20 mM), 1.2 µl $MgCl_2$ (25 mM), 2 µl 10x PCR buffer, 0.25 µl AmpliTaq, 2 pmol forward and 2 pmol reverse primer. The PCR was carried out in an AB9700 cycler according to the following profile: 1) 94°C for 2 min, 2) 35 cycles of 94°C for 30 sec, 56°C for 30 sec and 72°C for 1 min, 3) 72°C for 10 min. The PCR products were separated on 20 cm 8% polyacrylamide gels and stained with ethidium bromide. Gels were scored visually.

6.3.3 Marker development

Primer pairs were developed to amplify each of the 259 sequences described in Chapter 5, the thioredoxin gene and the SRK gene described in Chapter 3, using the software Primer Express v. 2.0 (Applied Biosystems). Also six of the primer pairs published by Hackauf and Wehling (2005), and which amplify sequences flanking Z in rye were used for mapping (Table 6.2).

Table 6.2 Overview of the 6 STS markers published by Hackauf and Wehling (2005) and used in this study.

STS marker	**Tentative annotation**	**Map position in rye**
TC116908	Ubiquitin-specific protease 22	Cosegregation with Z
TC89057	Glycerol kinase	0.5 cM from TC116908
TC101821	TPR domain, putative	1 cM from TC116908
TC32601	Beta-ketoacyl-acyl carrier protein synthase	3.5 cM from TC116908
TC108778	**Similar to oryzain alpha chain precursor**	3.5 cM from TC116908
TC87436	Plasma membrane P-type proton pump ATPase	5.4 cM from TC116908

All primer pairs were first checked for amplification efficiency and polymorphism using a reduced number of plants of each mapping population. For the ILGI population, the heterozygote parent and four progeny plants (one for each SI-genotype) were used. From the CLO-DvP population the two parents and five randomly chosen progeny plants were used. Primer pairs which were able to reveal clear fragment length polymorphisms were used to fingerprint the whole population. Some primer pairs revealed length polymorphisms only in the ILGI population, some others only in the CLO-DvP population and others in both.

6.3.4 Marker scoring

In a first step all segregating DNA-bands were scored as dominant markers, i.e. for presence or absence of the detected band. For the CLO-DvP population both parents were also scored and the DNA-fragments were divided into three groups based on their segregation type: 1) a- x --: informative for the mother (father is homozygous); 2) -- x a-: informative for the father (mother is homozygous) and 3) a- x a-: heterozygous in both parents. A first linkage analysis using Joinmap 3.0 (van Oojien and Voorrips 2001) was based on this information. After this first linkage analysis, the markers generated with the same primer pair, and mapping in the same linkage group were scored co-dominantly.

For the ILGI population, a BC1 type population, the expected segregation type of all the polymorphic markers is the same, namely: ab x aa and the genotype of an offspring is either “aa” or “ab”.

6.3.5 Linkage analysis and map construction

The newly developed EST markers were mapped in each mapping population, in combination with the marker information (RFLPs, AFLPs, SSRs and STSs) which was already available (described by Jones et al. (2002a) and Thorogood et al. (2002) for the ILGI population and by Muylle et al. (2005) for the CLO-DvP population). Marker segregation ratios were first checked by the χ^2 test available in the Joinmap 3.0 software. The linkage groups were numbered according to the order proposed by Jones et al. (2002a), and which follows the numbering accepted for the Triticeae.

ILGI population: the polymorphic markers developed in this research were added to the publicly available mapping dataset of the ILGI population

(http://ukcrop.net/perl/ace/search/FoggD). The SSR markers mapped in this population by Jensen et al. (2005) were also included in the linkage analysis. Map construction was carried out using Joinmap 3.0 (Van Ooijen and Voorrips 2001). Determination of the linkage groups (LG) was performed at LOD ratio thresholds of 5.0 with the exception of LG1 and LG3 where LOD ratio thresholds of 12 were necessary to separate the two linkage groups from each other.

CLO-DvP population: As we used the two-way pseudo-testcross procedure, the total dataset was divided into two separate datasets. One contained the markers polymorphic in the first parent (abxaa; abxcc; abxcd; abxab; abxac) and the other contained the markers polymorphic in the second parent (aaxab; aaxbc; abxab; abxac; abxcd). For each dataset a linkage map was constructed. The markers were grouped into linkage groups at a LOD threshold of 6 for the TC1 parental map and at a LOD threshold of 11 for the SB2 parental map using JoinMap 3.0 (van Oojien and Voorrips 2001).

Markers polymorphic in both parents and integrated in both parental linkage maps were used to identify and align the homologous linkage groups of the two parental maps.

Marker order was calculated at LOD=1.00 and recombination threshold value (REC) of 0.40. Map distances in centimorgans were calculated using the Kosambi function. In a second step, distorted markers and markers unmapped during the first round, were included in the map only if the map order was not drastically affected. In some cases a LOD=0.01 and REC=0.499 had to be used to integrate the markers (Muylle et al. 2005).

6.4 Results and discussion

6.4.1 DNA-marker analysis

For 29 out of 259 TDFs described in Chapter 5, no primer pairs could be developed, as the sequenced fragments were too short. For the remaining 230 TDFs, primer pairs were designed and used to screen the two mapping populations for length polymorphism. One primer pair amplified a part of the thioredoxin gene, one amplified a SRK gene homologue and five amplified genomic regions located in the neighbourhood of Z in rye. No amplification products were obtained with the primer pair TC87436 in any of the populations.

Of all these primer pairs 21% and 17%, were unable to amplify any DNA fragment in the ILGI and CLO-DvP populations respectively. Sixty-seven

percent (159) and 69% (165) of the primer pairs amplified monomorphic bands in the ILGI and the CLO-DvP populations respectively. The rest of the primer pairs (29 for ILGI and 32 for CLO-DvP) detected clear polymorphic bands and were used for amplification in the complete ILGI and/or CLO-DvP population. An example of a primer combination displaying polymorphisms is shown in Figure 6.1. Although the rate of successful amplification was rather high in both populations, most of the primers tested revealed PCR products which were monomorphic on all the plants of the mapping populations. Probably, sequence polymorphisms, resulting in restriction site polymorphisms are present in the amplified sequences and therefore can be used for the development of CAPS markers. This is however, rather time-consuming as several restriction enzymes have to be tested for each of the PCR products in order to identify the informative ones and was out of the scope of this PhD. Alternatively, SNP markers can be developed, but this requires sequencing of PCR products obtained from the parents of the mapping populations, which is also time-consuming and costly. This will be the subject of future research.

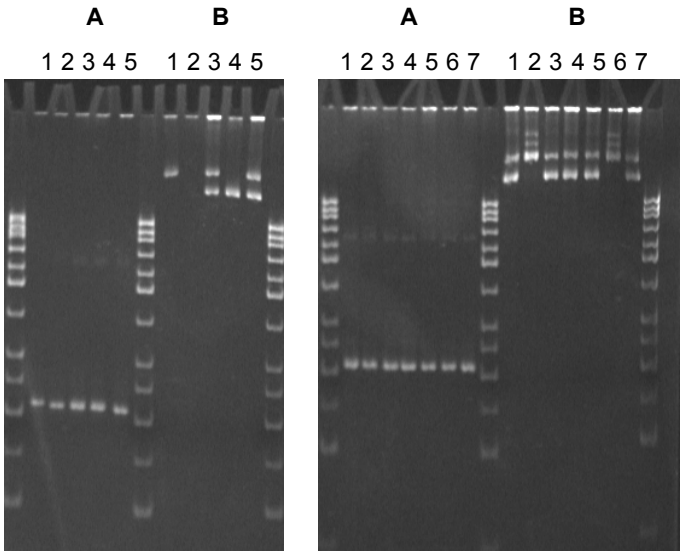


Figure 6.1 Example of two primer combinations on the test set from the ILGI and CLO-DvP population. For the ILGI population: (1): heterozygous parent; (2-5): progeny plants. For the CLO-DvP population: (1-2): parents; (3-7): progeny plants. The primer combination CDB178_280C (A) displayed no polymorphisms on neither of the populations. The primer combination CDB186_350 (B) displayed clear polymorphisms on both populations and was tested on the whole mapping population.

Of the 46 primer pairs which revealed clear polymorphisms (Table 6.3), 15 amplified polymorphic DNA-bands in both populations and 31 in only one of the mapping populations. Of the 29 primer pairs which revealed polymorphism in the ILGI population, 22 produced single locus amplification products while 6 primer pairs amplified two loci and one three loci (indicated by letters: a-c, example is given in Figure 6.2). Of the 32 primer pairs used in the CLO-DvP population, 20 amplified single locus products and 12 amplified more than one locus (eight primer pairs detect two loci and four primer pairs detect three loci) (Figure 6.2). Finally, a dataset of 37 segregating markers for the ILGI population and 48 for the CLO-DvP population was obtained for mapping.

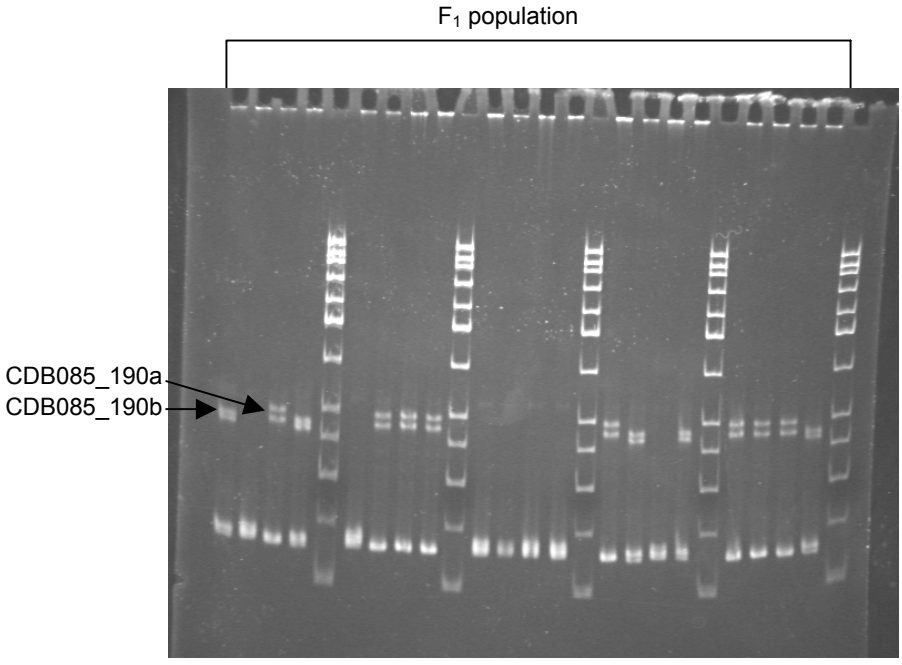


Figure 6.2 Example of primer combination CDB085_190, amplifying two loci in the CLO-DvP population, each scored separately. The 50 bp ladder was used as length marker.

Table 6.3 An overview of the primer pairs showing polymorphisms in one or both populations. P=Polymorphic/M=Monomorphic

EST	ILGI	CLO-DvP	EST	ILGI	CLO-DvP
CDB100_340B	P	M	CDB036_285	P	M
CDB109_170	M	P	CDB036_410A	M	P
CDB011_255A	P	M	CDB037_320	P	M
CDB114_370	P	P	CDB039_235B	M	P
CDB117_300	P	P	CDB039_350	M	P
CDB126_190B	M	P	CDB039_500	M	P
CDB013_285	P	M	CDB040_380	M	P
CDB140_300	P	P	CDB004_170	P	P
CDB014_180	M	P	CDB043_200F	M	P
CDB154_460	P	P	CDB046_165	M	P
CDB186_350	P	P	CDB051_370D	P	M
CDB187_305	P	P	CDB069_380D	M	P
CDB188_180	M	P	CDB008_240A	M	P
CDB202_280F	P	M	CDB084_370	P	M
CDB210_440	P	M	CDB085_190	M	P
CDB214_275	M	P	SRK	P	M
CDB022_230M	M	P	Thioredoxin	P	P
CDB230_195B	P	M	TC32601	P	P
CDB232_390	P	M	TC101821	P	P
CDB232_500A	P	P	TC116908	P	P
CDB246_175	M	P	TC89057	P	P
CDB248_270	P	P	TC108778	P	M
CDB026_440	P	P	**Total polymorphic**	**29**	**32**
CDB028_275C	P	M			

6.4.2 Segregation analysis

The law of segregation, which is the most fundamental law in Mendelian genetics, relies on the predictable transmission of alleles from a parent to its offspring, and on a predictable formation of genotypes from the transmitted alleles. Segregation distortion is the deviation of segregation ratios from the Mendelian expectation and has been reported in a wide range of plant species (Jenczewski et al. 1997), including *Lolium* spp. (Hayward et al.

1998, Bert et al. 1999, Jones et al. 2002 a and b, Armstead et al. 2002, Thorogood et al. 2002, Muylle et al. 2005).

Marker information available for these populations, including the markers generated in the context of this study and markers produced in previous studies and publicly available (AFLP, RFLP, SSR and STS) was combined for the construction of the respective maps. For each population the χ^2 test integrated in the Joinmap 3.0 was performed to test for Mendelian segregation ratios of the DNA-markers generated. Fifteen percent and 36% of the markers showed distorted segregation ($P<0.05$) for the ILGI population and the CLO-DvP population, respectively. In Jones et al. (2002a) 32% of the markers displayed distorted segregation ratios at $P<0.05$ (excluding the AFLPs). Armstead et al. (2002) reported 36% distorted markers at $P<0.05$ and Faville et al. (2004) mentioned 26% and 15% segregation distortion. These values are in the same range as the percentages obtained in the present study. Segregation distortions can be due to errors in genotyping and scoring and to biological reasons. Biological reasons for segregation distortion can be found in pollen tube growth, fertilization processes, distorting factors such as SI alleles and selective elimination of zygotes (Lu et al. 2002). Only the distorted AFLP markers were eliminated for the linkage analysis and map construction.

6.4.3 Linkage analysis and map construction

6.4.3.1. CLO-DvP population

The CLO-DvP map is based on the two-way pseudo-testcross approach. It contains RFLPs, STSs, AFLPs and SSRs, grouped into seven linkage groups (Muylle et al. 2005). A total of 545 loci were included in the present analysis and for each parental map seven major linkage groups were detected. The TC1 map contained 212 markers and the SB2 map 192 markers (results based on 'Map 3' version of Joinmap, in which most of the markers are forced to join a linkage group). In a second step the parental maps were integrated, using markers polymorphic in both parents.

The combined map contains 121 AFLP loci, 36 RFLP loci, 97 SSR loci, 46 STS markers and 36 EST markers. The map, generated from 252 individuals, spans 720 cM and consists of 336 loci (Figure 6.3). The mean distance between two consecutive loci is 2.1 cM. The length of the LGs varies from 88 cM till 133 cM.

The different marker types used to construct the map are distributed over the different linkage groups. A total of 121 AFLP markers were incorporated in the map. Although the AFLP markers are not very informative in a two-way pseudo-testcross, they fulfil an important role in expanding the linkage groups. They also fill in gaps between co-dominant markers (e.g. LG4). Most of the developed EST markers mapped to LG2, LG3 and LG4. One EST fragment (CDB187_305) detected multiple loci, which mapped to different linkage groups. This can be due to a duplication of the locus that is also recognised by the primers and amplified or to the amplification of a PCR product different in. Six EST fragments could not be scored co-dominantly. For example, for primer pair CDB188_180 two alleles were visualized. The marker seems to segregate as ab x ac, with 'b' a null allele. This class could not be entered in Joinmap 3.0 as the genotype aa could not be distinguished from genotype ab. Therefore, the two alleles were scored and mapped as separate markers. They map 4 cM apart and although they could represent the same locus, the possibility that they correspond to different loci cannot be ruled out.

Several regions of the maps contained markers with distorted segregation at $P<0.001$. Segregation distortion in *Lolium* has been reported by several authors: Bert et al. (1999) found a group of distorted markers in LG3, near the isozyme marker GOT/3, previously designated as linked to the SI-locus Z. Jones et al. (2002a) found distortion mainly in LG3 but also in LG4 and LG5. Thorogood et al. (2002) reported segregation distortion in LG3 and Faville et al. (2004) in LG2, LG3, LG4 and LG5. In this study numerous markers with segregation distortion were found in LG1 (8 markers), LG2 (8 markers) and LG7 (8 markers). Also in LG3 and LG4, five and six distorted markers were found, respectively. The high number of distorted markers in LG1 and LG2 can be due to the presence of the S- and Z-loci, respectively. Indeed, in LG2 most of the distorted markers were concentrated around Z, localized to the long arm end of chromosome 2 (Figure 6.3). In LG1 the distorted markers are especially concentrated around S (localized in the centromeric region of LG1). The consistency of segregation distortion on LG3 across different maps strongly suggests the presence in this genomic region of gene(s) affecting the viability after fertilization and may correspond to a genetic factor that interacts with the self-incompatibility S-locus on LG1 (Thorogood et al. 2002). Similarly to what has been found in the present study, segregation distortion on LG7 was also found by Armstead et al. (2004). This can be due to the presence of a QTL influencing heading date and seed set (Armstead et al. 2004).

6.4.3.2 ILGI population

For the ILGI population, the map positions of the S- and Z-loci are known (Thorogood et al. 2002). A total of 315 markers were included in the analysis. LG3 was strongly associated with LG1 and only using a LOD score > 12 was it possible to split them. The linkage between these two groups of markers was associated with significantly distorted segregation ratios for the majority of the markers on LG3 (Jones et al. 2002a). Thorogood et al. (2002) found that the S-locus on LG1 is associated with some markers on LG3, implicating that there are gene(s) on LG3 that interact with the S-locus (or a gene tightly linked to it).

It was not possible to obtain one linkage map for LG1 and LG7. They both split into two groups. In both cases, the smallest groups contain the newly added data (the EST markers developed in this research and SSR markers mapped by Jensen et al. (2005)). The reason why these markers group together and could not be integrated in the other part of the linkage group is difficult to find out, but one possibility is that some progeny plants were mislabelled between the different mapping experiments. This would result in each independent genotyping experiment in the identification of markers with strong linkages, but which show no linkage to previously mapped markers for the same linkage group.

The final map was constructed using less stringent conditions (REC=0.499 and LOD=0.01). It includes 106 AFLPs, 117 RFLPs, 29 SSRs, 3 isozyme markers and 37 EST markers (25 of them developed in this study). The linkage map shown in Figure 6.3 spans 852 cM and consists of 292 loci. The length of the linkage groups varies from 60 (not taking into account the smaller two sub groups) to 198 cM. These results are in the same range as found by Jones et al. (2002b) for the same mapping population, a genetic map containing 240 loci covering 811 cM on seven linkage groups.

Nine previously mapped markers showed distorted segregation at $P<0.001$ and they are all located on LG3, as was also mentioned by Thorogood et al. (2002). This can be due to the association of these markers with the S-locus on LG1 (see above).

6.4.3.3 Comparison with results of similar mapping experiments

Mapping of TDFs identified using cDNA-AFLP was previously achieved in cassava (Suárez et al. 2000) and rice (Zheng et al. 2004). In cassava (*Manihot esculenta* Crantz), Suárez et al. (2000) applied the cDNA-AFLP technique to roots and leaves of the parents of a genetic mapping population to identify genes controlling starch quantity and quality, disease and pest resistance. They identified more than 700 TDFs between the two parents. A subset of 50 TDFs was cloned and sequenced and finally six of them were mapped. In rice, Zheng et al. (2004) combined cDNA-AFLP and Bulked Segregant Analysis (BSA) to identify genes that control rice blast resistance in a doubled-haploid population derived from a cross between a resistant and a susceptible variety. They identified 12 TDFs, five of which mapped on chromosome 1, four on chromosome 2 and three of them were not mapped as they showed no polymorphisms between the parents. The results obtained are in the same range as in our experiment, of the 259 TDFs namely, 36 mapped in the CLO-DvP population and 25 in the ILGI population.

6.4.4 Alignment of the two maps

The integrated CLO-DvP map has been aligned with the ILGI map. The alignment was based on 8 RFLPs, 14 SSRs and 11 STSs (seven of them were developed during this research) that were common (Figure 6.3). Three inconsistent markers between the maps were observed (Table 6.4). This can be due to the detection of multiple loci by the respective RFLP, SSR and STS.

Table 6.4 Overview of the inconsistencies between the ILGI and CLO-DvP maps

Marker name	Marker type	ILGI	CLO-DvP
LP1	SSR	LG7	LG4
CDO459	RFLP	LG5	LG7
TC116908	STS	LG1 and LG6	LG2

LG1: For LG1 of the ILGI population, the mapping of the EST markers identified during this research was repeated but combined only with the SSR, RFLP and AFLP data obtained from Dr. Danny Thorogood (only a subset of the RFLP data as available in the Fogg database) and the SSRs published by Jensen et al. (2005) (Figure 6.4). LG1 still spliced into two

groups but in this map, it is clear that the thioredoxin marker maps in the neighbourhood of S (3 cM away), similar to that was mentioned by Langridge et al. (1999) for *P. coerulescens*. Using all the data available in the Fogg database, we were not able to map the thioredoxin near the S-locus. As explained above, this can be due to the mislabelling of some progeny plants in different experiments as all the recently added markers (from this study and from Jensen et al 2005) mapped together in a separate group. The thioredoxin marker was also mapped in the CLO-DvP population but in this case, it was located 19 cM away from the position predicted for S in this population (S and Z have not been mapped in this population but it is known that S is located in the neighbourhood of the OSE-3 marker). In the CLO-DvP population the CDB069_380D marker, which shows a down regulated expression profile in the self-pollinated pistils and is absent in the cross pollinated pistils (Figure 6.3) is about 10 cM away from the OSE-3 locus. This gene displayed homology to an ATP binding protein of the ABC transport system of bacteria (Q5HR53). In the ILGI population, four other EST markers mapped in LG1 (Figures 6.3 and 6.4 and Table 6.5).

Table 6.5 Overview of the fragments mapped on LG1 in the ILGI population.

Marker	Expression profile	Distance from S	Homology to	% homology
CDB100_340Bb	Down regulated in SP pistils and not expressed in CP pistils	4 cM (Figure 6.4)	ABC transport permease of bacteria (Q5WDD5)	35% in 72 aa
CDB202_280F	Not expressed in SP pistils and constitutively expressed in CP pistils	3 cM (Figure 6.3)	No homology	
CDB232_500A	Down regulated in SP and CP pistils	LG1a (Figure 6.4)	No homology to known sequences	
TC116908b	Not known	5 cM (Figure 6.3)	Putative gag-pol protein of *Arabisopsis* (Q6UUR5)	43% in 99 aa

The TC116908b marker mapped to LG1 in the ILGI, does not correspond to the gene mapped in the neighbourhood of Z by Hackhauf and Wehling (2005), as it maps to a different linkage group and displays no homology to ubiquitin specific protease-like proteins, as expected for the gene closely linked to Z. In fact the DNA band of larger size (indicated as TC116908a in Figure 6.5), which could not be mapped in the ILGI population, displays 60%

homology to an ubiquitin specific protease-like protein of *Arabidopsis* (Q9LEW0). TC116908a mapped to LG2 in the CLO-DvP population (see below).

LG2: Several markers derived from cDNA-AFLP fragments and the STS markers (TCxxxxx) published by Hackauf and Wehling (2005) mapped on LG2. According to Hackauf and Wehling (2005), TC116908a co-segregates with Z, TC101821 maps at 1 cM from Z and TC89057 maps at 0.5 cM from Z in rye. TC116908a displayed homology (50% in 70 amino acids) to an ubiquitin specific protease-like protein of *Arabidopsis* (Q9LEW0) and corresponds to the gene mapped by Hackauf and Wehling (2005) near Z in rye. Given that ubiquitination of proteins is emerging as a general mechanism apparently involved in different SI-systems (Chapter 2, Chapter 5), it is of great interest to further analyse this gene.

In the CLO-DvP population the locus detected by the primer pair CDB187_305 showed an interesting expression profile, namely, down regulated in the self-pollinated pistils and absent in the cross-pollinated pistils. This marker co-segregates with the STS marker TC116908. For CDB187_305 homology has been found to a putative cellular retinaldehyde-binding/triple protein. This protein is involved in intracellular protein transport and enables the directed movement of proteins in a cell, including the movement of proteins between specific compartments or structures within a cell, such as organelles of a eukaryotic cell (http://www.godatabase.org/cgi-bin/amigo/go.cgi?view=details&query=0006886).

Other markers, derived from cDNA-AFLP fragments map to LG2 in one or both populations but they are not entirely co-linear between the two populations and the results obtained in rye (Hackauf and Wehling 2005). This may be a result of sampling effects or chromosomal rearrangements. The cDNA-AFLP derived marker which maps closer to Z (CDB084_370a), is located 23 cM away from Z in the ILGI population. However, some of the markers of this LG show interesting expression profiles (Figure 6.3). The genes CDB154_460 and CDB037_320 are up regulated in the self-pollinated pistils and are not or constitutively expressed in the cross pollinated pistils. The gene CDB004_170 is expressed at 24 hours after either self or cross pollination and can be involved in the degeneration of the pistil. Homology with known genes was not found for any of these genes.

LG3: Eight markers and one mapped to LG3 in the CLO-DvP and the ILGI population, respectively. Four of the markers on CLO-DvP LG3 showed distorted segregation. In the ILGI population, one gene, CDB114_370b, with a down regulated expression profile in the self-pollinated pistils, maps near the isozyme locus GOT/3. In the study of Thorogood and Hayward (1992) joint segregation of GOT/3 with self-fertility was observed in the analysed backcross families. In addition, significant segregation distortion was observed at the GOT/3 locus in the selfed progeny of these families. By that time, it was already known that S was linked to the isozyme locus PGI/2 (Cornish et al. 1979) and these results were used as evidence that the Z locus was linked to the isozyme locus GOT/3 (Thorogood and Hayward 1992). However, in the paper of Thorogood et al. (2002) it was found that GOT/3 maps to LG3 and is therefore not linked to either S or Z. Although, there was still an association of the S-locus with some markers on LG3 indicating an interaction of genes on LG3 with the S-locus as the linkage was associated with significantly distorted segregation of the majority of the markers on LG3 and LOD scores as high as 13 had to be used to split the two linkage groups.

LG4: There is a good agreement between the two maps although some rearrangements are evident. Several markers, showing interesting expression profiles mapped to LG4 for both populations (Figure 6.3). Till now, nothing is known about the involvement of this linkage group in the self-incompatibility response, although it can be that some genes involved in the signal cascade, leading to inhibition of the pollen tube, map to this LG. For six fragments, homology to known genes has been found (Table 6.6). Description of the putative functions of these genes has been discussed in chapter 5. Also the marker SRK (Chapter 3), the key component of the SI-response in members of the Brassicaceae, mapped on LG4 in the ILGI population indicating that SRK probably does not represent a component of the SI recognition process in ryegrasses. The involvement of SRK in the signalling cascade can not be ruled out at this stage.

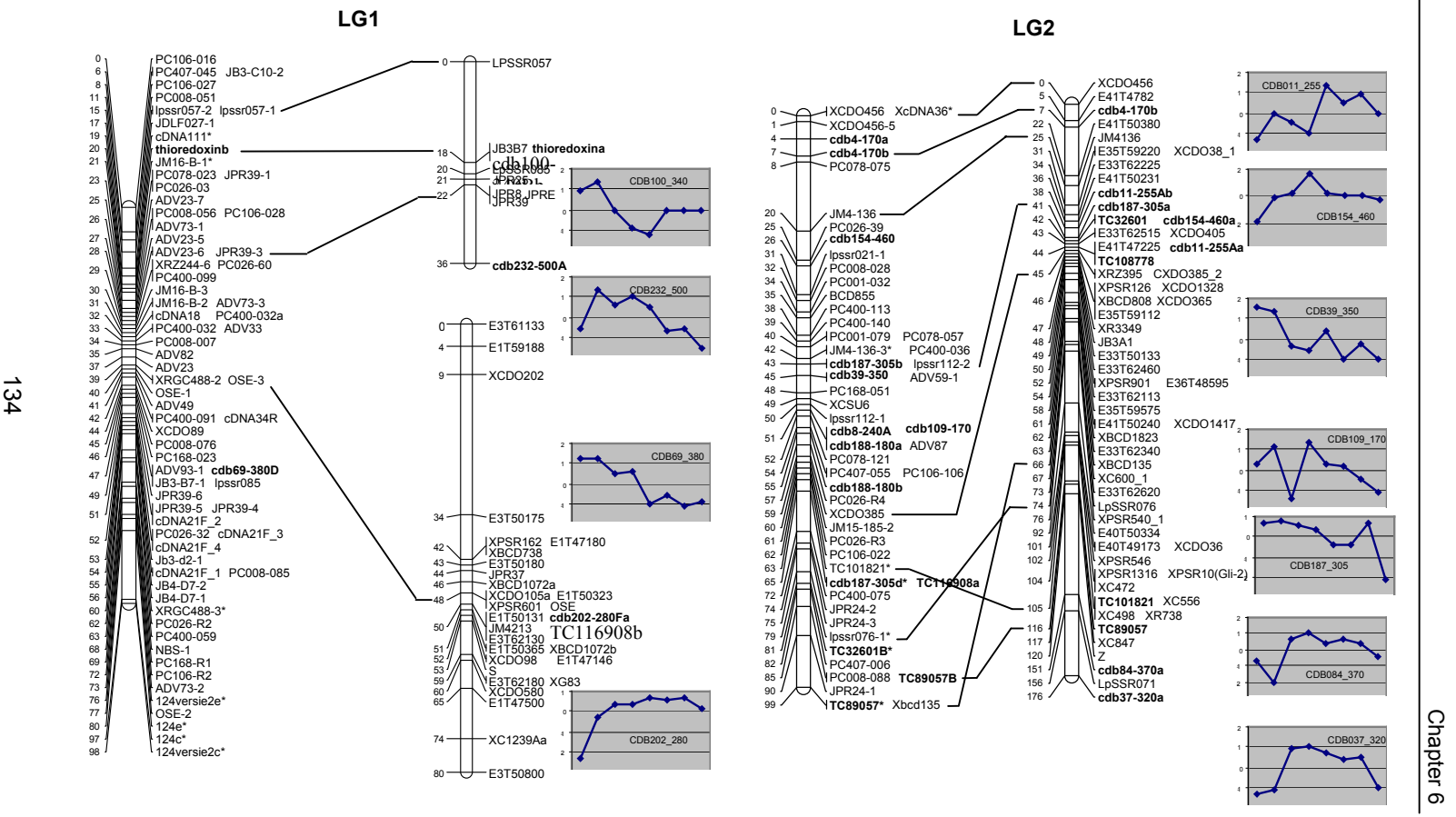
LG5: Recently, Thorogood et al. (2004), identified on this linkage group a self-fertility (T-) locus, functioning through interaction with the S-locus. This trait has not been mapped in the ILGI or CLO-DvP population but should be located in the neighbourhood of the CDO412-2 marker (personal communication, Dr. D. Thorogood). In the CLO-DvP population, marker CDB036_410 maps 3 cM away from the CDO412-2 marker. No putative function could be assigned for this gene.

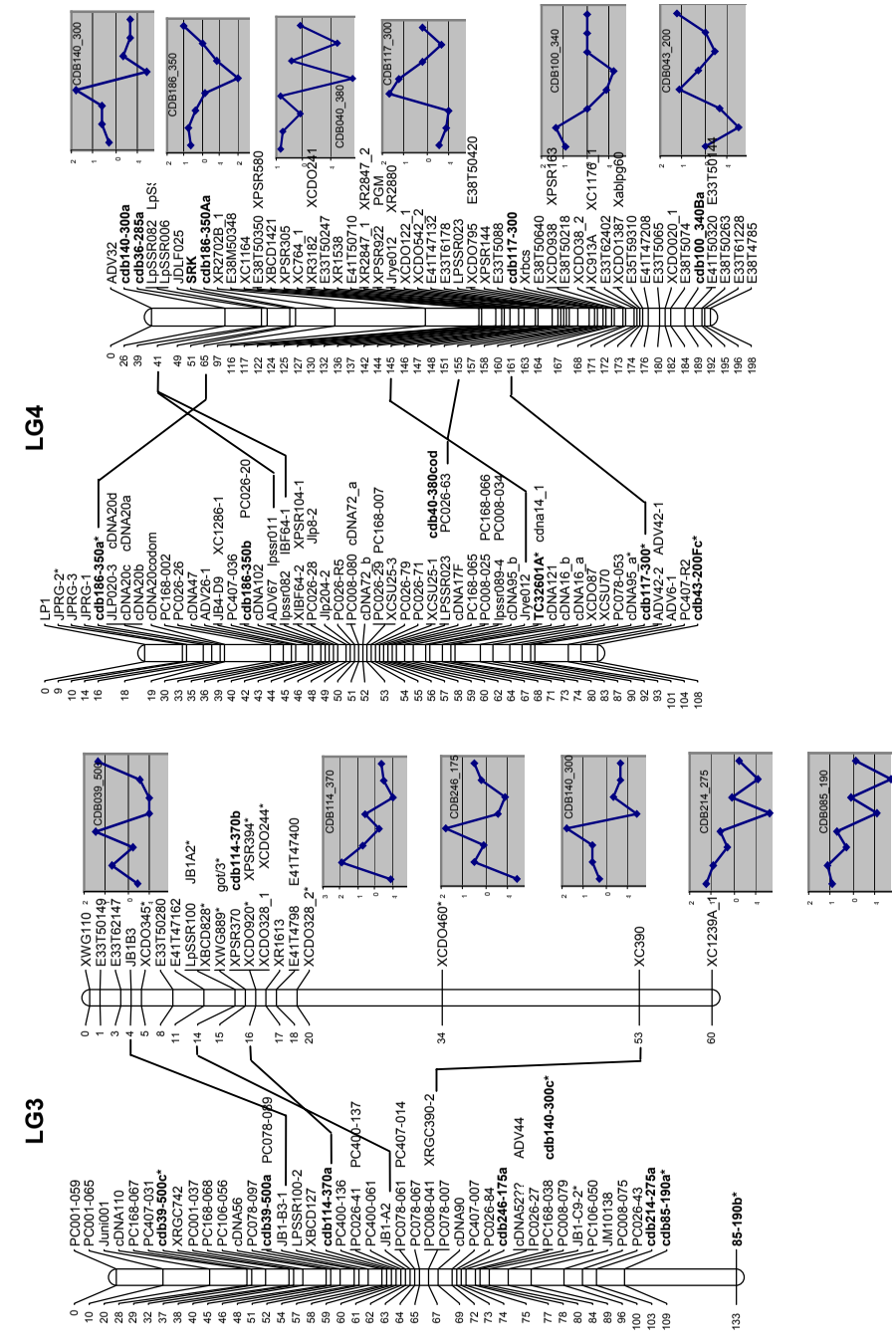
LG6: For LG6 the two maps had only one marker in common. Marker CDB026_440a, showing no homology to known genes, maps in the CLO-DvP population and two cDNA-AFLP derived markers map in the ILGI population. Marker CDB210_440 displayed homology to a serine carboxypeptidase, a protein with a general function, namely release of a C-terminal amino acid with broad specificity and marker CDB013_285b displayed homology to an amino acid transporter-like protein.

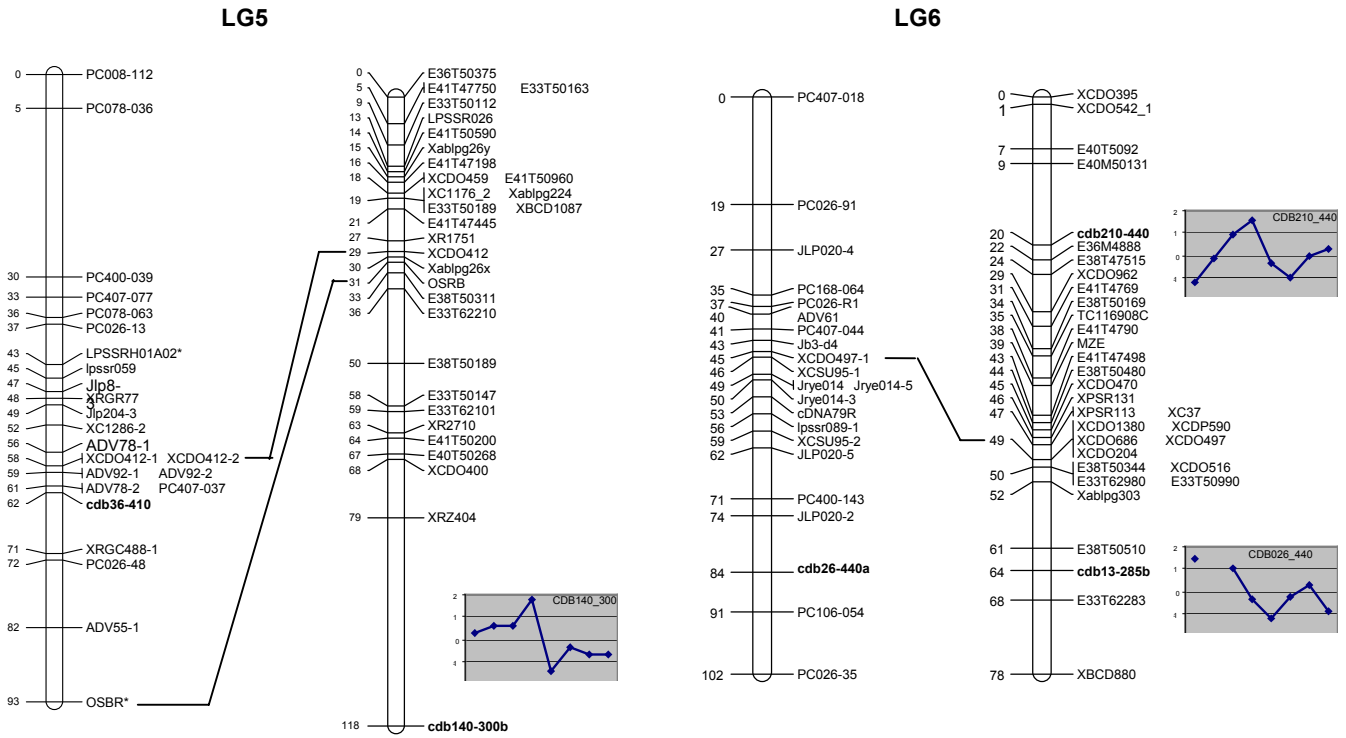
Table 6.6 Markers mapped on LG4 in the CLO-DvP and ILGI population.

Marker	Population	Homology
CDB036_285a	ILGI	No
CDB040_380	CLO-DvP	Actin
CDB043_200Fc	CLO-DvP	Hypothetical protein involved in protein ubiquitination
CDB140_300a	ILGI	No
CDB100_340Ba	ILGI	Oligopeptide ABC transporter permease
CDB117_300	CLO-DvP and ILGI	Oxide cyclase
CDB186_350a	CLO-DvP and ILGI	MADS box gene
SRK	ILGI	Kinase

LG7: Armstead et al. (2004) identified a QTL for heading-date on LG7 in a F_2 *L. perenne* mapping population. This QTL should be located in the neighbourhood of the markers CDO545 and OSW (personal communication Dr. D. Thorogood). In the CLO-DvP population, the marker, CDB248_270a mapped 1 cM away from OSW. In the ILGI population, on the other hand, this marker mapped around 20 cM away from OSW. No homology to known genes has been found for this marker or for the other markers which mapped to this LG (CDB046_165, in the CLO-DvP population and CDB028_275Ca and CDB230_195B in the ILGI population). Only the marker CDB187_305 displayed homology to an intracellular transport protein, but this fragment detected multiple loci, which mapped to different LGs (see also LG2).


LG1
PC106-016
PC407-045 JB3-C10-2
PC106-027
PC008-051
lpssr057-2 lpssr057-1
JDLF027-1
cDNA111*
thioredoxinb
JM16-B-1*
PC078-023 JPR39-1
PC026-03
ADV23-7
PC008-056 PC106-028
ADV73-1
ADV23-5
ADV23-6 JPR39-3
XRZ244-6 PC026-60
PC400-099
JM16-B-3
JM16-B-2 ADV73-3
cDNA18 PC400-032a
PC400-032 ADV33
PC008-007
ADV82
ADV23
XRGC488-2 OSE-3
OSE-1
ADV49
PC400-091 cDNA34R
XCDO89
PC008-076
PC168-023
ADV93-1 cdb69-380D
JB3-B7-1 lpssr085
JPR39-6
JPR39-5 JPR39-4
cDNA21F_2
PC026-32 cDNA21F_3
cDNA21F_4
Jb3-d2-1
cDNA21F_1 PC008-085
JB4-D7-2
JB4-D7-1
XRGC488-3*
PC026-R2
PC400-059
NBS-1
PC168-R1
PC106-R2
ADV73-2
124versie2e*
OSE-2
124e*
124c*
124versie2c*
LPSSR057
JB3B7 thioredoxina
cdb100-
JPR39
cdb232-500A
CDB100_340
CDB232_500
E3T61133
E1T59188
XCDO202
CDB69_380
E3T50175
XPSR162 E1T47180
XBCD738
E3T50180
JPR37
XBCD1072a
XCDO105a E1T50323
XPSR601 OSE
E1T50131 cdb202-280Fa
JM4213
E3T62130
E1T50365 XBCD1072b
XCDO98 E1T47146
S
E3T62180 XG83
XCDO580
E1T47500
TC116908b
XC1239Aa
CDB202_280
E3T50800
LG2
XCDO456 XcDNA36*
XCDO456-5
cdb4-170a
cdb4-170b
PC078-075
JM4-136
PC026-39
cdb154-460
lpssr021-1
PC008-028
PC001-032
BCD855
PC400-113
PC400-140
PC001-079 PC078-057
JM4-136-3* PC400-036
cdb187-305b lpssr112-2
cdb39-350 ADV59-1
PC168-051
XCSU6
lpssr112-1
cdb8-240A cdb109-170
cdb188-180a ADV87
PC078-121
PC407-055 PC106-106
cdb188-180b
PC026-R4
XCDO385
JM15-185-2
PC026-R3
PC106-022
TC101821*
cdb187-305d* TC116908a
PC400-075
JPR24-2
JPR24-3
lpssr076-1*
TC32601B*
PC407-006
PC008-088 TC89057B
JPR24-1
TC89057* Xbcd135
XCDO456
E41T4782
cdb4-170b
E41T50380
JM4136
E35T59220 XCDO38_1
E33T62225
E41T50231
cdb11-255Ab
cdb187-305a
TC32601 cdb154-460a
E33T62515 XCDO405
E41T47225 cdb11-255Aa
TC108778
XRZ395 CXDO385_2
XPSR126 XCDO1328
XBCD808 XCDO365
E35T59112
XR3349
JB3A1
E33T50133
E33T62460
XPSR901 E36T48595
E33T62113
E35T59575
E41T50240 XCDO1417
XBCD1823
E33T62340
XBCD135
XC600_1
E33T62620
LpSSR076
XPSR540_1
E40T50334
E40T49173 XCDO36
XPSR546
XPSR1316 XPSR10(Gli-2)
XC472
TC101821 XC556
XC498 XR738
TC89057
XC847
Z
cdb84-370a
LpSSR071
cdb37-320a
CDB011_255
CDB154_460
CDB39_350
CDB109_170
CDB187_305
CDB084_370
CDB037_320


LG4
LG3


LG5
0 PC008-112
5 PC078-036
30 PC400-039
33 PC407-077
36 PC078-063
37 PC026-13
43 LPSSRH01A02*
45 lpssr059
47 Jlp8-
48 XRGR77
49 Jlp204-3
52 XC1286-2
56 ADV78-1
58 XCDO412-1 XCDO412-2
59 ADV92-1 ADV92-2
61 ADV78-2 PC407-037
62 cdb36-410
71 XRGC488-1
72 PC026-48
82 ADV55-1
93 OSBR*
0 E36T50375
5 E41T47750 E33T50163
9 E33T50112
13 LPSSR026
14 E41T50590
15 Xablpg26y
16 E41T47198
18 XCDO459 E41T50960
19 XC1176_2 Xablpg224
E33T50189 XBCD1087
21 E41T47445
27 XR1751
29 XCDO412
30 Xablpg26x
31 OSRB
33 E38T50311
36 E33T62210
50 E38T50189
58 E33T50147
59 E33T62101
63 XR2710
64 E41T50200
67 E40T50268
68 XCDO400
79 XRZ404
118 cdb140-300b
CDB140_300
LG6
0 PC407-018
19 PC026-91
27 JLP020-4
35 PC168-064
37 PC026-R1
40 ADV61
41 PC407-044
43 Jb3-d4
45 XCDO497-1
46 XCSU95-1
49 Jrye014 Jrye014-5
50 Jrye014-3
53 cDNA79R
56 lpssr089-1
59 XCSU95-2
62 JLP020-5
71 PC400-143
74 JLP020-2
84 cdb26-440a
91 PC106-054
102 PC026-35
0 XCDO395
1 XCDO542_1
7 E40T5092
9 E40M50131
20 cdb210-440
22 E36M4888
24 E38T47515
29 XCDO962
31 E41T4769
34 E38T50169
35 TC116908C
38 E41T4790
39 MZE
43 E41T47498
44 E38T50480
45 XCDO470
46 XPSR131
47 XPSR113 XC37
XCDO1380 XCDP590
49 XCDO686 XCDO497
XCDO204
50 E38T50344 XCDO516
E33T62980 E33T50990
52 Xablpg303
61 E38T50510
64 cdb13-285b
68 E33T62283
78 XBCD880
CDB210_440
CDB026_440

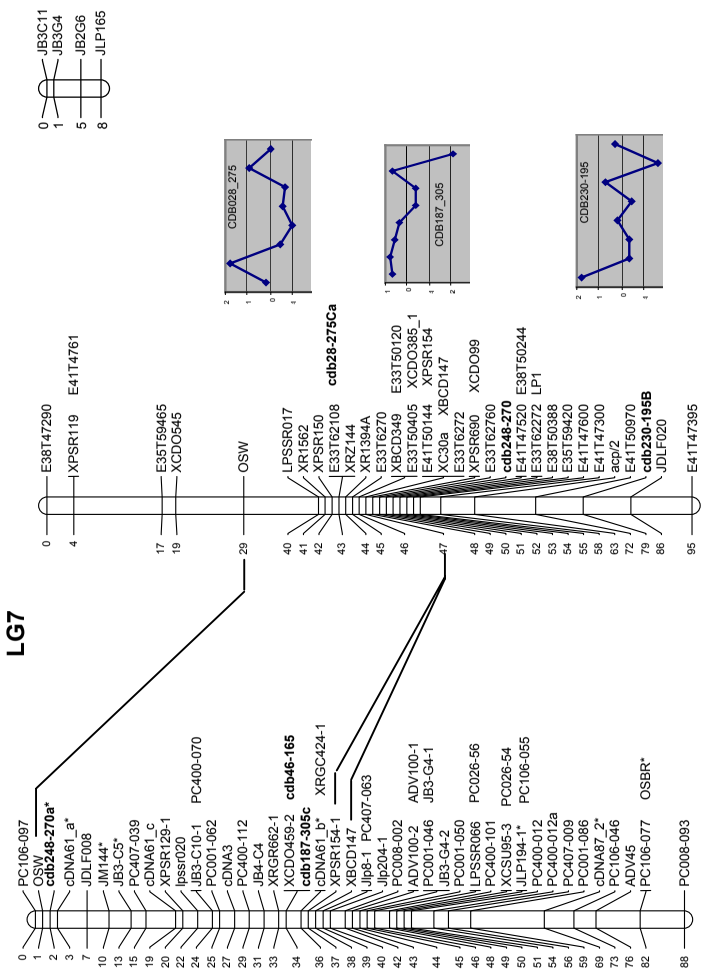


Figure 6.3. Alignment of the two genetic linkage maps, the CLO-DvP population (left) and ILGI population (right) based on SSRs, RFLPs, STSs and ESTs in common. AFLPs have prefix PC and ExxTxxxxx, RFLPs have prefix X, SSRs have a prefix J, Ipssr or ADVxx, isozyms are acp/2, GOT/3 and PGM and STS are indicated with names and the ESTs are indicated as CDBxxx-xxx. Distorted markers (p<0.001) are marked with *. Right: The expression profiles: NP, SP1h, SP4h, SP24h, CP1h, CP4h, and CP24h.

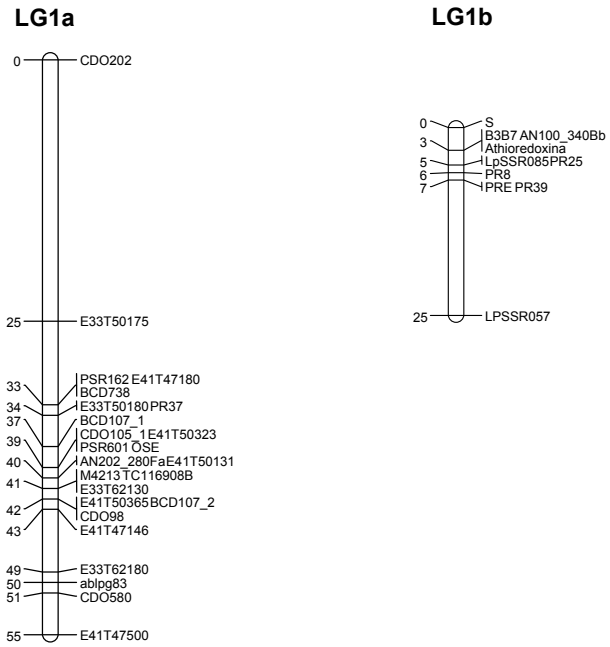


Figure 6.4 LG1 of the ILGI population obtained with the EST markers identified during this research combined with the SSR, RFLP and AFLP data obtained from Dr. Danny Thorogood (only a subset of the RFLP data as available in the Fogg database) and the SSRs published by Jensen et al. (2005).

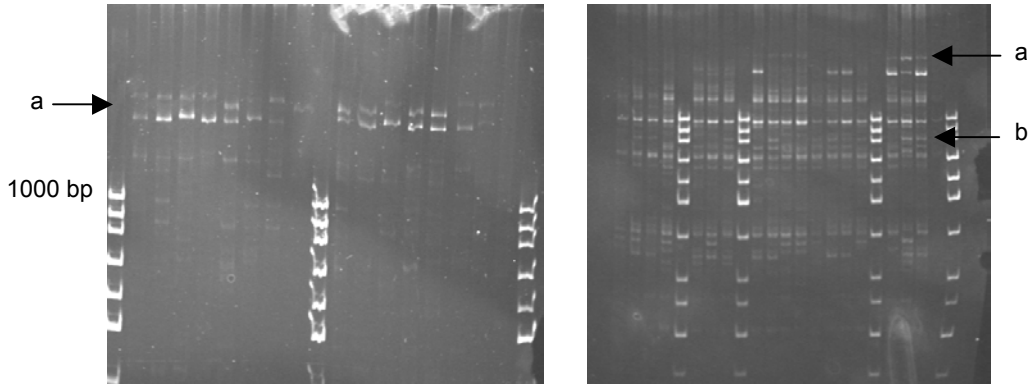


Figure 6.5 Results obtained for primer combination TC116908 on the CLO-DvP and ILGI population. In the CLO-DvP population the marker was scored codominantly, ab x cd; in the ILGI population it was scored as two loci but only TC116908b could be mapped (to LG1). The 50 bp DNA ladder was used as length marker.

6.5 Conclusions

We have successfully transformed some cDNA-AFLP fragments characterized in Chapter 5 into polymorphic markers. Also information recently published for rye was used to map markers derived from gene sequences around the Z-locus in *L. perenne*. The markers have been mapped in two segregating populations, in combination with available AFLP, RFLP, STS and SSR marker information. As the gene sequences from which the markers were derived are putatively involved in SI, the markers developed identified genomic regions involved in SI in *L. perenne*. As expected, the markers mapped over the different LGs with some of them around the S- and Z-loci in LG1 and LG2, respectively. In the ILGI population, the TDF derived marker CDB100_340Bb (down regulated in the self-pollinated pistils and absent in the cross-pollinated pistils), and thioredoxin mapped 3 cM away from S on LG1. CDB100_340Bb displayed homology to an ABC transport permease of bacteria. In LG2, where Z is located, several EST markers and the STS markers published by Hackauf and Wehling (2005) mapped. One of the mapped markers (TC116908a) putatively represents the Z-locus of *L. perenne*, but a thorough characterisation of this gene is required to eventually demonstrate this. In the CLO-DvP population, one additional marker (CDB187_305d) co-segregated with TC116908a. Homology to a cellular retinaldehyde-binding/triple function protein, a protein involved in intracellular transport processes has been found for CDB187_305d. The other EST markers developed in this study mapped across the different linkage groups with a higher concentration on LG3 and LG4.

The constructed CLO-DvP consensus map consisted of 336 loci (121 AFLP loci, 36 RFLP loci, 97 SSR loci, 46 STS markers in addition to the 36 EST markers developed during this research) distributed over the seven linkage groups and spanned 720 cM. The ILGI map consisted of 292 loci (106 AFLPs, 117 RFLPs, 29 SSRs, 3 isozyme markers, 13 STS markers and 24 ESTs, developed in this research) and spanned 852 cM. These results are in the same range as what was found in several previous studies in *L. perenne*. The alignment of the two maps was straightforward and was based on common RFLPs, SSRs and STSs. Only three inconsistencies (see Table 6.4) were observed but this could be due to the use of primer pairs detecting multiple loci.

Fifteen percent and 36% of the markers showed distorted segregation ($P<0.05$) for the ILGI and the CLO-DvP population, respectively. In the CLO-DvP population, segregation distortion was found at $P<0.001$ in LG1, LG2, LG3, LG4 and LG7. The high number of distorted markers in LG1 and LG2 is probably related to the action of S and Z. As found in previous studies (Thorogood et al. 2002), some markers located on LG3 showed strong association to markers located in LG1, in the neighbourhood of the S-locus. Probably, on LG3 genes are located which interact with the S-locus (or a gene tightly linked to it). Segregation distortion of loci on other linkage groups could have a similar cause, or be related to the presence of loci linked to viability.

As mentioned in the introduction, the use of such functional markers enables a rapid transfer of linkage information between related species. Due to the presence of high levels of synteny among the grasses, information about genes involved in the SI-response, generated in *L. perenne* should be easily transferred to other members of the Poaceae. Similarly, the information generated in rye could easily be extrapolated to *L. perenne* and the STS markers originally generated in rye mapped to the predicted LG in *L. perenne*, in the neighbourhood of the Z-locus. In a next step the identified regions involved in SI could be saturated with more markers to allow the application of map-based cloning to identify S and Z and additional genes involved in the signalling cascade, triggered after an SI-reaction.

Chapter 7

Microarray analysis of genes involved in self-incompatibility

7.1 Introduction

7.1.1 Microarray technology

DNA microarrays use between hundreds and thousands of DNA probes arrayed on a solid surface to interrogate the abundance and/or binding ability of DNA or RNA target molecules. The DNA probes that are used in a microarray can be amplified cDNA fragments or synthesized DNA oligonucleotides (Zhu 2003). Both cDNA microarrays and oligonucleotide microarrays are now widely used for monitoring gene expression on a large scale in a wide range of organisms, including plants (Breyne and Zabeau 2001, Aharoni and Vorst 2002). For some organisms such as yeast (Wodicka et al. 1997), a number of bacteria (Laub et al. 2000) and for *Arabidopsis*, microarrays comprising complete gene sets are available. For a few other organisms, the current generation of microarrays is limited to a subset of genes. Despite this limitation, these microarrays have been used successfully to identify genes involved in different biological processes, e.g. differences in gene expression after treatment with three antifungal compounds in wheat (Pasquer et al. 2005) or identification of genes involved in different physiological processes in the seeds of *Brassica* (Soeda et al. 2005).

In *Arabidopsis*, cDNA microarrays have been used to analyze gene expression in various biological processes, such as cell division (Zhang et al. 2005), cell cycle transitions (Beemster et al. 2005), metabolic pathways (Buckhout and Thimm 2003, Scheible et al. 2004), plant defense responses (Schenk et al. 2000, Vandenabeele et al. 2004) and environmental stress responses (Aharoni and Vorst 2002; Bray 2004). The analysis of the processes underlying fruit ripening in strawberries was the first application of microarrays in a non-model plant species (Aharoni et al. 2000).

The most important advantage of microarray-based technology is that large data sets from different experiments can be combined together in a single database. However, the principal limitation for many organisms is that only the fraction of the genes for which either DNA sequence information or a cDNA clone is available can be surveyed. Another important restriction is the difficulty in distinguishing among different transcripts, genes belonging to the same gene family. This is especially true for organisms, including plants, in which a substantial fraction of the genes are members of gene families (Breyne and Zabeau 2001). Another technical limitation is the large amount of RNA needed for hybridization. This problem can be overcome using new

technologies for linear amplification of mRNA samples. Different protocols and commercial kits are already available for the amplification of RNA (Puskás et al. 2002; Minerd 2005), which provide reproducible results. RNA amplification techniques are frequently used in microarray analysis as microarrays expression analysis demands large amounts of RNA which are not always available (Puskás et al. 2002).

In array-based methods such as oligonucleotide arrays and cDNA arrays, complex mixtures of fluorescently labeled polynucleotides (such as cDNA derived from mRNA) are hybridized to large numbers of individual elements (e.g. unique PCR products in cDNA microarrays) attached to a solid surface. As solid surfaces, nylon filters, silicon wafers and glass (microscope slides or micro-capillaries) can be used. Preparation of the microarray slides can be done either by spotting the synthesised probes to the support or by direct synthesis of the oligonucleotides on a surface using photolithography and combinatorial chemistry to fabricate high-density oligonucleotide microarrays. In a next step, the samples under study, typically mRNA preparations from two or more biological sources are fluorescently labelled and hybridised to the microarray slides. Hybridisation is then visualized by fluorescence detection. Fluorescence emission from the microarray is converted into a digital output for each dye and is stored as separate image files. Next, image analysis software is used for quantification of individual array elements (Aharoni and Vorst 2002). In a first step, the background fluorescence from foreign, non-plant array elements, such as yeast clones is subtracted from the raw data. Subsequently, the data are normalized to correct for channel-specific effects such as differences in quantum yield of the dyes and unequal labelling efficiencies of the samples and for differences in the amount of sample used. Several ways of normalization can be used (Aharoni and Vorst 2002). The overwhelming amount of data generated by microarray experiments can then be analysed, using statistical techniques.

7.1.2 The ryegrass microarray

A *Lolium* cDNA microarray has recently been constructed, consisting of cDNA probes derived from candidate genes putatively involved in different biological processes, such as disease resistance, cold tolerance, carbohydrate metabolism, cell wall composition, N use efficiency, protein stability, root formation, leaf expansion (in response to temperature and light), flowering and vernalization, self-incompatibility and fertility (Table 7.1).

These genes were identified by the partners of the European project "Development of ryegrass allele-specific (GRASP) markers for sustainable grassland improvement" (http://www.grasp-euv.dk/). These genes have been identified in several independent experiments, carried out by the different GRASP partners and using homology-based and differential expression techniques such as cDNA-AFLP or SSH (Suppression Subtractive Hybridisation). Printing and microarray preparation were carried out by one of the partners of the project (Plant Research International, PRI, NL), as described in materials and methods.

The main problem of this microarray is that the level of redundancy is unknown, as many of the cDNA probes have not been characterized yet. As a consequence, nothing is known about the 'representativity' of the ryegrass transcriptome which is covered by this microarray. However, this microarray represents, as far as we know, the first effort to construct a publicly available *L. perenne* microarray. The GRASP microarray is presently being used to screen trait-specific contrasting mRNA pools, supplied by the different partners of the project. The clones provided by DvP, the hybridisation conditions and the results obtained by DvP are presented in this chapter.

7.2 Objectives

We have used the GRASP microarray to perform hybridisations of cDNA samples derived from *L. perenne* pistils that were submitted to different fertilization treatments. The objectives were (i) to check the expression profiles of the TDFs putatively involved in the SI-response described in Chapters 4 and 5, and (ii) to determine whether any of the genes provided by other GRASP partners could be related to SI in ryegrass.

7.3 Material and methods

7.3.1 Preparation of the GRASP cDNA microarray

The GRASP cDNA microarray contains 4224 cDNAs (Table 7.1). As mentioned above, these cDNAs correspond to candidate genes involved in different biological processes and which were isolated using different approaches.

Part of the cDNA clones provided by DvP correspond to genes isolated from a *L. perenne* cDNA library of leaf and part of the cDNA clones correspond to the TDFs described in Chapters 4 and 5 of this thesis. cDNA-

AFLP bands were excised from the gel, reamplified and cloned in the TOPO TA cloning kit (Invitrogen) following the manufacturer's instructions. Positive clones were picked and dissolved in MQ water. Direct colony PCR reactions were performed with the universal M13 forward and reverse primers in a total volume of 100 µl with 5 µl of eluted DNA, 10 µl of 10 x PCR buffer (without $MgCl_2$), 3 µl of 50 mM $MgCl_2$, 10 µl of dNTP's (2 mM each), 1 µl of each primer (100 pmol/µl), 69.5 µl MQ and 0.5 µl Taq (5 U/µl). The PCR conditions were: 94°C 30 sec, 30 cycles of 30 sec 94°C, 30 sec 55°C and 2.5 min 72°C with a thermal cycler (9700, Applied Biosystems). Thereafter, the PCR products were checked on 1.5% agarose gels to check for presence of a single PCR product, and to determine its concentration (a minimum of 5 µg was required). Finally, the PCR products were purified in 96-well plates, using the QiaQuick 96 PCR purification Kit (Qiagen) but with 100 µl of 0.1 x TE (= 1 mM Tris/0.1 mM EDTA) (pH 8.0) instead of the elution buffer delivered with the kit. Samples were completely dried, resuspended in 12 µl of 5 x SSC (sodium citrate, printing buffer) and transferred to 384-well plates.

Table 7.1 Overview of the fragments spotted on the *L. perenne* cDNA microarray of the GRASP project.

Number of cDNAs contributed	GRASP Partner	Trait
768	DLF	Flowering time and vernalization
790	PRI	N use efficiency
458	DVP	SI and fertility + anonymous genes
672	IGER	Carbohydrates, cell wall composition and protein stability
576	AUN	Cold tolerance
576	LIA	Cold tolerance
384	INRA	Leaf expansion in response to temperature and light

The cDNA clones, positive controls (luciferase genes) and negative controls (yeast clones, that lacked cross-hybridization to plant genes) were anchored to GAP II amino-silated coated microscope glass slides (Corning Inc. Life Science, Acton, MA) using a 3-pin print head and a custom-built arraying robot. This work was done by PRI, using published methods (Aharoni et al. 2000). Each spot contained 0.5 to 10 ng of PCR product. The cDNA clones were arrayed in duplicate within a 1.7 cm^2 area. The hybridisation test showed that a part of clones was not printed correctly.

These clones (the first 768) were repeated four times. Finally, each glass slide consisted of 10 metarows and 3 metacolums (Figure 7.1).

After arraying, the slides were left to dry overnight. Before hybridisation the slides were rehydrated by holding them (array side down) for ca. 8 seconds over a bath of hot water (ca. 70°C) and snap-dried (array side up) on a 95-100°C hot plate (5 sec). The DNA was UV cross-linked by using a cross linker at 150 mJoules. Finally they were soaked twice in 0.2% SDS (Sodium Dedecyl Sulfate) for 2 min at room temperature (RT) with gentle agitation, soaked twice in MQ water for 2 min at room temperature by gentle agitation, boiled in MQ water for 2 min to allow DNA denaturation, dried, rinsed 3 x in 0.2% SDS for 1 min, rinsed once in MQ water for 1 min, submerged in boiling MQ water for 2 seconds, and dried again. Slides processed in this way can be stored dry in the dark for several months prior to hybridisation.

7.3.2 Sample preparation

Total RNA was isolated from self-pollinated (Lp58 x Lp58, progeny plant of the ILGI population) and cross-pollinated pistils 1h and 8h after pollination using the RNeasy mini kit from Qiagen. The cross-pollinated pistils were collected on genotype Lp58, after pollination with genotype T_2. T_2 was a plant chosen at random among the progeny of the cross $S_{22}6/16$ x $S_{22}6/56$ (see also 4.3.1). We could make use of six microarray slides in this experiment. Therefore, the following contrasts were tested:

1) Self-pollinated pistils 1 h after pollination / self-pollinated pistils 8 h after pollination (SP1h/SP8h)
2) Self-pollinated pistils 8 h after pollination / cross-pollinated pistils 8 h after pollination (SP8h/CP8h)
3) Cross-pollinated pistils 1 h after pollination / cross-pollinated pistils 8 h after pollination (CP1h/CP8h).

The RNA was amplified using the one-round amplification step of the RiboAmp RNA amplification kit (Arcturus). The RiboAmp kit achieves high yields of antisense RNA (aRNA) in a linear amplification process. The unlabeled aRNA can be processed using reverse transcription based methods of label incorporation to produce labelled cDNA for hybridisation onto microarray. In addition, amplification was necessary in our case, as controlled pollinations of *L. perenne* are time-consuming and laborious and therefore, only tiny amounts of plant materials were available for RNA

extraction. The samples were sent to PRI for hybridisation on the GRASP microarray.

The fluorescent labels Cyanine-3 (Cy3) and Cyanine-5 (Cy5) were used. Indirect fluorescent labeling of the mRNA was performed by first-strand cDNA synthesis in the presence of modified nucleotides (amino-allyl-dUTP) using a random hexamer as primer and according to standard protocols (Aharoni et al. 2000).

7.3.3 Hybridisation and scanning

Prior to hybridisation, the slides were incubated for 2 hours at 50°C in a solution of 50% formamide, 5 x Denhardt's reagent, 5 x SSC, 0.2% SDS and 0.1 mg/ml denaturated fish-DNA. One µg of the Cy5-labelled sample and 2 µg of the Cy3-labelled sample were mixed and applied on to the microarray. Each contrast was repeated in 'dyeswap'. The slides were placed in one of the chambers of the hybridisation station overnight at 50°C. After hybridisation, the array was washed by agitation in 1 x SSC containing 0.1% SDS and in 0.1 x SSC containing 0.1% SDS for 5 min each at room temperature. The array was then rinsed briefly in 0.1 x SSC, dried and scanned for fluorescence emission using a ScanArray Express HT (Perkin Elmer, Boston, MA). A separate scan was conducted for each of the two fluorescent dyes used for hybridisation. Using the ScanArray Express software (Perkin Elmer, Boston, MA), the integrated optical density of each individual probe on the array was calculated. Extraction of the data for each probe was performed using a grid constituted of defined circles fitting the size of each DNA spot and measuring the integrated fluorescent intensity inside each circle.

7.3.4 Data analysis

The spot recognition software ScanArray Express was then used to quantify the fluorescent intensities of the microarray spots and to perform the background correction and normalization of each spot. Normalization of the background corrected data was based on the LOWESS method (Quackenbush 2002). Only those genes with a spot flag of 3, ensuring an optimal spot recognition, were retained in the analysis. Only spots whose background corrected signals were greater than one standard deviation above the background for both channels were considered for further data analysis.

The expression data were first log-transformed and for each slide an 'MA-plot' was constructed. A MA-plot is a scatter plot of $M=\log_2(R/G)$ (where R = the background corrected, normalized red fluorescence of the dye Cy5 and G = the background corrected, normalized green fluorescence of the dye Cy3) and A (A = mean ($\log_2(R)$, $\log_2(G)$)). Based on these results, an absolute value of the $\log_2(R/G)$ higher than 1, was used as threshold to select differentially expressed genes.

Several quality controls were performed before the data values of the 'dyeswap' experiment and the replicated spots within the slide were averaged. First, the linear regression was calculated between the $\log_2(R/G)$ values of the 'dyeswap' hybridisations. Those genes which were strongly displaced from the regression line were removed from the dataset. For the rest of the genes, displaying similar values of $\log_2(R/G)$ in the dyeswap hybridisations, the $\log_2(R/G)$ values of the two 'dyeswap' hybridisations were averaged (this value was designated 'meanswap'). In a next step, we integrated the values of the different replicates within the slides. To do this we calculated the standard deviation and the range of 'meanswap' values of the replicates within the slide. Only those data points for which the range was smaller than 1, were averaged. Other data points were excluded from the analysis. Finally a dataset was constructed in which one value was obtained for each gene for each contrast. This strategy has been followed based on the contact with Dr. Paul Van Hummelen from the VIB microarray Facility (MAF-Leuven).

7.3.5 Real-time RT-PCR reactions

Real-time RT-PCR reactions were carried out for some of the TDFs described in Chapters 4 and 5, to confirm the expression profiles obtained for them in the cDNA-AFLP experiment and the results of the microarray hybridisation. Total RNA was isolated from self-pollinated and cross-pollinated pistils 1 h and 8 h after pollination, using the RNeasy mini kit (Qiagen). The obtained RNA was then amplified using the RiboAmp kit (Arcturus) and used for double stranded cDNA synthesis with a random hexamer primer according to standard protocols. All samples were prepared in duplicate. PCR reactions were performed in a total volume of 25 µl with 200 ng cDNA, 12.5 µl SYBRGreen Master mix (ABI), 0.75 µl of each (forward and reverse) primer (10 µM) and MQ water using a 7000 Sequence Detection System (Applied Biosystem). Amplifications were carried out at 50°C for 2min, 95°C for 10 min and 45 cycli of denaturation (95°C, 15 sec)

and annealing (60°C, 1 min). As primers, the gene specific primers have been used, developed for the mapping work (see Chapter 6).

7.4 Results

7.4.1 Expression analysis with cDNA microarray

Three contrasts, corresponding to six hybridisations, were analysed. In each hybridisation one mRNA population was labelled with Cy3 and the other with Cy5. The labelled targets were then mixed and hybridized simultaneously to the microarray. The image obtained for one of the hybridisations is shown in Figure 7.1. To minimize the inherent variability of the microarray assay and to ensure the reliability of the results, the genes were spotted in two replicates and a 'dyeswap' was performed. The 'dyeswap' experiment is a reciprocal labeling experiment with each pair of targets, using the same techniques of the first experiment except that the labels are exchanged (Aharoni and Vorst 2002). For example, in one hybridisation we labeled the SP1h sample with Cy5 and the SP8h sample with Cy3. In the corresponding dyeswap hybridisation, the SP1h sample was labeld with Cy3 and the SP8h sample with Cy5.

7.4.2 Detection of differentially expressed genes

The expression data were first background-corrected and then normalized using the LOWESS method by the ScanArray software. Per slide, spots with background-corrected values smaller than one standard deviation above the background value for one or the two channels were excluded from further analysis. In this way, we ensured that only spots displaying hybridisation signals clearly above the background levels were retained in the analysis. The expression data were then log-transformed, and for each spot the mean $\log_2$ (intensity) and the $\log_2$ (ratio) were calculated. Finally, only those spots with an absolute value of the $\log_2$(R/G) higher than 1 were considered to be up- or down-regulated (Figure 7.2). No clear guidelines exist for the drawing of thresholds which define which genes display differential expression for the conditions tested and we chose this threshold after visual inspection of the MA plots. The threshold chosen in this study (absolute value of Log_2(R/G) > 1) corresponds to a two-fold difference in signal detected for the two channels. Different threshold values have been used depending on the number of differentially expressed genes identified in the literature. Most of the time, genes that were assigned as up- or down-regulated displayed at least a two-fold difference in genes

expression (e.g. Aharoni et al. 2000, Goda et al. 2004, De Paepe et al. 2004, Zhang et al. 2005).

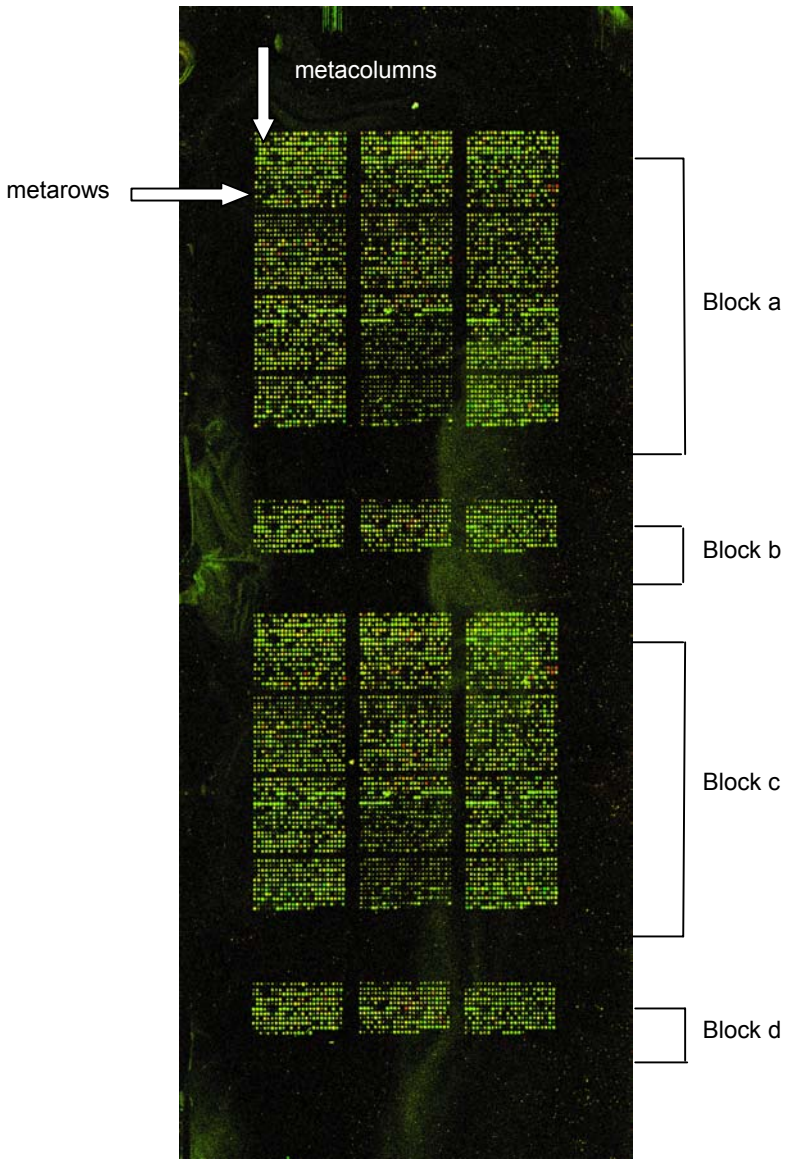


Figure 7.1 Scan of the microarray slide38, in which the contrast SP1h/SP8h was analysed. The SP1h sample was labeled with Cy5 and the SP8h sample with Cy3.

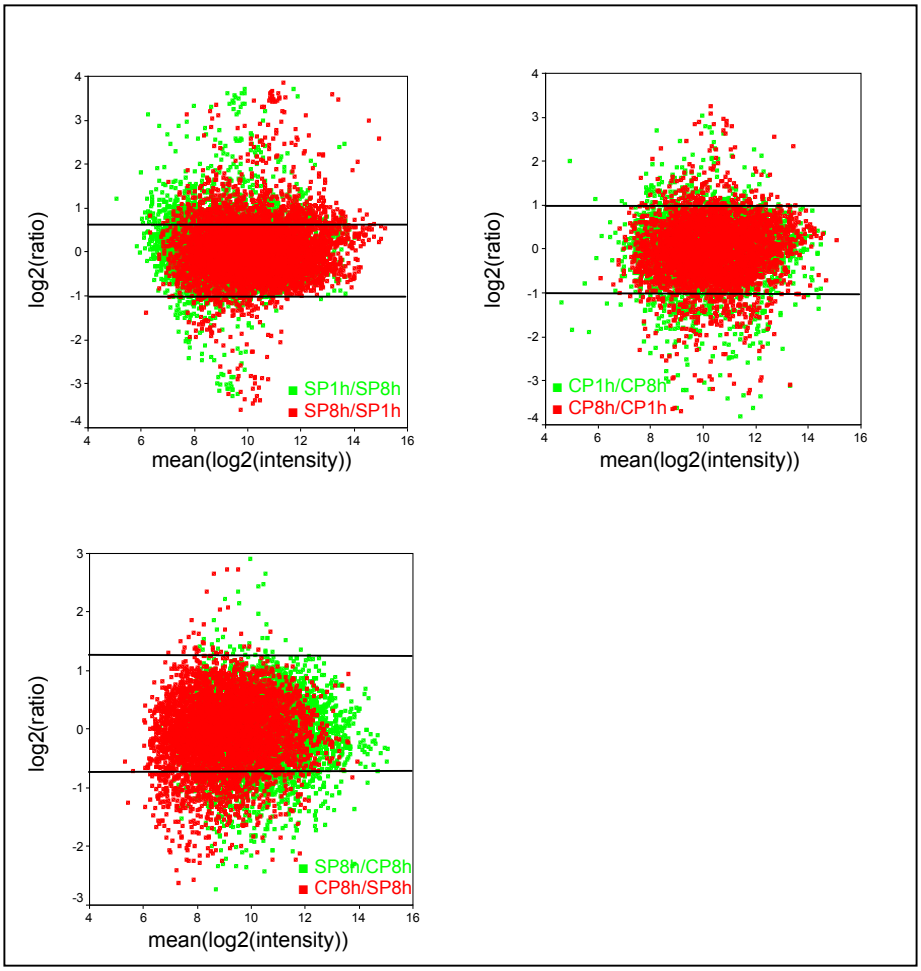


Figure 7.2 The MA-plots of the three different contrasts analysed using the microarray technique.

In a second step we checked whether the normalized values of each spot of the 'dyeswap' experiments could be averaged by calculating the regression between the log(ratio) values of the 'dyeswap' experiments. Three data points were excluded from the data as they were outliers (Figure 7.3), displaying clear differences between the two dyeswap hybridisations. For all the other data points the values of the dyeswap experiment were averaged, to calculate their 'meanswap' value.

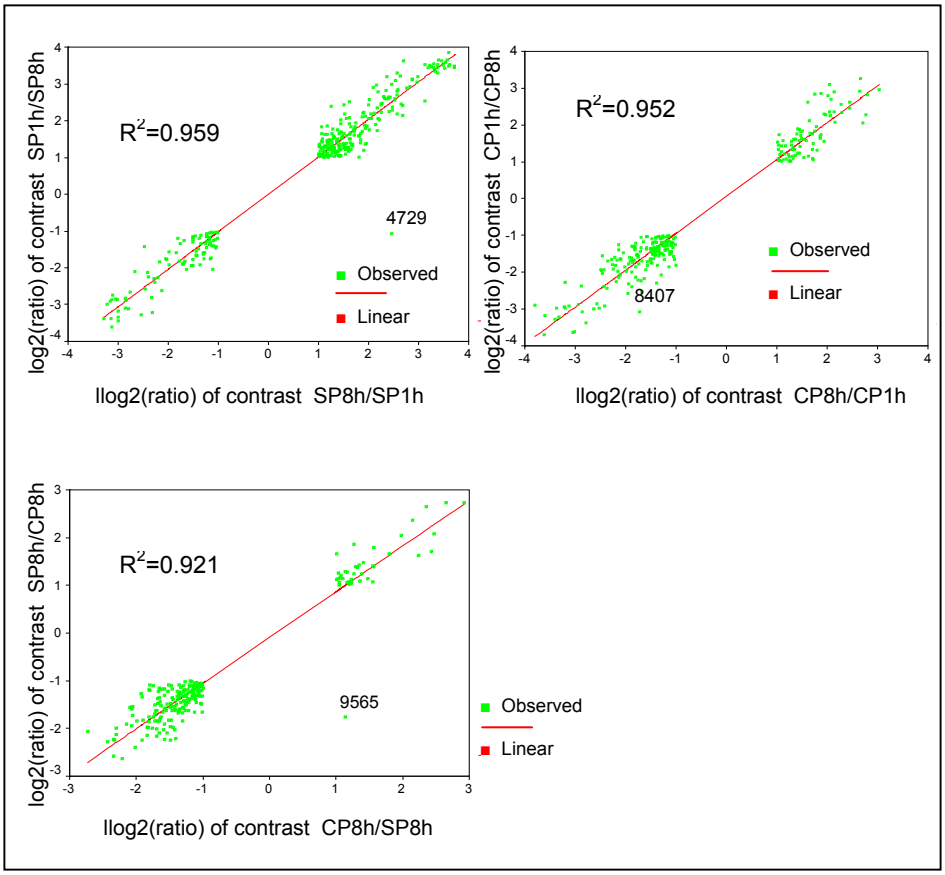


Figure 7.3 The regression analysis of the three different 'dyeswap' experiments of the microarray analysis. Only three genes, labeled by their ID number were removed from the further analysis. For all the other data points the values of the 'dyeswap' experiment were averaged.

In a further step we calculated the standard deviation and the range of the meanswap values of replicates within the experiments. In Table 7.2 an overview is given of the number of replicates per gene available per experiment. For all the data points except one (B17PRInr043, for which 2 replicates were available on slides 38-39 with a range of 5 and a standard deviation of 3.59), the range of the meanswap values was not higher than 1 and the standard deviation not higher than 0.6. Therefore, the values of the replicates were averaged to work further with. Also those genes for which no replicates were available (last row of table 7.2) were retained in the analysis.

It seems that we can trust these values of 'unreplicated' genes as for almost all the genes for which replicates were available, the values of the replicates were similar and could be averaged. Finally, for the SP1h/SP8h, CP1h/CP8h and SP8h/CP8h, 481, 506 and 430 genes, respectively, were considered to be down- or up-regulated.

Table 7.2 Overview of the amount of replicates available per experiment. For example, in the contrast SP1h/SP8h, for 26 genes 4 replicates were available, for 14 genes 3 replicates, for 183 genes 2 replicates and for 258 genes no replicates were available.

# Replicates	Slides 38-39: SP1h/SP8h	Slides 40-41: CP1h/CP8h	Slides 42-43: SP8h/CP8h
4	26	20	27
3	14	18	17
2	183	202	158
0	258	266	228
Total	481	506	430

Combining the data of the three contrasts analysed, we got a common dataset of 1067 genes which were differentially expressed in one or more conditions. As we are interested in the genes involved in the self-incompatibility response, we worked further with the genes which were up- and down-regulated in the SP1h/SP8h comparison and the genes which were up-regulated in the SP8h/CP8h comparison. The genes which were differentially expressed in the CP1h/CP8h comparison were also removed from the dataset, as they probably represented genes involved in the normal ripening process of the pistil but not in SI. The other genes, which probably play a role in cross-pollinations were ignored. Finally a dataset of 371 differentially expressed genes putatively involved in the SI-response was obtained.

No clear guidelines are available for data handling for results of cDNA microarray experiments. Therefore, we defined rather conservative selection criteria at the different steps, to ensure that only those genes which displayed clear differences between the conditions tested, and for which consistent results were obtained in the dyeswap hybridisations and in replicated spots (when available), were selected. However, the degree of replication in the microarray used and the number of conditions tested in this experiment were not optimal. A more detailed time-sequence, analyzed in a

microarray with a higher degree of replication would have been more appropriate. On the other hand, the use of the GRASP microarray represented a unique opportunity to validate the putatively SI-related candidate genes identified by cDNA-AFLP and eventually identify new candidates among the probes contributed by other partners.

The number of differentially expressed genes identified in microarray experiments, depends on the biological processes under study. In most studies only those genes for which replicated data are available and display consistent results are retained. For example, Schenk et al. (2000) selected a set of 705 ESTs which displayed a significant differential expression in response to one or more treatments but only those fragments displaying the similar differential expression profiles in both replicates were considered. Aharoni et al. (2000) used a microarray containing 1701 cDNA clones, spotted in duplicate. In total, 401 cDNA clones were identified as being differentially expressed. De Paepe et al. (2004) used the microarray technique in combination with cDNA-AFLP in *Arabidopsis* to determine a group of ethylene-regulated genes. Out of the 6008 genes present on the chip in duplicate, 214 genes were significantly and consistently differentially expressed. Endo et al. (2004) used a rice cDNA microarray containing 4304 clones to identify novel anther-specific genes. The clones were spotted in duplicate. A total of 396 cDNA clones displaying increased expression in the anthers compared to leaves were selected.

7.4.3 Comparison of cDNA-AFLP and microarray data

One of the objectives of the experiments presented in this Chapter was to eventually validate the expression profiles obtained by cDNA-AFLP. Out of the 371 genes selected in the microarray experiment, 20 were common with the cDNA-AFLP experiment. These genes were used to compare the agreement between the results of the cDNA-AFLP experiment and the microarray experiment. These 20 common genes are listed in Table 7.3 and their expression patterns in the microarray experiment and in the cDNA-AFLP experiment are shown in Figures 7.4.and 7.5 respectively. Group a in Figure 7.4 comprises transcripts that displayed higher expression levels in sample SP1h than in sample SP8h. Group b contains genes with higher expression levels in SP8h than in SP1h and the genes of group c displayed higher expression levels in sample SP8h than in CP8h. The results obtained for most transcripts (except three, CDB042_140, CDB053_250 and CDB099_210) in the microarray experiment were completely different to the

results of the cDNA-AFLP experiment. For example, in the cDNA-AFLP analysis the TDFs CDB021_290, CDB052_360 and CDB126_160 displayed lower levels of expression in sample SP8h than in sample SP1h (Figure 7.4), but exactly the opposite in the microarray analysis (higher expression levels in sample SP8h) (Figure 7.5**)**.

Table 7.3 Overview of the positive cDNA-AFLP genes, displaying an interesting expression profile concerning the SI, selected using the microarray analysis.

ID	SP1h/SP8h log2(ratio)	SP8h/CP8h log2(ratio)	
CDB008_240A	1.4622		*Lycopersicon esculentum* lipoxygenase tomloxA (tomloxA) gene, promoter region
CDB013_285	1.2912		Amino acid transporter-like, *Oryza*
CDB014_180	1.3443		No homology
CDB021_290A	-1.1440		IDI2, *Hordeum*
CDB031_290D	1.2990		Putative ABC transport system ATP-binding protein
CDB036_290	1.1109		Putative regulatory protein, *Oryza*
CDB036_315		1.1462	Germin B, H*ordeum*
CDB042_140C	1.0473		No homology
CDB042_215		1.2506	Putative regulatory protein, *Oryza*
CDB050_125	1.0667		No homology
CDB052_360	-1.0998	1.0234	Starch branching enzyme lib, *Hordeum*
CDB053_250		1.3671	Putative senescence-associated protein (Fragment), *Pisum*
CDB093_410	2.5103	2.1737	Putative phosphoesterase, *Oryza*
CDB099_210E	1.7716		No homology
CDB099_210H	2.6960		Putative SKD1 protein, *Oryza*
CDB126_160A	-1.1190		Unknown function
CDB168_260	1.1299		Putative auxin-responsive factor (ARF1), *Oryza*
CDB202_280F	1.2309		Ubiquitin carboxyl-terminal hydrolase 35 (EC 3.1.2.15) (Ubiquitin thiolesterase 35), human
CDB228_240	1.1777		Germin D, *Hordeum*
CDB256_225D		1.0219	Unknown function

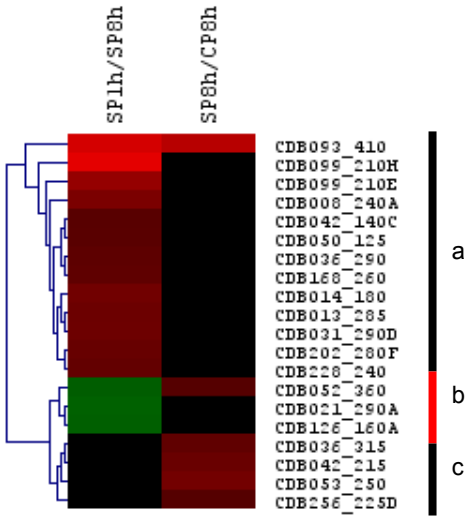


Figure 7.4 Expression profiles of the genes selected using the microarray analysis.

In order to determine which of the two experiments was 'correct', and to find out the origin of these discrepancies, we performed Real-time RT-PCR reactions on a subset of eight transcripts for which clear discrepancies were found between the two expression experiments (Table 7.4). mRNA from pistils at the same stages studied in the microarray analysis were used. In all eight cases tested high consistency was found with the cDNA-AFLP analysis but not with the microarray results. An example is shown in Figure 7.6. According to the Real-time RT-PCR, TDF CDB093_410 displays higher expression levels in sample SP8h than in sample SP1h and higher expression levels in CP8h than in CP1h, which is in contrast to what was found in the microarray analysis. These results agree with the cDNA-AFLP analysis (Figure 7.5). It is difficult to find now a reason for these discrepancies, but one possibility is that some of the samples of the microarray experiment were switched. The microarray analysis of new samples would shed light on this point. It has to be pointed out, that even if (some of) the samples were switched during the microarray experiment, the genes identified as differentially expressed would also be differentially expressed if no switching had happened. The only consequence is that an opposite differential expression would have been found. Therefore, we will not make any further allusion in the discussion which follows to relative differences found between samples in the microarray experiment.

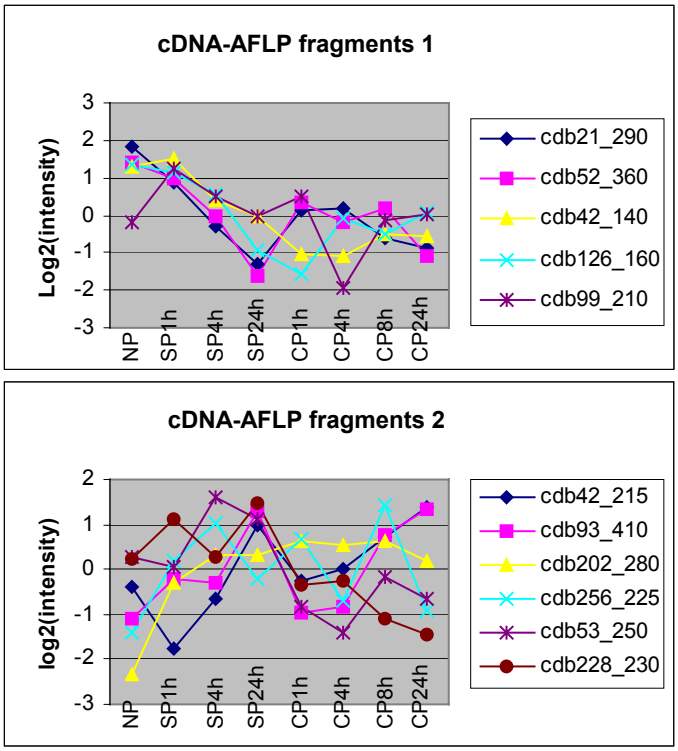


Figure 7.5 The cDNA-AFLP expression profiles of the selected fragments in the microarray analysis. NP: non-pollinated; SP: self-pollinated and CP: cross-pollinated.

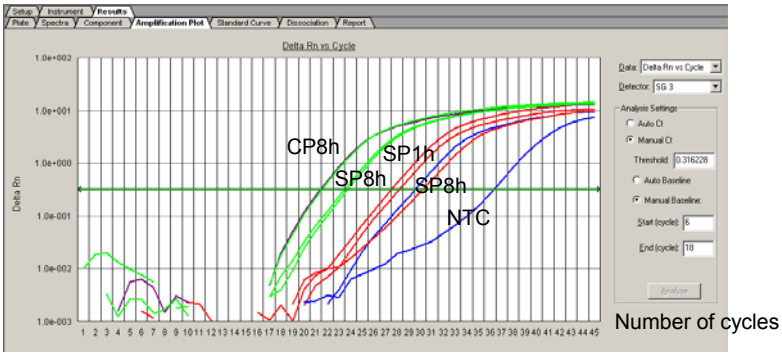


Figure 7.6 Example of amplification plots obtained by Real-time RT-PCR. The figure shows the results obtained for TDF CDB093_410 on 9 samples. SP1h and SP8h: self-pollinated pistils, 1h and 8h after pollination; CP1h and CP8h: cross-pollinated pistils 1h and 8h after pollination; NTC: negative control containing MQ water and no cDNA template. Sample SP8h displays a higher expression than sample SP1h. Sample CP8h displays also higher expression than CP1h. These results agree with the results of the cDNA-AFLP experiment.

Table 7.4 Results obtained from the Real-time RT-PCR. CT value: number of cycles needed to pass the threshold.

	CT value		CT value		CT value		CT value
CDB036_290				CDB093_410			
SP1h	27.09	SP1hbis	26.69	SP1h	27.42	SP1hbis	28.14
SP8h	25.66	SP8hbis	24.61	SP8h	23.37	SP8hbis	23.64
CP1h	25	CP1hbis	24.85	CP1h	29.49	CP1hbis	30.18
CP8h	23.76	CP8hbis	23.22	CP8h	21.31	CP8hbis	21.44
NTC	Undetermined			NTC	36.24		
CDB042_215				CDB101_440			
SP1h	24.82	SP1hbis	24.15	SP1h	24.33	SP1hbis	23.7
SP8h	23.29	SP8hbis	22.44	SP8h	23.52	SP8hbis	23.07
CP1h	24.31	CP1hbis	23.7	CP1h	21.25	CP1hbis	21.42
CP8h	20.94	CP8hbis	20.22	CP8h	21.75	CP8hbis	21.48
NTC	Undetermined			NTC	41.83		
CDB053_250				CDB126_160A			
SP1h	12.3	SP1hbis	13.19	SP1h	13.05	SP1hbis	14.07
SP8h	13.7	SP8hbis	14.25	SP8h	15.59	SP8hbis	16.12
CP1h	14.36	CP1hbis	15.18	CP1h	16.02	CP1hbis	16.82
CP8h	14.02	CP8hbis	13.96	CP8h	16.13	CP8hbis	16.05
NTC	33.96	NTC	34.08	NTC	43.57	NTC	41.76
CDB069_380F				CDB202_280F			
SP1h	21.56	SP1hbis	21.61	SP1h	27.72	SP1hbis	26.77
SP8h	23.77	SP8hbis	23.61	SP8h	25.95	SP8hbis	25.55
CP1h	22.42	CP1hbis	22.88	CP1h	25.7	CP1hbis	25.56
CP8h	23.02	CP8hbis	22.82	CP8h	25.16	CP8hbis	24.7
NTC	Undetermined			NTC	35.01		

7.4.4 Functional classification

The second aim of the work presented in this chapter was to eventually identify other genes, putatively involved in the SI-response among the clones contributed by other GRASP partners for the construction of the microarray. Therefore, for the 351 genes selected in the microarray hybridisations which were not common with the cDNA-AFLP analysis, a functional analysis was done as described in Chapter 5.

For 147 genes, no sequence information was available. Among the 204 genes for which sequence information was available, 8% displayed no homology to genes with known function. The rest of the selected genes were classified into 12 functional categories, based on the MIPS classification system (http://mips.gsf.de/proj/funcatDB/) (Figure 7.7), as described in Chapter 5. The unclassified category makes up the greatest fraction and consists of genes of unknown function and genes displaying homology to genes with known function but for which no gene ontology information was available. Similar categories to those identified and discussed in Chapter 5 were found, but the percentage of genes belonging to a certain cluster was slightly different (Table 7.5). More genes involved in energy processes, protein fate and protein synthesis were identified as differentially expressed and less genes involved in transcription, cell cycle and DNA processing were identified in this experiment. This is probably related to the fact that on the GRASP microarray numerous genes belonging to these categories were present.

The results indicate an overrepresentation of differentially SI-related genes in the following functional categories: metabolism, energy and protein synthesis. Twenty-eight genes displayed homology to genes involved in cellular metabolism processes such as biosynthesis of fatty acids, carbohydrates and amino acids. Two interesting homologies were found with genes which display homology to glyceraldehyde 3-phospahte dehydrogenase (GADPH), an enzyme that plays an important role in glycolysis and gluconeogenesis by reversibly catalyzing the oxidation and phosphorylation of D-glycerladehyde-3-phosphate using NADP as a cofactor. This enzyme binds to actin and may thus play a role in cytoskeleton assembly (http://www.ebi.ac.uk/interpro/IEntry?ac=IPR000173). GADPH displays also diverse non-glycolytic functions such as translocation of GADPH to the nucleus. This translocation acts as a signal for programmed cell death or apoptosis, a mechanism recently thought to be involved in the SI-response (Thomas and Franklin-Tong 2004).

Twenty-eight genes which play a role in energy processes were also found. Electron transport is any series of protein complexes and lipid-soluble messengers that convert the reductive potential of energized electrons into a cross-membrane proton gradient. This proton motive force is used to produce ATP. Another group of twenty genes showed homology to ribosomal proteins that catalyze mRNA-directed protein synthesis in all

organisms. Numerous biological processes needed the synthesis of new proteins, including SI.

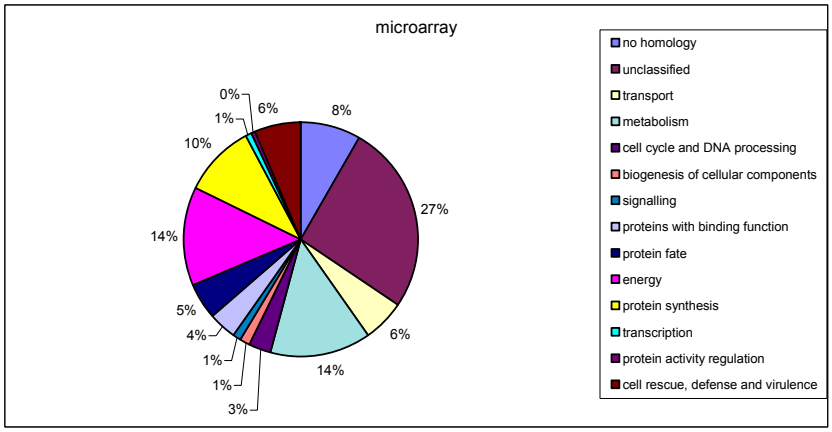


Figure 7.7 Functional classification of the genes probably involved in the SI-response in ryegrass, using the MIPS classification system.

Table 7.5 Comparison of the functional categories obtained in the cDNA-AFLP and microarray experiment.

	cDNA-AFLP (%)	Microarray (%)
No homology	18	8
Unclassified	25	27
Transport	8	6
Metabolism	11	14
Cell cycle and DNA processing	9	3
Biogenesis of cellular components	5	1
Signaling	3	1
Proteins with binding function	2	4
Protein fate	2	5
Energy	5	14
Protein synthesis	3	10
Transcription	5	1
Protein activity regulation	0	0
Cell rescue, defense and virulence	2	6

7.5 Conclusions

In the experiments described in this chapter the *L. perenne* GRASP microarray was used to confirm the expression profiles of the TDFs described in Chapters 4 and 5 and to eventually select additional genes involved in SI among the clones contributed by other GRASP partners to the microarray. As the number of slides available was limited, only three contrasts SP1h/SP8h, SP8h/CP8h and CP1h/CP8h could be analysed

We defined rather stringent selection criteria at the different steps of data handling, to ensure that only those genes which displayed clear differences between the conditions tested, and for which consistent results were obtained in the dyeswap hybridisations and in replicated spots (when available), were selected. Finally, a set of 371 differentially expressed genes was selected. This figure is in line with previous results obtained in microarray experiments with other organisms

The use of microarray hybridisation to confirm the results obtained in the cDNA-AFLP experiment would be less time-consuming than performing Real-time RT-PCR reactions for each individual TDF to confirm their expression patterns. Out of the 371 differentially expressed genes selected in the microarray analysis, 20 originated from the cDNA-AFLP experiment described in Chapter 4 and were used to compare the results of the two methodologies. Surprisingly, for 17 out of 20 transcripts, discrepant results were obtained by cDNA-AFLP and microarray analysis. Therefore, Real-time RT-PCR reactions were performed for a subset of 8 genes. In most cases, consistency was found among the results of the Real-time RT-PCR and those from the cDNA-AFLP, and no agreement with the microarray data. From these results we concluded that (some of) the samples used in the microarray might have been switched. In which way and for which samples this happened is impossible to find out and further research is still needed to confirm this. Either new microarray hybridisations oo Real-time RT-PCR analyses could be used.

With regards to the identification of other genes involved in the SI-response in *L. perenne*, the 204 genes which showed differential expression in the microarray experiment and for which sequence information was available, were assigned to functional categories according to the MIPS classification system. The categories found were similar to those previously described for the cDNA-AFLP TDFs, but with an overrepresentation of

metabolism, energy and protein synthesis. These categories are probably overrepresented in the GRASP microarray. This is not surprising if we take into account that the selection of cDNA clones to spot on the GRASP microarray was based on traits for which these functional processes are relevant (e.g. N use efficiency, carbohydrate content, cell wall composition and protein stability, cold tolerance and leaf expansion). For some of the clones spotted on the microarray, high differences in expression levels were found in the different pistil samples analysed. However, further investigation is necessary to obtain reliable data.

Part C

Conclusions and perspectives

Chapter 8

General conclusions and future perspectives

8.1 Introduction

In self-incompatible grass species, including *L. perenne*, the genetic control of the recognition between pistil and pollen is gametophytic, involving the action of two multiallelic loci (*S* and *Z*) located on different chromosomes. The products of the S- and Z-genes in the pistil and in the pollen interact in such a way that unrelated pollen (non-self pollen) is allowed to penetrate the stigma surface and eventually effect fertilization, while the penetration of related pollen (self-pollen) into the stigma is inhibited. However, it is still unknown which kind of molecules are involved in the initial recognition pollen/pistil and no clues are available about the processes which are activated after an SI-reaction and which finally result in the inhibition of the pollen tube growth.

The main objective of this thesis was then to get insights into the molecular basis of the SI-mechanism in *L. perenne*. We did not concentrate on the identification of the SI-components responsible for the recognition between pistil and pollen (*S* and *Z*), but the strategy followed was more general and aimed also at identifying genes involved in the signalling cascade activated by an SI-reaction.

Different approaches were applied to reach these objectives, resulting in the identification of a set of SI-related candidate genes. For some of these candidates the genome position was determined, resulting in the identification of genomic regions which control SI in *L. perenne*. Given the strong syntenic relationships within the Poaceae, the identification of the corresponding genomic regions in other species of the grass family is now possible.

8.2 General conclusions

8.2.1 Relation between the SI-mechanism of L. perenne and SI in other plant families

One of the first research questions formulated by the beginning of this investigation was whether the SI recognition in *L. perenne* was controlled by gene families known to be active in the SI-mechanisms of other plant families. We used primers derived from conserved sequences in S-RNase, SLG and SRK genes to interrogate the ryegrass genome for orthologues of these genes.

Primers derived from Solanaceae and Rosaceae S-RNases amplified in *L. perenne* DNA sequences with no homology to known S-RNases or to any other kind of RNases. Based on these results, we concluded that no S-RNase based SI-system is active in *L. perenne*. This is not surprising, as beyond the similarities with regards the genetic control of SI, very little SI-features are common for grasses and species displaying S-RNase based SI. For example, the arrest of pollen tube growth in grasses is very rapid, directly on the stigma surface, while in the species displaying an S-RNase based SI-system arrest of the pollen tube growth takes place within the transmitting tract of the style (Chapter 2).

From a dynamic point of view, the inhibition of the pollen tube in members of the grass family displays clear similarities with the system active in *Brassica* (cauliflower family). SI in *Brassica* is based on the action of a receptor kinase (SRK) on the female side and a cysteine-rich protein (SRC or SP11) on the male side (Chapter 2). Therefore, the female-component of the SI-recognition in *L. perenne* could also be a receptor kinase. Indeed, primer pairs derived from *Brassica* SRK genes amplified in *L. perenne* a DNA band with homology to (S-)receptor kinases. However, this kinase gene was expressed not only in *L. perenne* pistils, but also in vegetative tissues, and probably does not represent a component of the SI-recognition process in ryegrasses. This conclusion is also supported by its map position, as it mapped in LG4 and not in LG1 or LG2 of *L. perenne*, where *S* and *Z* are located. However, the involvement of the identified gene in the signalling cascade triggered by the SI-response can not be ruled out at this stage.

8.2.2. Exploitation of knowledge in other species of the Poaceae

For a certain time, it was believed that one thioredoxin gene identified in *Phalaris coerulescens* was the pollen product of the S-locus. We used information from *P. coerulescens* to amplify the homologue of this gene in *L. perenne.* Surprisingly, the *L. perenne* thioredoxin gene was not only expressed in reproductive tissues (non-pollinated and self-pollinated pistils), but also in leaves of ryegrass. These results contrasted with the observations of Li et al. (1994), who only found the transcription product of this gene in mature pollen. Shortly later, it was demonstrated that the thioredoxin gene of *P. coerulescens* was not the S-locus, but that it mapped 2 cM from the S-locus (Langridge et al. 1999). This finding agrees with the results obtained in the present thesis, as the marker derived from the thioredoxin gene mapped in the neighbourhood of S (3 cM away). Anyhow,

the identified *L. perenne* thioredoxin gene constitutes an interesting subject for future investigations, as thioredoxins have been shown to play important roles in the SI-response of *Brassica* (Cabrillac et al. 2001). It is therefore possible, that also in the grass family the characterised thioredoxin gene is involved in SI, in a similar way as in *Brassica*.

In a recent study, Hackauf and Wehling (2005) identified STS markers which mapped in the neighbourhood of Z in rye. One of the markers (TC116908) fully co-segregated with Z in a mapping population of 204 individuals. We made use of this information and determined the map position of these STS markers in the ryegrass genome. In our study, four of the markers tested mapped on the long arm end of LG2, where also Z is located according to Thorogood et al. (2002). We could not determine the map distance between Z and TC116908, as Z is mapped in the ILGI population and marker TC116908 was only polymorphic in the CLO-DvP population. The TC116908 markers of rye and *L. perenne* display homology to an ubiquitin-specific protease (50% in 70 amino acids for *L. perenne*). In the CLO-DvP population TC116908, cosegregates with marker CDB187-305. CDB187-305 was derived from a TDF displaying homology to a protein involved in intracellular transport. Future research in ryegrass should concentrate on the analysis of these interesting candidate genes.

8.2.3 Identification of novel L. perenne genes involved in the SI-response using techniques for the analysis of differential gene expression

Two techniques for the study of differential gene expression, cDNA-AFLP and cDNA microarray hybridisations, were combined in this investigation. The genome-wide expression analysis carried out by cDNA-AFLP allowed us to identify 479 TDFs as putatively related to SI in *L. perenne.* Our results demonstrate once more that cDNA-AFLP is a powerful technique to screen for genes controlling specific biological functions when no prior sequence information is available. The technique becomes very powerful when cDNA-AFLP expression profiles are quantified using dedicated software (AFLPQuantar-Pro, Keygene, NL). This allowed us to analyse the quantitative data using clustering methods. The two computation methods tested (AQBC and average linkage) clustered the TDFs in similar groups of co-regulated genes. The results of these analyses were used to select a subset of 180 differentially expressed cDNA-AFLP fragments for further characterisation.

The reamplification products of the selected fragments rendered finally a total of 259 different gene sequences. Comparison to the database of NCBI allowed us to classify the genes into 11 functional categories. Several of them included genes involved in general cellular processes. Homology was found also to classes of proteins known to be involved in fertilization and SI-processes in other plant families. From the results, it is possible to make a selection of interesting TDFs to work further with:

(1) CDB011_255B and CDB098_220A represent P-type ATPases, a superfamily of cation transport enzymes, including Ca^{2+}-transporting proteins
(2) actins (CDB015_200, CDB022_230I, CDB022_230L and CDB40_380)
(3) CDB016_320A and CDB016_320B, which represent GTP-binding proteins
(4) ubiquitin-related proteins (CDB022_310B, CDB043_200D and CDB211_300)
(5) the TDF homologue (CDB117_300) of an allene oxide cyclase gene

This choice is justified by several arguments. In the first place, calcium gradients and actin are important in pollen tube growth and SI triggers a fast and rapid rearrangement of the actin cytoskeleton of pollen tubes in *Papaver.* Calcium and GTP-binding proteins are also involved in signalling cascades important in different biological processes, probably also in the SI-response. Ubiquitin-related proteins represent also interesting objects for future research as the SI-response in several gametophytic and sporophytic systems involves ubiquitin related proteins and presumably protein degradation. Although the specific molecular mechanism is still unknown, it is currently accepted that proteins necessary for pollen tube growth are probably degraded in incompatible pollen tubes. Furthermore, an STS derived fragment in rye which cosegregates with Z, displays homology to an ubiquitin-specific protease gene. So it is beyond any doubt that the involvement of these TDFs in the SI-response in *L. perenne* constitutes an interesting target for future investigations. With regards to the choice of CDB117-300, allene oxide cyclases are important in jasmonic acid synthesis induced by mechanical wounding and as a defense against pathogens. Some of the processes activated in plants during fertilization and SI could be common to those triggered by pathogen attack. Although rather speculative at this stage, this hypothesis is very interesting if we take into account that

the growth of the pollen tube through the transmitting tract of the style is a process with clear similarities to the growth of fungal pathogens in plants. In support of this hypothesis, SRC, the pollen determinant of SI specificity in crucifers, is similar in structure to defensins, the molecules of innate immunity that present the first line of defence to microbial challenge in plants and animals (Nasrallah 2005).

Still, the main disadvantage of the cDNA-AFLP technique is the large amount of differentially expressed genes identified. Although the rate of false positives produced by mis-priming is probably low, false-positives are generated by co-amplifying unrelated sequences that may either underlie the targeted band or co-migrate very closely with the chosen fragment. This was certainly the case in this study, as for the 180 cDNA-AFLP bands which were excised from the gels, 34% corresponded to more than one gene sequence. Therefore, a detailed verification of the expression profile of each selected TDF is essential. This could be achieved by RT-PCR analyses, but in this study we used the GRASP microarray, which contains also the TDFs selected by cDNA-AFLP for confirmation of the expression profiles. In addition, as this microarray contains also cDNA-clones of genes involved in different biological processes, it offered us the possibility to eventually identify other SI-related genes. A 'dyeswap' experiment was performed with three contrasts: SP1h/SP8h, SP8h/CP8h and CP1h/CP8h. After data analysis, a set of 371 differentially expressed genes putatively involved in the SI-response, was selected. Surprisingly, the results from both techniques were contradictory. Therefore, Real-time RT-PCR reactions were performed for a subset of genes identified as differentially expressed by both techniques. Although the Real-time RT-PCR protocol followed did not allow a detailed analysis of differential gene expression, the qualitative data obtained demonstrated consistency between the results of the Real-time RT-PCR and those of the cDNA-AFLP experiment. Low agreement was therefore found with the microarray data. From these results we concluded that (some of) the samples used in the microarray might have been switched. In which way and for which samples this happened is impossible to find out and further research is still needed to confirm this. Either new microarray hybridisations or Real-time RT-PCR analyses will be used in the near future.

8.2.4 Genome regions involved in SI in L. perenne

Although the genes which control specific biological processes are not necessarily clustered in specific genomic regions, it is well known that S-loci in Brassicaceae and Solanaceae are located in chromosome positions where recombination is suppressed and which contain several genes interspersed with repetitive sequences and non-functional ORFs (Fukai et al. 2003, Wang et al. 2004). These regions contain not only the genes which code the components which determine the SI-specificity, but also other genes involved in SI and fertilization and genes with unrelated functions. Therefore, we determined the genome position of the genes identified using the homology-based search and the cDNA-AFLP. Sequence information was used to develop polymorphic markers which were then mapped together with the available AFLP, RFLP, STS and SSR markers in two mapping populations. The constructed CLO-DvP consensus map consisted of 336 loci (121 AFLP loci, 36 RFLP loci, 97 SSR loci, 46 STS markers in addition to the 36 EST markers developed during this research) distributed over the seven linkage groups and spanned 720 cM. The ILGI map, constructed using less stringent conditions (REC=0.499 and LOD=0.01), consisted of 292 loci (106 AFLPs, 117 RFLPs, 29 SSRs, 3 isozyme markers, 13 STS markers and 24 ESTs developed in this research) and spanned 852 cM. The alignment between the two maps was straightforward and based on common RFLPs, SSRs and STSs.

As the gene sequences from which the markers were derived are putatively involved in SI, the markers developed in this study identified genomic regions controlling SI in *L. perenne*. Given the high level of synteny among the grasses, information about genes involved in the SI-response, generated in *L. perenne* should be easily transferred to other members of the Poaceae. Similarly, in this study the information generated in rye could easily be extrapolated to *L. perenne* and the STS markers originally generated in rye mapped to the predicted LG in *L. perenne*, in the neighborhood of the Z-locus.

As expected, some of the mapped genes clustered around the S- and Z-loci in LG1 and LG2, respectively. CDB100-340Bb and CDB202-280F are located 4 cM and 3 cM away from S on LG1 respectively. In LG2, where Z is located, several EST markers and the STS markers published by Hackauf and Wehling (2005) mapped. However, only CDB187-305d mapped in the neighbourhood of Z and cosegregated with TC116908, the STS marker

which cosegregates with Z in rye. Other EST markers developed in this study mapped across the different linkage groups with a higher concentration on LG3 and LG4. Also in this study a strong association was identified between LG3 and LG1 which implicates the presence of genes on LG3 which may interact with SI-genes on LG1. The TDF derived marker CDB114-370 maps in this genomic region. Furthermore, the TDF derived marker CDB036-410 mapped in the neighbourhood of the self fertility (-T) locus identified by Thorogood et al. (2004) and marker CDB248-270 mapped in region of LG7 where a QTL for heading date was identified in a previous study (Armstead et al. 2004, Dr. D. Thorogood, personal communication).

8.2.5 Concluding remarks

Summarizing, the combination of several approaches allowed us to identify a set of genes putatively involved in the signalling cascade triggered by the SI-response, leading to the inhibition of the pollen tube growth. Combining information from the different sources, a selection has been made of the genes which should constitute the focus of future research, as presented in the previous sections. In Table 8.1 these genes are listed and all the information which is currently available for them has been summarised. In the next section perspectives for future research on these candidate genes are discussed and other research topics in the field of SI in ryegrasses are presented.

Table 8.1 Summary of candidate genes selected in this study for future research.

Candidate gene	Type of evidence	Homologue (Accession No.)	Degree of identity (aa)	Remarks
CDB011_255B	Function	P-type ATPase (Q94IN2)	94% in 71	
CDB015_200	Function Expression	Actin (Q947D0)	74% in 51	Down regulated during self-pollination and absent in CP pistils
CDB016_320A	Function	GTP-binding protein (Q6K2P6)	97% in 42	
CDB016_320B	Function	GTP-binding protein (Q6K2P6)	97% in 42	
CDB022_230I	Function Expression	Actin2 (P17298)	88% in 17	Up regulation till SP4h, then down regulation till SP24h and no expression in the CP pistils

Table 8.1 Continued

Candidate gene	Type of evidence	Homologue (Accession No.)	Degree of identity (aa)	Remarks
CDB022_230L	Function Expression	Actin2 (P17298)	88% in 17	Up regulation till SP4h, then down regulation till SP24h and no expression in the CP pistils
CDB022_310B	Function Expression	Ubiquitin protein ligase (O55176)	35% in 67	Down regulated during self-pollination and absent in CP pistils
CDB036_410	Map position	Unknown function		LG5, linked to SF locus (T)
CDB040_380	Function	Actin (Q947D0)	57% in 67	
CDB043_200D	Function	Ubiquitin-like protein (Q8L8S0)	86% in 22	Up regulation in the CP and SP pistils
CDB098_220A	Function Expression	ATPase2 (Q7RNG4)	32% in 59	Down regulated during self-pollination and absent in CP pistils
CDB100_340Bb	Map position	ABC transporter permease (QWDD5)	33% in 72	Linked to S
CDB114_370	Map position	Unknown function		LG3, interacts with S
CDB117_300	Function	Allene oxide cyclase (Q8L6H4)	97% in 46	
CDB187_305d	Map position	Cellular retinaldehyde-binding/triple function protein (Q5TKJ2)	50% in 77	Linked to Z
CDB202_280F	Map position	Unknown function		Linked to S
CDB211_300	Function Expression	F6I1.6 protein (involved in the ubiquitin cycle) (Q9FZ52)	41% in 36	Down regulated in the self-pollinated pistils and up regulated in the CP24h
CDB248_270	Map position	Unknown function		LG7, linked to heading date QTL
SRK	Function	S-locus receptor-like kinase (Q8LLI4)	83% in 149	
TC116908	Map position Function	Ubiquitin specific protease-like protein (Q9LEW0)	50% in 70	Linked to Z
Thioredoxin	Map position Function	Thioredoxin-like protein (AF159389)	85% in 705 bp	Linked to S

8.3 Future perspectives

During this PhD thesis we were able to identify different candidate genes putatively involved in the signaling cascade triggered by the SI-response of ryegrass but not specifically the S- and Z-genes, the two key components of the recognition step between pistil and pollen. To identify these genes an alternative strategy was necessary, and this is the objective of research which is currently going on in our laboratory. In this case we are using the cDNA-AFLP technique to compare pistil bulks differing in SI-genotypes. The ILGI population, for which the SI-genotypes of the offspring plants have been determined is being used. As the population is formed from a cross between a heterozygote plant (genotype: $S_{23}Z_{23}$) and a double haploid ($S_{11}Z_{11}$), only four different SI-genotypes are present in the offspring, namely, $S_{12}Z_{12}$, $S_{12}Z_{13}$, $S_{13}Z_{12}$ and $S_{13}Z_{13}$. Until now, different putative S-allele and Z-allele specific fragments have been identified. Further, we will use these fragments to perform a homology-based search, for annotating a putative function to and to convert them to polymorphic markers which will be mapped in the two different populations already used during this PhD to check if they cosegregate with S or Z. Given the set-up of the experiment and the objective, only those TDFs mapping in the neighborhood of S and Z will be selected for further characterization and analysis. These candidates could eventually be used to screen a yeast two-hybrid pistil cDNA library to look for proteins that interact with the S and/or Z and to unravel the signal transduction pathway triggered by the SI-response.

But even if the S- and Z-genes are not identified, another possibility is to use the information derived from this PhD thesis and the information obtained in the research described above to develop a map-based cloning strategy to identify the S- and Z-genes. In the first place we could saturate the regions around S and Z on LG1 and LG2, respectively with markers. Those molecular markers can then serve as a starting point for chromosome walking or jumping to the genes of interest.

The candidate genes listed in Table 8.1 should be investigated further. First of all, their expression profiles need to be confirmed. As mentioned above, Real-time RT-PCR approaches could be used. Those genes for which interesting differential expression behavior is confirmed, can then be used in reverse genetics approaches. In this case, mutations that disrupt the sequence and thereby their function can be used (Peters et al. 2003). Two insertional mutagenesis approaches for gene disruption have been used in a

restricted number of model species: Transferred DNA (T-DNA) (Krysan et al. 1999) and transposon tagging (Spuelman et al. 1999). Although these are powerful tools for the generation of novel mutants, they present some limitations, including the impossibility to study the function of duplicated genes and the multiple insertional nature of these approaches which frequently leads to concomitant disruption of several genes (Benedito et al. 2004). Alternative reverse approaches include antisense, cosupression and RNA interference strategies, through which genes are silenced in a specific and efficient manner (Wang and Wagner 2003). These approaches could eventually be used to determine the role of the different TDFs in SI.

Existing evidence suggests that the self-incompatible grasses share a common mechanism for SI. Therefore, the genes identified as being involved in the SI-response of *L. perenne* can contribute to a better understanding of the SI-response in other species of the Poaceae. Finally, it is of no doubt that a better understanding of the recognition and rejection processes between stigma and pollen would be of relevance for the development of improved breeding strategies. Ryegrass cultivars are synthetic varieties based on a small amount of selected clones, sometimes selected out of one family, which are polycrossed. The presence of an SI-mechanism in these species prevents the efficient production of inbred lines and hybrids. However, the production of hybrids would have several advantages, namely exploiting the heterosis effect. The temporarily breakdown of the SI-mechanism can be useful to develop hybrids. Furthermore, a better understanding of the genes involved in the SI-response can help to develop a PCR-based method to screen the plants used in a polycross for their SI-genotypes to ensure that only cross-compatible plants are used. If several plants with an identical SI-genotype are selected to form a polycross, the seed production will be very low. Also, non-random fertilization will lead to a non-panmictic population with inherent uniformity and stability problems.

References

Aharoni A., Keizer L.C.P., Bouwmeester H.J., Sun Z., Alvareze-Huerta M., Verhoeven H.A., Blaas J., van Houwelingen A.M.M.L., De Vos R.C.H., van der Voet H., Jansen R.C., Guis M., Mol J., Davis R.W., Schena M., van Tunen A. and O'Connell A.P. (2000). Identification of the *SAAT* Gene Involved in Strawberry Flavor Biogenesis by Use of DNA Microarrays. The Plant Cell 12: 647-661

Aharoni A. and Vorst O. (2002). DNA microarrays for functional plant genomics. Plant Molecular Biology 48: 99-118

Albertini A., Marconi G., Barcaccia G., Raggi L. and Falcinelli M. (2004). Isolation of candidate genes for apomixis in *Poa pratensis* L. Plant Molecular Biology 56: 879-894

Alm V., Fan C., Busso C.S., Devos K.M., Vollan K., Grieg Z. and Rognli O.A. (2003). A linkage map of meadow fescue (*Festuca pratensis*) and comparative mapping with other Poaceae species. Theoretical and Applied Genetics 108: 25-40

Altpeter F., Fang Y.-D., Xu J. and Ma X. (2003). Comparison of Transgene Expression Stability after *Agrobacteriuim*-mediated or Biolistic Gene Transfer into Perennial Ryegrass (*Lolium perenne* L.). Molecular Breeding of forage and turf. In: Hopkins A, Wang Z-Y, Mian R, Sledge M and Berker RE (Eds): Proceedings of the 3rd International Symposium, molecular breeding of forage and Turf, Dallas, Texas, an Ardmore, Oklahoma, USA, May 18-22, 2003, pp. 255-260

Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W. and Lipman D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein search programs. Nucleic Acids Research 25 (17): 3389-3402

Andersen J.R. and Lübberstedt T. (2003). Functional markers in plants. Trends in Plant Science: 8(11): 554-60

Armstead I.P., Turner L.B., Farrell M., Skot L., Gomez P., Montoya T., Donnison I.S., King I.P. and Humphreys M.O. (2004). Synteny between a major heading-date QTL in perennial ryegrass (*Lolium perenne* L.) and the Hd3 heading-date locus in rice. Theoretical and Applied Genetics 108: 822-828

Armstead I.P., Turner L.B., King I.P., Cairns A.J. and Humphreys M.O. (2002). Comparison and integration of genetic maps generated from F_2 and BC_1-type mapping populations in perennial ryegrass. Plant Breeding 121: 501-507

Bachem C.W.B., Horvath B., Trindade L., Claassens M., Davelaar E., Jordi W. and Visser R.G.F. (2001). A potato tuber-expressed mRNA with

homology to steroid dehydrogenases affects gibberllin levels and plant development. The Plant Journal 25(6): 595-604

Bachem C.W.B., van der Hoeven R.S., de Bruijn S.M., Vreugdenhil D., Zabeau M. and Visser R.G.F. (1996). Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. The Plant journal 9: 745-753

Barnabas B. and Fridvalsky L. (1984). Adhesion and germination of differentially treated maize pollen grains on stigma. Acta Botanica Hungary 30: 329-332

Baumann U., Juttner J., Bian X. and Langridge P. (2000). Self-incompatibility in the grasses. Annals of Botany 85 (suppl A): 203-209

Beddows A.R. (1931). Seed setting and flowering in various grasses. University College of Wales Series H. 12: 5-99

Beemster G.T., De Veylder L., Vercruysse S., West G., Rombaut D., Van Hummelen P., Galichet A., Gruissem W., Inze D. and Vuylsteke M. (2005). Genome-wide analysis of gene expression profiles associated with cell cycle transitions in growing organs of *Arabidopsis*. Plant Physiology 138 (2): 734-743

Benedito V.A., Visser P.B., Angenent G.C. and Krens F.A. (2004). The potential of virus-induced gene silencing for speeding up functional characterization of plant genes. Genetics and Molecular Research 3(3): 323-341

Bert P.F., Charmet G., Sourdille P., Hayward M.D. and Balfourier F. (1999). A high-density molecular map for ryegrass (*Lolium perenne*) using AFLP markers. Theoretical and Applied Genetics 99: 445-452

Bian X.-Y., Friedrich A., Bai J.-R., Baumann U., Hayman D.L., Barker S.J. and Langridge P. (2004). High-resolution mapping of the S and Z loci of *Phalaris coerulescens*. Genome 47: 918-930

Bray E.A. (2004). Genes commonly regulated by water-deficit stress in *Arabidopsis thaliana*. Journal of Experimental Botany 55(407):2331-41

Breyne P., Dreesen R., Cannoot B., Rombaut D., Vandepoele K., Rombauts S., Vanderhaeghen R., Inzé D. and Zabeau M. (2003). Quantitative cDNA-AFLP analysis for genome-wide expression studies. Molecular Genetics and Genomics 269: 173-179

Breyne P., Dreesen R., Vandepoele K., De Veylder L., Van Breusegem F., Callewaert L., Rombauts S., Raes B., Cannoot B., Engler G., Inzé D. and Zabeau M. (2002). Transcriptome analysis during cell division in plants. Proceedings of the National Academy of Sciences 12(99): 14825-14830

Breyne P. and Zabeau M. (2001). Genome-wide expression analysis of plant cell cycle modulated genes. Current Opinion in Plant Biology 4: 136-142

Broothaerts W., Janssens G.A., Proost P. and Broekaert W.F. (1995). CDNA cloning and molecular analysis of two self-incompatibility alleles from apple. Plant Molecular Biology 27: 499-511

Brugmans B., Fernandez del Carmen A., Bachem C.W.B., van Os H., van Eck H.J. and Visser R.G.F. (2002). A novel method for the construction of genome wide transcriptome maps. The Plant Journal 21(2): 211-222

Buckhout T.J. and Thimm O. (2003). Insights into metabolism obtained from microarray analysis. Current Opinion in Plant Biology 6: 288-296

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

Cai G., Casino C.D. and Cresti M. (2000). Cytoskeletal basis of organelle trafficking in the Angiosperm pollen tube. Annals of Botany 85 (suppl A): 69-77

Cai G., Moscatelli A. and Cresti M. (1997). Cytoskeletal organization and pollen tube growth. Trends in Plant Science 2 (3): 86-91

Cato S.A., Gardner R.C., Kent J. and Richardson T.E. (2001). A rapid PCR-based method for genetically mapping ESTs. Theoretical and Applied Genetics 102: 296-306

Cheung A.Y., Wang H. and Wu H. (1995). A floral transmitting tissue-specific glycoprotein attracts pollen tubes and stimulate their growth. Cell 82: 383-393

Chomczynski P. and Sacchi N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Analytical Biochemistry. 162: 156-159

Clark G.B., Thompson G. and Roux S.J. (2001). Signal transduction mechanisms in plants: an overview. Plant Molecular Biology 80(2): 170-177

Cogan N.O., Smith K.F., Yamada T., Francki M.G., Vecchies A.C., Jones E.S., Spangenberg G.C. and Forster J.W. (2005). QTL analysis and comparative genomics of herbage quality traits in perennial ryegrass (*Lolium perenne* L.). Theoretical and Applied Genetics 110(2): 364-80

Connor H.E. (1979). Self-incompatibility in diploid *Lolium perenne* L. Heredity 43: 95-106

Cooper B. (2001). Collateral gene expression changes induced by distinct plant viruses during hypersensitive resistance reaction in *Chenopodium amaranticolor*. The Plant Journal 26(3): 339-349

Cornish M.A., Hayward M.D. and Lawrence M.J. (1979). Self-incompatibility in ryegrass. 1. Genetic control in diploid *Lolium perenne* L. Heredity 43: 95-106

Creelman R.A., Bell E. and Mullet J.E. (1992). Involvement of a Lipoxygenase-Like Enzyme in Abscisic Acid Biosynthesis. Plant Physiology 99(3): 1258–1260

Creelman R.A. and Mullet J.E. (1995). Jasmonic Acid distribution and action in plants: regulation during development and response to biotic and abiotic stress. Proceedings of the National Academy of Sciences 92: 4114-4119

Cruz-Garcia F., Hancock C.N. and McClure B. (2003). S-RNase complexes and pollen rejection. Journal of Experimental Botany 54 (380): 123-130

De Graaf B.H.J., Knuiman B.A., Derksen J. and Mariani C. (2003). Characterization and localization of the transmitting tissue-specific PELPIII proteins of *Nicotiana tabacum*. Journal of Experimental Botany 54 (380): 55-63

Delorme V., Giranton J.-L., Hatzfeld Y., Friry A., Haizmann P., Ariza M.J., Dumas C., Gaude T. and Cock J.M. (1995). Characterization of the S locus genes, SLG and SRK, of the *Brassica* S_3 haplotype: identification of a membrane-localized protein encoded by the S locus receptor kinase gene. The Plant Journal 7(3): 429-440

De Nettancourt D. (1977). Incompatibility in angiosperms. Vol. 3 of the Series: “Monographs on Theoretical and Applied Genetics”, Springer.

De Nettancourt D. (2001), Incompatibility and incongruity in wild and cultivated plants. Springer-Verslag Berlin, Heidelberg, New York, 322pp

De Paepe A., Vuylsteke M., Van Hummelen P., Zabeau M. and Van Der Straeten D. (2004). Transcriptional profiling by cDNA-AFLP and microarray analysis reveals novel insights into the early response to ethylene in *Arabidopsis*. The Plant Journal 39: 537-559

De Smet F., Mathys J., Marchal K., Thijs G., De Moor B. and Moreau Y. (2002). Adaptive quality-based clustering of gene expression profiles. Bioinformatics 18(5): 735-746

Dixit R., Nasrallah M.E. and Nasrallah J.B. (2000). Post-transcriptional maturation of the S-receptor kinase of *Brassica* correlates with co-expression of the S-locus glycoprotein in the stigmas of two *Brassica*

strains and in transgenic tobacco plants. Plant Physiology 124: 297-311

Dixit R., Rizzo C., Narallah M. and Nasrallah J.B. (2001). The *Brassica* MIP-MOD gene encodes a functional water channel that is expressed in the stigma epidermis. Plant Molecular Biology 45: 51-62

Dowd P.E., McCubbin A.G., Wang X., Verica J.A., Tsukamoto T., Ando T. and Kao T.-h. (2000). Use of *Petunia inflata* as a model for the study of Solanaceous type self-incompatibility. Annals of Botany 85 (suppl A): 87-93

Dumas C. and Gaude T. (1982). Sécrétions et biologie florale. II. Leur rôle dans l'adhésion et la reconnaissance pollen-stigmate. Données récentes, hypothèses et notion d'immunité végétale. Bull. Soc. Bot. France 129: 89-104

Durrant W.E., Rowland O., Piedras P., Hammond-Kosack K.E. and Jones J.D.G. (2000). CDNA-AFLP reveals a striking overlap in Race-specific resistance and wound response gene expression profiles. The Plant Cell 12: 963-977

Eckey C., Korell M., Leib K., Biedenkopf D., Jansen C., Langen G. and Kogel K.-H. (2004). Identification of powdery mildew-induced barley genes by cDNA-AFLP: functional assessment of an early expressed MAP kinase. Plant Molecular Biology 55: 1-15

Elleman C.J., Franklin-Tong V.E. and Dickinson H.G. (1992). Pollination in species with dry stigmas: the nature of the early stigmatic response and the pathway taken by pollen tubes. New Phytologist 121: 413-424

Endo M., Tsuchiya T., Saito H., Matsubara H., Hakozaki H., Masuko H., Kamada M., Higahitani A., Takahashi H., Fukuda H., Demura T. and Watanabe M. (2004). Identification and molecular characterization of novel anther-specific genes in *Oryza sativa* L. by using cDNA-microarray. Genes and Genetics Systems 79: 213-226

Faville M.J., Vecchies A.C., Schreiber M., Drayton M.C., Hughes L.J., Jones E.S., Gythridge K.M., Smith K.F., Sawbridge T., Spangenberg G.C., Bryan G.T. and Forster J.W. (2004). Functionally associated molecular genetic marker map construction in perennial ryegrass (*Lolium perenne* L.). Theoretical and Applied Genetics 110 (1): 12-32

Fearon C.H., Cornish M.A., Hayward M.D. and Lawrence M.J. (1994). Self-incompatibility in ryegrass. X. Number and frequency of alleles in a natural population of Lolium perenne L. Heredity 73: 254-261

Forster J.W., Jones E.S., Kölliker R., Drayton M.C., Dupal M.P., Guthridge K.M. and Smith K.F. (2001). Application of DNA profiling to an outbreeding forage species. In: Henry R.J. (ed.). Plant genotyping: the DNA fingerprinting of plants, Wallingford, UK, pp: 299-320

Franklin-Tong V.E. (1999). Signaling and the modulation of pollen tube growth. The Plant Cell 11: 727-738

Franklin-Tong V.E., Atwal K.K., Howell E.C., Lawrence M.J. and Franklin F.C.H. (1991). Self-incompatibility in *Papaver rhoeas*: there is no evidence for the involvement of stigmatic ribonuclease activity. Plant Cell Environmental 14: 423-429

Franklin-Tong V.E. and Franklin F.C.H. (2003). The different mechanisms of gametophytic self-incompatibility. Philosophical Transactions: Biological Sciences 358: 1025-1032

Franklin-Tong V.E., Holdaway-Clarke T.L., Straatman K.R., Kunkel J.G. and Hepler P.K. (2002). Involvement of extracellular calcium influx in the self-incompatibility response of *Papaver rhoeas*. The Plant Journal 29 (3): 333-345

Franklin-Tong V.E., Ride J.P. and Franklin F.C.H. (1995). Recombinant stigmatic self-incompatibility (S-) protein elicits a Ca^{2+} transient in pollen of *Papaver rhoeas*. Plant Journal 8: 299-307

Franklin-Tong V.E., Ride J.P., Read N.D., Trewawas A.J. and Franklin F.C.H. (1993). The self-incompatibility response in *Papaver rhoeas* is mediated by cytosolic free calcium. Plant Journal 4:163-177

Fukai E, Fujimoto and Nishio T. (2003). Genomic organization of the S core region and the S flanking regions of a class-II S haplotype in *Brassica rapa*. Molecular Gene Genomics 269: 361-369

Fuong F.T., Voylokov A.V. and Smirnov V.G. (1993). Genetic studies of self-fertility in rye (*Secale cereale* L.) 2. The search for isozyme marker genes linked to self-incompatibility loci. Theoretical and Applied Genetics 83: 619-623

Fusaro G., Dsgupta P., Rastogi S., Joshi B. and Chellappan S. (2003). Prohibitin induces the transcriptional activity of p53 and is exported from the nucleus upon apoptotic signalling. The Journal of Biological Chemistry 278 (48): 47853-47861

Gertz A. and Wricke G. (1989). Linkage between the incompatibility locus Z and a beta-glucosidase locus in rye. Plant Breeding 102: 255-259

Giddings G.D., Sackville Hamilton N.R. and Hayward M.D. (1997a). The release of genetically modified grasses. Part 1: pollen dispersal to traps in *Lolium perenne*. Theoretical and Applied Genetics 94: 1000-1006

Giddings G.D., Sackville Hamilton N.R. and Hayward M.D. (1997b). The release of genetically modified grasses. Part 2: the influence of wind direction on pollen dispersal. Theoretical and Applied Genetics 94: 1007-1014

Giddings G.D. (2000). Moddeling the spread of pollen from Lolium perenne. The implications for the release of wind-pollinated transgenics. Theoretical and Applied Genetics 100: 971-974

Goda H., Sawa S., Asami T., Fujioka S., Shimada Y. and Yoshida S (2004). Comprehensive comparison of auxin-regulated and bassinosteroid-regulated genes in *Arabidopsis*. Plant Physiology 134: 1555-1573

Goupil P., Mahamoud Y.S., Poulain J., Windels D., Crete P., Huss B.and Rambour S. (2003). cDNA-AFLP display for the isolation of differentially expressed genes during chicory root development. Journal of Plant Physiology 160 (3): 303-309

Grasgids (2000). Grassen voor sport, gazon, recreatie, berm, dijk en golfgreen. Nederlandse vereniging voor Zaaizaad en Plantgoed, 46pp

Grattapaglia D. (2004). Integrating genomics into *Eucalyptus* breeding. Genetics and Molecular Research 3 (3): 369-379

Grattapaglia D. and Sederoff R. (1994). Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137: 1121-1137

Gray J.W. and Collins C. (2000). Genome changes and gene expression in human solid tumors. Carcinogenesis 21: 443-452

Hackauf B. and Wehling P. (2005). Approaching the self-incompatibility locus Z in rye (*Secale cereale* L.) via comparitive genetics. Theoretical and Applied Genetics 110 (5): 832-845

Hannaway D., Fransen Z., Cropper J., Teel M., Chaney M., Griggs T., Halse R., Hart P., Hansen D., Klinger R. and Lane W. (1999a). Annual ryegrass (*L. multiflorum*), http://eesc.orst.edu/agcomwebfile/edmat/html/PNW/PNW501/PNW501.html

Hannaway D., Fransen Z., Cropper J., Teel M., Chaney M., Griggs T., Halse R., Hart P., Hansen D., Klinger R. and Lane W. (1999b). Perennial ryegrass (*L. perenne*), http://eesc.orst.edu/agcomwebfile/edmat/html/PNW/PNW503/PNW503.html

Hayman D.L. (1956). The genetic control of self-incompatibility in *Phalaris coerulescens*. Australian Journal of Biological Sciences 9: 321-331

Hayman D.L. and Richter J. (1992). Mutations affecting self-incompatibility in *Phalaris coerulescens* Desf. (Poaceae). Heredity 68: 459-503

Hayward M.D., McAdam N.J., Jones J.G., Evans C., Evans M., Forster J.W., Ustin A., Hossain K.G., Quader B., Stammers M. and Will J.A.K. (1994). Genetic markers and the selection of quantitative traits in forage grasses. Euphytica 77: 269-275

Hayward M.D., Forster J.W., Jones J.G., Dolstra O., Evans C., McAdam N.J., Hossain K.G., Stammers M., Will J., Humphreys M.O. and Evans G.M. (1998). Genetic anamysis of *Lolium*. I. Identification of linkage groups and the establishment of a genetic map. Plant Breeding 117: 451-455

Hearn M.J., Franklin F.C.H. and Ride J.P. (1996). Identification of a membrane glycoprotein in pollen of *Papaver rhoeas* which binds stigmatic self-incompatibility (S-) proteins. The Plant Journal 9: 467-475

Heizmann P., Luu D.T. and Dumas C. (2000). Pollen-stigma adhesion in the *Brassicaceae*. Annals of Botany 85 (suppl A): 23-27

Henikoff S. and Comai L. (2003). Single-nucleotide mutations for plant functional genomics. Annual Review of Plant Biology 54: 375-401

Hepler P.K., Vidali L. and Cheung A.Y. (2001). Polarized cell growth in higher plants. Annual Review of Cell and Developmental Biology 17: 159-187

Heslop-Harrison Y. (1982). Pollen stigma interaction and cross-incompatibility in the grasses. Science 161: 230-237

Higgins ABC Transporters: From microorganisms to man. Annual Review of Cell Biology 8: 67-113

Hiscock S.J. and Mclnnis S.M. (2003). Pollen recognition and rejection during the sporophytic self-incompatibility response: *Brassica* and beyond. Trends in Plant Science 8 (12): 606-613

Hogenboom N.G. (1973). A model for incongruity in intimate partner relationships. Euphytica 22:219-233

Huff D.R. (1997). RAPD characterisation of heterogeneous perennial ryegrass cultivars. Crop science 37: 557-564

Hülskamp M., Kopczak S.D., Horejsi T.F., Kihl B.K. and Pruitt R.E. (1995). Identification of genes required for pollen stigma recognition in *Arabidopsis thaliana*. The Plant Journal 8: 703-715

Humphreys M.O. (2000). Breeding methods for forage and amenity grasses. In: Spangenberg G. (eds.). Molecular breeding of forage crops. Proceedings of the 2nd International Symposium, molecular breeding of forage and turf, Lorne and Hamilton, Victoria, Australia, November 9-24, 2000, pp. 41-50

Humphreys M.W., Humphreys J., Donnison I., King I.P., Thomas H.M., Ghesquière M., Durand J.-L., Rognli O.A., Zwierzykowski Z. and Rapacz M. (2003). Molecular breeding and functional genomics for tolerance to abiotic stress. In: Hopkins A., Wang Z.-Y., Mian R., Sledge M. and Barker R.E. (eds.). Molecular breeding of forage and turf. Proceedings of the 3rd International Symposium, molecular breeding of forage and turf, Dallas, Texas and Ardmore, Oklahoma, U.S.A., May 18-22, 2003, pp. 61-80

Inoue M., Gao Z., Hirata M., Fujimori M. and Cai H. (2004). Construction of a high-density linkage map of Italian ryegrass (*Lolium multiflorum* L.) using restricted fragment length polymorphism, amplified fragment length polymorphism, and telomeric repeat associated sequence markers. Genome 47: 57-65

Ioerger T.R., Gohlke J.R., Xu B. and Kao T.-h. (1991). Primary structural features of the self-incompatibility protein in *Solanaceae*. Sexual Plant Reproduction 4: 81-87

Jack T., Brockman L.L., Meyerowitz E.M. (1992). The homeotic gene APETALA3 of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. Cell 68:683

Jansen J. and van Ooijen J.W. (2003). Joinmap® 3.0, Genetic Linkage mapping. Plant Research International, Wageningen, The Netherlands

Jenczewski E., Ghererdi M., Bonnin I., Prosperi J.M., Olivieri I. and Huguet T. (1997). Insight on segregation distortion in two intraspecific crosses between annual species of Medicago (*Leguminosae*). Theoretical and Applied Genetics 94: 682-691

Jensen L.B., Muylle H., Arens P., Andersen C.H., Holm P.B., Ghesquiere M., Julie B., Lübberstedt T., Nielsen K.K., De Riek J., Roldán-Ruiz I., Roulund N., Taylor C., Vosman B. and Barre P. (2005). Development and mapping of a public reference set of SSR markers in *Lolium perenne* L. Molecular Ecology Notes (submitted)

Johnsen M.A. and Preuss D. (2002). Plotting a course: multiple signals guide pollen tubes to their targets. Developmental Cell 2: 273-281

Jones E.S., Dupal M.P., Kölliker R., Drayton M.C. and Foster J.W. (2001). Development and characterization of simple sequence repeat (SSR) markers for perennial ryegrass (*Lolium perenne* L.). Theoretical and Applied Genetics 102: 405-415

Jones E.S., Mahoney N.L., Hayward M.D., Armstead I.P., Jonas J.G. and Humphreys M.O. (2002a). An enhanced molecular marker-based map of perennial ryegrass (*Lolium perenne* L.) reveals comparative relationships with other *Poaceae* genomes. Genome 45: 282-295

Jones E.S., Dupal MP, Dumsday JL, Hughes LJ and Forster JW (2002b). An SSR-based genetic map for perennial ryegrass (*L. perenne* L.). Theoretical and Applied Genetics 105: 577-584

Jordan N.D., Franklin F.C.H. and Franklin-Tong V.E. (2000). Evidence for DNA fragmentation triggered in the self-incompatibility response in pollen of *Papaver rhoeas*. The Plant Journal 23: 471-479

Jordan N.D., Kakeda K., Conner A., Ride J.P., Franklin-Tong V.E. and Franklin F.C.H. (1999). S-protein mutants indicate a functional role for SBP in the self-incompatibility reaction of *Papaver rhoeas*. The Plant Journal 20: 119-126

Kareiva P. (1993). Transgenic plants on trial. Nature 363: 580-581

Kachroo A., Schopfer C.R., Nasrallah M.E. and Nasrallah J.B. (2001). Allele-specific receptor-ligand interactions in *Brassica* self-incompatibility. Science 293:1824-1826

Kemp B.P., Beeching J.R. and Cooper R.M. (2005). cDNA-AFLP reveals genes differentially expressed during the hypersensitive response of cassava. Molecular Plant Pathology 6(2): 113-123

Kowyama Y., Kakeda K., Kondo K., Imada T. and Hattori T. (1996). A putative receptor protein kinase gene in *Ipomoea trifida*. The Plant Cell Physiology 37: 681-685

Kowyama Y., Kakeda K., Nakano R. and Hattori T. (1995). SLG/SRK-like genes are expressed in the reproductive tissues of *Ipomoea trifida*. Sexual Plant Reproduction 8: 333-338

Krysan P.J., Young J.C. and Sussman M.R. (1999). T-DNA as an insertional mutagen in *Arabidopsis*. Plant Cell 11: 2283-2290

Kubik C., Meyer W.A. and Gaut B.S. (1999). Assessing the abundance and polymorphism of simple sequence repeats in perennial ryegrass. Crop Science 39: 1136-1141

Kubik C., Sawkins M., Meyer W.A. and Gaut B.S. (2001). Genetic diversity in seven perennial ryegrass (*Lolium perenne* L.) cultivars based on SSR markers. Crop Science 41: 1565-1572

Kumar L.S. (1999). DNA markers in plant improvement: an overview. Biotechnology advances 17: 143-182

Kuranandaa B., Huang S. and Kao T.-h. (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

Lai Z., Ma W., Han B., Liang L., Zhang Y., Hong G. and 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

Langridge P., Baumann U. and Juttner J. (1999). Revisiting and revising the self-incompatibility genetics of *Phalaris coerulescens*. The Plant Cell 11: 1826

Laub M.T., McAdams H.H., Feldblyum T., Fraser C.M. and Shapiro L. (2000). Global analysis of the genetic network controlling a bacterial cell cycle. Science 290: 2144-2148

Leach C.R. (1988). Detection and estimation of linkage for a co-dominant structural gene locus linked to a gametophytic self-incompatibility locus. Theoretical and Applied Genetics 75: 882-888

Lee H.-S., Huang S. and Kao T.-h. (1994). S proteins control rejection of incompatible pollen in *Petunia inflata*. Nature 367: 560-563

Lem P. and Lallemand J. (2003). Grass consensus STS markers: an efficient approach for detecting polymorphism in *Lolium*. Theoretical and Applied Genetics 107: 1113-1122

Lewis E.J., Humphreys M.W. and Caton M.F. (1980). Chromosomal location of two isozyme loci using primary trisomics. Theoretical and Applied Genetics 57: 237-239

Li X., Niedl J., Hayman D. and Langridge P. (1994). Cloning of a self-incompatibility gene from the pollen of the grass *Phalaris coerulescens*. The Plant Cell 6: 1923-1932

Li X., Niedl J., Hayman D. and Langridge P. (1995). Thioredoxin activity in the C terminus of *Phalaris* S-proteins. The Plant Journal 8: 133-138

Li X., Peach N., Niedl J., Hayman D. and Langridge P. (1997). Self-incompatibility in the grass: evolutionary relationship of the S-gene from Phalaris coerulescens. Plant Molecular Biology 34: 223-232

Liang P. and Pardee A.B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257: 967-971

Lord E.M. (2003). Adhesion and guidance in compatible pollination. Journal of Experimental Botany 54: 47-54

Lorkowski S. and Cullen P. (2002). Analysing gene expression: A handbook of methods, possibilities and pitfalls. Wiley-VCH

Lu H., Romereo-Severson J. and Bernardo R. (2002). Chromosomal regions associated with segregation distortion in maize. Theoretical and Applied Genetics 105: 622-628

Lübberstedt T., Andreasen B.S., Holm P.B. et al. (2003). Development of Ryegrass Allele-specific (GRASP) Markers for Sustainable Grassland Improvement – A New EU Framework V Project. Czech J. Genet. Plant Breeding 39: 125-128

Lundqvist A. (1954). Studies on self-sterility in rye, *Secale cereale* L. Hereditas 40: 278-294

Lundqvist A. (1961). A rapid method for the analysis of incompatibility in the grasses. Hereditas 47: 705-707

Luo H., Hu Q., Neslon K., Longo C. and Kausch P. (2003). Controlling transgene escape in genetically modified grasses. In: Hopkins A., Wang Z.-Y., Mian R., Sledge M. and Barker R.E. (eds.). Molecular breeding of forage and turf. Proceedings of the 3rd International Symposium, molecular breeding of forage and turf, Dallas, Texas and Ardmore, Oklahoma, U.S.A., May 18-22, 2003, pp. 245-254

Luu D.-T., Qin X., Laublin G., Yang Q., Morse D. and Cappadocia M. (2001). Rejection of S-heteroalllelic pollen by a dual-specific S-RNase in *Solanum chacoense* predicts a multimeric SI pollen component. Genetics 159: 329-335

Maiko I., Zhensheng G., Maroko H., Masahiro F. and Hongwei C. (2004). Construction of a high-density linkage map of Italian ryegrass (*Lolium multiflorum* L.) using restriction fragment length polymorphism, amplified fragment length polymorphism and telomeric repeat associated sequence markers. Genome 47: 57-64

Malho R. and Trewavas A.J. (1996). Localized apical increases of cytosolic free calcium control pollen tube orientation. The Plant cell 8: 1938-1949

Maliepaard C., Jansen J. and van Ooijen J.W. (1997). Linkage analysis in a full-sib family of an outbreeding plant species : overview and consequences for applications. Genetical Research 70: 237-250

Maliepaard C., Alston F.H., Van Arkel G., Brown L.M., Chevreau E., Dunemann F., Evans K.M., Gardiner S., Guilford P., Van Heusden A.W., Janse J., Laurens F., Lynn J.R., Manganaris A.G., Den Nijs A.P.M., Periam N., Rikkerink E., Roche P., Ryder C., Sansavini S., Schmidt H., Tartarini S., Verhaegh J.J., Vrielink-Van Ginkel M. and King G.J. (1998). Aligning male and female linkage maps of apple

(*Malus pumila* Mill.) using multi-allelic markers. Theoretical and Applied Genetics 97: 60-73

Mao C., Keke Y., Yang L., Zheng B., Wu Y., Liu F. and Wu P. (2004). Identification of aluminium-regulated genes by cDNA-AFLP in rice (*Oryza sativa* L.): aluminium-regulated genes for the metabolism of cell wall components. Journal of Experimental Botany 55(394): 137-143

Martin F.W. (1959). Staining and observing pollen tubes in the style by means of fluorescence . Stain Technology 34: 125-128

Mayfield J.A., Fiebig A., Johnstone S.E. and Preuss D. (2001). Gene families from the *Arabidopsis thaliana* pollen coat proteome. Science 292: 2482-2485

McCubbin A.G. and Kao T.-h. (1996). Self-incompatibility and pollen rejection in angiosperms. In: Jain S.M., Sopory S.K. and Veilleux R.E. (eds). *In vitro* haploid production in higher plants. Dordrecht, Kluwer Academic Publishers

McCubbin A.G. and Kao T.-h. (2000). Molecular recognition and response in pollen and pistil interactions. Annual Review of Cell and Developmental Biology 16: 333-364

McCLure B.A., Cruz-Garcia F., Beecher B. and Sulaman W. (2000). Factors affecting inter- and intra-specific pollen rejection in *Nicotiana*. Annals of Botany 85 (suppl A): 113-123

McClure B.A., Mou B., Canevascini S. and Bernatzky R. (1999). A small asparagines-rich protein required for S-allele-specifc pollen rejection in *Nicotiana*. Proceedings of the National Academy of Sciences, USA 96: 13548-13553

Minerd J. and Minerd A. (2005). Whither the ideal RNA amplification kit? The Scientist 19(5): 30-34

Mishima M., Takayama S., Sasaki K.I., Jee J.G., Kojima C., Isogai A. and Shirkawa (2003). Structure of the Male Determinant Factor for *Brassica* Self-incompatibility. Journal of Biological Chemistry 278: 36389-36395

Mullis K.B. and Faloona F.A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Methods in Enzymology. 155: 335-50

Murfett J., Atherton T., Mou B., Gasser C. and McClure B. (1994). S RNase expressed in transgenic *Nicotiana* causes S allele specific pollen rejection. Nature 367: 563-566

Murphy D.J., Hernendez-Pinzon I., Patel K., Hope R.G. and McLauchlan J. (2000). New insights into the mechanisms of lipid-body biogenesis in

plants and other organisms. Biochemical Society Transactions 28: 710-711

Muylle H., Baert J., Van Bockstaele E., Pertijs J. and Roldán-Ruiz I. (2005). Four QTLs determine crown rust (*Puccinia coronata* f. sp. *Lolii*) resistance in a perennial ryegrass (*Lolium perenne*) population. Euphytica (submitted)

Nasrallah J.B. (2005). Recognition and rejection of self in plant self-incompatibility: comparisons to animal histocompatibility. Trend in Immunology 26 (8): 412-4418

Nasrallah J.B., Kao T. -h., Chen C.H., Goldberg M.L. and Nasrallah M.E. (1987). Amino-acid sequence of glycoproteins encoded three alleles of the S-locus of *Brassica oleraceae*. Nature 326: 617-619.

Nasrallah J.B., Kao T.-h., Goldberg M.L. and Nasrallah M.E. (1985). A cDNA clone encoding an S-locus specific glycoprotein from *Brassica oleraceae*. Nature 318: 617-618

Nasrallah J.B., Rundle S. and Nasrallah M.E. (1994). Genetic evidence for the requirement of the *Brassica* S-locus receptor kinase gene in the self-incompatibility response. The Plant Journal 5: 373-384

Nederlandse Vereniging voor Zaaizaad en Plantgoed, Proefstation recreatie Sport Gras (2000).Grasgids

Newbigin E. and Uyenoyama M.K. (2005). The evolutionay dynamics of self-incompatibility systems. Trends in Genetics 21: 500-505

Parsell D.A. and Lindquist S. (1993). The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. Annual Review of Genetics 27: 437-96.

Paterson A.H. (1996). Mapping genes responsible for differences in phenotype. In: Paterson, A.H. (ed.); Genome mapping in plants: 41-54

Pasquer F., Isidore E., Zarn J. and Keller B. (2005). Specific patterns of changes in wheat gene expression after treatment with three antifungal compounds. Plant Molecular Biology 57(5): 693-707

Patnaik D. and Khurana P. (2001). Germins and germin like proteins: an overview. Indian Journal of Experimental Biology 39(3): 191-200

Peakall R., Gilmore S., Keys W., Morgante M. and Rafalski A. (1998). Cross-species amplification of soybean (Glycine max) simple sequence repeats (SSRs) xithin the genus and other legume genera: implications for the transferability of SSRs in plants. Molecular Biology and Evolution 15: 1275-1287

Peters J.L., Cnudde F. and Gerats T. (2003). Forward genetics and map-based cloning approaches. Trends in Plant Science 8(10): 1360-1385

Pierson W.R. and Lipman D.J. (1988). Improved tools for biological sequence comparison. Proceedings of the National Academy of Sciences (USA) 85, 2444-2448.

Puskás L.G., Zvara A., Hackler L. Jr. and Van Hummelene P. (2002). RNA amplification results in reproducible microarray data with slight ration bias. Biotechniques 32: 1330-1340

Qiao H., Wang H., Zhao L., Zhou J., Huang J., Zhang Y. and Xue Y. (2004). The F-box protein AhSLF-S_2 physically interact with S-RNases that may be inhibited by the Ubiquitin/26S proteosome pathway of protein degradation during compatible pollination in Antirrhium. The Plant Cell (16): 582-595

Quackenbush J. (2002). Microarray data normalization and transformation. Nature Genetics 32(supplement): 496-501

Reheul D. and Ghesquiere A. (1996). Breeding perennial ryegrass with better crown rust resistance. Plant Breeding 115: 465-469

Reijans M., Lascaris R., Groeneger A.O., Wittenberg A., Wesselink E., van Oeveren J., de Wit E., Boorsma A., Voetdijk B., van der Spek H., Grivell L.A. and Simons G. (2003). Quantitative comparison of cDNA-AFLP, microarrays, and GeneChip expression data in *Saccharomyces cerevisiae*. Genomics 82: 606-618

Ride J.P., Davies E.M., Franklin F.C.H. and Marshall D.F. (1999). Analysis of *Arabidopsis* genome sequence reveals a large new gene family in plants. Plant Molecular Biology 39: 927-932

Roldán-Ruiz I., Dendauw J., Van Bockstaele E., Depicker A. and De Loose M. (2000). AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). Molecular Breeding 6: 125-134

Roldán-Ruiz I., van Eeuwijk F.A., Gilliland T.J., Dubreuil P., Dillman C., Lallemand J., De Loose M. and Baril C.P. (2001). A comparative study of molecular and morphological methods of describing relationships between perennial ryegrass (*Lolium perenne* L.) varieties. Theoretical and Applied Genetics 103: 1138-1150

Rudd J.J. and Franklin-Tong V.E. (2003). Signals and targets of the self-incompatibility response in pollen of *Papaver rhoeas*. Journal of Experimental Botany 54 (380): 141-148

Rudd J.J., Franklin F.C.H. and Franklin-Tong V.E. (1997). Ca^{2+}-independent phosphorylation of a 68 kD pollen protein is stimulated by the self-incompatibility response in *Papaver rhoeas*. The Plant Journal 12: 507-514

Rudd J.J., Franklin F.C.H., Lord J.M. and Franklin-Tong V.E. (1996). Increased phosphorylation of a 26-kD pollen protein is induced by

self-incompatibility response in *Papaver rhoeas*. The Plant Cell 8: 713-724

Rudd J.J., Osman K., Franklin F.C.H. and Franklin-Tong V.E. (2003). Activation of a putative Map kinase in pollen is stimulated by the self-incompatibility (SI) response. FEBS Letters 547: 223-227

Saeed A.I., Sharov V., White J., Li J., Liang W., Bhagabati N., Braisted J., Klapa M., Currier T., Thiagarajan M., Sturn A., Snuffin M., Rezantsev A., Popov D., Ryltsov A., Kostukovich E., Borisovsky I., Liu Z., Vinsavich A., Trush V., Quackenbush J. (2003). A free open-source system for microarray data management and analysis. Biotechniques 34(2): 374-378 WWW.tigr.org/software/tm4/

Sakamoto K., Kusaba M. and Nishio T. (1998). Polymorphism of the S-locus glycoprotein gene (SLG) and the S-locus related gene (SLR1) in *Raphanus sativus* L. and self-incompatible ornamental plants in the *Brassicaceae*. Molecular Genetics and Genenomics 258: 397-403

Sambrook J., Fritsch E.F. and Maniatis T. (1989). Molecular cloning: a laboratory manual (2nd edition). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Samuelian S., Kleine M., Ruyter-Spira C.P., Klein-Lankhorst R.M. and Jung C. (2004). Cloning and functional analyses of a gene from sugar beet up-regulated upon cyst nematode infection. Plant Molecular Biology 54: 147-156

Sanger F., Nicklen S. and Coulson A.R. (1977). DNA sequencing with chain terminating inhibitors. Proceedings of the National Academy of Sciences (USA) 74: 5463-5467

Shanker S., Salazar R.W., Taliercio E.W. and Chourey P.S. (1995). Cloning and characterisation of full-length cDNA encoding cell-wall invertase from maize. Plant Physiology 108: 873-874

Schäfer S. and Schmülling T. (2002). The CRK1 receptor-like kinase gene of tobacco is negatively regulated by cytokinin. Plant Molecular Biology 50: 155-166

Scheible W.-R., Morcuende R., Czechowski T., Fritz C., Osuna D., Palacios-Rojas N., Schindelasch D., Thimm O., Udvardi M.K. and Stitt M. (2004). Genome-wide Reprogramming of Primary and Secundary Metabolism, Protein Synthesis, Cellular Growth processes, and the Regulatory Infrastructure of Arabidopsis in response to Nitrogen. Genome Analysis 136: 2483-2499

Schopfer C.R., Nasrallah M.E. and Nasrallah J.B. (1999). The male determinant of self-incompatibility in *Brassica*. Science 286:1697-1700

Shenk P.M., Kazan K., Wilson I., Anderson J.P., Richmond T., Somerville S.C. and Manners J.M. (2000). Coordinated plant defense responses

in *Arabidopsis* revealed by microarray analysis. PNAS 97(21): 11655-11660

Shiba H., Takayama S., Iwano M., Shimosato H., Funato M., Nakagawa T., Suzuki G., Watanabe M., Hinata K. and Isogai A. (2001). An anther expressed gene SP11 determines the pollen S-specificity in the self-incompatibility of *Brassica rapa*. Plant Physiology 125: 2095-2103

Sijacic P., Wang X., Skirpan A.L., Wang Y., Dowd P.E., McCubbin A.G., Huang S. and Kao T.-h. (2004). Identification of the pollen determinant of S-RNase-mediated self-incompatibility. Nature (429): 302-305

Sim S., Chang T., Curley J., Warnke S.E., Barker R.E. and Jung G. (2005). Chromosomal rearrangements differentiating the ryegrass genome from the Triticaceae, oat and rice genomes using common heterologous RFLP probes. Theoretical and Applied Genetics 110: 1011-1019

Sims T.L. and Ordanic M. (2001). Identification of a S-ribonuclease-binding protein in *Petunis hybrida*. Plant Molecular Biology 47: 771-783

Snowdon R.J. and Friedt W. (2004). Molecular markers in *Brassica* oilseed breeding: current status and future possibilities. Plant Breeding 123: 1-8

Snowmann B.N., Geitmann A., Clarke S.R., Staiger C.J., Franklin C.H., Emons A.-MC. and Franklin-Tong V.E. (2000). Signalling and the cytoskeleton on pollen tubes of *Papaver rhoeas*. Annals of Botany 85 (suppl A): 49-57

Snowmann B.N., Kovar D.R., Shevchenko G., Franklin-Tong V.E. and Staiger C.J. (2002). Signal-mediated depolymeriation of actin in pollen during the self-incompatibility response. The Plant Cell 14: 2613-2626

Soeda Y., Konings M.C., Vorst O., van Houwelingen A.M., Stoopen G.M., Maliepaard C.A., Kodde J, Bino R.J, Groot S.P. and van der Geest A.H. (2004). Gene expression programs during *Brassica oleracea* seed maturation, osmopriming, and germination are indicators of progression of the germination process and the stress tolerance level. Plant Physiology 137 (1): 354-368

Sonneveld T., Tobutt K.R., Vaughan S.P. and Robbins T.P. (2005). Loss of pollen-S function in two self-compatible selections of *Prunus avium* is associated with deletion/mutation of an S haplotype-specific F-box gene. The Plant Cell 17: 37-51

Spangenberg G., Kalla R., Lidgett A., Sawbridge T., Ong E.K. and John U. (2001). Breeding forage plants in the genome Era. In: Sprangenberg G (Ed.) Molecular Breeding of Forage Crops. Proceedings of the 2nd International Symposium, molecular breeding of forage crops, Lorne and Hamilton, Victoria, Australia, November 19-24, 2000, pp. 1-39

Speulman E., Metz P.L.J., van Arkel G., Hekkert P.T.L., Stiekema W.J. and Pereira A. (1999). A two-component enhancer-inhibitor transposon mutagenesis system for functional analysis of the *Arabidopsis* genome. The Plant Cell 11: 1853-1866

Staiger C.J. and Franklin-Tong V.E. (2003). The actin cystoskeleton is a target of the self-incompatibility response in *Papaver rhoeas*. Journal of Experimental Botany 54 (380): 103-113

Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E. and Nasrallah J.B. (1991). Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleraceae*. Proceedings of the National Academy of Sciences (USA) 88: 8816-8820.

Stephenson A.J., Doughty J., Elleman C.J., Hiscock S.J. and Dickinsom H.G. (1997). The male determinant of self-incompatiblity in *Brassica oleraceae* is located in the pollen-coating. The Plant Journal 12: 1351-1359

Stone S.L., Anderson E.M., Mullen R.T. and Goring D.R. (2003). ARC1 Is an E3 Ubiquitin Ligase and Promotes the Ubiquitination of Proteins during the Rejection of Self-Incompatible *Brassica* Pollen. The Plant Cell 15: 885-898

Stone S.L., Arnoldo M. and Goring D.R. (1999). A breakdown of *Brassica* self-incompatibility in ARC1 antisense transgenic plants. Science 286: 1729-1731

Strausberg R.L. (2001). The cancer genome anatomy project: new resources for reading the molecular signatures of cancer. Journal of Pathology 195: 31-40

Suárez M.C., Bernal A., Gutiérrez J., Tohme J. and Fregene M. (2000). Developing expressed sequence tags (ESTs) from polymorphic transcript-derived fragments (TDFs) in cassa (*Manihot esculenta Crantz*). Genome 43: 62-67

Suzuski G., Kakizaki T., Takada Y., Shiba H., Takayama S., Isogai A. and Watanabe M. (2003). The S haplotypes lacking SLG in the genome of *Brassica rapa*. The Plant Cell Reports 21 (9): 911-915

Takasaki T., Hatakeyama K., Suzuki G., Watanabe M., Isogai A. and Hinata K. (2000). The S receptor kinase determines self-incompatibility in *Brassica* stigmas. Nature 403: 913-916

Takayama S., Shiba H., Iwano M., Shimosato H., Che F.S., Kai N., Watanabe M., Suzuki G., Hinata K. and Isogai A. (2000). The pollen determinant of self-incompatibility in *Brassica campestris*. Proceedings of the National Academy of Sciences (USA) 97: 1920-1925.

Takayama S., Shimosato H., Shiba H., Funato M., Che F.S., Watanabe M., Iwano M. and Isogai A. (2001). Direct ligand-receptor complex interaction controls *Brassica* self-incompatibility. Nature 413: 534-538

Tatsuta T., Model K. and Langer T. (2005). Formation of membrane-bound ring complexes by prohibitin in mitochondria. Molecular Biology of the Cell 16: 248-259

Tavazoie S., Hughes J.D., Campbell M.J., Cho R.J. and Church G.M. (1999). Systematic determination of genetic network architecture. Nature Genetics 22: 281-285

Taylor C., Madsen K., Borg S., Moller M.G., Boelt B. and Holm P.B. (2001). The development of sequence-tagged sites (STSs) in *L. perenne* L.: the application of primer sets derived from other genera. Theoretical and Applied Genetics 103: 648-658

Thomas S.G. and Franklin-Tong V.E. (2004). Self-incompatibility triggers programmed cell death in *Papaver* pollen. Nature 429: 305-309

Thomas S., Osman K., de Graaf B.H.J., Shevchenko G., Wheeler M., Franklin F.C.H. and Franklin-Tong V.E. (2003). Investigating mechanism involved in the self-incompatibility response in *Papaver rhoeas*. Philosophical Transactions: Biological Sciences 358: 1033-1036

Thorogood D. (1991). Self-fertility in *Lolium*. PhD thesis, University of Wales

Thorogood D., Armstead I.P., Turner L.B., Humphreys M.O. and Hayward M.D. (2004). Identification and mode of action of self-compatibility in *Lolium perenne* L. Heredity 94 (3): 356-363

Thorogood D. and Hayward M.D. (1991). The genetic control of self-incompatibility in an inbred line of *Lolium perenne* L. Heredity 67: 175-181

Thorogood D. and Hayward M.D. (1992). Self-incompatibility in *Lolium temulentum* L.: its genetic control and transfer into *L. perenne* L. and *L. multiflorum* L. Heredity 68: 71-78

Thorogood D., Kaiser W.J., Jones J.G. and Armstead I. (2002). Self-incomaptibility in ryegrass 12. Genotyping and mapping the S and Z loci of *Lolium perenne* L. Heredity 88: 385-390

Tsai D.-S., Lee H.-S., Post L.C., Kreiling K.M. and Kao T.-h. (1992). Sequence of an S-protein of *Lycopersicon peruvianum* and

comparison with other Solanaceous S-proteins. Sexual Plant Reproduction 5: 256-263

Van Bockstaele E. (1998). Breeding and cultivar development in perennial ryegrass (*L. perenne* L.). Biologish Jaarboek Dodoneae 66: 146-156

Van Bockstaele E. (academiejaar 1998-1999). Cursus, 'Toegepaste plantenveredeling'. Faculteit Landbouwkundige en Toegepaste Biologische Wetenschappen, Gent, 159pp

Vandenabeele S., Vanderauwera S., Vuylsteke M., Rombauts S., Langebartels C., Seidlitz H.K., Zabeau M., Van Montagu M., Inzé D. and Van Breusegem F. (2004). Catalase deficiency drastically affects gene expression induced by high light in *Arabidopsis thaliana* The Plant Journal 39: 45-58

Van Ooijen J.W. and Voorrips R.E. (2001). Joinmap® 3.0, software for the calculation of genetic linkage maps. Plant research International, Wageningen, The Netherlands

Velculescu V.E., Zhang L., Vogelstein B. and Kinzler K.W. (1995). Serial analysis of gene expression. Science 270: 484-487

Verica J.A., McCubbin A.G. and Kao T.-h. (1998). Are the hypervariable regions of S-RNases sufficient for allele-specific recognition of pollen? The Plant Cell 10: 314-316

Vos P., Hogers R., Bleeker M., Reijans M., van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M. and Zabeau M. (1995). AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407-4414

Voylokov A.V., Fuong F.T. and Smirnov V.G. (1993). Genetic studies of self-fertility in rye (*Secale cereale* L.) 1. The identification of genotypes of self-fertile lines for the S_f alleles of self-incompatiblity genes. Theoretical and Applied Genetics 83:616-618

Wang Y., Tsukamoto T., Yi K.-w., Wang X., Huang S., McCubbin A.G. and Kao T.-h. (2004). Chromosome Walking in the *Petunia Inflata* Self-Incompatibility (*S*-) Locus and Gene Identification in an 881-kb Contig Containing S_2 *–RNase.* Plant Molecular Biology 54: 727-742

Wang E. and Wagner G.J. (2003). Elucidation of the functions of genes central to diterpene metabolism in tobacco trichomes using post-transcriptional gene silencing. Planta 216: 686-691

Warnke S.E., Barker R.E., Jung G., Sim S.-C, Mian M.A.R., Sasha M.C., Brilman L.A., Dupal M.P. and Forster J.W. (2004). Genetic linkage mapping of an annual x perennial ryegrass population. Theoretical and Applied Genetics 109: 294-304

Wehling P., Hackauf B. and Wricke G. (1994a). Phosphorylation of pollen grains in relation to self-incompatibility in rye (*Secale cereale* L.). Sexual Plant Reproduction 7: 67-75

Wehling P., Hackauf B. and Wricke G. (1994b). Identification of S-locus linked PCR fragments in rye (*Secale cereale* L) by denaturing gradient gel electophoresis. The Plant Journal 5(6): 891-893

Weising K., Beyerman B., Ramser J. and Kahl C. (1991). Plant DNA fingerprinting with radioactive and dioxigenated oligonucleotide probes complementary to simple repetitive DNA sequences. Electroforesis 8: 805-814

Welsh J., Chada K., Dalal S.S., Cheng R., Ralph D. and McClelland M. (1992). Arbitrarily primed PCR fingerprinting of RNA. Nucleic. Acids Research. 20: 4965-4970

Wheeler M.J., Franklin-Tong V.E. and Franklin F.C.H. (2001). The molecular and genetic basis of pollen-pistil interactions. New phytologist 151: 565-584

Wilhelmi L.K. and Preuss D. (1997). Pollen tube guidance in flowering plants. Plant Physiology 113: 307-312

Wodicka L., Dong H., Mittman M., Ho M.H. and Lockhart D.J. (1997). Genome-wide expression monitoring in *Saccharomyces cerevisiae*. Nature Biotechnology 15: 1359-1367

Wolters-Arts M., Lush W.M. and Mariani C. (1998). Lipids are required for directional pollen-tube growth. Nature 392: 818-821

Wu H.-m., Wang H. and Cheung A.Y. (1995). A pollen tube growth stimulatory glycoprotein is deglycosylated by pollen tubes and display a glycosylation gradient in the flower. Cell 82: 395-403

Xu L., Liu F., Lechner E., Genschik P., Crosby W.L., Ma H., Peng W., Huang D. and Xie D. (2002). The SCF^{col1} ubiquitin-ligase complexes are required for jasmonate response in *Arabidopsis*. The Plant Cell 14: 1919-1935

Yamane H., Ushijima K., Sassa H. and Tao R. (2003). The use of the S haplotype-specific F-box protein gene, SFB, as a molecular marker for S-haplotypes and self-compatibility in Japenese apricot (*Prunus mume*). Theoretical and Applied Genetics 107 (8): 1357-1361

Yih Bih F., Wu S.S.H., Ratnayake C., Walling L.L., Nothnagel E.A., and Huang A.H.C. (1999). The predominant protein on the surface of Maize pollen is an Endoxylanase synthesized by a tapetum mRNA with a Long 5' Leader. Journal of Biological Chemistry 274 (32): 22884-22894

Yu Y.G., Buss G.R. and Maroof M.A.S. (1996). Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. Proceedings of the national academy of sciences of the USA 93: 11751-11756

Zinkl G.M. and Preuss D. (2000). Dissecting *Arabidopsis* Pollen-stigma interactions reveals novel mechanisms that confer mating specificity. Annals of Botany 85 (suppl A): 15-21

Zhang W.-K., Shen Y.-G., He X.-J., Du B.-X., Xie Z.-M., Luo G.-Z., Zhang J.-S. and Chen S.-Y. (2005). Characterization of a novel cell cycle-related gene from *Arabidopsis*. Journal of Experimental Botany 56 (413): 807-816

Zhao T.-Y., Meeley R.B. and Downie B. (2003). Aberrant processing of a Maize GALACTINOL SYNTHASE transcript is caused by heat stress. Plant Science 165: 245-256

Zheng X., Chen X., Zhang X., Lin Z., Shang J., Xu J., Zhai W. and Zhu L. (2004). Isolation and identification of a gene in response to rice blast disease in rice. Plant Molecular Biology 54: 99-109

Zhu T. (2003). Global analysis of gene expression using GeneChip microarrays. Current Opinion in Plant Biology 6: 418-425

Curriculum vitae

Personalia

Naam:	Van Daele Inge
Woonplaats:	Gontrodeheirweg 171 9090 Melle
Telefoon:	09/2110378
Email:	inge.vandaele@belgacom.net
Nationaliteit:	Belg
Geboren:	te Leuven op 14/02/1974
Status:	samenwonend en moeder van Lien en Michiel Lemmens

Studies

Middelbaar onderwijs:

Tot 1992: Humaniora ASO richting Latijn-wetenschappen, Sint-Pietersinstituut, Gent

1992-1993: voorbereidend jaar wiskunde, Sint-Amandusinstituut, Gent

Universitair onderwijs:

1993-1994: eerste kandidatuur polyvalente wetenschappen, VUB, onderscheiding
1994-1995: tweede kandidatuur bio-ir, VUB, onderscheiding
1995-1996: eerste jaar bio-ir, optie cel- en genbiotechnologie, VUB, onderscheiding
1996-1997: tweede jaar bio-ir, optie cel- en genbiotechnologie, VUB, onderscheiding
1997-1998: derde jaar bio-ir, optie cel- en genbiotechnologie, VUB, grote onderscheiding

Scriptie: Isolatie van genen betrokken bij de zelf-incompatibiliteitsrespons in raaigras (*Lolium* spp.).

Promotor: Prof. Dr. Ir. Michel Jacobs

Beroepservaring

1998-2003 (augustus): IWT-bursaal aan de vakgroep plantaardige productie, UG, met als thesisonderwerp: “Het moleculair mechanisme van zelf-incompatibiliteit bij raaigras (*Lolium* spp.)” in samenwerking met het Departement voor Plantengenetica en Veredeling (DVP, CLO-Gent) onder begeleiding van Dr. Isabel Roldán-Ruiz en Prof. Dr. Ir. Erik Van Bockstaele.

Aug 2003-heden: Wetenschappelijk medewerker aan het CLO-DVP, labo biotechnologie, als doctoraatsstudent met hetzelfde thesisonderwerp.

Publicaties

Publicaties in wetenschappelijke tijdsschriften zonder referee

Van Daele I., Van Bockstaele E., De Loose M. and Roldán-Ruiz I. (1999) Development of pollen tubes during fertilization in ryegrass. Proceedings of Workshop on Gametic embryogenesis in Monocots, Cost action 824. Jokioinen, Finland: 10-13 june 1999, 21-22

Van Daele I., De Loose M., Van Bockstaele E, Roldán-Ruiz Isabel (1999). Molecular analysis of self-incompatibility in ryegrass. Med. Fac. Landbouww. Univ. Gent, 64/5b: 447-449

Maras M., Van Daele I., De Loose M., Baert J., Van Bockstaele E., Roldán-Ruiz I. (2000). Characterisation of the self-incompatibility system in chicory (*Chicorium intybus* L.). Med. Fac. Landbouww. Univ. Gent, 65/3b: 467-469

Van Daele I., De Loose M., Van Bockstaele E, Roldán-Ruiz I. (2000). Molecular analysis of self-incompatibility in ryegrass. Med. Fac. Landbouww. Univ. Gent, 65/3b: 443-446

Van Daele I., Van Bockstaele E, Roldán-Ruiz I. (2003).

Identification and mapping of SI-related genes in *Lolium perenne*. Proceedings of the 3rd International Symposium, Molecular Breeding of Forage and Turf, Dallas (Texas) (11), 127-131.

Van Daele I., Van Bockstaele E, Roldán-Ruiz I. (2003).
Identification and mapping of SI-related genes in *Lolium perenne* (2003). Proceedings of the 7th International Congres of Plant Molecular Biology (ISPMB), 23-28 juni 2003, Barcelona, Spanje, p 148

Uitgebreide verslagen en technische rapporten

IWT aanvraag 1st termijn: "Het moleculair mechanisme van zelf-incompatibiliteit bij raaigras (*Lolium* spp.)", September 1998, 30 pp

IWT aanvraag specialisatiebeurs tweede termijn: "Het moleculair mechanisme van zelf-incompatibiliteit bij raaigras (*Lolium* spp.)", mei 2000, 37 pp

IWT verslag eerste drie beursjaren (1998-2001): "Het moleculair mechanisme van zelf-incompatibiliteit bij raaigras (*Lolium* spp.)", mei 2001, 20 pp

IWT eindverslag (1998-2003): "Het moleculair mechanisme van zelf-incompatibiliteit bij raaigras (*Lolium* spp.)", februari 2003, 20 pp
Activiteitsverslagen CLO, Departement Plantengenetica en Veredeling: 1999, 2000, 2001, 2001, 2003

Development of ryegrass allele-specific (GRASP) markers for sustainable grassland improvement, Interim Progress report, march 2003, QLRT-2001-00862

Development of ryegrass allele-specific (GRASP) markers for sustainable grassland improvement, First Annual Progress report for the period 1.10.2002-30.09.2003, QLRT-2001-00862

Development of ryegrass allele-specific (GRASP) markers for sustainable grassland improvement, Interim Progress report, march 2004, QLRT-2001-00862

Development of ryegrass allele-specific (GRASP) markers for sustainable grassland improvement, Second Annual Progress report for the period 1.10.2003-30.09.2004, QLRT-2001-00862

IWT Tussentijds verslag van het onderzoeksproject: “Reproductie- en zelfincompatibiliteitsonderzoek bij raaigras, januari-december 2004

Development of ryegrass allele-specific (GRASP) markers for sustainable grassland improvement, Interim Progress report, march 2005, QLRT-2001-00862

Development of ryegrass allele-specific (GRASP) markers for sustainable grassland improvement, Second Annual Progress report for the period 1.10.2004-30.09.2005, QLRT-2001-00862

Congressen

Met voordracht:

Cost 824, “Gametic embryogenesis in monocots”, 10-13 juni 1999, Jokioinen, Finland

Met poster:

Forum for Applied Biotechnology (FAB), 24-25 september 1998, Brugge, België

PHd Symposium, 07 oktober 1998, FLTBW, UGent

Forum for Applied Biotechnology (FAB), 22-23 september 1999, Gent, België

PHd sympsosium, 13 oktober 1999, FLTBW, UGent

Forum for Applied Biotechnology (FAB), 28-29 september 2000, Brugge, België

7th International Congres of Plant Molecular Biology (ISPMB), 23-28 juni 2003, Barcelona, Spanje

Passieve deelname:

Forum for Applied Biotechnology (FAB), 24-25 september 2001, Gent, België

Eucarpia: 20th international symposium section ornamentals: strategies for new ornamentals. EUCARPIA, 3-6 juli 2001, Melle, België

Andere activiteiten

Deelname aan studiedagen en opleidingen

4th Accelerating Gene Discovery and Mutation Detection Seminar (Perkin Elmer), 12 november 1998, Brasschaat

Westburg/Clontech seminar 'cDNA Expression Arrays-Now and Beyond', 25 november 1998, Brussel

Cursus: "Toegepaste plantenveredeling" van Prof. Dr. ir. E. Van Bockstaele
Academiejaar 1998-1999, FLTBW, vakgroep Plantaardige Productie, UGent

1ste international meeting of the F.W.O. Research Community on "Plant Growth Regulators: metabolism, function and mode of action" in samenwerking met BeNePhyt, 8 januari 1999, U.I.A., Antwerpen

48ste Post-universitaire studiedag: "GMO's in de voeding", 1 december 1999, FLTBW, UGent

Cursus postuniversitair onderwijs: "Epigenetics" van Prof. A. Depicker
Academiejaar 1999-2000, Faculteit Wetenschappen, vakgroep Moleculaire Genetica, UGent

ICES-Course in Bio-Informatics, Module 2: Bio-statistics: Instituut voor permanente vorming
Academiejaar 2000-2001, Ugent

Cursus postuniversitair onderwijs: "Functional genomics" van Prof. M. Zabeau
Academiejaar 2001-2002, Faculteit Wetenschappen, vakgroep Moleculaire Genetica, UGent

4th international LI-COR user meeting, 21 november 2002, Leusden (NL),

Vormingsdagen georganiseerd door het Belgisch EMBnet Knooppunt, 13-14 januari 2003, ULB, Brussel

ABI 7000 Basic Customer Training, 28-29 januari 2003, Nieuwerkerk aan de Ijssel

Cursus postuniversitair onderwijs: “Molecular Techniques in Plant Breeding” van Dr. I. Roldán-Ruiz
Academiejaar 2002-2003, Faculteit Wetenschappen, Ugent
Cursus: “Introduction to Bioinformatics”, 29 april 2005

Buitenlandse stages

Stage in het ‘Institute of Grassland and Environmental Research (IGER)’, Aberystwyth, Wales, 10-16 juni 2000
Bepaling van de incompatibiliteitsrelaties tussen *Lolium perenne* nakomelingen

Deelname aan vergaderingen met actieve bijdrage al dan niet in het kader van projecten

Bezoek aan het zaadbedrijf DSV in Thule en aan het "Institut fur Zuchtungsmethodik" van het Bundensansalt fur Zuchtungsorschung (BAZ) in Gross Lusewitz, 9-10 maart 1999, Duitsland
Bespreking van het zelf-incompatibiliteitsonderzoek in raaigras en rogge

Bezoek aan het veredelingsbedrijf Advanta, 9 september , Rilland, Nederland
Bespreking van het zelf-incompatibiliteitsonderzoek in raaigras

Tweede werkvergadering van het EU project GRASP, 24-25 maart 2003; Plant Research International (PRI), Wageningen (NL.)

Derde werkvergadering van het EU project GRASP, 19-21 oktober 2003; Institut National de la Recherche Agronomique (INRA), Lusignan, Frankrijk

Vierde werkvergadering van het EU project GRASP, 24-26 oktober 2004; Institute of Grassland and Environmental Research (IGER), Aberystwyth, UK

Vijfde werkvergadering van het EU project GRASP, 06-08 april 2005; Danish Institute of Agricultural Sciences, Research Centre Flakkebjerg, Slagelse, Denemarken

Begeleiden van scripties

Coppens Ingeborg, academiejaar 1999-2000; Opsporen van genen betrokken in de zelf-incompatibliteitsrespons van raaigras en cichorei. Eindwerk Katholieke Hogeschool Sint-Lieven, Gent, 48 pp

Externe dienstverlening

Organisatie werkwinkel van DVP op de KVLV studiedag: Kennismakingsdag biotechnologie beter begrijpen, 18 november 2000, Leuven