Investigating *R* gene evolution by meiotic recombination using synthetic gene clusters in Arabidopsis

Jian Sun

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

Doctor of Philosophy

In

Department of Plant Pathology, Physiology and Weed Science

John G. Jelesko, Chair

John M. McDowell, Co-Chair

Elizabeth A. Grabau

Boris A. Vinatzer

Eric P. Beers

April 30th, 2008

Blacksburg, Virginia

Keywords: *R* gene, evolution, meiotic, somatic, unequal crossing-over, recombination, synthetic gene cluster

Investigating *R* gene evolution by meiotic recombination using synthetic gene clusters in Arabidopsis

Jian Sun

(ABSTRACT)

Plant gene families organized as linked clusters are capable of evolving by a process of unequal crossing-over. This results in the formation of chimeric genes that may impart a novel function. However, the frequency and functional consequences of these unequal cross-over events are poorly characterized. Plant disease resistance genes (*R* genes) are frequently organized as gene clusters. In this study, I constructed a reconfigurable synthetic *RPP1* (for resistance to *Paranospora parasitica*) gene cluster (*synthRPP1*) to model *R* gene evolution by meiotic recombination. This experimental design utilizes a gain-of-luciferase phenotype (luc^+) to identify and isolate recombinant *R* genes and uses two alternatively marked alleles to distinguish and measure different types of meiotic recombination (intra- vs. inter-chromosomal). Two putative single copy transgenic plant lines containing the *synthRPP1* gene cluster were generated. The *synthRPP1* gene clusters in these lines were reconfigured *in vivo* by two kinds of site-specific recombination systems (CRE/*Lox*, FLP/*FRT*) to generate two alternative versions of the *synthRPP1* gene clusters *in vivo*. These lines, as well as others being developed, will be used in future genetic crosses to identify and characterize plants expressing chimeric *RPP1* genes. My second area of research was to use a previously developed synthetic *RBCSB* gene cluster (*synthRBCSB*) to investigate the relative frequency of meiotic unequal crossing over between paralogous genes located on either homologous chromosomes (homozygous lines) or sister chromatids (hemizygous lines). In contrast to published somatic recombination

frequencies using a different reporter gene system, no statistically significant difference of meiotic unequal crossing over was observed between homo- and hemi-zygous *synthRBCSB* lines. This result suggests that meiotic unequal crossing-over between paralogs located on homologous chromosomes occurs at about the same frequency as paralogs located on sister chromatids. To investigate the rate of somatic recombination in *synthRBCSB* lines, a QRT-PCR method was developed to estimate the frequency of somatic recombination. Preliminary results suggest that the somatic recombination frequency was about 10,000 fold higher than meiotic recombination in the same generation. Moreover, two of five cloned chimeric genes that formed by somatic recombination indicated a different distribution of resolution sites than those observed in meiotic recombination. This finding suggests there are significant differences in both the frequency and character of somatic versus meiotic unequal crossing-over between paralogous genes in Arabidopsis.

Acknowledgements

I would like to thank all of the persons who have helped or encouraged me to realize one of the biggest career goals in my life and make this achievement. Special thanks will be given to my major co-advisors Dr. John Jelesko and Dr. John McDowell. Dr. Jelesko is a great advisor and teacher, who taught me not only the science, but also how to be a successful scientist. He was always patient, explaining every detail that I didn't understand, especially in the first year when my English was not good. He often trained me to think about questions and then guided me to the correct direction instead of just let me "do this or that", from which I benefited a lot. He allowed me to try new things, make mistakes, learn from them, and was always supportive financially and emotionally. Dr. McDowell is also a great advisor. He gave me a lot of critical thinking about my projects and experiments. He helped me solve a lot of problems. Here I express my earnest appreciation to them.

I am also grateful for my committee members, Dr. Elizabeth Grabau, Dr. Eric Beers and Dr. Boris, Vinatzer. I wish to thank them for many helpful suggestions that helped me stay on the right track and keep moving.

I'd like to thank Crystal Gilbert for her prior work on the *RPP1* recombination project, for the many transgenic lines made by her and for the help of letting me get familiar with lab equipment and facilities.

Thanks to Michael Weaver in The Sainsbury Laboratory, Norwich, UK for kindly providing the source of *RPP1* genes, Dr. Paul Sadowski in University of Toronto, Toronto, CA for providing FLP enzyme and the corresponding plasmid. Dr. Nigel J. Kilby in Boreal Plant Breeding Ltd., Finland for providing the *HSP-FLP* transgenic Arabidopsis seeds.

Thanks to the NIH for financial support.

Table of Contents

List of figures and tables

Chapter One: Background and Justification

1.1. Meiotic unequal crossing-over

Meiotic recombination is a process by which chromosomes exchange information and is necessary for proper segregation of homologous chromosomes to opposite poles at meiosis. During meiosis, two alleles of a particular gene may align and recombine. In most cases, the resulting recombinants retain an equal number of parental alleles and this is referred to as recombination or crossing-over (Tartof, 1988; Meselson & Radding, 1975). Less frequently there is a net gain of one allele with a corresponding loss of another allele. These are called gene conversion events. Both of these are normal meiotic recombination events (Smith, 1974) (Figure 1).

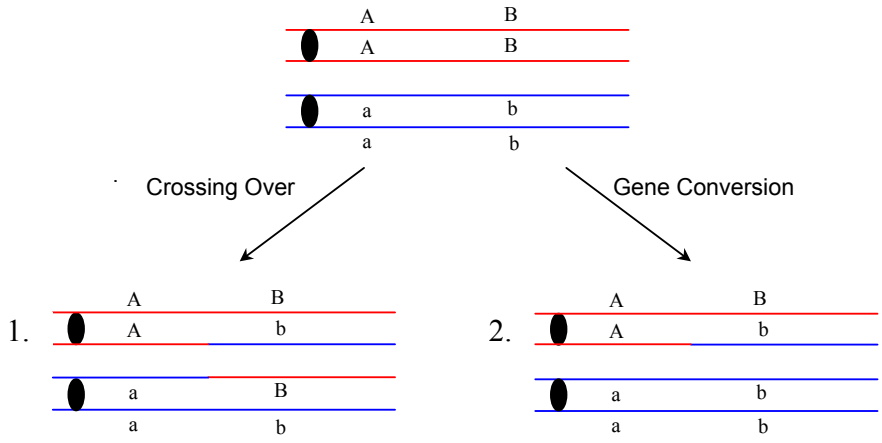


Figure 1: Meiotic recombination. "A" and "a", "B" and "b" are two pairs of allelic genes. Red lines and blue lines indicate DNA molecules. Black dots indicate centromere.

However, some plant multigene families are organized as a gene cluster as shown in Figure 2. Homologous genes that formed by duplication within a given species are called paralogs.

Paralogous genes generally have high degree of sequence similarity between them. This paralogous gene cluster arrangement provides the possibility of misalignment of two non-identical paralogs in the gene cluster during meiosis. If this misalignment occurs and the recombination machinery resolves them, it will result in two unequally exchanged chromosomes. Specifically, it will create a deletion on one chromosome, a duplication on the other chromosome, and two reciprocal recombinant genes (Figure 2). This type of recombination is significant because it is a potential mechanism for producing novel chimeric genes from pre-existing homologous genes.

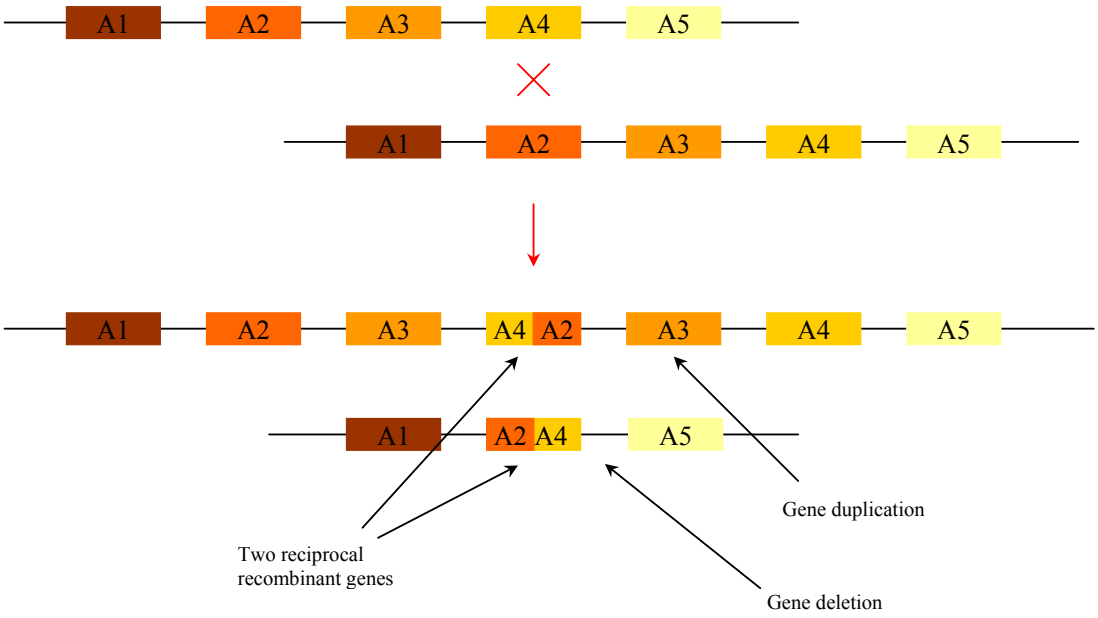


Figure 2: Meiotic unequal crossing-over. A1-A5 represent different paralogous genes within a gene cluster.

Many observations support meiotic unequal crossing-over between paralogs within a gene cluster resulting in expansion/contraction of the gene families. These include *KIR* genes (Martin *et al.*, 2003), *Hox* genes (Ruddle *et al.*, 1994), amylase genes (Gumucio *et al.*, 1988), SMS (Shaw *et al.*, 2002), *NF1* genes (Lopez Correa *et al.*, 2000) and *CYP21* gene

(Tusie-Luna & White, 1995) in humans and immune system gene families (Hughes & Yeager, 1997) in humans and other vertebrates. In most cases, it is difficult to determine when and how meiotic unequal crossing-over gave rise to the gene duplications, deletions, and novel genes.

Sequence comparison suggested that unequal crossing-over might be a mechanism of gene evolution (Martin *et al.*, 2003). Restriction map analysis (Lopes *et al.*, 1996), flanking marker analysis (Buard *et al.*, 2000), and PCR based analysis (Tusie-Luna & White, 1995) provide some details about meiotic unequal crossing-over (such as the recombination frequency) at very unstable loci. Genetic screening based upon an altered phenotype tends to be biased toward the identification of the deletion products because these confer a complete loss of function. On the other hand, gene duplications and chimeric recombinant genes often go unnoticed unless the locus is either susceptible to gene dosage effects or a novel and predictable phenotype is conferred by the newly formed chimeric genes.

1.2. Early observations of unequal crossing-over

In contrast to animals, plants are relatively insensitive to gene dosage effects. For this reason, mutant phenotypes caused by unequal crossing-over were initially discovered in insect and animal models, in which gene dosage effects are more apparent. The first examples of unequal crossing-over were characterized at *bar* (Sturtevant, 1925) and bobbed (*bb*) (Ritossa *et al.*, 1966) loci of Drosophila, which result in changes in eye shape (i.e. narrow vs. round) and a short-bristle phenotype, respectively. In the case of the *bb* locus, bristle length is reversible. The short bristle phenotype *bb* is associated with a "reduction" of ribosomal DNA (rDNA) (i.e. deletion), whereas a "magnification" or increase in rDNA (i.e. gene duplications caused by unequal crossing-over at the *bb* locus) would revert back to a wild type *BB* phenotype (Tartof, 1974; Tartof, 1988) .

Unequal crossing-over events leading to congenital diseases have been reported within human gene clusters. A good example of a human disease associated with unequal

crossing-over is human color blindness (Nathans *et al.*, 1986b). Human color vision is controlled by the red- and green-pigment encoding genes. In a color-normal male, the genes are arranged as an array with one red pigment gene followed by a varying number of green pigment genes, ranging from none to multiple copies of the green pigment gene. Intragenic (within genes) and intergenic (between genes) unequal crossing-over result in partial or complete deletions resulting in various color vision deficiencies (Nathans *et al.*, 1986a). Another example of human diseases related to unequal crossing-over is homozygous familial hypercholesterolemia, that results from the unequal crossing-over between two repetitive Alu sequences in the low density lipoprotein receptor gene resulting in a 7.8 kb deletion (Lehrman *et al.*, 1987). A similar situation is also observed in human thalassemias (disease caused by an imbalance of the α- and β-globin chains). Human α- and β-globin genes are organized as gene clusters. The α gene cluster consists of 5 α-like genes occupying an approximately 30 kb region on the short arm of chromosome 16. In α-thalassemias, large deletions were found in α genes, resulting in reduced production of α chains. The long duplications of DNA in α genes were potential target for unequal crossing-over. “Patchwork” arrangements (i.e., a gene with interspersed segments arising from different genes) of some mutants further suggested the occurrence of unequal crossing-over. In β-thalassemias, deletions/insertions and “fused” hemoglobin genes are found, indicating the occurrence of unequal crossing-over (Collins & Weissman, 1984).

The above genetic disorders in flies and humans provide evidence for unequal crossing-over. However, these observations were based on observable altered phenotypes resulting from the deletion products resulting from unequal crossing-over. Even in the examples above, not all of the products of unequal crossing-over are detectable. For example, gene duplication and chimeric genes may not cause an observable altered phenotype. So, most studies are biased towards the identification of the deletion alleles. In comparative genomic type studies of extant gene clusters that show evidence of unequal crossing-over, it is difficult to know when the recombinant genes formed and thus estimates of the frequency of unequal crossing-over are extremely difficult. However, investigations in yeast, described below, have provided many important insights about contemporary unequal crossing-over.

1.3. Unequal crossing over in yeast

Saccharomyces cereviseae is a simple unicellular eukaryotic organism and its life cycle includes both a diploid and haploid life cycle. After meiosis, the four gametes are enclosed in the ascus and there is no additional cell divisions following meiosis. The separation and culture of the four individual gametes allows further characterization of the genetic markers by so-called tetrad analysis. Tetrad analysis is a powerful tool to study meiotic recombination and gene conversion events because it allows the recovery and analysis of each of the four gametes derived from a single meiosis event.

Jackson (Jackson & Fink, 1985) studied meiotic recombination using duplicated *HIS4* fragments in yeast. The most frequent recombination observed is interchromosomal reciprocal recombination (recombination between the same gene on homologous chromosomes), which is as high as 41%. The second most frequent type is unequal interchromosomal recombination (recombination between paralogous genes on homologous chromosomes), which occurs at 17%. However, unequal intrachromosomal recombination (recombination between paralogous genes on sister chromatids) is much less frequent, occurring in only 1% of the tetrads. Interchromosomal and intrachromosomal gene conversions are also observed and occur with approximately equal frequencies ranging from 1 to 5%. Complex recombination such as double recombination events occurs with a frequency of 2.8%. These data indicate that different types of unequal crossing-over occur at different frequencies in yeast.

An important advance in our mechanistic understanding homologous recombination was the development of the Holliday junction model. Holliday (Holliday, 1964) proposed a model to explain the aberrant 3:1 segregation ratio (any heterozygous locus normally yields an expected 2:2 Mendelian ratio in tetrad analysis) observed in fungi. The model starts from two homologous DNA duplexes, each corresponding to one of the four chromatids present in meiosis. After DNA replication, single strand nicks are produced which unravel to form single strands. These single strands of DNA can invade the opposite chromosome and pair with the complementary strand to form a hybrid DNA intermediate. Strands

separate and exchange along the DNA molecules to extend the hybrid molecules. After strand annealing occurs, the chromatids form a chiasma-like intermediate called a Holliday junction. Alternative orientation of breakage and ligation of the opposite DNA strands (i.e. resolution) give different recombination results. Depending upon the orientation of resolution in the Holliday junction, it will result in either exchange of flanking markers (i.e. a cross over) or short patches of interspersed sequence of flanking markers (i.e. gene conversion). The Holliday model initiates by hypothetical single-strand nicks. However, Orr-Weaver (Orr-Weaver *et al.*, 1981) found that the frequency of recombination between genes on chromosomes and plasmids can be enhanced as much as 1,000 fold by the introduction of double-strand breaks in the homologous regions. This brought forward the idea of double-strand break repair model.

Szostak and colleagues (Szostak *et al.*, 1983) proposed the double strand break (DSB) model for meiotic recombination. The model begins with a DSB on one chromosome. The break enlarges by deleting from both 5' ends, leaving 3' single strand tails. Then both 3' single strands invade a homologous DNA duplex and the invading ssDNA molecule pairs with complementary DNA sequences. The annealed 3' ends then replaces the other strand by elongation via DNA synthesis to form an enlarging D loop. When the DNA synthesis fills the gap and the DNA religates, two Holliday junctions are formed. Alternate patterns of resolution of the two Holliday junctions will lead to two possible recombination outcomes: i) a non-crossover (i.e. gene conversion) or ii) a crossing-over. Thus, the DSB model for DNA recombination accounts for the formation of both cross-overs and gene conversion events.

Studies in yeast provide many important insights about meiotic recombination, such as the frequencies of intra- vs. inter-chromosomal reciprocal recombination and gene conversion (Jackson & Fink, 1985). However, one must be careful when trying to extend results from yeast to multicellular organisms, especially higher plants, where gene organization is dramatically different from that in yeast. In plants, genes are often organized in gene clusters, whereas gene clusters are quite rare in yeast. Likewise, the overall recombination frequency at a particular gene is much lower in plants than in yeast: ~1% in yeast (Jackson

& Fink, 1985) compared to ~10^{-6} in Arabidopsis (Jelesko *et al.*, 1999). Also, the artificial *HIS* gene duplication used in the yeast studies is not a naturally occurring endogenous gene cluster and therefore may not reflect the real situation of meiotic recombination of an endogenous gene cluster.

There are different opinions about how meiotic recombination affects gene cluster evolution. One opinion is that unequal crossing-over and gene conversion tend to homogenize gene clusters (Smith, 1974; Hickey *et al.*, 1991; Ohta, 2000; Gumucio *et al.*, 1988). Examples for this include the rDNA gene cluster and histone gene cluster in which the members of these gene clusters are more similar within species than between species, resulting in so-called concerted evolution. In contrast, reports coming from plant *R* gene clusters (Michelmore & Meyers, 1998) such as the major cluster disease resistance genes in lettuce (Kuang *et al.*, 2004) and the *Cf-4/9* locus in tomato (Parniske *et al.*, 1997) suggest that unequal meiotic recombination does not occur frequently enough to homogenize the gene cluster and tends to generate new chimeric genes. These studies all relied on comparative genomics of extant gene clusters. However, what is needed are empirical measurements of both the pattern and rate of meiotic recombination within extant gene clusters.

1.4. Unequal crossing over in plants

Unequal crossing-over in plants is of particular interest because many plant disease resistance (*R*) genes are organized in large gene clusters, which facilitates misalignment and unequal crossing-over between different paralogs (Dixon *et al.*, 1998; Ellis *et al.*, 1999). Evidence from several different plants species suggest that unequal crossing-over in *R* gene clusters has occurred. For example, the variation of leucine rich repeat (LRR) copy number within the *Cf-2/Cf-5* gene family generates new pathogen recognition specificities in tomato (Dixon *et al.*, 1998). In maize, the high rate of intragenic recombination and mispairing between genes within the *rp1* complex contributes to the diversity of rust resistance specificity (Hulbert, 1997). In lettuce, the related but divergent family of *RGC2*

genes may generate new resistance phenotypes by unequal recombination (Kuang *et al.*, 2004). Unequal crossing-over is also observed in Arabidopsis. Several alleles at the *RPP8* locus (McDowell *et al.*, 1998) are consistent with unequal crossing-over during the diversification of the *RPP8* locus genes.

These examples all support the action of unequal crossing-over in plants. However, because of the uncertainty about the frequency of unequal crossing-over events and a lack of obvious phenotype of recombinant alleles, it is often difficult to observe/detect/capture them in natural plant populations. Alternatively, it is useful to employ artificially constructed DNA repeats or partially overlapping reporter gene fragments to model DNA recombination processes. Artificially constructed DNA repeats have generally been comprised of two inactive and overlapping pieces of a single reporter gene. Thus, when DNA recombination occurs between the overlapping DNA segments, it reconstructs the full-length reporter gene and results in some reporter gene activity. Many reporter genes have been used in DNA recombination studies. For example, DNA recombination has been used to reconstitute an intact Neomycin phosphotransferase gene (*nptII*), which confers resistance to kanamycin (Peterhans *et al.*, 1990; Baur *et al.*, 1990; Assaad & Signer, 1992). Defective beta-glucuronidase (*GUS*) gene is also used to detect homologous recombination, by assaying for assaying for a colorometric histochemical dye (Puchta *et al.*, 1995; Schmidt-Puchta *et al.*, 2004). Similarly, partial overlaping defective *LUC* gene repeats were used to investigate somatic and meiotic recombination in tobacco by assaying for *in vivo* bioluminescence (Kovalchuk *et al.*, 2003). Even a defective partial dimer of Cauliflower Mosaic Virus (CaMV) genome has been used as reporter gene. In this case, when homologous recombination reconstitutes the intact viral gene, replication of the virus ensues leading to viral symptoms as a marker of a recombination event (Swoboda *et al.*, 1993). The homologous overlapping truncated reporter gene segments can be either direct or inverted orientation.

Reporter gene-based approaches have been used to estimate recombination frequencies in a variety of plant tissues. Using two direct repeats of truncated *nptII* gene fragments, Assaad and Signer (1992) and Peterhans *et al.* (1990) measured the recombination rates in

Arabidopsis and tobacco by scoring kanamycin resistant (Km^r) seedlings or Km^r calli. They found that the intrachromosomal recombination occurs at a frequency of approximately ~10^{-5} per division. Similarly, Gal *et al.* (1991) and Swoboda *et al.* (1993) used the direct repeats of alternatively truncated cauliflower mosaic virus (CaMV) sequences to study somatic homologous recombination leading to reconstitution of a replicating virus, that was used as a "reporter" of recombination. Swoboda reported a maximal somatic frequency of 10^{-6} in *Brassica napus*. Besides direct homologous repeats, inverted *nptII* duplications of reporter gene in tobacco were used as well, Tovar and Lichtenstein (1992) estimated a somatic recombination rate of 3×10^{-5} to 10^{-6} per cell.

These somatic recombination events exist in two different contexts, either extrachromosomal (ECR) or intrachromosomal (ICR) recombination. Extrachromosomal recombination (ECR) is recombination between input DNAs such as plasmids that are not stably integrated into plant chromosome. These extracrhomosomal DNAs were introduced into plant cells by either transfecting plant protoplasts with a plasmid or by transformation of Arabidopsis with binary vectors with truncated reporter gene repeats. Somatic ECR occurs mainly in the early period after transformation and can be measured by transient reporter gene expression assay. In general, somatic ECR is efficient, with a frequency of 10^{-1} to 10^{-3} (Puchta *et al.*, 1992; Puchta *et al.*, 1994)). Offringa *et al*. (1990) estimated a homologous recombination rate of 3×10^{-4} between an extrachromosomal T-DNA with a defective *nptII* gene at integrated into the plant chromosome. They also estimated a recombination rate of 1-4% between two co-transformed T-DNAs in tobacco protoplasts. Similarly, Hrouda and Paszkowski (1994) showed low efficiency (10^{-4}) of ECR between a chromosomal locus containing modified *nptII* gene and extrachromosomal homologous *nptII* sequence. Puchta and Hohn (1991) co-transfected tobacco protoplasts with plasmids carrying non-allelic mutations in *GUS* repeats and monitored *GUS* activity (i.e. recombination rates). They found the recombination rates depended on the length of homologous overlap and proposed that the mechanism of ECR is probably via the single strand annealing model rather than the double strand break repair model. Because of different types of extrachromosomal DNA substrates were used in these different studies, there are widely variable estimates of recombination frequencies. ECR is likely a measure

of somatic recombination associated with DNA damage.

Although extrachromosomal recombination experiments provide empirical estimates of somatic recombination rates in plants, a more relevant view of natural DNA recombination is one that examines homologous recombination between DNA sequesnces located on intact chromosomes (inter/intra chromosomal reaction, or ICR), because naturally occurring recombination within gene clusters is intrachromosomal. Somatic ICR is defined as DNA recombination between homologous sequences located either within or between two homologous chromosomes in somatic tissues. ICR occurs less frequently than ECR, usually one event in 10^5 to 10^7 cells (Puchta *et al.*, 1994). ICR can take place as (i) intermolecular recombination between sister chromatids or homologous chromosomes, or (ii) as an intramolecular event between adjacent DNA repeats (either direct or inverted) located on the same chromosome. To study ICR, Peterhans *et al*. (1990) and Tovar *et al*. (1992) inserted a pair of defective *nptII* genes positioned as inverted repeats into tobacco genome as a stable transgene. Kanamycin resistant calli were observed at a frequency between 10^{-4}-10^{-6}. Swoboda *et al*. (1994) and Puchta *et al*. (1995) also studied ICR using defective *GUS* direct repeats in Arabidopsis and tobacco genomes, respectively. Overall somatic recombination frequency is in the range 10^{-6}-10^{-7}, but varies in different organs of different transgenic lines and between even different individuals. Gal *et al*. (1991) and Swoboda *et al*. (1993) used CaMV sequence to do a similar study and found the recombination frequency increased with increasing transgene copy number and the maximal frequency was of the order of 10^{-6} per cell division (Swoboda *et al.*, 1993).

Because of the ease of transformation and selection, many studies measured recombination within somatic tissues. However, these products of somatic recombination generally can't be inherited in the next generation, thus they have less significance than meiotic recombination events in which the recombinant alleles are passed to the next generation. Relatively few studies have focused on meiotic recombination. Two early reports used a pair of defective inverted *nptII* repeats to measure meiotic recombination frequency. Tovar and Lichtenstein (1992) estimated meiotic recombination at between $\sim 3\times10^{-5}$ to 10^{-6}. Similarly, Assaad and Signer (1992) used two non-overlapping truncated *nptII* direct

repeats to measure the meiotic recombination rate in stable transgenic Arabidopsis plants. The meiotic recombination rate was estimated on the order of 10^{-6}-10^{-5} per meiotic division (22 $Kan^{+}/9.3\times10^{5}$). Based on these limited data, similar meiotic recombination rates were observed in different experiments. However, these quantitative estimates of meiotic recombination frequencies were based upon only one or two putative meiotic recombinants. Thus, the reproducibility of these estimates was not established.

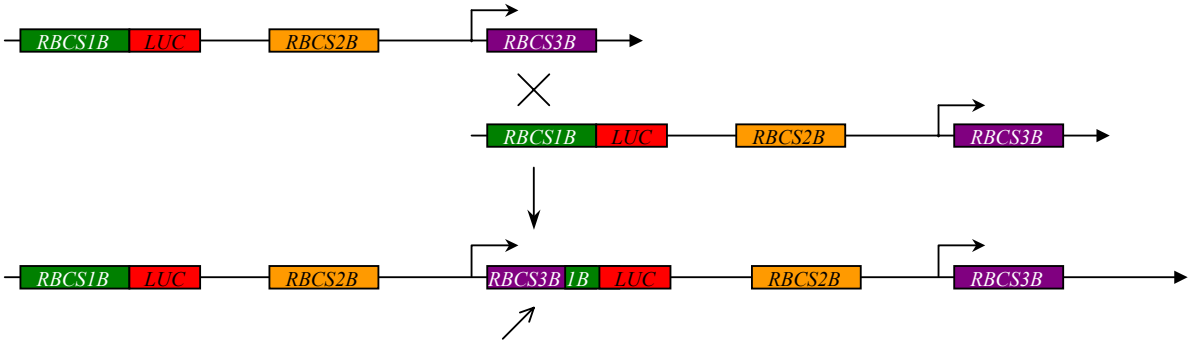
Recombination studies using artificial gene duplications of a reporter gene have certain limitations. First, the reporter genes are of bacterial origin and have different gene structure than endogenous plant genes. Specifically, they did not contain introns. Thus they may behave in different ways than do endogenous plant genes. Second, they were composed of relatively short regions of homology (<1 kb) that may affect the recombination efficiency. Third, they didn't create chimeric genes (i.e. only restored the original reporter gene) and therefore don't provide information about recombination resolution sites.

1.5. Using synthetic gene cluster to investigate unequal crossing over

An alternative system for measuring DNA recombination in plants was developed that uses a reporter gene to monitor recombination between *bona fide* plant genes. This approach uses paralogous plant genes that are naturally present as a gene cluster and therefore serve as natural substrates for recombination. This approach is called synthetic gene cluster technology. A synthetic gene cluster is a modified naturally occurring gene cluster in which one of the paralogous genes is both inactived and fused in-frame to a reporter gene. This chimeric gene is the target for DNA recombination events. Recombination between a linked active paralogous gene and the inactive paralog-reporter gene fusion results in the formation of an activated chimeric reporter gene and expression of reporter gene activity. The use of synthetic gene cluster technology has several advantages. First, the genes are naturally occurring and should behave similar to natural meiotic recombination. Second, the paralogs are not 100% identical, which allows the mapping of recombination resolution

(i.e. cross-over) sites. Third, the gene cluster arrangement allows unequal crossing-over between non-allelic paralogs in different orientations resulting in altered organization of the gene cluster (i.e. inversion, deletion and duplication).

The first example of a synthetic gene cluster was the Arabidopsis *RBCSB* synthetic gene cluster (Jelesko *et al.*, 1999). It is derived from a naturally occurring tandem array of three *RBCSB* genes (*RBCS1B, RBCS2B* and *RBCS3B*). The 5' region of the *RBCS1B* gene was deleted and the 3' region fused in-frame to the luciferase gene (i.e. *ΔRBCS1B:LUC* gene fusion as shown in Figure 3). Unequal crossing-over between *ΔRBCS1B:LUC* and *RBCS2B* (or *RBCS3B*) would result in an active *RBCS2B(3B)/RBCS1B:LUC* chimeric gene fusion. The active *RBCS3B* gene promoter resulted in expression of a *RBCS3B/1B:LUC* gene fusion conferring luciferase activity to the plant (Figure 3). This bioluminescence can be easily identified using a low light video imaging system. The gain-of-luciferase phenotype enables the identification and isolation of rare recombination events from large number of unrecombined seedlings. The screen is fast and efficient. It is easy to screen over 7,500 five day-old seedlings every ten minutes. This screen is also sensitive enough to detect extremely low levels of recombinant chimeric gene-luciferase gene fusion expression, i.e. as low as 23 photons in 20 minutes.



RBCS1B::LUC gene fusion is activated when recombination between *RBCS3B* and *RBCS1B* occurs

Figure 3: *RBCSB* genes recombination and Luciferase gene screening method

Using a synthetic gene cluster, Jelesko *et al.* (1999; 2004) performed detailed studies on meiotic intergenic unequal crossing-over between Arabidopsis sister chromatids. They observed reproducible unequal crossing-over frequencies of approximately 3×10^{-6} between paralogous *RBCS1B* and *RBCS3B* genes. Morover, an observed non-uniform distribution of the recombination resolution sites brought forward the hypothesis that meiotic recombination tends to happen in a region where relatively high sequence similarity transitions to a region of low sequence similarity.

These attributes of synthetic gene cluster technology provide many opportunities to examine unequal crossing-over between paralogous genes organized as gene clusters, including plant disease resistance genes that are believed to evolve in part by unequal crossing-over.

1.6. *R* gene evolution and unequal crossing over

1.6.1. Disease resistance in plants

Plants utilize a variety of mechanisms to defend against pathogens. These mechanisms include non-specific passive mechanisms and specific recognition defenses. Passive mechanisms include waxy cuticular layers and anti-microbial phytoalexins. Most plant pathogens that are not highly adapted to a particular plant species are blocked by these physical and chemical obstacles.

However, some pathogens can circumvent these defense barriers and attack plants. The plant innate immune system, on the other hand, has evolved highly specific mechanisms to recognize these pathogens. The recognition of a pathogen by the plant can trigger signal transduction cascades that result in a specific defense response, usually localized cell death (Figure 4). A specific defense response against a pathogen is initially triggered by the recognition of pathogen by a specific R protein that is encoded by a constitutively expressed "*R* gene". The pathogen determinant recognized by the R protein is referred to as

the avirulence product, encoded by a particular *avr* (avirulence) gene. Only when an *avr* gene is present in the pathogen and the corresponding *R* gene is present in plant, can this specific recognition be triggered. This model of interaction requires both a specific *R* gene in the plant genome and a corresponding *avr* gene in pathogen (Flor, 1942).

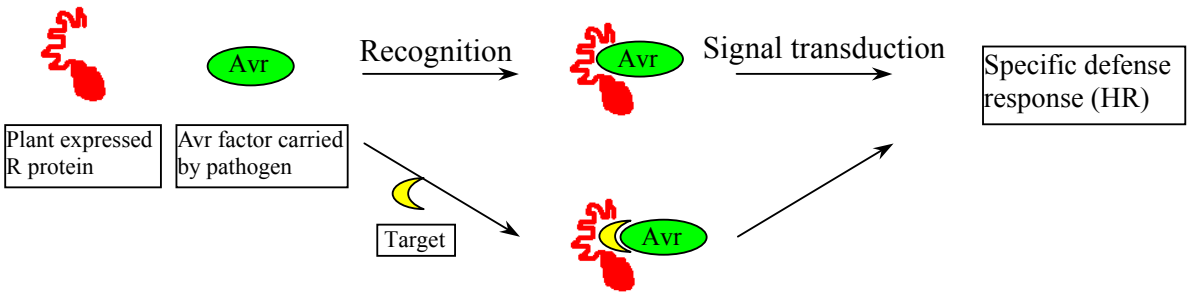


Figure 4: Specific R protein mediated defense response

In some cases, R protein recognizes the Avr protein directly to trigger defenses response (Jia *et al.*, 2000). However, in other cases, it doesn't. The interaction between an R protein and the cognate Avr protein can be mediated by a third protein (Target) instead of direct interaction. The Avr protein binds to the target to form a complex and causes a conformational change of the target. Then the target can be recognized by the R protein to trigger resistance defenses. This model of target mediating the recognition of R protein and Avr is known as the "Guard Hypothesis" (Van der Biezen & Jones, 1998).

Thespecific defense response induced by the recognition of an Avr by an R protein usually results in the hypersensitive response (HR). HR is a rapid and active resistance response against pathogens causing localized cell death in and around the location of infection to restrict the further invasion of pathogens (Hammond-Kosack & Jones, 1996; Delledonne *et al.*, 1998). When the resistance proteins detect the pathogens, the cells change the pH and ion permeability, form lesions (Orlandi *et al.*, 1992), and produce reactive oxygen species

(e.g., superoxide anions), callose, lignin and other compounds (Hammond-Kosack & Jones, 1996). These physiological changes and chemical toxins can prevent the pathogen's further invasion, if not enough to kill the pathogen.

1.6.2. R gene structure, function, and diversity

A large variety of *R* genes have been identified in the past few years and the R proteins they encode specify resistance to bacterial, viral, fungal, oomycete and even some nematodes and insect pests. This large number of R proteins can be divided into five classes.

The largest class of R proteins is composed of a "nucleotide-binding site (NB) domain with a leucine-rich repeat" (LRR) domain (Dangl & Jones, 2001). The basic structural features of this class are highly variable carboxy terminal LRRs and a conserved nucleotide-binding site (NB). Sequence comparisons and domain swapping suggested that *R* gene specificity largely resides in the LRR region (Botella *et al.*, 1998; Dixon *et al.*, 1998; Parniske *et al.*, 1997; Dixon *et al.*, 1996). LRR domain was found to have homologs involved in protein-protein interaction or ligand binding in many other proteins (Jones & Jones, 1996; Baudino *et al.*, 2001). The NB domain is the site of ATP or GTP binding. The NB-LRR class can be subdivided based on the N-terminal structural features: one class (TIR-NB-LRR) has a N-terminal with homology to the intracellular signaling domains of the *drosophila* Toll and mammalian interleukin-1 (IL-1) receptors and the other class (CC-NB-LRR) has a putative coiled-coil domain (CC). The functions of TIR and CC domains may be involved in the interactions between proteins mediating signal transduction.

The other four classes of R proteins are diverse in structure and function. They include i) extracellular LRR domains associated with the cell membrane (such as Cf-5 (Jones *et al.*, 1994; Dixon *et al.*, 1996)), ii) cytoplasmic kinases (such as Pto (Tang *et al.*, 1996)), and iii) extracellular LRR and cytoplasmic kinase domains connected by a trans-membrane domain (such as Xa21 (Song *et al.*, 1997)). One R protein, RPW8 (Xiao *et al.*, 2001), was found to have no homology to known proteins. RPW8 consists of a possible membrane associated domain and a possible coiled-coil cytoplasmic domain. The functions of the R proteins in

these classes are also diverse. Some of them even participate in other biological process such as development.

Much interest has been focused on the *R* gene structure to determine R protein structure-function relationships. A large number of *R* genes encode contain one or more LRR domains. Several lines of evidences suggest that the recognition specificity mainly resides in the LRR region. Exchange of LRRs in some cases can alter *R* gene specificity (Ellis *et al.*, 1999). Even though the exact structure elements contributing to recognition specificity are not completely understood, the study of LRR diversity and specificity will provide more detailed insight of how R protein recognize pathogen Avr determinant.

1.6.3. Rapid evolution of R genes

R and Avr proteins are encoded by corresponding *R* genes and *Avr* genes. In order to recognize Avr factors, the plant genome needs to evolve and express corresponding *R* genes. This relationship between a *R* gene and *avr* gene is the gene-for-gene hypothesis, which was first proposed by Flor (Flor, 1942).

The complete Arabidopsis genome sequence analysis revealed that there are ~150 sequences with homology to the NB-LRR class of *R* genes (Dangl & Jones, 2001). The number of *R* genes seems to be small compared to the large number of possible pathogen-encoded ligands. R proteins usually have narrow recognition specificity, capable of recognizing one or few Avr determinants. Even if one R protein recognized more than one Avr protein, it is still not able to recognize highly divergent and rapidly evolving pathogens. Moreover, in order to circumvent plant *R* gene recognition, pathogens mutate the *avr* gene to produce novel Avr proteins that can't be recognized by existing R proteins. So how can this small number of *R* genes recognize a large number of diverse pathogens? The organization of plant *R* genes provides important clues to *R* gene evolution. *R* genes are frequently organized in gene clusters. Thus, *R* genes in linked clusters have the potential to create new genes by unequal crossing-over between paralogous *R* genes within the cluster.

Gene clusters are a common phenomenon in plant genomes (Meyers *et al.*, 1998; Michelmore & Meyers, 1998; Baumgarten *et al.*, 2003). Some *R* gene loci consist of a functional *R* gene and one or more genes that are non-functional gene or have no detectable function, such as the *RPP8/RPH8A* locus in Arabidopsis (McDowell *et al.*, 1998). Most characterized *R* genes belong to complex *R* gene loci. These complex *R* gene loci are comprised of multiple paralogs. Some of these paralogs have detectable resistance functions such as the *RPP1* locus in Arabidopsis, the *cf-4/9* loci in tomato and the *Xa21* locus in rice. Four *Hcr9* genes (*cf-4*, *cf-9*, *9E*, and *9A* or *9B*) are found at the *Cf-4/9* cluster, the products of which recognize at least three different *avr* factors of *C. fulvum* (Parniske *et al.*, 1997). Many of other complex *R* gene loci contain paralogs without obvious resistance functions such as the *Rp1* loci (Collins *et al.*, 1999) and *Rpp5* loci (Parker *et al.*, 1997). These non-functional *R* gene paralogs may recognize as yet uncharacterized pathogens or those that have been functional in the evolutionary past.

Genes organized in gene clusters are likely to be more dynamic than single genes, because they can undergo gene rearrangement and recombination. For example, In *Zea mays*, the *R-r* complex encodes a tissue-specific transcriptional activator controlling anthocyanin pigment biosynthesis (Robbins *et al.*, 1991). The complex is composed of two distinct components: a *P* component (a single *R* gene (not to be confused with a disease resistance gene) conferring pigmentation of plant part) and a *S* subcomplex (consisting of a truncated inactive *R* gene (*Q*) and two functional *R* genes *S1* and *S2*). Loss-of-function of either *P* or *S* is often (10^{-3} to 10^{-4}) found with the exchange of flanking markers. This is explained by the tandemly arranged gene duplications carrying *P* and *S* components that mispair and recombine during meiosis.

Leister *et al*. (1998) also reported the rapid reorganization of resistance gene homologs in cereal genomes. When they mapped monocot *R*-like genes, they found mixed gene clusters harboring highly dissimilar *R*-like genes and the intraspecific variation of *R* gene copy number. When using probes that show interspecific hybridization, NBS-LRR loci often

show nonsyntenic map locations. This evidence strongly suggested a rapid rearrangement of NBS-LRR genes.

1.6.4. Effect of unequal crossing over on R gene evolution

Several lines of evidence suggest that unequal crossing-over and gene conversion are influencing *R* gene cluster evolution (Chin *et al.*, 2001; Robbins *et al.*, 1991; Parniske *et al.*, 1997; Hulbert, 1997; Song *et al.*, 1997; McDowell *et al.*, 1998; Dixon *et al.*, 1998). For example, in Arabidopsis, the *RPP8* locus in Landsberg erecta (Ler-0) imposes resistance to *Hyaloperonospora parasitica* isolate Emco5. The *RPP8*-Ler locus encodes an NBS-LRR protein imposing resistance to *Hyaloperonospora parasitica* isolate Emco5 and a homologous *R* gene, *RPH8A*, with unknown resistance specificity. In contrast, the *RPP8* locus in Col-0 contains a single chimeric gene (McDowell *et al.*, 1998). The Col-0 chimeric *RPP8* gene was likely derived from an early unequal crossing-over event between *RPP8*-Ler and *RPH8A* ancestors within a Ler-like haplotype. Similar situations are found in the *RPP8/HRT* locus in *Dijon-17* (Cooley *et al.*, 2000) and *C24* (Takahashi *et al.*, 2002). Both of the *RPP8/HRT* loci contain one gene that seems to be a chimera of two ancestors. The *RPP8/HRT* locus in *Dijon-17* confers resistance to both viral (Turnip crinkle virus) and oomycete pathogens. The *RPP8/HRT* locus in *C24* confers resistance to yellow strain of cucumber mosaic virus. This suggested that in spite of the high sequence similarity, the sequence exchange between closely related *R* genes can generate novel recognition specificity to completely different pathogens.

The study of the maize *Rp1* locus strongly supported the hypothesis that unequal crossing-over between *R* gene paralogs creates new pathogen recognition specificity (Hulbert, 1997). The *rp1* complex is a cluster of rust (*Puccinia sorghi*) resistance genes in maize. Sixteen different genes in the *rp1* locus were identified by using an extensive collection of rust biotypes. Most (14 out of 16) of them mapped to a single locus. Early observations showed that both susceptible and double-resistance progeny could be obtained from the crosses of *rp1* homozygotes or heterozygotes suggesting that recombination caused gene rearrangements. Studies of the flanking RFLP markers from the progenies from *Rp1*

homozygotes with heterozygous flanking markers revealed non-parental combinations of the flanking markers within the susceptible progenies, indicating crossing-over was involved.

Another *R* gene locus shows that unequal crossing-over is involved in *R* gene evolution and formation of new resistance specificities. Parniske and colleagues (Parniske *et al.*, 1997) reported novel disease resistance specificities were generated by sequence exchange between tandemly repeated *Cf* genes in tomato, which confer resistance to *C. fulvum*. *Cf-4* and *Cf-9* are close related *R* genes. The *Cf* genes in *Cf-4/9* transheterozygous tomato shows dramatic meiotic instability. Five disease-sensitive progenies were identified from 7,500 progenies of a *Cf-4/9* transheterozygous cross to *Cf0* corresponding a frequency of 7×10^{-4}.

The opinions regarding how unequal crossing-over effects *R* gene clusters still remains controversial. Unequal crossing-over occurs frequently in some *R* gene clusters (*cf4/9*, *rp1*). Obviously the frequent recombination between paralogs in the gene cluster continuously generates novel chimeric genes, which may acquire novel resistance specificity. In this view, unequal crossing-over tends to diversify the paralogs within gene clusters. On the other hand, there are opinions that frequent unequal crossing-over and gene conversion in gene cluster tends to homogenize gene family (Smith, 1974; Hickey *et al.*, 1991) and slows the divergence of the paralogs within gene clusters.

However, at other *R* gene clusters unequal recombination and gene conversion are both rare. In the *Dm3* loci, recombination is 18-fold lower than genome average (Chin *et al.*, 2001). Orthologous *Pto* genes (in different species) are more similar than paralogous genes (duplications within a species) suggesting unequal recombination and gene conversion is rare at *Pto* loci (Michelmore & Meyers, 1998). The low frequency recombination is not enough to contribute to divergence in gene clusters. The primary mechanisms generating variation in those *R* gene loci may be different. Therefore, the unequal recombination frequency is an important clue indicating which mechanism influences particular *R* gene cluster evolution.

However, there is still very little information about *R* gene cluster recombination rates so far. Most of the studies come from *R* gene sequence comparisons, or the exchange of RFLP markers. The direct measurement of the recombination frequency within a R gene cluster would be very informative. In particular, it would clarify our understanding of what mechanism is the primary force driving the evolution of a particular *R* gene cluster, i.e. whether it diversifies or homogenizes gene clusters.

1.7. Arabidopsis *RPP1* genes and their recognition specificity

RPP1 specifies resistance to several isolates of the oomycete *Hyaloperonospora parasitica* (downy mildew) in Arabidopsis. The *RPP1* locus maps to the bottom arm of chromosome 3 in the same locus with *RPP10* and *RPP14* (Botella *et al.*, 1998). In Arabidopsis ecotype Wassilewskija, the *RPP1* locus contains four genes and three of them (*RPP1*-WsA, *RPP1*-WsB, *RPP1*-WsC) encode functional NBS-LRR class R proteins, specifying resistance to different isolates of the oomycete *Hyaloperonospora parasitica* (Figure 5). These three

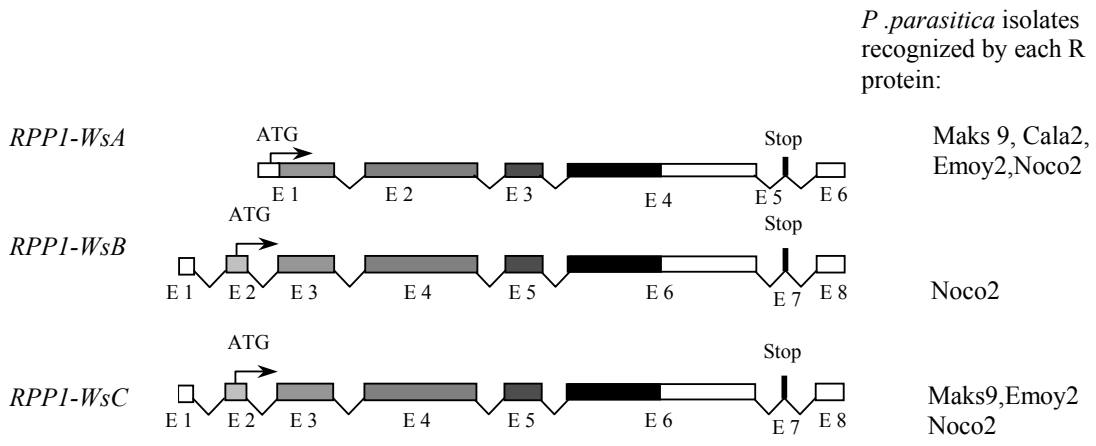


Figure 5: The gene structure of *RPP1* family members and *P. parasitica* recognition specificity (Botella *et al.*, 1998). Exons shown in the same color share high homology.

*RPP1*s share high degree of nucleic acid (>91%) and protein similarity (>85%). The protein structures of *RPP1A (B or C)* and their *Hyaloperonospora parasitica* recognition specificities are shown in Figure 5. These genes are comprised of 6 to 8 exons. The transcription start (ATG) and stop coden were shown on the gene. Because of these characters, three *RPP1* genes were selected to investigate the evolution of *R* gene clusters using synthetic gene cluster technology.

1.8. Using synthetic *RPP1* gene cluster technology to model *R* gene evolution by meiotic unequal crossing over

As discussed above, synthetic gene cluster technology is a method used to investigate meiotic recombination of paralogous genes organized in gene clusters. This method allows rapid and accurate measurement of low frequency unequal crossing-over events (Jelesko *et al.*, 1999; Jelesko *et al.*, 2004). The development of synthetic *RPP1* gene clusters will provide novel opportunities to measure recombination frequencies between different paralogs within *R* gene clusters over the course of a single generation. In contrast to the first generation of synthetic gene clusters, the synthetic *RPP1* gene clusters will utilize *in vivo* site specific recombination to generate alternative alleles from a single transgenic synthetic *RPP1* gene cluster. This feature will enable the unambiguous detection of several different recombination fates that are believed to influence the evolution of R gene clusters. This design should provide unprecedented detailed insights into how unequal crossing-over affects *R* gene cluster evolution.

Figure 6: Structure of the *synthRPP1.1* gene cluster.

Figure 6 shows the *synthRPP1.1* gene cluster that was developed for investigating *R* gene cluster evolution. After being reconfigured *in vivo* to generate two alternative *synthRPP1* alleles, recombination frequencies between *RPP1A* and *RPP1C*, between two allelic *RPP1C*, between *RPP1C* and *RPP1B* will be measured. Because of the opposite direction of *RPP1A* gene relative to *RPP1C::LUC*, the recombination between *RPP1A* and *RPP1C* will result in an inversion as well as two recombinant chimeric *RPP1* genes. The recombination of two non-allelic *RPP1C* genes and the unequal crossing-over between *RPP1C* and *RPP1B* gene represent intra- vs. inter-genic recombination events. Thus, by quantifying the observed recombination rates of intra- and inter-genic recombination, this will provide a quantitative measure of the relative contribution of each type of meiotic recombination to *RPP1* gene evolution.

Another advantage of using these *synthRPP1* gene clusters is that chimeric recombinant *R* genes generated from unequal crossing-over should be easily identified and isolated. The chimeric *RPP1* genes formed by these different recombination events should be readily identified by a gain-of-luciferase phenotype, rather than an altered disease resistance profile. Using oligonucleotide primers flanking the chimeric recombinant *R* gene, long and accurate PCR (LA-PCR) can be used to amplify the chimeric gene. Thus, the chimeric gene can be cloned and sequenced. Based upon its DNA sequence, the recombination resolution breakpoints will be determined by comparing the chimeric *RPP1* genes with their parental sequences. The chimeric *R* gene then can be transformed back into Arabidopsis CW84 to test pathogen resistance specificity, using a large collection of *Hyaloperonospora parasitica* isolates. Thus, a synthetic RPP1 gene cluster could be used to model both unequal crossing-over processes as well as evaluating the chimeric *RPP1* genes production of novel R proteins with altered pathogen recognition specificities.

Chapter Two: Materials and Methods:

2.1. Vector construction:

2.1.1. Subcloning of RPP1-WsA and RPP1-WsB genes

To facilitate the cloning of *RPP1-WsA* gene, an 83-bp cassette containing multiple restriction sites was synthesized by PCR and was inserted into pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA), resulting in plasmid pJS01. The 83-bp cassette was amplified from two oligonucleotide primers: oJS01 (5'GGCGGCCGCCCTCAGCGGTACCGCGA TCGCGAAGTTCCTATTCTCTAGAAAGT-3') and oJS02 (5'-CCCATGGCTCGAGCC GCGGGAAGTTCCTATACTTTCTAGAG-3') that have 14-bp complementary 3' ends. The full-length 83-bp cassette (5'-GGCGGCCGCCCTCAGCGGTACCGCGATCGCGA AGTTCCTATTCTCTAGAAAGTATAGGAACTTCCCGCGGCTCGAGCCATGGG-3') contained *Not*I, *Kpn*I, *FRT*, *Sac*II, *Xho*I, and *Nco*I sites respectively. Genomic *RPP1-WsA* gene was released from plasmid pAMC65-p (Botella *et al.*, 1998) by *Kpn*I. A 7.5 Kb *Kpn*I fragment containing the *RPP1-WsA* gene was cloned into the *Kpn*I site of the subcloned 83-bp cassette (see Figure 7), resulting in plasmids pJS06 and pJS07. Confirmation of both possible orientations was determined by *Xho*I restriction fragments polymorphisms. The ligation boundaries of pJS06 and pJS07 were sequenced using M13 forward and M13 reverse primers to confirm the orientation of the subcloned 7.5 Kb *Kpn*I fragments.

Similar to the 83-bp cassette, a 36 bp cassette (5'-GGCGCGCCGACTATGCATGGCC GGCCCTGAGGTACC-3') was synthesized from two 3' complimentary oligos: oJS03 (5'-GGCGCGCCGACTATGCATGGC-3') and oJS04 (5'-GGTACCTCAGGGCCGGCCAT GCATAG-3') by PCR and cloned into pCR2.1-TOPO vector resulting in pJS02. The cassette contained *Asc*I, *Nsi*I and *Fse*I restriction sites. A 11 Kb *Pst*I fragment containing genomic *RPP1-WsB* gene was cut from plasmid ws-cos69.12 (Botella *et al.*, 1998) and inserted into the *Nsi*I site of pJS02 (see Figure 7). Orientation of the insertion was determinated by sequencing using M13 forward and reverse primers. A plasmid with the orientation of 5'-*Fse*I-*RPP1B*-*Asc*I-3' was selected and named pJS05.

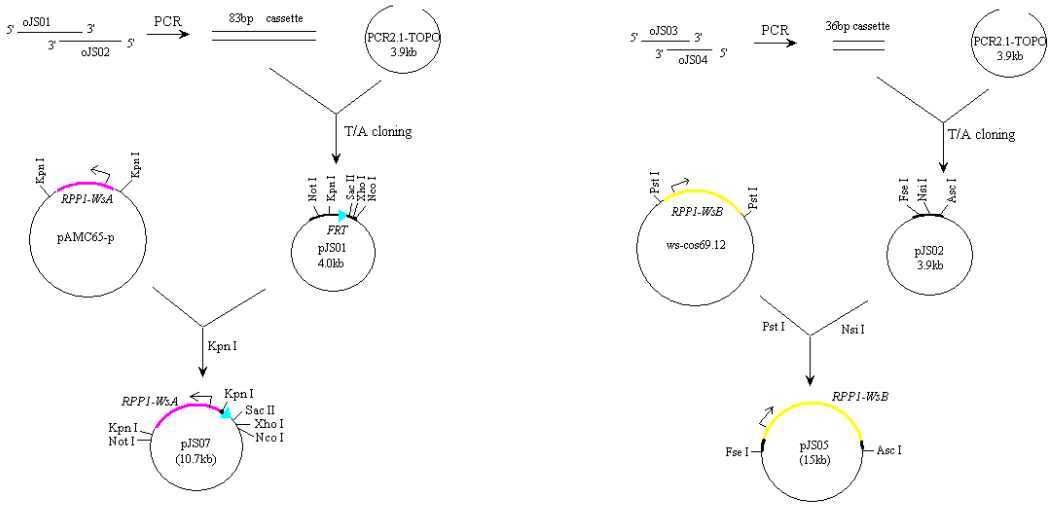


Figure 7: Subcloning of *RPP1-WsA* and *RPP1-WsB* genes:

2.1.2. Site-specific mutagenesis

Plasmid p75C-1 (Botella *et al.*, 1998) was digested with *Sac*II and *Xho*I. A 6.7 Kb *Sac*II-*Xho*I fragment (containing an intact 5'-end of *RPP1-WsC*, but lacking the transcriptional terminator, translational stop codon, and 700 bp of *RPP1-WsC* coding sequence) was subcloned into pBluescript SK(+), producing plasamid pJS10. A 4.0 Kb *BamH*I-*Pst*I subsequence of pJS10 was subcloned into the *BamH*I-*Pst*I sites of pBluescript SK(+), producing pJS11. In order to introduce two *FRT* sites into the *RPP1-WsC* gene, two complementary oligonucleotides oJS09(5'-CGGTGAGATTTTCATATTTTACTGAA GTTCCTATTCTCTAGAAAGTATAGGAACTTCCCTCACTATTTTTAAGTTTTGC-3', underlined is *FRT* sequence)and oJS10 (5'-GCAAAACTTAAAAATAGTGAGGGAAG TTCCTATACTTTCTAGAGAATAGGAACTTCAGTAAAATATGAAAATCTCACCG-3') containing underlined *FRT* sequences flanked by 22-bp and 21-bp sequences of intron 1 of *RPP1-WsC* were used in a QuikChange (Stratagene, Cedar Creek, TX) reactions using plasmid pJS11 as template. After extension by the DNA polymerase in the QuikChange reaction mix, newly synthesized pJS11 acquired a *FRT* sequence in the 5' untranslated leader region (~CATATTTTACT(*FRT*)CCTCACTA~) of the subcloned *RPP1C-WsC* gene.

FRT containing clones were identified by PCR using oJS11 (13 bp at 3' anneals to *FRT*, 5'-GATCAAGATTTAAGATGGCCTTTA-3') and oJS14 (~400-bp downstream, 5'-TCATATTTTACTGAAGTTCCTATTC-3'). A positive clone producing a 420-bp band was sequenced throughout the 2.5 Kb *BamH*I/*Pac*I region spanning the *FRT* insertion to ensure the *FRT* insertion and no additional mutations were introduced by the QuikChange procedure. This plasmid was named pJS12. Fragments with this *FRT* insertion will be denoted by "*".

2.1.3. Construction of a FRT-RPP1C-lox-LUC-lox-nos gene fusion

The region from *Xho*I to the 3' end of the coding sequence of *RPP1-WsC* gene (700-bp long) was amplified by PCR using oligonucleotide primers oJS08 (5'-CCCTCGAGAAA CTAGATTGTTGCT-3', spanning the *Xho*I site as underlined) and oJS06 (5'-CCATGGC AGAAACACCTGAAACCCATTAGGA-3') using plasmid p75C-1 as template DNA. The translation stop codon "TAG" was changed in oJS06 in order to produce an in-frame gene fusion to the luciferase gene in later cloning steps. An additional "C" (red) was added to keep retain an extended *RPP1-WsC* open reading frame. A *Nco*I (underlined) site was also added to facilitate cloning. This *Xho*I-*Nco*I fragment of the *RPP1-WsC* gene was then subcloned into the *Xho*I-*Nco*I site of plasmid pJS0X(?), producing pJS09. The *Not*I-*Nco*I fragment was excised from pJS09 and ligated into the *Not*I-*Nco*I sites of pBW56 (John McDowell, personal communication) that contains 3' *lox-LUC-lox-nos* sequences in-frame with the *Nco*I site of pBW56, thereby producing plasmid pVJ2.

The 2.5-Kb *BamH*I-*Pac*I* fragment from pJS12 was subcloned into the *BamH*I-*Pac*I sites of pJS10, replacing the wild-type *BamH*I-*Pac*I region of *RPP1-WsC* in pJS10, to generate pJS13.

To make the *FRT-RPP1C-lox-LUC-lox-nos* gene fusion, the *Sac*II-*Xho*I fragment of pJS13 was liberated and subcloned into the *Sac*II-*Xho*I sites of pVJ2, producing pJS14.

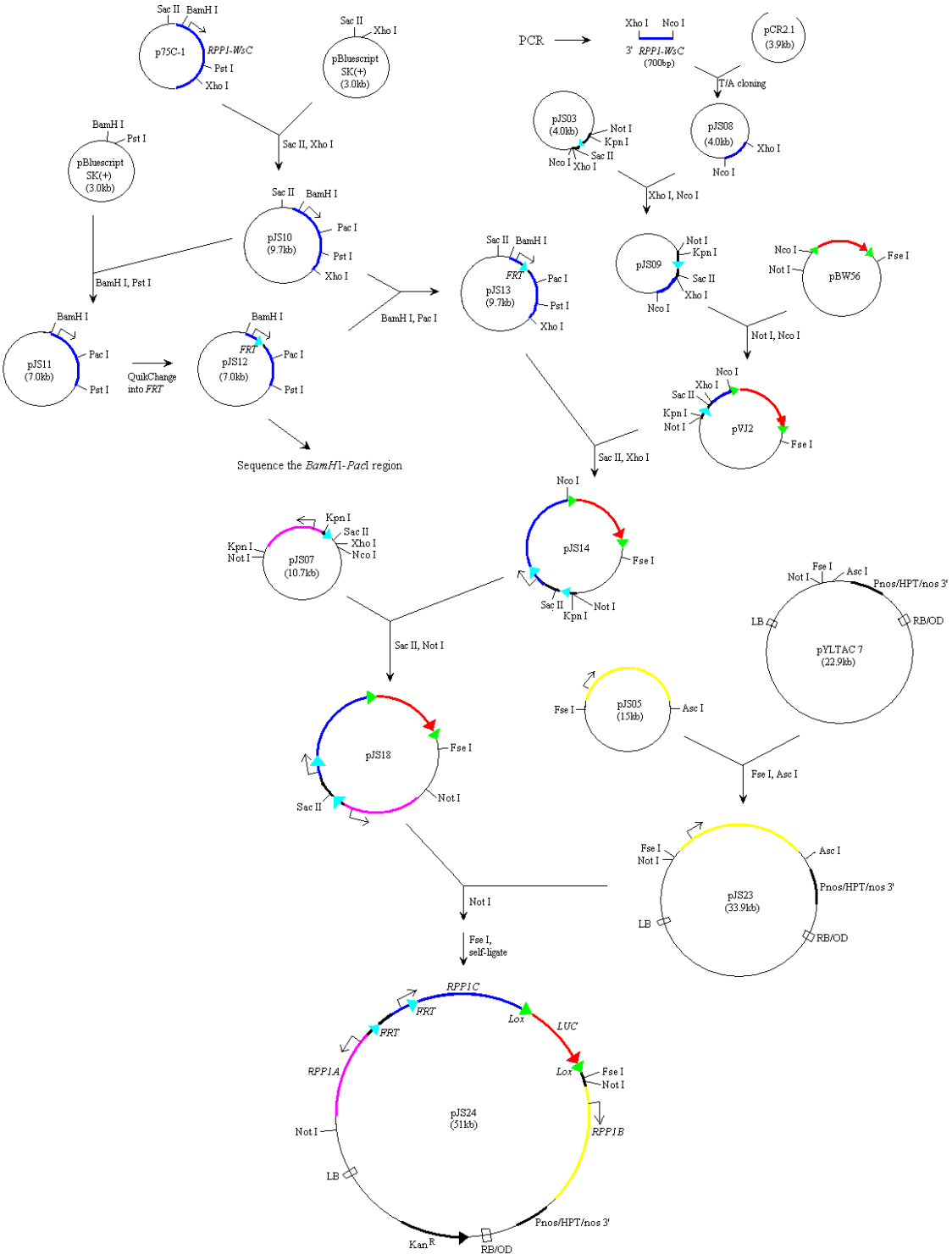


Figure 8: Construction of *synthRPP1* gene cluster and cloning into binary vector pYLTAC7. Shown here is only the construction of pJS24 containing

synthRPP1.1 (*RPP1A* gene is in opposite orientation to *RPP1C:LUC* and *RPP1B* gene). Plasmid pJS26 containing *synthRPP1.2* is not shown, in which *RPP1A* is in the same orientation with *RPP1C:LUC* and *RPP1B* gene.

2.1.4. Construction of synthRPP1.1 and synthRPP1.2 gene cluster

To make the *synthRPP1.1* gene cluster, The *RPP1-WsA* gene was firstly cut out from pJS07 as a *Not*I-*Sac*II fragment and subcloned into the *Not*I-*Sac*II sites of pJS14, producing pJS18 (*RPP1-WsA* is 5' to and in the opposite orientation relative to *RPP1-WsC**). Secondly, *RPP1-WsB* gene was excised from pJS05 as a *Asc*I-*Fse*I and subcloned into the *Asc*I-*Fse*I sites of binary vector pYLTAC7(Liu *et al.*, 1999), producing pJS21. Thirdly, pJS18 was linearized by *Not*I and subcloned into the *Not*I site of pJS21, producing pJS23. Finally, pJS23 was digested by *Fse*I and self-religated to remove the pBluescript SK(+) vector sequence, thus producing the final construct pJS24. The gene cluster on pJS24 was named *synthRPP1.1* (see Figure 8). Similarly, *RPP1-WsA* gene was cut out from pJS07, of which the *RPP1-WsA* gene is in opposite orientation relative to that of pJS06, and was constructed into pJS26 following the exact same procedures as pJS24. Thus pJS26 had the other version of *synthRPP1* with *RPP1-WsA* gene in the same orientation relative to *RPP1-WsC*:LUC* and *RPP1B* gene. The gene cluster on it was named *synthRPP1.2*.

2.2. Generation of transgenic plants:

Plasmids pJS24 and pJS26 were introduced into *Agrobacterium tumefaciens* strain GV3101 by electroporation using MicroPulser™ (Bio-Rad) with parameters of 2.2kV/0.1 cm with a 5 ms pulse.The transformants were selected by kanamycin resistance. Six week-old Arabidopsis plants CW84 were used for Agrobacterium transformation. CW84 is a hybrid of the Col and Ws accessions, that was selected for susceptibility to *Haloparanospora parasitica* isolates isolates Noco2, Emoy2, and Cala2 and Maks9 (described in (Botella *et al.*, 1998)). This line is homozgous for the disease-sensitive alleles at *RPP1/2/4/10/14* loci,

which will facilitate the characterization of recombinant *RPP1* genes recombining within the *synthRPP1* gene cluster. This line was transformed by *Agrobacterium* carrying pJS24 or pJS26 using the floral dip method (Clough & Bent, 1998). The floral dipping was repeated three times at 7-day intervals. Plants were covered with lids to keep moisture over night after each dipping treatment. The plants were grown to maturity and seeds were collected, some of which were comprised of T1 generation plants.

The T1 seeds were surface sterilized and planted onto ½ MS solid media (Sigma-Aldrich, St. Louis) containing 30 μg/ml hygromycin sulfate in petri plates. The plates were wrapped with aluminum foil and kept at 4° C for 2 days, then were exposed to weak light (on work bench) for a couple of hours to promote germination. Then the plates were moved into the dark at room temperature for 5 days to select transformants. Hygromycin resistant transformants have elongated hypocotyls (up to 2 cm) that stand upright on the media, compared with the short hypocotyls of the sensitive seedlings that lay down on the media. The hygromycin resistant T1 seedlings were transplanted into soil SUNSHINE^{TM} Mix (Sun Gro Horticulture Inc., Bellevue, WA) for growth (16 h, 24^{o} C light and 8 hr, 22^{o} C dark).

2.3. Southern blot assay to determine transgene copy number:

Genomic DNA (gDNA) was isolated from the leaves of T1 plants using the cetyltrimethylammonium bromide (CTAB) method originally described by Murray (Murray & Thompson, 1980), but modified by others to use as a routine method in Jelesko's lab). Approximately, 260 ng ~ 1.6 ug of gDNA from each sample was cut with *Pst*I, precipitated with ethanol, and loaded on a 0.8% agarose gel. The samples were run at 1 V/cm for 16 hr to separate the DNA fragments. A 620 bp DIG (deoxydigoxigenin) labeled probe (F. Hoffmann-La Roche Ltd, Basel, Switzerland) specific for the hygromycin gene was amplified by PCR from the plasmid pJS24 using oligos oJS40 (5'-ACCAATGCGGAGCATATACG-3') and oJS41 (5'-GTTTATCGGCACTTTGCATCG-3') and was hybridized to the gDNA at 42 °C. *Kpn*I cuts 5' to the hygromycin gene and then at an unknown site in the gDNA to produce a >3 Kb band. Thus, transgenic lines containing a

single T-DNA insert should produce one hygromycin gene fragment greater than 3 Kb in length.

A left border (LB) hybridization probe was developed based upon the *SacB* gene. *Kpn*I cuts ~1.7 Kb away from the LB boundary (see Figure 9). A hybridization probe targeting the *SacB* gene was generated using oJS42 (5'-ATCAGTTACCGTGTTTGAAG-3') and oJS43 (5'-GACAACTCAATCGACAGC-3') and then used to hybridize to the gDNA as described above. Lines containing single intact transferred T-DNA should produce a single band >1.7 Kb.

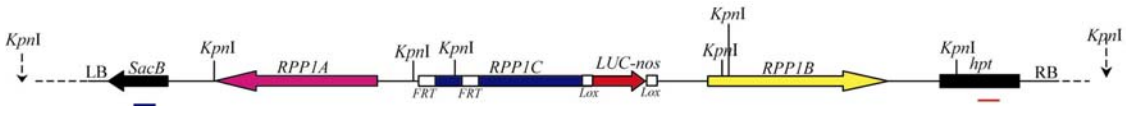


Figure 9: the binding positions of *hpt* and *SacB* gene probes to the *synthRPP1* gene cluster and their resulting hybridization bands. Red bar: *hpt* gene probe. Blue bar: *SacB* gene probe. Dashed lines at both ends of *synthRPP1* were the genomic DNA. The dashed arrows indicate the next *Kpn*I site at unknown position giving rise to random size hybridization bands.

To make sure the transgene has intact *RPP1* genes that didn't undergo gene rearrangement, *RPP1* gene-specific probes were designed and used to check if the corresponding gene were present. Probes that targeting 5'-UTR of *RPP1-WsA*, *RPP1-WsB* and *RPP1-WsC* were generated using oligos oJS44 (5'-CTGAACATATTCTTAAGAACTC-3') and oJS45 (5'-GTCTTTTCATCGCAACTTTTC-3'); oJS48 (5'-GAGGTATGATAATCAGTTCC-3') and oJS49 (5'-CTCCACTGTTCCTAATTTAC-3'); and oJS46 (5'-GAATGCCAGTCTCA ACGG-3') and oJS47 (5'-ACTAGGATGCTTTAATCTCTC-3'), respectively.

2.4. Imaging of seedlings

Arabidopsis seeds were germinated on ½ MS media without sucrose (2 days at 4° C in the dark, followed by 5 days at room temperature under 16 hr light/day). One-week old seedlings were sprayed with 1 mM luciferin solution (containing 0.01% TritonX-100) and were kept in dark for 20 min before imaging. Then the seedlings were imaged using the Hamamatsu VIM50 single photon counting camera (Hamamatsu Co., Hamamatsu City, Japan) for twenty minutes as previously described (Jelesko *et al.*, 1999).

2.5. Plant crosses

Arabidopsis pollen donor and receipt plants containing the corresponding transgenes were grown on soil for six to eight weeks with one plant in each 2-inch pot to make the crossing easier. One cross usually yield ~20 F1 seeds. Each plant had 9 - 12 crosses. The pollen recipient plants were first emasculated. The pollen donor plants did not need any treatment. The flowers of pollen recipient plants were prepared and crossed as follows. When the bolts were 10 - 20 cm tall, three or four of the oldest yet still closed floral buds were emasculated and all other open flowers, immature flower buds, and maturing siliques were removed. The emasculated flower buds were covered with a paper bag or kept the plants in separate space to avoid pollen contamination and the carpels were allowed grow two more days before pollination, until the papillae of the stigma fully expanded. When the papillae were fully expanded, open flowers from the pollen donor plants containing anthers shedding pollen were touched to the emasculated carpels of the pollen recipient plants. The pollinated flowers were covered with paper bags again or moved to a clean space to let them grow. Two to three days later, the fertilized carpel can be seen elongating, indicating a successful cross. The paper bags were then removed and the plants were allowed to grow to maturity. The seeds harvested from the fertilized siliques were defined as F1 seeds.

2.6. *In vivo* reconfiguration of *synthRPP1* by FLP or CRE to generate alternative alleles

The primary transgenic lines AtJS24.13 and AtJS24.33 containing transgenic *synthRPP1* gene clusters were crossed to transgenic Arabidopsis lines expressing either the FLP or CRE recombinase in order to reconfigure the *synthRPP1* gene cluster *in vivo* into two alternative luc$^-$ alleles, *synthRPP1* $^{\Delta RPP1C\,\mathrm{prom}}$ and *synthRPP1* $^{\Delta LUC}$ respectively (see Figure 13). The FLP recombinase expressing line was provided by Kilby (1995), which is the *FLP* gene driven by soybean *Gmhsp* 17.6L heat-shock inducible promoter. The CRE expressing line was CW84 transformed with plasmid pCLG9. The homozygous line AtCLG9.5b.4 was used in this experiment. Transgenic lines AtJS24.13 and AtJS24.33 were used as pollen donors, and FLP (CRE) expressing plants were used as pollen recipients in the crosses. F1 seeds were collected.

To generate the *synthRPP1* $^{\Delta RPP1C\,\mathrm{prom}}$ allele, AtJS24.13 × FLP F1 seeds were germinated and grown in soil SUNSHINE^{TM} Mix under 8 hr florescent light, 16 hr dark conditions (short days). When the F1 plants had inflorescences with 5 - 8 flower buds, the open flowers were removed and the remaining buds were heat shocked at 37° C for 4 hr on 5 successive days to induce FLP gene expression. The flowers were allowed to mature and the F2 seeds were collected. Then the F2 plants were screened by PCR to identify the FLP reconfigured form of *synthRPP1* (*synthRPP1* $^{\Delta RPP1C\,\mathrm{prom}}$) allele. The PCR was performed using two oligos: oJGJ201 (5'AGAGTTCCAAAGTCCAGCAAACAACAACAAACAA CAACGG-3') and oJGJ202 (GAACAGCTCAAGCTCATTGCAGAACCCATGAGAAC TTAGA) with a crude gDNA preparation (a routine quick plant gDNA extraction method used in John McDowell's lab. One flower bud or one small leaf is ground in extraction buffer: 200mM Tris pH7.5, 25mM EDTA pH7.5, 250mM NaCl, 0.5% SDS. After spinning down the plant tissue debris, the supernatant is precipitated by equal volume of isopropyl alcohol for 10 min, the resultant pellet is then dissolved in water) from the F2 plants leaves. The presence or absence of *FLP* gene was determined by PCR using oligos oJS103 (5'-ACGGAACAGCAATCAAGAGAGC-3') and oJS104 (5'-AGTCAACTCCGTTAGGCC

CTTC-3'). The *FLP* gene produces an 800 bp PCR product. Homozygous *synthRPP1* $^{\Delta RPP1C\,\text{prom}}$ lines lacking the *FLP* gene were identified in the F3 population of F2 plants that produced a 300 bp *synthRPP1* $^{\Delta RPP1C\,\text{prom}}$ specific band using the same PCR procedure. The results were confirmed by another PCR using oJS118 (5'-CAGTGAATGAATGATCTTC CTGATTCTAAGAGTGTCAATAGT-3') and oJS119 (5'-GTGATAAAATCTTAAGCAA AACTTAAAAATAGTGAGGGAAGTTCC-3'), which produces a 1.1 Kb band from primary non-reconfigure *synthRPP1* allele, but no amplification from the reconfigured allele.

Similarly, the T2 plant of AtJS24.13 was crossed to the homozygous CRE expressing line AtCLG9.5b.4 to develop the *synthRPP1* $^{\Delta LUC}$ allele *in vivo*. The F1 seeds were germinated and grown as above. The F2 population was screened as follows. Two oligos, oJS107 (5'-GGAGTCGCCTTTTCCTACAACC-3') and oJS108 (5'-AATGTAGCTCGTTGTTG GACCG-3'), flanking the two *lox* sites were used to amplify in the crude gDNA from F2 plants leaves. The primary non-reconfigured *synthRPP1* allele produced a 2.7 Kb band and the CRE reconfigured allele produced a 1.0 Kb band. The presence or absence of *CRE* gene was determined by PCR using oJS105 (5'-GTCGATGCAACGAGTGATGAGG-3') and oJS106 (5'-CCATGAGTGAACGAACCTGGTC-3'), which produced an 800 bp band from *CRE* gene. Homozygous *synthRPP1* $^{\Delta LUC}$ lines were selected by hygromycin selection and southern blot in F2 population. Probe that targets 5'-UTR of *RPP1-WsB* was generated by PCR using oJS48 (5'-GAGGTATGATAATCAGTTCC-3') and oJS49 (5'-CTCCAC TGTTCCTAATTTAC -3'). Genomic DNA was digested with *Nco*I. The non-recombined allele produced 3.5 kb band and recombined allele produced 1.8 kb band.

2.7. Development of *in vitro* recombined *synthRPP1*$^{\Delta 1C\,promoter}$ allele

Plasmid pJS24 and pJS26 were incubated with purified FLP recombinase (a gift provided by Dr. Paul Sadowski at University of Toronto) to produce *synthRPP1*$^{\Delta RPP1C\,prom}$ *in vitro*.

One microgram of plasmid DNA was incubated with purified FLP recombinase in the reaction buffer (50mM Tris-Cl pH7.4, 10mM $MgCl_2$, 30mM NaCl) at 30^{o} C for 15 min, then heated to 67^{o} C for 10 min to deactivate the FLP enzyme. The successful recombination of pJS24 and pJS26 was confirmed by PCR using oJGJ201 and oJGJ202 as above. The recombined pJS24 (named pJS27) and recombined pJS26 (named pJS28) were used to transform *E. coli* DH5 α . The resultant clones were screened by colony PCR using oJGJ201 and oJGJ202 to select against non-recombined plasmid. Plasmid pJS27 and pJS28 DNA were then extracted from *E. coli* and used to transform *A. tumefaciens* strain GV3101 by electroporation, as described above. Finally the Agrobacteria containing corresponding plasmid were used to transform Arabidopsis CW84 plants by floral dipping, as outlined above.

Transgenic T1 plants containing *in vitro* FLP recombined gene cluster *synthRPP1*$^{\Delta 1C\ promoter}$ from above pJS27 were screened for single copy transgenic lines as described in section 2.3 by southern blotting. The single copy transgenic lines were allowed to self creating subsequent T2 and T3 generations, respectively. Selection of homozygous T2 lines was conducted by examining the hygromycin resistance segregation ratios in the T3 generation.

2.8. Development of hemizygous *synthRBCSB-1* seed population

To compare the meiotic recombination between homozygous and hemizygous *synthRBCSB1* (Jelesko *et al.*, 2004) populations, the homozygous T3 transgenic Arabidopsis line *synthRBCSB1.10* (also called AtJGJ203.10) was crossed to Arabidopsis Columbia-0. More than 200 crosses were done to produce around 3,000 F1 seeds. The resultant F1 seeds were hemizygous for *synthRBCSB1* locus. At the same time, homozygous T3 plants were allowed to self-fertilize to produce matched homozygous T4 seeds.

The hemizygous F1 seeds and homozygous T4 seeds were germinated and grown in a SUNSHINE MIX 1® potting soil in Conviron® environmental growth chamber (Manitoba, Canada) with 16 hr light, 24^{o} C and 22^{o} C dark), producing about 10 million seeds for each line.

2.9. Luciferase activity screening of *synthRBCSB-1* homo- and hemizygous lines

Seeds were sowed on 20 × 20 cm chromatography paper wetted with 1X Hoagland's solution in 25 × 25 cm^2 flats, then were cold treated at 4^{o} C for 2 nights before being placed onto growth racks for germination (16 hr light, 24° C days and 22° C in dark). Five-day old seedlings were sprayed with 1 μM luciferin with 0.01% Triton-X100. Twenty minutes later, the seedlings were imaged using the Hamamatsu VIM50 single photon counting camera as previously described (Jelesko *et al.*, 1999).

2.10. Cloning meiotic recombinant *RBCSB* genes

Luciferase positive seedlings identified in section 2.10 were isolated and transplanted onto soil and grown to maturity. Genomic DNA was isolated from the leaves or flowers for amplifying the recombinant chimeric *RBCSB* genes using oJGJ13 (5'-CAAAAGAAA GATAAGATAAGGGTGTCAA-3') and oJGJ14 (5'-CCTTTCTTTATGTTTTTGGCG TCTTC-3') in PCR reactions. The resultant PCR products were then subcloned into pCR2.1-TOPO® vector for sequencing.

2.11. Amplifying recombinant *RBCSB* genes from somatic tissues

Genomic DNA was purified from homo- and hemi-zygous *synthRBCSB1.10* (see 2.8) leaves, stems, and flower buds using the CTAB method (Murray & Thompson, 1980). Nested PCR reactions using primer pairs oJGJ13 & oJGJ14 followed by oligonucleotide primers oJS112 (5'- AATCCTGTGGCAGTAAACGACG -3') and oJS113 (5'-GTTTTT GGCGTCTTCCAGCTTG-3') were used to amplify the recombinant genes. The first round PCR used oJGJ13 and oJGJ14 (94^{o} C, 2 min; 94^{o} C, 30 sec, 53^{o} C, 30 sec; 72^{o} C, 1.5min for 30 cycles; with a 72^{o} C, 2 min final extension). One ul of this was used in a second round of PCR using oJS112 with oJS113 (94 ^{o}C, 2 min; 94^{o} C, 30 sec, 57^{o} C, 30 sec, 72^{o} C, 1 min, for 35 cycles; with a 72^{o} C, 2 min final extension). The PCR products were cloned into pCR2.1-TOPO® vector for sequencing at the VBI Core Facility (Blacksburg, VA).

2.12. Metagenomic analysis:

In order to determine which *RBCSB* gene recombined with *RBCS1B:LUC* gene fusion and to map the recombination resolution sites, the DNA sequences of the recombinant *RBCSB* genes were aligned with their parental genes (*RBCS1B*, *RBCS2B* and *RBCS3B* of ecotype K85). The sequences of these genes (also called *ATS1B*, *ATS2B* and *ATS3B* genes) can be found in Genbank (accession X14564). The recombinant gene sequences were aligned with *RBCS1B/2B/3B* genes using program MegAlign® of Lasergene® using the Clustal W method. Polymorphic nucleotides were used to determine the recombination templates and recombination resolution sites were estimated as previously described (Jelesko *et al.*, 1999).

2.13. Quantitative real-time PCR to detect chimeric *RBCSB* genes formed during somatic recombination

A nested PCR approach was used to detect chimeric *RBCSB2B/1B:LUC* and *RBCS3B/1B:LUC* genes that formed by somatic recombination in various plant tissues. Genomic DNA was isolated from leaves, stems, or flowers of either hemizygous F1 or homozygous T4 *synthRBCSB1.10* plants (see 2.8) and was used in a first round of PCR using oJS123 (5'-GCGACCTTATTTCCAAAGAAAGTG-3') and oJGJ14 (94^{o} C 2 min; 94 ^{o} C 20 sec, 53^{o} C 20 sec, 72^{o} C 2 min, 20 cycles; 72^{o} C 2 min). Then 2 ul of the first PCR reaction mixture was used as template in Quantitative Real Time PCR reactions using primers oJS112 and oJS113 with the SYBR® green master mix (Applied Biosystem, Foster City). Genomic DNA from a *synthRBCSB* recombinant luc^{+} plant XJS1.2 (containing a stable inherited *RBCS3B/1B* recombinant gene, see chapter 5) was used to make a 10-fold serial dilution series to serve as standard curve (6.5×10^{5} - 10^{2}). The total genomic DNA input of the dilution series were added to 110 ng of Col-0 genomic DNA. The sample genomic DNA inputs were also 110 ng. The QRT-PCR reactions were performed in ABI 7500 (Applied Biosystem, Foster City) real-time PCR system with the following amplification conditions: 50^{o} C, 2 min; 95^{o} C, 10 min; 95^{o} C, 20 sec denaturation, 57^{o} C, 20 sec annealing, 60^{o} C, 2 min extension, 79 ^{o}C, 40 sec fluorescent measurement for a total of 40 cycles. Fluorescence data were collected at the 79^{o} C, for 40 seconds step. Using XJS1.2 genomic DNA as standard, the Ct value of each sample can be converted into the absolute recombinant genome copy number in each sample. The lower the Ct value, the more recombinant genome copies present. The redundancy of the recombinant genome in the sample (i.e. recombination rate) can be calculated by recombinant gene copy number divided by the total genome copy number calculated from the total initial input genomic DNA quantity (estimated from a UV Nanodrop spectrophometer (Coleman Technologies, Inc., Orlando, FL).

The total haploid genome equivalents (haploid gDNA copy number) were calculated by [(starting gDNA)/(0.17pg per genome)] for hemizygous lines and twice that for

homozygous lines. The recombination frequency was calculated by (Quantity/haploid gDNA copy number).

Chapter Three: Development of synthetic *RPP1* (*synthRPP1*) transgenic Arabidopsis lines

3.1. Introduction:

Various studies indicate that plant disease resistance genes organized within clusters can recombine to produce recombinant *R* genes with altered or new recognition specificities (Dixon *et al.*, 1996; McDowell *et al.*, 1998; Hulbert, 1997; Dixon *et al.*, 1998). Studies of the maize *rp1* locus provided good evidence for recombination playing an important role in creating genetic diversity. The *Rp1* complex is a cluster of rust (*Puccinia sorghi*) disease resistance genes in maize (Hulbert, 1997). Sixteen different genes at the *rp1* locus were identified by using an extensive collection of rust biotypes. In this locus, both intergenic gene reassortment and intragenic cross-over events were revealed by restriction fragment length polymorphism (RFLP) analysis. However, some recombinant progeny showed resistance to at least one rust biotype to which both parents were susceptible (Hulbert, 1997).These results suggested that physical changes at the *rp1* locus resulted in novel pathogen resistance specificities and are consistent with the predicted effects of unequal crossing-over between parental *rp1* genes resulting in chimeric *R* genes with new rust recognition specificities. Unfortunately, few other molecular details of the recombinant *rp1* loci were characterized.

Parniske *et al* (1997) also reported novel disease resistance specificities generated by sequence exchange between tandemly repeated *Cf* genes in tomato that confer resistance to the fungus *C. fulvum*. The*Cf-4* and *Cf-9* are closely related *R* genes, present at the approximately same locus. Interestingly, the *Cf* genes in *Cf-4/9* transheterozygous tomato show increased meiotic instability, suggesting that non-allelic genes at the same locus might stimulate meiotic unequal crossing-over and novel chimeric *R* genes with new resistance specificities.

Sequence comparison of the *RPP8* locus in different ecotypes of Arabidopsis also suggests that recombination may be an important mechanism for generating new *R* gene resistance specificities (McDowell *et al.*, 1998). The *RPP8* locus of Arabidopsis ecotype Landsberg erecta (Ler-0) is comprised of a tandem array of two *R* genes (McDowell *et al.*, 1998). The *RPP8*-Ler gene is a functional *R* gene conferring resistance to *Haloparanospora parasitica* isolate Emco5. The second *R* gene (*RPH8A*) is quite similar to *RPP8*-Ler, but its function has not been demonstrated. In contrast, the *RPP8* locus in Col-0 contains a single chimeric gene (McDowell *et al.*, 1998) at this locus. Moreover, the Col-0 chimeric *RPP8* gene was likely derived from a prior unequal crossing-over event between *RPP8*-Ler and *RPH8A* ancestors within a Ler-like haplotype (McDowell *et al.*, 1998). Similar Col-0 –like gene configurations are found at the *RPP8*/*HRT* locus in Dijon-17 (Cooley *et al.*, 2000) and C24 (Takahashi *et al.*, 2002). Each *RPP8/HRT* loci contains one gene that seems to be a chimera of two ancestors and have completely different pathogen recognition specificities. The *RPP8/HRT* locus in Dijon-17 confers resistance to both viral (Turnip crinkle virus) and oomycete pathogens. On the other hand, the *RPP8/HRT* locus in C24 confers resistance to a yellow strain of cucumber mosaic virus. These data suggest that in spite of the high sequence similarity, the sequence exchange between closely related *R* genes can generate novel recognition specificity to completely different pathogens. Although the chimeric structure of the *RPP8/HRT* locus is consistent with unequal crossing-over, there are additional polymorphic bases that arose from random mutation events. Thus, one can not conclude that recombination alone was responsible for generating the new resistance specificities at the *RPP8/HRT* locus. To differentiate which of these two mechanisms is responsible for generating new pathogen resistance specificities, it would require the identification of primary recombinant loci and testing these for altered/novel pathogen recognition specificities.

It is believed that the pathogen recognition specificities of plant R proteins reside in LRR domains. Thus, recombination between two *R* genes within the DNA region corresponding to LRR domains would likely be significant in *R* gene evolution. Consistent with this hypothesis, alterations in LRR repeat numbers correlate with altered resistance specificities. In the tomato *Cf-2/Cf-5* locus, six of the seven homologs vary in LRR copy number (Dixon

et al., 1998) and pathogen recognition specificity (Dixon *et al.*, 1998). Thus, it is inferred that unequal crossing-over between LRR coding regions resulted in different LRR copy number and concomitant different pathogen recognition specificities.

However, none of the above studies provide direct evidence of meiotic recombination generating a particular novel recognition specificity from one round of recombination. Nor did these studies provide important details about relative recombination frequencies of intra- vs. inter-genic recombination events. This state of affairs is partially due to the difficulties associated with identifying recombinants. For example, not all recombinant alleles may result in readily observable phenotypes, particularly when recombination results in gene duplications, because plants are generally not susceptible to increased gene dosage effects. Likewise, if a new pathogen resistance specificity is formed but that pathogen is not included in a phenotypic screen, then one would overlook/underestimate both the frequency and impact of unequal crossing-over.

Therefore, a new method needs to be employed to identify recombinant loci that are independent of the function and phenotype of the recombinant gene. To this end, synthetic gene cluster technology offers several advantages over phenotypic screening approaches. Briefly, a synthetic gene cluster is comprised of an inactive (usually a deleted promoter) target gene fused in-frame to a reporter gene with one or more expressed paralogous genes linked together (Figure 10). During meiosis, a rare misalignment and unequal cross-over between paralogous genes (gene 1 and gene 2 in Figure 10) yields a

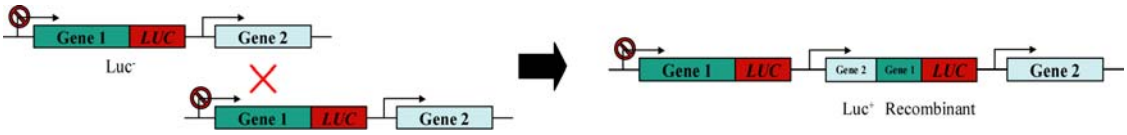


Figure 10: Use synthetic gene cluster technology to screen for recombinants:

"recombinant gene 2/gene 1-reporter gene" fusion driven by an active promoter derived from the gene 2. This chimeric reporter gene fusion will confer a luciferase positive phenotype (bioluminescence) in seedlings that inherit the chimeric gene. Thus, luciferase activity becomes a proxy for the formation of a novel chimeric gene-reporter gene fusion. This approach effectively identifies chimeric gene formation by unequal crossing-over without assaying for the biological activity of the original gene product. Thus, one can isolate chimeric genes of interest and then subsequently test the phenotype of the novel chimeric genes. This approach enables the whole spectrum of possible functional outcomes: no change, loss-of-function, altered function, and gain-of-function. This approach was first described by Jelesko *et al.* (1999) utilizing a synthetic *RBCSB* (*synthRBCSB1*) locus in Arabidopsis.

3.2. Experimental Design:

3.2.1. Introduction to RPP1 genes

This study used three plant disease resistance genes from the *Arabidopsis thaliana* ecotype *Ws-0 RPP1* locus to construct a synthetic *R* gene cluster to investigate meiotic recombination. They share a high degree of sequence similarity (91-92% of nucleotide sequence identity and 84-87% of protein sequence identity). The endogenous *RPP1-WsC* gene is 5' of *RPP1-WsB* gene and they are organized as a tandem array. On the other hand, *RPP1-WsA* is 5' of *RPP1Ws-C* on the genetic linkage map. However, neither the distance nor the orientation of *RPP1-WsA* relative to *RPP1Ws-C* is currently known. The pathogen recognition specificities of each of the *RPP1* genes were determined with a panel of *Hyaloperonosporaparasitica* isolates. They have overlapping but non-identical resistance specificities. *RPP1-WsA* recognizes isolates Maks9, Cala2, Emoy2 and Noco2. *RPP1-WsB* only recognizes isolate Noco2. *RPP1-WsC* recognizes isolates Maks9, Emoy2 and Noco2 (Botella *et al.*, 1998).

There are several reasons why *RPP1* genes were chosen for this study. First, the *RPP1* genes in Ws-0 show active evolutionary dynamics. Molecular analysis suggested that the three homologous genes arose from gene duplication events following by sequence divergence (Botella *et al.*, 1998). There is evidence of sequence exchange between *RPP1* genes and this sequence shuffling could be an important evolutionary mechanism to create new recognition specificity from existing *RPP1* genes. In particular, the predicted *RPP1* gene products are TIR-NBS-LRR encoding genes that have a distinct C-terminal leucine-rich repeats (LRRs) and a central nucleotide-binding site with an additional N-terminal domain that is similar to Toll/interleukin-1 receptor (TIR). Previous studies have showed that recombination in LRR regions created new recognition specificity (Hulbert, 1997; Parniske *et al.*, 1997). Secondly, the *RPP1* genes are several kilobases long and share a high degree of sequence identity, and therefore should provide efficient recombination substrates. Sequence identity is a critical factor for homologous recombination (Chen & Jinks-Robertson, 1999; Metzenberg *et al.*, 1991). Besides high sequence identity, there are also many sequence polymorphisms, which allow resolution of recombination breakpoints (Gal *et al.*, 1991; Jelesko *et al.*, 1999). Thirdly, the pathogen recognition specificities of these three genes were determined and they recognize overlapping subsets of *Hyaloperonospora parasitica* isolates. These overlapping pathogen resistance specificities should allow new insights into the structure function relationships of chimeric *RPP1* gene architecture and pathogen resistance specificity.

3.2.2. A RPP1C:LUC gene fusion confers luciferase activity in transgenic Arabidopsis plants

Synthetic gene cluster technology identifies chimeric genes that form by *in vivo* recombination based upon a gain-of-luciferase activity phenotype. The firefly luciferase gene fused to the Agrobacterium nopaline synthase transcriptional terminator (*LUC-nos*) was used as reporter gene in this study because it offers excellent sensitivity when used in conjunction with sensitive photonics. For several reasons that will be outlined later, it was

desirable to use the *RPP1C* gene as a target for recombination. Therefore, an in-frame *RPP1C:LUC* gene was created, in which a fully functional *RPP1C* gene was fused in-frame to the *LUC-nos* gene positioned as a 3' gene fusion. This fusion gene was used to construct *synthRPP1* gene cluster (discussed later). When expressed in stable *synthRPP1* transgenic Arabidopsis plants (AtJS24), the *RPP1C:LUC* gene fusion resulted in plants displaying *in vivo* luciferase activity (luc^+), see Figure 11 and Figure 18 below. This was an important finding because it demonstrated that a full length RPP1C-LUC chimeric protein would retain luciferase activity. These results also provided reasonable assurances that chimeric *RPP1* genes fused to *LUC* should also result in a luc^+ phenotype, thereby making the genetic screen feasible. This result was important for this whole project because it enables the screening of recombinant *RPP1* genes based upon the gain-of-function of luciferase activity.

3.2.3. Construction of a reconfigurable synthetic RPP1 gene cluster

Using the *RPP1C:LUC* construct as a target gene for recombination, a novel reconfigurable synthetic *RPP1* gene cluster (*synthRPP1.1*) was generated to investigate meiotic recombination.

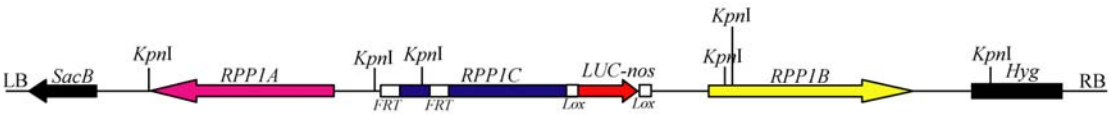


Figure 11: T-DNA region of pJS24 showing synth*RPP1.1* gene cluster:

The three *RPP1* genes were assembled into a novel reconfigurable synthetic gene cluster, designated synth*RPP1.1*. Figure 11 illustrates the composition and organization of *synthRPP1.1*. A detailed explanation of how the *synthRPP1.1* locus was assembled from existing *RPP1* genes can be found in Chapter 2 Materials and Methods. The objective was to make a synthetic *RPP1* gene cluster that reflected the wild type organization of the endogenous *RPP1* gene cluster (i.e. gene order and, when known, orientation). However,

the intergenic regions between them were substantially shortened to several hundred base pairs to facilitate the cloning and transformation into plants. The *synthRPP1.1* gene cluster contained the intact *RPP1-WsC* gene (including the *RPP1C* promoter) fused in-frame to the Firefly luciferase gene. Transgenic plants containing the *RPP1C:LUC* transgene resulted in luciferase positive seedlings. The *RPP1B* gene was cloned 3' to *RPP1:LUC* in the same orientation. In contrast, the *RPP1A* gene was positioned 5' to *RPP1:LUC*. Because the precise orientation of the *RPP1A* gene relative to the *RPP1C* gene is not known, two *synthRPP1* gene clusters were generated, each differing in their relative orientation of *RPP1A* to *RPP1C:LUC*. Plasmid pJS24 has the *RPP1A* and *RPP1C:LUC* genes oriented in opposite directions (*synthRPP1.1*), whereas pJS26 has them oriented in the same orientation (*synthRPP1.2*). Figure 11 shows the *synthRPP1.1* locus. The reason for constructing two types of *synthRPP1* was in part because we do not know the endogenous orientation. Moreover, by having two *synthRPP1* loci that differ in the relative orientation between *RPP1A* and *RPP1C:LUC* we will be able to model quite different types of genomic recombination between these two genes (i.e. deletion/insertions from direct orientation versus inversions generated from the inverted orientation). Clustered *R* genes are often found in inverted orientation, so it is important to understand how this arrangement affects the dynamics of recombination.

Before explaining the novel attributes of the *synthRPP1* gene clusters, it is important to discuss the complexity of possible of intra- and inter-genic as well as intra- and interchromosomal recombination that can occur between three genes organized as a gene cluster. Figure 12 illustrates two distinct types of unequal crossing-over within the simplest type of gene cluster: a tandem gene duplication.

Different types of meiotic recombination may produce the same chimeric genes, making the assessment of the different types of recombination difficult. For example, recombination between sister chromatids can give rise to the same chimeric gene as recombination between the same genes located on homologous chromosomes. The only way to distinguish between these recombination processes is to recover all the gametes produced by a single recombination event. While this is readily feasible for many species

of yeast that package all haploid gametes in a single package, it is not feasible for most other eukaryotes that disperse the four meiotic products during gamete development.

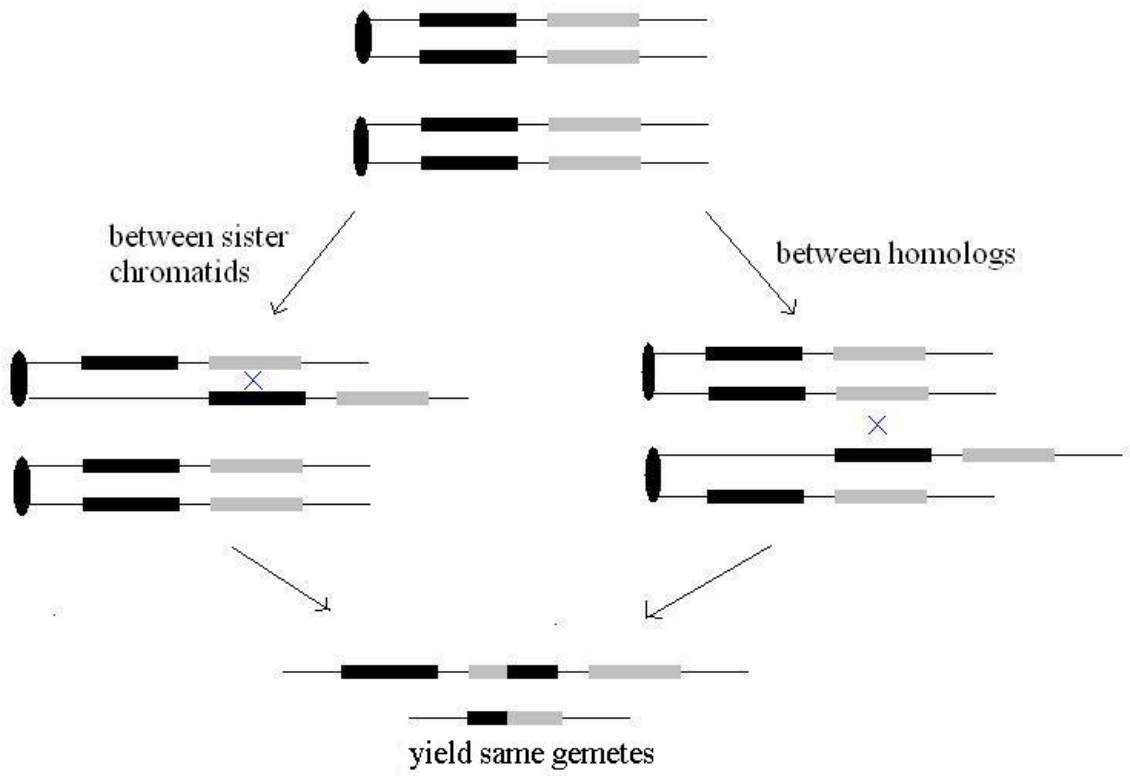


Figure 12: unequal crossing-over between sister chromatids and between homogous chromosomes yield the same type of gemetes.

However, if one can differentially mark each of the gene clusters residing on each of the homologous chromosomes, then one can determine whether a specific chimeric gene formed by recombination between sister chromatids or between homologous chromosomes. Therefore, a mechanism for generating two alternatively marked alleles for the target *RPP1C:LUC* gene was employed. Moreover, the alternatively marked *RPP1C:LUC* alleles must be at the same chromosomal location. This approach is superior to making two alternatively marked *synthRPP1* gene clusters *in vitro* and make two independent transgenic lines, because each of the randomly inserted transgenes would be located on

different chromosomes or in different locations on the same plant chromosome and therefore would not properly align during meiosis.

Therefore, an alternative approach was used that allowed the *in vivo* reconfiguration of a single *RPP1C:LUC* transgene into two alternative marked alleles, both at the same chromosomal location. The *in vivo* reconfiguration can be performed using two different site-specific recombination systems (*FRT*/FLP and *lox*/CRE) to develop two alternative alleles from one primary single copy transgenic line (see Figure 13). In my *synthRPP1*

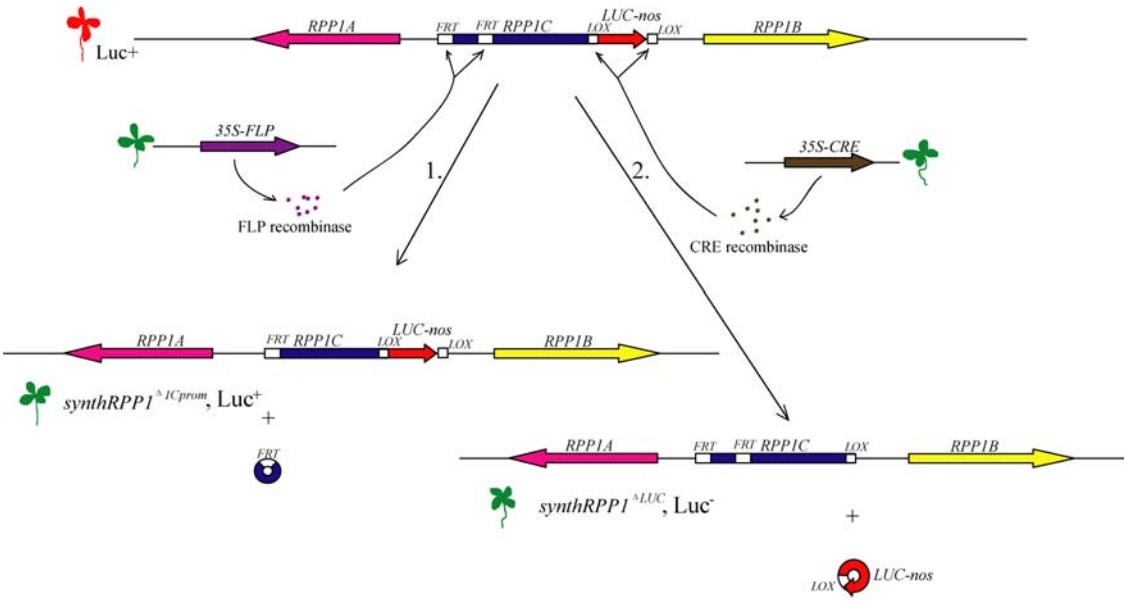


Figure 13: *In vivo FRT*/FLP and *lox*/CRE site specific recombination to reconfigure *synthRPP1* to generate alternative alleles *synthRPP1* $^{\Delta 1Cprom}$ and *synthRPP1* $^{\Delta LUC}$. (1): The primary transgenic lines were crossed with homozygous *35S-FLP* transgenic lines. In the F1 plants, FLP recombinase was expressed and acted on the two *FRT* sites to slice out the flanked sequence, producing allele, *synthRPP1* $^{\Delta 1Cprom}$. (2): Similarly, the primary transgenic lines were crossed with homozygous 35S-CRE transgenic lines to generate the other alternative allele, *synthRPP1* $^{\Delta LUC}$.

constructs, two *FRT* sites oriented in direct repeats 3.3 kb apart flanking the promoter of *RPP1-WsC* gene were introduced. In the presence of the FLP recombinase enzyme, the DNA segments bounded by the *FRT* sites will be excised as a circular DNA molecule, resulting in the deletion of the *RPP1C* promoter and inactivation of the *RPP1C:LUC* reporter gene. The FLP/*FRT* recombinase system was previously shown to function in plant cells enabling *in vivo* FLP/*FRT* recombination (Luo *et al.*, 2000; Lyznik *et al.*, 1996; Davies *et al.*, 1999). Therefore, by crossing a *synthRPP1* gene cluster transgenic line to a line that produces FLP enzyme one could produce an $RPP1C^{prom-}$ allele, resulting in a luciferase negative phenotype (luc^-). The reconfigured $RPP1C^{prom-}$ allele will enable the $RPP1C^{prom-}$*:LUC* gene serve as a target for homologous recombination with *RPP1A* or *RPP1B* genes.

The second alternative allele was generated by a different site-specific recombinase system called CRE/*lox*. CRE is a site-specific recombinase derived from Phage P1 and acts upon two 32 bp imperfect repeats called *lox* sites. Two *lox* sites were introduced that flank the luciferase gene (*LUC*). Specifically, the linker region between the in-frame *RPP1C:LUC* contained a *lox* sequence such that 15 amino acids were added to the *R* gene *LUC* fusion. Another *lox* site was incorporated downstream of the *LUC* stop codon in direct repeat orientation (i.e. *RPP1C-lox-LUC-lox*). Thus, in the presence of the CRE recombinase, the entire *LUC* coding region will be excised into a circular DNA molecule that lacking an origin of replication would be lost during cell division, resulting in a luc^- phenotype. As with the FLP recombinase system, crossing a *synthRPP1* line with a transgenic line expressing the CRE enzyme, should reconfigure the *synthRPP1* gene cluster into the $synthRPP1^{\Delta LUC}$ allele, conferring a luc^- phenotype (see Figure 13).

The ability to generate alternative marked *synthRPP1* alleles will allow the identification of a variety of meiotic recombination events. Initially, primary transgenic lines should have active *RPP1-WsC-LUC* gene fusion showing a luciferase positive phenotype (luc^+). Once it is reconfigured *in vivo* by either deleting the *RPP1-WsC* promoter ($RPP1C^{\Delta prom}$ allele) or the *LUC* gene ($synthRPP1^{\Delta LUC}$), the plant will show a luciferase negative phenotype (luc^-). Thus, crossing plants containing the $synthRPP1^{\Delta LUC}$ allele (luc^- phenotype) with those

containing the *RPP1C*$^{\Delta prom}$ allele (luc^{-} phenotype) will result in a heterozygous population for each of the alternative alleles. These plants would then be capable of several types of meiotic recombination resulting in a luciferase positive phenotype (see Figure 14). One is unequal crossing-over between sister *synthRPP1*$^{\Delta 1Cprom}$ chromatids (scenario A, in Figure 14), resulting in a recombinant *RPP1-LUC* fusion gene linked with a promoterless *RPP1C:LUC* fusion gene. The second type of recombination is homologous recombination between two *RPP1C* genes, reconstituting the original *RPP1C:LUC* luc^{+} primary transgene (senario B. in Figure 14). The third type is unequal crossing-over between homologous chromosomes (scenario C in Figure 14), resulting in a recombinant *RPP1C-LUC* fusion gene linked with another *RPP1* gene. The luc^{+} plants can be easily identified in the F2 generation and propagated for further analysis of the recombinant chimeric genes.

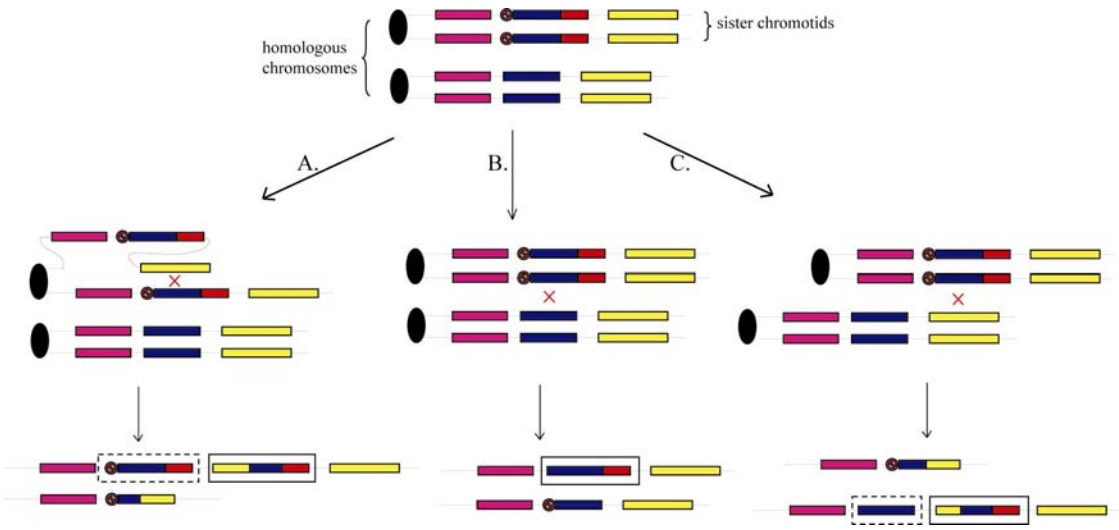


Figure 14: Different types of recombination resulting in luc^{+} phenotype. (A): Unequal crossing-over between sister chromatids. (B): Recombination between homologous chromosomes. (C): Unequal crossing-over between homologous chromosomes. Pink box: *RPP1A* gene. Blue box with a ⊘ : promoterless *RPP1C*. Red box: *LUC* gene. Yellow box: *RPP1B* gene. The genes bounded by a solid box are the chimeric recombinant genes that give a

luc$^+$ phenotype. The genes bounded by a dashed box diagnostic genes to determine which type of recombination gives rise to the luc$^+$ recombinant.

Once the luc$^+$ recombinants are identified and isolated, primers based on the conserved region of *RPP1* genes can be designed and used to amplify the recombinant genes. Sequencing and subsequent comparison of parental polymorphic bases allows one to determine the region where recombination resolution occurred. Moreover, the three different types of meiotic recombination gave rise to the activated chimeric *RPP1C:LUC* gene fusion can be distinguished because in each of the three different scenarios, there are different linked genes in the resulting gene cluster. For example, in scenario A of Figure 14 (intra-chromosomal, intergenic), there is a linked *RPP1C*$^{\Delta 1Cprom}$*:LUC* gene (in dashed box) adjacent to the activated recombinant *RPP1B/C*:*LUC* gene fusion. Whereas in scenario B (inter-chromosomal, intragenic), the original *RPP1C:LUC* gene was reconstituted by homologous recombination. Scenario C (inter-chromosomal, intergenic) has a *RPP1C* $^{\Delta LUC}$ gene linked to the activated recombinant chimeric *RPP1B/1C*:*LUC* gene fusion.

Once the luc$^+$ recombinants are identified and isolated, oligonucleotide primers corresponding to conserved regions of *RPP1* genes can be designed and used to amplify the recombinant *RPP1(A/B)/C:LUC* genes. These amplified recombinant chimeric *RPP1* genes can be subcloned and fully sequenced. Comparison of the distribution of polymorphic bases between both the recombinant *RPP1* genes and the parental genes will allow the determination of apparent meiotic recombination resolution sites (i.e. cross-overs) and more complicated gene conversion tracts (Jelesko *et al.*, 1999; Jelesko *et al.*, 2004; Jelesko *et al.*, 2005).

3.3. Results:

3.3.1. Construction of recombinant synthRPP1.1 and synthRPP1.2 gene clusters

The two *synthRPP1* gene clusters were then subcloned into the plant binary vector pYLTAC7 (Liu *et al.*, 1999) producing plasmids pJS24 and pJS26 (Figure 8), respectively. The two versions of *synthRPP1* gene cluster will allow the study of different types of recombination within *synthRPP1* gene cluster. The *RPP1C:LUC* gene fusion will serve as a target for recombination in which the linked *RPP1A* and *RPP1B* genes can recombine during meiosis. The two versions of *synthRPP1* with the *RPP1A* gene in opposite orientations will allow the study of both intra-chromosomal and inter-chromosomal recombination events. Specifically, the *RPP1A* gene on *synthRPP1.2* is in the opposite orientation relative to the *RPP1C:LUC* gene and therefore it can loop back and pair with the *RPP1C:LUC* gene in intra-chromosomal recombination events, resulting the inversion of the intervening DNA segment. Whereas the *RPP1A* gene on *synthRPP1.1* is oriented in the same direction as *RPP1C:LUC* and can undergo unequal crossing-over with either sister chromatids (intra-chromosomal) or homologous chromosomes (inter-chromosomal) recombination events.

3.3.2. Confirming in vitro FLP/FRT and CRE/lox recombination with pJS24 and pJS26

In order to confirm the *FRT-RPP1C promoter-FRT* fragment was capable of being deleted by the FLP recombinase, *in vitro* FLP/*FRT* reactions were performed using plasmid pJS24 and pJS26 as substrate and incubated with purified FLP recombinase. The *in vitro* reactions were sampled and used as template DNA in PCR amplification reactions using oligonucleotide primers that flank the two *FRT*. The non-recombined DNA template should produce a ~4 kb PCR amplification product, whereas the FLP-mediated reaction should

generate a smaller ~1 kb PCR product. As shown in Figure 15A, after incubation with the FLP enzyme, PCR produced a shorter (1.0 kb) amplification product compared to the control band (4 kb). These results demonstrated that the introduced *FRT* sites were recognized by the FLP recombinase resulting in the deletion of the 3.0 kb *RPP1C* promoter region *in vitro*, and should undergo the same reaction *in vivo*. Moreover, a plasmid (pJS28) containing an *in vitro* FLP reconfigured *synthRPP1C* was isolated and used to transform Arabidopsis plants (see Chapter 4).

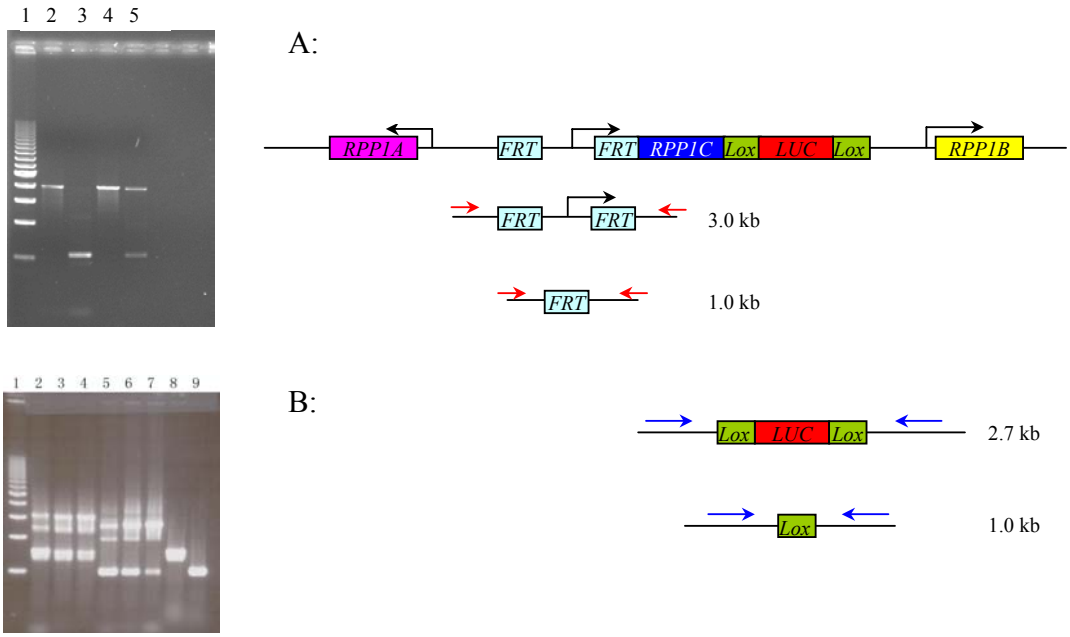


Figure 15: PCR products of in vitro FLP and CRE recombined synthRPP1 gene clusters. (A): FLP recombination. Lane1: 1 kb marker. Lanes 2 and 4: ~4 kb products before recombination (lane 2: pJS24, lane 4: pJS26). Lanes 3 and 5: ~1 kb products after recombination (lane 3:pJS24, lane 5:pJS26). Red arrows: primer pair used to amplify the fragment spanning the FRT sites. (B): CRE recombination. Lane 1: 1 kb marker. Lane 5 - 7: 1 kb and 2.7 kb products amplified from a mix of pJS24 and CRE recombined pJS24. Lane 2~4, 8, 9 were something else.

Similarly, pJS24 was inoculated with CRE recombinase to test the capability of deleting the *LUC* gene. Primers flanking the *lox* sites were used to amplify the region between the *lox* sites and thus assess whether the *LUC* gene was deleted. As shown in Figure 15B, when plasmid pJS24 was used as template in the PCR reaction there was a 2.7 kb amplification product indicating an intact *RPP1C-lox-LUC-lox* gene was present. However, when reactions that contained both pJS24 and the CRE enzyme together were used as template DNA in the PCR reactions, there was a diagnostic 1 kb amplification product, indicating a $RPP1C^{\Delta LUC}$ gene. Therefore, the *RPP1C-lox-LUC-lox* gene should also be competent for *in vivo* CRE/*lox* reactions as well.

Having demonstrated that the engineered *FRT* and *lox* sites produced the expected alternative *RPP1C:LUC* alleles, the next step is to create single copy transgenic Arabidopsis lines in the CW84 accession. CW84 is a plant line selected from a Col-0 × Ws-0 mapping cross and was homozygous for the disease-sensitive alleles from Ws-0 at *RPP2* and *RPP4* and from Col-0 at *RPP1/10/14*. In other words CW84 is susceptible to the *H. parasitica* isolates recognized by the Ws-0 *RPP1* genes present in *synthRPP1.1* and *synthRPP1.2* gene clusters.

The *synthRPP1.1* and *synthRPP1.2* gene clusters were introduced into Arabidopsis line CW84 using *Agrobacterium tumefaciens* strain GV3101 using the floral dip transformation procedure (see Chapter 2 Materials and Methods). The integrity of plasmids pJS24 and pJS26 in *A. tumefaciens* was verified by restriction enzyme fragment pattern analysis of the plasmids that were isolated from the *A. tumefaciens* cultures used to transform Arabidopsis CW84 (Figure 28).

3.3.3. Development of transgenic Arabidopsis lines containing single copy synthRPP1 gene clusters.

Transgenic T1 plants were initially selected for hygromycin resistance (Hyg^r) imparted by the *HPT* gene from pYLTAC7 located in the T-DNA region. A total of 115 Hyg^r CW84

plants were isolated. Multiple copy *synthRPP1* inserts (either at one or multiple locations) would greatly confound our estimates of both recombination frequencies and different types of DNA recombination within the *synthRPP1* loci. Therefore, it was essential to identify transgenic plants containing one and only one complete copy of the *synthRPP1* gene cluster.

Multiple copies of T-DNA inserts often insert into a single chromosomal locus. These can take the form of either inverted repeats joined at the T-DNA right borders (RB) RB-RB or joined at the left borders (LB) LB-LB. Alternatively the T-DNAs can integrate as direct repeats in either RB-LB or LB-RB orientations (Jorgensen, 1987; De Buck *et al.*, 1999; De Neve *et al.*, 1997; Krizkova & Hrouda, 1998). In addition, the T-DNA transfer process may be interrupted resulting in partially transferred T-DNAs. To select for lines with a single copy intact T-DNA insertion, genomic DNA was isolated from T1 transgenic plants, digested with *Kpn*I, blotted to nylon membranes, hybridized first with a *HPT* gene hybridization probe, stripped, and subsequently hybridized with a *SacB* gene hybridization probe. The *HPT* gene is located close to the RB and the *SacB* gene is located proximal to the LB (Figure 11 and Figure 16A). *Kpn*I cuts twice inside *synthRPP1.1* transgene and at adjacent *Kpn*I sites in the genomic DNA (Figure 16A). Because T-DNA inserts at random chromosomal locations, the position of the proximal *Kpn*I sites in the genomic DNA are unpredictable. Nevertheless, lines with single copy transgene should produce a single hybridizing *Kpn*I band with each hybridization probe, whereas multiple *synthRPP1* transgenes should produce two or more hybridizing *Kpn*I bands with each hybridization probe.

Transgenic lines containing two very closely linked T-DNA inserts should yield predictable sized probe-specific hybridizing bands depending upon the orientation of the two T-DNA inserts relative to one another. For example, multiple insertion lines with RB-RB conformation should produce a predicted ~4.4 kb *Kpn*I fragment when hybridized with the *HPT* probe and two variable-sized bands when hybridized to the *SacB* probe (Figure 16A). Similarly, two linked T-DNA inserts with a LB-LB conformation will produce two variable size bands when hybridized with the *HPT* probe and one 2.4 kb *SacB* hybridizing band.

Only in rare occasions will the variable size bands generated by a single copy insert by chance be the same size as the 4.5 kb or 2.4 kb bands diagnostic for RB-RB and LB-LB configurations. So, in general, the presence of only one *Kpn*I *SacB*-hybridizing band and one *HPT*-hybridizing band (other than the predicted 4.4 kb *HPT*-hybridizing or 2.4 kb *SacB*-hybridizing bands diagnostic for very closely linked insertions) should indicate that a particular line contains a single copy T-DNA insertion.

Only one line (AtJS24.13) from a total of 115 transgenic plants was identified that produced a single hybridizing band with either the *HPT* or *SacB* probes (Figure 16B). The red arrow in the left panel of Figure 16B (left panel) showed a single *HPT* -hybridization band in line AtJS24.13. The fact that this single HPT-hybridizing band was close to the predicted 4.4 kb RB-RB *Kpn*I fragment was not fully appreciated at this early stage of the project. AtJS24.13 produced only one *SacB* hybridizing band (Figure 16B, right panel), which was far larger than the 2.4 kb band predicted for a LB-LB double insertion. Thus, AtJS24.13 was initially characterized as a single insertion *synthRPP1.1* transgenic line.

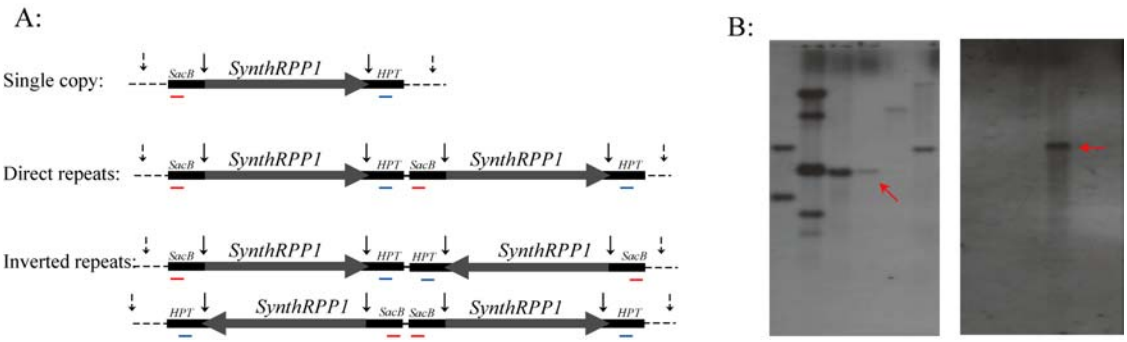


Figure 16: Southern blot to determine transgene copy number. (A): Strategy for determining transgene copy number based on the number of hybridization bands. Short vertical arrows indicate *Kpn*I cut site. Thick lines and arrows indicate the transgenes. Dashed lines indicate Arabidopsis genome. Red and blue bar indicate the positions that *SacB* gene probe and *HPT* gene probe bind, respectively. The *HPT* or *SacB* hybridizing bands were released by the *Kpn*I sites in the transgene (solid

vertical arrows) and another *Kpn*I site somewhere in the adjacent genome DNA (dashed vertical arrows), resulting a random size hybridization fragment due to the random insertion of T-DNA. Genomic DNA was hybridized to *HPT* and *SacB* gene probes respectively. A single copy line produces only one hybridizing band in each blot. Multi-copy lines (direct or inverted repeats) produce more than one hybridizing band in at least one of the two blots. (B): Southern blot result. Left panel: blot with *HPT* gene probe. Right panel: blot with *SacB* gene probe. Blots were performed at 42° C for 16 hr and exposed to Kodak Biomax® MS film for 20 minutes. Red arrows point to the single band on both blots, indicating that line was a single copy insertion transgenic line, denoted AtJS24.13.

To confirm the presence of the three intact *RPP1* genes in line AtJS24.13, DIG-labeled hybridization probes specific for *RPP1A*, *RPP1B* and *RPP1C* gene were generated by PCR based upon the 5'-UTR regions where each of the sequences significantly diverge. The same membrane used for the *HPT* and *SacB* hybridizations was stripped and hybridized with a mixture of the three *RPP1* gene probes. Predicted 7.5 kb, 8.1 kb, and a 2.0 kb hybridizing fragments characteristic of *RPP1-WsA*, *RPP1-WsB,* and *RPP1-WsC* genes (respectively) were observed. These results indicated the presence of three intact *RPP1* genes (Figure 17). There was also a fourth band with high molecular weight in all of the lines, which could be resulted from endogenous *RPP1* gene copy. Thus, Line AtJS24.13 was judged to be a single copy full-length *synthRPP1.1* transgene and was used for my subsequent studies.

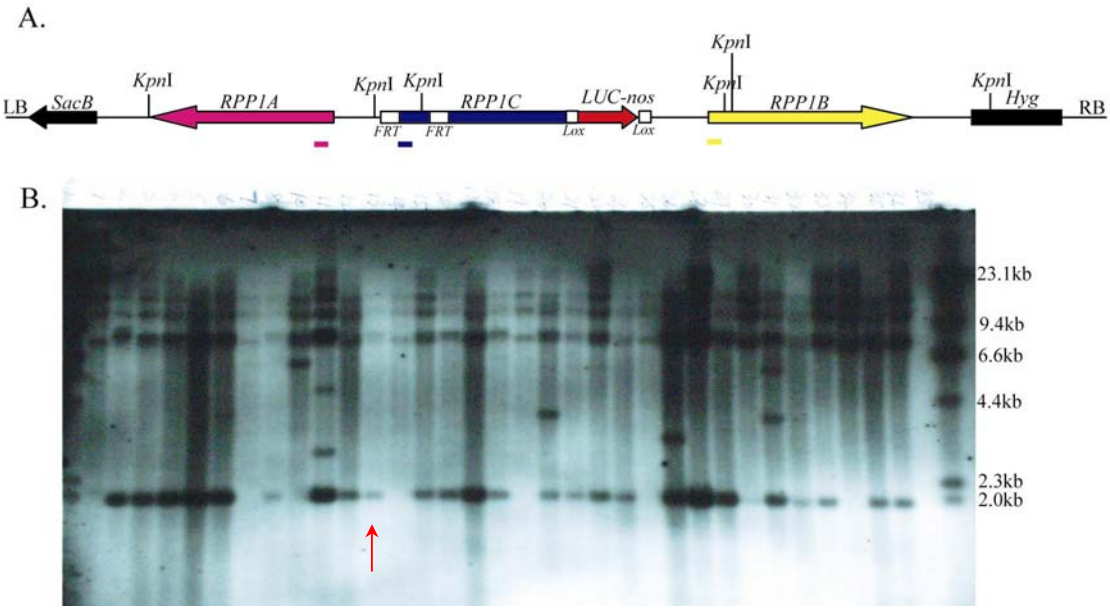


Figure 17: Southern blot to confirm the presence of intact *RPP1A*, *RPP1B* and *RPP1C* genes in the pJS24 transgenic lines.(A): the positions of the probes specific to *RPP1A*, *RPP1B* and *RPP1C* genes. Red bar, *RPP1A* gene probe; blue bar, *RPP1C* gene probe; and yellow bar, *RPP1B* gene probe. Genomic DNA was digested with *Kpn*I and probed with a mixture containing all three hybridization probes. Lines containing the intact construct were predicted to generate a 2.0 kb, a 7.5 kb and a 8.1 kb bands. Panel B, Southern blot results. The 7.5 kb *RPP1A* hybridizing band and the 8.1 kb *RPP1B* hybridizing bands were too close and not well separated in this gel. Some lines only had the 7.5 kb band, others had only the 8.1 kb band, and yet others had both bands. The two or more hybridizing bands greater than 9.4 kb were present in all of the transgenic lines and probably represent homologous *R* genes that are not *RPP1* orthologs.

3.3.4. AtJS24.13 containing a synthRPP1 gene cluster yielded transgenic plants with luciferase activity:

Line AtJS24.13 was a *synthRPP1.1* transgenic line with an apparent single copy intact transgene. This line contained the intact *FRT-RPP1C promoter-FRT-RPP1C-lox-LUC-lox* gene fusion and should produce plants with luciferase activity. T2 seeds of line AtJS24.13 were germinated on soil and the 7-day old seedlings were sprayed with luciferin and imaged using single photon video imaging equipment. Weak *in vivo* luciferase activity was detected both in the T1 plant (data not shown) and in T2 seedlings (Figure 18). The low in vivo luciferase activity may have resulted from inherent weak promoter activity of the *RPP1C:LUC* transgene and/or suboptimal conditions for normal *RPP1C* gene expression.

Transgenic lines containing a *RPP8:LUC* gene fusion showed significantly increased *in vivo* luciferase activity after the plants were subjected to a heat shock (Mohr, 2005). Therefore, in order to increase the *in vivo* luciferase expression levels of the *RPP1:LUC* gene fusion, 7 day-old AtJS24.46 and AtJS24.47 T2 seedlings were subjected to varying lengths of 37°C heat-shock (16, 20, 24, and 28 hours) followed by single photon video imaging. A 24 hr heat-shock resulted in the detection of 100% luc^+ seedlings (55 seedlings

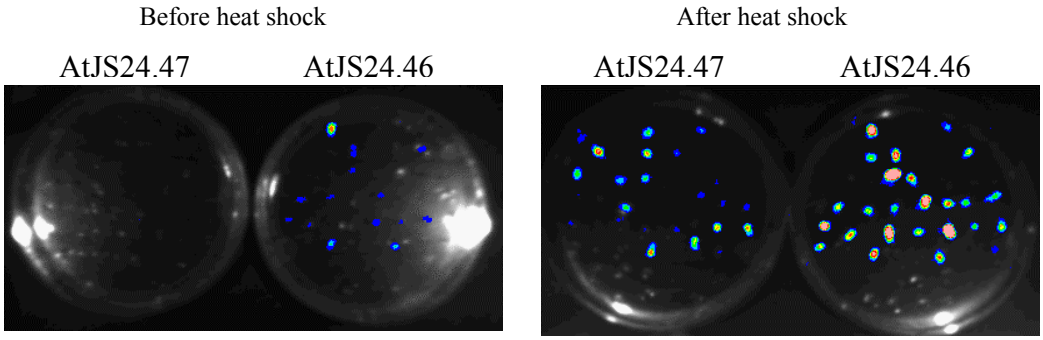


Figure 18: Luciferase activity test. One-week old seedlings were imaged as described in the materials and methods. The blue or red spots represent luciferase activity, which were superimposed onto the image of the seedling. Two different *synthRPP1* transgenic lines (AtJS24.46 and AtJS24.47) were tested here.

total). Thus, a 24 hr 37°C heat-shock before luciferase imaging was set as a standard procedure for identifying active *RPP1C:LUC* genes and gene fusions (Figure 18). Once a standard protocol was developed to yield optimal *RPP1C:LUC* expression levels of *in vivo* luciferase activity, the primary *synthRPP1.1* transgene in AtJS24.13 was then reconfigured *in vivo* to generate two alternative alleles at this single transgenic locus.

3.3.5. Generating AtJS24.13 with an alternative synthRPP1$^{\Delta LUC}$ allele

To generate the *synthRPP1*$^{\Delta LUC}$ allele homozygous T2 plants of AtJS24.13 were crossed as pollen donor to the homozygous CaMV*35S-CRE* transgenic line AtCLG9.5b.4 to form a F1 population hemizygous for both transgenes. In the F1 progeny, the CRE recombinase was expressed it should interact with the direct *lox* repeats flanking the *LUC* gene. This will delete the *LUC* gene from the *synthRPP1.1* transgene and create the *synthRPP1*$^{\Delta LUC}$ allele. These *synthRPP1*$^{\Delta LUC}$ alleles should be passed to the F2 generation and F2 plants homozygous for the *synthRPP1*$^{\Delta LUC}$ allele would have a luciferase negative phenotype (luc^{-}), whereas F2 plants containing one or more non-recombined *synthRPP1.1* alleles should show a luciferase positive phenotype (luc^{+}). More than fifty F2 seeds were germinated on ½ MS media and the 7 day-old seedlings were heat shocked and imaged for luciferase activity. All 50 plants showed a luc^{-} phenotype. The luc^{-} phenotype may be the result of: 1) *in vivo* deletion of *LUC* gene by CRE recombinase, 2) segregation away of the hemizygous *synthRPP1.1* gene, or 3) weak luciferase activity that was under the threshold of detection. Therefore, to identify bona fide *synthRPP1*$^{\Delta LUC}$ alleles a PCR screen was employed using primers that flank the *LUC* gene (Figure 15B). Using genomic DNA from F2 flower buds as template, 24 F2 plants produced a diagnostic 1.0 kb CRE-specific recombined band in the all fifty F2 plants screened. No non-recombined 2.7 kb PCR fragments were amplified (Figure 19A). These results suggest that *in vivo* CRE/*lox* recombination was extremely efficient in the F1 plants, thereby converting all *RPP1C:LUC* genes into *RPP1*$^{\Delta LUC}$ genes *in vivo*.

To avoid unwanted CRE activity that could cause unwanted deletion of the *LUC* gene from the alternative *synthRPP1*$^{\Delta 1Cprom}$ allele in subsequent crosses, the *CRE* gene needed to be segregated away from the *synthRPP1*$^{\Delta LUC}$ allele. Moreover, homozygous *synthRPP1*$^{\Delta LUC}$ lines needed to be selected. Both of these objectives were accomplished by screening the F3 segregating population for the presence of both the *CRE* gene and the presence of the *synthRPP1*$^{\Delta LUC}$ allele. Thirteen F3 lines did not amplify the *CRE* -specific fragment in PCR reactions using *CRE*-specific oligonucleotide primers (data not shown), of which line AtJS24.13-1.6 was selected.

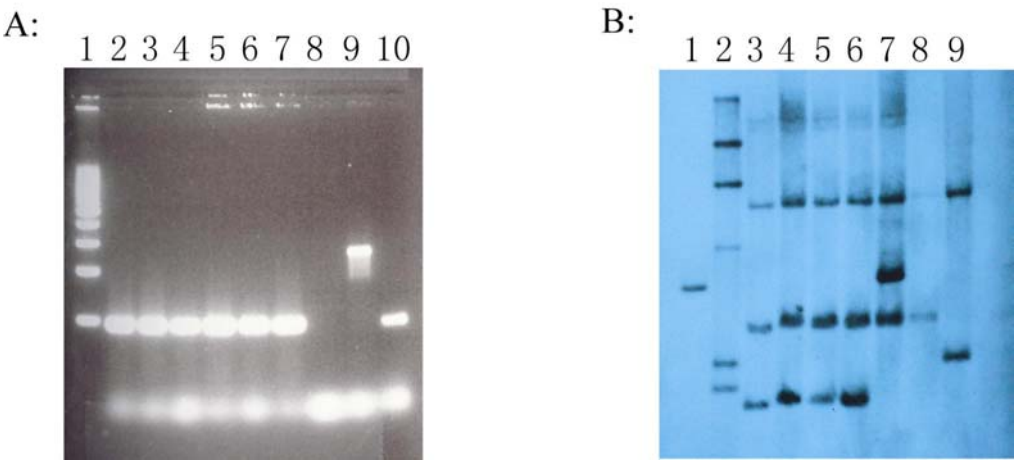


Figure 19: Generating alternative allele *synthRPP1*$^{\Delta LUC}$.Panel A, PCR to select for CRE recombined allele. Lane 1, 1 kb marker; lanes 2 - 8, PCR of genomic DNA from different F2 lines, most producing 1.0 kb characteristic band from CRE/*lox* recombined allele; lane 9, CRE non-recombined plasmid pJS24 as control, amplifying 2.7 kb band; Lane 10, CRE-recombined plasmid pJS24 as control, amplifying 1.0 kb band. Panel B Southern blot assay to confirm the CRE deletion of *LUC* gene to generate the *synthRPP1*$^{\Delta LUC}$ allele. Genomic DNA were digested with *Nco*I and blotted with a probe comprised of the 5'-UTR of the *RPP1B* gene. Lane 1: plasmid pJS24. Lane 2: λDNA (HindIII) marker. Lane 3-6: F2 lines 1.6, 1.9, 5.10 and 13.4. Lane 7: AtJS24.13. Lane 8: CW84. Lane 9: HSP-FLP.

A hygromycin resistance test on F3 seeds of line AtJS24.13-1.6 showed 12 resistant vs. 0 sensitive. The absence of any hygromycin sensitive seedlings in the F3 generation suggested this line was homozygous for the *synthRPP1* transgene (either recombined *synthRPP1*$^{\Delta LUC}$ allele or original non-recombined *synthRPP1.1*). Finally, Southern blot analysis produced a predicted 1.8 kb CRE-recombined band and did not show a 3.5 kb predicted non-recombined band (Figure 19B). These segregation studies, PCR, and Southern blot results together confirmed the genotype of line AtJS24.13-1.6 as homozygous for the *synthRPP1*$^{\Delta LUC}$ allele.

3.3.6. Attempt to generate an in vivo recombined synthRPP1 $^{\Delta RPP1C\ prom}$ *and realization that AtJS24.13 contained two very closely linked synthRPP1 loci*

In order to generate an *in vivo synthRPP1* $^{\Delta RPP1C\ \text{prom}}$ allele by FLP/*FRT* site-specific recombination, several T2 AtJS24.13 plants were crossed as the pollen donor to a homozygous CaMV*35S-FLP* transgenic line AtCLG4. Genomic DNA isolated from leaves or flowers of F2 plants was subjected to PCR screening to select for a FLP recombined allele *synthRPP1* $^{\Delta RPP1C\ \text{prom}}$ using the same method outlined in the *in vitro* FLP/*FRT* assays shown above (3.3.2 Confirming in vitro FLP/*FRT* and CRE/*lox* recombination with pJS24 and pJS26). As shown in Figure 15A, the initial non-recombined *synthRPP1.1* transgene produced a characteristic 3.0 kb band, whereas FLP recombined *synthRPP1* $^{\Delta RPP1C\ \text{prom}}$ allele produced a 300 bp band. Amplification of a 300 bp band in leaves (Figure 20A) suggested that the FLP gene was active and resulted in deletion of *RPP1C* promoter in leaf tissue. However, a 300 bp amplification product was not observed when gDNA isolated from flowers of more than 700 F2 plants (Figure 20B). These results suggested that the FLP protein/gene was active in leaf tissue, but was not active in floral tissues that give rise to the next generation (Figure 20C). Thus, gDNA from floral tissue was a more accurate indicator of a germinally inherited *synthRPP1* $^{\Delta RPP1C\ \text{prom}}$ allele, than leaf tissue where somatic FLP/*FRT* reactions led to a false positive result. The exact reason for the lack of

effective FLP recombinase activity in floral tissues was not determined. This phenomenon may have been caused either by an inactive CaMV35S promoter, or an inactive/unstable

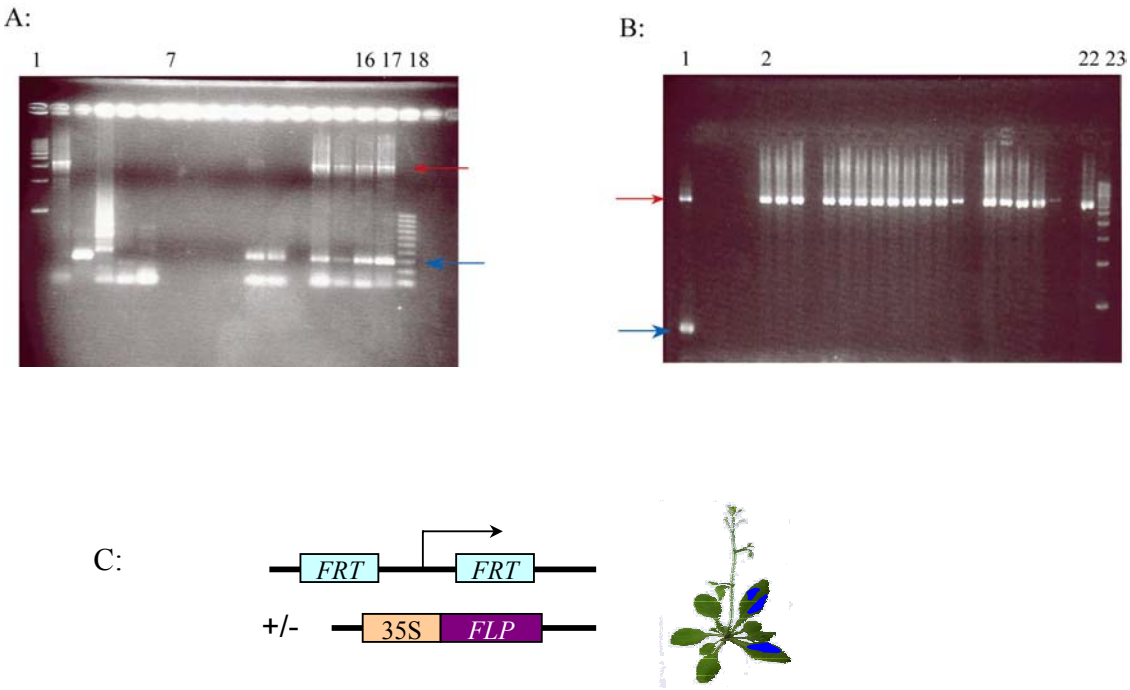


Figure 20: PCR screen for stably inherited FLP recombined *synthRPP1*$^{\Delta RPP1C\text{prom}}$.alleles.Panel A, PCR results from leaf genomic DNA. Lane 1, 1 kb marker; lanes 7 - 16, PCR products amplified from leaf genomic DNA of individual plants; lane 17, PCR products amplified from recombined and non-recombined pJS24 plasmid mix; lane 18, 100 bp marker. Panel B, PCR results using flower genomic DNA. Lane 1, PCR products amplified from recombined and non-recombined pJS24 plasmid mixture; lanes 2-22, PCR products amplified from flowers from F2 plants; lane 23, 500 bp marker. Red arrows indicate the 3.3 kb band amplified from non-recombined allele. Blue arrows indicate the 300 bp band amplified from the FLP recombined *synthRPP1* $^{\Delta RPP1C\text{ prom}}$ allele. Panel C, genotype of F2 non-recombined plants with FLP/*FRT* recombined sectors on leaves.

FLP recombinase protein in reproductive tissue. In any case, after screening flower genomic DNA from over 700 F2 plants it was apparent that this represented a significant obstacle that required another solution.

It was previously reported that germinally transmitted *FRT* recombinants could be isolated using the soybean *Gmhsp* 17.6L heat-shock promoter to drive the expression of the *FLP* gene (*HSP-FLP*) in Arabidopsis (Kilby *et al.*, 1995). Therefore, I obtained seeds that were homozygous for the *HSP-FLP* transgene. Homozygous AtJS24.13 T2 plants were crossed to homozygous *HSP-FLP* plants. When the T1 plants had a maximum number of un-open flower heads (~6 weeks old), all siliques were removed and the F1 plants were heat shocked at 37^{o} C for 4 hrs to induce *FLP* gene expression. This procedure was repeated for five consecutive days. Then the heat-shocked flowers were allowed to mature and set seed. Genomic DNA was isolated from floral tissue of the F2 plants and was used as template in PCR reactions with oligonucleotide primers that flank the *FRT* sites in the *synthRPP1.1* gene cluster in AtJS24.13. Two F2 plants (lines 11.1 and 21.5) out of 250 produced a 300 bp PCR fragment (Figure 21A). Because somatic FLP activity was very low in reproductive tissue (see Figure 20 B) and no detectable FLP expression without heat shock (data not shown), the amplification of a 300 bp band suggested the *synthRPP1*$^{\Delta RPP1Cprom}$ allele was generated in the F1 generation and stably inherited in the F2 plants. Line AtJS24.13-11.1 was chosen for further study. It is noteworthy that line AtJS24.13-11.1 also amplified the 3.3 kb non-recombined band (Figure 21B), suggesting its genotype was perhaps heterozygous at the transgenic locus (*synthRPP1*/ *synthRPP1*$^{\Delta RPP1C\ prom}$).

This finding, along with the fact that a line without the *FLP* gene was needed for subsequent crosses, made it necessary to examine the F3 population for lines that lacked the *HSP-FLP* gene and were homozygous for the *synthRPP1*$^{\Delta RPP1C\ prom}$ allele. Genomic DNA was isolated from F3 lines derived from line AtJS24.13-11.1 and used in conjunction with *FLP* gene specific primers in PCR reactions. Lines containing the *HSP-FLP* gene should produce a 800 bp diagnostic band (Figure 21C). Therefore, 50 F3 AtJS24.13-11.1 plants were screened by PCR for the absence of the *HSP-FLP* gene. Eleven F3 plants did

not produce a diagnostic 800 bp *HSP-FLP* band (Figure 21D), suggesting that the *HSP-FLP* gene had segregated away. These eleven plants were subsequently used in PCR reactions using primers flanking the two *FRT* sites. Two lines (2.1 and 5.9, Figure 21E) produced only the 300 bp amplification product that is diagnostic for the *synthRPP1*$^{\Delta RPP1C\ prom}$ allele, suggesting these plants were homozygous for the *synthRPP1*$^{\Delta RPP1C\ prom}$ allele. However, it was noted that the 3.3 kb amplification products seemed to amplify inefficiently in the other PCR reactions.

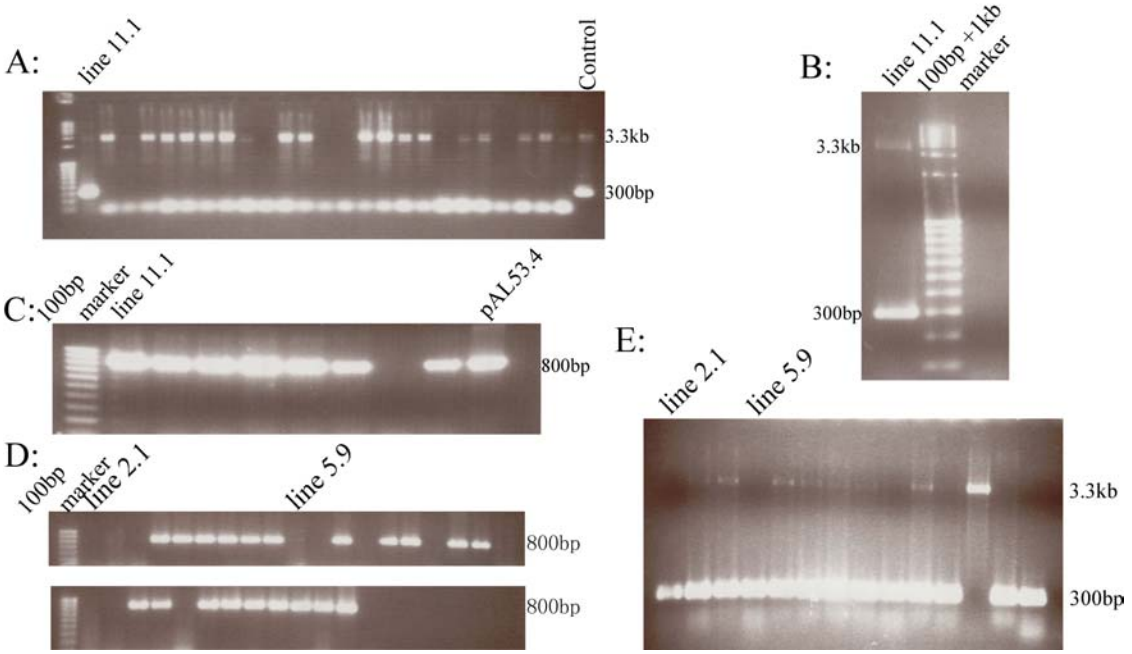


Figure 21: Attempt to develop homozygous *synthRPP1*$^{\Delta RPP1C\ prom}$ lines without the *HSP-FLP* gene construct. Panel A: PCR amplification products from gDNA of F2 plants used to identify *synthRPP1*$^{\Delta RPP1C\ prom}$ alleles. Control, *in vitro* FLP recombined plasmid pJS24 with non-recombined pJS24 amplifying both a 300 bp and 3.3 kb bands. First left lane: 1kb+100bp marker. The other lanes were individual F2 lines of the cross. Panel B: line 11.1 amplified both a 300 bp and 3.3 kb bands. Panel C: PCR amplification of *FLP* gene in individual F2 lines of the cross. line 11.1 showed a 800 bp *FLP* gene amplification product, indicating that it should be a hemizygous line; plasmid

pAL53.3 containing a *FLP* gene was used as a positive PCR amplification control. Panel D: genomic DNA from F3 line 11.1 plants was assayed for the presence of *FLP* gene using PCR. Panel E: using the same PCR assay as in panel B to select for homozygous *synthRPP1*$^{\Delta RPP1C\ prom}$ lines form the 11 homozygous *FLP*-null lines isolated from panel D.

3.3.7. Line AtJS24.13 contains two copies of transgenes, probably an intact and a partial in the form of RB-RB.

The two apparent homozygous *synthRPP1*$^{\Delta RPP1C\ \text{prom}}$ lines 2.1 and 5.9 described above were selected by PCR based on the apparent absence of a 3.3 kb band and the presence of the 300 bp band, representing the original *synthRPP1* allele and the recombined *synthRPP1*$^{\Delta RPP1C\ \text{prom}}$ allele, respectively. However, in a heterozygous line the amplification efficiency of the 3.3 kb fragment is much lower than the 300 bp band (data not shown). In some cases may result in a false negative result because of the inefficient amplification of the 3.3 kb band. To avoid the confounding problem of differential PCR amplification efficiency in a heterozygous genotype, Southern blot analysis was used to confirm the zygosity of the *synthRPP1*$^{\Delta RPP1C\ \text{prom}}$ allele in these two lines. A hybridization probe targeting *LUC* gene was designed and gDNA from lines 3.3 (hemizygous line, used as a control) and 5.9 were digested with *Kpn*I. The original non-recombined *synthRPP1* allele was predicted to produce an 8.1 kb *LUC*-hybridizing band and the FLP-recombined *synthRPP1*$^{\Delta RPP1C\ \text{prom}}$ allele should produce a 7.0 kb *LUC*-hybridizing band. Unfortunately, as observed in Figure 22A, line 5.9 produced both predicted hybridizing bands, suggesting it was heterozygous for the both the *synthRPP1*$^{\Delta RPP1C\ \text{prom}}$ allele and the original non-recombined *synthRPP1* allele. Therefore, the genomic DNA of 41 F4 plants from the presumed segregating F3 line 5.9 were examined by Southern blot analysis to identify a homozygous *synthRPP1*$^{\Delta RPP1C}$ plant. If the *synthRPP1*$^{\Delta RPP1C\ \text{prom}}$ allele and the original non-recombined *synthRPP1* alleles were each on homlogous chromosomes, then 25% of the progeny should be homozygous for the non-recombined *synthRPP1* allele and display only the 8.1 kb *LUC*-hybridizing band. Similarly, 25% should be homozygous for only the

synthRPP1$^{\Delta RPP1C\ prom}$ allele and display only the 7.0 kb *LUC*-hybridizing band, and 50% should be heterozygous for both alleles and show both *LUC*-hybridizing bands. However, of the 41 F4 plants, 39 produced both 8.1 kb and 7.0 kb *LUC*-hybridizing bands, whereas the remaining plants did not display any *LUC*-hybridizing bands (Figure 22B). The observed segregation ratio was

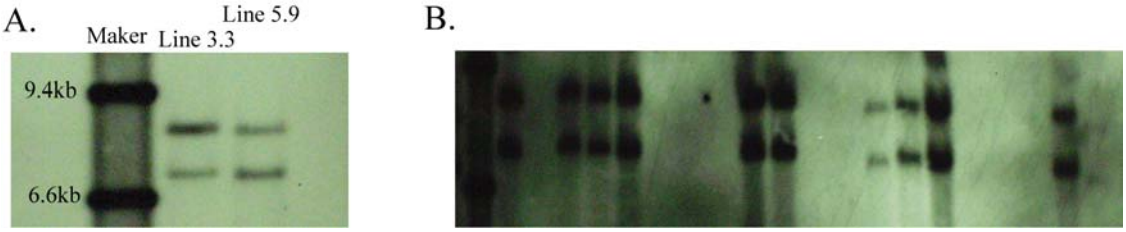


Figure 22: Determine the genotype of FLP reconfigured F3 line 5.9. (A): Southern blot of line 5.9, line 3.3 produced both recombined and non-recombined bands in PCR screen and was used here as control. (B): southern blot of F4 plants of line 5.9. Blank lanes here were due to absence of genomic DNA.

inconsistent with the prediction of the *synthRPP1*$^{\Delta RPP1C\ prom}$ gene and a non-recombined *synthRPP1* allele at the same locus of two homologous chromosomes. Moreover, none of the F4 plants displayed either a single 8.1 kb or single 7.0 kb *LUC*-hybridizing band. This segregation pattern strongly suggested that the *synthRPP1*$^{\Delta RPP1C}$ allele was closely linked to a second non-recombined *synthRPP1* transgene on the same chromosome.

This hypothesis was further supported by the observed segregation of the luciferase activity phenotype. The non-recombined *synthRPP1* transgene confers luciferase activity to one week-old seedlings, whereas the *synthRPP1*$^{\Delta RPP1C}$ allele should not confer a luc^{+} phenotype because of the deletion of the *RPP1C* promoter. If the above F3 line 5.9 and line 11.1 plants were heterozygous for the *synthRPP1*$^{\Delta RPP1C}$ and original *synthRPP1* alleles, then

approximately 25% of the resulting F4 progeny should be homozygous for the *synthRPP1*$^{\Delta RPP1C}$ allele and show a luc^{-} luciferase phenotype. More than 400 F4 line 5.9 seedlings and more than 500 F4 line 11.1 seedlings were screened for *in vivo* luciferase activity. Only 3 of 400 F4 line 5.9 plants and 14 of 500 F4 line 11.1 plants showed a luc^{-} phenotype. These observed frequencies of the luc^{-} phenotype in the presumed F4 segregating populations were far below the predicted 25% luc^{-} phenotype, expected for the *synthRPP1*$^{\Delta RPP1C}$ and *synthRPP1* alleles located at the same locus on homologous chromosomes. In other words, the vast majority of F4 luc^{+} plants suggested the 300 bp PCR marker for the *synthRPP1*$^{\Delta RPP1C}$ allele was genetically very tightly linked to the luc^{+} phenotype conferred by a non-recombined *synthRPP1* allele. The relatively few F4 luc^{-} plants were likely due to a low false negative rate consistently observed in the original *synthRPP1* transgenic lines. In summary, the luc+/- segregation data suggested that the *synthRPP1*$^{\Delta RPP1Cprom}$ allele is very closely linked to a non-FLP recombined *synthRPP1* transgene. These results were an unwelcome surprise, especially in light of our original Southern blot studies suggesting that AtJS24.13 was a single copy transgenic line.

The *syntRPP1* transgenic line AtJS24.13 that was used to produce line 5.9 was previously screened by Southern blot to select for a single transgenic insert and it was believed to contain single copy. Line AtJS24.13 was selected because it produced only one band when hybridized with either a *HPT* probe or a *SacB* probe (Figure 16). In order to explain this discrepancy, the initial Southern blot data (Figure 23) was re-examined. With the benefit of hindsight, it was apparent that the single *HPT*-hybridizing band was approximately 4.4 kb and many of the multiple copy transgenic *synthRPP1* lines also displayed a comparable 4.4 kb *HPT*-hybridizing band. Moreover, these 4.4 kb *HPT*-hybridizing bands are consistent with two *synthRPP1* transgenes closely linked at their right border (RB) ends as inverted RB-RB repeats.

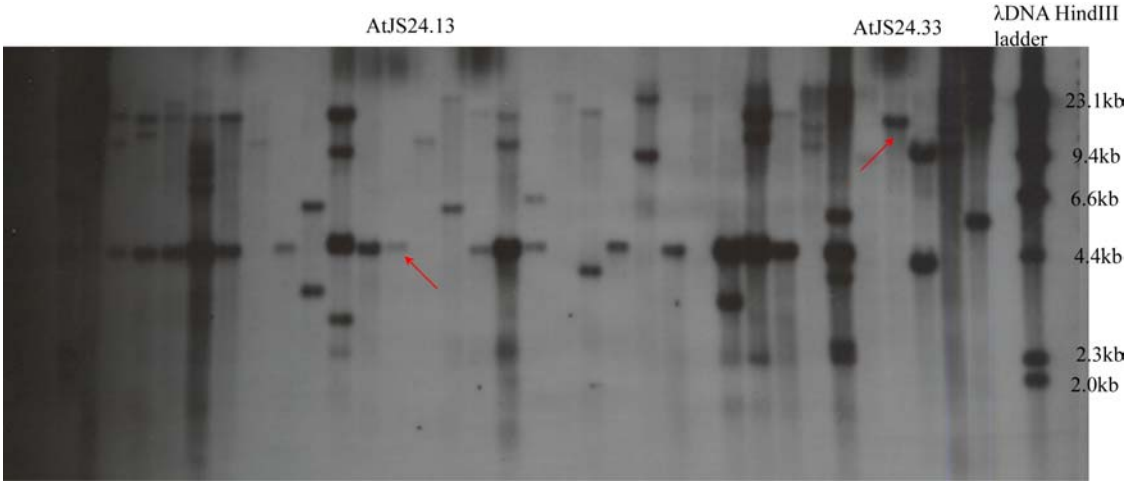


Figure 23: Transgenic line AtJS24.13 contains two copy of transgene in the form of RB-RB. Red arrow points to the ~4.4kb band generated by RB-RB configuration of AtJS24.13, which also presents in many of the other lines.

However, if line AtJS24.13 originally contained two inverted copies of the *synthRPP1.1* transgene then another conflict emerges. The Southern blot using the *SacB* gene probe should have also produced two *SacB*-hybridizing bands corresponding to the two copies of transgenes. However, only one *SacB*-hybridizing band was apparent in Figure 16B. One possible explanation is that one *synthRPP1.1* copy was intact at the LB and the putative *synthRPP1.1* insertion had an intact *RPP1C:LUC* gene, but lacked the *SacB* gene at the LB end. This would lead to the false conclusion that there was one and only one *synthRPP1.1* T-DNA insertion in AtJS24.13.

So, it was hypothesized that there were two copies of tightly linked *synthRPP1.1* transgenes in AtJS24.13 in the form of RB-RB, one intact and one partial. The partial copy contained the full length *RPP1C:LUC,* but missed some or all of the *RPP1A* gene and the lacked the entire *SacB* gene and thus was not detected by hybridization to *SacB* gene probe. During *in vivo* FLP recombination only one of the *RPP1C:LUC* gene fusions recombined

to produce a *synthRPP1*$^{\Delta RPP1Cprom}$ allele, while the other linked *RPP1C:LUC* allele remained intact, conferring a luc^{+} phenotype to this line (Figure 24).

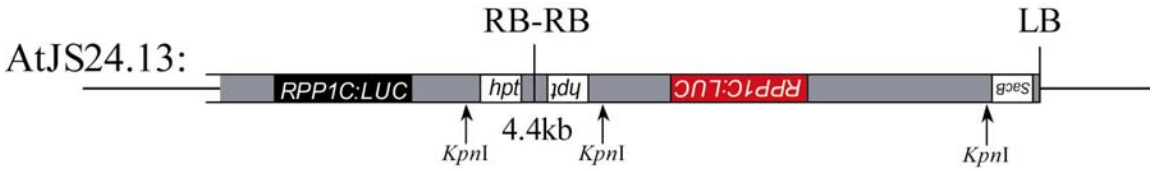


Figure 24: The hypothesized initial *synthRPP1.1* transgene configuration in line AtJS24.13.

Given that this project necessitates a single copy *synthRPP1.1* transgenic line, no further experiments were carried out using transgenic line AtJS24.13.

3.3.8. Line AtJS24.33 contains a single copy, partial transferred synthRPP1 gene cluster

Besides line AtJS24.13, another single copy *synthRPP1* transgenic line, AtJS24.33 was identified by Southern blot analysis (see Section 3.3.3, Development of transgenic Arabidopsis lines containing single copy *synthRPP1* gene clusters). The initial criterion for a single copy T-DNA insert was based upon the production of one and only one band when hybridizing to *HPT* gene probe and one band when hybridizing to the *SacB* gene probe (see Figure 16). Moreover, the single *HPT* hybridizing band shouldn't be 4.4 kb. Line AtJS24.33 produced a single hybridizing band with *HPT* gene probe that was not 4.4 kb (see Figure 23). However, it failed to show any *SacB* hybridizing fragment (see Figure 25).

Figure 25: AtJS24 transgenic lines hybridizing to *SacB* gene probe. Red arrow indicates the lane that has AtJS24.33 gDNA and no hybridizating band was observed in this lane.

Since AtJS24.33 displayed a luc^{+} phenotype, it suggested that the *RPP1C:LUC* gene fusion was intact. Thus, it was hypothesized that line AtJS24.33 had a single incomplete T-DNA insertion and the T-DNA transfer was interrupted somewhere between the intact *RPP1C:LUC* gene fusion and the *SacB* gene. If this hypothesis was correct, then line AtJS24.33 might still useful if it contained full length *RPP1C:LUC* and *RPP1B* genes, and it might also contain sufficient 5' *RPP1*A sequences to allow recombination with a *synthRPP1*$^{\Delta RPP1C}$ allele, resulting in the activation of a putative *RPP1A/1C:LUC* fusion.

Southern blot analysis was used to evaluate how much of the *RPP1A* gene was present in transgenic line AtJS24.33. As shown in Figure 26A, several restriction enzyme digestions were performed that would yield specific *RPP1A*-hybridizing fragments depending upon how much of the *RPP1A* gene was present in line AtJS24.33. Specifically, if the entire *RPP1A* gene was intact it should produce a 1.3 *Hind*III, a 6.0 kb *Nsi*I, a 5.3 *Bgl*II, a 7.4 *Bts*I and a 7.5 *Kpn*I *RPP1A*-hybridizing fragment when digested with these enzymes. Of these predicted fragments, only a 1.3 kb *RPP1A Hind*III-hybridizing genomic DNA fragment

was observed (Figure 26). The *Nsi*I, *Bgl*II, and *Kpn*I *RPP1A*-hybridizing bands had different sizes and the *Bts*I band was absent, probably because

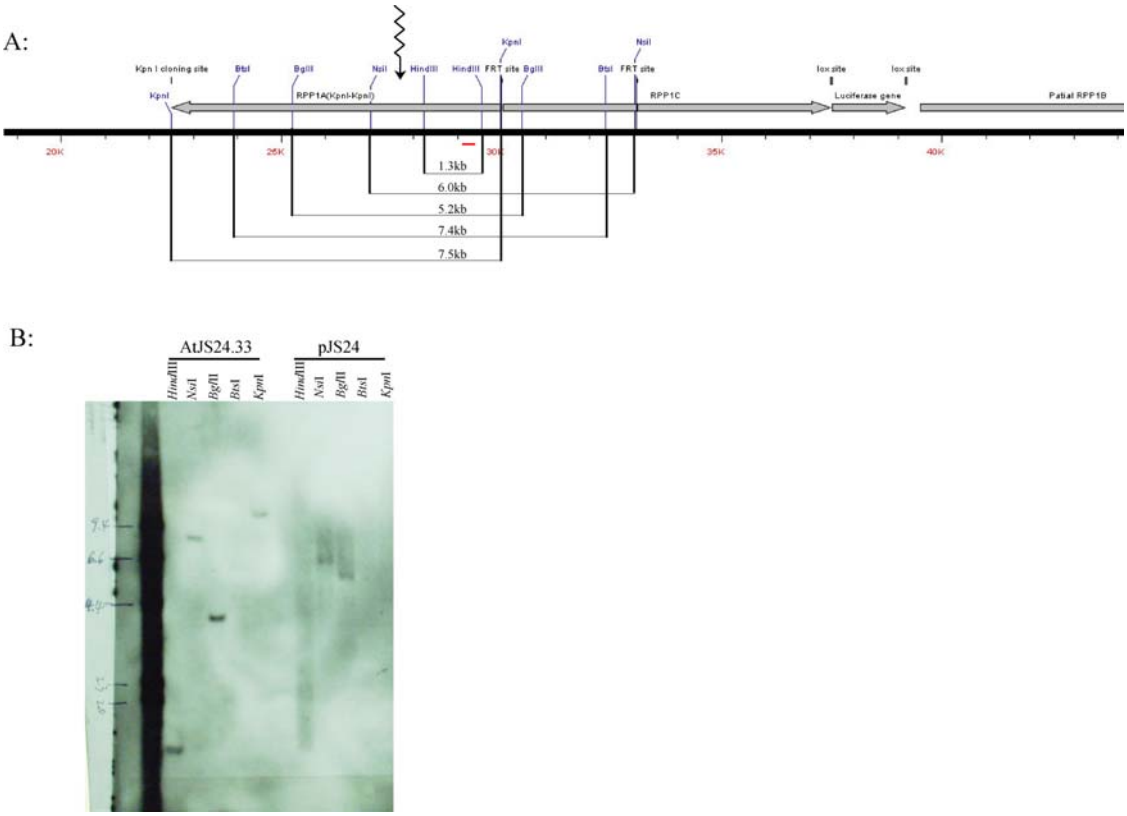


Figure 26: Southern blot to determine how much of the RPP1A gene was transferred in line AtJS24.33. Panel A, restriction map of the RPP1A region in the synthRPP1.1 gene cluster. Short red bar indicates the position of the RPP1A gene-specific probe used in the Southern blot. Panel B; Southern blot results, pJS24 was used as a positive control.

the small hybridizing fragments ran off the gel. From this result it was concluded that the T-DNA transfer beginning at the RB was interrupted somewhere between the 5' *Nsi*I and the 5' *Hind*III restriction sites shown with a wavy arrow in Figure 26B. Therefore, most of the 3' region of the *RPP1A* gene (5' to the wavy arrow, including the *SacB* gene) was missing in transgenic line AtJS24.33. Consistent with this interpretation, the original

Southern blot analysis used to determine transgene copy number didn't produce a *SacB* hybridizing fragment (data not shown) suggesting that the *synthRPP1.1* insertion in this line was incomplete. The transcription start codon (ATG) of *RPP1A* gene is located between these 5' *Nsi*I and the 5' *Hind*III sites. However, the exact end of the *RPP1A* gene fragment was not more accurately determined. Therefore, it is formally possible that a short stretch of *RPP1A* coding sequence was transferred in line AtJS24.33. If this is the case, there may be a sufficient amount of the 5' coding region of the *RPP1A* gene to allow homologous recombination with the *RPP1C:LUC* gene, allowing intra-molecular recombination. With that said, even if there is not a sufficient amount of the *RPP1A* gene present, this line can be still used as a two gene synthetic gene cluster (*RPP1C:LUC* and *RPP1B*) and would be valuable for investigating the recombination activities between *R* genes.

3.3.9. Additional efforts to generate transgenic lines containing a single intact synthRPP1 transgene

Generating a single copy full-length *synthRPP1* transgenic line is crucial to this project. Therefore, Arabidopsis CW84 was re-transformed with plasmid pJS24 using *A. tumefaciens*-mediated floral dip transformation. Thirty-one hygromycin-resistant T1 transformants were selected on ½ MS media with 30 ug/ml hygromycin. Genomic DNA will purified from these transformed plants and subjected to Southern blot analysis using hygromycin and *SacB* gene probes as described above.

3.4. Conclusion and Discussion:

This chapter describes the development of two novel synthetic *RPP1* gene clusters suitable for advanced studies of meiotic recombination between closely linked paralogous *RPP1* genes. Assembly of these clusters was very complicated, and represents a very important

step in the development of this project. These *synthRPP1* gene clusters were designed such that two alternative *synthRPP1* alleles could be generated from any single *synthRPP1* transgene insertion using *in vivo* site specific recombination reactions. These alternatively marked alleles could be used in genetic crosses to identify a variety of meiotic recombination events leading to chimeric RPP1 genes that could be identified by a gain-of-luciferase phenotype. To this end, two initial transgenic *synthRPP1.1* lines were generated (AtJS24.13 and AtJS24.33). Using line AtJS24.13 in vivo FRT and CRE reactions were intiated and produced their respective alternative alleles at the AtJS24.13 transgenic locus. This demonstrated that alternative alleles at a *synthRPP1* locus could be generated, which is another important step towards fulfilling the long-term goals of this project. However, during subsequent characterization it was discovered that the initial transgenic locus was comprised of one fully intact and one partial *synthRPP1.1* loci jointed as inverted repeats at their right border sequences. The AtJS24.33 synthRPP1.1 locus was shown to be a partial T-DNA insertion comprised of mostly an intact *RPP1C:LUC* and *RPP1B* genes. More transgenic AtJS24 lines were generated in hope of identifying a single copy synthRPP1.1 insertion.

Obtaining a single copy *synthRPP1* transgenic locus was essential for the investigation of meiotic recombination. As illustrated in Figure 14, the genetic organization of the *synthRPP1* locus associated with a chimeric *RPP1C:LUC* gene will indicate exactly which type of meiotic recombination gave rise to the chimeric *RPP1C:LUC* gene. However, if there were more than one *RPP1A* or *RPP1B* genes present in the original T-DNA insertion (i.e. two closely linked T-DNA inserts) it would greatly complicate the determination of both the composition and organization of the recombined gene cluster. This would introduce considerable uncertainty into assigning a particular type of meiotic recombination to a particular activated chimeric *RPP1C:LUC* gene. For the vast majority of transgenic research, it is not essential to generate one and only one T-DNA insertion to assess the role of the transgene. For example, either genetic complementation studies or gene knockdown studies are not adversely affected by multiple linked transgenes. In over-expression studies one occasionally obtains lines that show co-suppression of the endogenous gene. These co-suppression effects are now believed to be due to two or more

linked tail-to-tail transgenes that form a region of double stranded RNA transcripts that activate RNAi mediated gene silencing. During these studies and those of Dr. Stacey Simon, approximately one out of every 100 trangenic lines appears to contain one and only one transgene copy.

Initially, a Southern blot strategy was used to select for the single copy transgenic lines (see Figure 16) by examining the organization of the T-DNA ends. At that time, it was assumed that all transgenic lines contained at least full-length T-DNA transfers. However, the results with line AtJS24.33 showed that partial T-DNA transfer also occurred, resulting in a truncated *RPP1A* transgene (Figure 25). Consistent with this interpretation, AtJS24.33 didn't produce the downstream *SacB* gene (data not shown).

Line AtJS24.13 was initially considered a single-copy transgenic line because it produced a single *HPT*- and a single *SacB*-hybridizing fragment. However, the size of *HPT*-hybridizing band was 4.4 kb, which was in agreement with the calculated size of a RB-RB inverted repeat. Consistent with the hypothesis that this 4.4 kb *HPT* hybridizing band was generated by two RB-RB linked T-DNA inserts, many of the other transgenic lines (Figure 23) that showed clear evidence for multiple T-DNA copies (i.e. more than one *HPT* hybridizing bands) frequently had a 4.4 kb *HPT* hybridizing band. Our experience with multiple linked T-DNA inserts is in accord with previous reports that T-DNA often integrates as multicopy arrays in the form of head-to-head, tail-to-tail or tandem arrays (De Buck *et al.*, 1999; De Neve *et al.*, 1997; Jorgensen, 1987).

With the clarity of hindsight, it now seems obvious that an *HPT*-hybridizing band that is also present in clear multicopy lines should have been of concern. However, the single *SacB*-hybridizing band was also consistent with the countervailing interpretation of a single copy T-DNA. Moreover, it took the cumulative weight of extensive genetic segregation data showing linkage of the luc^{+} phenotype to the PCR-based *synthRPP1*$^{\Delta RPP1C\,\mathrm{prom}}$ allele to determine that the transgenic locus in AtJS24.13 was a complex RB-RB fusion with one of the T-DNAs being comprised of a partial T-DNA transfer (i.e. deletion of part of *RPP1A* and the second *SacB* gene).

Although AtJS24.13 was not suitable for our eventual recombination studies, this line did demonstrate several important proof-of-principle concepts of the synthRPP1 gene cluster design. Firstly, this research demonstrated that two alternative alleles of the *synthRPP1.1* transgene could be formed using *in vivo* sight specific recombination. While the *in vivo* CRE/*lox* reactions were remarkably efficient, the *in vivo* FLP/*FRT* site specific reactions were prone to several complicating factors. For example, the FLP/*FRT* were efficient in leaves, but very inefficient in floral tissues. In the F2 progeny of the *CaMV35S-FLP* X AtJS24.13 lines the FLP activity could be assayed in leaves but not in floral tissue. Although the exact cause of this phenomenon was not determined, this observation could be put to good use by only assaying floral tissue for identification of bona fide *synthRPP1*$^{\Delta RPP1Cprom}$ alleles. Secondly, the use of a *HSP-FLP* transgene to drive FLP/*FRT* site specific recombination events suitable for germinal transmission of *synthRPP1*$^{\Delta RPP1Cprom}$ alleles was clearly established. Thus, the basic tenets of developing an *in vivo* reconfigurable *synthRPP1* gene cluster were shown to be feasible. This work lays essential ground work for future studies with these *synthRPP1.1* and *synthRPP1.2* gene clusters to investigate and model *R* gene evolution by meiotic unequal crossing-over.

Chapter Four: Development of *in vitro* FLP reconfigured synthetic *RPP1* (*synthRPP1*$^{\Delta 1Cprom}$) lines.

4.1. Introduction:

Because of the delays associated with generating the *in vivo* reconfigured synthetic *RPP1* lines described in Chapter 3, a different *in vitro* FLP reconfigured synthetic *RPP1* construct was developed to measure meiotic recombination between a *RPP1C:LUC*$^{\Delta RPP1Cprom}$ allele and linked *RPP1A* and *RPP1B* genes (Figure 27). This synthetic *RPP1* gene cluster was called *synthRPP1.3* to differentiate it from those discussed in Chapter 3.

Instead of performing *in vivo* FLP/*FRT* recombination reactions post plant transformation, the *synthRPP1.3* gene cluster employed *in vitro* FLP/*FRT* recombination prior to plant transformation to generate the inactive *RPP1C:LUC*$^{\Delta 1Cprom}$ allele (Figure 27). In this respect, the *synthRPP1.3* gene cluster was conceptually similar to the *synthRBCSB*1 gene cluster (Jelesko *et al.*, 1999), in which the *ΔRBSC1B:LUC* gene fusion is lacking all 5' promoter elements required for expression of the *RBSC1B:LUC* gene fusion resulting in initial luc$^-$ transgenic plants. However, instead of using restriction enzymes to delete the *RPP1C* promoter region, *in vitro* FLP/*FRT* site-specific recombination could be used to delete the *RPP1C* promoter from the original *synthRPP1.1* gene cluster.

Compared to the *in vivo* reconfigurable *synthRPP1.1* and *synthRPP1.2* gene clusters, the disadvantage of this experimental approach is that there were no alternatively marked alleles to distinguish meiotic recombination between sister chromatids vs. recombination between homologous chromosomes. Nevertheless, like the *synthRBCSB1* synthetic gene cluster, this approach should rapidly generate initial luc$^-$ *synthRPP1.3* transgenic lines that can be selfed to generate *RPP1A/1C:LUC* and *RPP1B/1C:LUC* chimeric *R* genes that confer a gain-of-luciferase (luc$^+$) phenotype. Thus, this approach should rapidly provide information on overall meiotic recombination frequency of *RPP1B* to *RPP1C* vs. *RPP1A* to *RPP1C* (gene orientation effect) because extensive in vivo reconfiguration steps are

obviated. Moreover, this will provide an initial platform of rapidly generating recombinant *RPP1* genes for structure function investigations.

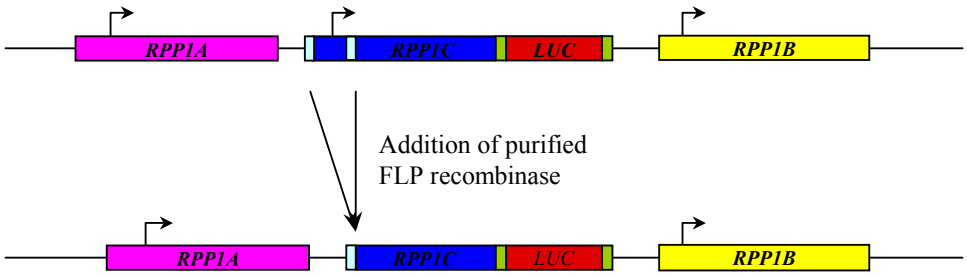


Figure 27: *In vitro* FLP reconfiguration of *synthRPP1.1* to produce a *synthRPP1.3* $^{\Delta 1Cprom}$ construct. Blue boxes indicate *FRT* sites and green boxes indicate *lox* sites.

4.2. Experimental design:

The objectives of this line of experimentation was to first generate a *synthRPP1.3* gene cluster containing a *RPP1* $^{\Delta 1Cprom}$ allele *in vitro* using FLP/*FRT* recombination (Figure 27). This recombinant DNA construct will be used to generate single copy *synthRPP1.3* luc^{-} Arabidopsis lines. Once these lines are generated they can be used to identify meiotic recombinants with a luc^{+} phenotype indicating the formation of chimeric *RPP1A/1C:LUC* or *RPP1B/1C:LUC* alleles. These recombinant *RPP1* genes will demonstrate proof of concept that one can model natural *R*-gene evolutionary processes by identifying recombinant *R* genes in a single round of meiotic recombination.

4.3. Results:

4.3.1. In vitro FLP recombination of pJS26 and plant transformation

The *in vitro* FLP recombination was performed by adding purified FLP recombinase with plasmid pJS26 containing the *synthRPP1.2* gene cluster. Successful FLP/*FRT* site-specific recombination was tested by PCR using primers (oJGJ201 and oJGJ202) flanking the two *FRT* sites (see Figure 15A in Chapter 3). Aliquots from these *in vitro* pJS26 FLP/*FRT* reactions were used to transform *E. coli* DH5alpha and single colonies were screened by PCR using oligonucleotide primers oJGJ201 and oJGJ202. Colonies displaying a 3.3 kb product indicated no FLP/*FRT* recombination, whereas colonies displaying a 300 bp PCR product indicated deletion of the *RPP1C* promoter and formation of a $RPP1C^{\Delta 1C\,prom}$ allele. One clone with $synthRPP1^{\Delta 1Cprom}$ allele was selected and the plasmid was named pJS28. Thus, plasmid pJS28 contained the in vivo reconfigured synthRPP1.3 gene cluster containing a $RPP1^{\Delta 1Cprom}$ allele.

Plasmid pJS28 was introduced into *Agrobacterium tumefaciens* GV3101 using electroporation. *A. tumefaciens* containing pJS28 was used to transform Arabidopsis CW84 by floral dipping. To confirm construct integrity after transfer to *A. tumefaciens,* plasmid pJS28 was extracted out of the Agrobacterium culture used in the floral dipping transformation reactions and transformed back into *E.coli*. Then pJS28 was isolated from *E.coli* and subjected to various restriction enzyme digestions with control pJS28 samples that didn't undergo this process. The digestion showed that the FLP recombination resulted in a restriction band shift compared to pJS26 and there were no changes in restriction fragment pattern in pJS28 with or without being transformed into Agrobacterium (see Figure 28).

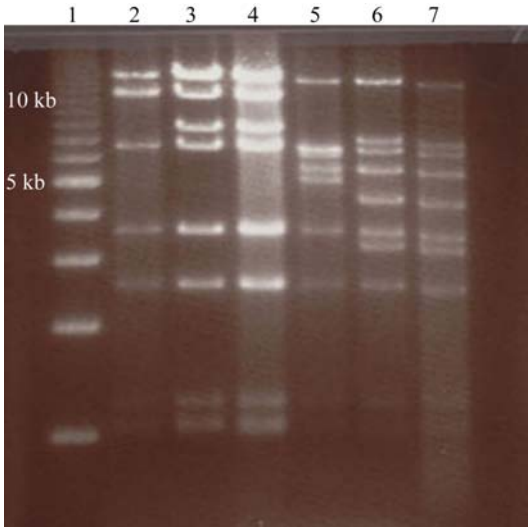


Figure 28: *Pst*I digestion to confirm in vitro FLP recombination and plasmid pJS28 integrity after transfer into Agrobacterium.Lane 1, 1 kb marker; lane 2, pJS24 (*Pst*I); lane 3, pJS27 FLP reconfigured pJS24 (*Pst*I, without going into Agrobacterium); lane 4, pJS27 (*Pst*I, extracted from Agrobacterium); lane 5, pJS26 (*Pst*I); lane 6, pJS28 FLP reconfigured pJS26 (*Pst*I, without going into Agrobacterium); lane 7, pJS28 (*Pst*I, extracted from Agrobacterium). The altered banding patterns were due soley to *in vitro* FLP/*FRT* recombination events and not unequal crossing over in Agrobacterium.

4.3.2. Selection for Arabidopsis lines that have a single-copy synthRPP1.3 transgene:

More than 100 hygromycin-resistant *synthRPP1.3* Arabidopsis primary transformants were selected. The hygromycin-resistant T1 seedlings were transferred to soil for growth. Genomic DNA was purified from the leaves and flowers of T1 transformant plants and subjected to Southern blot analysis as previously described in Chapter 3 to identify single-copy *synthRPP1.3* transgenic lines.

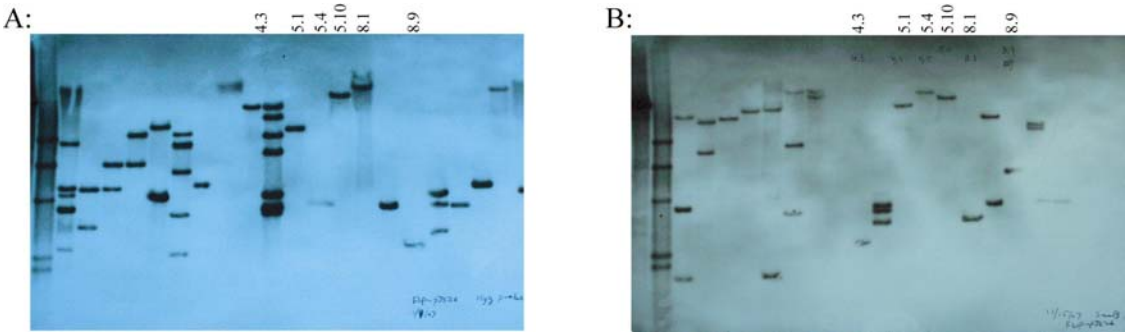


Figure 29: Southern blot to select for single copy *synthRPP1.3* transgenic lines.Panel A, gDNA was hybridized to the *HPT* gene probe. Panel B, gDNA was hybrized to the *SacB* gene probe.

Genomic DNA from over 100 T1 transformants were digested with *Kpn*I and hybridized to *HPT* and *SacB* DIG-hybridization probes (see Figure 29A and 29B). Six transgenic lines (lines 4.3, 5.1, 5.4, 5.10, 8.1, and 8.9) demonstrated both a single *HPT* and a single *SacB* hybridizing band. Line 5.1 was selected for subsequent studies because the *HPT* gene hybridizing band was not a: i) 4.4 kb band, indicative of a RB-RB tandem T-DNA insertion, or ii) 2.4 kb band that indicative of a LB-LB tandem T-DNA insertion, and iii) not a 3.4 kb band indicative of a LR-RB direct repeat (see Section 3.3.3 "Development of transgenic Arabidopsis lines containing single copy *synthRPP1* gene clusters" for explanation),. Based upon these more rigorous Southern blot criteria, AtJS28 line 5.1 was identified as a single copy *synthRPP1.3* T-DNA transgenic line.

4.3.3. Develop homozygous lines with synthRPP1.3

By first principles, a single copy transgene in the T1 generation must be a hemizygous genetic locus. Thus, AtJS28 line 5.1 was hemizygous for *synthRPP1.3* locus. To increase the likelihood of obtaining meiotic recombinants at the *synthRPP1.3* locus, it is desirable to work with homozygous lines. Therefore, the AtJS28 line 5.1 T1 plant was allowed to self

fertilize, giving rise to a segregating T2 population. T2 seeds were germinated and individual T2 plants were allowed to self fertilize, giving rise to T3 seeds. T3 seeds from individual T2 plants were tested for hygromycin resistance segregation. T2 plants were identified as those that produced T3 seeds with 100% hygromycin resistant T3 progeny. Five T2 plants were identified as homozygous for *synthRPP1.3* transgene and their T3 seed bulks were saved as a source of homozygous single copy *synthRPP1.3* homozygous seed stocks.

Approximately one thousand T3 AtJS28 line 5.1 homozygous seeds were germinated to generate a large homozygous *synthRPP1.3* T4 population in which unequal crossing-over between either the *RPP1A* or *RPP1B* genes would misalign and recombine with the linked *RPP1C:LUC*$^{\Delta 1Cprom}$ allele to form activated recombinant chimeric *RPP1(A or B)/1C:LUC* genes that would confer a luc^{+} phenotype in the T5 generation. This T5 generation is currently growing and screening of this population will be performed by other lab members working in the Jelesko and McDowell laboratories.

4.4. Conclusion and Discussion

This chapter describes the development of an *in vivo* FLP reconfigured *synthRPP1.3* gene cluster that is suitable for rapidly identifying recombinant chimeric *RPP1(A or B)/1C:LUC* genes that confer a gain-of-luciferase phenotype. Efficient *in vitro* FLP/*FRT* reactions led to a *synthRPP1.3* gene cluster that was used to generate several apparent single copy transgenic lines by Southern blot analysis. One of these lines was used to develop a large T5 homozygous seed bulk, suitable for future luciferase screening.

The purpose of developing *in vitro* reconfigured promoterless *synthRPP1* transgenic lines was to rapidly generate estimates of the frequency of unequal crossing-over between *RPP1* paralogs and isolation of recombinant chimeric *RPP1(A or B)/1C:LUC* genes. This approach bypasses the time-consuming and rather inefficient *in vivo* FLP/*FRT* and CRE/*lox* reconfiguration steps and subsequent genetics required for the *synthRPP1.1* and

synthRPP1.2 gene clusters described in Chapter 3. Like the *synthRBCSB1* gene cluster, the *synthRPP1.3* gene cluster will not be able to distinguish between sister chromatids and homologous chromosome exchanges. Nevertheless, it should rapidly allow us to test if chimeric *RPP1(A or B)/1B:LUC* genes can be identified in gain-of-luciferase screens, what their frequency of formation is, and mapping of recombination resolution sites between the *RPP1* paralogs. Success in these three specific objectives will firmly establish the utility of using *synthRPP1* gene clusters to model the evolution of *R* gene clusters. These results will also generate much detailed information about meiotic unequal crossing-over and recombinant *R* gene recognition specificity.

Chapter Five: Investigating the role of zygosity on meiotic and somatic recombination within a *synthRBCSB* gene cluster

5.1. Introduction:

Synthetic gene cluster technology was originally developed using a *synthRBCSB* gene cluster and was used to study meiotic recombination between paralogous genes located on sister chromatids. The *synthRBCSB* gene cluster is comprised of three paralogous Arabidopsis *RBCS1B*, *RBCS2B* and *RBCS3B* genes (Figure 30). They belong to a closely linked multigene family encoding the small subunit of ribulose-1, 5-bisphosphate carboxylase. The three genes are arranged in tandem repeats. The target for recombination is the *RBCS1B* gene that is fused in-frame to the *LUC* gene. The promoter of *RBCS1B* gene is deleted as well as part of exon 1 (designated *ΔRBCS1B:LUC*). The remaining *RBCS2B* or *RBCS3B* genes are positioned 3' to the *ΔRBCS1B:LUC* target locus at the same relative distances as the endogenous *RBCSB* locus. This transgene construct confers a luc^- phenotype to transgenic plants. However, once recombination occurs between the inactive *ΔRBCS1B:LUC* gene and either the *RBCS2B* or *RBCS3B* genes, the resulted chimeric gene-*LUC* gene fusion is then driven by the active promoter of the *RBCS2B* or *RBCS3B* gene and these alleles will be passed to the next generation and identified therein as plants with a luc^+ phenotype.

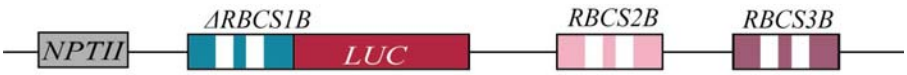


Figure 30: Structure of *synthRBCSB* gene. Firefly luciferase-*nos* gene was fused 3' in-frame to *RBCS1B* gene. The promoter and part of the first exon of the *RBCS1B* gene was deleted (*ΔRBCS1B:LUC)*. The *RBCS2B* and *RBCS3B* genes are cloned 3' to the *ΔRBCS1B:LUC* gene. White boxes in *RBCSB* genes indicate introns.

The most accurate estimates of the relative impact of meiotic unequal crossing-over between genes located on either homologous chromosomes or sister chromatids were measured in yeast. Jackson and Fink (1985) showed that meiotic unequal crossing-over between homologous chromosomes occurs at least 30 fold more frequently relative to sister chromatid exchanges (Jackson & Fink, 1985). There are no equivalent reports directly comparing the rate of meiotic unequal crossing-over between homologous chromosomes vs. sister chromatid exchanges in plants. One relevant, although not parallel experiment, was the measurement of spontaneous stress-induced somatic intrachromosomal recombination frequencies of *NPTII* gene repeats in hemi- and homo-zygous transgenic tobacco protoplast (Lebel *et al.*, 1993). These showed a somatic recombination frequency of 9×10^{-5} in hemizygous protoplasts and 19×10^{-5} in homozygous protoplasts, suggested an additive effect in the homozygous lines. This is supported by the observation of somatic recombination between homozygous tandem disrupted beta-glucuronidase or luciferase reporter genes in direct repeat orientation were twice as frequent as that between lines with hemizygous constructs (i.e. sister chromatids) (Molinier *et al.*, 2004). Therefore, in plants all measures of the effect of zygosity were performed in somatic tissues. Thus, it is not clear if similar phenomenon occurs during meiosis.

In order to study if zygosity has an impact on meiotic unequal crossing-over, large populations of homozygous and hemizygous *synthRBCSB* lines were developed and meiotic recombination frequencies in F2 populations were measured.

5.2. Experimental design and results:

5.2.1. Meiotic unequal crossing-over in hemi- and homo-zygous synthRBCSB populations

One way to estimate different rates of unequal crossing-over between homologous chromosomes vs. sister chromatids is to measure the rate of sister chromatid exchange at the *synthRBCSB* locus when comparable locus is present on the homologous chromosome

(i.e. a hemizygous population) and compare that to the rate of unequal crossing-over when the *synthRBCSB* locus is present on both homologous chromosomes. In the case of a *synthRBCSB* homozygous line, two types of meiotic unequal crossing-over are possible: sister chromatid exchange and exchange between homologous chromosomes. Unfortunately, a limitation of this experimental design is that one can not differentiate between these two types of meiotic recombination (see Figure 12). Nevertheless, if meiotic recombination between homologous chromosomes is additive or much greater than sister chromatid exchanges, then the observed frequency of unequal crossing-over in the homozygous lines will be significantly greater than that in the hemizygous population. To test this hypothesis, large populations of matched hemizygous *synthRBCSB* and homozygous *synthRBCSB* plants were generated.

A T3 homozygous *synthRBCSB1-10* transgenic line (Jelesko *et al.*, 2004) was used in this experiment. To develop a large population of hemizygous *synthRBCSB*, the homozygous T3 *synthRBCSB1-10* plants were crossed to wild-type Arabidopsis Columbia-0. More than 200 crosses were performed in order to generate approximately 3,000 hemizygous F1 seeds. This F1 population was sufficiently large to yield an F2 population of approximately 10 million F2 seed. Similarly, to generate the homozygous *synthRBCSB* population, approximately 20 homozygous T3 plants were self-fertilized to produce more than 3,000 homozygous T4 *synthRBCSB* seeds and 10 million T5 seeds. In order to minimize experimental error due to environmental conditions, the hemi- and homozygous populations were grown at the same time in a Conviron growth chamber. During meiosis in the hemizygous F1 and homozygous T4 plants, the *synthRBCSB* gene cluster should undergo unequal crossing-over, giving rise to recombinant chimeric *RBCS(2B/3B)/1B:LUC* alleles with activated luciferase reporter genes. These alleles should be passed through the gametes to the subsequent generations (F2 for hemizygous population, and T5 for the homozygous population, respectively).

The progeny from the homo- and hemizygous plants were then germinated on filter paper wetted with Hoagland's solution (see Chapter 2 Materials and Methods). Five-day old seedlings were imaged for a luc^+ phenotype using the VIM50 single photon video imaging

system (Hamamatsu Co., Hamamatsu City, Japan) (see Figure 31). Luciferase positive seedlings were isolated and transplanted in soil and grown to maturity.

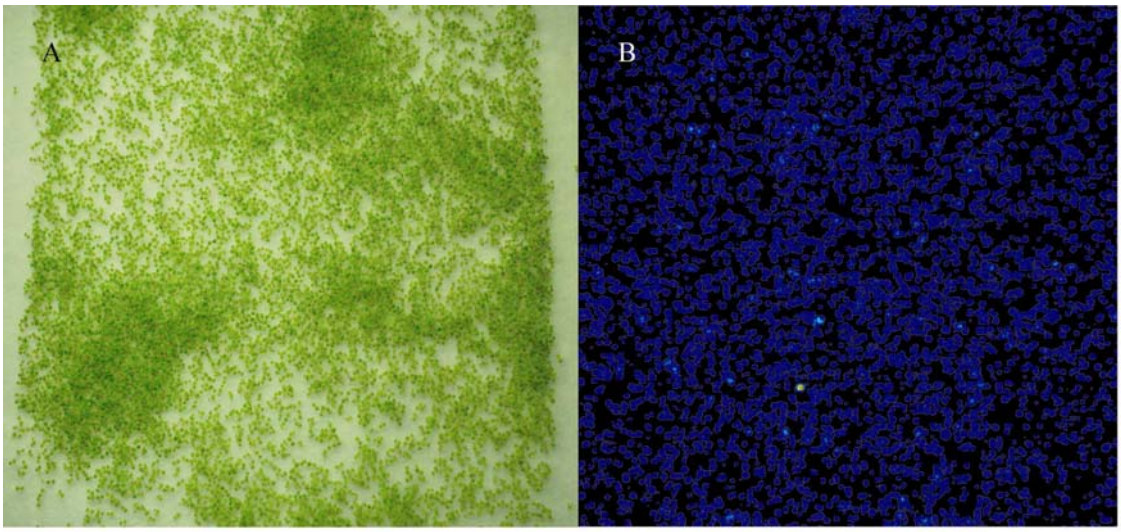


Figure 31: Screening of homo- and hemi-synthRBCSB seedlings for Luciferase activity. Panel A, Color digital image of 5-day old seedlings germinating on chromatograph paper wetted with Hoaglands solution in a 20 cm × 25 cm square flat. Panel B, Photon-counting image of the same field as panel A. The red spot represented a bioluminescence from a single luc+ seedling. Blue spots indicated background "heat photons".

Two Luc^+ recombinants were identified in the homozygous T5 population and two luc^+ seedlings were identified in the F2 population from the hemizygous plants, see Table 1. Thus, there was no significant difference in the observed frequency of unequal crossing-over between the parental populations.

Line	luc+	Total screened	Frequency
F1 hemizygous population	2	4.0 million	5.0×10^{-7}
T4 homozygous population	2	3.8 million	5.3×10^{-7}

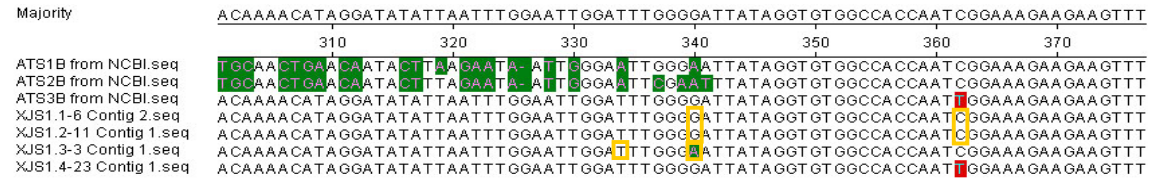
Table 1: Luc^{+} meiotic recombinants of synth*RBCSB*.

The four luc^{+} seedlings were isolated, transplanted to soil, and grown to maturity.

5.2.2. Determine recombination templates and breakpoints

To identify which of the parental RBCSB genes underwent meiotic unequal crossing over and to map the location of the meiotic recombination resolution sites, genomic DNA was purified from the leaves and used in PCR reactions to amplify the luc^{+} recombinant chimeric *RBCSB* genes. One of the oligonucleotide primers corresponding to a highly conserved sequence in the 5' region of all three *RBCSB* genes and one primer specific to the *LUC* gene (oJGJ13 and oJGJ14, respectively) were used to amplify the recombinant genes sequences by PCR. The ~900 bp PCR products were then subcloned into the pCR2.1-TOPO vector for sequencing using M13 forward and M13 reverse primers. To determine the location of the unequal cross-over resolution sites in each of the recombinant luc^{+} alleles, the recombinant chimeric genes were aligned to one another as well as the parental *ATS1B* (*RBCS1B*), *ATS2B* (*RBCS2B*) and *ATS3B* (*RBCS3B*) genes.

A:



B:

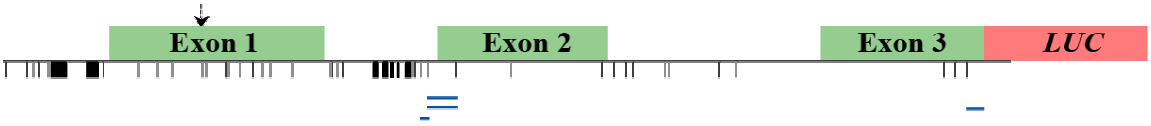


Figure 32: Alignment of the recombinant genes with parental *RBCS1B, RBCS 2B and RBCS 3B* genes used to determine recombination templates and breakpoints. Panel A, alignment of recombinant chimeric genes (XJS1.x) with parental *ATS1B* (*RBCS1B*). *ATS2B* (*RBCS2B*), and *ATS3B* (*RBCS3B*) DNA sequences. Yellow boxes indicate the nucleotides polymorphisms that bound the interval where the DNA sequences from one parental genotype transitions to another (thereby defining the region where recombination resolution must have occurred.. Panel B, a representation of where the apparent resolution sites occurred within the RBCSB gene organization. Short vertical lines indicate nucleotides polymorphisms between *RBCS1B* and *RBCS3B* genes. Blue bars indicate the interval where recombination resolution occurred. The arrow in exon 1 is the *PflM*I site where the *RBCS1B* sequence starts. The length and position of each component is in scale except for the length of *LUC* gene.

The four recombinant chimeric genes were all *RBCS3B/1B* chimeric genes (Figure 32A). No *RBCS2B/1B* chimeric genes were identified in this screen. This was consistent with the previous data (Jelesko *et al.*, 1999). The recombination breakpoints were shown in Figure 32B.

5.3. Conclusion and Discussion:

The experiments described in Chapter 5 were designed to address the question of whether zygosity affects meiotic unequal crossing-over frequency at the *synthRBCSB* locus. A synthRBCSB1-10 transgenic line was used to generate a large hemi-zygous and homozygous populations that were allowed to undergo meiotic unequal crossing-over. Approximately 3.8 to 4.0 million progeny seedlings were assayed for *in vivo* luciferase activity and two luc^+ seedlings were identified and isolated from each population. These results did not support the hypothesis that the frequency of meiotic unequal crossing-over between homologous chromosomes was either additive or more frequent than the frequency of meiotic sister chromatid exchanges.

Somatic recombination between two halves of a transgenic reporter gene is approximately two times more frequent in a homozygous transgenic line (i.e. between homologous chromosomes) compared to a hemizygous transgenic line (i.e. between sister chromatids) (Lebel *et al.*, 1993; Molinier *et al.*, 2004). So it was of great interest to determine if a similar effect occurs during meiotic recombination between homologous chromosomes compared to sister chromatids. Our data yielded two luc^+ recombinants in 4.0 million progeny from a hemizygous line and two luc^+ recombinants in 3.8 million progeny from a homozygous population of the transgenic line (Table 1). Thus, no significant difference was observed in meiotic unequal crossing over in the *synthRBCSB.1* line during meiotic recombination in the homozygous and hemizygous populations. At face value, this might indicate that meiotic unequal crossing-over between paralogous genes on homologous chromosomes occurs as frequently as unequal crossing-over between the same paralogous genes located on sister chromatids. However, this conclusion should be considered with some degree of caution, because the relatively few observed luc^+ recombinants in each population precludes rigorous statistical support and might represent stochastic events.

The absolute estimate of unequal crossing-over frequency at the *synthRBCSB1-10* transgenic locus was quite different from previous reports of this same transgenic line.

Previously, the frequency of meiotic unequal crossing-over in hemizygous lines of the *synthRBCSB* gene cluster was estimated at $\sim 3 \times 10^{-6}$ (Jelesko *et al.*, 1999; Jelesko *et al.*, 2004). However, these previous estimates of meiotic sister chromatid exchange are almost 10 times higher than those reported in Table 1. The two previous estimates of *synthRBCSB* meiotic unequal crossing-over between sister chromatids were all based upon hemizygous populations that were generated at the same time in the same greenhouse conditions by Prof. Jelesko when he was postdoctoral fellow at UC Berkeley. While the environmental conditions of a greenhouse are relatively more stable than plant growth during field conditions, there is still substantial variability in both temperature and especially lighting. For example, during the course of a day the temperature and light quality in a greenhouse can vary substantially depending upon the external weather conditions. Moreover, the difference in light quality and fluence can vary from day to day and even hour to hour. This amounts to a substantial degree of experimental variability. In contrast, in this study the homozygous and heterozygous *synthRBCSB* populations were germinated and grown in environmental growth chambers where the temperature and light conditions were well controlled. Consistent with this assertion, many kinds of environment stresses (e.g. temperature, biotic, oxidative stress, etc) affect overall somatic recombination frequency (Kovalchuk *et al.*, 2003; Lebel *et al.*, 1993; Ries *et al.*, 2000a; Ries *et al.*, 2000b). Thus, the estimates of meiotic unequal crossing over reported here may represent the lower range of possible meiotic unequal crossing-over rates at the *synthRBCSB* locus.

The four *RBCSB* recombinant genes isolated from luc^+ plants were all *RBCS3B/1B* chimeric genes. No *RBCS2B/1B* gene was isolated. These results are consistent with previous reports using the same transgenic *synthRBCSB1-10* line. There were several possible reasons for the absence of luc^+ plants with a *RBCS2B/1B:LUC* chimeric gene. One formal possibility is that there was recombination between *RBCS2B* and *ΔRBCS1B:LUC* resulting in chimeric *RBCS2B/1B:LUC* genes, but not resulting in expression of a chimeric protein due to an inadvertently truncated *RBCS2B* promoter during the cloning of the *synthRBCSB* gene cluster (Jelesko *et al.*, 2004). Another possibility is that a *RBCS2B/1B:LUC* fusion would lead to a non-functional or misfolded protein fusion. However, given the high degree of protein similarity between the *RBCS2B* and *RBCS3B*

proteins (96.7%) and given that *RBCS3B/1B:LUC* chimeric proteins have luciferase enzyme activity, this formal possibility seems unlikely.

A more likely explanation has to due with the fact that in order to maintain the relative distance between the three *RBCSB* genes in the *synthRBCSB* gene cluster, only 133 bp of DNA sequence 5' of the *RBCS2B* ATG start codon is present in the *synthRBCSB* gene cluster. Thus, we speculate that this region lacks sufficient promoter activity to drive transcription of the cloned *RBCS2B* gene within the *synthRBCSB* gene cluster. If this is the case, it would be at odds with the requirement for the reconstitution of a functional promoter associated with a putative recombinant *RBCS2B/1B:LUC* reporter gene. Therefore, the dearth of luc^+ plants with a recombinant *RBCS2B/1B:LUC* gene is consistent with the hypothesis that the *RBCS2B* gene within the *synthRBCSB* gene cluster has a severely truncated promoter that is not capable of effectively initiating transcription of the *RBCS2B* gene within the *synthRBCSB1* gene cluster. With that said, this does not preclude the possibility that recombinant *RBCS2B/1B:LUC* genes can form by unequal crossing-over, but rather these recombinant chimeric genes would go undetected because they would not confer a luc^+ phenotype. In fact, this prediction was directly tested and the results of these experiments are described in the next chapter of this dissertation.

Chapter Six: Somatic recombination of a *synthRBCSB* gene cluster

6.1. Introduction:

Research using reporter gene-based assays of homologous recombination in plants have measured either somatic recombination or meiotic recombination, but rarely both types of homologous recombination in a single generation. Overlapping deletions of either *GUS* or *LUC* reporter genes are extensively used to measure the frequency of somatic recombination in plants subjected to a variety of abiotic and biotic stressors (Molinier *et al.*, 2004; Kovalchuk *et al.*, 2003; Swoboda *et al.*, 1993). In these experiments, somatic recombination rates are calculated as the total number of recombinant luc^+ spots per juvenile plant. While these measurements provide statistically significant measures of recombination frequency, they do not provide absolute measures of recombination rates because the total number of cells per plant is not the same from plant to plant. Another disadvantage of utilizing overlapping segments of a reporter gene is the absence of DNA polymorphisms within the overlapping segments of the reporter gene. This makes it impossible to map the recombination resolution sites within reconstituted reporter genes. Thus, one cannot determine the character of the somatic recombination events leading to the reconstitution of the reporter genes used to assay recombination frequencies.

In contrast, recent measures of meiotic recombination used unequal crossing-over between paralogous plant genes within *synthRBCSB* gene clusters to measure meiotic recombination rates. In these experiments, the appearance of luc^+ seedlings is used to estimate recombination rates that occurred during meiotic unequal crossing-over in the previous generation (Jelesko *et al.*, 1999; Jelesko *et al.*, 2004). The advantage of this approach is that each seedling is the product of two distinct gametes formed by two independent meiotic recombination events. Therefore, this measure of recombination frequency is inherently normalized to the number of cells undergoing meiosis. This is in marked contrast to the above measures of somatic recombination that standardize the number of

recombinant events (i.e. spots) per whole plant. Thus, it is uninformative to directly compare these two different measurements of recombination frequency because they are normalized to completely different scales (i.e. per plant vs. per meiosis).

One theoretical solution to this problem is to develop a measure of somatic recombination that is normalized to a per-cell scale. This would require the estimation of the total number of cells in an Arabidopsis plant at a specific age. Unfotunately, it is technically difficult to effectively sample all the cells in a whole plant. This is further confounded by the fact that there is variation in plant-to-plant vigor, overall plant size, and therefore cell number. Another complicating factor is the phenomenon of endoreduplication in which many cells have undergone additional rounds of chromosomal duplication without cell division and therefore have increased chromosomal ploidy levels on a per cell basis.

A more optimal solution would be to normalize somatic recombination rates to a per genome equivalent scale. This would require isolation of a defined amount of genomic DNA from a given somatic tissue and quantify both: i) the absolute number of genome equivalents present and ii) the absolute number of recombinant *RBCS3B/1B:LUC* genes present. The ratio of the absolute number of recombinant chimeric (e.g. *RBCS3B/1B:LUC*) gene copies relative to the total number of genome equivalents would be an estimate of the somatic recombination frequency in that tissue. One advantange of a per genome equivalent normalization scale is that it accounts for all genomes regardless of the chromosomal endoreduplication status of the cells. A somatic recombination frequency based on a per genome equivalent scale could be compared to 50% of a meiotic recombination frequency based on a per seedling estimate, because all luc^+ seedlings are derived from two haploid gamates (each derived from either a diploid micro- or mega-sporocyte). Thus, normalization of somatic recombination frequency to a per genome equivalent scale provides an opportunity for meaningful comparisons to meiotic unequal crossing-over frequencies derived from the same populations.

Quantitative Real Time PCR (QRT-PCR) is the assay of choice for estimating the absolute copy number of particular DNA sequence in a sample. QRT-PCR is analytically sensitive

over a wide dynamic range (approximately five orders of magnitude) and one can develop gene-specific oligonucleotide primers to amplify target gene sequences. Using a standard curve of known amounts of a cloned gene of interested and derived Ct values, one can generate an accurate estimate of the absolute number of gene-specific DNA molecules present in any genomic DNA sample. Therefore, one can estimate the absolute copy number of any target gene, including the chimeric *RBCS2B/1B:LUC* or chimeric *RBCS3B/1B:LUC* genes that are generated by unequal crossing over within somatic tissue..

One interesting result from Chapter 5 and from previous *synthRBCSB* meiotic recombination research (Jelesko *et al.*, 1999; Jelesko *et al.*, 2004) was that all luc^{+} recombinant alleles were *RBCS3B/1B:LUC* chimeric genes. In other words, they are the products of unequal crossing-over between the linked *RBCS3B* and *ΔRBCS1B:LUC* genes. There is a conspicuous absence of *RBCS2B/1B:LUC* alleles isolated from these screens. As discussed in Chapter 5, this does not necessarily indicate that the *RBCS2B* gene cannot recombine with *ΔRBCS1B:LUC*. Rather, it is most likely due to a non-functional truncated *RBCS2B* promoter that obviates a gain-of-luciferase phenotype; Therefore, chimeric *RBCS2B/1B:LUC* genes cannot be identified in the *in vivo* luciferase activity screen. On the other hand, the posited *RBCS2B/1B:LUC* chimeric genes should be detectable by carefully designed QRT-PCR assays.

The large F1 hemi- and T4 homozygous *synthRBCSB* populations used to investigate meiotic unequal crossing-over in the Chapter 5 were an excellent resource to investigate the frequency and character of somatic recombination using a QRT-PCR approach. In short, the proportion of recombinant *RBCS2B/1B:LUC* and *RBCS3B/1B:LUC* alleles relative to the total number of genome equivalents would comprise a measure of somatic unequal crossing over in these plant populations. Since the number of recombinant chimeric genes could be estimated by QRT-PCR methods, the resulting PCR amplification products could also be readily subcloned and sequenced as part of a meta-genomics analysis to determine relative recombination frequencies between the target *Δ1B:LUC* gene and the *RBCS2B* or *RBCS3B* genes. This metagenomics sequence data would also allow mapping of recombination resolution sites. These somatic recombination data could be compared to

the meiotic recombination data from the same plant populations described in Chapter 5, and thereby identify any significantly different patterns between somatic and meiotic unequal crossing-over at the *synthRBCSB* gene cluster.

6.2. Experimental design and results:

6.2.1. Amplification and cloning of somatic recombinant RBCS2B/1B:LUC and RBCS3B/1B:LUC genes

In order to study somatic recombination at the *synthRBCSB* gene cluster, QRT-PCR experiments were designed to monitor unequal crossing-over between the paralogous *RBCS2B* or *RBCS3B* genes with the *ΔRBCS1B:LUC* target gene in a transgenic *synthRBCSB* gene cluster. Genomic DNA was isolated from three different tissues (leaf, stem, and floral bud clusters) from F1 hemizygous *synthRBCSB* plants and T4 homozygous *synthRBCSB* plants that were previously developed to measure meiotic unequal crossing-over frequencies (see Chapter 5). A nested PCR approach was implemented because of the formal possibility that somatic unequal crossing-over frequencies might be so low that a single round of PCR would not be sufficiently sensitive to detectably amplify rare templates.

In order to amplify both *RBCS2B/1B:LUC* and *RBCS3B/1B:LUC* chimeric genes, two nested oligonucleotide primer pairs were developed that should amplify both predicted chimeric genes. The oligonucleotide primers needed to selectively amplify the recombinant chimeric *RBCS2B/1B:LUC* and/or *RBCS3B/1B:LUC* genes the formed within the transgenic *synthRBCSB* transgene, without amplifying any of the endogenous *RBCSB* genes located on Chromosome 5. This specificity was achieved by designing the upstream oligonucleotide primers to DNA sequences that are well conserved in the promoters of all three *RBCSB* genes, while the down stream primers bound to *LUC* specific sequences. Nested PCR reactions used oligonucleotide primer pair oJGJ13 and oJGJ14 in the first PCR reaction and oJS112 and oJS113 (see Section 2.11) in the second PCR reaction. Using

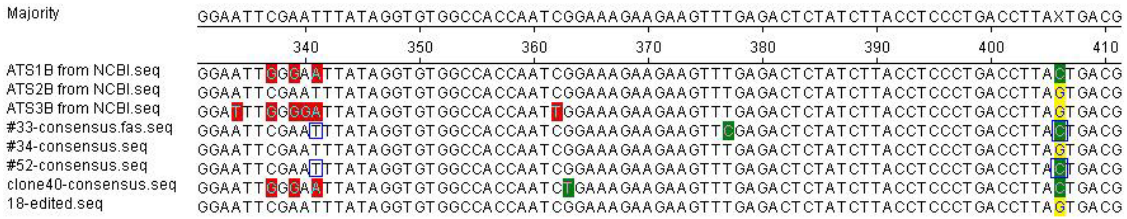
gDNA isolated from homozygous T4 generation *synthRBCSB* plants in nested PCR reactions an expected ~900 bp band was amplified from gDNA from leaves, bolts (stems), and flower buds (Figure 33A).

The PCR amplification products were subcloned and five random clones containing the ~900 bp fragment were sequenced. DNA multiple sequence alignments indicated that four clones were comprised of *RBCS2B/1B:LUC* chimeric genes and one clone was a *RBCS3B/1B:LUC* chimeric gene (Figure 33B and C). This was the first report of a chimeric *RBCS2B/1B:LUC* gene forming within the *synthRBCSB* gene cluster.

The multiple sequence alignments were also used to map the location of the apparent somatic recombination resolution site(s). Figure 33C illustrates that three of the four *RBCS2B/1B:LUC* resolution sites mapped within exon-2. This result is interesting because the analogous exon-2 region in *RBCS3B/1B:LUC* meiotic recombination is a documented meiotic recombination cold spot (Jelesko *et al.*, 2004). The resolution site of the single *RBCS3B/1B:LUC* recombinant mapped to a region in exon-1 that has not been observed in meiotic *RBCS3B/1B:LUC* recombinants and is considered a meiotic recombination cold spot. Thus, four of the five somatic recombination resolution sites did not map to locations that are commonly observed during meiotic recombination.

A.

B.



C.

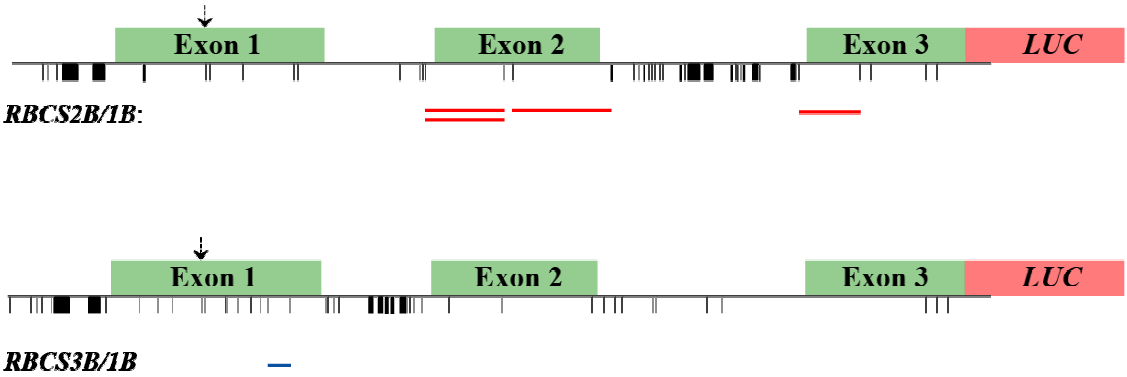


Figure 33: Characterization of recombinant chimeric *RBCSB* genes that formed by somatic recombination. Panel A, agarose gel of amplification products from gDNA isolated from different homozygous *synthRBCSB* plant tissues. Template gDNA used for PCR reactions. Lanes 1 – 6, leaf; lanes 7 – 8, stem; lanes 9 – 10, flower; lane 11 gDNA from a luc^{+} meiotic recombinant plant XJS1.1 used as a positive control; and lane 12, Col-0 gDNA used as a negative control. Panel B, a portion of a multiple DNA sequence alignment used to identify the chimeric genes and approximate location of recombination resolution sites. Panel C, the position of recombination resolution sites for the five cloned recombinant chimeric genes. Blue and red bars indicate where somatic recombination resolution occurred in each clone.

Short vertical lines indicate nucleotides polymorphisms. The length and position of each component is to scale except for the *LUC* gene.

6.2.2. Experiments to rule out the possibility of PCR template switch

The atypical somatic recombination resolution sites (relative to the more well characterized meiotic resolution sites) raised the possibility that these apparent single cross-over events were artifacts of PCR amplification caused by template switching during the PCR reactions. To rule out the possibility that these amplification products were the artifacts of PCR template switch, PCR experiments were performed using *RBCS2B, RBCS3B* genes, and *ΔRBCS1B* genes residing on separate plasmids. In these experiments, no PCR amplification products should form unless there some amplification products derived one plasmid anneal with a different plasmid, see Figure 34. Template switching during PCR is not an effective

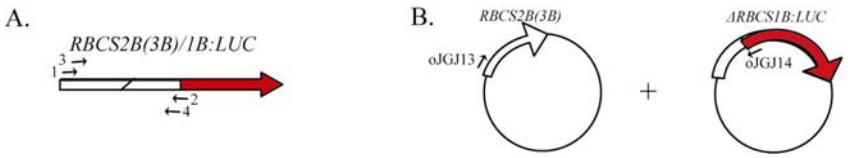


Figure 34: Template switch experiment. Panel A, the positions of nested oligonucleotide primers to amplify recombinant chimeric *RBCS2B(3B)/1B:LUC* genes: 1, oJGJ13; 2, oJGJ14; 3, oJS113; and 4, oJS114. Slashed line indicates the position of resolution site. Red arrow indicates *LUC* gene. Panel B, plasmid pairs used in control template switch PCR reactions: *ΔRBCS1B:LUC* gene (pJGJ196), and either *RBCS2B* gene (pJGJ189) or *RBCS3B* gene (pJGJ192).

reaction and is favored at high template concentrations. However, when template concentrations are below a critical concentration template switching should not occur. Therefore, a serial dilution series of equal amounts of each plasmid template were used to establish the minimum template concentration requirement for template switching to occur. Therefore, below this critical template concentration value no template switching will occur and any amplified chimeric genes should be derived from bona fide chimeric genes.

Nested PCR reactions were performed with serial dilutions of plasmid pJGJ189 (*RBCSB2B*) or plasmid pJGJ192 (*RBCSB3B*) combined with plasmid pJGJ196 (*ΔRBCS1B:LUC*). Since none of the three plasmids contained a *RBCS(2B or 3B)* promoter positioned 5' to the *LUC* gene, no PCR products should be produced unless template switching was occurring, which would result in a 900 bp amplification product. As illustrated in Figure 35, at plasmid concentrations greater than 30 million copies, 900 bp PCR amplification products were produced, indicating that PCR template switching occurred.

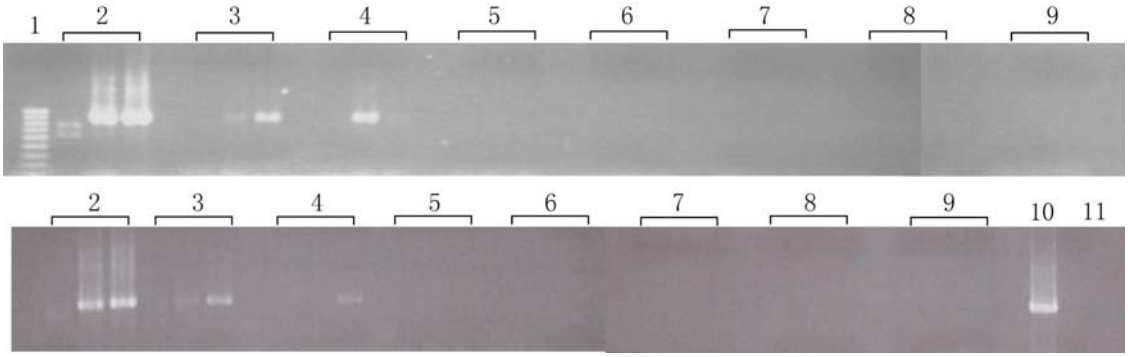


Figure 35: Template switching experiment. 1: 100bp maker. 2~9: PCR with 10X template dilution corresponding to that in table 1. The order of the three lanes in each group was pJGJ196 (*RBCS1B*), pJGJ196 + pJGJ189 (*RBCS1B* + *RBCS2B*), pJGJ196 + pJGJ192 (*RBCS1B*+*RBCS3B*). 10: *RBCS3B*/*1B* chimeric gene as positive control. 11: No template negative control. Shown here were two replicates of the four.

Number of molecules added into PCR reaction		pJGJ196				pJGJ196+pJGJ189				pJGJ196+pJGJ192			
		R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4
2	3,000,000,000	-	-	-	-	+	+	+	+	+	+	+	+
3	300,000,000	-	-	-	-	+	+	+	+	+	+	+	+
4	30,000,000	-	-	-	-	-	+	-	-	+	+	+	+
5	3,000,000	-	-	-	-	-	-	-	-	-	-	-	-
6	300,000	-	-	-	-	-	-	-	-	-	-	-	-
7	30,000	-	-	-	-	-	-	-	-	-	-	-	-
8	3,000	-	-	-	-	-	-	-	-	-	-	-	-
9	300	-	-	-	-	-	-	-	-	-	-	-	-

Table 2: a summary of template switch experiment. R1-R4: four independent replicates. "+" indicates template switch occurred to produce a ~900bp band. "-" indicates no amplification product.

However, as shown in Table 2, when plasmid template concentrations were less than 3,000,000 copies no PCR amplification products were observed. The amount of genomic DNA (6 - 20ng) used in the nested PCR amplification reactions that amplified the putative chimeric *RBCS2B/1B:LUC* and *RBCS3B/1B:LUC* genes from homozygous *synthRBCSB* plants (Figure 35) was estimated at between 35,000 - 120,000 haploid genome equivalents. Because these amounts of gDNA equivalents were much less than the observed threshold value of 3,000,000 DNA copies required to generate template switching during the control nested PCR reactions, it was concluded that the nested PCR amplification products generated from *synthRBCSB* plant gDNA represented bona fide recombinant chimeric alleles that formed by somatic recombination. These results validated the utility of using a nested PCR approach to amplify and isolate chimeric *RBCS(2B* or *3B)/1B:LUC* genes that formed by somatic recombination.

6.2.3. Using Quantitative Real-Time PCR to estimate somatic recombination frequencies

To quantify the number of chimeric *RBCS2B/1B:LUC* and *RBCS3B/1B:LUC* gene copies in a given sample a two step nested QRT-PCR approach was utilized. Using gDNA from homozygous *synthRBCSB* leaves, stems, and flowers the first round of regular PCR was performed using oligonucleotide primer pair oJS123 and oJGJ14. An aliquot of each first PCR reaction was used in a subsequent round of QRT-PCR reactions using oligonucleotide primers oJS112 and oJS113 (see Chapter 2, Materials and methods). A standard curve of haploid genome equivalents was generated using a ten-fold serial dilution series of gDNA isolated from a hemizygous line containing a *RBCS3B/1B:LUC* luc^+ plant (XJS1.2). Figure 36A illustrates the QRT-PCR amplification plots for the gDNA isolated from homozygous *synthRBCSB* tissues. These QRT-PCR estimates of gene copy number were the sum of *RBCS2B/1B:LUC* and *RBCS3B/1B:LUC* recombinant chimeric gene copies in each sample. Figure 36B shows the standard curve used to estimate the number of chimeric *RBCS2B(3B)/1B:LUC* DNA molecules in each sample. The standard curve gave an acceptable R^2 of 0.98.

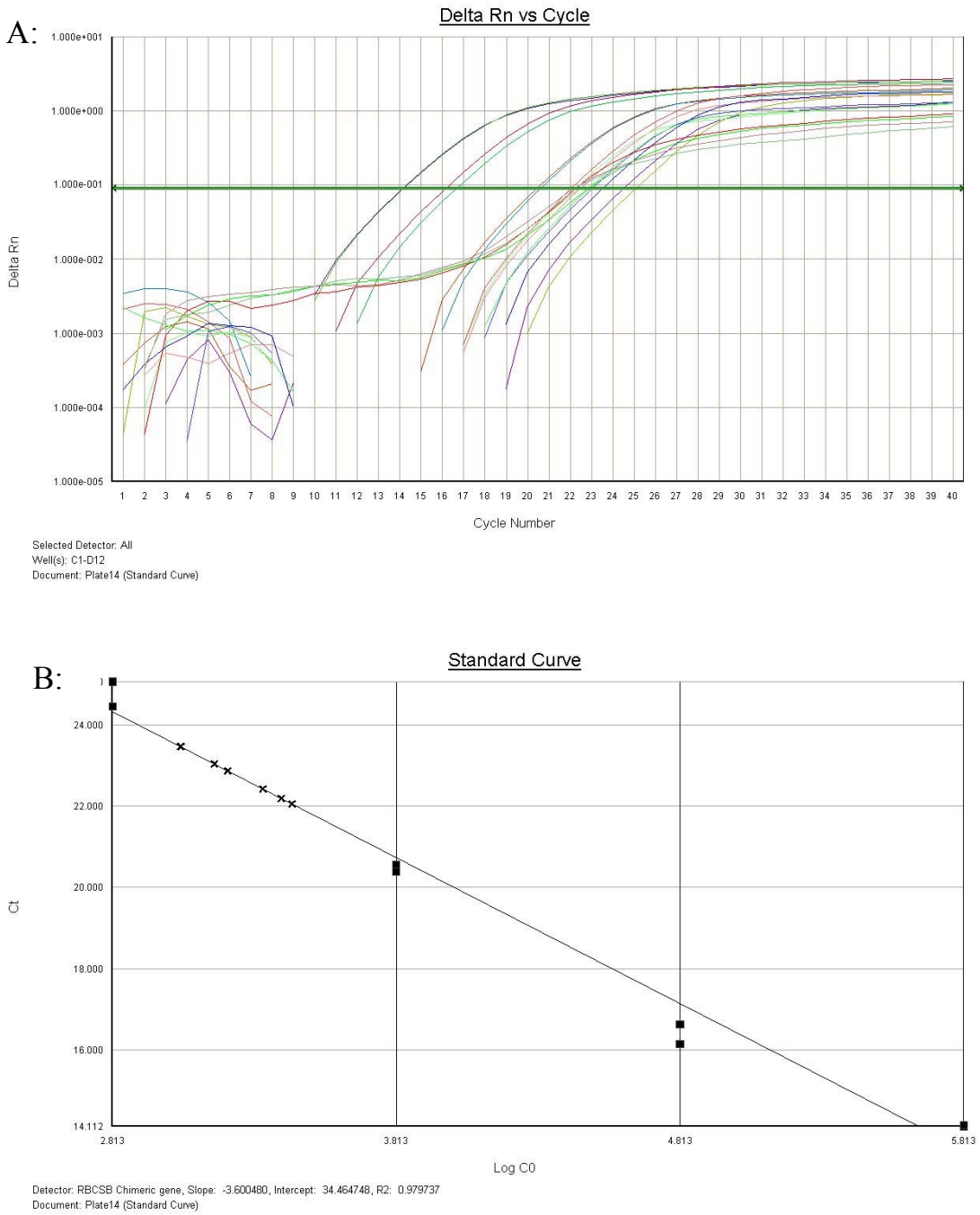
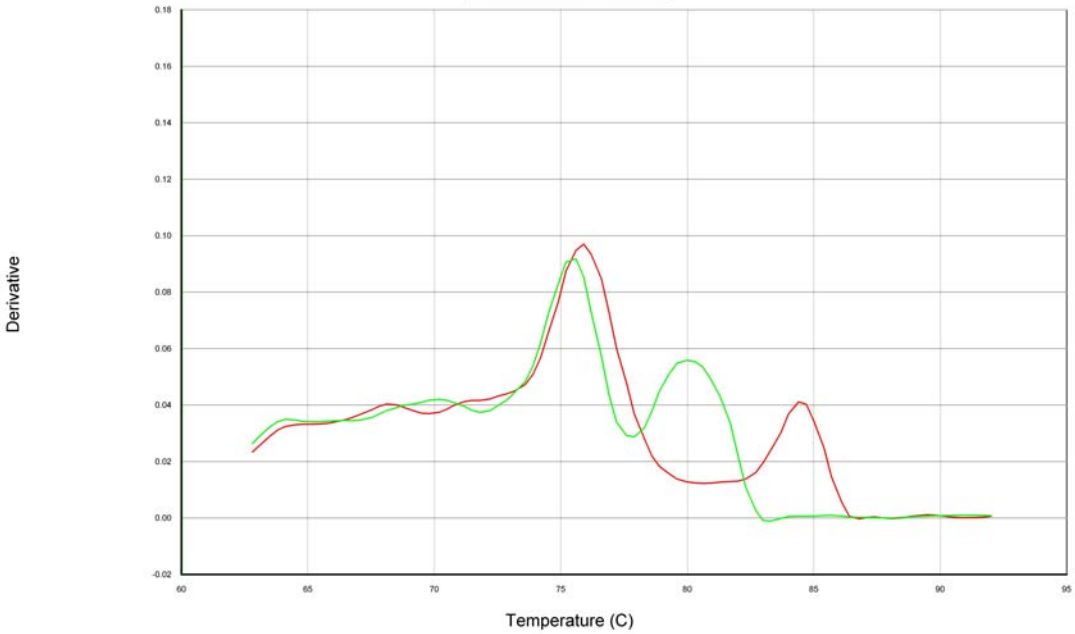


Figure 36: QRT-PCR of *RBCS(2B/3B)/1B:LUC* recombinant chimeric genes that formed by somatic unequal crossing-over. Panel A, The amplification plots of QRT-PCR standards and samples. Panel B, Standard curve of Ct vs.

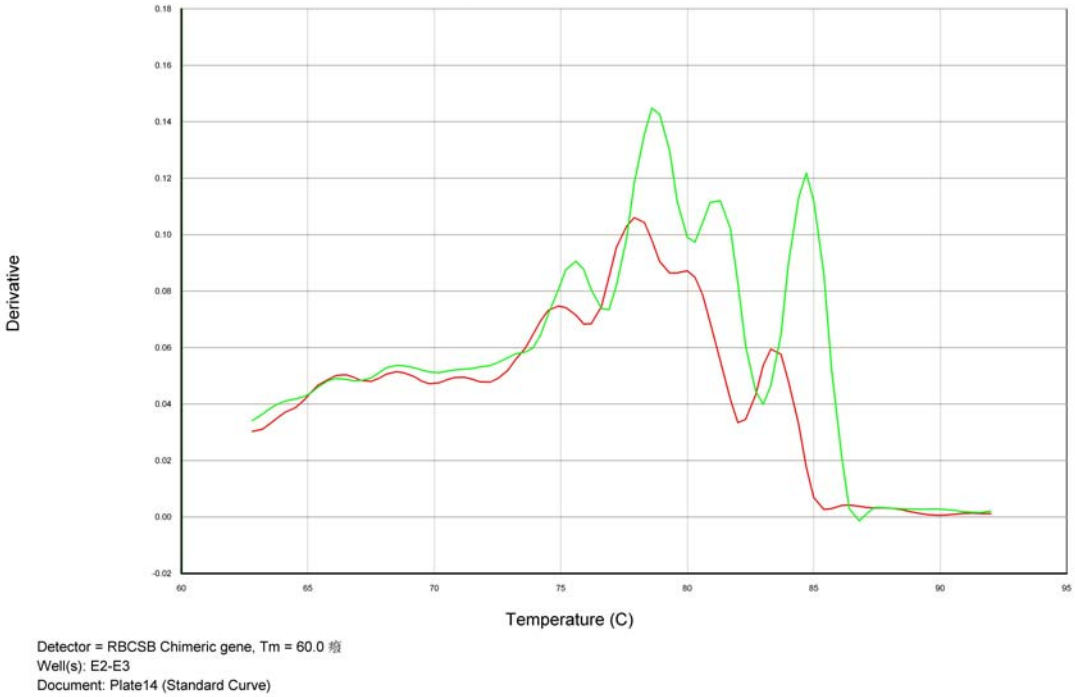
log XJS1.2 genome equivalents and estimated combined *RBCS(2B/3B)/1B:LUC* gene copy number (R^2=0.98).

All samples containing a suspected *RBCS(2B or 3B)/1B:LUC* recombinant chimeric gene showed disassociation curves with four peaks, Figure 37. This brought forward the question if these peaks represented four different amplicons or a single long PCR product (~900 bp) with complex melting properties. To address this question, control QRT-PCR experiments were performed using different single template DNA samples (e.g. cloned chimeric recombinant genes or gDNA from a stable meiotic *RBCS3B/1B:LUC* recombinant) or combinations of these DNA templates. Control PCR reactions lacking input template DNA showed amplification products that were less than 100 bp in length (panel I lanes 31 and 32) and a consistent disassociation peak at approximately 75^{o}C (panel A). Thus, the 75^{o}C disassociation peak identified background primer dimers. All QRT-PCR reactions showed evidence (75^{o}C disassociation optima) of this background primer dimer amplification product (panels A – H). QRT-PCR reactions using cloned *RBCS2B/1B* or *RBCS3B/1B* genes (panels C and E, respectively) as template DNA showed a complex melt curve pattern, comprised of three additional peaks (79 ^{o}C , 81 ^{o}C, 84.7 ^{o}C). However, when the QRT-PCR products were run on an agarose gel (panel I, lanes 2–24), they all showed a single dominant amplification product at the expected 900 bps. This demonstrated that the complex melt pattern was due to a single PCR amplification product produced from the cloned chimeric *RBCS(2B or 3B)/1B:LUC* genes. Similarly, when these *RBCS2B/1B* or *RBCS3B/1B* gene containing plasmids were mixed together with an excess of (110ng) Col-0 genomic DNA, the same melt curve patterns were observed (panels D and E) and a dominant 900 bp PCR product was produced (panel I, lanes 25 and 26). These results demonstrated that the long PCR products had complex melting kinetics and there were no synergistic effects (i.e. template switching) when combined with Arabidopsis genomic DNA. The contribution of primer dimer products on the quantification of the 900 bp amplification products was eliminated by reading the fluorescent data at 79^{o}C, at which the primer dimer dissociated (panel A) and was not contributing a SYBER Green fluorescent

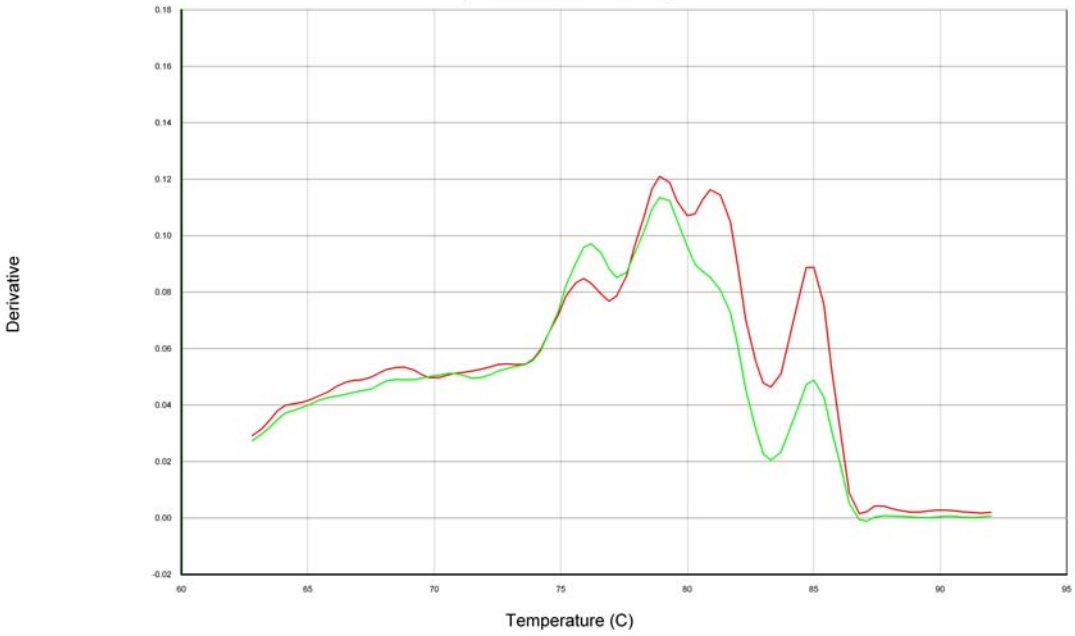
A: No template control



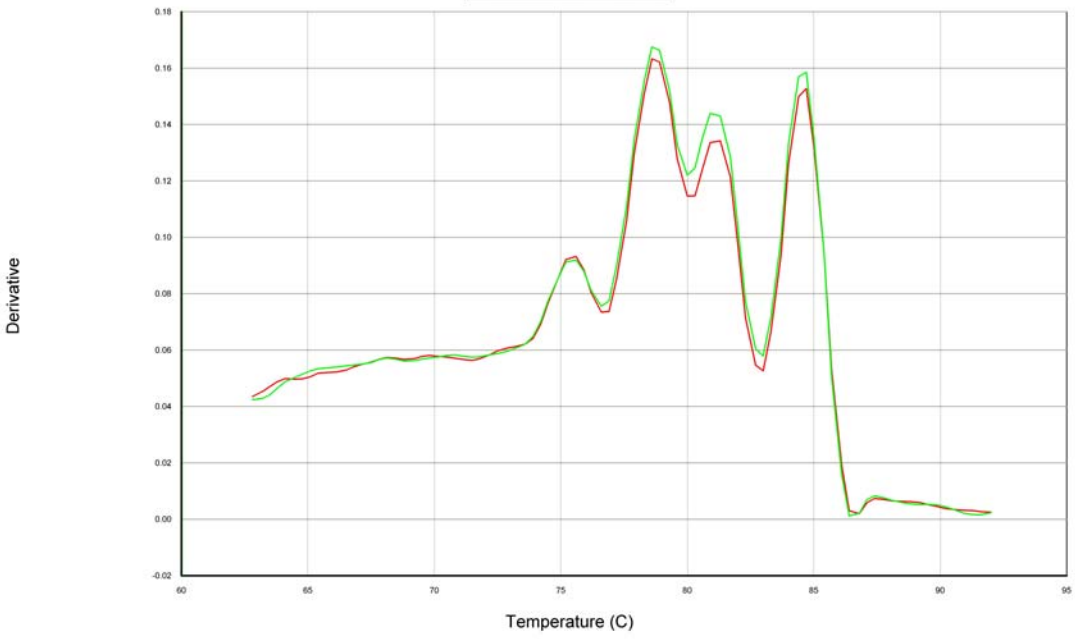
B: pJS29 (*RBCS2B/1B* recombinant gene in pCR2.1)



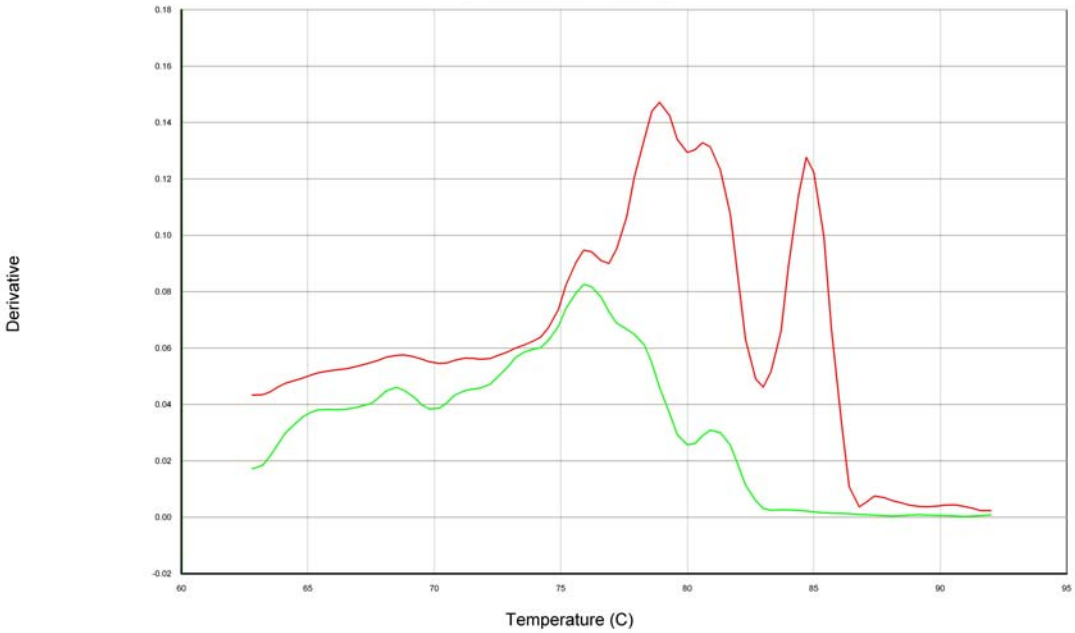
C: pJS30 (*RBCS3B/1B* recombinant gene in pCR2.1)



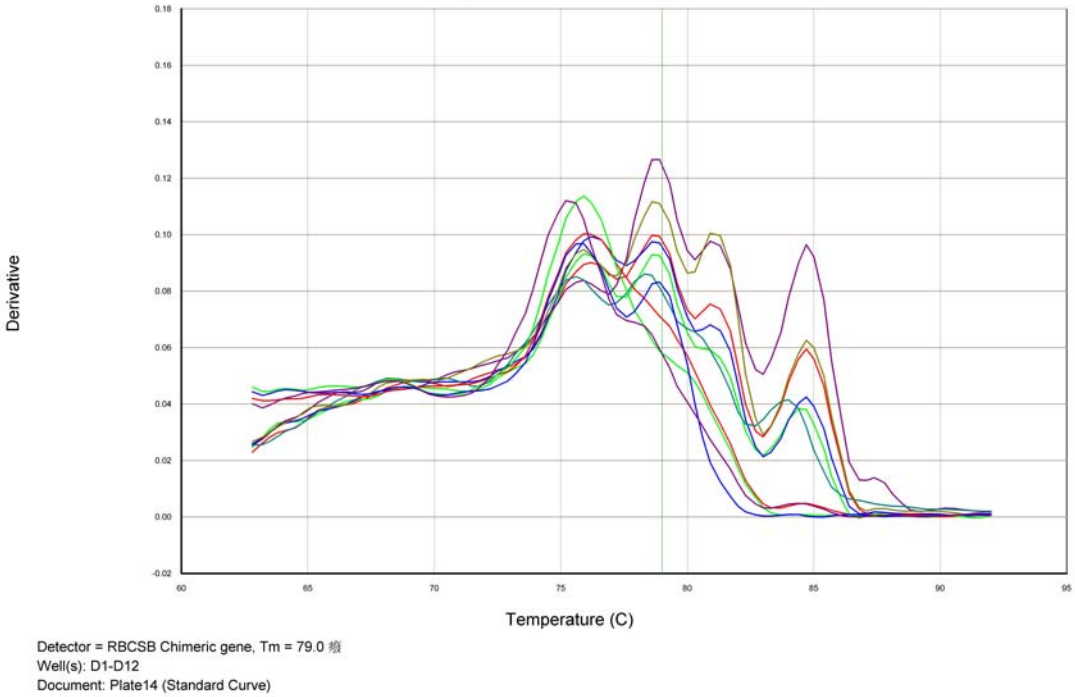
D: pJS29 + Col-0 gDNA



E: pJS30 + Col-0 gDNA

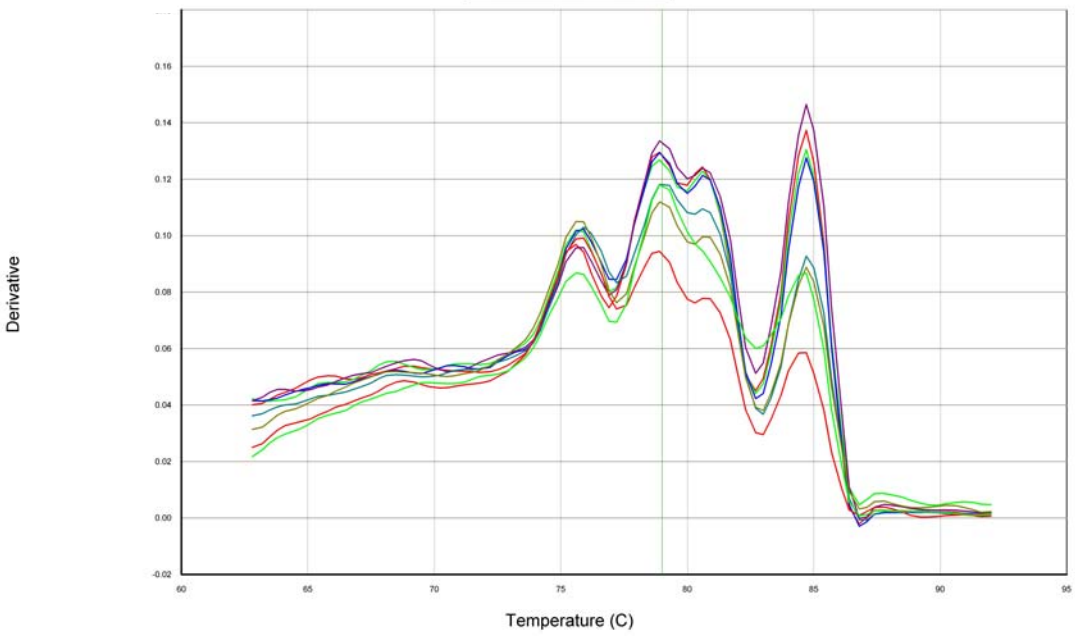


F: gDNA of *synthRBCSB* plants



XJS1.2 gDNA as standards

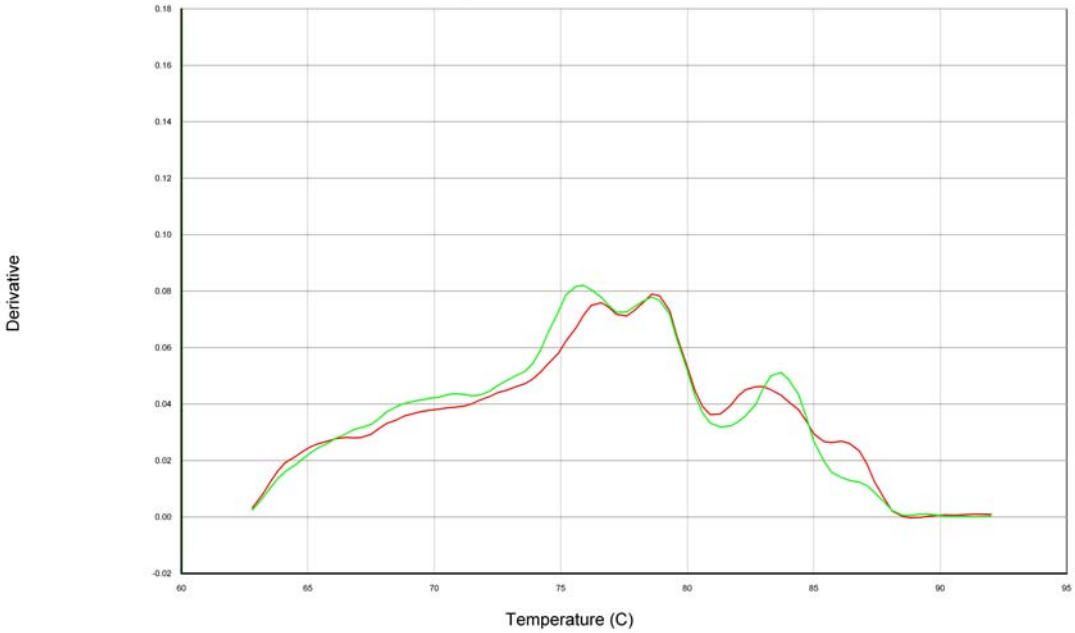
G:



Detector = RBCSB Chimeric gene, Tm = 79.0 癈
Well(s): C1-C12
Document: Plate14 (Standard Curve)

Col-0 gDNA as negative control

H:



Detector = RBCSB Chimeric gene, Tm = 60.0 癈
Well(s): F2-F3
Document: Plate14 (Standard Curve)

I: QRT-PCR products on agrose gel

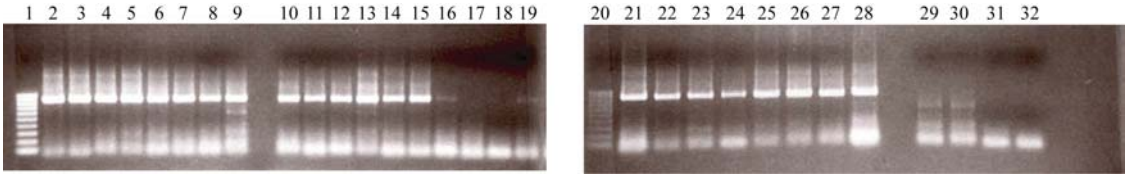


Figure 37: The dissociation curves of the QRT-PCR product.
(A)~(H): dissociation curve of PCR product. (I): the products on the gel. (A): No template control (lane 31, 32 in panel I). (B): pJS29 (cloned *RBCS2B/1B* gene in pCR2.1 vector (Invitrogen®)), corresponding to lane 21, 22 in panel I. (C): pJS30 (cloned *RBCS3B/1B* gene in pCR2.1 vector, corresponding to lane 23, 24 in panel I. (D): B+110ng Col-0 genomic DNA as template, corresponding to lane 25, 26 in panel I. (E): C+110ng Col-0 genomic DNA as template, corresponding to lane 27, 28 in panel I. (F): genomic DNA from hemi- and homo-zygous *synthRBCSB* plant tissues as template. The products corresponded to lane 10~19. 10~17 were amplified from different tissues of one homozygous plant (10~16 from individual leaves and 17 from flower cluster). 18, 19 were amplified from the stem and flower cluster of a hemizygous plant. (G): XJS1.2 as template for making standard curve, the products corresponding to lane 2~9 in panel I. (H): Col-0 as template (lane 29, 30 in panel I). Lane 1, 20: 100bp marker. The additional vertical lines in panel F and G indicate temperate 79^{o}C.

signal. The somatic recombination frequencies were calculated from the estimates the number of chimeric *RBCS2B/1B:LUC* and *RBCS3B/1B:LUC* gene copies in each sample devided by the estimated haploid genome equivalents based upon the total amount of gDNA added to the first PCR reaction. Table 3 shows the calculations used to estimate the fraction of recombinant chimeric *RBCSB* genes relative to the total haploid genome equivalents. Based on these data, the somatic recombination frequencies from leaves of a single homozygous *synthRBCSB* plant ranged from 8.7×10^{-4} to 2.2×10^{-3}. This suggested

little leaf-to-leaf variation in the frequency of somatic unequal crossing-over. This level of somatic unequal crossing-over was at least three orders of magnitude higher than meiotic recombination in the same T4 homozygous population (see Table 1). Interestingly, the QRT-PCR estimates of recombinant chimeric *RBCS(2B* or *3B)/1B:LUC* copy number in stems and flowers were below calibration range of the standard curve used to estimate gene copy number. Consistent with this result, when visualized on an agarose gel the intensity

Sample name	Sample type	C_t value	quantity	Starting gDNA (ng/ul)	hapliod gDNA copy number	Recomb. Frequency
XJS1.2	Standard	14.1123	6.50e+05			
XJS1.2		14.1458	6.50e+05			
XJS1.2		16.6229	6.50e+04			
XJS1.2		16.1344	6.50e+04			
XJS1.2		20.3857	6.50e+03			
XJS1.2		20.5547	6.50e+03			
XJS1.2		24.4637	650			
XJS1.2		25.07	650			
Homo2_1	Leaf gDNA	22.4309	2.20e+03	110	1.29e+06	1.70e-03
Homo2_2		22.1956	2.56e+03	110	1.29e+06	1.98e-03
Homo2_3		23.0433	1.49e+03	110	1.29e+06	1.15e-03
Homo2_4		22.0605	2.79e+03	110	1.29e+06	2.15e-03
Homo2_6		22.8765	1.65e+03	110	1.29e+06	1.28e-03
Homo2_7		23.471	1.13e+03	110	1.29e+06	8.74e-04
Homo2_8		Undetermined		110	1.29e+06	
Homo2_F1	Flower gDNA	Undetermined		110	1.29e+06	
Hemi1_S1	Stem gDNA	Undetermined		110	1.29e+06	
Hemi1_F1	Flower gDNA	Undetermined		110	1.29e+06	
Col-0	Col-0 gDNA	29.8777	18.7935			
Col-0		29.8868	18.6847			
NTC	No template control	33	2.55165			
NTC		32.7851	2.92748			

Table 3: Estimation of somatic recombination frequencies in various tissues. Genomic DNA was isolated from leaves, stems and flowers and then assayed by QRT-PCR.

of the 900 bp PCR product produced from stems and flowers (Figure 37 panel I, lanes 16 – 19) was substantially less than that produced from leaves (Figure 37 panel I, lanes 10–19). These results suggest that somatic recombination rates were substantially lower in tissues (inflorescence stems and flower buds) that had transitioned to a reproductive developmental program.

6.3. Conclusion and discussion:

A PCR-based method was developed and used investigate somatic recombination with in a *synthRBCSB* gene cluster. Using this PCR method, recombinant chimeric *RBCS2B/1B:LUC* genes and *RBCS3B/1B:LUC* genes were amplified from leaves, stems and flower tissues from both hemi- and homo-zygous *synthRBCSB* transgenic plants. This method was expanded to a two-step procedure that included QRT-PCR estimates of the absolute copy number of combined *RBCS(2B or 3B)/1B:LUC* genes in a given gDNA sample. These measurements provided a quantitative estimate of somatic recombination frequency in these tissues. The levels of somatic recombination in leaves were on the order of three-orders of magnitude greater than meiotic recombination measured in the same T4 homozygous population. In contrast, somatic recombination levels appeared to be significantly lower in tissues that had differentiated from a foliar to a reproductive developmental program.

As noted in Chapter 5, a recombinant *RBCS2B/1B* gene had never been isolated from previous meiotic recombination screens that utilized a gain-of-luciferase phenotype at the *synthRBCSB* gene cluster. The successful amplification of *RBCS2B/1B* gene demonstrated that unequal crossing-over occurred between *RBCS2B* and *RBCS1B* genes in somatic tissues. These data support the hypothesis that the failure to identify *RBCS2B/1B:LUC* genes was probably due to an inactive *RBCS2B* promoter, rather than some inability to undergo unequal crossing-over.

The estimated somatic recombination frequency ranged from $8.7x10^{-4}$ to $2.2x10^{-3}$ per genome equivalent. Somatic recombination frequency varies greatly in different organisms, ranging from 10^{-4} to 10^{-7} (Puchta *et al.*, 1994; Peterhans *et al.*, 1990; Tovar & Lichtenstein, 1992; Swoboda *et al.*, 1994; Swoboda *et al.*, 1993; Gal *et al.*, 1991; Kovalchuk *et al.*, 2003; Jelesko *et al.*, 2005). Compared with these data, the somatic recombination frequency measured within *synthRBCSB* gene cluster was a higher than previous reports using overlapping segments of a reporter gene. Interestingly, the observed somatic recombination frequencies at this stably integrated *synthRBCSB1* gene cluster were in reasonable agreement with previously published somatic unequal crossing-over frequencies of a similar *synthRBCSB3* gene cluster (comprised of a *ΔRBCS1B:LUC-RBCS3B* two-gene cluster) associated with T-DNA integration/transformation (Jelesko *et al.*, 2005). The reproducibility of somatic recombination frequencies in six leaves from a single homozygous *synthRBCSB1* line, suggests (but does not prove) that there is probably not a lot of variability in the frequency of somatic unequal crossing-over from leaf to leaf on the same plant.

The QRT-PCR assay measuring the number of recombinant chimeric genes did not differentiate between the scenario where an early somatic recombination event that formed a large sector of cells within a leaf, from a scenario with many independent somatic recombination events. The former scenario would result in the production of many cells all containing the same recombination resolution site, whereas the latter scenario would produce many cells each with different recombination resolution sites. Random cloning and sequencing the PCR amplification products (metagenomics) provided important, albeit limited, initial insights into this question. Of the five cloned and sequenced recombinant chimeric genes, four had different recombination resolution sites. This suggests that there were at least four independent somatic unequal cross-over events in that leaf. These results in conjunction with the reproducible estimates of total somatic recombination frequencies from leaf to leaf, suggests that multiple independent somatic recombination events were detected by this assay.

The predominance of isolated *RBCS2B/1B* chimeric genes could be due to two factors. One factor could be a higher PCR amplification efficiency of the chimeric *RBCS2B/1B* gene relative to the *RBCS3B/1B* gene. Another possible factor could be a higher recombination frequency between *RBCS2B* and *ΔRBCS1B:LUC*, relative to *RBCS3B* and *ΔRBCS1B:LUC*. Considering that the *RBCS2B* gene was much closer (approximately 3 kb closer) to the *ΔRBCS1B:LUC* gene the amount of chromosomal misalignment that must occur for mispairing between the *RBCS2B* and *ΔRBCS1B:LUC* genes was probably less than the degree of localized chromosomal misalignment required for mispairing between *RBCS3B* with *ΔRBCS1B:LUC* (Figure 30). Thus, *RBCS2B* probably recombined more frequently with *ΔRBCS1B:LUC* because there were less barriers to misalignment over a smaller region of chromatin.

One of the objectives of this work was to obtain measurements of both somatic and meiotic recombination frequencies from the same population of plants. To this end, the results from Table 3 and Table 1, indicate that somatic unequal crossing-over in T4 homozygous leaves ($2.2x10^{-3}$) was over 4,000 times more frequent than meiotic unequal crossing-over in flowers ($5.3x10^{-7}$)of the same population of plants. This is the first report that provided a direct comparison of somatic recombination frequency with meiotic recombination frequency at the same genetic locus in the same generation of plants. These results suggest that there are large quantitative differences in between the frequency of somatic recombination and meiotic recombination. This has important implications for genome stability. The relatively high somatic unequal crossing-over frequencies observed in these data and other published data are consistent with the role that homologous recombination plays in DNA repair processes, whereby UV induced double strand breaks can be repaired by homologous sequences on the homologous chromosome. However, too much DNA unequal crossing-over between paralogous genes could give rise to detrimental chromosomal rearrangements that include deletions, inversions, and transversions with resulting chromosome instability due to dicentric or acentric chromosomes. These detrimental effects would be most acute in tissues that undergo multiple rounds of cell division. However, they would be less of a concern in cells that have terminally differentiated and are no longer dividing. In short, genome instability in leaves would not

lead to heritable chromosomal defects, but would remain localized to an eventually terminal tissue type. However, a high rate of somatic unequal crossing-over in tissues that are actively dividing could result in chromosomal instabilities that would likely result in compromised cell function or lethality. Thus, it would be advantageous to suppress somatic recombination rates in all but terminally differentiated cells. To this end, the estimates of somatic unequal crossing-over were not equivalent in all tissue types tested.

Three genomic DNA samples from stem and flower tissues didn't produce a detectable amplification signal in the QRT-PCR assay (Table 3). Consistent with very low levels of detection, an agarose gene containing the QRT-PCR reactions (Figure 37 panel I, lanes 17 - 19) showed either little or no PCR amplification products, suggesting a reduced number of chimeric RBCSB genes in in these tissues.. These results suggest that there were much higher somatic somatic recombination rates in leaves compared to stem and flowers. Nevertheless, the stem and flower tissues are mostly composed of somatic cells. These results together suggests that somatic recombination is subject to developmental regulation, such that somatic recombination is suppressed in tissues leading to or associated with the formation of gametes. This finding may also help to explain why transgenic plant lines with overlapping *GUS* or *LUC* transgene segments that show high somatic recombination rates (Puchta *et al.*, 1995; Puchta *et al.*, 1994; Swoboda *et al.*, 1994; Tovar & Lichtenstein, 1992; Jelesko *et al.*, 2004; Kovalchuk *et al.*, 2003) do not produce progeny with a comparable proportion of germinally inherited recombinant alleles. Put differently, plants with high somatic recombination rates in leaves do not result in a proportionally high frequency of germinal transmission of recombined reporter genes. The finding that somatic recombination rates are lower in reproductive tissues has important evolutionary considerations because it insures that overall genome stability is maintained from generation to generation. However, the comparatively low rate of meiotic unequal crossing-over, nevertheless, allows for the evolution of new chimeric genes that are germinally transmitted.

References:

Assaad, F.F. & Signer, E.R. (1992) Somatic and germinal recombination of a direct repeat in Arabidopsis. *Genetics* **132**(2), 553-66.

Baudino, S., Hansen, S., Brettschneider, R., Hecht, V.F., Dresselhaus, T., Lorz, H., Dumas, C. & Rogowsky, P.M. (2001) Molecular characterisation of two novel maize LRR receptor-like kinases, which belong to the SERK gene family. *Planta* **213**(1), 1-10.

Baumgarten, A., Cannon, S., Spangler, R. & May, G. (2003) Genome-level evolution of resistance genes in Arabidopsis thaliana. *Genetics* **165**(1), 309-19.

Baur, M., Potrykus, I. & Paszkowski, J. (1990) Intermolecular homologous recombination in plants. *Mol Cell Biol* **10**(2), 492-500.

Botella, M.A., Parker, J.E., Frost, L.N., Bittner-Eddy, P.D., Beynon, J.L., Daniels, M.J., Holub, E.B. & Jones, J.D. (1998) Three genes of the Arabidopsis RPP1 complex resistance locus recognize distinct Peronospora parasitica avirulence determinants. *Plant Cell* **10**(11), 1847-60.

Buard, J., Shone, A.C. & Jeffreys, A.J. (2000) Meiotic recombination and flanking marker exchange at the highly unstable human minisatellite CEB1 (D2S90). *Am J Hum Genet* **67**(2), 333-44.

Chen, W. & Jinks-Robertson, S. (1999) The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. *Genetics* **151**(4), 1299-313.

Chin, D.B., Arroyo-Garcia, R., Ochoa, O.E., Kesseli, R.V., Lavelle, D.O. & Michelmore, R.W. (2001) Recombination and spontaneous mutation at the major cluster of resistance genes in lettuce (Lactuca sativa). *Genetics* **157**(2), 831-49.

Clough, S.J. & Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* **16**(6), 735-43.

Collins, F.S. & Weissman, S.M. (1984) The molecular genetics of human hemoglobin. *Prog Nucleic Acid Res Mol Biol* **31**, 315-462.

Collins, N., Drake, J., Ayliffe, M., Sun, Q., Ellis, J., Hulbert, S. & Pryor, T. (1999) Molecular characterization of the maize Rp1-D rust resistance haplotype and its mutants. *Plant Cell* **11**(7), 1365-76.

Cooley, M.B., Pathirana, S., Wu, H.J., Kachroo, P. & Klessig, D.F. (2000) Members of the Arabidopsis HRT/RPP8 family of resistance genes confer resistance to both viral and oomycete pathogens. *Plant Cell* **12**(5), 663-76.

Dangl, J.L. & Jones, J.D. (2001) Plant pathogens and integrated defence responses to infection. *Nature* **411**(6839), 826-33.

Davies, G., Kilby, N., Riou-Khamlichi, C. & Murray, J. (1999) Somatic and germinal inheritance of an FLP-mediated deletion in transgenic tobacco. *J. Exp. Bot.* **50**(338), 1447-1456.

De Buck, S., Jacobs, A., Van Montagu, M. & Depicker, A. (1999) The DNA sequences of T-DNA junctions suggest that complex T-DNA loci are formed by a recombination process resembling T-DNA integration. *Plant J* **20**(3), 295-304.

De Neve, M., De Buck, S., Jacobs, A., Van Montagu, M. & Depicker, A. (1997) T-DNA integration patterns in co-transformed plant cells suggest that T-DNA repeats originate from co-integration of separate T-DNAs. *Plant J* **11**(1), 15-29.

Delledonne, M., Xia, Y., Dixon, R.A. & Lamb, C. (1998) Nitric oxide functions as a signal in plant disease resistance. *Nature* **394**(6693), 585-8.

Dixon, M.S., Hatzixanthis, K., Jones, D.A., Harrison, K. & Jones, J.D. (1998) The tomato Cf-5 disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. *Plant Cell* **10**(11), 1915-25.

Dixon, M.S., Jones, D.A., Keddie, J.S., Thomas, C.M., Harrison, K. & Jones, J.D. (1996) The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell* **84**(3), 451-9.

Ellis, J.G., Lawrence, G.J., Luck, J.E. & Dodds, P.N. (1999) Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. *Plant Cell* **11**(3), 495-506.

Flor, H.H. (1942) Inheritance of pathogenicity in Melampsora lini. *Phytopathology* **32**, 653-669.

Gal, S., Pisan, B., Hohn, T., Grimsley, N. & Hohn, B. (1991) Genomic homologous recombination in planta. *Embo J* **10**(6), 1571-8.

Gumucio, D.L., Wiebauer, K., Caldwell, R.M., Samuelson, L.C. & Meisler, M.H. (1988) Concerted evolution of human amylase genes. *Mol Cell Biol* **8**(3), 1197-205.

Hammond-Kosack, K.E. & Jones, J.D. (1996) Resistance gene-dependent plant defense responses. *Plant Cell* **8**(10), 1773-91.

Hickey, D.A., Bally-Cuif, L., Abukashawa, S., Payant, V. & Benkel, B.F. (1991) Concerted evolution of duplicated protein-coding genes in Drosophila. *Proc Natl Acad Sci U S A* **88**(5), 1611-5.

Holliday, R. (1964) A mechanism for gene conversion in fungi. *Genet. Res. Camb.* **5**, 282-304.

Hrouda, M. & Paszkowski, J. (1994) High fidelity extrachromosomal recombination and gene targeting in plants. *Mol Gen Genet* **243**(1), 106-11.

Hughes, A.L. & Yeager, M. (1997) Molecular evolution of the vertebrate immune system. *Bioessays* **19**(9), 777-86.

Hulbert, S.H. (1997) Structure and evolution of the rp1 complex conferring rust resistance in maize. *Annu Rev Phytopathol* **35**, 293-310.

Jackson, J.A. & Fink, G.R. (1985) Meiotic recombination between duplicated genetic elements in Saccharomyces cerevisiae. *Genetics* **109**(2), 303-32.

Jelesko, J.G., Carter, K., Kinoshita, Y. & Gruissem, W. (2005) Frequency and character of alternative somatic recombination fates of paralogous genes during T-DNA integration. *Mol Genet Genomics* **274**(2), 91-102.

Jelesko, J.G., Carter, K., Thompson, W., Kinoshita, Y. & Gruissem, W. (2004) Meiotic recombination between paralogous *RBCSB* genes on sister chromatids of *Arabidopsis thaliana*. *Genetics* **166**(2), 947-57.

Jelesko, J.G., Harper, R., Furuya, M. & Gruissem, W. (1999) Rare germinal unequal crossing-over leading to recombinant gene formation and gene duplication in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **96**(18), 10302-7.

Jia, Y., McAdams, S.A., Bryan, G.T., Hershey, H.P. & Valent, B. (2000) Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *Embo J* **19**(15), 4004-14.

Jones, D.A. & Jones, J.D.G. (1996) The roles of leucine rich repeats in plant defences. *Adv. Bot. Res. Adv. Plant Pathol.* **24**, 90-167.

Jones, D.A., Thomas, C.M., Hammond-Kosack, K.E., Balint-Kurti, P.J. & Jones, J.D.G. (1994) Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* **266**, 789-793.

Jorgensen, R., Snyder, C and Jones, JDG (1987) T-DNA is organized predominantly in inverted repeat structures in plants transformed with Agrobacterium tumefaciens C58 derivatives *Molecular and General Genetics* **207**, 471.

Kilby, N.J., Davies, G.J. & Snaith, M.R. (1995) FLP recombinase in transgenic plants: constitutive activity in stably transformed tobacco and generation of marked cell clones in Arabidopsis. *Plant J* **8**(5), 637-52.

Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M. & Hohn, B. (2003) Pathogen-induced systemic plant signal triggers DNA rearrangements. *Nature* **423**(6941), 760-2.

Krizkova, L. & Hrouda, M. (1998) Direct repeats of T-DNA integrated in tobacco chromosome: characterization of junction regions. *Plant J* **16**(6), 673-80.

Kuang, H., Woo, S.S., Meyers, B.C., Nevo, E. & Michelmore, R.W. (2004) Multiple genetic processes result in heterogeneous rates of evolution within the major cluster disease resistance genes in lettuce. *Plant Cell* **16**(11), 2870-94.

Lebel, E.G., Masson, J., Bogucki, A. & Paszkowski, J. (1993) Stress-induced intrachromosomal recombination in plant somatic cells. *Proc Natl Acad Sci U S A* **90**(2), 422-6.

Lehrman, M.A., Russell, D.W., Goldstein, J.L. & Brown, M.S. (1987) Alu-Alu recombination deletes splice acceptor sites and produces secreted low density lipoprotein receptor in a subject with familial hypercholesterolemia. *J Biol Chem* **262**(7), 3354-61.

Leister, D., Kurth, J., Laurie, D.A., Yano, M., Sasaki, T., Devos, K., Graner, A. & Schulze-Lefert, P. (1998) Rapid reorganization of resistance gene homologues in cereal genomes. *Proc Natl Acad Sci U S A* **95**(1), 370-5.

Liu, Y.G., Shirano, Y., Fukaki, H., Yanai, Y., Tasaka, M., Tabata, S. & Shibata, D. (1999) Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning. *Proc Natl Acad Sci U S A* **96**(11), 6535-40.

Lopes, J., LeGuern, E., Gouider, R., Tardieu, S., Abbas, N., Birouk, N., Gugenheim, M., Bouche, P., Agid, Y. & Brice, A. (1996) Recombination hot spot in a 3.2-kb region of the Charcot-Marie-Tooth type 1A repeat sequences: new tools for molecular diagnosis of hereditary neuropathy with liability to pressure palsies and of Charcot-Marie-Tooth type 1A. French CMT Collaborative Research Group. *Am J Hum Genet* **58**(6), 1223-30.

Lopez Correa, C., Brems, H., Lazaro, C., Marynen, P. & Legius, E. (2000) Unequal meiotic crossover: a frequent cause of NF1 microdeletions. *Am J Hum Genet* **66**(6), 1969-74.

Luo, H., Lyznik, L.A., Gidoni, D. & Hodges, T.K. (2000) FLP-mediated recombination for use in hybrid plant production. *Plant J* **23**(3), 423-30.

Lyznik, L.A., Rao, K.V. & Hodges, T.K. (1996) FLP-mediated recombination of FRT sites in the maize genome. *Nucleic Acids Res* **24**(19), 3784-9.

Martin, M.P., Bashirova, A., Traherne, J., Trowsdale, J. & Carrington, M. (2003) Cutting edge: expansion of the KIR locus by unequal crossing over. *J Immunol* **171**(5), 2192-5.

McDowell, J.M., Dhandaydham, M., Long, T.A., Aarts, M.G., Goff, S., Holub, E.B. & Dangl, J.L. (1998) Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the RPP8 locus of Arabidopsis. *Plant Cell* **10**(11), 1861-74.

Meselson, M.S. & Radding, C.M. (1975) A general model for genetic recombination. *Proc Natl Acad Sci U S A* **72**(1), 358-61.

Metzenberg, A.B., Wurzer, G., Huisman, T.H. & Smithies, O. (1991) Homology requirements for unequal crossing over in humans. *Genetics* **128**(1), 143-61.

Meyers, B.C., Chin, D.B., Shen, K.A., Sivaramakrishnan, S., Lavelle, D.O., Zhang, Z. & Michelmore, R.W. (1998) The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. *Plant Cell* **10**(11), 1817-32.

Michelmore, R.W. & Meyers, B.C. (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res* **8**(11), 1113-30.

Mohr, T.J. (2005) Dissecting transcriptional regulation of RPP8 in arabidopsis thaliana. *Virginia Tech library database, Electronic Theses and Dissertations* **etd-06292005-140540**.

Molinier, J., Ries, G., Bonhoeffer, S. & Hohn, B. (2004) Interchromatid and interhomolog recombination in Arabidopsis thaliana. *Plant Cell* **16**(2), 342-52.

Murray, M.G. & Thompson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* **8**(19), 4321-5.

Nathans, J., Piantanida, T.P., Eddy, R.L., Shows, T.B. & Hogness, D.S. (1986a) Molecular genetics of inherited variation in human color vision. *Science* **232**(4747), 203-10.

Nathans, J., Thomas, D. & Hogness, D.S. (1986b) Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science* **232**(4747), 193-202.

Offringa, R., de Groot, M.J., Haagsman, H.J., Does, M.P., van den Elzen, P.J. & Hooykaas, P.J. (1990) Extrachromosomal homologous recombination and gene targeting in plant cells after Agrobacterium mediated transformation. *Embo J* **9**(10), 3077-84.

Ohta, T. (2000) Evolution of gene families. *Gene* **259**(1-2), 45-52.

Orlandi, E.W., Hutcheson, S.W. & Jacyn Baker, C. (1992) Early physiological responses associated with race-specific recognition in soybean lead tissue and cell suspension treated with *Pseudomonas. syringae pv. glycinea*. *Physiol. Mol. Plant Pathol.* **40**, 173-180.

Orr-Weaver, T.L., Szostak, J.W. & Rothstein, R.J. (1981) Yeast transformation: a model system for the study of recombination. *Proc Natl Acad Sci U S A* **78**(10), 6354-8.

Parker, J.E., Coleman, M.J., Szabo, V., Frost, L.N., Schmidt, R., van der Biezen, E., Moores, T., Dean, C., Daniels, M.J. & Jones, J.D.G. (1997) The Arabidopsis downy mildew resistance gene *Rpp5* shares similarity to the Toll and interleukin-1 receptors with *N* and *L6*. *Plant Cell* **9**, 879-894.

Parniske, M., Hammond-Kosack, K.E., Golstein, C., Thomas, C.M., Jones, D.A., Harrison, K., Wulff, B.B.H. & Jones, J.D.G. (1997) Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. *Cell* **91**, 821-832.

Peterhans, A., Schlupmann, H., Basse, C. & Paszkowski, J. (1990) Intrachromosomal recombination in plants. *Embo J* **9**(11), 3437-45.

Puchta, H. & Hohn, B. (1991) The mechanism of extrachromosomal homologous DNA recombination in plant cells. *Mol Gen Genet* **230**(1-2), 1-7.

Puchta, H., Kocher, S. & Hohn, B. (1992) Extrachromosomal homologous DNA recombination in plant cells is fast and is not affected by CpG methylation. *Mol Cell Biol* **12**(8), 3372-9.

Puchta, H., Swoboda, P., Gal, S., Blot, M. & Hohn, B. (1995) Somatic intrachromosomal homologous recombination events in populations of plant siblings. *Plant Mol Biol* **28**(2), 281-92.

Puchta, H., Swoboda, P. & Hohn, B. (1994) Homologous recombination in plants. *Cellular and Molecular Life Sciences (CMLS)* **50**(3), 277-284.

Ries, G., Buchholz, G., Frohnmeyer, H. & Hohn, B. (2000a) UV-damage-mediated induction of homologous recombination in Arabidopsis is dependent on photosynthetically active radiation. *Proc Natl Acad Sci U S A* **97**(24), 13425-9.

Ries, G., Heller, W., Puchta, H., Sandermann, H., Seidlitz, H.K. & Hohn, B. (2000b) Elevated UV-B radiation reduces genome stability in plants. *Nature* **406**(6791), 98-101.

Ritossa, F.M., Atwood, K.C. & Spiegelman, S. (1966) A molecular explanation of the bobbed mutants of Drosophila as partial deficiencies of "ribosomal" DNA. *Genetics* **54**(3), 819-34.

Robbins, T.P., Walker, E.L., Kermicle, J.L., Alleman, M. & Dellaporta, S.L. (1991) Meiotic instability of the R-r complex arising from displaced intragenic exchange and intrachromosomal rearrangement. *Genetics* **129**(1), 271-83.

Ruddle, F.H., Bartels, J.L., Bentley, K.L., Kappen, C., Murtha, M.T. & Pendleton, J.W. (1994) Evolution of Hox genes. *Annu Rev Genet* **28**, 423-42.

Schmidt-Puchta, W., Orel, N., Kyryk, A. & Puchta, H. (2004) Intrachromosomal homologous recombination in Arabidopsis thaliana. *Methods Mol Biol* **262**, 25-34.

Shaw, C.J., Bi, W. & Lupski, J.R. (2002) Genetic proof of unequal meiotic crossovers in reciprocal deletion and duplication of 17p11.2. *Am J Hum Genet* **71**(5), 1072-81.

Smith, G.P. (1974) Unequal crossover and the evolution of multigene families. *Cold Spring Harb Symp Quant Biol* **38**, 507-13.

Song, W.Y., Pi, L.Y., Wang, G.L., Gardner, J., Holsten, T. & Ronald, P.C. (1997) Evolution of the rice Xa21 disease resistance gene family. *Plant Cell* **9**(8), 1279-87.

Sturtevant, A.H. (1925) The effects of unequal crossing-over at the BAR locus in Drosophila. *Genetics* **10**, 117-147.

Swoboda, P., Gal, S., Hohn, B. & Puchta, H. (1994) Intrachromosomal homologous recombination in whole plants. *Embo J* **13**(2), 484-9.

Swoboda, P., Hohn, B. & Gal, S. (1993) Somatic homologous recombination in planta: the recombination frequency is dependent on the allelic state of recombining sequences and may be influenced by genomic position effects. *Mol Gen Genet* **237**(1-2), 33-40.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J. & Stahl, F.W. (1983) The double-strand-break repair model for recombination. *Cell* **33**(1), 25-35.

Takahashi, H., Miller, J., Nozaki, Y., Takeda, M., Shah, J., Hase, S., Ikegami, M., Ehara, Y. & Dinesh-Kumar, S.P. (2002) RCY1, an Arabidopsis thaliana RPP8/HRT family resistance gene, conferring resistance to cucumber mosaic virus requires salicylic acid, ethylene and a novel signal transduction mechanism. *Plant J* **32**(5), 655-67.

Tang, X., Frederick, R.D., Zhou, J., Halterman, D.A., Jia, Y. & Martin, G.B. (1996) Physical interaction of avrPto and the Pto kinase defines a recognition event involved in plant disease resistance. *Science* **274**, 2060-2063.

Tartof, K.D. (1974) Unequal mitotic sister chromatin exchange as the mechanism of ribosomal RNA gene magnification. *Proc Natl Acad Sci U S A* **71**(4), 1272-6.

Tartof, K.D. (1988) Unequal crossing over then and now. *Genetics* **120**(1), 1-6.

Tovar, J. & Lichtenstein, C. (1992) Somatic and Meiotic Chromosomal Recombination between Inverted Duplications in Transgenic Tobacco Plants. *Plant Cell* **4**(3), 319-332.

Tusie-Luna, M.T. & White, P.C. (1995) Gene conversions and unequal crossovers between CYP21 (steroid 21-hydroxylase gene) and CYP21P involve different mechanisms. *Proc Natl Acad Sci U S A* **92**(23), 10796-800.

Van der Biezen, E.A. & Jones, J.D.G. (1998) Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochem. Sci.* **12**, 454-456.

Xiao, S., Ellwood, S., Calis, O., Patrick, E., Li, T., Coleman, M. & Turner, J.G. (2001) Broad-spectrum mildew resistance in Arabidopsis thaliana mediated by RPW8. *Science* **291**(5501), 118-20.