A potato large-insert library for isolation of candidate loci for late blight resistance and studies on their genome organization

Inauguraldissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von
Rosa Angela Castillo Ruiz
aus Lima, Perú

Köln, 2002

Die vorliegende Arbeit wurde am Max-Planck-Institute für Züchtungsforschung in Köln durchgeführt

Berichterstatter : **PD Dr. habil. Christiane Gebhardt**
Max-Planck-Insitute für Züchtungsforschung
Prof.Dr.Martin Hülskamp
Institut für Botanik, Universität zu Köln

Abbreviations

BAC	Bacterial Artificial Chromosome
bp	base pair
BSA	bovine serum albumin
cDNA	complementary deoxyribonucleid acid
CAPS	cleaved amplified polymorphic sequence
cM	centiMorgans
dATP	deoxyadenosinetriphosphate
dCTP	deoxycytosinetriphosphate
dGTP	deoxyguanosinetriphosphate
dTTP	deoxythymidinetriphosphate
E. coli	*Escherichia coli*
EDTA	ethylenediamine tetraacetic acid
EST	expressed sequence tag
EtBr	ethidium bromide
IPTG	isopropylthio-ß-? -galactopyranoside
kb	kilo base
LG	linkage group
LOD	logarithm of odds
M	Molar
min	minute(s)
ORF	open reading frame
PAL	Phenylalanine ammonia-lyase
PCR	polymerase chain reaction
PFGE	Pulsed field gel electrophoresis
P.infestans	*Phytophthora infestans*
PPO	Polyphenol oxidase
PR-5	Pathogenesis related protein(thaumatin-like)
QTL	quantitative trait loci
RFLP	restriction fragment length polymorphism
s	second(s)
SCAR	sequence characterized amplified region
SDS	sodium dodecyl sulfate
SNP	single nucleotide polymorphism
SSCP	single strand conformational polymorphism
TMV	Tobacco mosaic virus
X-gal	5-bromo-4chloro-3-indolyl-ß-D-galactopyranoside

I Introduction

In trying to understand the genetic basis of field or partial resistance to *P. infestans* in potato, genetic analyses have correlated resistant phenotypes with DNA markers linked to the genes contributing to resistance (QTLs =quantitative trait loci). Those studies suggest that partial resistance appears to be controlled by a mixture of race-specific and race non-specific loci of varying effect. Besides, some QTLs co-localize with defense related genes (Leonards-Schippers et al. 1994, Trognitz et al. 2002). However, the identity of the genes behind the QTLs is still unknown. The DNA markers tightly linked to a QTL can be used in combination with large insert libraries in what is called the map-based cloning approach. Currently, this approach represents the most feasible route for cloning genes underlying traits known only by phenotype (O'Sullivan et al. 2001). In this study, a large insert library was constructed from a diploid potato hybrid. The hybrid was derived from the cross *S. phureja* x dihaploid *tuberosum*. This cross has been recently used in the mapping of genetic loci that contribute to *P. infestans* resistance (Ghislain et al. 2001). QTL effects where detected on LG VII and XII as a contribution from both parents, totaling up to 16% and 43% of the phenotypic variance, respectively. The strong QTL effect on LGXII was localized in a region where no major gene or QTL for *P. infestans* resistance has been reported. The construction of a library with this genetic material will facilitate the dissection of the QTL once a finer map of the region is available.

To demonstrate the advantages on having a large insert potato library, the candidate gene approach was used to clone the genetic counterparts of some cDNAs and RFLP markers located on potato chromosome XII. Besides, the clones were used to characterize gene families like *PAL* and *PR-5*. The specific objectives of the present study were: 1) develop a large insert library providing at least 3-4 haploid genome equivalents, 2) characterize the library for insert size and 3) screen the library with probes homologous to genes involved in defense responses. The importance of a large insert potato library and the rationale used in the screening will be outlined in the following sections.

1.1 The potato genome

The cultivated potato (*Solanum tuberosum* ssp. *tuberosum*) belongs to the family *Solanaceae*, genus *Solanum*, section *Tuberarium*, consisting of a number of species and species hybrids of which approximately 225 are tuber bearing. The basic chromosome number is x = 12 and the species range from diploids (2n = 2x = 24) to hexaploids (2n = 6x = 72) (Hawkes 1994). The nuclear content of the tetraploid potato has been determined in 3.31-3.86 pg/2C[(1)] and 1597-1862 Mbp. This makes potato to have a medium genome size, ten times larger than *Arabidopsis* (145 Mbp) and ten times smaller than wheat (15 966 Mbp) (Arumuganathan and Earle 1991).

Over the last 15 years several genetic maps of potato have been developed based on RFLP, AFLP, SSR and other PCR-based markers. (Bonierbale et al. 1988; Gebhardt et al. 1989, 1991; Tanksley et al. 1992; Jacobs et al. 1995; Perez et al. 1999 and Ghislain et al. 2001). The utility of genetic maps of potato has been demonstrated (reviewed by Gebhardt and Valkonen, 2001), many single genes and QTLs that confer resistance to major classes of plant pathogens (viruses, bacteria, nematodes, fungi and insects) have been mapped. The comparison of potato and tomato (*Lycopersicon*) revealed relatively few differences between both genomes, except for paracentric inversions on chromosomes 5, 9, 10, 11 and 12 (Bonierbale et al. 1988; Gebhardt et al. 1989, 1991; Tanksley et al. 1992). The comparison of tuber bearing and non-tuber bearing *Solanum* species revealed overall synteny. However, multiple translocations were identified on chromosomes 2 and 8 (Perez et al. 1999). The genomes of tomato, potato and pepper (*Capsicum*) have been also compared. RFLP markers showed homology between the three genomes (Livingstone et al. 1999). The synteny among species in the *Solanaceae* allow for direct comparative mapping of major traits, such as disease resistance. Several cross-generic clusters of resistance genes have been identified and genes conferring resistance to the same major pathogen group (e.g. virus, nematode, fungus, bacteria) were found in corresponding positions. Whether this reflects shared biology or components involved in plant-pathogen interactions remains to be seen (Grube et al. 2000).

[(1)] 1C is the DNA amount of unreplicated haploid genome of the species.

1.2 Large-insert DNA libraries

Large-insert DNA libraries are essential tools for many genomic studies: Physical mapping of genomic regions, specific chromosomes or the whole genome of a species; analysis of genome organization and evolution of complex genes or multi-gene families; and the establishment of the relationship between genetic and physical distances (Deng et al. 2001). Plant YAC libraries have been constructed for several organisms including: *Arabidopsis*, maize, tomato, barley and potato (Grill and Somerville 1991; Edwards et al. 1992; Martin et al. 1992; Kleine et al. 1993; Leister et al. 1997). These libraries have been used for a number of studies but their general use has been limited by the high frequency of chimeric and unstable clones. Today, the most commonly used system for constructing large insert libraries in plants is the bacterial artificial chromosome (BAC) system. In contrast to YAC, BAC vectors allow the cloning and stable-maintenance of large DNA fragments in *E. coli*. BAC inserts are present as a single copy/cell and the upper limit for cloning capacity in BACs is about 400 kb (Shizuya et al., 1992). BAC libraries became popular because they are relatively easy to develop (Figure 1.1), easy to handle and exhibit a low frequency of chimeric clones. BAC ends can be easily isolated and sequenced, and overlap relationships of clones easily and reliably established by fingerprinting, even at a whole-genome scale (Ding et al. 1999). Furthermore, BAC clones of interest may be readily subcloned and the entire sequence determined (Boysen et al. 1997). BAC clones are suitable for fluorescence *in situ* hybridization (FISH) analyses that enable direct gene localization on chromosomes (Jiang et al. 1995). Finally, BAC libraries are an important part of map base cloning efforts (O'Sullivan et al. 2001).

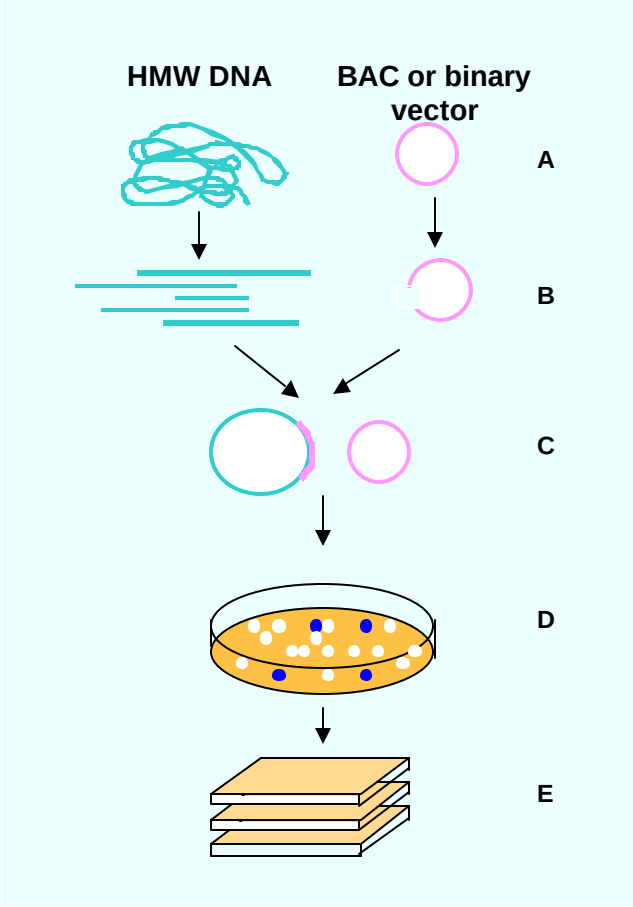


Figure 1.1 Construction of a large-insert DNA library. **A**. High molecular Weight (HMW) DNA is embedded in agarose and the vector is purified. **B**. Partial digestion of the HMW DNA and complete digestion of the vector. **C**. Ligation of the fragments with the vector and transformation into *E. coli* by electroporation. **D**. Plating in selective medium, white colonies: recombinants and blue colonies: non-recombinants. **E**. Storage of individual clones in 384 well plates.

BAC libraries have been constructed for two model plant systems: *Arabidopsis thaliana* (Wang et al. 1996) and *Medicago truncatula* (Nam et al. 1999) and for many crop plants including sorghum, rice, lettuce, soybean, apple, tomato, bean, sugarcane, barley, sunflower, citrus, maize, peach and papaya (Woo et al. 1994; Wang et al. 1995; Frijters et al. 1997; Danesh et al. 1998; Vinatzer et al. 1998; Hamilton et al. 1999; Vanhouten et al. 1999; Tomkins et al. 1999b; Yu et al. 2000; Gentzbittel et al. 2001; Deng et al. 2001; O'Sullivan et al. 2001; Wang et al. 2001 and Ming et al. 2001). Almost all the published libraries have been constructed using the pBeloBAC11 vector (Shizuya et al. 1992) or derivatives like pECSBAC1, pECSBAC4 and pBAC/SACB1 (Frijters et al. 1996; Bendahmane 1999). Hamilton et al. (1999) reported the construction of a tomato BIBAC library. BIBAC2 is a binary vector and is able to replicate in *E. coli* and *A. tumefasciens* making possible to introduce large insert DNA clones into plant cells (Hamilton et al. 1996). However,

there are only few reports of transformation of large DNA fragments (> 70 kb) into a plant genome (Hamilton et al. 1996; Liu et al. 1999).

Quanzhou and Zhang (1998) demonstrated that is possible to clone large fragments of eukaryotic DNA in bacteria using conventional plasmid-based vectors. They used two binary vectors, pCLD04541 and pSLJ1711 (Jones et al. 1992), and obtained highly stable clones ranging in insert sizes from 40 to 310 kb. In contrast to BACs, the clones were shown to be present at least 4-5 copies/cell. This in an advantage in the recovery of inserts using small volumes of culture. These conventional vectors have the characteristics of BAC vectors and besides they can transfer large-insert DNA fragments from *E. coli* to *Agrobacterium* sp. and then to plants. pCLD04541 has been successfully used to transfer DNA fragments of about 20 kb into plants (Bent et al. 1994; Brommonschenkel and Tanksley 1997). A soybean library was constructed with pCLD04541 (Meksem et al. 2000). In this study, pCLD04541 was used as cloning vector to construct the potato library. Since pCLD04541 is a conventional vector, the present library is a large-insert library and not a BAC or BIBAC library. However, for simplicity the clones of the library will be referred as BAC clones to indicate that they contain large inserts.

Plant BAC libraries have been used for a variety of applications: In *Arabidopsis*, are being used to develop sequence-ready physical maps (Bevan et al. 1998). In addition, BAC libraries have been used in map-based gene cloning of disease resistance genes (Song et al. 1995; Ballvora et al. 2002). Many of them have been constructed with the future purpose of cloning disease resistance genes in different crop species (Vinatzer et al. 1998; Deng et al. 2000; Allouis et al. 2001), to identify the sex determination gene in papaya (Ming et al. 2001) or genes involved in the domestication of sorghum (Lin et al. 1999). Moreover, BAC libraries are useful for examining genomic structure. The genomic structure of the a1-sh2 region of maize, sorghum and rice has been extensively examined through the sequencing of BAC clones containing these genes (Chen et al. 1997, 1998; Tikhonov et al. 1999). BAC libraries have also been used to develop physical maps for genomic regions containing resistance loci (Yang et al. 1997; Nakamura et al. 1997; Marek and Shoemaker, 1997; Yang et al. 1998; Tomkins et al. 1999b; Folkertsma et al. 1999).

More recently, a rice BAC library was used in chromosome landing[2] at the *Xa4* locus (Wang et al. 2001). BAC libraries are also valuable tools to characterize specific regions of genomes, for example a barley BAC library is being used in the physical mapping of gene-rich regions of barley (Yu et al. 2000).

1.3 *P. infestans* and the late blight disease

P. infestans (Mont.) de Bary caused the Irish potato famine of 1846 and still remains as a destructive pathogen responsible for multi-billion dollar losses in potato and tomato production (Kamoun et al. 2001). *P. infestans* is grouped in the class of Oomycetes, which behave like fungi in that their main body consists of a network of hyphae that grow at the tip, and in propagation via spores. Molecular phylogenetic analysis has demonstrated that Oomycetes evolved completely independent from the true fungi and are more closely related to golden-brown algae and *heterokont* algae in the eukaryotic crown group of *Stramenophiles*, kingdom *Protoctista* (Govers 2001). It was believed that *P. infestans* had its origin in Mexico. However, PCR amplification of the pathogen from historic specimens has shown that all the known modern mtDNA haplotypes of *P. infestans* do not occur in Mexico. This might suggest an alternative center of origin for the pathogen, perhaps in South America, which is the ancestral home of the potato (Ristaino et al. 2001).

P. infestans is a diploid heterothallic organism with two mating types (A1 and A2). Before 1980, A1 was the predominant mating type outside of Mexico (Goodwin et al. 1994). Over the following 10-15 years, the A2 mating type appeared throughout the world (Fry et al. 1993; Goodwin et al. 1994). Where both A1 and A2 mating types are present, there is the possibility of sexual recombination and increased genetic diversity among progeny. The oospore represents a possible new source of inoculum because of its capacity for long-term survival under harsh conditions (Drenth et al. 1993). Metalaxyl was the only curative fungicide[3] available to growers until resistance to it became widespread in *P. infestans*

[2] Chromosome landing has been defined as the selective enrichment for DNA markers within the sub-cM region around the target gene (Tanksley et al. 1995).

[3] *P. infestans* is not classified in the Mycota kingdom anymore. However, the terms fungicide or antifungal will be used to refer to the capability of killing and/or inhibiting true fungi and oomycetes like *Phytophthora* and *Pythium* species.

populations at about the same time A2 was detected outside of Mexico (Deahl et al. 1993; Goodwin et al. 1996).

1.4 Mapping of *P. infestans* resistance loci in potato

P. infestans exists as distinct races. To date, 11 race-specific late blight resistance genes (R genes) have been identified as originating from *S. demissum* (Malcolmson and Black 1966; Ewing et al. 2000). Five of these 11 R genes have been mapped to the potato molecular map, *R1* to chromosome 5 (Leonards-Schippers et al. 1992), R2 to chromosome 4 (Li et al. 1998), and *R3*, *R6* and *R7* to chromosome 11 (El-Kharbotly et al. 1994, 1996). Recently, Kulh et al. (2001) characterized and mapped *Rpi1*, a late blight resistance locus to chromosome 7. The *P. infestans* isolate used in the evaluations possessed the avirulence gene corresponding to *R9*, indicating that *Rpi1* could correspond to *R9*.

In addition to race-specific resistance genes, plants also possess other, race-non-specific resistance genes, which are partially effective against all races of the pathogen (minor gene resistance) (Lyon G. htt://scri.sari.ac.uk/TiPP/Celltxt.htr). QTL mapping of quantitative resistance to *P. infestans* has been pursued in potato. In a F1 progeny from a cross between non-inbred diploid parents, 11 QTLs on 9 potato chromosomes were identified as contributing to the resistance (Leonards-Shippers et al. 1994). The most significantly segregating locus detected in this study was located on chromosome 4. Two other regions correspond to segments on chromosomes 5 and 12. QTLs on chromosome 5 and 12 co-localize with *R1* and Potato Virus X (PVX) resistance genes *Rx1* and *Rx2*, respectively (Leonards-Shippers et al. 1994). Oberghagemann et al. (1999) found a reproducible QTL effect linked to RFLP marker GP179 on chromosome 5 using five hybrid populations grown in different locations. The QTL effect was related to foliage resistance. Meyer et al. (1998) have reported a QTL on chromosome 8 to be responsible for a significant expression of resistance to late blight in tetraploid potato[(4)]. QTL mapping experiments involving wild potato species as source of *P. infestans* resistance have also been reported (Ewing et al. 2000; Sandbrink et al. 2000 and Naess et al. 2000). For example, Naess et al. (2000) using the BC_1 progeny of *S.*

[(4)] This QTL has now been relocated to the short arm on chromosome 4 (De Jong personal communication reported in Naess *et al.* 2000)

bulbocastanum and *S. tuberosum* mapped a single dominant gene or a tightly linked cluster of genes on chromosome 8. The nature of resistance seems to depend on different loci in every wild potato material tested. Recently, Ghislain et al. (2001) reported QTLs associated with *P. infestans* resistance from a hybrid population between the native cultivated species *S. phureja* and dihaploid *tuberosum*. *S. phureja* is a diploid potato species that shows horizontal resistance to late blight and good agronomical characteristics. In this experiment, two major QTLs were detected on chromosomes VII and XII as a contribution from both parents, totaling up to 16% and 43% of the phenotypic variance, respectively. Such a strong QTL effect on LGXII has not been reported previously. Unfortunately, the QTL expands to a large portion of the chromosomes of both parental lines and a fine map of the region is not available (Figure 1.2). For this reason, the candidate gene approach was used for screening the potato library with markers mapping on LGXII.

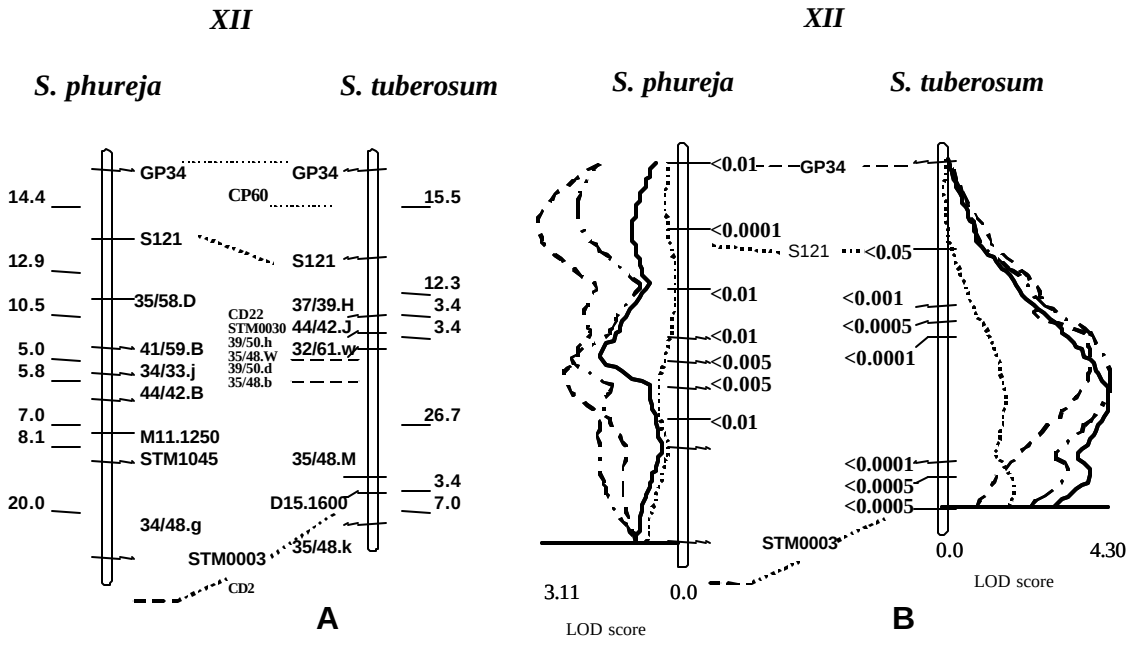


Figure 1.2 A. Genetic LGXII of *S. phureja* and dihaploid *S. tuberosum*. Genetic distances (cM) and markers are indicated. **B.** QTL location on LGXII of *S. phureja* and *S. tuberosum*. QTLs are represented as LOD-score plots. Bridging markers are repeated from figure A. (adapted from Ghislain et al. 2001).

1.5 Candidate gene approach

The map based cloning strategy requires the identification of markers tightly link to the trait of interest through high-resolution mapping, construction and

screening of large-insert libraries, and the search for the gene by using complementation tests. An alternative to the high-resolution mapping is the candidate gene approach (CG) (Wang et al. 2001). CG analysis is based on the hypothesis that known function genes (the candidate genes) could correspond to the loci controlling traits of interest. CGs refer either to cloned genes presumed to affect a given trait (functional CGs) or to genes suggested by their close proximity on linkage maps to loci controlling the trait (positional CGs) (Pflieger et al. 2001a). In this way, some resistance QTLs previously mapped may correspond to defense response (DR) genes (Pflieger et al. 1999).

The defense responses at biochemical level include the synthesis and deposition of phenolic compounds and proteins in the cell wall, rapid localized cell collapse and death (hypersensitive response, HR), accumulation of antimicrobial phytoalexins, and the synthesis of pathogenesis-related protein (PR proteins) comprising a wide range of different plant defense proteins (Kombrink and Sommsich, 1995) (Figure 1.3). Many of the genes involved in these biochemical pathways have been cloned and some have been mapped in QTL regions for resistance (Leonards Shippers et al. 1994; Bonierbale et al. 1994; Byrne et al. 1996; Faris et al. 1999; Pflieger et al. 2001b; Trognitz et al. 2002; Wang et al. 2001). Therefore, DR genes can be used in genetic analysis as genetic markers searching for correlation with phenotypic data (Faris et al., 1999). The final validation of a CG will be provided through physiological analyses, genetic transformation and/or sexual complementation (Pflieger et al. 2001a).

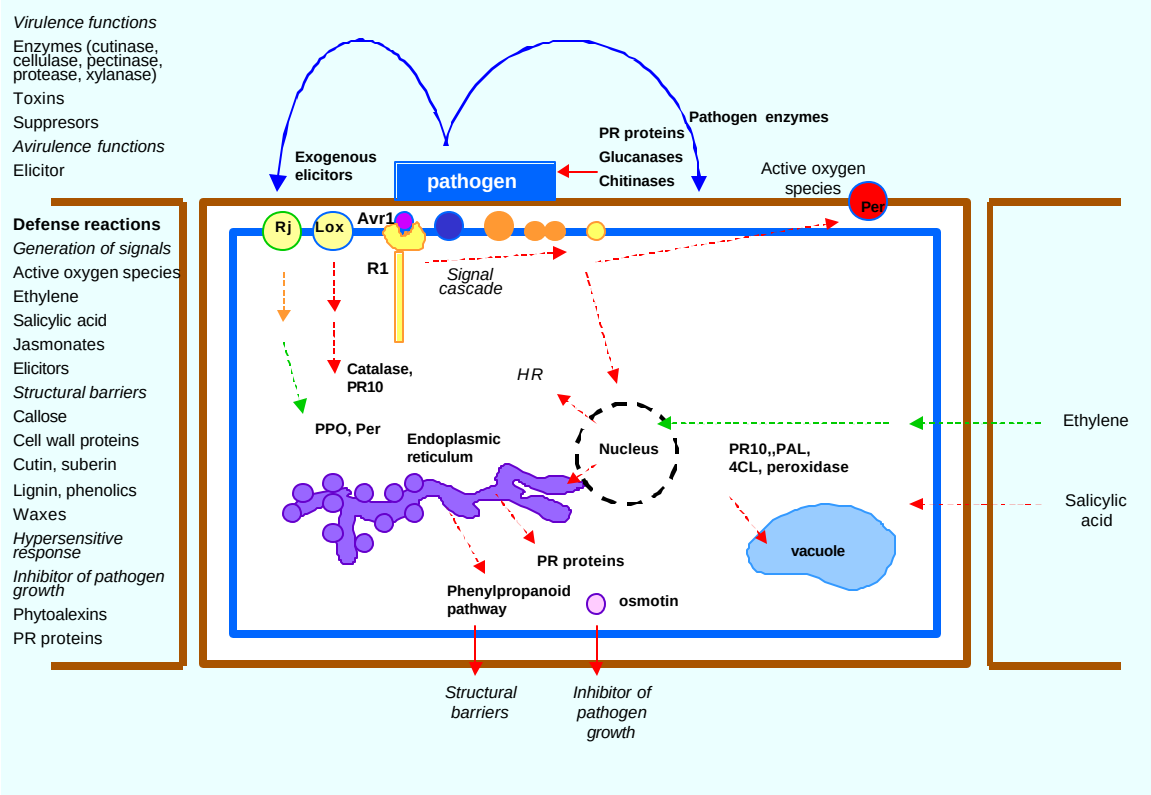


Figure 1.3 Schematic representation and summary of mechanisms involved in plant-pathogen interactions. R1: resistance gene, Avr1: avirulence gene 1, Rj: Jasmonate receptor, Lox: lipoxigenase, PPO: polyphenoloxidase, Per: peroxidase, HR: hypersensitive response, PAL: phenylalanine ammonia-lyase, 4CL: 4-coumarate:CoA ligase, PR: pathogenesis related proteins. (modified based on Lyon G. http://scri.sari.ac.uk/TiPP/Celltxt.htr).

From the gene markers mapping on potato LGXII (Table 1.1) *NtPR-5* and CP66 were chosen among others as probes to clone their genomic counterparts. *NtPR-5* belong to the pathogenesis related proteins (PR). On the other hand, CP66 shows homology to polyphenol oxidases.

Table 1.1 Candidate genes of potato LGXII.

Loci on BC916[2] map	Position on BC916[2] (cM)	Species	Function assignment based on sequence similarity
NtPR-5	19	*N. tabacum*	Pathogenesis related protein PR-5
Pot kin	34	*S. tuberosum*	Protein kinase
CP118		*S. tuberosum*	Shaggy-like kinase
CP66	56	*S. tuberosum*	Polyphenol oxidase

1.6 PR-5 protein family (Thaumatin-like protein)

The term pathogenesis-related (PR) proteins, was coined in 1980 to define a group of ubiquitous plant polypeptides that accumulate in pathological or related situations (Cutt and Klessing, 1992). PR proteins were described first on the basis of their accumulation in tobacco leaves infected with TMV (van Loon and van Kammen 1970). Currently, they are grouped in 11 families (Table 1.2) (van Loon et al., 1994). The accumulation of PR proteins is often not restricted to infected areas but also occurs, with some delay, in uninfected parts of the plant (van Loon 1985). This implies that PR proteins may play an important role in enhancing the levels of resistance to secondary challenges by pathogens, a phenomenon that is referred to as systemic acquired resistance or SAR (Somssich 1994).

Table 1.2 Recognized families of pathogenesis related proteins (adapted from van Loon, 1999).

Family	Type member	Properties
PR-1	Tobacco PR-1a	Unknown
PR-2	Tobacco PR-2	β 1,3-glucanase
PR-3	Tobacco P, Q	Chitinase type I, II, IV, V, VI, VII
PR-4	Tobacco R	Chitinase type I, II
PR-5	Tobacco S	Thaumatin-like
PR-6	Tomato inhibitor I	Proteinase-inhibitor
PR-7	Tomato P_{69}	Endoproteinase
PR-8	Cucumber chitinase	Chitinase type III
PR-9	Tomato "lignin forming peroxidase"	Peroxidase
PR-10	Parsley PR1	Ribonuclease-like
PR-11	Tobacco class V chitinase	Chitinase, type I

The PR-5 family includes a quite diverse group of proteins. The first PR-5 was reported by Cornelissen et al. (1986) as one of the TMV induced mRNAs of cv. Samsun NN tobacco. The mRNA encodes a protein with 65% identity to thaumatin, the sweet taste protein from *Thaumatoccus daniellii* (Iyengar et al. 1979). The tobacco protein was called PR-S. PR-S had an acidic pI and it was present in the intercellular fluids of TMV tobacco leaves (Pierpoint et al. 1987). Since Cornelissen (1986), many thaumatin- like genes have been described and cloned from different plant genera (van Kan et al. 1989; Pierpoint et al. 1990; Woloshuk et al. 1991; Zhu et al. 1995a and 1995b; Cheong et al. 1997; Hu et al. 1997; Barre et al. 2000; Darby et al. 2000; Newton et al. 2000; Skadsen et al. 2000; Shih et al. 2001). Their role in

plant defense was suggested by the finding that zeamatin, a maize seed thaumatin-like protein showed antifungal activity against *C. albicans, N. crassa* and *T. reesei* (Roberts et al. 1990). Woloshuk et al. (1991) isolated proteins with inhibitory effect towards *P. infestans* from TMV tobacco and tomato infected leaves. The proteins were called tobacco AP24 and tomato AP24, respectively. They resembled thaumatin, but especially osmotin, a protein present in osmotic stressed tobacco cells (Singh et al. 1985). For this reason, they are referred as osmotins or OLP (osmotin-like proteins). They differ from the intercellular PR-5 in that they are vacuolar and with a basic pI. Osmotin is active *in vitro* against a variety of fungi: *C. beticola, T. reesei, C. albicans* and *N. crassa* (Vigers et al. 1992). On the other hand, the purified thaumatin-like PR-S from TMV inoculated tobacco leaves had no antifungal activity against *P. infestans,* but it showed a strong inhibition of *C. beticola* (Vigers et al. 1992). The sweet thaumatin did not show antifungal activity (Vigers et al. 1992). Koiwa et al. (1994) reported a third subclass of PR-5 proteins, a tobacco neutral PR-5. The tobacco neutral PR-5 has been called PR-5d and it has shown antifungal activity against *N. crassa, T. reesei, C. miyabeanus, F. oxysporum* and *A. solani* (Koiwa et al. 1997).

A common feature of the PR-5 family is the conservation of cysteine residues. There are 16 cysteine residues highly conserved among most of the members of the PR-5 family. The cysteine residues form 8 disulfide bridges that are thought to be important for the functional structure of the protein. Vigers et al. (1992) postulated that the regions that have conserved amino acid sequences are catalytically important, while those regions of the sequence that differ comprise the recognition portion which confers specificity. The mechanism by which thaumatin-like proteins exert their antifungal activity is not clear. It was proposed that they insert into the lipid bilayer creating transmembrane pores and thus enhancing permeability (Vigers et al. 1992; Anzlovar et al. 1998). Thaumatin-like proteins with these properties are also known as permeatins. On the other hand, Batalia et al. (1996) suggested that PR-5 may interact with other components that facilitate its uptake, or it may modulate osmoregulators causing a rapid influx of water leading to hyphal rupture. The fact that the *Arabidopsis* receptor kinase PR-5K has a receptor domain homologous to PR-5 (Wang et al. 1996) indicates that PR-5 has the ability to interact

specifically with ligand molecules, perhaps those of a pathogen. Studies on the effects of osmotin on the growth of *Saccharomyces cerevisiae* indicated that there is a specific membrane receptor for osmotin (Yun et al. 1997). Koiwa et al. (1999) reported a common motif in antifungal thaumatin-like proteins. By doing a crystal structure study of tobacco PR-5d, zeamatin and the sweet thaumatin they found a negatively charged surface cleft present in antifungal thaumatin-like proteins but not in thaumatin. Glucanase activity has also been reported for some thaumatin-like proteins (Grenier et al. 1999) including: a barley leaf stress-related permatin, two tomato fruit osmotins (AP24 and NP24) a cherry fruit and two tobacco stigma proteins. In the same experiment no activity was found with thaumatin, zeamatin, tobacco PR-R (PR-S) among others.

1.7 Polyphenol oxidases

Polyphenol oxidases (PPOs) in plants are nuclear encoded enzymes. They are located in or on the internal membranes of plastids in both photosynthetic and non-photosynthetic tissues of angiosperms. PPOs catalyze the oxidation of mono and *o*-diphenols to *o*-diquinones (Hunt et al. 1993). A 59 kDa PPO is the predominant protein constituent of foliar glandular trichomes of the wild potato *S. berthaultii*. This enzyme appears to be responsible for oxidative polymerization of glandular trichome exudate that results in insect entrapment and resistance of *S. berthaultii* to insect pests (Kowalski et al. 1992). A QTL for trichome-mediated resistance was located on chromosome 8 of a hybrid population of *S. tuberosum* x *S. berthaultii* (Bonierbale et al. 1994) at the location to which structural genes of the *PPO* gene family have been mapped (Tanksley et al. 1992). The function of PPO within leaves or other plant organs is not well understood. PPO appears to be a small gene family in potato and each gene has a specific spatial and temporal pattern of expression (Thygesen et al. 1995).

II Materials and methods

2.1 Materials

2.1.1 Chemicals, enzymes, oligonucleotides and kits

Laboratory reagents were obtained from the following companies: Biozym (Epicenter USA), Invitrogen-Life Technologies (Freiburg), Sigma (Deisenhofen), Merck (Darsmstadt), Difco (USA), Roth (Karlsruhe), Serva (Heidelberg). Filter 3MM paper was obtained from Whatman (England). dNTPs were from Pharmacia. Radioisotope [α^{32}P]-dCTP (10µCi/µl) was purchased from Amersham Buchler (Braunschweig). The restriction enzymes were obtained from Roche (Mannheim) and Biolabs (England). Gelase was obtained from Biozym (Epicenter USA), Klenow and Shrimp phosphatase were purchased from Roche (Mannheim), and T4 DNA ligase was purchased from Invitrogen-Life Technologies (Freiburg). *Taq*-Polymerase was purchased from Invitrogen-Life Technologies (Freiburg). Oligonucleotides were custom synthesized by MWG-Biotech (München) and Metabion (Martinsried). Kits were obtained from Qiagen (Hilden) including Gel-extraction kit (Qiaex II and Qiaquick), plasmid Midi and Maxi Kit (QiaPrep), RNeasy Plant Mini kit. SMART™ Race cDNA Amplification Kit was from Clontech (CA, USA).

2.1.2 Plant material

Potato diploid line PD59 was used for the construction of the library. This line is one of the F_1 offspring of the PD population (Ghislain et al. 2001) showing late blight field resistance. *In vitro* plants of line PD59 were obtained from the International Potato Center (CIP, Lima-Peru). The plants were propagated by stem cuttings in MS medium. Rooted plantlets were planted in pots and transferred to the greenhouse under normal daylight conditions.

Two diploid mapping populations were used PD and F1840. The PD population was derived from a diploid inter-specific cross between *S. phureja* (CHS-625 line) and dihaploid *S. tuberosum* ssp. *tuberosum* (PS-3 line). This population consists of 246 individuals and segregates for field resistance to late blight. Both parents are susceptible to avirulent isolates of *P. infestans* (Race 0) and thus do not express any R gene conferring race specific resistance (Ghislain et al. 2001). F1840 is a population of 92 individuals from the cross of H82.337/49 x H80.696/4.

H82.337/49 was a diploid potato breeding line, while H80.696/4 was a hybrid between *S. tuberosum* ssp. *tuberosum* and *S. spegazzini* (Gebhardt et al. 1991; Leister et al. 1996).

DNA from the following species: *S. etuberosum, S. morelliforme, S. lignicaule, S. chacoense, S. megistracolobum, S. maglia, S. verrucosum, S. leptophyes, S. canasense, S. sparsipilum, S. bukasovii, S. kurtzianum, S. microdontum, S. hondelmanii, S. oplicense, S. neorossii, S. gourlayi, S. alandiae, S. vernei. S. berthaultii, S. spegazzinii, S. brevicaule, S. phureja, S. andigena, S. tuberosum* cv. Desirée, *S. stenotomum, S. acaule, S. stoloniferum, S. demissum, S. dulcamara, S. nigrum,* tomato (*Lycopersicon esculentum*), pepper (*Capsicum annuum*), wild tomato (*Lycopersicon* sp.), tobacco (*Nicotiana tabacum*) and eggplant (*Solanum melongena*).

2.1.3 Bacterial strain

The genotype of *E. coli* DH10B: F^- *mcr*A Δ(*mrr-hsd*RMS-mcrBC) ϕ80d*lac*ZΔM15 Δ*lac*X74 deoR *rec*A1 *end*A1 *ara*D139 Δ(*ara, leu*)7697 *gal*U *gal*K λ^- *rps*L *nup*G (DH10B *E. coli* ElectroMax competent cells Life Technologies).

2.1.4 *P. infestans* strains

P. infestans race 4 and complex race 1-11 were kindly provided by Dr. Gieffers (Max Planck Institute, Cologne).

2.1.5 Cloning vectors

pCDL04541 (Jones *et al*. 1992). pCLD04541 is a binary vector capable of replicating in both *E. coli* and *A. tumefaciens*. It carries the genes for tetracyclin and kanamycin resistance and β- galactosidase for blue and white selection (Figure. 2.1).

pBluescript II $SK^+/-$ (Stratagene, Heidelberg).

2.1.6 Standard primers

T3 primer	5' AATTAACCCTCACTAAAGGG 3'
T7 primer	5' GTAATACGACTCACTATAGGGC 3'

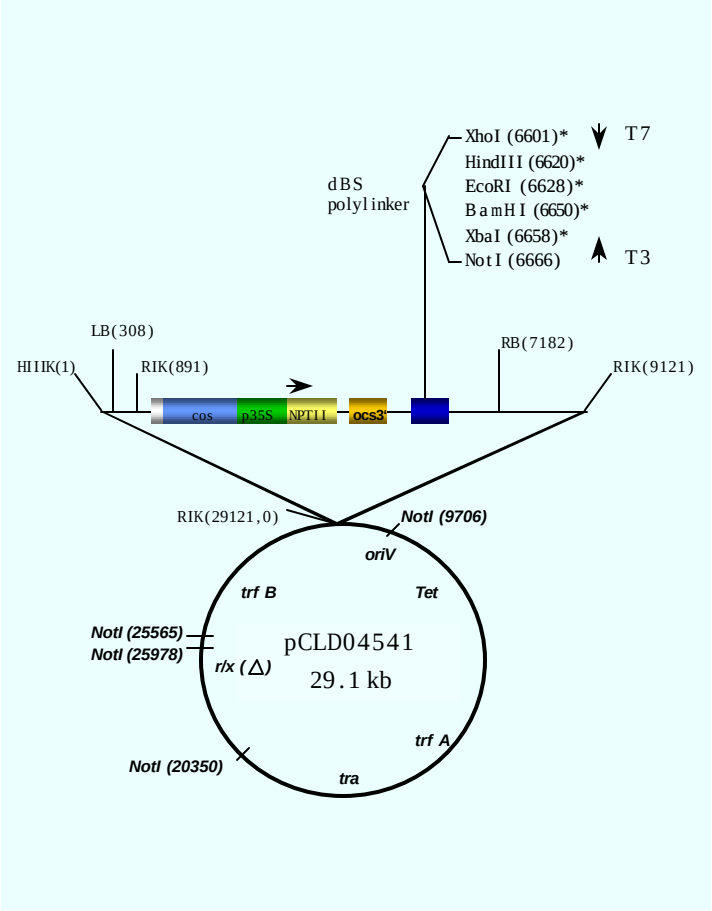


Figure 2.1 The structure of pCLD04541 (Jones *et al*. 1992). The T-DNA was cloned into the *Eco*RI site of pRK290 (20 kb) (Ditta et al. 1980), p35 signifies cauliflower mosaic virus (CaMV) 35S RNA gene promoter, pdBS for dark Bluescript polylinker, *oriV for* origin of replication, *trfA* and *trfB* genes for trans acting replication, r/x signifies relaxation complex site, and *tra* regions contain genes for conjugal transfer. HIIIK signifies a *Hin*dIII site filled with Klenow polymerase and dNTPs, RIK indicates an *Eco*RI site similarly treated, LB signifies T-DNA left border repeat sequence and RB signifies the right border. *ocs 3'* indicates octopine synthase 3' end. NPT indicates neomycin phosphotransferase. Tetracyclin resistance gene is indicated by *tet*. Gene and promoter orientations are indicated by arrows. * -are the unique sites for cloning.

2.1.7 Probes used for screening of the BAC library

Probe	Reference
S. tuberosum PAL cDNA	Fritzemeier et al., 1987
N. tabacum PR-5 cDNA	Payne et al. 1988
PCR product of osmotin sequences: Forward: 5'- gag gta cgc aac acc tgt cca tac -3' Reverse: 5'-agg gga aat ttg ggc tag taa cac -3	Accession numbers: X65701, S40046, X61679, AF093743, X72927, X72928, X67244 and M21346
RFLP marker CP66	Gebhardt et al. 1991
PCR product of *PPO* sequences: Forward: 5'-ccc tct caa ctt ttc ctc cac gga-3' Reverse: 5'-cta gta taa ctt ccc gca aac tct-3'	Accession numbers: M95197, M95196, STU22921, STU22922

2.1.8 Media, buffers and solutions commonly used

LB plates (1 liter) 10 g Tryptone
5 g Yeast extract
10 g NaCl
20 g Agar

LB medium (1 liter) 10 g Tryptone
5 g Yeast extract
10 g NaCl

T1E10 pH 8.0 (1 liter) 10mM Tris-HCl, pH 8.0
1mM EDTA, pH 8.0

50 x TAE 2M Tris-HCl
50mM EDTA
57.1ml glacial acetic acid

10X TBE (1 liter) 108 g Tris base
55 g Boric acid
40 ml EDTA 0.5M, pH 8.0

Depurination solution 250mM HCl

Denaturation solution 1.5M NaCl
0.5M NaOH

Neutralizing solution 1.5M NaCl
0.5M Tris HCl, pH 7.5

20 x SSC 0.3M Na_3 citrate, 3M NaCl, pH 7.0

2.2 Methods

2.2.1 Isolation of genomic DNA for PCR

10 mg freeze-dried leaf material was grinded under liquid nitrogen to fine powder. Total genomic DNA was extracted using DNeasy™ Plant Mini Kit (Qiagen) following the supplier's instructions. Extracted DNA was quantified as described in Sect. 2.2.4.

2.2.2 Isolation of genomic DNA for Southern analysis

2 g of fresh leaf material was crushed in liquid nitrogen with mortar and pestle and transferred to a 50 ml falcon tube. 20 ml of CTAB extraction buffer (100mM Tris pH 9.5, 1.4M NaCl, 20mM EDTA, 2% CTAB, 1% PEG 6000) plus 50 µl β-Mercaptoethanol were added and incubated at 74 °C for 20 min. 1 volume of chloroform / isoamylalcohol (24:1) was added and the mixture was centrifuged for 10 min at 5000 x g at 4°C. The supernatant was transferred to a new tube. RNase (Qiagen) (400 µg/ml) was added and incubated for 30 min at 37 °C. The same volume of water was added and the pH was adjusted to 7.0 with HCl. The solution was applied to an equilibrated Qiagen – tip 100 and the protocol for Midi Prep (Qiagen) was followed. Extracted DNA was quantified as described in Sect. 2.2.4.

2.2.3 Isolation of plasmid DNA

25 ml overnight-cultured bacteria were harvested by centrifugation and plasmid DNA was extracted using QIAprep® Midiprep Kit following the supplier's protocol. The concentration of plasmid DNA was mesured as in Sect. 2.2.4.

2.2.4 Concentration measurement of DNA

The quality and quantity of DNA were measured by comparison of band-intensity on ethidium bromide stained agarose gels with a DNA molecular weight standard, or alternatively, by the absorbance at ? = 260 nm.

2.2.5 Southern gel blot analysis

8 µg genomic DNA were digested for 6 hours with 40U of restriction enzyme in the corresponding buffer at 37 °C. Digested DNA was mixed with 1/10 th volume of 3M Na-acetate and 3 volumes 70% ethanol, precipitated and collected by

centrifugation at 13 000 rpm for 30 min. The pellet was washed twice with 70% ethanol and air-dried, followed by adding 30 µl Tris 10mM. DNA fragments were separated on 0.9% agarose gel. The electrophoresis was run in 1 x TAE as running buffer for 20 hours at 2V/cm. The gel was processed for 10 min in depurination solution, 30 min in denaturation solution and 30 min in neutralizing solution. Transfer of DNA fragments onto Nylon-membrane (N^+ Hybond, Amersham) was performed with 20 x SSC buffer using a conventional capillarity system (Sambrook et al. 1989). DNA was covalently cross-linked to the membrane by Stratalinker (Stratagene) using 1200 µl joules x 100 followed by drying at 80 °C for at least 1 hour.

2.2.6 Radioactive labeling and purification of DNA probes

Sephadex buffer: 10mM Tris-HCl, pH 7.5; 1mM EDTA; 0.3mM NaCl; 0.1 % SDS.

Sephadex- G50 suspension: 100 ml Sephadex buffer, 2 g Sephadex-G50.

50-100 ng DNA were denaturated in 12µl at 95 °C for 5 min and chilled immediately on ice. The random primer labeling method (Feinberg and Vogelstein 1984) was used to label the DNA. 4µl of 5 x OLB buffer (0.1 mM dATP, dCTP, dGTP, dTTP; 1 M HEPES, pH 6.6; 1 mg/ml Hexadesoxyribonucleotide pdN6 (Pharmacia); 0.1 M Tris-HCl, pH 8.0; 3.7 x 10^{-5} M MgCl2; 0.35% (v/v) β-mercaptoethanol) were added to the denaturated DNA, 3 µl of [α^{32} P] - dCTP (10 µCi/µl) and 2 units of Klenow enzyme (Roche) were added in a total reaction volume of 20 µl. The labeling reaction was left at room temperature overnight or at 37 °C for 1h.

The separation of labeled DNA-fragments from unincorporated nucleotides was done using a 1 ml-Sephadex G-50 column. The 1 ml syringe was filled with sterile glass wool at the bottom and G-50 Sephadex equilibrated in buffer (G-50 Sephadex suspension). The labeling reaction was mixed with 300 µl Sephadex suspension and eluted through the column. A second elution with the same amount of Sephadex suspension was performed and the second elution was collected. Labelling efficiency was checked using the radioactivity counter. The DNA probe was mixed with 20 µl (15 mg/ml) of herring sperm and denaturated at 95 °C for 5 min. After denaturation the probe was immediately added to the filters.

2.2.7 Hybridization

Hybridization buffer: 7% SDS, 0.12M NaHPO, 1mM EDTA, pH 7.0

20 x SSPE: 3M NaCl, 0.2M NaHPO, 20mM EDTA, pH 7.0

2 x SSPE washing buffer: 2X SSPE, 0.1% SDS

Southern Blot Washing Buffer: 5mM NaHPO, 1mM EDTA, 0.2% SDS, pH7.0

Hybridizations were performed in a hybridization oven (Bachofer, Reutlingen, Germany) in glass tubes (30 cm x 3.5 cm in diameter) at 65 °C. The membranes were rolled up around a plastic stick, placed inside the tube and unrolled around the inner wall of the tube. 20 ml hybridization buffer plus 200µl (10mg/ml) denaturated herring sperm DNA were added to the tube for pre-hybridization. For pre-hybridization the tubes were incubated at 65 °C for at least 4 hours. After adding the denaturated probe, the hybridization was performed overnight under the same conditions as pre-hybridization.

After hybridization the filters were washed in the hybridization oven at 65 °C in the same tube with 30 ml of pre-heated 2 x SSPE washing solution, three times 15 min each. Then the filters were taken out from the tube and washed two times in a tank containing Southern Blot Washing buffer at room temperature for at least 1 hour. The filters were wrapped in thin plastic foil (Saran film) and placed in cassettes with intensifying Trimax-screen. Kodak X-OMAT AR5 film was placed on the filter and exposed at – 70 °C for 10 days before developing.

2.2.8 Standard PCR

All PCR (Polymerase Chain Reaction) reactions were performed either in a heating-lid thermo-cycler (Biometra, Göttingen) or in a heating-lid thermo-cycler Tetrad (MJ Research, USA) depending on the number of reactions.

PCR reaction (25 µl):	
10X buffer	2.5µl
dNTPs, 2.5mM	2µl
MgCl, 50mM	1.25µl
Primers, 10pM	
Forward	0.75µl
Reverse	0.75µl
DMSO	0.5µl
H_2O	17.35 µl
*Taq*Polymerase, 5U/µl	0.4 µl
Genomic DNA, 25 ng/µl	1µl

Amplification conditions:

96 °C denature 6 min

94 °C denature 20 s

50-62 °C annealing 30 s

72 °C extension 1-1 min 40 s

39 cycles

72 °C extension 6 min

10 °C hold

PCR fragments used for labeling as probes, were gel extracted using QIAquick Gel Extraction Kit (Qiagen). The extraction procedures recommended by the supplier were followed. The concentration of extracted DNA was estimated as in 2.2.4.

2.2.9 Development of genetic markers

2.2.9.1 Design of primers for genetic markers

Primers were selected to have good characteristics for amplification with optimal annealing temperature of 55 – 60 °C and primer length between 22 and 24 bases.

2.2.9.2 Identification of sequence characterize amplified region (SCAR)

PCR amplification from parental lines was analyzed by electrophoresis on 1.5% -2% agarose gel in 1 x TAE. The DNA fragments were visualized by EtBr staining.

2.2.9.3 Identification of cleaved amplified polymorphic sequences (CAPS)

Monomorphic PCR products were digested with restriction endonucleases. The restriction map of the PCR products was obtained either with WebCutter (http://www.firstmarket.com/cutter/cut2.html) or Jellyfish Version 1.3 (Bioware) and enzymes producing fragments larger than 100bp were used. These restriction

enzymes were mainly 4-bp cutter including *Alu*I, *Ava*II, *Rsa*I, *Taq*I, *Hae*III, *Mse*I, and *Mbo*I. Five units of restriction enzyme, mixed with the reaction buffer, were added to 12 µl of the PCR reaction to final volume of 20 µl. The digests were incubated for 4 hours. CAPS polymorphisms were analysed by electrophoresis on 1.5- 3% agarose gel in 1 x TAE. Gene Ruler DNA ladder Mix (Fermentas) was used as a size marker. DNA fragments were visualized by ethidium bromide staining.

2.2.9.4 Identification of single strand conformational polymorphism (SSCP)

When CAPS markers did not show polymorphism between parental lines, the PCR products were digested with frequent 4-bp cutter like *Alu*I, *Mse*I and *TaqI*, run on 0.5% MDE gels following the supplier's protocol (BMA) and silver stained as described Sanguinetti et al. (1994).

2.2.10 Generation of segregation data and linkage analysis

Fragments were scored as 0, 1 and 2 for absence, presence and missing value, respectively. The segregation data was then added to the existing RFLP framework data sets of the F1840 population using Dr. C. Gebhardt's MAPRF software.

2.2.11 *P. infestans* infection tests

Rye Agar Medium: 200 g rye. Fill up to 1 liter with Millipore water and autoclave. Filter the liquid, and add 39 g/liter of potato dextrose agar. Fill up to 1liter and autoclave again.

P. infestans race 4 and a complex race 1-11 were used to infect potato line PD59. The mycelia were propagated on rye agar medium and grown for at least one week. Before infection, mycelia were grown on susceptible Desirée plants to induce virulence. Zoospores were induced by adding sterile deionized water to the mycelium and incubating at 4 °C for two hours. The zoospores were counted under the microscope and dilutions 5 x 10^4 zoospores/ml were used. 500-1000 zoospores were inoculated per leaf. Plants were grown in the greenhouse for at least 4 weeks and young leaflets (approx. 1 cm diameter) were detached and put on trays over wet filter paper. The inoculated leaves were kept in a chamber with 12 hours of light at 17 °C followed by 12 hours of darkness at 10 °C.

2.2.12 Isolation of total RNA

Infected or uninfected leaf material was harvested in liquid nitrogen and kept at – 70 °C until use. The frozen leaves were crushed with tungsten balls. RNA was isolated using the RNeasy Plant Mini kit (Qiagen) following the supplier protocol. RNA concentration was measured with a Smart spectrophotometer. Quality was assayed by running a 1.2% formaldehyde gel and visualized with ethidium bromide staining.

2.2.13 Northern gel blot analysis

10 x FA Gel buffer: 0.2M 3-[N-Morpholino]propanesulfonic acid] (MOPS free acid), 50mM Sodium acetate, 10mM EDTA, pH to 7.0 with NaOH.

1.2% FA Gel (100ml): 1.2 g agarose, 10 ml 10X FA gel buffer, RNase-free water to 100ml. Microwave to melt agarose. Cool to 65ºC in a waterbath. Add 1.8ml of 37% (12.3M) formaldehyde and 1µl of ethidium bromide.

5 x RNA loading buffer: 16 µl saturated bromophenol blue, 80 µl 0.5M EDTA, pH 8.0, 720 µl 37% formaldehyde, 2 ml 100% glycerol, 3084 µl formamide, 4 ml 10X FA.

20 µg of total RNA were loaded on a 1.2% formaldehyde gel. The gel was blotted as in section 2.2.5. The hybridization and washing conditions were the same as in the section 2.2.7. The filters were wrapped in thin plastic foil (Saran film) and exposed on a Phosphoscreen for 4 days. Fragment patterns were visualized using a Storm scanner 860 (Molecular Dynamics) and the IQ software.

2.2.14 Synthesis of 5' cDNA

The RNA was isolated as in section 2.2.12 with the exception that RNAse free DNAse was included in the procedure. The synthesis of cDNA was done using the SMART Race cDNA Amplification Kit (Clontech, CA, USA) following the supplier's protocol.

2.2.15 RT-PCR

For RT-PCR experiments 25 ng of synthesized cDNA was used for every reaction with standard PCR conditions (section 2.2.8).

2.2.16 Methods used for the constructing and screening the BAC library (BB library)

2.2.16.1 High Molecular Weight (HMW) DNA isolation

High molecular weight (HMW) DNA was prepared from nuclei or protoplasts of the line PD59. The plants were put in darkness for 3 days before the tissue was harvested to decrease the starch content in the leaves.

Preparation of HMW DNA from Plant Nuclei The protocols were taken from http://www.genome.clemson.edu/protocols written by Choi and Wing, and http://hbz.tamu.edu written by Zhang H.

10 x Homogenization buffer (HB) stock (1 liter): 100mM Tris base, 800mM KCl, 100mM EDTA, 10mM Spermidine-HCl, 10mM Spermine-HCl, pH 9.4-9.5 adjusted with NaOH.

1 x Homogenization buffer (HB) (1 liter): 100 ml 10X HB stock, 0.5M Sucrose.

20% Triton-X 100 (100 ml): 10 ml 10X HB stock, 0.5M Sucrose, 20 ml Triton X-100

Wash buffer (1 liter): 100 ml 10X HB stock, 0.5M Sucrose, 25 ml 20% Triton-X100. Before use, add β-mercaptoethanol to 0.15%.

About 15 g frozen leaves were grinded into a fine powder under liquid nitrogen with a mortar and pestle and transferred into an ice-cold 500 ml beaker containing 200 ml ice-cold wash buffer. The contents were stirred for 10 minutes on ice and filtered through 2 layers of Cheesecloth (Fisher) into an ice-cold 250 ml centrifuge bottle. The filtrate was centrifuged with a fixed-angle rotor in a Beckman centrifuge at 1,800 g for 20 minutes at 4°C. The supernatant was discarded and the pellet was resuspended in 30 ml of icecold Wash buffer. To remove the remaining debris in the suspension, the resuspended nuclei were filtered into a 50 ml centrifuge tube through two layers of Miracloth (Calbiochem) by gravity. The filtrate was centrifuged at 57 g (500 rpm) for 2 minutes at 4°C to remove intact cells and tissue residues. The supernatant was transferred into a fresh centrifuge tube and centrifuged at 1,800 g, for 15 minutes at 4°C in a swinging bucket rotor (Heraeus). The pellet was washed two more times by resuspension in wash buffer followed by centrifugation at 1,800 g, for 15 minutes at 4°C. After the final wash, the nuclei was resuspended in 1 ml of 1 x HB.

Preparation of HMW DNA from protoplasts

MS K3: 4.5 g MS salt, 0.4M Sucrose, 100 mg Inositol, 250 mg Xylose, 1 ml MS vitamins, pH 5.6. Sterilize by filtration.

Wash buffer 5 (W5): 154mM NaCl, 125mM $CaCl_2$, 5mM KCl, 5mM Glucose, pH 5.6-6.0. Keep at 4°C.

Protoplasts enzyme solution: 100ml MS K3 Stock, 0.1 g Cellulase R10, 0.5 g Macerozym R10, pH 5,6

MaMg: 450mM Mannitol, 15mM $MgCl_2$, 1g MES, pH 5.5- 6.7. Keep at 4°C.

Potato leaves were cut into thin slices with a sterile razor blade and put in 14cm petri dishes with enough Protoplast Enzyme Solution to cover the leaves. The petri dishes were left in a dark room at 28°C overnight. On the next day, the leaves were removed and the liquid containing protoplasts was filtered using a sieve. The protoplasts were pooled and transferred in a 50 ml falcon tube. The solution was centrifuged at 800 rpm for 5 min at 4°C to separate the cell debris from the floating protoplasts. After that the supernatant was transferred into a new falcon tube and filled with W5 buffer. The protoplasts were centrifuged at 500 rpm for 5 min at 4°C and checked under the microscope. The cells were resuspended with MaMg at 2 x 10^7 protoplasts/ml concentration.

2.2.16.2 Embedding of the nuclei or protoplasts in agarose plugs

Na- lauryl sarcosine (2%) (1 liter): 20 g Na- lauryl sarcosine

Lysis buffer (1 liter): 500 ml 0.5M EDTA pH 9-9.3, 500 ml Na-lauryl sarcosine 2%. Before use, add Proteinase K to 0.1 mg/ml.

PMSF (phenylmethylsulfonyl fluoride): 100mM PMSF dissolve in isopropanol

1% low-melting point agarose (Biozyme Epicenter) was molten in 1 x HB (for nuclei preparation) or MaMg buffer (for protoplasts preparation) and kept in a 45°C water bath. The nuclei or protoplasts were pre-warmed at 45°C for 5 minutes and then mixed with an equal volume of the agarose solution. The mixture was aliquot into ice-cold plug molds (75 µl) on ice using a pipette with cut off tip. When the agarose was completely solidified, the plugs were transferred into 10 volumes of lysis buffer. The plugs were incubated in the lysis buffer at 50°C with gentle shaking for 48 hours changing the buffer once after 24 hours. The plugs were

washed in 0.5M EDTA pH 9.0-9.3 at 50°C for one hour, once in 0.05M EDTA pH 8.0 on ice for one hour and stored in 0.05 M EDTA pH8.0 at 4°C.

2.2.16.3 Partial digestion of HMW DNA and fragment purification

Digestion of DNA agarose plugs

Before digestion, the agarose plugs were washed three times in 10-20 volumes of ice cold TE plus 0.1M PMSF and three times in 20 volumes of ice cold TE on ice, one hour each wash. The washed plugs were stored at 4°C. For digestion of individual agarose plugs, the following solutions were mixed in a sterile micro centrifuge tube:

10 x restriction enzyme buffer:..... 15 µl

100 x BSA (10mg/ml):..... 1.5 µl

40 mM Spermidine: ..... 6 µl

H_2O:..... up to 150 µl

The plugs were cut in 8 pieces and incubated with the buffer mix on ice for one hour. The enzyme was added and incubated on ice for 10 min. After that, the samples were transferred to 37°C and incubated for 20 min. Optimal partial digestion conditions were set up for every batch of nuclei or protoplasts in order to get most of the fragments in a size range of 100-500 kb. The digestion reactions were stopped by adding $1/10^{th}$ volume of 0.5M EDTA pH 8.0 and kept on ice.

Pulsed Field Gel Electrophoresis (PFGE)

The digested HMW DNA was loaded into a 1% low melting point agarose gel in 0.5 x TBE. Mid range and Lambda ladder PFGE markers were also included (Biolabs). The wells were sealed with molten agarose. Pulsed field gel electrophoresis was performed in a CHEF- DR III (BioRad) using the following conditions: Initial time: 60s, final time: 90s, time: 19 hours, voltage: 6 V/cm, angle: 120° and temperature: 14°C. After the run, the gel was stained with ethidium bromide and visualized under UV.

Size selection and purification

Two methods were used to recover DNA from agarose slices containing DNA in the range of 300-500 kb: electro elution and enzymatic digestion.

The electro-elution of DNA fragments from gel slices consists on sitting the gel pieces containing the DNA in dialysis bags with a small amount of

electrophoresis buffer (100 µl of 0.5 x TBE) and exposed it to an electric field (125 V for 22 hours at 4°C) in order to produce the movement of DNA molecules. The negative DNA molecules will be attracted to the positive electrode and in this way they will run out from the gel. When they are in solution the DNA can be concentrated and visualized in an agarose gel. Alternatively, the size-selected fragments were cut from the gel, equilibrated with 1 x Gelase buffer plus 60 mM NaCl, melted and digested with Gelase (Biozym, Epicenter). 1U of Gelase was used to digest 100 mg of gel at 45°C for 1 hour.

2.2.16.4 BAC vector preparation

Cell growth and plasmid isolation

The *E. coli* strain harboring pCLD04541 was streaked onto an LB agar plate and grown at 37°C overnight. 5 ml LB media were inoculated with a single colony from the overnight culture and grown for 8 hours at 37°C. 2.5 liter culture flasks containing 500 ml of LB medium and 15 mg/l tetracyclin were inoculated with 0.5 ml of the 8 hours culture. In total, 4 flasks were inoculated and the cultures were grown with shaking (225 rpm) at 37°C overnight to an OD of 1-1.2. The cells were harvest by centrifugation at 4°C for 15 minutes at 6000 g in a Beckman centrifuge (Beckman, USA). The vector was extracted with Plasmid Maxi Kit (Qiagen) following the supplier's protocol.

$CsCl_2$ gradient purification

The pellets obtained with the Maxi prep were pooled and resuspended in 7.5 ml 0.1M Tris pH 8.0 to which 7.9 g of $CsCl_2$ were added. The solution was transferred to a 15 ml ultracentrifuge tube, 375 µl of ethidium bromide (10 mg/ml) were added and the tube was filled to the top with $CsCl_2$/ 0.1M Tris 1:1 (w/w) and paraffine at the top before sealing. A balance tube was filled with $CsCl_2$/ 0.1M Tris 1:1(w/w). The sample was run in a Kontron Ultracentrifuge, Rotor TV 850 at 43 000 rpm at 20°C for 16 hours. After the run the closed circular plasmid band was collected (Sambrook et al. 1989) and transferred to a 5 ml ultracentrifuge tube. The sample was run a second time in a Rotor TV 865 at 45 000 rpm at 20°C for 16 hours. After the second run the plasmid was collected as before and washed several times with isoamyl alcohol/20 x SSC 1:1 to remove the ethidium bromide. The plasmid was precipitated with 1/10th volumes of 3M Na-acetate pH 5.6 and 0.7 volumes of

isopropanol. The precipitated plasmid was washed twice with 90% and 70% ethanol, dried and resuspended in 10mM Tris-HCl, pH 8.0.

Digestion and dephosphorylation of the plasmid

The vector was completely digested with *Hin*dIII, purified with phenol/chloroform 1:1 and precipitated with 1/10 volumes of 3M Na Acetate pH 5.6 and 3 volumes of 70% ethanol. The final pellet was resuspended in TE. The plasmid concentration was measured as in section 2.2.4. Prior to ligation, the plasmid was dephosphorylated with Shrimp phosphatase enzyme (Roche). 100-500 ng of plasmid were incubated with 1U of enzyme at 37°C for 1 hour. The enzyme was heat inactivated at 65°C for 30 min.

2.2.16.5 Ligation and transformation

Size selected DNA was ligated with dephosphorylated vector in a molar ratio of 1 to 3 (size-selected DNA to vector DNA). 100 ng of DNA were ligated to the same amount of plasmid in a 110 µl volume using 6 units of T4 DNA ligase (Life Technologies). The ligations were incubated at 16°C in a water bath for at least 24 hours. Salt was removed from the ligation using 30 000NW filters (Millipore).

1 µl ligation was mixed with 25 µl DH10B *E. coli* cells. The mixture was electroporated using Gene Pulser (BioRad) with the following conditions: Capacitance: 25 µF, voltage: 1.8 kV, resistance: 100 Ω. 600 µl of SOC medium (Life Technologies) were added to the cells and incubated for 50 min at 37°C. Dilutions of the cells were plated in 14 cm petri dishes containing LB medium with tetracyclin, IPTG and X-gal and grown at 37°C for 24 hours.

2.2.16.6 BAC insert isolation

In order to check the size of the inserts of the BACs, clones were picked at random and grown in 5 ml LB. Alkaline lysis minipreps (Sambrook et al. 1989) including a phenol/chloroform purification step were done to isolate BAC plasmid DNA. Approximately 250 ng of BAC plasmid DNA were digested with 5 U of *Not*I and run on a 1% low melting point agarose gel by PFGE using the following conditions: Initial time: 5s, final time: 15s, running time: 16 hours, voltage: 6 V/cm, angle: 120° and temperature: 14°C.

2.2.16.7 BAC clones arraying and storage

LB freezing medium: 36mM K_2HPO_4, 13.2mM KH_2, 1.7mM citrate, 0.4mM $MgSO_4$, 6.8mM $(NH_4)_2SO_4$, 7% v/v glycerol, LB.

Multiple transformations

Multiple transformations were carried out and plated on 22 x 22 cm plates (Genetix) containing 250 ml of LB with tetracyclin, IPTG and X-gal. In order to evenly distribute the clones on the plates, glass balls were used for plating the bacteria. 6 balls of 1 cm diameter were used for every plate. They were washed with 70% ethanol and autoclaved. The clones were grown for 28 hours, after that the plates were put at 4°C to increase the contrast between blue and white clones.

Colony picking

Insert containing colonies were picked mainly robotically using the Q-Pix robot (Genetix, UK) and the supplier's software. The software can be configured in order to select between white/blue colonies by a threshold that considers blue clones as background and allows the specification of size, proximity and shape of the clones that will be picked. Once located using a CCD camera, colonies are picked from culture trays by a 96 pin Picking Head and inoculated into pre-filled 384 well plates (Genetix, UK) containing freezing medium and 15mg/liter tetracyclin. The well plates were incubated at 37 °C overnight and stored at -70°C.

2.2.17 BAC library screening

2.2.17.1 Replica

Three replicas were made from each original 384 well plate using the BioGrid robot (BioRobotics, USA). The replicas were stored in different -70 °C freezers.

2.2.17.2 Filter spotting

High-density colony filters for hybridization-based screening of the library were prepared from one of the replicas. 72 well plates were spotted per nylon filter (PALL, Biodyne, UK) of 22 x 22 cm in a 5 x 5 pattern using the BioGrid robot (BioRobotics, USA) (Figure 2.2). This pattern allowed each colony to be represented twice on each filter. The complete library of 144 well plates was spotted in two filters. As control, the clone pSW1 (PE Biosystems, Foster City, CA, USA) was also spotted. The clones were grown at 37ºC overnight. On the next day, the filters were

treated in the following way: 5 min in 0.1% SDS, 5 min in denaturation buffer and 5 min in neutralization buffer. The filters were dried and baked for 2 hours at 80ºC.

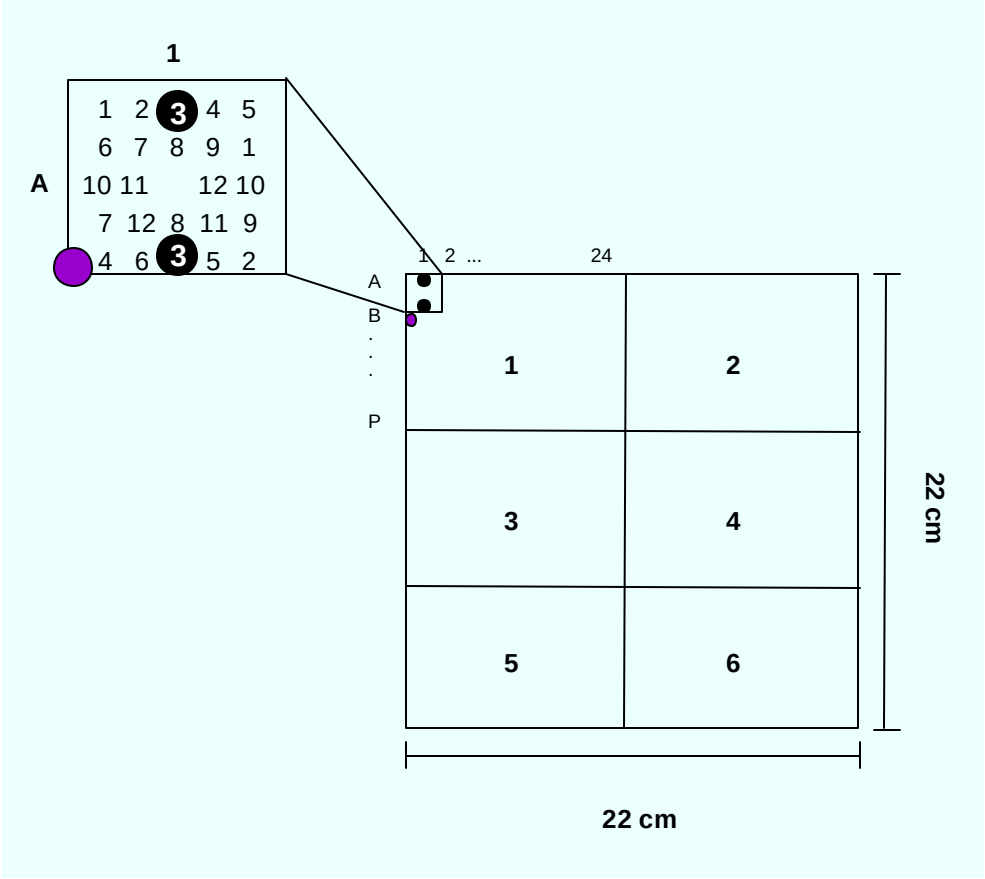


Figure 2.2 Library spotting. A 22 x 22 cm filter carries the clones of 72 well plates (384 wells plate). Therefore, the filter has 27 648 clones. Each filter is divided in six fields (1,2, 3, 4, 5 and 6). Each field contains clones corresponding to 12 well plates spotted in a 5 x 5 array where each clone is spotted twice in a specific pattern according to the number of the well plate. For example, the pattern shown corresponds to the well plate number 3. The clone is in the row A and column 1. The clone pSW1 is spotted in all the fields in the same position (violet spot at the left corner).

2.2.17.3 Filter hybridization

The same conditions as described in section 2.2.6 and 2.2.7 were used for labeling, hybridization and washes of the filters. 300 pg of pSW1 were added to the labeling reaction. Filters were exposed for two days using a Trimax Screen.

2.2.17.4 Nomenclature of the BACs

For naming the clones, the number of the well plate goes in first place followed for the letter of the row and finally the number of the column where the double spot appears. In order to differentiate BACs from this library and BACs from a library present also in the laboratory (BAC library BA), all the clones have BB (BAC library BB) at the first position of the name.

2.2.18 BAC insert subcloning into pBluescript $K^+/-$

1 µg of BAC plasmid DNA was digested with *Hin*dIII, *Eco*RV, *Xho*I or *Not*I (see results) and checked by electrophoresis on a 0.8% agarose gel. The pBluescript vector was completely digested with the same enzyme. Digested BAC plasmid and pBluescript were purified from agarose gel with Quiaquick columns. The plasmid was dephosphorylated with Shrimp phosphatase (Roche) and ligated to the digested BAC plasmid DNA in 10 µl reaction volume with 2U of T4 DNA ligase at 16°C overnight. The ligation was transformed into *E. coli* DH10B electrocompetent cells (Life Technologies), recovered with S.O.C medium and grown in LB medium containing 100 mg/liter ampicillin. Some transformant clones were taken from the plates and used as template for PCR screening.

2.2.19 Sequencing of PCR products, plasmids and BACs

PCR products were separated on a 1.5% agarose gel and single bands were cut from the gel. DNA was recovered using the QIAquick Gel Extraction Kit (Qiagen). Custom sequencing was performed by the ADIS service unit of MPIZ. It is based on the ABIPRISM Dye terminator cycle Sequencing Ready Reaction Kit (Applied Biosystems) and sequencing on an Applied Biosystems 377 DNA Sequencer (Applied Biosystems).

2.2.20 Sequence analysis

Sequence analysis was performed using the software package GCG (Wisconsin University, Version 9.1, UNIX, September 1997). Updated versions of the sequence database GenBank(NCBI) were searched for sequence homology of BAC ends and subcloned BACs using the programs BLASTN and BLASTX. The sequences of the subcloned BACs were asambled and analyzed using Jellyfish 1.3 (Bioware).

2.2.21 Mapping of BACs and candidate genes

The mapping of the genetic markers was done in the genetic map frame of the potato population F1840. RFLP filters were hybridized with the PCR products of *PR-5* and *PPO* genes. Additionally, PCR based genetic markers derived from the BACs and BAC sublcones sequences were mapped using the same population (collaboration with C. Gebhardt). Some markers were also mapped in the PD population (collaboration with M. Ghislain).

III Results

3.1 Construction of the BB BAC library

3.1.1 HMW DNA from nuclei compared to HMW DNA from protoplasts

Nuclei and protoplasts were isolated from potato leaves using the procedures described in material and methods. The DNA concentration and quality obtained with both methods were compared in order to choose the most efficient one for the construction of the library. The concentration of DNA in nuclei and protoplast plugs was determined by comparing digested potato genomic DNA (isolated in a standard way) with digested nuclei and protoplasts. The samples were digested with *Hin*dIII and run on a 0.9% agarose gel. Different dilutions of the standard DNA allowed the estimation of DNA concentration in nuclei and protoplasts plugs (Figure 3.1). Protoplasts dilutions of 2 x 10^7/ml had a concentration 3 µg of DNA per 75 µl of agarose plug. Nuclei isolated from 15 g of leaves and resuspended in 1 ml of 1 x HB buffer had a concentration of 1.5 µg of DNA per 75 µl of agarose plug. The quality of DNA in nuclei and protoplasts was measured by the amount of DNA partially digested under the same conditions. Figure 3.2 shows that a higher amount of protoplast DNA is digested in relation to nuclei. The digestion conditions seemed to be more optimal for protoplasts than for nuclei for getting fragments in the range of 100-300 kb. It can be seen in both nuclei and protoplasts controls that the amount of sheared DNA is minimal. Based on the higher concentration of DNA in protoplasts and due to that protoplast plugs were easier to handle than nuclei plugs, protoplasts were chosen for constructing the library.

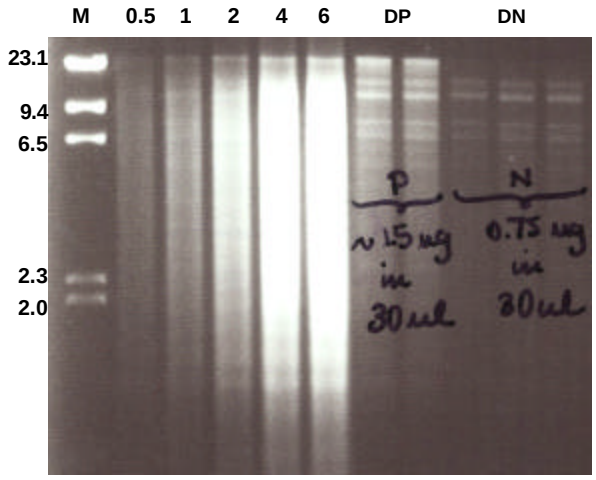


Figure 3.1 Concentration of HMW DNA from protoplasts and nuclei. M: *Hin*dIII lambda ladder, the sizes are indicated in kb. From 0.5 to 6 µg of digested potato DNA. DP: 30 µl of completely digested protoplasts plugs, DN: 30 µl of completely digested nuclei plugs.

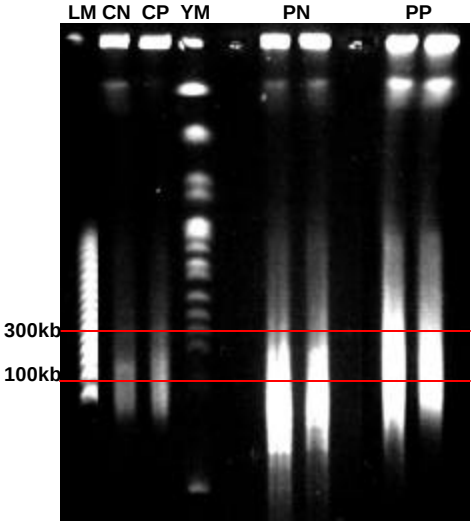


Figure 3.2 PFGE of partially digested HMW DNA from protoplasts and nuclei. LM: Lambda PFG Marker, CN: nuclei DNA without enzyme, CP: protoplasts DNA without enzyme, YM: Yeast Chromosome PFG Marker, PN: partially digested nuclei DNA, PP: partially digested protoplast DNA. The DNA digested in the range of 100-300 kb is indicated.

3.1.2 Standardization of partial digestion conditions

In order to get a high concentration of digested fragments in the range of 100-500 kb, optimal conditions for partial digestion were set up considering several factors.

Restriction enzyme

From the five cloning sites of the pCLD04541 vector (see Figure 2.1): *Hin*dIII, *Eco*RI, and *Bam*HI, were tested for partial digestion on potato protoplasts. The three enzymes produced similar results (Figure 3.3). However, *Hin*dIII was chosen for construction of the library.

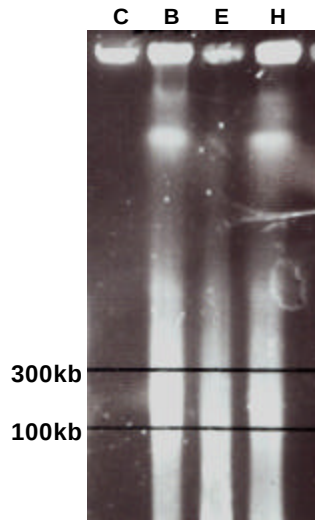


Figure 3.3 PFGE of partially digested HMW DNA using different enzymes. C: undigested DNA, lines 2 to 4: DNA partially digested with *Bam*HI, *Eco*RI and *Hin*dIII, respectively. Digested fragments in the range of 100-300 kb are indicated.

Concentration of enzyme

To set up the optimum amount of *Hin*dIII, 4, 8, 16 and 50 units of enzyme were added to 75 µl DNA agarose plugs and incubated at 37°C for 30 min. The figure 3.4A shows that the optimal digestion conditions were between 4 and 8 units. It also shows that the DNA was completely digested with 50 units of *Hin*dIII. To determine a narrower optimization window, 4, 6, 8 and 12 units of *Hin*dIII were added to 75 µl plugs and incubated at 37°C for 30 min. Figure 3.4B shows that using 6 units of enzyme most of the DNA ranged between 100 to 500 kb.

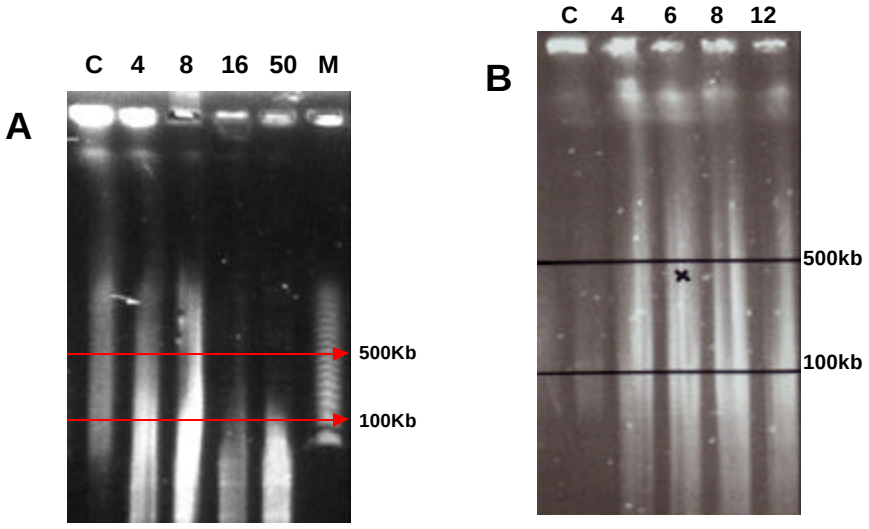


Figure 3.4 PFGE of partially digested HMW DNA using different concentrations of enzyme. **A.** Wide window. C: potato DNA without enzyme. Lanes 2-5: potato DNA partially digested with *Hin*dIII using 4, 8, 16 and 50 units of enzyme, respectively. M: Lambda PFG marker. **B**. Narrow window. C: potato DNA without enzyme. Lanes 2-5: potato DNA partially digested with *Hin*dIII using 4, 6, 8 and 12 units of enzyme.

Time of incubation with the restriction enzyme

Using 6 units of enzyme the time of incubation was optimized. The DNA was digested for 10, 20 and 30 minutes. The optimal time was determined in 20 minutes (Figure 3.5).

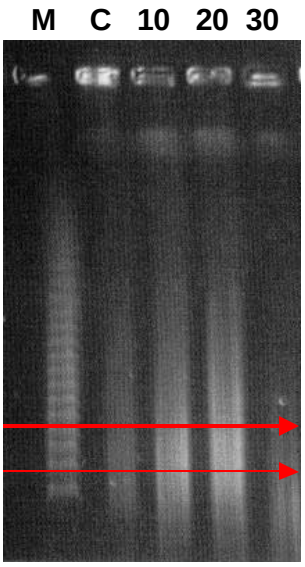


Figure 3.5 PFGE of partially digested HMW DNA using different times of incubation. M: Lambda PFG Marker, C: undigested DNA, lanes 3-5: potato DNA partially digested with *Hin*dIII for 10, 20 and 30 minutes.

Diffusion of buffer and enzyme

To facilitate the diffusion of the restriction enzyme into the agarose, the plugs were sliced in 2, 4 and 8 pieces prior to digestion. The efficiency of digestion was increased when the plugs where chopped in 8 pieces.

3.1.3 Standardization of DNA size selection

Several settings for PFGE running parameters were assed based on references and the manual user's guide of the CHEF DR-III system (BioRad). A better separation of fragments between 100-300 kb was observed using the following program: Initial time 60s, final time 90s and a running time of 18 hours. 6V/cm and an included angle of 120° were standard settings for the size range used (50-1000kb). All the gels were run at 14°C.

3.1.4 Standardization of gel purification

After PFGE, the DNA fragments in the range of 100-300 kb were excised from the gel and recovered using two different methods:

Electro-elution

It was possible to electro-elute DNA fragments in the range of 50-100 kb. Fragments of larger size could not run out of the agarose and be recovered. Figure 3.6 shows different size fragments electro-eluted, concentrated in a smaller volume and run on a 1% PFGE. For comparison, the fragments of the same size cut from the first PFGE and without any modification were also included in the gel.

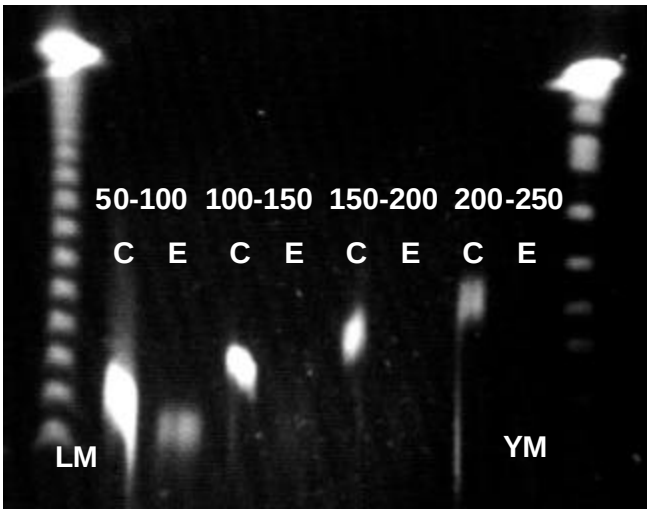


Figure 3.6 PFGE of different HMW DNA size fractions after electro-elution. The sizes are indicated in kb. LM: Lambda PFG Marker, C: Control fragments without modification, E: Electro-eluted fragments. YM: YM: Yeast Chromosome PFG Marker.

Gelase digestion

The agarose slices containing digested DNA were melted at 65°C for 5 min, which was sufficient for the complete melting of 100 mg agarose slices. The gel slices were equilibrated in Gelase buffer (40 mM Bis-Tris pH 6.0 and 40 mM NaCl) before the gel was molten. Since NaCl has a DNA protective ability, a higher concentration of the salt was used. More DNA was recovered from gel slices incubated with 100 mM NaCl than from slices that were incubated with a 40 mM NaCl concentration. The pH of the buffer also was considered as another factor for protecting DNA during Gelase digestion. 10 mM Tris pH 8.0 was used as an alternative to Gelase buffer of pH 6.0 but no differences were observed. Figure 3.7 shows a PFGE with partially digested DNA of different size ranges that were

treated with Gelase. The gel slices were cut from the first PFGE, incubated in Gelase buffer (100 mM NaCl), molten and digested with Gelase. Digested gel slices were compared to fragments of the same size cut from the PFGE without any treatment. From the figure it can be seen that fragments up 300 kb can be almost completely recovered. However, DNA fragments larger than 150 kb run slightly lower that the control meaning that some degradation of the DNA has occurred.

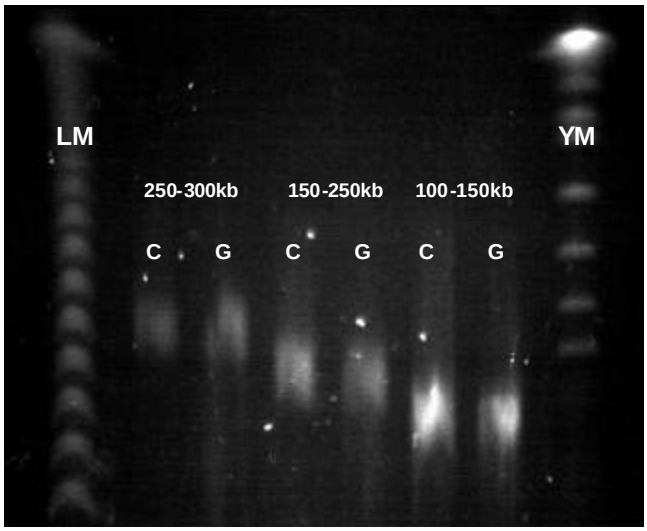


Figure 3.7 PFGE of different HMW DNA size fractions after gelase purification. LM: Lambda PFG Marker, C: untreated fractions, G: fractions treated with 100 mM NaCl and 40 mM Bis-Tris pH 6 as equilibration buffer and Gelase. YM: Yeast chromosome PFG Marker. The size of the fractions is indicated.

HMW DNA concentration

HMW DNA cut from PFGE and treated with Gelase was quantified by running on a 0.9% agarose gel. To determine the DNA concentration, different lambda dilutions were loaded on the same gel. The concentration ranged from 0.5 to 2 ng of DNA per µl. Fragments larger than 50 kb, run slightly higher than the lambda marker in a normal agarose gel. The DNA was used without further modifications.

3.1.5 Ligation and transformation

Ligation conditions

Gel slices containing different size DNA fractions (Fraction 1: 250-300 kb, Fraction 2: 150-250 kb and Fraction 3: 100-150kb) were treated and kept separately in independent ligation reactions. Before adding ligase buffer and T4 ligase enzyme, the plasmid and insert DNA were mixed and incubated at 50°C for 10 min to separate sticky ends. A large volume of ligation (110 µl) was used in order to avoid self-ligation of genomic DNA. The molar ratio of insert:plasmid used was 1:3. Assuming an average insert size of 100 kb for all fractions (1, 2 and 3) and the size of the vector about 30 kb, 100 ng of insert were ligated to 100 ng of plasmid. Prior to electroporation, salts were removed to avoid arcing.

Transformation conditions

Some factors like voltage and time were tested in order to improve the transformation efficiency[5] and to increase the size of BACs being transformed. Voltages from 1.3 to 1.9 kV were tested. Transformations using 1.8kV were most efficient. The inserts showed an average size of 80kb and a transformation efficiency of up to 900 clones/ng DNA (Figure 3.8). The size of the inserts was determined as explained in section 3.1.6. Independent transformations of the same ligation reaction gave different transformant / non transformant ratios. From $^1/_3$ to $^1/_2$ of non transformants were observed and these fractions correlate with the age of the ligation reaction. The high number of transformant clones compensated however the number of non- transformants. A ligation reaction could be kept at -20°C up to 2 months, after longer periods less transformant clones were obtained and shorter inserts were observed.

[5] Transformation efficiency was calculated as the number of transformant clones per ng of DNA insert used in the ligation.

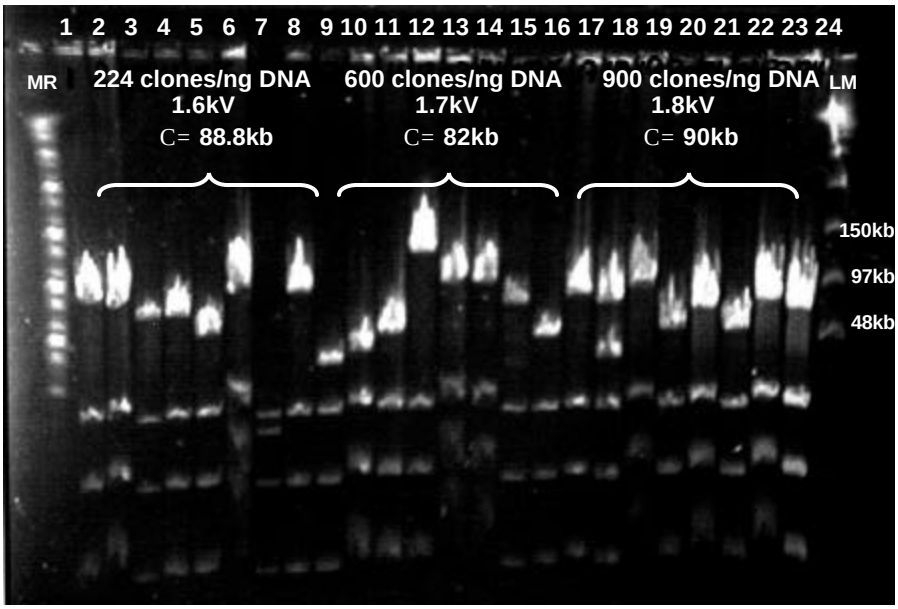


Figure 3.8 PFGE of BAC clones transformed using different voltage conditions. The transformation efficiency and the average insert size are indicated for every voltage used. MR: Mid Range Marker, 1-8: clones transformed using 1.6 kV, 9-16: clones transformed using 1.7kV and 17-24: clones transformed using 1.8 kV, LM: Lambda PFG Marker. Line 7 does not contain insert.

Another parameter considered was the pulse time of the electric shock. The time of electroshock depends on the electric resistance. A resistance of 200Ω is often used in E. coli standard transformations, and gives a pulse time of 4.2-4.6 sec. It was decided to use a resistance of 100Ω that gives a pulse time of 2.2-2.6 sec. This means the cells are stressed a shorter time and that may condition higher number of transformant cells.

3.1.6 Determination of insert size of the BAC clones

Different ligations from different size fractions were transformed independently. Clones were picked at random in order to check the insert size of the ligations. Standard alkaline lysis minipreps (Sambrook et al. 1989) were done from randomly selected transformant clones. 5 ml of culture inoculated with a single colony produced mostly 1 μg of DNA. Aproximately 250 ng plasmid DNA were digested with 4U of *Not*I at 37°C for 4 hours. *Not*I is an 8-cutter and most insertions of 100 kb did not contain a *Not*I site. The vector has 5 *Not*I sites (see Figure 2.1), and the fragments released after digestion have a size of 12, 8, 6, 3 and 0.4 kb.

When a BAC is digested with *Not*I and run in a PFGE, the insert appears as the largest band including 8 kb of the vector followed by 3 smaller bands (12, 6 and 3 kb) corresponding to the vector (Figure 3.9). The highest average insert size was obtained with Fraction 3 (80kb). Smaller insert sizes were obtained with Fraction 1 and Fraction 2.

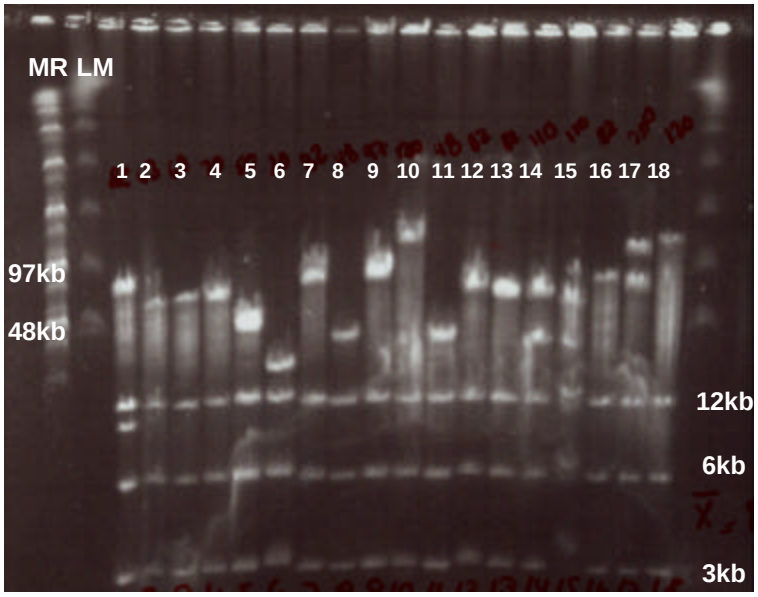


Figure 3.9 Analysis of 18 random BAC clones by PFGE. MR: Mid Range Marker, LM: Lambda PFG Marker, 1-18: *Not*I digested BAC clones. The inserts are the higher bands on every track and the three lower bands of 12, 6 and 3 kb are the digested vector. Some inserts show internal *Not*I sites. The clones correspond to the Fraction 3.

3.1.7 Insert size distribution

90 random clones were picked. The BAC DNA was isolated, digested with *Not*I and run on PFGE. The distribution of the insert sizes of these BACs is shown in figure 3.10. All the ligations used for transformations were done with fractions containing fragments larger than 100 kb. However 30% of the clones show insert sizes from 25 to 75 kb, 40% had sizes from 75-100kb, 15% showed sizes from 100-125kb and 10% had sizes larger than 125kb.

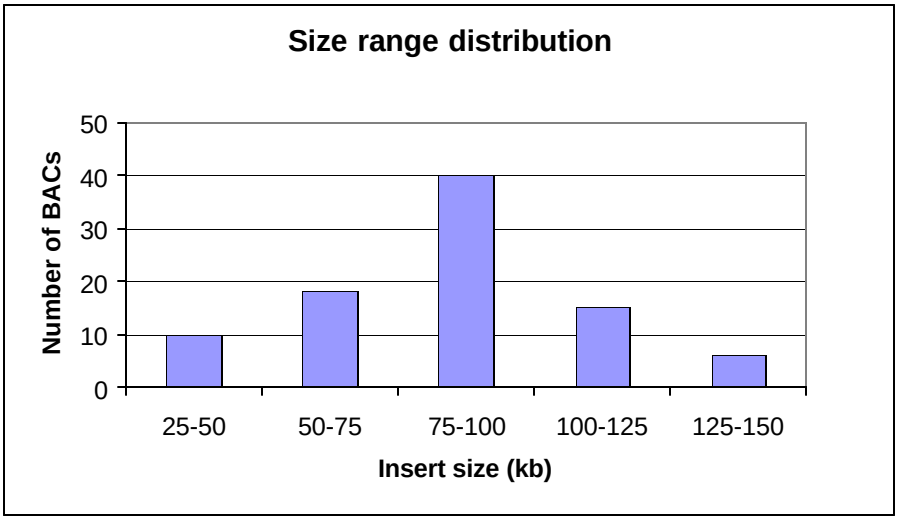


Figure 3.10 Size insert distribution of 90 potato BAC clones. Ninety BAC clones were randomly selected, digested with *Not*I and run on a PFGE. The figure shows an average insert size in the range of 75-100kb.

3.1.8 BAC clones arranging and storage

Statistical evaluation of the number of clones needed to contain a target sequence or gene

The number of clones needed in the library was determined using the following equation (Clarke and Carbon, 1976):

N = ln (1- P) / ln (1 – L/G) where,

N = number of clones in the library
P = probability to contain the target gene
L = length of average clone insert in base pairs
G= haploid genome length in base pairs

Considering a probability of 99% to contain the target gene, the potato haploid genome in ~850 Mbp (2n=2x=24) (Arumuganathan and Earle 1991) and an average insert size of the BACs of 80 kb, about 48 000 clones are needed which represent 3-4 haploid genome equivalents.

Number of clones in the library

144 well plates (384 well plates) were inoculated. Theoretically, they represent 56, 496 clones. However, not all the clones inoculated in the media of the plates grew. In average, 350 clones were present in every plate making approximately 50 400 clones in the library.

3.2 BAC library screening: The *PAL* gene family

In order to test the coverage of the library as well as the hybridization and stringency conditions, the screening of the library was started with a cDNA from *PAL* (*phenylalanine ammonia-lyase*). This gene belongs to a gene family. Therefore, several positive clones were expected. The same procedures followed for genomic Southern blot were used for labeling of the probe, hybridization and washing conditions of the filters (see Material and Methods). Figure 3.11 shows the autoradiogram of the filter 1 (well plates 1-72) after hybridization with a *PAL* probe.

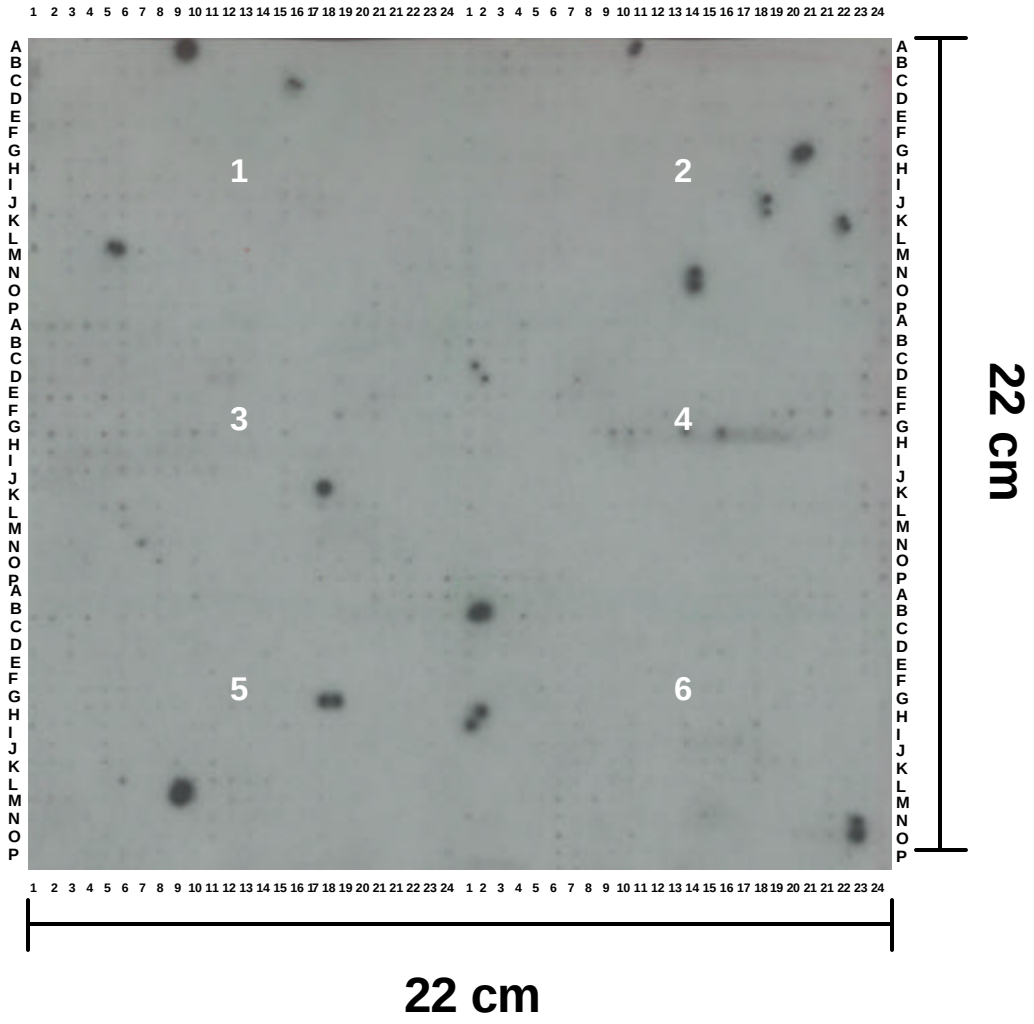


Figure 3.11 Autoradiogram of the Filter 1 (well plates 1-72) after hybridization with *PAL*. The background signal is due to the Control clone that has been spotted in all the fields. This background facilitates the identification of positive clones, which can be recognized as the double spots on the film. The number of the fields (1 to 6) is indicated as well the columns and rows for each field.

29 positive BAC clones containing sequences with homology to *PAL* were identified on the filters and picked: BB8b9, BB11d15, BB11m5, BB15k17, BB19b10, BB21l21, BB24h19, BB28h13, BB28j17, BB31j20, BB32o5, BB36m13, BB38d1, BB58f17, BB64g1, BB70f3, BB72a1, BB86e9, BB86h9, BB88e13, BB88m2, BB96g8, BB101k19, BB104p18, BB112h11, BB113e8, BB122b1, BB126b9, BB139h9. Plasmid DNA of these BACs was isolated and digested with *Hin*dIII. Southern blot with the *Hin*dIII fragments was carried out and the blot was hybridized with *PAL*

(Figure 3.12). The results showed that most of the BACs have 2 or 3 bands. There were overlaps between BACs[6] and some BACs were redundant[7]. Taken together it can be concluded that members of the PAL gene family cluster. There were approximately 16 different bands with a frequency from 1 to 6 and an average of 3.4. This number is in agreement with the estimated redundancy of 3-4 times the haploid genome of the library.

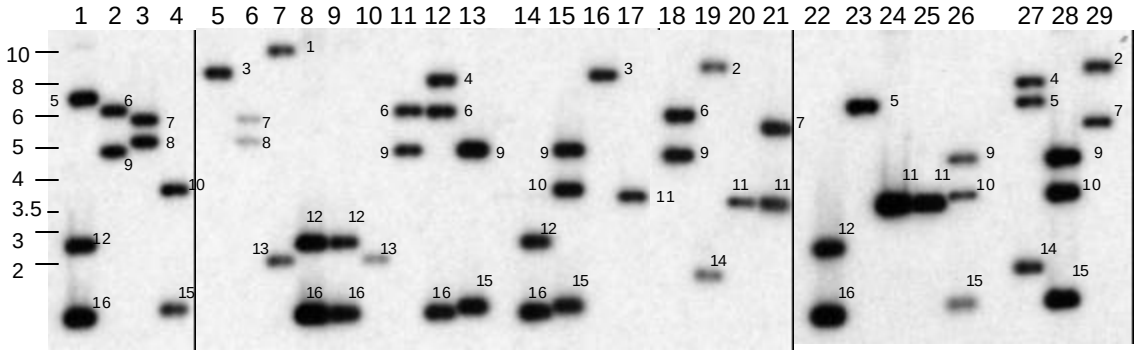


Figure 3.12 Autoradiogram of a Southern blot from BAC clones with homology to *PAL*. Lanes 1 to 29 indicate clones BB8b9(1), BB11d15(2), BB11m5(3), BB15k17(4), BB19b10(5), BB21l21(6), BB24h19(7), BB28h13(8), BB28j17(9), BB31j20(10), BB32o5(11), BB36m13(12), BB38d1(13), BB58f17(14), BB64g1(15), BB70f3(16), BB72a1(17), BB86e9(18), BB86h9(19), BB88e13(20), BB88m2(21), BB96g8(22), BB101k19(23), BB104p18(24), BB112h11(25), BB113e8(26), BB122b1(27), BB126b9(28) and BB139h9(29). Every different hybridizing band has been numbered (1 to 16). The sizes in kb are indicated at the left.

17 *PAL* clones representing the different band patterns were digested with *Not*I and separated by PFGE. The sizes of the clones are the following: BB8b9 (60kb), BB11m5 (80kb), BB15k17 (80kb), BB19b10 (70kb), BB31j20 (80kb), BB32o5 (100kb), BB36m13 (80kb), BB58f17 (40kb), BB64g1 (100kb), BB86e9 (60kb), BB86h9 (40kb), BB88m2 (50kb), BB101k19 (80kb), BB112h11 (60kb), BB122b1 (100kb), BB126b9 (100kb) and BB139h9 (60kb).

3.3 The *PR-5* gene family in potato

The BB BAC library was further screened with probes for the *PR-5* or thaumatin-like gene family. This gene family has basic, neutral and acidic members and it is reported that the different members do not cross hybridize (Velazhahan et al., 1999). For this reason, two probes, one acidic and one basic were used. The

[6] Two or more clones are overlapping when they share at least one hybridizing band.
[7] Two or more clones are redundant when they share all the hybridizing bands.

acidic group was represented by *NtPR-5*, a cDNA from tobacco codifying for intercellular PR-R major form (X12739) (Payne et al. 1988). A single clone was picked with this probe: BB43d14. This probe has been mapped in the population BC9162 to LG XII (Gebhardt et al., 2001) and this was the reason to use it as a candidate gene for the QTL on LGXII of the PD population (Ghislain et al. 2001). On the other hand, the basic group includes osmotin-like genes and the probe used was a PCR product derived from osmotin primers (see Materials and Methods). 6 BAC clones were fished with this probe: BB18h19, BB41d14, BB48h15, BB82a11, BB86e10 and BB99n12. The probe was also used to hybridize RFLP filters from the F1840 population to determine the map position. The RFLP experiment gave as result an osmotin locus on LG VIII.

The BAC plasmids were isolated, digested with *Not*I and run on PFGE to determine the size of the inserts (Figure 3.13).

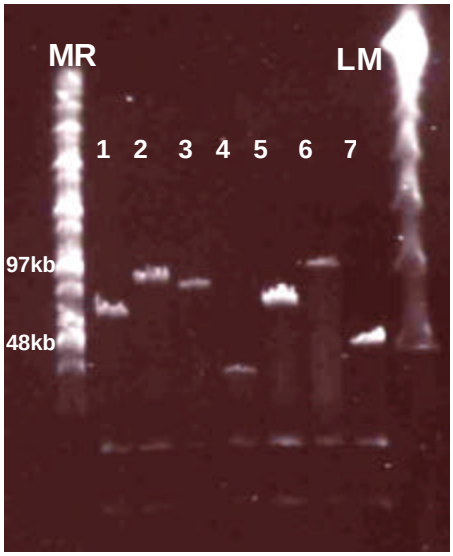


Figure 3.13 PFGE of *PR-5* like clones. MR: Mid Range Marker. Lanes 1 to 7, BACs BB41d14, BB48h15, BB18h19, BB99n12, BB82a11, BB86e10 and BB43d14 after *Not*I digestion. LM: Lambda marker. The size of the inserts was determined by comparing them to the size markers.

3.3.1 Physical mapping of the *PR-5* gene family in potato

3.3.1.1 *PR-5* cluster on LG XII

A Southern gel blot with DNA of PD59 and BB43d14 was done using *Hin*dIII restriction. The hybridization with *Nt**PR-5*** showed that the genomic DNA pattern is nearly identical to the BAC DNA pattern (Figure 3.14). This indicates that most of the genes with homology to *NtPR-5* are clustered in the BAC BB43d14.

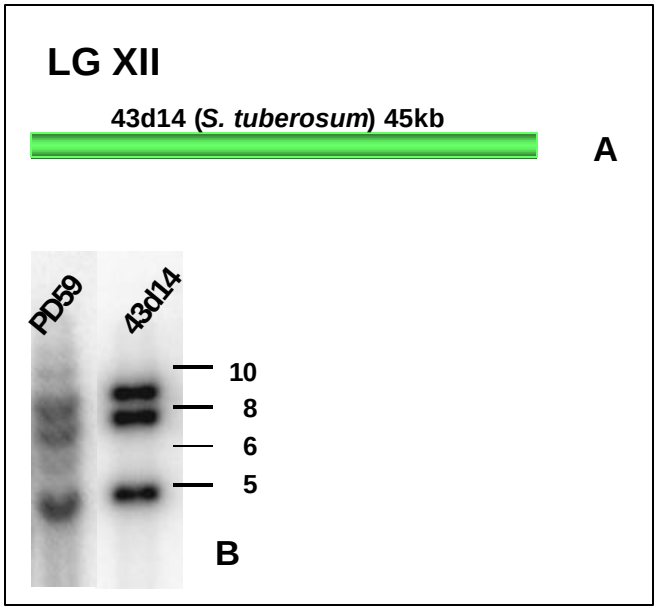


Figure 3.14 Intercellular *PR-5* genes in potato. **A.** Schematic representation of BAC BB43d14 indicating the size and the parental line from which is derived. **B**. Genomic blot from PD59 and BAC BB43d14 digested with *Hin*dIII and hybridized with *NtPR-5.* The sizes are indicated in kb on the right.

0.5 kb sequence from the T7 end of the BAC BB43d14 showed 89% homology to E22 (X15224) and 81% homology to E2 (X15223). E22 and E2 are the genes codifying for PR-S (PR-R) major (X12739) and minor form, respectively (van Kan et al. 1989). Both proteins are present in the intercellular fluids of tobacco cells after infection with TMV (Pierpoint et al. 1987). Further direct sequencing into the T7 end of the BAC failed due to the presence of several highly homologous copies in the BAC. 1.9 kb sequence from the T3 end of the BAC showed an open reading frame (ORF) with 50% identity to a putative retroelement of *A. thaliana* (NP_179258). To get more sequence information of acidic members of *PR-5*-like genes in potato, the BAC was subcloned into pBluescript. Since the sequence of E2 has two *Hin*dIII restriction sites and E22 has one *Hin*dIII site, it was decided to use enzymes that do not cut these genes internally. *Eco*RV, *Xho*I and a combination of *Xho*I/*Not*I were used to generate subclones with complete ORFs. The screening of positive clones was done by PCR using the primers E22f (5'- gat ggt agt ggc cga ggc aaa tg-3') and E22r (5-' tcc agc agg aca tgt aaa caa act-3'). These primers amplify part of the ORF of E22 and E2 and do not amplify osmotin-like genes. The primers were designed based on the high sequence similarity between tobacco E22 and the T7 end of BB43d14. For this reason, the sequences obtained with these primers will be called

E22-like genes. In this way, an *Eco*RV subclone (**319**) of 7 kb, a *Xho*I subclone (**321**) of 7 kb and a *Xho*I/*Not*I subclone (**21**) of 4 kb were obtained. The subclones were firstly sequenced with T3 and T7 primers and these sequences were used to designed new primers and walk along the inserts.

Partial sequences of the subclones were obtained and searched for ORFs using Jellyfish Version 1.3 (Bioware) and blasted using the BLASTX 2.2.1 program of the NCBI (GenBank) to find similarities with *PR-5* genes. Three ORFs with homology to *PR-5* genes were found. One ORF was present at the T3 end of 319, the ORF was called **319b**. 319b is 690 nucleotides long and the translated ORF is 227 amino acids. When the ends of 21 were analyzed it was noticed that the T7 end sequence was identical to the T3 end sequence of 319, suggesting that 21 was an internal fragment of 319. Indeed, both sequences correspond to the T7 end of the BAC BB43d14. The T3 end of 21 showed a second sequence with similarity to *PR-5* genes. The sequence was missing part of the N-terminus. Subclone 319 was used as template to get the sequence upstream the T3 end of 21. For this purpose, a sequencing primer was designed 50 bp from the T3 end of 21 in the reverse orientation. The sequencing provided a 50 bp sequence identical to the one of 21 together with the N-terminus that was missing. Therefore, it was concluded that 21 was an internal fragment of 319. The complete ORF was obtained and was called 21a. 21a is 683 nucleotides and 229 amino acids long.

Up to 1 kb sequence of both ends of subclone 321 did not present any similarity to *PR-5* genes. From the PCR screening it was known that at least one E22-like sequence was present in 321**.** For this reason, it was decided to use the complementary sequences of E22f and E22r: E23f (5'-ca ttt gcc tcg gcc act acc atc-3') and E23r (5'-agt ttg ttt aca tgt cct gct gga-3') respectively as sequencing primers to get the upstream and downstream sequences of the gene copy. In this way a third ORF was found, this was called **321c** with 741 nucleotides and **246** amino acids. Later, 2.4 kb from the T7 end of 321 a fourth *PR-5*- like sequence was found, this was called **321d. 321d** had a shorter N-terminus and no starting codon was found. Therefore, **321d** seems to be a pseudogene. 321d is next to 321c in the same orientation. The position of the ORFs and their orientations in the subclones is shown in figure 3.15.

The similarities of the *PR-5* like sequences of subclones 319, 21 and 321 are summarized in Table 3.1[8]. The degree of identity between **21a**, **319b** and **321c** is shown in Table 3.2.

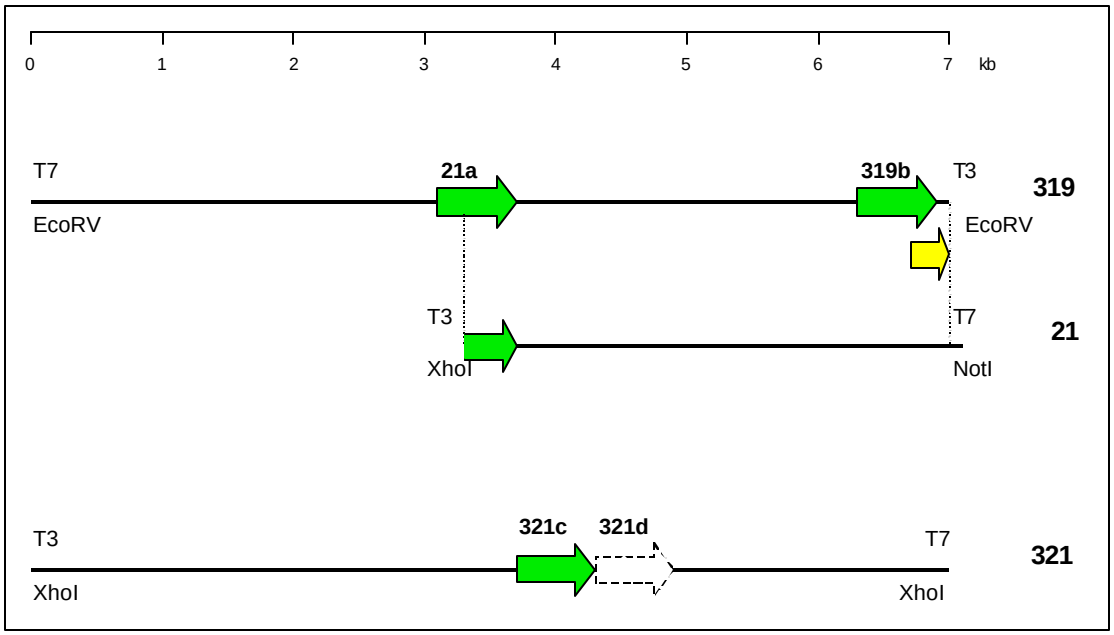


Figure 3.15 Subclones of BAC BB43d14 with sequence similarity to *PR-5* genes. The size in kb of the subclones is indicated by the scale. T3 and T7 indicate the position of the T3 and T7 sequencing primers in the subclones. The T3 end of subclone 319 is identical to the T7 end of subclone 21 and both correspond to the T7 end of BAC BB43d14. The position of subclone 21 in subclone 319 is shown by the dashed lines. The solid green arrows indicate the size and orientation of the ORFs with homology to *PR-5* genes. The solid yellow arrow indicates a smaller ORF with no homology to known proteins. The dashed arrow under 321d indicates that this is a pseudgene with a short N-terminus. The enzymes used for subcloning are indicated.

[8] The sequences are in Appendix 7.3.

Table 3.1 Similarity of *PR-5*- like sequences derived from BAC BB43d14.

Subclone (total lenght in kb)	**Subclone end (read sequence in bp)**	**ORF (amino acid length)**	**Identity at aminoacid level using BLASTX**
319 (7)	T3 (2564)	319b (227) 319b1 (111)	85% X12739 PR-5 *N. tabacum* -
	T7 (1170)	-	-
21 (4)	T3 (1500)	21a (229)	60% AF378571 thaumatin-like protein *Sambucus nigra*
321 (7)	T3 (1000)	-	-
	T7 (4779)	321c (246)	77% X12739 PR-5 *N. tabacum*
		321d *	62% AF378571 thaumatin-like protein *Sambucus nigra*

* Sequence similarity to *PR-5* genes with short N-terminus.

Table 3.2 Identity between 21a, 319b and 321c at amino acid level.

	21a	319b	321c
21a	-	54.9	54.1
319b		-	78
321c			-

The ORF of **21a**, **319b** and **321c** do not have introns. Amino acid sequence comparison of **21a, 319b, 321c** and **321d** is shown in figure 3.16. **319b** and **21a** show a short C-terminus similar to tobacco E22 and E2 genes (van Kan et al. 1989). On other hand, **321c** shows a longer C-terminus in relation to **319b** and **21a**. The C-terminus is more similar to osmotin-like genes. Osmotin-like genes have longer C-terminus.

```
      1                                                          50
319b  MHFLK.FFPL FVFLY.FGQY YLYVTHAATF DITNRCTYPV WAAASPGGGR
321c  MHFLK.FSPL FVFLY.F... .LSVTHAATF DITNRCTYPV WAAASPGGGR
 21a  MRFIKSFFI. ..SLYIIT.S HISIQANAII DIQNNCPFTV WAAAVPGGGR
321d  ~~~~~~IFVL KHELHQISLF YDSAEADTTI DIQNNCPFTV WAAAVPGGGR

      51                                                        100
319b  RLDSGQTWNL N...VNPGTI QARIWGRTNC NFDGSGRGKC ETGDCNGLLE
321c  RLDLGQTWNI N...VNPGTT MGRIWGRTNC NFDGSGRGKC QTGDCNGRLE
 21a  QLEYGDTWKI EPDMTTSSTK KGRIWGRTNC NFDSLGHSQC QTGDCNGLLE
321d  RLEYGGTWKI E.....ANTK NGRIWGRTNC NFNTLGKGQC QTGDCNGLLE

      101                                                       150
319b  CQGY.GSPPN TLAEFALNQP NNLDFVDISL VDGFNIPMEF SPINGG.CR.
321c  CQGF.GTVPN TLAEFALNQP NNLDFVDISL VDGFNIPMEF SPINGG.CR.
 21a  CQTFSNTPPN TLAENALNQY NNNDFFDLSL VEGFNVPMEF SPVFADECST
321d  CQAY.GTPPN TLAEYAINQF NNNDFLGISL VDGFNIPMEF SPVSADECSV

      151                                                       200
319b  NLLCNAPIND QCPNELRTPG GCNNPCTVFK TNEFCCTNGP GSCGPTDLSR
321c  KIRCSADING QCPSELRAPG GCNNPCTVFK MKEFCCTNGP GSCGPTDYSK
 21a  RIRCTADIIG QCPNELRTPG GCNNPCTIFK TDEYCCTS.. GNCGPTNYSR
321d  RITCTADIIG QCPNELRTPG GCNNPCTVFK TAKYCCTS.. GTCGSTKYSR

      201                                                       250
319b  FFKQRCPDAY SYPQDDPTSL FTCPAGTNYK VVFCP*~~~~ ~~~~~~~~~~
321c  FFKQRCPDAY SYPLDDPTSM FTCHAGTNYK VVFCPGNTFT PFNFFYLSRL
 21a  FFKDRCSTSF SYPADDLSSL FTCPSGTTYR VVFCP*~~~~ ~~~~~~~~~~
321d  FFKDKCSTSY SYSLDSSTSA FFCPSGTSYR VVFCP*~~~~ ~~~~~~~~~~

      251
319b  ~~~~~~~~~
321c  DLTNLLRK*
 21a  ~~~~~~~~~
321d  ~~~~~~~~~
```

Figure 3.16 Amino acid sequence comparison of *PR-5*- like sequences from BB43d14. 21a, 319b and 321c have complete ORFs. 321d has homology to *PR-5* genes but the translated frame has a short N-terminus in relation to the other genes. The 16 cysteine residues are conserved in all *PR-5*- like sequences and they are indicated in yellow boxes. * indicates the stop codon.

3.3.1.2 Promoter structure of E22-like genes

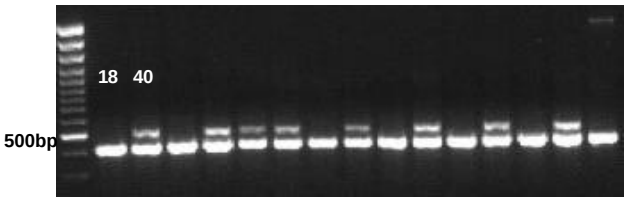
Partial promoter sequences of the copies **21a**, **319b** and **321c** were searched for *cis*-acting elements. Since **319b** and **321c** show homology to E22 and **21a** is less than 2 kb apart from **319b**, up 1 kb upstream of the starting codon of the three genes were compared to the promoter region of tobacco E22. The structure of the genes encoding thaumatin-like proteins has been studied in E22 and E2 tobacco genes (van Kan et al., 1989). A putative TATA box identical to the one of E22 was found in copy **319b** (Figure 3.17). Copies **21a** and **321c** did not present any recognizable TATA box at the predicted positions. Up to 60 bp upstream sequence of **319b**, **321c** and E22 showed to be conserved. Further no conservation was found. However, a short stretch of the promoter sequence of **321c** was found to share homology to E22. When **319b** and **321c** where compared to **21a**, also no conserved sequence was found. E22 and E2 show higher homology in the untranslated sequences than potato genes **319b**, **321c** and **21a**.

```
 321c -881  TAGTGaCGGCTTGACTTaCAaGTCAAGAgGCAAGATTTT  -842
  E22 -564  TAGTGtCGGCTTGACTTgCAgGTCAAGAaGCAAGATTTT  -526

319b  -74  AATtAGCTATTTAAACCCgT  -55
 E22  -36  gATcAGCTATTTAAACCCaT  -17


319b  -54  AGATTGTGTCTCaAAATAAACACATTCTCAAGTTTAACAAcAcAAAAAAAA..TG  +3
 E22  -16  AGAT..aGTCTCCAAATAAACACATTCTCAAGTTTAcaAAAAgAAAAAAAAAATG +37
321c  -60  tGAaaGTGTCTtCAAATAAACACATTCTCAAaTTgAacAtAcacAcAcAtAAAAG  -5
```

Figure 3.17 Conserved untranslated sequences of E22-like genes in potato. The putative TATA box from E22 and 319b are bold and in light blue boxes. 60 bp sequence upstream the starting codon of 319b and 321c and –16 to 32 of E22 showed to be conserved. The translation initiation of E22 occurs at the position +35. The conserved nucleotides are in upper cases.

3.3.1.3 Mapping of BB43d14

The BAC BB43d14 was mapped in the F1840 population with a SCAR marker derived from primers E22f and E22r. As expected, the marker mapped to LGXII (Figure 3.18). The *NtPR-5* probe was previously mapped to LGXII in the BC916[2] population (Gebhardt et al. 2001).



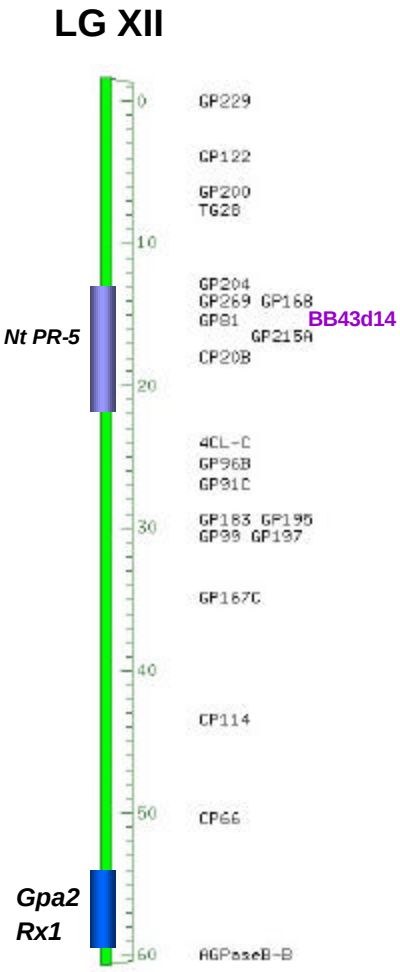


Figure 3.18 Mapping of BAC BB43d14 in the population F1840. Above, the segregation of the SCAR marker E22 in some individuals of the F1840 population. To the right, LG XII showing the position of BB43d14 and the RFLP marker *Nt*PR-5. The lilac box indicates that this is a defense related gene. At the bottom of the chromosome, the blue box indicates the position of two resistance genes *Gpa2* (resistance to *Globodera pallida*) and *Rx1* (resistance to PVX virus). The position of some RLFP markers (GP: genomic potato, CP: cDNA potato and TG: tomato genomic) and the gene 4CL (4-Coumarate:CoA Ligase) are also indicated.

The next step was to map BB43d14 in the PD population. The SCAR marker used in the F1840 population did not segregate in the PD population. However, it was thought that a CAPS marker could be derived from the PCR sequences. For this reason, the PCR products of primers E22f and E22r on genomic DNA of PD59 and the parental lines *S. phureja* (phu) and dihaploid *S. tuberosum* (tbr) were sequenced. Sequence alignment of the PCR products of phu, tbr, PD59, 319 and 321 is shown in figure 3.19. The genomic sequences showed overlapping signals (N) indicating the amplification of more than one *PR-5* gene. The overlapping signals occurred where **321c** and **319b** are polymorphic. It can be concluded that E22f and

E22r amplify two copies in the potato genomic DNA: **319b** and **321c**. No marker could be derived from the polymorphic nucleotides.

```
          1                                                          50
 phuE22   GYTAGAATGT CAAGGTTYTG GTASTSCACC CAAYACTTTA GCYGAATTTG
PD59E22   GYTAGAATGT CAAGGTTYTG GTASTSYACC CAAYACTTTA GCYGAATTTG
 tbrE22   GYTAGAATGT CAAGGTTYTG GTASTSYACC CAAYACTTTA GCYGAATTTG
 319E22   GTTAGAATGT CAAGGTTATG GTAGTCCACC CAACACTTTA GCCGAATTTG
 321E22   GCTAGAATGT CAAGGTTTTG GTACTGTACC CAATACTTTA GCTGAATTTG

          51                                                        100
 phuE22   CWCTWAAYCA GCCCAATAAC TTGGAYTTTG TCGATATTTC TCTAGTCGAT
PD59E22   CWCTWAAYCA GCCCAATAAC TTGGAYTTTG TCGATATTTC TCTAGTCGAT
 tbrE22   CWCTWAAYCA GCCCAATAAC TTGGAYTTTG TCGATATTTC TCTAGTCGAT
 319E22   CTCTAAATCA GCCCAATAAC TTGGATTTTG TCGATATTTC TCTAGTCGAT
 321E22   CACTTAACCA GCCCAATAAC TTGGACTTTG TCGATATTTC TCTAGTCGAT

          101                                                       150
 phuE22   GGATTCAACA TYCCTATGGA ATTCAGCCCA ATCAACGGTG GGTGTCGYAA
PD59E22   GGATTCAAYA TYCCTATGGA ATTCAGCCCA ATCAACGGTG GGTGTCGYAA
 tbrE22   GGATTCAAYA TYCCTATGGA ATTCAGCCCA ATCAACGGTG GGTGTCGYAA
 319E22   GGATTCAACA TTCCTATGGA ATTCAGCCCA ATCAACGGTG GGTGTCGTAA
 321E22   GGATTCAATA TCCCTATGGA ATTCAGCCCA ATCAACGGTG GGTGTCGCAA

          151
 phuE22   WMTCYKATGC WMCGCACMTA TTAAYG
PD59E22   WMTCYKATGC WMMGCACMTA TTAAYG
 tbrE22   WMTCYKATGC WMMGCACMTA TTAAYG
 319E22   TCTCTTATGC AACGCACCTA TTAACG
 321E22   AATCCGATGC TCAGCAGATA TTAATG
```

Figure 3.19 Sequence alignment of E22-like genes from phu, tbr, and PD59. The alignment corresponds to partial PCR sequences derived from phu (*S.phureja*), tbr (*S. tuberosum*), PD59, 319 and 321**.** Overlapping signals occur in genomic templates in the positions where **319b** and **321c** are polymorphic. The single letter code has been used in positions where two nucleotides are present Y = C or T, W = A or T, S = C or G, K = G or T and M = A or C. In yellow boxes, nucleotides that are unique to 319E22 and in green boxes nucleotides that are unique to 321E22.

New primer pairs were developed, therefore using the T3 end sequence from BAC clone BB43d14. The amplification was carried out with the parental lines phu and tbr. The PCR products were sequenced and no polymorphisms were found. Later, promoter sequences of the **21a**, **319b** and **321c** copies were used with the same purpose. *S. phureja* and *S. tuberosum* turned out to be highly homozygous in these loci. Few SNPs were detected, however, in the parental lines and in two cases

it was possible to convert the polymorphisms into CAPS markers[9]. Figure 3.20 shows one of the cases in which a SNP was converted into the CAPS marker 319prf. In both cases, the segregating marker was derived from *S. phureja* (Figure 3.21). SNPs present in *S. tuberosum* were not sufficient for developing allele-specific primers in order to map the BAC in the male parental line.

```
      1                                                          50
PD59  TATTTTCCAC CTAAAAGAAA TTTCTATCTA AGCCTATCTA GATAACAAAA
 tbr  TATTTTCCAC CTAAAAGAAA TTTCTATCTA AGCCTATCTA GATAACAAAA
 phu  TATTTTCCAC CTAAAAGAAA TTTCTATCTA AGCCTATCTA GATAACAAAA

        AvaII

      51                                                         100
PD59  GGACCAACAT TAAAATTTCA ATAAATCAAG GACTATTACA CAAATTTCCA
 tbr  GGACCAACAT TAAAATTTCA ATAAATCAAG GACTATTACA CAAATTTCCA
 phu  GGACCAACAT TAAAATTTCA ATAAATCAAG GACTATTACA CAAATTTCCA
           T
```

Figure 3.20 Alignment of partial PCR sequences from PD59, tbr and phu obtained with primers from the promoter region of copy **319b**. A single overlapping signal indicating two nucleotide bases (C and T) was converted into a segregating marker using *Ava*II.

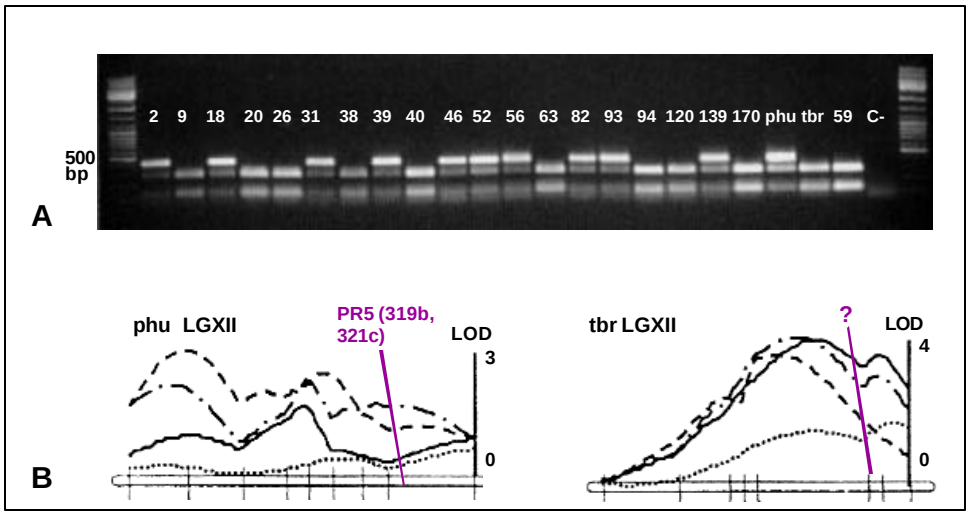


Figure 3.21 Mapping position of clone BB43d14 in *S. phureja*. **A**. The figure shows the segregation of the marker 319pr in some individuals of the PD population. **B**. LG XII from S. *phureja* and *S. tuberosum* with LOD score curves for late blight resistance are depicted. The position of the BAC is indicated as PR-5 **(319b, 321c)** because markers derived from 319b and 321c mapped to the same position. A question mark indicates the putative position of the same loci in *S. tuberosum*.

[9] The sequences of the primers are listed in the Appendix 7.2.

The mapping of BB43d14 in the PD population showed that PR-5 genes do not overlap with the QTL effect from *S. phureja* on LG XII. On the other hand, correlation between markers from both species suggests that it might fall in the peak of the QTL of *S. tuberosum* (Ghislain, personal communication). This hypothesis has to be tested and it is a point to be discussed.

3.3.1.4 Osmotin-like BACs

BACs BB18h19, BB41d14, BB48h15, BB82a11, BB86e10 and BB99n12 were further characterized. Both ends of the BACs were sequenced with the standard primers T3 and T7. The sequences were blasted using the programs BLASTN 2.2.1 and/or BLASTX 2.2.1 of the NCBI(GenBank). The results of the blast search are presented in Table 3.3.

Table 3.3 Sequence similarity of T3 and T7 ends of osmotin-like BACs.

BAC end	Blasted sequence (bp)	Species	Accesion number with highest similarity	E Value	Score	Identity %	Function assigment based on sequence similarity
BB18h19T3	1000 *	*A. thaliana*	NP_194070	$3e^{-09}$	43	61	putative protein
BB18h19T7	960 ***	-	-	-	-	-	-
BB41d14T3	600 **	*S. commersonii* *L. esculentum*	X72927 AF093743	e^{-136} $5e^{-20}$	490 105	94 84	genes for osmotin-like protein
BB41d14T7	616 ***	-	-	-	-	-	-
BB48h15T3	1528 *	*A. thaliana*	AAL32696	e^{-108}	393	67	ATP-dependent RNA helicase
BB48h15T7	779 *	*A. thaliana*	NP_194054	$6e^{-10}$	43	37	putative protein
BB82a11T3	635 ***	-	-	-	-	-	-
BB82a11T7	1521 ***	-	-	-	-	-	-
BB86e10T3	1351 ***	-	-	-	-	-	-
BB86e10T7	972 ***	-	-	-	-	-	-
BB99n12T3	1128 *	*O. sativa*	T50649	$4e^{-07}$	57	71	elicitor responsive gene
BB99n12T7	710 ***	-	-	-	-	-	-

* Homologies generated by BLASTX.
** Homologies generated by BLASTN.
*** Not significative homology founded neither with BLASTN or BLASTX.

The sequences were used to designed primer pairs from every BAC end[10]. PCR amplification was carried out with the different primer pairs using as template all the osmotin-like BACs. Presence of a PCR product in a different template indicated a possible overlap. After amplification with all the primer combinations, the orientation of the overlapping clones could be inferred. Moreover, the PCR products were sequence and aligned to confirm the overlaps and to show the parental line from which the BACs were derived. With this information two small contigs were assembled (Figure 3.22). BACs BB41d14, BB48h15, BB18h19 and BB99n12 form one group (Contig A). BACs BB82a11 and 86e10 form a second group (Contig B). Contig A expands around 100 kb and Contig B expands around 120 kb.

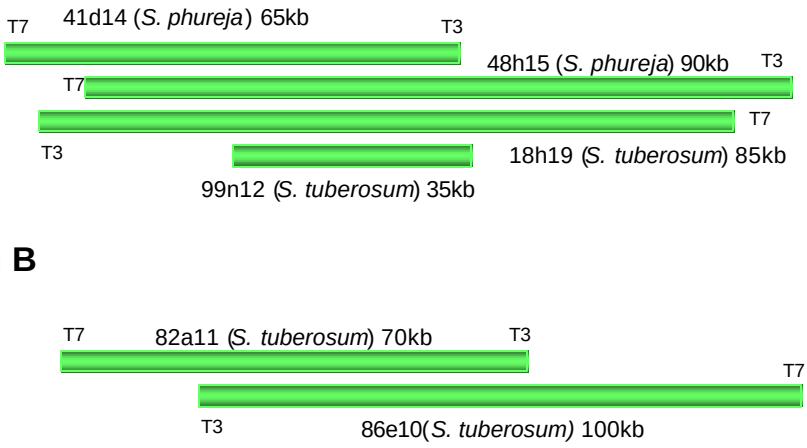


Figure 3.22 Schematic representation of overlapping BACs with homology to osmotin-like genes. The overlapping clones were group in two contigs: Contig A and Contig B. The orientation of the BACs in the contigs could be inferred for all of them but one, BAC 99n12.

3.3.1.5 Genetic mapping of the osmotin contigs

The primers designed from the osmotin-like BACs ends were used to develop genetic markers (SCAR, CAPS and SSCP) (as indicated in Material and Methods) in

[10] The sequences of the primers are listed in Appendix 7.1.

order to map **Contig A** and **Contig B**. Table 3.4 shows the primer combinations and restriction enzyme generating segregating fragments in the PD and/or F1840 populations that were used for mapping.

Table 3.4 Genetic mapping of Contigs A and B.

Contig	Primer pair	Restriction enzyme	Mapping in PD population	Mapping in F1840 population
A	18h19T3for 18h19T3rev	*Alu*I, *Mbo*I	- -	LG VIII **CAPS A**
A	48h15T3for 48h15T3rev	*Mse*I	-	LG VIII **SSCP A**
B	82a11T7for 82a11T7rev	-	-	LG VI **SCAR**
B	86e10T3for 86e10T3rev	*Xho*I	-	LG IV and VI
B	68311 68321	*Alu*I	LG XI	LG XI

Using the F1840 population, the **Contig A** was mapped to the same position as the *osmotin* RFLP marker. **CAPS A** and **SSCP A** were located at the ends of the Contig A. One recombinant was present between the two markers. This allowed the orientation of the cluster on the long arm of chromosome VIII and also allowed to calculate the relation between genetic and physical distance for this genomic region (1cM/100kb). On the other hand, mapping of Contig B was ambiguous due to that markers derived from the ends of BB82a11 and BB86e10 mapped to different loci in LG IV and VI. Using for mapping a subclone of BAC BB82a11 (68311), Contig B was mapped to LG XI. The map position of Contig B was confirmed by mapping the same marker in the PD population (Ghislain M., personal communication). The map positions are indicated in figure 3.23.

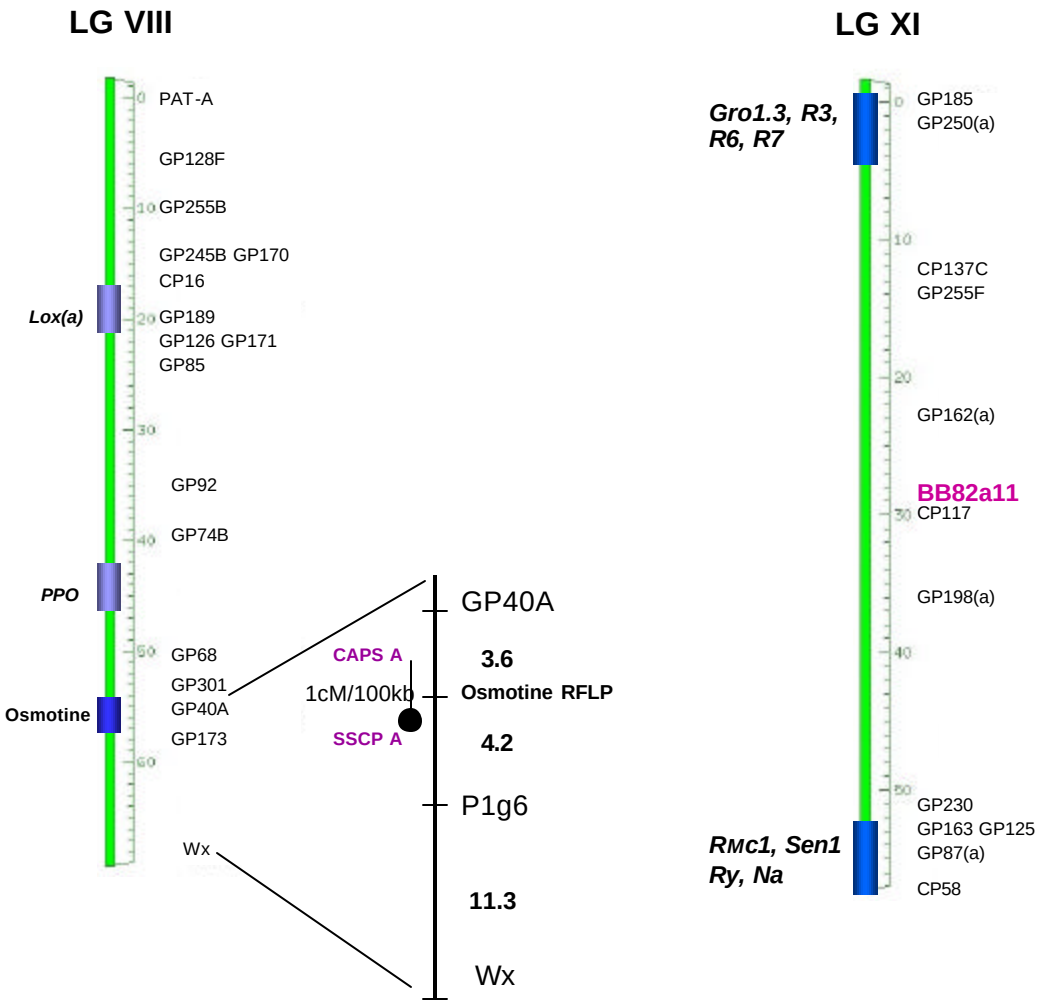


Figure 3.23 Map positions of c*smotin*- like genes in potato. On the left site is the schematic representation of LG VIII. The lilac boxes indicate two defense related genes: *Lox* (lipoxigenase) and *PPO* (polyphenol oxidase) and the blue box indicates the genetic position of the *osmotin* cluster. The map position of some RFLP markers (GP: genomic potato, CP: cDNA potato and TG: tomato genomic) and genes PAT-A, patatin gene and Wx, waxi locus is also indicated. Next to LG VIII a more detailed representation of the osmotin cluster delimited by two genetic markers: **CAPS A** and **SSCP A**. On the right, LG XI and the position of the clone BB82a11 containing an osmotin-like gene. The blue boxes indicate resistance genes. At the top, resistance genes against nematodes (*Gro1.3*) and *P. infestans (R3, R6 and R7.* At the bottom, resistance genes against bacteria (*Sen1*), virus (*Ry* and *Na*) and nematodes (*RMc1*). The orientation of the chromosomes was based on the publication of Dong et al. (2000). The genetic distances are indicated in cM.

3.3.1.6 Copy number of osmotin- like genes in potato

In order to compare the number of *osmotin*-like genes in potato with the osmotin-like genes present in the BACs, a Southern gel blot was done with DNA from linePD59 and BACs BB18h19, BB41d14, BB48h15, BB82a11, BB86e10 and BB99n12. The DNAs were digested with *Hin*dIII and the hybridization was carried

out with the osmotin probe (Figure 13.24). The Southern gel blot showed that osmotin-like genes are a small gene family in potato with at least 5 copies. A similar number has been reported for *S. commersonii* by Zhu et al. (1995a). Almost all the copies are clustered in the Contig A. BACs BB18h19 and BB48h15 are the more informative since they carry most of the copies and they represent homologous chromosomes from *S. tuberosum* and *S. phureja*, respectively. BACs BB82a11 and BB86e10 (Contig B) are carrying a single copy of the gene. Together, these BACs are carrying most of the copies of *osmotin*-like genes of potato.

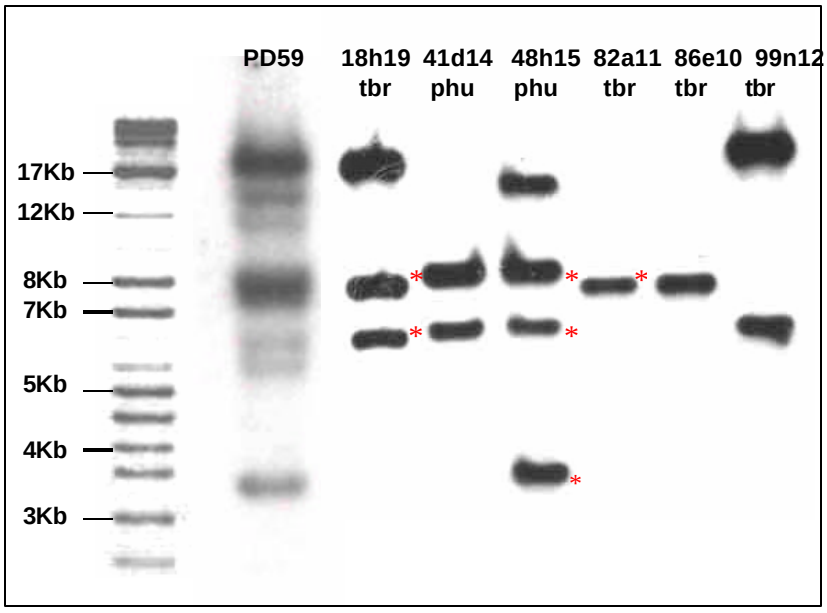


Figure 3.24 Autoradiogram of a genomic Southern gel blot containing DNA from PD59, BB18h19, BB41d14, BB48h15, BB82a11, BB86e10 and BB99n12. The parental origin of the BACs is indicated as tbr (*S. tuberosum*) or phu (*S. phureja*). The DNA was digested with *Hin*dIII and the blot was hybridized with osmotin. At the left the HMW marker indicate the size of the fragments. The * indicates fragments that were subcloned into pBluescript.

3.3.1.7 Subcloning of osmotin-like BACs

Some of the hybridizing fragments were subcloned in order to know more about the genomic structure of the osmotin genes in potato. BB18h19, BB48h15 and BB82a11 were chosen as representative BACs. These BACs were digested with *Hin*dIII, the fragments were purified, ligated to pBluescript and transformed into *E. coli* DH10B cells. The screening of positive clones was done by PCR using the osmotin primers. Two fragments were subcloned from BB18h19, one of 7 kb and

another one of approximately 6 kb (subclones **17** and **137** respectively). Three fragments were subcloned from BB48h15, corresponding to the fragments of 8, 6 and 3.4 kb (subclones **517**, **57** and **534** respectively). From BB82a11 (Contig B) a fragment of 6.8kb was subcloned (subclone **68**). All the subclones were sequenced by primer walking. The sequences were analyzed using Jellyfish Version 1.3 (Bioware) to find putative ORFs. The ORFs were blasted using the NCBI(Genbank) database and searched for homology to osmotin-like genes. The sequences obtained and the results of the blast search are summarized in Table 3.5[11].

Table 3.5 Subclones from osmotin-like BACs.

Subclone (total lenght)	**Subclone end (read sequence in bp)**	**ORF (amino acid length)**	**Identity at amino acid level using BLASTX**
17 (7 kb)	17T3 1009	-	-
	17T7 886	-	-
137 (5674 bp)	137T3 3333	-	-
	137T7 3580	137d' 247	99% X67244 osmotin *S. commersonii*
		137d'1 344	-
517 (8 kb)	517T3 2723	517b 243	90% AJ277064 *PR-5 L. sculentum*
		517b1 335	-
	517T7 3225	517c 247	99% X72927 osmotin *S. commersonii*
57 (5829 bp)	57T3 3300	-	-
	57T7 3445	57d 247	100% X67244 osmotin *S. commersonii*
		57d1 302	-
534 (3433 bp)	534T3 2394	534a 243	91% X72927 osmotin *S. commersonii*
	534T7 1744	-	-
68 (6823bp)	68T3 3800	68e 220	84% X67121 osmotin *S. commersonii*
		68e1 314	-
	68T7 4117	-	-

[11] The sequences are in Appendix 7.3.

The analysis of the sequences 517T3 and 534T7 indicated that subclones 517 and 534 could be joined at these ends forming a single fragment. To confirm this idea, two primers were designed from 517 T3 and 534 T7 ends. Both of them were 200 bp apart from the ends in reverse orientation. Amplification on BAC BB48h15 and PD59 DNA gave a PCR product of ~400 bp. The PCR products were sequenced and aligned. The sequences were identical confirming that subclones 534 and 517 together form a fragment of 11.4 kb without a gap in between. Three ORFs homologous to osmotin- like genes were found in the fragment of 11.4 kb, they have been called copies **534a**, **517b** and **517c**. These copies are 243, 243 and 247 amino acids long, respectively. The sequence from clone 57 revealed a fourth ORF with homology to *osmotin*, this was called copy **57d** and is 247 amino acids long. The sequence from clone 137 also showed an ORF with homology to *PR-5* -like genes, this was called **137d'.** As shown in Table 3.4, **137d'** (*S. tuberosum*) and **57d** (*S. phureja*) are nearly identical to pA35, one of the osmotin genes of *S. commersonii* (X67244). These sequences might represent, therefore, orthologous genes (Figure 3.25). **57d** and **137d'** share 97.7% homology at the nucleotide level and they are 99.2% identical at amino acid level. Sequence comparison of 0.5 kb of the promoter region of **57d** and **137d'** showed 97.2% homology. And sequence comparison of 0.5 kb downstream the gene from **57d** and **137d'** showed 94.8% homology.

```
           1                                                          50
phu(57d)   MSHLTTCLVF FLLAFVTYTN ASGVFEVHNN CPYTVWAAAT PIGGGRRLER
comm(pA35) MSHLTTCLVF FLLAFVTYTN ASGVFEVHNN CPYTVWAAAT PIGGGRRLER
tbr(137d') MSHLTTCLVF FLLAFVTYTN ASGVFEVHNN CPYTVWAAAT PIGGGRRLER

           51                                                         100
phu(57d)   GQSWWFWAPP GTKMARIWGR TNCNFDGAGR GWCQTGDCGG VLECKGWGKP
comm(pA35) GQSWWFWAPP GTKMARIWGR TNCNFDGAGR GWCQTGDCGG VLECKGWGKP
tbr(137d') GQSWWFWAPP GTKMARIWGR TNCNFDGAGR GWCETGDCGG VLECKGWGKP

           101                                                        150
phu(57d)   PNTLAEYALN QFSNLDFWDI SVIDGFNIPM SFGPTNPGPG KCHPIQCVAN
comm(pA35) PNTLAEYALN QFSNLDFWDI SVIDGFNIPM SFGPTNPGPG KCHPIQCVAN
tbr(137d'  PNTLAEYALN QFSNLDFWDI SVIDGFNIPM SFGPTNPGPG KCHPIQCVAN

           151                                                        200
phu(57d)   INGECPGSLR VPGGCNNPCT TFGGQQYCCT QGPCGPTDLS RFFKQRCPDA
comm(pA35) INGECPGSLR VPGGCNNPCT TFGGQQYCCT QGPCGPTDLS RFFKQRCPDA
tbr(137d') INGECPGSLR VPGGCNNPCT TFGGQQYCCT QGPCGPTDLS RFFKQRCPDA

           201                                                        250
phu(57)    YSYPQDDPTS TFTCQSWTTD YKVMFCPYGS THNETTNFPL EMPTSTLEVA*
comm(pA35) YSYPQDDPTS TFTCQSWTTD YKVMFCPYGS THNETTNFPL EMPTSTLEVA*
tbr(137d') YSYPQDDPTS TFTCQSWTTD YKVMFCPYGS THNETTNFPL EMPTSTLELA*
```

Figure 3.25 Amino acid sequence alignment from three putative orthologous osmotin-like genes. phu(57d) is the copy **57d** derived from *S.phureja*, comm(pA35) is an osmotin gene from *S. commersonii* and tbr(137d') is the copy **137d'** derived from *S. tuberosum*. 57d and pA35 are identical. The non-identical amino acids belonging to tbr are indicated in green boxes. The * indicates the stop codon of the ORF.

All **534a**, **517b**, **517c**, **57d** and **137d**' ORFs have the 16 cysteine residues that are highly conserved in the *thaumatin* family. The five copies have a longer C-terminus when compared to E22-like genes. According to their C-terminus sequence, theoretically they are vacuolar PR-5 proteins (Melchers et al. 1993).

The subclone 68 had a smaller ORF of **220** amino acids with 84% homology to osmotin, this was called **68e**. Compared to **534a**, **517b**, **517c, 57d** and **137d'** this copy has an internal deletion of 10 amino acids**.** The deletion involves the cysteine residues 11 and 12**.** The N-terminus is longer by four amino acids and the C-terminus is shorter. Sequence alignment of **534a**, **517b**, **517c**, **57d** and **68e** is shown in figure 3.26.

```
      1                                                          50
517c  ~~~~MSHLT. TCLVFFLLAF VTYTYASGVF EVHNNCPYTV WAAAVPVGGG
 57d  ~~~~MSHLT. TCLVFFLLAF VTYTNASGVF EVHNNCPYTV WAAATPIGGG
517b  ~~~~MGYL.R SSFIFSLLAF VTYTYAA.TI EVRNNCPYTV WAASTPIGGG
534a  ~~~~MGYL.R SSFVFSLLAF VTFTYAV.TI EVRNNCPYIV WAASTPIGGG
 68e  MSTNMDYLIR SSSVFFLLSF VTYTYST.SF EVRNNCPYTI WAASTPIGGG

      51                                                        100
517c  RRLERGQSWW FWAPPGTKMA RIWGRTNCNF DGAGRGGCQT GDCGGVLDCK
 57d  RRLERGQSWW FWAPPGTKMA RIWGRTNCNF DGAGRGWCQT GDCGGVLECK
517b  RRLNKGQTWV INAPRGTKMA RIWGRTGCNF NAAGRGSCQT GDCGGVLQCT
534a  RRLDRGQTWV IDAPRGTTMA RIWGRTGCNF DGAGRGSCMT GDCGGVLQCT
 68e  RRLDSGQTWV IDAPRGTKMA RIWGRTNCNF DGAGRGSCQT GDCGGVLQCT

      101                                                       150
517c  GWGKPPNTLA EYALNQFGNL DFWDISVIDG FNIPMSFGPT KPGPGKCHGI
 57d  GWGKPPNTLA EYALNQFSNL DFWDISVIDG FNIPMSFGPT NPGPGKCHPI
517b  GWGKPPNTLA EYALDQFSNL DFWDISLVDG FNIPMTFAPT KPSAGKCHAI
534a  GWGKPPNTLA EYALDQFSNL DFWDISLVDG FNIPMTFAPT NPSGGKCHAI
 68e  GWGKSPNTLA EYALNQFSNL DFWDISLIEG FNIPMTFAPT NPSGGKCHAV

      151                                                       200
517c  QCTANINGEC PGSLRVPGGC NNPCTTFGGQ QYCCNHGPCG PTDLSRFFKQ
 57d  QCVANINGEC PGSLRVPGGC NNPCTTFGGQ QYCCTQGPCG PTDLSRFFKQ
517b  HCTANINGEC PRALKVPGGC NNPCTTFGGQ QYCCTQGPCG PTELSKFFKK
534a  QCTANINGEC PGQLKVPGGC NNPCTTFGGQ QYCCTQGPCG PTFFSNFFKQ
 68e  QCTANINGEC PNQLKVSGGC NNPCTI.... ......GQCG PNDLSRFFKQ

      201                                                       250
517c  RCPDAYSYPQ DDPTSTFTCQ SWTTDYKIMF CPYGSTHNET TNFPLEMPTS
 57d  RCPDAYSYPQ DDPTSTFTCQ SWTTDYKVMF CPYGSTHNET TNFPLEMPTS
517b  RCPDAYSYPQ DDPTSTFTCP GGSTNYKVVF CPNGVA...D PNFPLEMPAS
534a  RCPDAYSYPQ DDPTSTFTCP SGSTNYRVVF CPNGVA...G PNFPLEMRAS
 68e  RCPDAYSYPQ DDPTSLFACP SDSTNYRVVF C*~~~~~~~~ ~~~~~~~~~~

      251
517c  ILEVA*~
 57d  TLEVA*~
517b  TDEVAK*
534a  TDEVAK*
 68e  ~~~~~~~
```

Figure 3.26 Amino acid sequence alignment of osmotin-like genes. 534a, 517b, 517c and 57d are derived from *S. phureja* and they mapped to LG VIII. 68e is derived from *S. tuberosum* and it mapped to LG XI. The 16 cysteine residues highly conserved in the *PR-5* gene family are indicated in yellow boxes.

The degree of identity of **534a**, **517b**, **517c**, **57d** and **68e** is indicated in the Table 3.6.

Table 3.6 Identity between 534a, 517b, 517c, 57d and 68 at amino acid level.

	534a	517b	517c	57d	68
534a	-	90.3	69.8	75.7	73.4
517b		-	73.4	75	70.6
517c			-	94.8	64.6
57d				-	65.8
68e					-

Copies **534a** and **517b** have the same length and orientation. They are 90.3% identical at the amino acid level. Sequence alignment of 0.5 kb upstream the ORFs show 65% homology and the alignment of 0.5kb downstream the ORFs show 49.5% homology. Copy **57d** has the same length as copy **517c** and they are 94.8% identical at amino acid level. 0.5 kb sequence upstream the ORFs shows 58% homology and 0.5 kb downstream regions of the copies share 59.1% homology. All the osmotin-like copies analyzed did not present introns as is expected from *thaumatin*-like genes and all but one (68e) have the 16 cysteine residues highly conserved in this gene family. A schematic representation of the organization of the osmotin copies is depicted in figure 3.27.

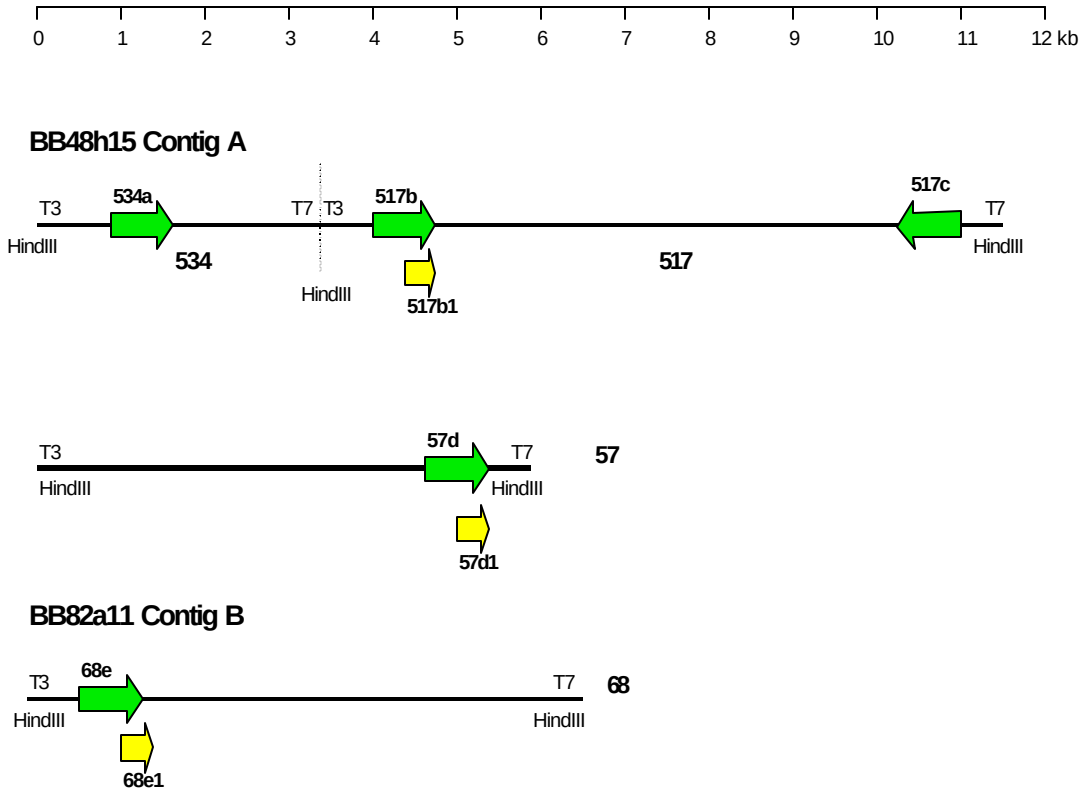


Figure 3.27 Osmotin-like subclones derived from BACs BB48h15 and BB82a11. The scale indicates the size of the subclones and ORFs. Subclones 534 and 517 form a fragment of 11.4 kb, the position where they joined is indicated by the dashed line. Green solid arrows indicated the orientation and extension of the ORFs with homology to osmotin-like genes. Yellow solid arrows indicate the orientation and extension of additional ORFs, these smaller ORFs do not show any homology to known genes or proteins. T3 and T7 indicate the position of these primers in the subclones.

Smaller additional ORFs were found in some subclones that did not show any homology to known genes (Table 3.5).

3.3.1.8 Promoter structure of *osmotin*-like genes

Promoter sequences of *osmotin*-like genes were searched for *cis*-acting elements that may function in the activation of *PR-5* genes. 300 bp sequence upstream of the start codon of copies **534a**, **517b**, **517c**, **57d** and **137d'** were compared to the promoter of tobacco osmotin (Figure 3.28). Several *cis* elements have been correlated to the promoter activity in a study of the structure of the

osmotin promoter (Liu et al., 1995). The results of the comparison showed that TATA and CAAT boxes were highly conserved in all osmotin-like genes. A 6-bp 5'-GCCGCC-3' sequence 5bp away from the TATA box was also present in all copies. This motif is present in certain ethylene-inducible PR genes (Zhu et al., 1995a). **137d'** and **57d** have a second 5'-GCCGCC-3' sequence in their promoter region. The osmotin gene also present the tandem repeat and a third GCCGCC sequence in reverse orientation (GGCGGC). This reverse sequence was found in **517c**, **517b** and **534a**. Since none of these motifs was recognized in the promoter of the osmotin-like of subclone **68e**, this copy was not included in the figure 3.28.

```
   137d' -264  tctGCGGCttattatctcaa  -245
    57d  -264  tctGCGaCttattatctcaa  -245
   517c  -259  ttGGCGGCtttttgtttcaa  -240
   517b  -194  taGGCGGCttctta.tatat  -175
   534a  -187  taGGCGGCtcttat.acatt  -168
osmotin  -163  taGGCGGCtcttatgtttaa  -144

   137d' -170  taGCCGCCtagatctatcttatt  -148
    57d  -170  taGCCGCCtagatctatcttatt  -148
   517c  -164  tacctttataactaaagtggcag  -142
   517b  -163  cctttaccaacgcatattttaat  -141
   534a  -156  cgtttaccaaatcatattgtaat  -134
osmotin  -143  gcGCCGCCtccatctttgccaaa  -121

   137d' -98  tattgtCCATattgcaaaaaGCCGCCac  -70
    57d  -98  tattgtCCATattgcaaaaaGCCGCCac  -70
   517c  -89  taatgtCCATagtgaaaaaaGCCGCCat  -61
   517b  -88  aaaaatCCATagtgccaaaaGCCGCCat  -60
   534a  -84  taatatCCATagtgccaaaaGCCGCCat  -56
osmotin  -67  taatatCCATagtacgaaaaGCCGCCat  -39

   137d' -68  cccTATATAaaccccctactcaattgtgt  -40
    57d  -68  cccTATATAaaccccctactcaattgtgt  -40
   517c  -59  cacTATATAaaccccctagctcgatatgc  -31
   517b  -58  cccTATATA.aagtaccaaaaaataatt  -32
   534a  -54  cccTATATAtaaagtaccaaaaataatt  -26
osmotin  -37  cccTATATAaa....ccactaaacaatt  -13
```

Figure 3. 28 Comparison of the conserved 5' flanking sequences of a tobacco osmotin gene with potato osmotin copies **534a**, **517b**, **517c**, **57d** and **137d**'. Putative CAAT and TATA boxes are in upper cases and boldface. The 6-bp 5'-GCCGCC-3' putative ethylene responsive motif in the top strand are shaded in light blue. The gray shaded 5'-GGCGGC-3' is complementary to the 5'-GCCGCC-3' motif shaded in yellow. The first nucleotide position of osmotin corresponds to the start of the translation. The first positions indicated for the potato osmotins correspond to the start codon because the transcription start sites have not yet been determined.

3.3.2 Transcription of PR-5 genes in potato

RNA isolated from leaves of greenhouse grown PD59 plants was used to synthesize complementary DNA strands. A RT-PCR experiment was carried out using primers corresponding to the ORF from osmotin-like genes (*osmotin* primers) and E22-like genes (E22f and E22r primers). In both cases, very clear amplification products were obtained (Figure 3.29). There was no difference in size between PCR products derived from genomic and cDNA as is expected from genes that do not have introns. As a control for genomic contamination of the cDNA, a PCR was carried out using 319pr primers. The primers amplify part of the promoter of copy **319b**. No product was obtained, indicating that the cDNA was free from genomic DNA (Data not shown).

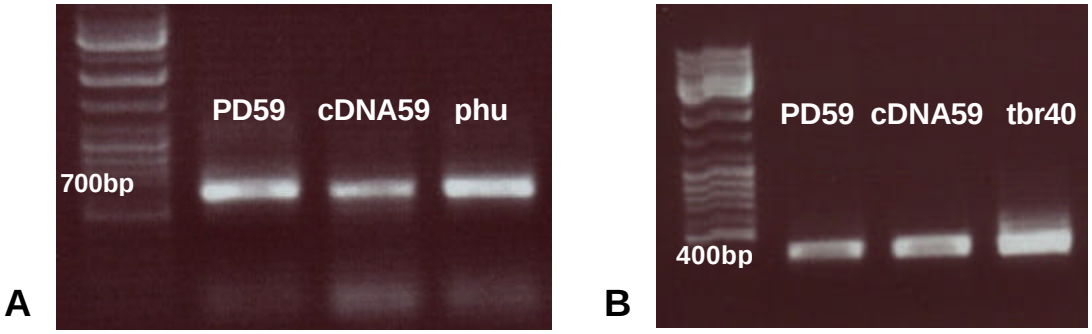


Figure 3.29 RT-PCR of *PR5*- like genes. **A.** PCR products from genomic DNA of line PD59, cDNA from PD59 and genomic DNA of phu (*S. phureja*) derived from osmotin primers. **B.** PCR products from genomic DNA of line PD59, cDNA from PD59 and genomic DNA from tbr (*S. tuberosum* line 40) derived from E22 primers.

The PCR products were sequenced. The sequence of cDNAPD59 amplified using *osmotin* primers showed overlapping signals (N) indicating that more than one copy was transcribed. It was not possible to identify the genes that are corresponding to the transcribed sequence. On the other hand, the sequence derived from E22 primers showed a single gene. The sequence corresponded to the copy **319b.**

A Northen gel blot experiment was carried out with leaves infected with *P. infestans*. PD59 leaves were detached from greenhouse plants and infected with

P.infestans race 4 (see Matherials and Methods). RNA was isolated from infected leaves that were collected after 18 and 24 hours. The RNA samples were run and blotted. The filter was firstly hybridized with the PCR product derived from E22 primers. As control, a second hybridization was carried out using a fragment of the potato actin gene. The results of the experiment are shown in figure 3.30A. There is no significant difference to be seen. There was no induction of the gene family. Individual members may still be induced. An accumulation of Protein C (the N-terminus sequence has been published and is identical to copy **319b**) has been reported in salicylic acid-induced leaves of potato (Pierpoint et al., 1990). The increment of protein was also observed in aging leaves. In tobacco, the proteins corresponding to genes E22 and E2 were accumulated after TMV infection (Pierpoint et al., 1992).

The infection progress on leaves of PD59, *S. tuberosum* cv. Desirée (susceptible) and a non-host resistant *S. nigrum* were followed after *P. infestans* race 4 inoculation. The signs of infection were observed after 7, 14 and 18 days. A typical compatible interaction was observed between PD59 and *P. infestans* race 4. The same response was observed in the positive control, Desirée (data not shown). After 14 days, PD59 and Desirée leaves were completely covered by mycelium. On the other hand, *S. nigrum* leaves did not show any macroscopic sign even after 18 days (Figure 3.30 B).

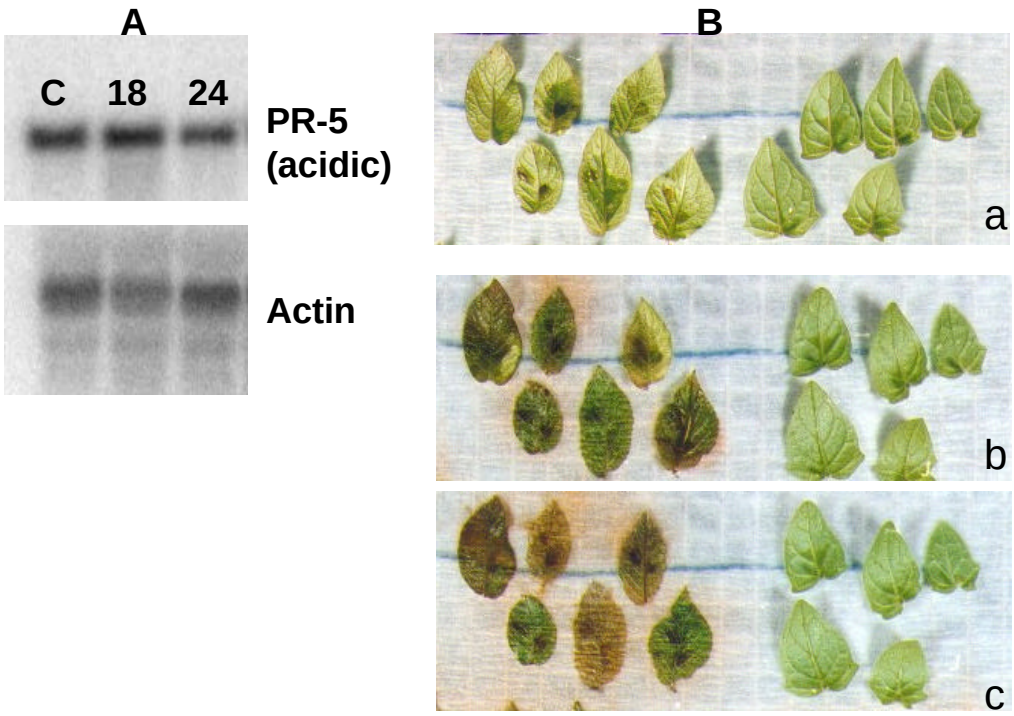


Figure 3.30 A. Intercellular *PR-5* expression under infection. At the top, *PR-5* expression after 18 and 24 hours after *P.infestans* inoculation. C indicates the control, RNA from uninfected leaves. **B.** Progress of *P. infestans* infection. In panels a, b and c: To the left PD59 and to the right *S. nigrum* showing macroscopic signs of infection after 7, 14 and 18 days respectively.

3.4 The *PR-5* gene family in *Solanaceae* species

3.4.1 Conservation of E22-like and osmotin-like genes among *Solanaceae*

Potato species are grouped in the Section Petota. This Section is divided in two subsections: *Estolonifera* and *Potatoe*. Subsection *Estolonifera* groups species related to potato to some extent but they do no bear stolons or tubers. Subsection *Potatoe* contains the true potatoes (Hawkes 1991). The subsections are divided in series. DNA from different potato species (Table 3.7) was amplified using primers

from PR-5 genes from LG VIII (osmotin primers) and LG XII (E22 primers). Additionally, primers amplifying anonymous sequences physically close to *PR-5* genes were also tested. The results are shown in figure 3.31.

Table 3.7 Systematic classification of potato species (based on Hawkes 1991). The * indicates that resistance to *P. infestans* has been described for those species.

Subsection	**Series**	**Species**
ESTOLONIFERA		*S. etuberosum*
POTATOE	*I MORELLIFORMIA*	*S. morelliforme*
	VII LIGNICAULIA	*S. lignicaule*
	IX YUNGASENSA	*S. chacoense*
	X MEGISTACROLOBA	*S. megistacrolobum*
	XV MAGLIA	*S. maglia*
	XVI TUBEROSA	Wild species: *S. verrucosum *, S. leptophyes, S. canasense,* *S. sparsipilum, S.bukasovii, S. kurtzianum,* *S. microdontum *, S. hondelmannii, S. oplocense* *S. neorossii, S. gourlayi, S. alandiae, S. vernei *,* *S. berthaultii *, S. spegazzinii, S. brevicaule* Cultivated species: *S. phureja *, S. andigena *, S. tuberosum,* *S. stenotomum, S. acaule, S. stoloniferum,* *S. demissum **

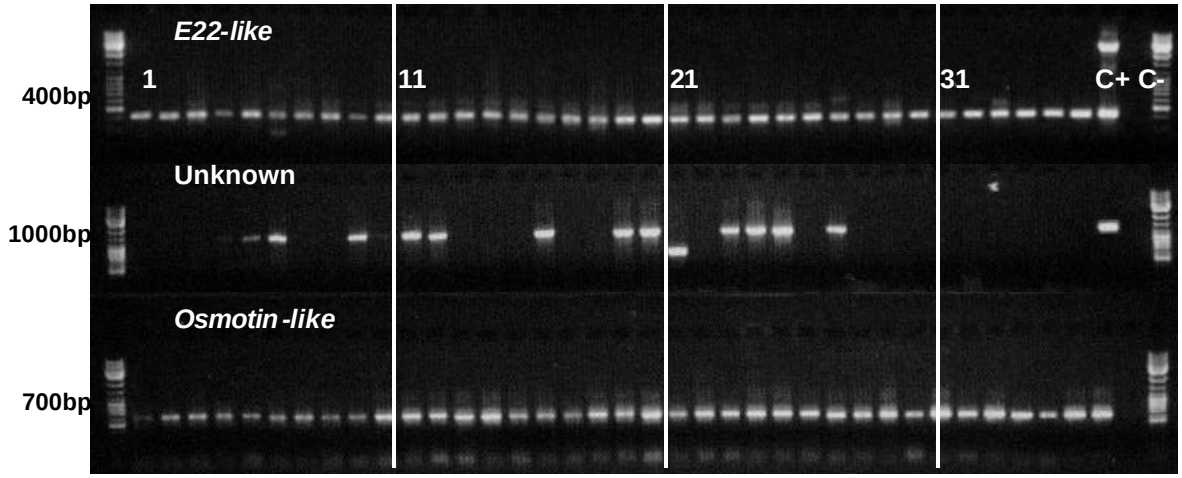


Figure 3.31 Conservation of *PR-5*-like genes in *Solanaceae*. In the three panels the order from 1 to 39: *S. etuberosum* (1), *S. morelliforme* (2), *S. lignicaule* (3), *S. chacoense* (4), *S. megistracolobum* (5), *S. maglia* (6), *S. verrucosum* (7), *S. leptophyes* (8), *S. canasense* (9), *S. sparsipilum* (10), *S.bukasovii* (11), *S. kurtzianum* (12), *S. microdontum* (13), *S. hondelmannii* (14), *S. oplocense* (15), *S. neorossii* (16), *S. gourlayi* (17), *S. alandiae* (18), *S. vernei* (19), *S. berthaultii* (20), *S. spegazzinii* (21), *S. brevicaule* (22), *S. phureja* (23), *S. andigena* (24), *S. tuberosum* (25), *S. stenotomum* (26), *S. acaule* (27), *S. stoloniferum* (28), *S. demissum* (29), *S. dulcamara* (30), *S. nigrum* (31), *tomato* (32), *Lycorpersicon sp.* (33), *C. annuum* (34), *N. tabacum* (35), *S. melongena* (36), the positive control PD59DNA (37) and the negative control (38). In the first panel, DNA amplified with primers corresponding to E22-like genes. In the middle panel, DNA amplified with primers corresponding to BB43d14 T3 end, an anonymous sequence. Bottom panel DNA, amplified with primers corresponding to *osmotin*-like genes.

In all potato species and relatives, PCR products corresponding to *PR-5* genes (E22 and osmotin) were amplified. The sizes of the products were identical to the positive control for E22 and osmotin-like genes. In the reactions were also included DNA from *S. dulcamara*, *S. nigrum*, *Lycopersicon sp*. and *Nicotiana tabacum* and other *Solanaceae* crop species like *Capsicum annum* (pepper), *Lycopersicon esculentum* (tomato) and *Solanum melongena* (eggplant). Using primers that amplify anonymous sequences from the BAC BB43d14, the amplification products were absent in 1/2 of the potato samples and no product was obtained from *S. dulcamara*, *S. nigrum* and the other *Solanaceae sp.* used. Those sequences seem to

be less conserved in potato species. Other primer combinations derived from 48h15T3 (with homology to an ATP-dependent RNA helicase from *A. thaliana*), 48h15T7 (with homology to a putative protein from *A. thaliana*) and 86e10T3 (no homology found) gave amplification products in all potato species (data not shown). From the primers that were tested no marker was found that could indicate correlation with the resistance to *P. infestans* present in some potato species.

The PCR products corresponding to osmotin and E22-like genes were sequenced. The sequences from o*smotin* showed overlapping signals indicating the presence of several copies in the genome of most *Solanaceae* species as it occurs in *S. commersonii*, *S. tuberosum* and *S. phureja*. The *E22*-like sequences showed a lower number of overlaps. E22-like sequences with clear peaks were aligned (Figure 3.32). *Solanum* species show in most of the cases the presence of two copies highly homologous to **319b** and **321c**. The predominant products seem to be the orthologous sequences of **319b.** With few exceptions, the *Solanum* species used were diploid. One triploid (*S. maglia*) and three tetraploids (*S. acaule*, *S. tuberosum* subsp. *andigena*, *S. tuberosum* subsp. *tuberosum*, and *S. tuberosum* cv. Desirée) did not seem to have a higher number of different alleles with respect to the diploids. On the other hand, the sequences of E22-like genes in pepper, eggplant, tomato, *S. nigrum*, *S. dulcamara* and *N. tabacum* showed a single sequence in the electropherograms. Pepper and tobacco E22-like genes showed higher homology to copy **321c** than to copy **319b**.

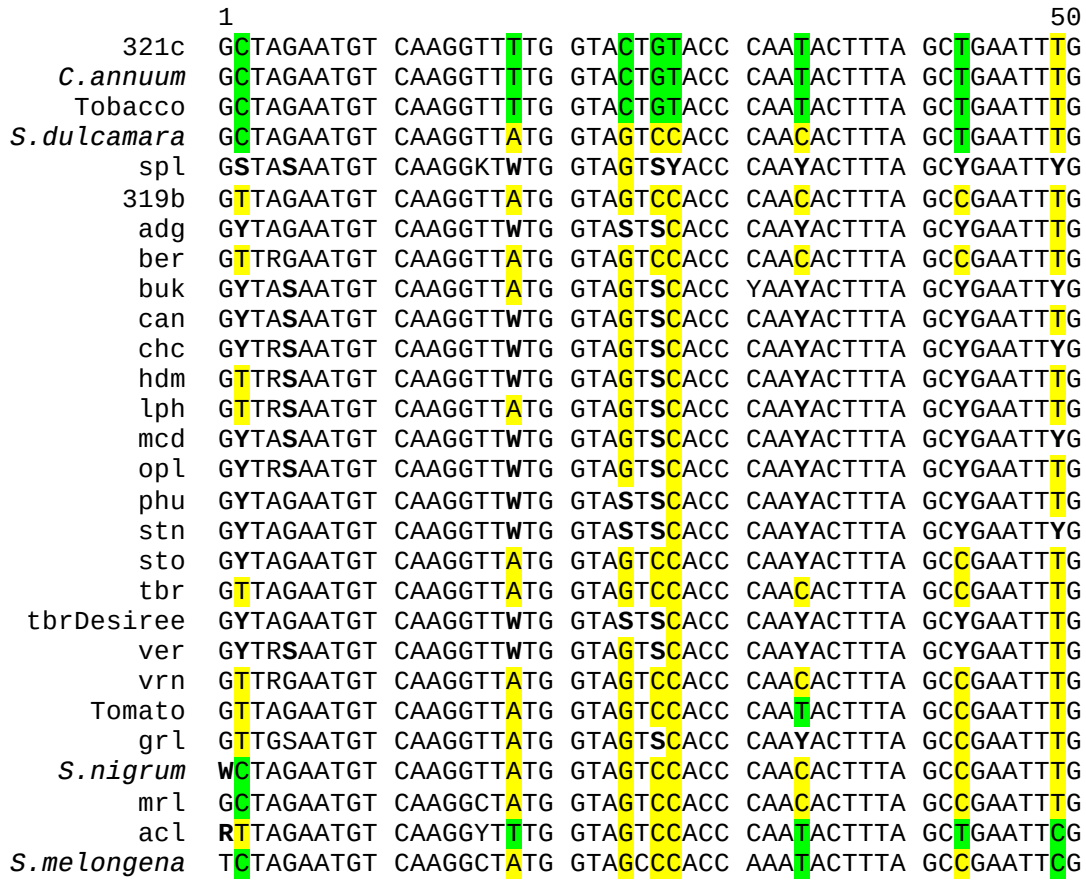


Figure 3.32 Polymorphisms of E22- like genes in *Solanaceae*. Partial sequences of PCR products derived from E22f and E22r primers. Overlapping signals occur mostly in positions where copies **319b** and **321c** are polymorphic. The single letter code has been used when two nucleotides are present, where Y = C or T, S = C or G, R = A or G, K = G or T and W = A or T. Some species showed to have a single sequence. In these cases, the most frequent bases in all species analyzed are indicated in yellow and the most infrequent ones are indicated in green. The nomenclature is the following: spl (*S.sparsipilum*l), adg (*S.andigena*), ber (*S.berthaultii*), buk (*S.bukasovii*), can *(S.canasense),* chc (*S.chacoense*), hdm (*S.hondelmanii*), lph (*S.leptophyes*), mcd (*S.microdontum*), opl (*S.oplacense*), phu (*S.phureja),* stn *(S.stenotomum*), sto (*S. stoloniferum)*, tbr (*S. tuberosum)*, ver (*S.verrucosum*), vrn (*S. vernei),* grl (*S. gourlayi*), mrl (*S.molleriforme*l) and acl (*S. acaule).*

3.4.2 Copy number of *PR-5* genes in *Solanaceae*

In order to compare the gene number of osmotin-like genes among *Solanaceae*, a Southern blot was done with DNA from different members of the family. The results are shown in figure 3.33. Osmotin sequences are conserved (more than 60% homology) in *Solanaceae* and they cross-hybridized. Low stringency conditions (1X SSP, 0.1% SDS) were used in the Southern blot washes. Compared to potato, the number of osmotin copies seems to be lower in *Lycopersicon sp*. and S. *melongena*.

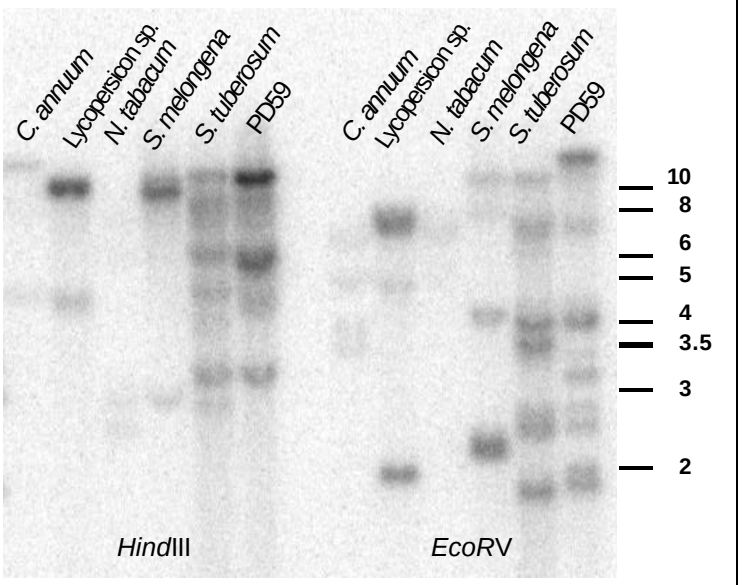


Figure 3.33 Autoradiogram of a genomic blot containing DNA from different *Solanaceae* species and PD59. The DNA was digested with *Hin*dIII and *Eco*RV and the blot was hybridized with osmotin. The relative sizes are indicated in kb.

The same blot was hybridized with a PCR probe derived from E22 primers. Cross hybridization was observed in all species. The number of loci of this class of gene seems to be lower in *C. annuum*, *Lycopersicon sp*., *N. tabacum* and *S. melongena* than in potato (Figure 3.34).

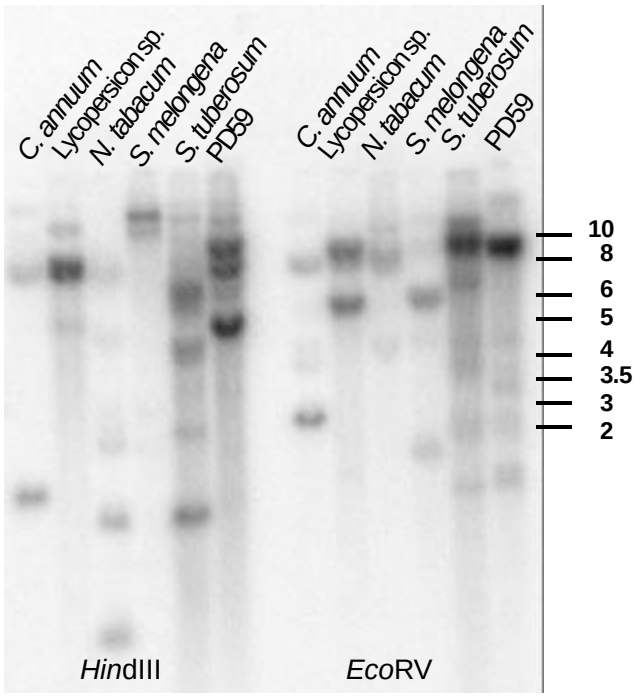


Figure 3.34 Autoradiogram of a genomic blot containing DNA from different *Solanaceae* species and PD59. The DNA was digested with *Hind*III and *Eco*RV and the blot was hybridized with E22-like sequence. The relative sizes are indicated in kb.

The PCR sequences of E22 -like from *C. annum*, *N. tabacum*, *S. melongena*, *S. dulcamara*, *L. esculentum* and *S. tuberosum* were translated into the frame with homology to PR-5 proteins. The amino acid sequences were aligned and a cladogram was generated with them (Figure 3.35 A and B). The cladogram shows the relations between E22-like genes of *Solanaceae* species.

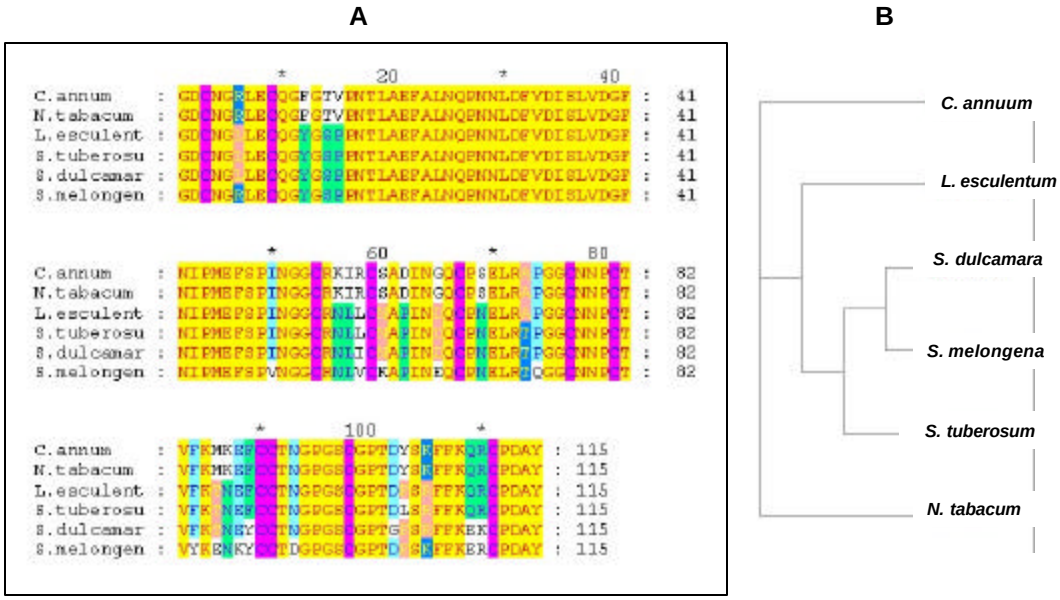


Figure 3.35 **A.** Amino acid sequence alignment of translated PCR products with homology to E22-like genes. The degree of identity is 78.2%. **B**. Gene tree generated with E22-like sequences.

3.4.3 Phylogenetic relationships of *PR-5* genes among *Solanaceae*

There are acidic, neutral, and basic PR-5 proteins. All three have been identified in tobacco (Koiwa et al. 1994). The sequences of tobacco PR-5a or PR-S (acidic), PR-5c (basic) and PR-5d (neutral) were compared to PR-5 *Solanaceae* sequences from the Genebank and to the potato genes from this study (**534a**, **517b**, **517c**, **57d**, **137d'**, **68e**, **319b**, **21a** and **321c**). A multi sequence alignment was done using the Pileup program from GCG. The alignment was used to construct a distance matrix (Kimura) and generate a phylogram by the neighbor joining method (Figure 3.36). The analysis grouped the sequences in three main branches that cluster acidic, neutral and basic genes. Members of the three classes were found among the potato genes. Analogs could also be identified. As it was proposed before, *S. phureja* (57d), *S. tuberosum* 137d' and *S. commersonii* (pA35) seem to be orthologous genes. *S. phureja* (517b) and *S. commersonii* (pOSML81) form another pair of analogs and probably they are orthogous genes.

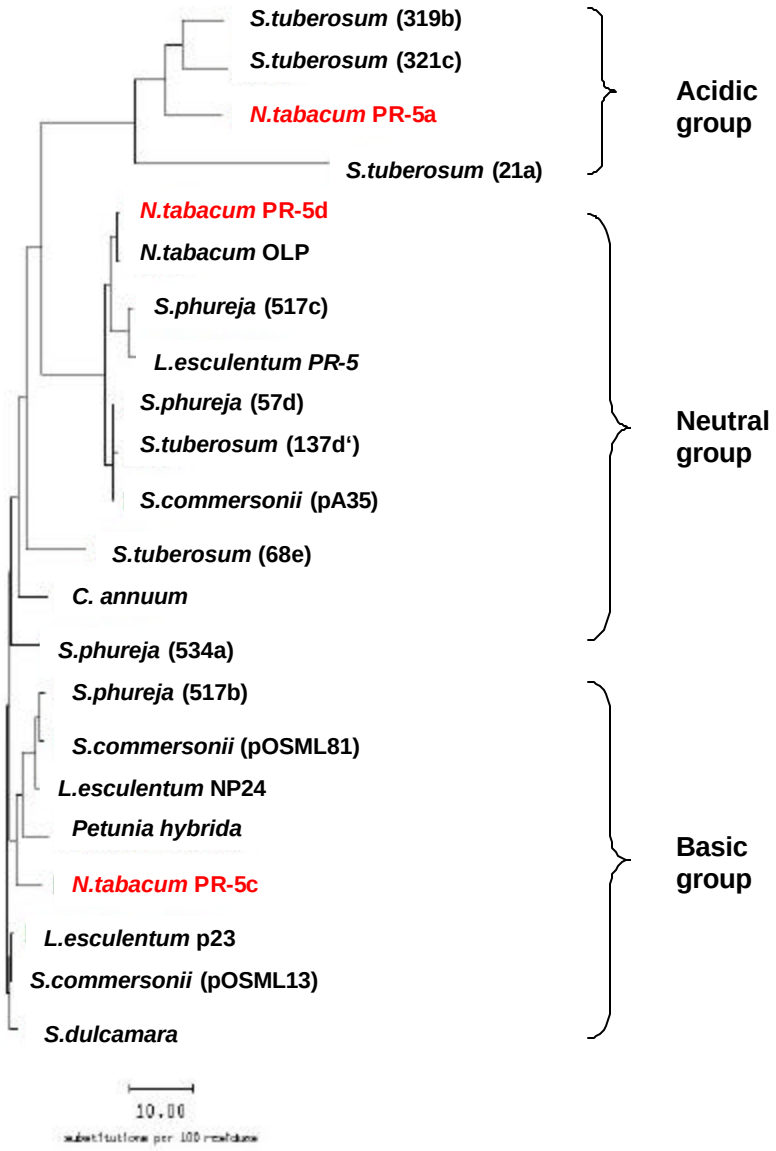


Figure 3.36 Phylogram of *S. tuberosum* (319b), *S. tuberosum* (321c), *N. tabacum* **PR-5a**, *S. tuberosum* (21a), *S. phureja* 517c gene, *L. esculentum* PR-5 (AJ277064), *S. phureja* (57d), *S.tuberosum* (137d'), *S. commersonii* pA35 (X67244), *N. tabacum* **PR-5d** (D76437), *N. tabacum* osmotin-like protein (M64081), *Capsicum annuum* (AF297646), *S. phureja* (517b), *S. commersonii* pOSML81 (X72927), *L. esculentum* **NP24** (AF093743), *N. tabacum* **PR-5c** (X6570), *Petunia hybrida* (AF376058), *S. phureja* (534a), *S. tuberosum* (68e), *S.dulcamara* cryoprotective osmotin-like protein (AY007309), S. *commersonii* pOSML13 (X72928) and *L.esculentum* p23 (X70787). Tobacco PR-5a, PR-5d and PR-5c are in red letters to indicate they are the representative members of acidic, neutral and basic PR-5 proteins, respectively.

3.4.4 In-silico mapping of E22- like genes in *Solanaceae*

E22-like genes are clustered on potato LG XII. The PCR sequences of E22-like genes from pepper and tomato indicated that *PR-5* genes of acidic nature are present in these two species too. The putative map positions of the sequences were identified in tomato and pepper by correlating genetic markers already mapped in the three species (Figure 3.37). The potato chromosome XII is syntenic to the same chromosome of tomato with the exception that a paracentric inversion involves the short arm of the chromosome (Bonierbale *et al*. 1988; Tanksley *et al*. 1992). Therefore, the tomato sequence was putatively positioned in tomato chromosome XII. On the other hand, the corresponding syntenic region in pepper is the long arm of chromosome IX. The pepper *PR-5* sequence was positioned there. The frames of the linkage groups were taken from the Solgenes database at Cornell (http://ars-genome.corne.edu/cgi-bin/WebAce/webace?db=solgenes). The orientation of the linkage groups was based on the publication of Livingstone *et al*. (1999) and Grube *et al*. (2000).

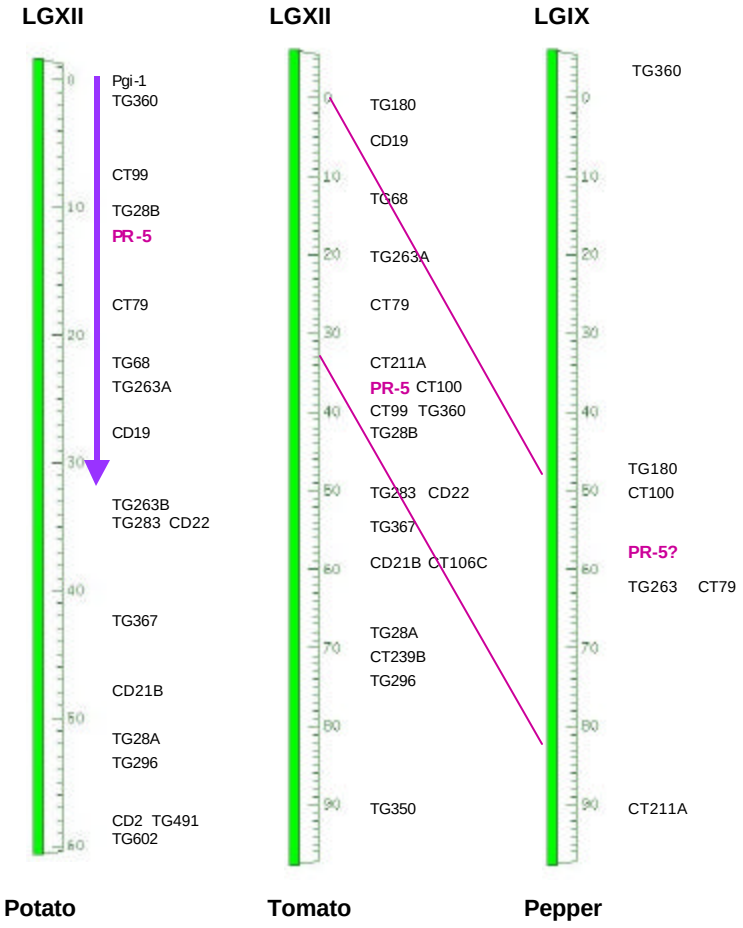


Figure 3.37 In silico mapping of *PR-5* like genes in tomato, potato and pepper. In the middle is the chromosome XII from tomato. Potato LG XII presents an inversion with relation to tomato LGXII that involves the short arm of the chromosome (pointing arrow) and tomato *PR-5* should be in the inverted portion of the chromosome. The short arm of tomato LGXII corresponds to the long arm of pepper LG IX and the *PR-5*- like gene has been located in the middle part of LGIX. *PR-5* indicated the relative position of the genes in the three species.

3.5 *PR-5* genes in other plant species

Due to their potential use in molecular breeding programs PR-5 -like genes have been cloned and described for almost all important crop plants. The invariant position of some amino acids that are shared by all PR-5 proteins, suggests that the genes coding for these proteins have descended from a common ancestral gene. A sub-group of *PR-5* genes from monocots are substantially smaller in size due to internal deletion of about one-fourth of the amino acids. However, they are also

induced by pathogens and show antifungal activities (Rebmann et al. 1991; Frendo et al. 1992). Koiwa et al. (1999) reported a common motif in antifungal thaumatin-like proteins. By doing a crystal structure study of tobacco PR-5d, zeamatin and thaumatin, they found a negatively charged surface cleft present in the antifungal thaumatin-like proteins but not in the sweet thaumatin. Besides, in PR-5d and zeamatin, two phenylalanine residues, Phe91 and Phe96, form a locally hydrophobic patch at the edge of the cleft region. These residues are conserved among the antifungal PR-5 proteins, but in thaumatin tyrosine residues have replaced them. A multiple sequence alignment of thaumatin-like genes retrieved from the Genebank was done using the Pileup program of GCG (Figure 3.38) including the potato homologous **517b**, **517c** and **319b**. For the alignment, the predicted mature protein sequences were considered based on the publication of Koiwa et al. (1999). With few exceptions, most of the thaumatin-like sequences analyzed have the amino acids that form the acidic cleft. However, not all have Phe91 and Phe96 residues, in some they have been replace by tyrosines and in others they have been replaced by proline, glycine, asparagine or tryptophan.

Figure 3.38 Multiple sequence alignment of *Secale cereale* (AF099671), *Triticum aestivum* (X58394), *Oryza sativa* (X68197), *Avena sativa* (U57787), *Hordeum vulgare* (AF355457), *Beta vulgaris* (BVU309172), *Zea mais* **zeamatin** (AAA92882), *Nicotiana tabacum* **PR-5d** (D76437)), *Solanum phureja* 517c gene, *Lycopersicon esculentum* **NP24** (AF093743), *Nicotiana tabacum* **PR-5c** (X6570), *Petunia hybrida* (AF376058), *Capsicum annuum* (AF297646), *Lycopersicon esculentum* p23 gen (X70787), *S. phureja* 517b gene, *Fragaria ananasa* (AF199508), *S. tuberosum* 319b, *Nicotiana tabacum* **PR-5a** (PR-S) (X12739), *Vitis vinifera* (AF227324), *Cicer arietum (*AJ010501)*)*, *Sambucus nigra* (AF378571), *Thaumatoccocus daniellii* (PO2884), *Atriplex nummularia* (M84468), *Malus domestica* (JC7201), *Prunus avium* (U32440), *Castana sativa* (CSA242828), *Pyrus serotina* (AB006009), *Daucus carota* (AAL47574), *Arabidopsis thaliana* (U71244) and *Benincasa hispida* (AF175270). Residues forming the acidic cleft are in yellow and those corresponding to Phe91 and Phe96 are in green. No identical amino acids in the conserved positions are in purple.

```
                51                                  81
    S.cereale   ..VPAGTQAG RIWARTGCSF NG.GTGSCQT GDCG.GQLSC S.LSGRPPAT
  T.aestivum    ..VPAGTQAG RIWARTGCSF NG.GSGSCQT GDCG.GQLSC S.LSGRPPAT
    O.sativa    ..VPAGTSSG RVWGRTGCSF DGSGRGSCAT GDCA.GALSC T.LSGQKPLT
    A.sativa    ...PAGTRGA RVWPRTGCTF DASGRGHCVT GDCG.GALAC R.VSGQQPAT
   H.vulgare    ...PAGTAGA RVWPRTGCTF DGSGRGRCIT GDCG.GALAC R.VSGQQPTT
   B.vulgaris   .YANPGQTAA RIWARTGCRY NGQNGLICNT GDCG.GQFHC TGY.GRGPNT
      Z.mais    .TAPAGTTAA RIWARTGCKF DASGRGSCRT GDCG.GVLQC TGY.GRAPNT
  TobaccoPR5d   ..APPGTKMA RIWGRTNCNF DGAGRGWCQT GDCG.GVLEC KGW.GKPPNT
   Potato517c   ..APPGTKMA RIWGRTNCNF DGAGRGGCQT GDCG.GVLDC KGW.GKPPNT
  TomatoNP24    ..APRGTKMA RIWGRTGCNF NAAGRGTCQT GDCG.GVLQC TGW.GKPPNT
  TobaccoPR5c   ..APRGTKMA RVWGRTNCNF NAAGRGTCQT GDCG.GVLQC TGW.GKPPNT
   P.hybrida    ..APAGTKMA RIWGRTKCNF NAAGRGSCQT GDCG.GVLQC TGW.GKPPNT
    C.annuum    ..APPGTAMA RIWGRTNCNF DGSGRGSCQT GDCG.GVLQC TGW.GKPPNT
   Tomatop23    ..APRGTKMA RIWGRTNCNF DGAGRGSCQT GDCG.GVLQC TGW.GKPPNT
   Potato517b   ..APRGTTMA RIWGRTGCNF DGAGRGSCMT GDCG.GVLQC TGW.GKPPNT
    F.ananasa   ..VAAGTKGA RIWPRTNCNF DGAGRGRCQT GDCG.GLLQC QGY.GQPPNT
   Potato319b   ..VNPGTIQA RIWGRTNCNF DGSGRGKCET GDCN.GLLEC QGY.GSPPNT
  TobaccoPR5a   ..VNPGTVQA RIWGRTNCNF DGSGRGNCET GDCN.GMLEC QGY.GKPPNT
   V.vinifera   ..VNPGTTNA RIWGRTSCTF DANGRGKCET GDCN.GLLEC QGY.GSPPNT
    C.arietum   ..VNAGTSMA RIWGRTGCNF DGSGRGRCET GDCT.GGLQC TGW.GVPPNT
      S.nigra   ..VPAGTRGA RIWARTNCNF DGAGRGRCQT GDCN.GLLSC QAY.GAPPNT
    T.danielli  ..VEPGTKGG KIWARTDCYF DDSGRGICRT GDCG.GLLQC KRF.GRPPTT
  A.nummularia  LPARAG.... .VWGRTGCTS NGGNNLQCTT GGCG.TLFDC GMNSGAPPLT
   M.domestica  ..APSPWS.G RFWGRTRCST DAAGKFTCET ADCGSGQVAC NGAGAVPPAT
      P.avium   ..TPVPWN.G RFWARTGCST DASGKFVCAT ADCASGQVMC NGNGAIPPAT
    C.sativa    ..VQAPWK.G RFWARTRCTT N.SGKFTCET ADCSTGQVAC NGNGAIPPAS
   P.serotina   ..VQAPWS.G RFWGRSHCSI DSSGKFKCST GDCGSGQISC NGAGASPPAS
     D.carota   ..IPAPWS.G RIWARTFCAA ......TCLT GECGKGTGPC SGAGGAPPVT
   A.thaliana   ..LPPLWS.G RFWGRHGCTF DGSGRGRCAT GDCG.GSLTC NGAGGSPPAT
    B.hispida   ..APYQHWSG RVWARTGCVG DLN.YLTCQT GDCG.GKLEC NGAGGKTPAT

                      91    96                                  131
    S.cereale   LAEFTIGGGS TQDF..YDIS VIDGFNLAMD FSCSTGDALQ CRDPSCPP..
  T.aestivum    LAEYTIGGGS TQDF..YDIS VIDGFNLAMD FSCSTGDALQ CRDPSCPP..
    O.sativa    LAEFTIGG.. SQDF..YDLS VIDGYNVAMS FSCSSGVTVT CRDSRC....
    A.sativa    LAEYTLGKGG AKDF..FDLS VIDGFNVPMS FQPV..GGAA CRGATCAADI
   H.vulgare    LAEYTLGQGA NKDF..FDLS VIDGFNVPMS FEPV...GAS CRAARCATDI
   B.vulgaris   LAEYALKQWN DLDF..FDIS LVDGFNVPMS FLPT..N.GC NRGPTCAADL
      Z.mais    LAEYALKQFN NLDF..FDIS LIDGFNVPMS FLPDGGS.GC SRGPRCAVDV
  TobaccoPR5D   LAEYALNQFS NLDF..WDIS VIDGFNIPMS FGPTKPGPGK CHGIQCTANI
   Potato517c   LAEYALNQFG NLDF..WDIS VIDGFNIPMS FGPTKPGPGK CHGIQCTANI
  TomatoNP24    LAEYALDQFS NLDF..WDIS LVDGFNIPMT FAPTKPSGGK CHAIHCTANI
  TobaccoPR5C   LAEYALDQFS GLDF..WDIS LVDGFNIPMT FAPTNPSGGK CHAIHCTANI
   P.hybrida    LAEYALNQFG NLDF..WDIS LVDGFNIPMT FAPTKPSAGK CHPIHCTANI
    C.annuum    LAEYALNQFN NLDF..WDIS LVDGFNIPMT FAPTNPSGGK CHAIQCTANI
   Tomatop23    LAEYALDQFS NLDF..WDIS LVDGFNIPMT FAPTNPSGGK CHAIHCTANI
   Potato517b   LAEYALDQFS NLDF..WDIS LVDGFNIPMT FAPTNPSGGK CHAIQCTANI
    F.ananasa   LAEYALNQYM NRDF..YDIS LIDGFNVPMD FSPVS..NGC TRGIRCTADI
   Potato319b   LAEFALNQPN NLDF..VDIS LVDGFNIPME FSPIN..GGC .RNLLCNAPI
  TobaccoPR5a   LAEFALNQPN Q.DF..VDIS LVDGFNIPME FSPTN..GGC .RNLRCTAPI
   V.vinifera   LAEFALNQPN NLDY..IDIS LVDGFNIPMD FS......GC .RGIQCSVDI
    C.arietum   LAEFALNQYG NLDF..YDIS LVDGFNIPMD FFPIN..GGC HK.ISCTADI
      S.nigra   LAEYALNQFN NLDF..FDIS LVDGFNVAMD FSP.T..GGC ARGIQCTADI
    T.danielli  LAEFSLNQYG K.DY..IDIS NIKGFNVPMD FSPTTRG... CRGVRCAADI
  A.nummularia  IAEYTLT..N TLDT..IDIS LVDGFNVPMS F......GGC PNSPSCASNI
   M.domestica  LVEITIAANG GQDY..YDVS LVDGFNLPMS VAP.QGGTGE CKPSSCPANV
      P.avium   LAEFNIPAGG GQDF..YDVS LVDGFNLPMS VTP.QGGTGD CKTASCPANV
    C.sativa    LVEINIAANR GMDF..YDVS LVDGYNLPVS VAT.RGGTGD CKATSCRANV
   P.serotina   LVELTLATNG GQDF..YDVS LVDGFNLPIK LAP.RGGSGD CNSTSCAANI
     D.carota   LVEFTLNGDG GKDF..YDVS NVDGFNLPVS ITP.E..NSP CATTSCAANI
   A.thaliana   LAEITLGQE. .LDF..YDVS LVDGYNLAMS IMP.LKGRGQ CSYAGCVSDL
    B.hispida   LAQFSLH.HG HKDFSSYGVS LVDGFNIPLT ITP.HEGHGV CPVVGCKANL
```

```
                  141                                                 181
    S.cereale  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   T.aestivum  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     O.sativa  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     A.sativa  TKQCPQELKV .....A.G.. GCASACGKFG GDTYCCRG.. QFTDKCPPTN
    H.vulgare  TKECLKELQV .....P.G.. GCASACGKFG GDTYCCRG.. QFEHNCPPTN
   B.vulgaris  NGPCPGPLRA .....N.G.. GCNNA..... .......... ......YNYS
      Z.mais   NARCPAELRQ .....D.G.. VCNNACPVFK KDEYCCVG.. SAANDCHPTN
  TobaccoPR5d  NGECPGSLRV .....P.G.. GCNNPCTTFG GQQYCCT... QGP..CGPTE
   Potato517c  NGECPGSLRV .....P.G.. GCNNPCTTFG GQQYCCN... HGP..CGPTD
   TomatoNP24  NGECPRALKV .....P.G.. GCNNPCTTFG GQQYCCT... QGP..CGPTE
  TobaccoPR5c  NGECPRELRV .....P.G.. GCNNPCTTFG GQQYCCT... QGP..CGPTF
    P.hybrida  NGECPGALRV .....P.G.. GCNNPCTTFG GQQYCCT... QGP..CGPTD
    C.annuum   NGECPGSLRV .....P.G.. GCNNPCTTFG GQQYCCT... QGP..CGPTE
    Tomatop23  NGECPGSLRV .....P.G.. GCNNPCTTFG GQQYCCT... QGP..CGPTD
   Potato517b  NGECPGQLKV .....P.G.. GCNNPCTTFG GQQYCCT... QGP..CGPTF
    F.ananasa  NGQCPAQLRA .....P.G.. GCNNACTVSK TDQYCCN... S..GHCGPTD
   Potato319b  NDQCPNELRT .....P.G.. GCNNPCTVFK TNEFCCT... NGPGSCGPTD
  TobaccoPR5a  NEQCPAQLKT .....Q.G.. GCNNPCTVIK TNEFCCT... NGPGSCGPTD
    V.vinifera NGQCPSELKA .....P.G.. GCNNPCTVFK TNEYCCT... DGPGSCGPTT
    C.arietum  NGQCPNELRA .....Q.G.. GCNNPCTVYK TNEYCCT... NGQGSCGPTN
      S.nigra  NGQCPNELRA .....P.G.. GCNNPCTVYR TNEYCCT... NGQGTCGPTN
    T.danielli VGQCPAKLKA .....PGG.. GCNDACTVFQ TSEYCCTT.. ...GKCGPTE
 A.nummularia  LDSCPSDLKV ........NG GCLSACNKYS TDEYCCRG.. QYEKNCPPNK
  M.domestica  NKVCPAPLQV KAADG.S.VI SCKSACLAFG DSKYCCTPPN NTPETCPPTE
      P.avium  NAVCPSELQK KGSDG.S.VV ACLSACVKFG TPQYCCTPPQ NTPETCPPTN
     C.sativa  NAVCPAELQV KGSDA.S.VL ACKSACTAFN QPQYCCTGAF DTARTCPATK
    P.serotina NTVCPAELSD KGSDG.S.VI GCKSACLALN QPQYCCTGAY GTPDTCPPTD
     D.carota  NEGCPAGQEV KGPDG.A.TV GCKSACAVTN KPEDCCTGEF NNAEKCKPSA
    A.thaliana NRMCPVGLQV RSRNG.KRVL ACKSACSAFN SPQYCCTGTF GNPLTCKPTS
    B.hispida  LQTCPRELQV HAPQRYGQVI ACKSGCEAFN TDALCCRGHY NSPQTCKASS

                  191                                       221
    S.cereale  ~~~~~~~~~P QAYQHPNDMA T....HAC.R GNSNYQITFC P
   T.aestivum  ~~~~~~~~~P QAYQHPNDVA T....HAC.S GNNNYQITFC P
     O.sativa  ~~~~~~~~~P DAYLFPEDN. TKT  HAC.S GNSNYQVVFC P
     A.sativa  YSKFFKGKCP DAYSYAKDDQ TST..FTCPV G.TNYQIVLC P~~
    H.vulgare  YSKFFKGKCP DAYSYAKDDQ TST..FTCPA G.TNYQIVLC P~~
   B.vulgaris  YSRFFKGRCP DAYSYPQDDA TSM..YSCPS G.TNYKVTFC P~~
      Z.mais   YSRYFKGQCP DAYSYPKDDA TST..FTCPA G.TNYKVVFC P~~
  TobaccoPR5d  LSRWFKQRCP DAYSYPQDDP TST..FTCTS WTTDYKVMFC PYG
   Potato517c  LSRFFKQRCP DAYSYPQDDP TST..FTCQS WTTDYKIMFC P~~
   TomatoNP24  LSKFFKKRCP DAYSYPQDDP TST..FTCPG GSTNYRVVFC P~~
  TobaccoPR5c  FSKFFKQRCP DAYSYPQDDP TST..FTCPG GSTNYRVIFC P~~
    P.hybrida  LSKFFKKRCP DAYSYPQDDP TST..FTCPS GSTNYKVVFC P~~
    C.annuum   LSKFFKKRCP DAYSYPQDDA TST..FTCPS GSTNYRVVFC P~~
    Tomatop23  LSRFFKQRCP DAYSYPQDDP TST..FTCPS GSTNYRVVFC P~~
   Potato517b  FSNFFKQRCP DAYSYPQDDP TST..FTCPS GSTNYRVVFC P~~
    F.ananasa  YSRFFKSRCP DAYSYPKDDA TSTVLFTCPG G.TNYRVVFC P~~
   Potato319b  LSRFFKQRCP DAYSYPQDDP TS..LFTCPA G.TNYKVVFC P~~
  TobaccoPR5a  LSRFFKARCP DAYSYPQDDP PS..LFTCPP G.TNYRVVFC P~~
    V.vinifera YSKFFKDRCP DAYSYPQDDK TS..LFTCPS G.TNYKVTFC P~~
    C.arietum  FSTFFKDRCH DAYSYPQDDP TS..TFTCPA G.SNYKVVFC P~~
      S.nigra  FSRFFKERCR DAYSYPQDDP TS..TFTCPG G.TNYRVVFC P~~
    T.danielli YSRFFKRLCP DAFSYVLDKP TTV...TCP. GSSNYRVTFC PTA
 A.nummularia  YSMIFKGLCP QAYSYAKDDQ SST..FTCPS G.TNYVVTFC P~~
  M.domestica  YSEIFEKQCP QAYSYAYDDK NST..FTC.S GGPDYVITFC P~~
      P.avium  YSEIFHNACP DAYSYAYDDK RGT..FTC.N GGPNYAITFC P~~
     C.sativa  YSRIFKQQCP QAYSYAYDDS TST..FTC.S GAPDYVITFC P~~
    P.serotina FSKVFKNQCP QAYSYAYDDK SST..FTC.F GGPNYEITFC P~~
     D.carota  SSKYFKGKCP QAYSYAYDDK SST~~~~~~~ ~~~~~~~~~~ ~~~
    A.thaliana YSKIFKVACP KAYSYAYDDP TSI..ATC.S KA.NYIVTFC P~~
    B.hispida  CSLFFKHACP STFTYAHDTP SLM..HEC.A APRELKVIFC H~~
```

3.6 *PPO* gene family

Probes with homology to *PPO* (polyphenol oxidase) genes were also used in the screening of the BB BAC library (see Material and Methods). Ten BAC clones with sequence homology to *PPO* were fished: BB1g23, BB21f9, BB45n14, BB48e10, BB63h23, BB66i21, BB66p8, BB91j6, BB98e8, BB104a9. The ends of the BACs were sequenced with the standard primers T3 and T7. The sequences were blasted using the program BLASTX 2.2.1 of the NCBI(GenBank) . The results of the blast search are presented in Table 3.8

Table 3.8 Sequence similarity of T3 and T7 ends of *PPO*-like BACs.

BAC end	Blasted seq (bp)	Species	Accession number	E value	Score	% Identity	Function
BB1g23T7	814 *	*S.tuberosum*	M95197	e^{-141}	273	91	propolyphenol oxidase
BB1g23T3	415 *	*A. thaliana*	NM_114023	$3e^{-10}$	63.9	36	unknown protein
BB21f9T3	443 ***	-	-	-	-	-	-
BB21f9T7	509 *	*N.tabacum*	X80830	$5e^{-67}$	253	72	integrase
BB45n14T3	863 ***	-	-	-	-	-	-
BB48e8T3	415 *	*S.tuberosum*	U22921	$2e^{-42}$	170	62	polyphenol oxidase
BB48e8T7	811 *	*A.thaliana*	NM_102255	$2e^{-7}$	57	35	unknown protein
BB66i21T3	512 *	*S.tuberosum*	U22922	$9e^{-55}$	213	57	polyphenol oxidase
BB66i21T7	517 *	*S.tuberosum*	M95197	$4e^{-14}$	77.8	47	propolyphenol oxidase
BB63h23T3	416 *	*A.thaliana*	AP002071	$9e^{-22}$	102	40	putative non LTR-retroelement
BB63h23T7	386 *	*S.tuberosum*	U22922	7.7	29	78	polyphenol oxidase
BB66p8T3	445 *	*S.tuberosum*	M95197	0.001	42.7	45	propolyphenol oxidase
BB66p8T7	497 ***	-	-	-	-	-	-
BB91j6T3	790 ***	-	-	-	-	-	-
BB91j6T7	360 ***	-	-	-	-	-	-
BB98e8T3	424 ***	-	-	-	-	-	-

BB98e8T7	414 ***	-	-	-	-	-	-
BB104a9T7	411 ***	-	-	-	-	-	-

* Homologies generated by BLASTX.
***Not significative homology.

The sequences of the end of the BACs were used to design primers. PCR amplification using the different primer pairs was carried out on *PPO* BAC DNAs and DNA from *S. phureja*, *S. tuberosum* and PD59. The presence of PCR products on different BAC templates indicated overlaps among BACs. Sequencing and alignment of PCR products confirm the relationship between BACs and revealed the parental line from which the clones were derived. Based on the PCR experiments and in a Southern blot analysis (Figure 3.39), the BAC clones were organized in four groups (A, B, C and D) (Figure 3.40).

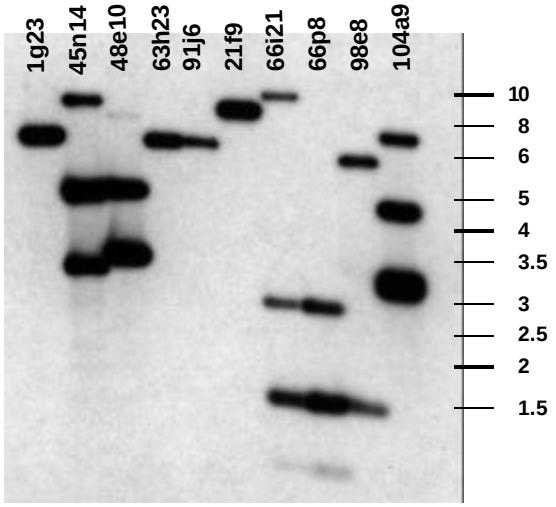


Figure 3.39 Autoradiogram of a Southern blot gel with *PPO*-like BACs hybridized with *PPO* probe. The common bands among different clones suggest possible overlaps. The sizes are indicated in kb.

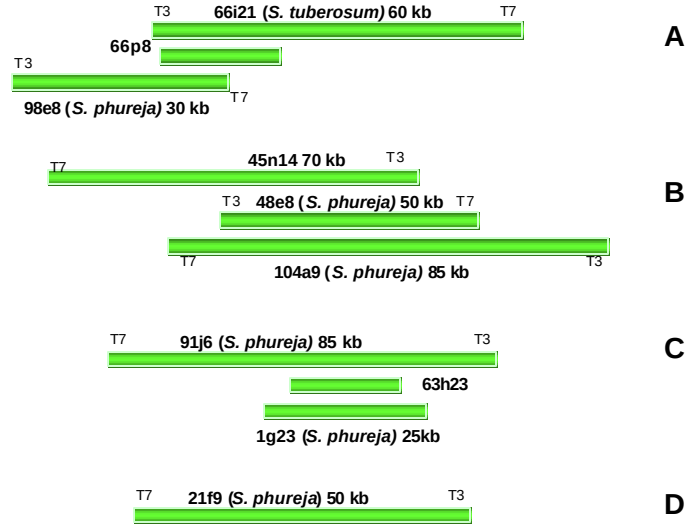


Figure 3.40. Schematic representation of BACs with sequence homology to *PPO* genes. The clones are organized in four groups A, B, C and D.

A SSCP marker derived from the clone BB91j6 (Group C) was mapped in the F1840 population to LG VIII. The position corresponds to the location of the PPO potato family (Tanksley et al. 1992). Groups A, B and D could represent different loci and they remain to be mapped.

Figure 3.41 shows a southern blot of *Solanaceae* species and some of the *PPO* BACs hybridized with the PCR *PPO* probe. Most of the hybridizing fragments present in the BACs could be recognized in the genomic DNA of PD59. However, due to the background not all BAC fragments were identified in the genomic DNA (fragments between 8 and 10 kb).

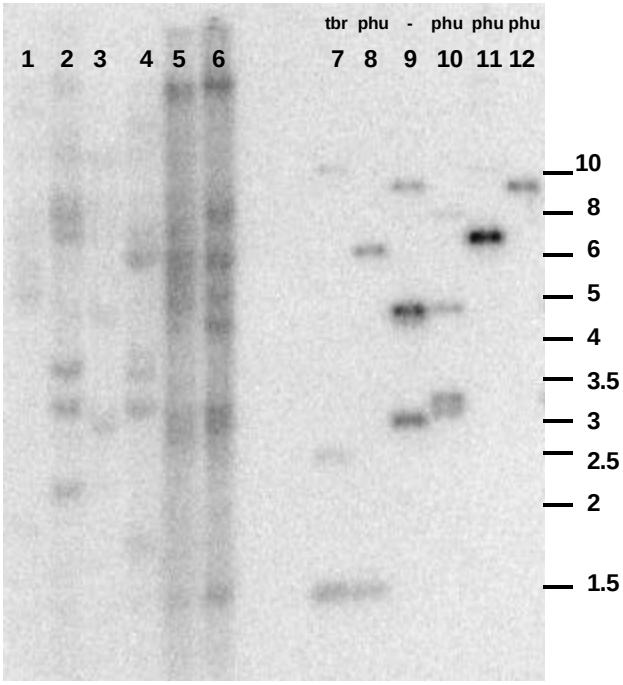


Figure 3.41 Autoradiogram of DNA blot hybridized with *PPO* probe. Lines 1 to 6 correspond to *C. annuum*, *Lycopersicon sp., N. tabacum*, *S. melongena*, *S. tuberosum* and PD59. Lines 7 to 12: BB66i21(7), BB98e8(8), BB45n14(9), BB104a9(10), BB91j6(11) and BB21f9(12). The parental line from which the BACs are derived is indicated, tbr (S*. tuberosum)* and phu (*S. phureja)*. The origin of BB45n14 could not be inferred. The relative sizes are indicated in kb.

IV Discussion

4.1 Construction of a potato large insert library

The most important requisite for constructing a large insert library is the amount and quality of high molecular weight (HMW) DNA. In this study, HMW DNA was obtained from nuclei and protoplasts. Both methods were compared to choose the most efficient one for the construction of the library. The isolation of nuclei was based on the protocol of Zhang et al. (1995) that has been used in the construction of most BAC libraries from different plant species (Wang et al. 1995; Frijters et al. 1997; Danesh et al. 1998; Vinatzer et al. 1998; Hamilton et al. 1999; Vanhouten et al. 1999; Nam et al. 1999; Tomkins et al. 1999a, 1999b; Yu et al. 2000; Gentzbittel et al. 2001; Deng et al. 2001; Meksem et al. 2001; Wang et al. 2001 and Ming et al. 2001). This method extracts and precipitates nuclei in a fast way but does not purify them. The purity of HMW DNA has a significant impact on the degree of digestion (Ming et al. 2001). Frijters et al. (1997) reported significant shearing of the DNA from isolated nuclei. However, in the present study the shearing of the DNA was minimal and comparable to the shearing produced in HMW DNA from protoplasts. An advantage of the nuclei method is that the amount of chloroplastic and mitochondrial DNA is very low. Wang et al. (1995) reported less than 0.3% of chloroplast and mitochondrial clones for a sorghum library, while Tomkins et al. (1999a) reported 1.85% of chloroplast DNA for a soybean library.

On the other hand, protoplast preparation yields DNA of very high quality. However, it must be first standardized for each plant species and the procedure is time consuming. The protocol for potato protoplasts was already established in this laboratory and this was an advantage. Few libraries have been constructed using DNA from protoplasts (Woo et al. 1994; Nakamura et al. 1997; O'Sullivan et al. 2001). A disadvantage of using protoplasts for library constructions is that the amount of organellar DNA is higher than in nuclei preparations. Woo et al. (1994) constructed a BAC library for sorghum from protoplasts and the chloroplastic DNA content calculated was 14%. Nakamura et al. (1997) estimated 7% chloroplastic content for a rice BAC library constructed using protoplasts. The amount of chloroplastic DNA is not a limitation when a library is representative. In the present study, the amount of organellar DNA has not been measured. However,

positive clones have been obtained with different probes indicating that the library is representative.

After isolation, nuclei and protoplasts were embedded in agarose plugs to prevent degradation and to create an interchange matrix for further procedures. In this stage the plugs can be pre-run on a PFGE to eliminate the sheared DNA. This procedure was not used because the amount of degraded DNA was minimal and it was better to avoid as much as possible extra steps that involve manipulation of DNA. To facilitate the diffusion of the restriction enzyme into the agarose, the plugs were sliced prior to digestion. Mechanical shearing due to slicing has not been reported (Wang and Schwartz, 1993; Allouis et al. 2001). Nuclei plugs were softer and this made them difficult to handle. Protoplasts plugs had a harder consistency and were ideal, therefore, for getting equal slices for partial digestion. In this study, nuclei isolated from 15g of leaves yielded a concentration of DNA of 1.5 μg/plug when the nuclei were resuspended in 1ml of buffer. The concentration of DNA per plug used in library constructions ranges from 1 μg/plug (Meksem et al. 2000) to 15 μg/plug (Frijters et al. 1997). Taking into account that the initial concentration of DNA is reduced during the whole procedure, it was decided to use protoplasts for their higher DNA concentration (3 μg/plug) and their better consistency.

The multiple steps involved in the construction of a library make it difficult to reproduce results even though using the same conditions. This is particularly true for the partial digestions, which are a pivotal step. Partial digestions must be optimized for every species. Moreover, the concentration and quality of each different preparation of nuclei or protoplasts has to be checked. The units of enzyme required for partial digestion and the time of incubation varies tremendously in the literature. For example, Meksem et al. (2000) used 0.1 to 2 U of *Hin*dIII per 1 μg of DNA and 20 min at 37ºC for soybean nuclei. On the other hand, Lin et al. (1999) used 18-30 U of *Hin*dIII per 1.2 μg of DNA and 5 min at 37°C for sorghum nuclei. In general, the units of enzyme are inversely proportional to the time of incubation but both depend on the diffusion of the enzyme into the plugs. For diffusion, the sliced plugs are incubated with the enzyme on ice for a period of time that has to be tested. A longer diffusion time will require less units of enzyme and less incubation time.

In this study, the optimization windows suggested by Choi and Wing (http://www.genome.clemson.edu/protocols) were used. The objective was to get the highest amount of digested DNA in the range of 100 to 300 kb. From the results of the wide and narrow optimization windows it was concluded that 4 to 6 U of *Hin*dIII were required for digesting 3 µg of DNA with 20 min of incubation at 37°C. The diffusion time of the enzyme was one hour. These conditions gave similar results in all batches of protoplasts. Some authors prefer to use different concentrations of enzyme and, after PFGE, fragments of the aimed size were cut from the gel without previous optimization. For example, Nam et al. (1999) used concentrations of 0.5, 1, 2 and 4 U of *Hin*dIII to digest *Medicago truncatula* nuclei and after PFGE, the slices containing DNA ranging in size from 200 kb to 500 kb were excised and directly used for ligation. The disadvantage of this approach is that is not possible to know if the slice that has been excised contains enough amount of DNA.

Especially critical was removing the selected fragments from the pulse-field gel. Two methods are commonly being used: Electroelution and agarose digestion. Both were assayed and compared in order to choose the most convenient one. Electroelution is a good alternative to enzymatic gel solubilization. This procedure does not involve any mechanical or chemical potential risk for the DNA. Strong et al. (1997) reported the improvement of PAC and BAC cloning by electroelution of fragments of up 250 kb. Few libraries have been constructed using this method (Deng et al., 2000, Wang et al., 2001; O'Sullivan et al., 2001). In the case of potato, size markers of up 800kb were successfully electroeluted but the largest fragment of potato DNA electroeluted was less than 80 kb in size. An explanation may be the higher purity of the markers and the matrix where they were embedded. For this reason, they could move faster and more efficient than the potato DNA. Due to the results, electro-elution was not used.

Other libraries have been constructed using a combination of electroelution and agarose digestion (Tomkins et al., 1999a and 1999b and Yu et al., 2000). The enzymatic digestion of the agarose has as main advantage that all DNA fragments are removed from the gel and no DNA is lost. Agarose digestion using Gelase has been more popular in the construction of genomic libraries (Woo et al. 1994, Wang et al. 1995, Nakamura et al. 1997, Frijters et al. 1997, Vinatzer et al. 1998; Hamilton

et al. 1999; Vanhouten et al. 1999; Nam et al. 1999, Salimath et al. 1999, Lin et al. 1999, Folkerstma et al. 1999 and Ming et al. 2001). However, the digestion of agarose using gelase requires a melting step that compromises the integrity of the DNA. Prior to melting, the agarose slices are incubated with Gelase buffer. The buffer includes 40 mM NaCl among other components. NaCl has a DNA protective activity and some investigators used a high salt concentration to reduce the heat-induced damage (Amemiya et al., 1996). The initial concentration was increased to 100mM. Even though certain reduction on size of the DNA was observed during the gelase digestion, improvement was obtained.

In most cases so far reported including the present work, the average size of the insertions obtained is lower than the size of the DNA selected (Danesh et al. 1998; O'Sullivan et al. 2001). This is due to the fact that smaller DNA fragments are preferentially transformed. Second-size selection has been reported to effectively reduce the number of small insert clones in the library by eliminating smaller 'trapped' DNA fragments from the size-selected gel slice (Woo et al. 1994; Zhang et al. 1996; Frijters et al. 1997). The method of second size selection was assessed (data not shown). Fragments between 100 and 300 kb were cut and run for a second time using: initial pulse 5s, final pulse 5s for 5 hours. Although a compression zone of fragments bigger than 100 kb was visualized, the efficiency of ligation and transformation was extremely low and insertions were much smaller than inserts selected after one PFGE running. Unsatisfactory results with this method have been reported previously (Danesh et al. 1998; Vanhouten et al. 1999; Allouis et al. 2001; Ming et al. 2001). Vanhouten et al. (1999) also observed that after second size selection the number of false positives was higher. Therefore, the DNA fragments for the present library were selected only once on PFGE. This procedure resulted in 30% of the clones with insert DNA smaller than 75 kb. However, keeping the number of manipulations to a minimum avoided the degradation of DNA giving higher transformation efficiency what in return compensated for the smaller sizes of inserts.

In the present study, a concentration of 1 ng of insert DNA per µl of ligation was used. This concentration yielded a transformation efficiency up to 900 transformant clones per µl of ligation. The opinion about the amount of DNA to be

ligated varies among the authors. Frijters et al. (1997) believed that during ligation a high DNA concentration could lead to chimeric clones. They used 0.25ng of insert DNA per µl of ligation obtaining usually less than 100 transformant clones per µl of ligation. However, Vinatzer et al. (1998) used 4ng of insert DNA per µl of ligation obtaining approximated 400 transformant clones per µl of ligation. Moreover, they suggested that a high DNA concentration in the ligation mix allows a significant reduction of the cost of library construction, because of the relatively small number of transformation events necessary. The range of the ratio between vector and insert DNA is also quite diverse among the authors. While Folkerstma et al. (1999) used a ratio of 2:1, Nam et al. (1999) used a ratio of 25:1. In both cases, pBeloBAC11 was used as the cloning vector. In general, authors used ratios between 5:1 and 10:1 for small vectors like pBeloBAC11 (7.4 kb). When using the large vector BIBAC (22 kb), Hamilton et al. 1999 used a vector:insert ratio of 3:1 and for pCLD04541 (29kb), Meksem et al. (2000) used a 4:1 ratio. In this study, pCLD04541 was used with a ratio of 3:1.

A Gene Pulser (BioRad) was used for transformations with the following settings: 100Ω, 1.8kV and 25µF. Yang et al. (1997) reported that this device was not suitable for transformation of BACs with large inserts because the pulse time cannot be adjusted. As it was explained, the pulse time depends on the resistance. When decreasing the resistance from 200Ω to 100Ω, the pulse time decreased to the half (from 4.8 to 2.4 sec) and this yielded good transformation efficiencies together with acceptable insert sizes. The largest clone obtained had an insert size of 200 kb. The reduction of the electroporation field strength has also been shown to increase the efficiency with which large DNA fragments could be introduced into *E. coli* (Frijters et al. 1997). Frijters reported 1.3 kV as the best gradient of voltage for transforming large pieces of DNA. Voltages from 1.3kV to 1.9kV were assessed in this study. 1.3 kV showed extremely low transformation efficiency although larger inserts were obtained. The best results were obtained with 1.8kV.

In order to produce the highest average insert size of BACs, independent ligations and transformations with different fragment sizes were conducted. The results among different authors did not lead to clear conclusions. For example, Allouis et al. (2001) observed that transformations with DNA fragments in the 250-

300 kb range resulted in fewer transformants containing, on average, inserts that were smaller than those obtained with DNA in the 150-250 kb range. However, Vinatzer et al. (1998) reported that when a slice was cut from 150-350 kb region, the average insert size of the resulting clones was too small. If, on the other hand, the size selected fragments were too large (440-550kb), they could not efficiently be transformed and only small insertions were obtained. Moreover, Ming et al. (2001) reported that DNA fragments ranging from 50 to 125 kb showed the highest ligation and transformation efficiencies. No positive clone was generated from DNA fragments larger than 200 kb. However, when they repeated the optimal partial digestion conditions, transformations showed the highest ligation and transformation efficiencies for DNA fragments larger than 175 kb and between 75 and 125 kb. Vinatzer et al. (1998) suggested an explanation for these contradictory results: "only a very restricted gel region enables the construction of a library with acceptable insert sizes. The localization of this region is unpredictable because even slight changes in DNA concentration seem to influence its mobility during electrophoresis, so the DNA size marker does not provide exact indication of fragment size". Frijters et al. (1997) also commented that if the pulse field gels are overloaded, there will be excessive trapping of smaller DNA fragments which then co-migrate with larger DNA fragments on a PFG. From the experience of this study, it is concluded that the DNA fraction that enables to construct a BAC library is a matter of trial and error and having enough amount of DNA will allow doing the necessary trials to obtain it.

The analysis of BAC clones by *Not*I digestion followed by PFGE showed that the majority of potato DNA inserts was typically present as single *Not*I fragments. Thus, the potato genome apparently contains few *Not*I sites (< 1 site/ 230 kb), a feature commonly observed in the genomes of other dicot species (Choi et. al. 1995; Frijters et al. 1997; Danesh et al. 1998; Nam et al. 1999). This is in contrast to the results obtained with monocot species (Woo et al. 1994; Wang et al. 1995; Zhang et al. 1996; Moullet et al. 1999; Tomkins et al. 1999; Allouis et al. 2001). For example, Allouis et al. (2001) reported an average of about one *Not*I site every 46 kb in *Pennisetum glaucum*.

The sizes of the inserts recovered from positive BAC clones identified by library screening for specific genes, were smaller than those of the clones sampled

randomly from the library (see the size of *PAL* BAC clones in page 45, the size of *PR-5* BAC clones in figures 3.13 and the size of *PPO* BAC clones in figure 3.40). This observation has been previously reported (Woo et al. 1994; Zhang et al. 1996; Tomkins et al. 1999a; Nam et al. 1999). Tomkins et al. (1999a) suggested that in the case of monocot genomes there are more *Hin*dIII sites in regions containing single copy sequences. Moreover, Nam et al. (1999) gave as potential explanation the presence of disproportionately high numbers of *Hin*dIII sites in genomic regions containing the target genes. They obtained sizes of 25 and 7 kb for inserts containing the gene *ENOD40* in a *Medicago truncatula* library. The gene contains four *Hin*dIII sites in the coding and immediate flanking regions. In this study, the genes that were targets of the screening contain at least one *Hin*dIII site either in the flanking or in the coding region (*PAL*, *PR-5* and *PPO* genes). The effect of this was observed when analyzing the end sequences of the BAC clones. For example, in clones with homology to *PR-5* genes, the T3 end of BB41d14 showed to be part of the flanking region of the gene, the same happened with the T7 end of BB43d14. Several end sequences of BAC clones carrying *PPO* genes were part of the flanking or coding regions of these genes (BB48e8, BB66i21, BB66p8 and BB98e8).

In cases in which the genes of interest are rich in sites for the enzyme that is being used to construct the library, partial digestions are not efficient because, even if short digestion times are allowed, that would be enough to restrict sites that are close to each other. Besides, the inserts are size-selected meaning that the target genes will be always counter selected as the discarded smaller size fractions. To avoid this inconvenience, mechanical shearing of the DNA is suggested as an alternative to enzymatic digestions. However, the use of mechanical force to break the DNA complicates the procedure because the DNA is more exposed to degradation and ligation of adaptors is required.

It has been suggested that libraries with insert sizes of less than 100 kb are not suitable for chromosome walking (Cai et al. 1995). However, with the advent of high-throughput mapping methods, such as AFLPs and SSRs in association with chromosome landing (Tanksley et al. 1995), it may be possible to directly land on the target gene, thus eliminating the time-consuming chromosomal walking step. For applications such as the ones done in this study (cloning of the genomic counterparts

of cDNAs, ESTs and RFLPs) the main limiting factor was the number of BAC clones rather than the size of the inserts (Gentzbittel et al. 2001). The mixture of small and large insert clones in a library is seen more as an advantage than a disadvantage. They may facilitate the characterization of genomic targets by reducing the demand of extensive subcloning (Wang et al. 1996). For example, the size of the *PAL* BAC clones analyzed by PFGE ranged from 40 to 100 kb (there are possibly smaller clones among the other *PAL* clones). The subcloning would be facilitated with the smaller BACs. On the other hand, larger clones would make it easier to look for the neighboring regions of the *PAL* genes. Nevertheless, it would be advisable to increase the number of clones in the library using partial digests with an alternative restriction enzyme such as *Eco*RI or *Bam*HI. It may provide a useful complement to the present *Hin*dIII library by increasing genome coverage and the total number of large insert clones.

4.2 Applications of the potato library

Characterization of gene families like *PAL* and *PR-5* using the BAC library has shown that a large potato insert library is essential for some applications like physical mapping of genomic regions and analysis of genome organization of multi-gene families. Clones containing sequences with homology to *PAL*, osmotin-like genes, intercellular *PR-5* like genes, and *PPO* genes have been isolated. Moreover, complete gene sequences of some osmotin-like genes and E22- like genes have been obtained. The sequences of the gene copies were compared and analyzed to formulate hypothesis about the meaning of its divergence.

4.2.1 *PAL* family

In this study, 29 genomic clones of *PAL* cDNA were obtained. *PAL* genes are differentially expressed after infection of potato leaves with *P. infestans* (Taylor et al. 1990). The Southern gel blot analysis of the *PAL* clones showed that more than one gene seems to be present in most of the clones. Therefore, *PAL* genes are in clusters. The exact number of *PAL* genes in the potato genome was not estimated. However, there were up 16 different hybridizing bands. Joos and Hahlbrock (1992) reported 40-50 copies per potato haploid genome (2n=4x=48). For a diploid species (2n=2x=24) these numbers could be reduced to the half (20-25 copies). 20-25 copies

seem a high number compared to the results of the Southern analysis. Joos and Halbrock (1992) estimated the copy number of *PAL* by titration series based on a haploid potato genome size of 2100 Mbp (Bennet and Smith 1976). This size is larger than the size considered in this study (1500-1800 Mbp). (Arumuganathan and Earle 1991).

The *PAL* clones can be further characterized. Five *PAL* loci have been mapped: *PAL* **a**, **c** on LG IX, *PAL* **d, e** on LG X and *PAL* **f** on LG III (Leonards-Schippers et al. 1994 and Trognitz et al. 2002). It would be interesting to map the *PAL* clones that represent different loci and correlate them with the loci that have been already mapped. QTL effects of resistance to *P. infestans* on LG III and LG IX co-localize with the position of two *PAL* loci, *PAL* f (Trognitz et al. 2002) and *PAL* a, c (Leonards-Schippers et al. 1994). Moreover, Geffroy et al. (2000) found that *PAL* co-localizes with QTL in *Phaseolus vulgaris* for resistance to *Colletotrichum lindemuthianum*. The cloning and overexpression of *PAL* genes corresponding to LGIII and LG IX in transgenic susceptible potato plants may reveal whether these genes are contributing to the expression of resistance. The difference between *PAL* genes will support the hypothesis of Leonards-Schippers et al. (1994) that allelic variants of pathogenesis related genes like *PAL* can cause differences in the quantitative resistance response to *P. infestans*.

4.2.2 PR-5 genes and proteins

The cDNA of PR-S (tobacco intercellular PR-5) was used to screen the potato library and one clone, BB43d14, was obtained. Southern blot analysis showed that the genome of Samsun NN tobacco contains two genes for the PR-S protein (van Huijsduijnen et al. 1987). They have been named E22 and E2, and they are highly homologous in their coding regions (van Kan et al. 1989). Pierpoint and coworkers (1987) identified two isoforms of the thaumatin-like protein of tobacco, which are expressed in a 3:2 ratio. These major and minor forms are encoded by E22 and E2, respectively. In this study, a Southern blot analysis of *Eco*RV digested tobacco and potato DNA using the potato counterpart of PR-S (PCR fragment of E22-like gene) as probe, showed in tobacco two bands corresponding to E22 and E2. The potato genomic DNA also showed two bands of similar size possibly containing the counterparts of E22 and E2.

The E22 homolog present in subclone 319b may be the potato counterpart of one of the isoforms of PR-S. Blast results indicated that 319b is more homologous to E22 than to E2. This is in agreement with the findings of Pierpoint et al. (1990). They described the appearance of an acidic thaumatin-like protein in intercellular fluids (IF) of potato (*S. tuberosum*) leaves during aging and after treatment with salycilate. The protein was called Protein C. From the sequence of the 58 N-terminal amino acids they concluded that Protein C is composed of a single isoform that is about 86% identical in sequence to the two isoforms of the tobacco protein. In the present study, a RT-PCR experiment showed the constitutive transcription of PR-S genes in potato lines PD59 and 40 (*S. tuberosum*, data not shown). The analysis of the sequences revealed that the transcripts in both genotypes corresponded to a single gene, 319b. Comparison of the 58 N- terminal amino acid sequence of Protein C with 319b showed 100% identity. From this result, it was concluded that 319b is the gene of Protein C. The other PR-S genes in potato can be functional but the conditions under which they are expressed are not yet known. RT-PCR experiments could be carried out with *P. infestans* infected or stressed potato leaves cDNAs to check the *PR-5* genes that are expressed.

Pierpoint et al. (1992) observed that the intercellular PR-5 or PR-S protein was induced in several species of *Nicotiana* by virus infection, and occurred constitutively in an interspecific hybrid. They did not detect intercellular PR-5 in tomato IF. They suggested that intercellular PR-5 is not part of the tomato PR proteins since tomato produces osmotin (AP24) and NP24 after infection with fungi and viroids. The present study has shown by Southern analysis and PCR experiments that *PR-5* (osmotin-like and intercellular) genes are present in all potato species, in *S. dulcamara*, *S. nigrum*, tomato, pepper and eggplant. The degree of conservation of intercellular *PR-5* genes among *Solanaceae* species indicates that they are part of an old and common stress response as proposed by Pierpoint et al. (1992). However, with the exception of tobacco and potato, none of the corresponding mRNAs or proteins from these species has been isolated so far. Therefore, their role in such a response remains to be clarified. A way to elucidate their function is the use of a large battery of pathogens in bioassay tests to determine the activity of these proteins.

A Northern gel blot experiment with *P. infestans* infected potato leaves did not show a clear induction of transcript level of intercellular *PR-5* genes. Probably, this was due to the fact that the level of transcription in the control leaves was already induced. Besides, a short time period was examined, 18 and 24 hours. Presence of the Protein C has been reported in potato leaves sprayed with salicylate after 5-10 days and in old plants (Pierpoint et al. 1990). The RNA was prepared from old plants (approximately 60 days) growing under unsterile conditions. Younger plants growing under sterile conditions should be used in future experiments.

PD59 shows field resistance to *P. infestans* and no R gene has been found in the parental lines *S. phureja* CHS-625 and dihaploid *S. tuberosum* PS-3. Therefore, it was expected that PD59 leaves infected with *P. infestans* would show a typical compatible interaction. Since, in the field complex races of the pathogen are present, the tests were firstly done using a complex race (1-11). However, the inoculum did not show any sign of infection even in susceptible Desirée leaves. Therefore, it was decided to use a simple race. Race 4 grew faster and was more efficient on infection. PD59 showed the same response as the positive control of the infection tests, the susceptible *S. tuberosum* cv. Desirée. Physiological and environmental conditions might play a very important role on conditioning field resistance as reported by Collins et al. (1999). Although PD59 grew normally in the greenhouse, the optimal conditions for potato plants that usually grow under short days may not be optimal. The QTL experiment of Ghislain et al. (2001) studied loci associated with *P. infestans* resistance from plants growing under short days. Besides, the infection tests were done in covered trays to provide the pathogen the optimal conditions for infection. Vleeshouwers et al. (1999) showed that detached leaves incubated in covered trays at high relative humidity were more susceptible than detached leaves kept in open trays or leaves on intact plants when challenged with *P. infestans*.

BB43d14 is carrying all PR-S like genes of potato in less than 35kb. This implies that markers derived from this clone could be use to detect any effect on resistance from this class of genes. The BAC clone was mapped in one of the parental lines, *S. phureja*. The mapping position showed that the clone does not

overlap with the QTL effect on LGXII from *S. phureja* for *P. infestans* resistance. Although no marker could be obtained for the other parental line (dihaploid *S. tuberosum*), the correlation of markers between both lines suggests that potato PR-S could be part of the QTL effect of *S. tuberosum* in the same chromosome. The analysis of flanking and coding sequences of genes 21a, 319b and 321c in both parental lines have shown that both species are highly homologous; therefore the translated proteins must be nearly identical as well as their properties. Moreover, the clone BB43d14 was derived from *S. tuberosum*, and the gene 319b is identical to the N-terminus of potato Protein C (Pierpoint et al. 1992) (see above). The activity of Protein C against fungi or oomycetes has not been tested. Protein C is 86% homologous to tobacco PR-S, this suggests that the activity of both proteins could be similar. Tobacco PR-S did not show any effect against *P. infestans*. However, the non-identical amino acids could represent a difference in activity as suggested by Vigers et al. (1992).

Zhu et al. (1995a and 1995b) reported the constitutive accumulation of three osmotin-like protein mRNAs (pA13, pA35 and pA81) in cell cultures, stems, roots and flowers of potato (*S. commersonnii*). Treatment with ABA, low temperature and NaCl increased the accumulation of all three mRNAs. On the other hand, salycilate and wounding resulted in only a moderate increase in the levels of pA13 and pA81 but not pA35 mRNAs. Infection with *P. infestans* activated strong and non-systemic the accumulation of all three mRNAs after 4 days of inoculation. This suggested that osmotin induction following pathogen attack might be induced by signals other than salycilate. Indeed, several osmotin have been shown to respond to other defence-related signals such as ethylene or jasmonate (Jung et al. 1995; Sato et al. 1996; Schweizer et al. 1998). The accumulation of osmotin proteins, however, was detected only in *P. infestans* infected tissues but not in plants treated with ABA, SA, NaCl, low temperature, or wounding. The ethylene responsive element 5'-GCCGCC-3' was found in the promoters of all osmotin-like genes analyzed. It was present at least once in 534a, 517b, 517c and 137d'. 57d and 137d' had two 5'-GCCGCC-3' sequences in tandem. A reverse orientation of the sequence was found in 534a, 517b and 517c. It is interesting to note that the promoters of these genes all have one copy of the element in the same position but the second copy is present in tandem or

in reverse orientation but not both. No ethylene responsive element was found in PR-S like genes. This element is not present either in tobacco PR-S genes. A putative TATA element has been recognized in 319b that is identical to the tobacco PR-S genes E22 and E2. However, no recognizable promoter sequences have been found in 21a and 321c. This might imply that they have unknown promoter elements or that they are pseudo-genes like the truncated copy 321d.

The genes corresponding to pA13 and pA81, OSML13 and OSML81 respectively, where contained in a lambda genomic clone arranged in the same orientation. **534a** and **517b** are in the same orientation in less than 5 kb. It seemed that OSML13 and OSML81 were the analogs of 534a and 517b. The phylogenetic analysis has shown that OSML81 and 517b are analogs. However, OSML13 and 534a are in different branches (see figure 3.36). Therefore, they are not analogs. The present study showed that several osmotin genes that were previously cloned from *S. commersonii* (Zhu et al. 1995a and 1995b) are clustered. 534a, 517b and 517c are within a fragment of 11.4kb and it is possible that 57d is also next to them. An observation that has been reported for defense related genes (Li et al. 1999; Wang et al. 2001): "map locations indicated that defense related gene loci are not randomly distributed throughout the wheat and the rice genomes but are rather located in clusters or in distal gene-rich regions of the chromosomes".

According to the results of Melchers et al. (1993), the targeting information for vacuolar PR proteins resides in between 6 and 21 amino acids beyond the C-terminus of the intercellular PR proteins and is removed during or after transport to the plant vacuole. Liu et al. (1996) reported that in plants over-expressing a C-truncated osmotin gene for the terminal 20 amino acids, osmotin was totally secreted into the extra-cellular matrix. However, the truncated osmotin purified from transgenic tobacco plants retained antifungal activity and potato plants that over-expressed the truncated osmotin protein exhibited resistance to *P. infestans*. In the present study, all gene copies obtained from LGVIII: 534a, 517b, 517c, 57d and 137d' have a C-terminus with homology to osmotin genes, which are vacoular. The gene from LGXI, 68e, has a C-terminus more similar to PR-S genes which are intercellular even though the N-terminus shows higher homology to osmotin-like genes. On the other hand, PR-S like genes of LGXII: 319b and 21a seem to be the genes of intercellular proteins. However, 321c has a long C-terminus. This gene, if

it is transcribed and translated, would correspond to a vacuolar protein, even though its sequence show higher homology to intercellular PR-5 genes.

4.2.3 Physical distance

The relationship between genetic and physical distance for the osmotin cluster in LGVIII could be calculated. One recombinant was present between markers CAPSA and SSCPA in 100 individuals analyzed from the F1840 population. Therefore, the genetic distance is 1cM. The markers are approximately 90kb apart. The calculated relationship between genetic and physical distance is 9 x 10^4bp/cM. On the other hand, the RFLP linkage map length of F1840 is 1000cM (Dr. C. Gebhardt, personal communication). The relationship between genetic and physical distance is 10^6 bp/cM. The relationship between genetic and physical distance calculated for the gene cluster is much smaller than the calculated from the map length.

4.3 Future prospects

Considering that a binary vector (pCLD04541) was used to construct the library, other applications can be considered. In potato, like in other crop species, important agronomical characteristics are under the control of QTLs. Genes underlying QTLs may be organized in clusters like R-genes (Staskawicz et al. 1995). With a large insert library those loci could be finally dissected and sequence of the clones could reveal the identity of the genes behind the QTLs. Moreover, the binary vector containing the QTL could be transferred to the plant via *Agrobacterium*. In theory, susceptible phenotypes requiring more than just one gene but gene families acting in concert could be complemented by transformation (Meksem et al. 2000). Here, it is important to mention that Trognitz et al. (2001) detected a QTL effect to late blight in the mapping position of osmotin on LGVIII. They used the mapping population of Ghislain et al. (2001). This study has shown that osmotin-like genes are clustered on LGVIII where at least 5 genes are located within less than 90 kb. Clones BB18h19 (*S. tuberosum*) and BB48h15 (*S. phureja*) belong to homologous chromosomes and carry all genes of this cluster. Although over-expression of tobacco osmotin and OSML13 resulted in a delay of development of *P. infestans* disease symptoms in *S. tuberosum* and *S. commersonii* (Liu et al. 1994; Zhu et al.

1996) the antifungal activity of the other, closely related potato proteins have not yet been determined. Other studies reported that the constitutive over expression of a rice thaumatin-like gene delayed the symptoms of fungal diseases in transgenics wheat and rice plants (Chen et al. 1999; Datta et al. 1999). It is suggested that resistance response depends on the simultaneous over expression of several osmotin genes. By over expressing defense genes the timing of a natural host defense mechanism is modified becoming more efficient (Zhu et al. 1996). This hypothesis could be tested transforming one of the BACs into a susceptible potato plant. *PAL* genes are also in clusters and the same approach can be used with *PAL* BAC clones mapping to LGIII and LG IX where QTL effects to *P. infestans* co-localized with *PAL genes* (Leonards-Schippers et al. 1994; Trognitz et al. 2002).

The fact that QTL experiments with different wild potato species show different loci contributing to resistance to *P. infestans* (Ewing et al. 2000; Sandbrink et al. 2000 and Naess et al. 2000) and that osmotin has shown to have a inhibitory effect on this pathogen make us to wonder if osmotin is contributing to the defense responses in other plant material grown under different environmental conditions. Markers derived from the osmotin cluster can be used to test association between phenotypic resistance and osmotin genes. If not linkage is found; this could be due to the origin of the plant material, the conditions under which it is grown and the constitutive level of expression of these genes. *S. phureja* is a native species from South America. All potato production in the Southern hemisphere is done under short day conditions. Under these conditions, osmotin might be playing an important role in plant defense. On the other hand, under long day conditions the expression of defense responses might depend in other genes but osmotin.

There are many markers in potato LGXII that could be used for future screening of the library. However, for a candidate gene approach a better genetic resolution of QTL in the region is necessary. Potato chromosome XII carries resistance genes (*Rx1* and *Rx2*) against PVX and a QTL for *P. infestans* resistance is also present in the same region. The resistance hot spot is located at a distal end of the chromosome far from the QTL effect of the PD population.

V Summary

QTL mapping of quantitative resistance to *P. infestans* has been pursued in potato and several loci contributing to the resistance have been identified (Leonards-Schippers et al. 1994; Meyer et al. 1998; Ewing et al. 2000; Sandbrink et al. 2000 and Naess et al. 2000). Ghislain et al. (2001) detected two major QTL effects on linkage groups VIII and XII using a hybrid cross between *S. phureja* x dihaploid *S. tuberosum*. The strong QTL effect on linkage group XII was localized in a region where no major gene or QTL for *P. infestans* has been reported. So far, none of the QTLs for *P. infestans* resistance has been cloned and the genes behind the QTLs are still unknown. Map based cloning has proved to be a promising approach for cloning genes and QTLs. This approach requires DNA markers tightly linked to the trait in combination with large insert libraries.

In this study, a large insert library was constructed from one of the resistant hybrids of the population analysed by Ghislain et al. (2001). The inserts have been cloned into the binary vector pCLD04541. The clones can be used for plant transformation via *A. tumefaciens*. The library contains approximately 50 000 clones with an average insert size of 80 kb. The coverage of the library has been calculated to be 3-4 times the haploid potato genome. The construction of a large insert library with this genetic material will facilitate the cloning and dissection of the genes underlying the QTL once a more precise map of the region is available.

The library was used to clone members of three defense related gene families: *PAL*(*Phenylalanine-Ammonia-Lyase*), *PR-5* (acidic and basic osmotin) and *PPO(Polyphenol oxidase)*. 29, 7 and 10 BAC clones containing *PAL, PR-5* and *PPO* genes, respectively were identified. The *PR-5* BAC clones were further characterized. 6 BACs containing osmotin-like sequences were grouped in two small contigs: Contig A (100kb) and Contig B (120kb). DNA markers derived from the contigs identified two genetic loci on linkage groups VIII (Contig A) and XI (Contig B). Southern gel blot analysis of genomic and BAC DNA showed that potato has at least 5 copies of osmotin genes. Most of the osmotin genes are clustered on linkage group VIII in less than 90 kb whereas only one gene is present on linkage group XI. Interestingly, Trognitz et al. (2001) using the population of Ghislain et al. (2001) reported correlation between a QTL effect derived from *S. tuberosum* and an osmotin

RFLP marker on linkage group VIII. On the other hand, one BAC with sequence similarity to acidic PR-5 showed that acidic members of PR-5 in potato are clustered on linkage group XII, where at least 3 copies are present in less than 35 kb. *S. phureja* and *S. tuberosum* are highly homozygous at these loci which made it difficult to develop specific markers. However, two CAPS markers were derived and mapped in the maternal line *S. phureja*, revealing that the position of the genes do not overlap with the QTL effect.

Fragments of the *PR-5* BACs inserts were subcloned into pBluescript. Sequence analysis of the subclones identified 6 osmotin-like genes and 4 acidic PR-5 genes, including a copy with a truncated sequence at the N-terminus. One of the acidic PR-5 genes had 100% identity to the partial sequence of potato Protein C (Pierpoint et al. 1990). All the *PR-5* ORF did not contain introns and all except the one from linkage group XI had the 16 cysteine residues highly conserved in the PR-5 family. PCR and Southern gel blot analysis demonstrated that *PR-5* genes are present in all members of the *Solanaceae* family tested. A phylogenetic analysis of PR-5 sequences of *Solanaceae* from the Genebank and the genes described in this study clustered the sequences in three main branches of acidic, neutral and basic genes.

Deutsche Zusammenfassung

In der Kartoffel wurde die QTL-Kartierung der quantitativen Resitenz gegen *P. Infestans* in den vergangenen Jahren intensiv vorangetrieben. Mehrere Loci mit Einfluss auf dieses Merkmal konnten identifiziert werden (Leonards-Schippers et al. 1994; Meyer et al. 1998; Ewing et al. 2000; Sandbrink et al. 2000 and Naess et al. 2000). So fanden Ghislain et al. (2001) unter Verwendung einer Hybridkreuzung aus *S. phureja* und der dihaploiden *S. tuberosum* zwei Haupt-QTL-Effekte auf den Kopplungsgruppen VIII and XII. Der starke QTL-Effekt auf der Kopplungsgruppe XII wurde in einer Region lokalisiert, für welche bis zu dieser Untersuchung keine Hauptgene oder QTLs beschrieben worden waren. Bisher wurden noch keine QTLs für *P. infestans* Resistenz kloniert und somit sind die Gene, welche für dieses Merkmal verantwortlich sind, unbekannt. ‚Map based cloning' hat sich als ein vielversprechender Ansatz zum Klonieren von Genen und QTLs erwiesen. Erforderlich für diese Vorgehensweise sind zum Einen DNA-Marker, welche genetisch eng mit dem Merkmal gelcoppelt sind, und zum anderen eine genomischer Bibliothek mit großen Insertionen.

Im Rahmen dieser Arbeit wurde eine solche Bibliothek mit großen Insertionen hergestellt. Hierbei wurde eine resistente Hybride aus der Population, welche Ghislain et al. (2001) analysiert hatten, genutzt. Die Insertionen wurden in den binären Vektor pCLD04541 kloniert. Diese Klone können für Pflanzentransformationen mittels *A. tumefaciens* genutzt werden. Die BAC Bibliothek umfasst ca. 50 000 Klone, rechnerisch das 3 bis 4 fache des haploiden Kartoffelgenoms. Sobald eine detailliertere Karte der QTL Region vorhanden ist, wird diese Bibliothek, erstellt mit spezifischen genetischen Material, es ermöglichen, dass die Gene, welche sich hinter dem QTL verbergen, kloniert und analyiert werden können.

Die Bibliothek wurde genutzt, um Mitglieder von drei ‚defense related'-Genfamilien zu klonieren: *PAL* (*Phenylalanine-Ammonium-Lyase*), *PR-5* (saüre und basische Osmotine) und *PPO*(*Polyphenol oxidase*). 29, 7 und 10 BAC-Klone wurden mit entsprechenden *Pal*, *PR-5* und *PPO* Sonden identifiziert. Die *PR-5* BAC-Klone wurden eingehender charakterisiert. 6 BACs mit PR-5-ähnlichen Sequenzen wurden

in zwei kleinen Contigs guppiert: Contig A (100kb) und Contig B (120kb). Mit Hilfe von DNA-Markern, welche auf Basis der Contigs entwickelt wurden, konnten zwei genetische Loci auf den Kopplungsgruppen VIII (Contig A) und XI (Contig B) identifiziert werden. ‚Southern gel blot'- Analysen von genomischer und BAC-DNA zeigten, dass die Kartoffel mindestens 5 Osmotingene enthält. Die meisten der Osmotingene sind in einem Cluster, welches weniger als 90 kb umfasst, auf der Kopplungsgruppe VIII angeordnet; auf der Kopplungsgruppe XI befindet sich dahingegen lediglich ein Osmotingen. In diesem Zusammenhang ist es interessant, dass Trognitz et al. (2001) bei Arbeiten an der Population von Ghislain et al. (2001) eine Korrelation zwischen einem QTL-Effekt, und einem Osmotin RFLP-Marker auf Linkage-Gruppe VIII feststellten. Andererseits zeigte ein weiterer BAC, welcher Sequenzähnlichkeit mit sauren PR-5 Genen aufweist, dass saure Mitglieder der PR-5 Familie bei der Kartoffel in einem Cluster auf Kopplungsgruppe XII angeordnet sind. In diesem Cluster mit einer Grösse von weniger als 35 kb sind mindestens drei Kopien enthalten. Da *S. phureja* and *S. tuberosum* an diesem Locus sehr homozygot sind, war es schwierig, spezifische Marker zu entwickeln. Dennoch konnten in der mütterlichen Linie *S. phureja* zwei CAPS-Marker abgeleitet werden. Mit diesen Marken konnte gezeigt werden, dass die Position der Gene nicht mit dem QTL-Effekt übereinstimmt.

Fragmente von Insertionen der PR-5 BACs wurden in pBluescript subkloniert. Durch Sequenzanalyse dieser Subklone wurden 6 Osmotin-ähnliche Genen und 4 saure PR-5 Gene identifiziert, wobei eine Kopie einen unvollständigen N-Terminus aufwies. Eines der sauren PR-5 Gene hatte eine 100%-ige Übereinstimmung mit einem Teil der Sequenz des Kartoffelproteins C (Pierpoint et al. 1990). Alle PR-5 ORFs enthielten keine Introns und alle, mit Ausnahme des Gens auf Kopplungsgruppe XI, hatten 16 Cystein Reste, die in der PR-5 Familie konserviert sind. Mit Hilfe von PCR und ‚Southern gel blot'- Analysen konnte gezeigt werden, dass PR-5 Gene in allen untersuchten Solanaceae Arten vorhanden sind. In phylogenetischen Untersuchungen von PR-5 Sequenzen der *Solanaceae* Arten, bei welchen sowohl Daten aus der Genbank als auch die in dieser Arbeit beschriebenen Gene eingingen, gruppierten die Sequenzen in drei Hauptästen mit sauren, neutralen und basischen PR-5 Genen.

VI References

Allouis S., Qi X., Lindup S., Gale M. and Devos K. (2001). Construction of a BAC library of pearl millet *Pennisetum glaucum*. *Theor Appl Genet* 102:1200-1205

Amemiya C., Ohta T. and Litman G, (1996). In Birren, B. and Lai, E. (eds) Nonmammalian Genome Analysis: A practical approach. Academic Press, New York, pp. 223-256

Anzlovar S., Serra M., Dermastia M. and Menestrina G. (1998). Membrane permealibilizing activity of pathogenesis-related protein linusitin from flax seed. *Mol Plant Microb Interact* 7:610-617

Arumuganathan K. and Earle E. (1991). Nuclear DNA content of some important plants species. *Plant Mol Biol Reporter* 9(3):208-218

Ballvora A., Ercolano M., Weiss J., Meksem K., Bormann C., Oberhagemann P., Salamini F. and Gebhardt C. (2002). The *R1* Gene for Potato Resistance to Late Blight (*Phytophthora infestans*) belong to the Leucine Zipper /NBS/LRR Class of Plant Resistance Genes. *The Plant Journal* (in press).

Barre A., Peumans W., Menu-Bouaouiche L., Van Damme E., May G., Fernandez Herrera A., Van Leuven F. and Rougé P. (2000). Purification and structural analysis of an abundant thaumatine-like protein from ripe banana fruit. *Planta* 211: 791-799

Batalia M., Monzingo A., Ernst S., Roberts W. and Robertus J. (1996). The crystal structure of the antifungal protein zeamatin, a member of the thaumatin-like PR-5 protein family. *Nat Struct Biol* 3:19-23

Bendahmane A. (1999). Zero-background plasmid vector for BAC library construction. *BioTechniques* 26:228-232

Beneth and Smith (1976). Nuclear DNA amounts in angiosperms. Phil. Trans. R. Soc. Lond. B 274:227-274

Bent A., Kunkel B., Dahlbeck D., Brown K., Schmidt R., Giraudat J., Leung J. and Staskawicz B. (1994). *RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. *Science* 265: 1856-1860

Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C., Bergkamp R., Dirkse W., Van Staveren M., Stiekema W., Drost L., Ridley P., Hudson S., Patel K., Murphy G., Piffanelli P., Wedler H., Wedler E., Wambutt R., Weitzenegger T., Pohl T., Terryn N., Gielen J., Villarroel R., De Clerck R., Van Montagu M., Lecharny A., Auburg S., Gy I, Kreis M., Lao N., Kavanagh T., Hempel S., Kotter P., Entian K., Rieger M., Montfort, Pons M., Puigdomenech P., Douka A., Voukelatou E., Milioni D., Hatzopoulos P., Piravandi E., Obermaier B., Hilbert H., Dusterhoft A., Moores T., Jones J., Eneva T., Palme K., Benes V., Rechman S., Ansorge W., Cooke R., Berger C., Delseny M., Voet M., Volckaert G., Mewes H., Klosterman S., Schueller C., Chalwatzis N (1998). Analysis of 1.9 Mb of contigous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature* 391:485-488

Bonierbale M., Plaisted R. and Tanksley S. (1988). RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics* 120:1095-1103

Bonierbale M., Plaisted R. Pineda O. and Tanksley S. (1994). QTL analysis of trichome-mediated insect resistance in potato. *Theor Appl Genet* 87:973-987

Boysen C., Simon M. and Hood L. (1997). Analysis of the 1.1-Mb human alpha/delta T-cell receptor locus with bacterial artificial chromosome clones. *Genome Research* 7:330-338

Brommonschenkel S. and Tanksley S. (1997). Map-based cloning of the tomato genomic region that spans the Sw-5 tospovirus resistance gene in tomato. *Mol Gen Genet* 256:121-126

Byrne P., McMullen M., Snook M., Musket T., Theuri J., Widstrom N., Wiseman B. and Coe E. (1996). Quantitative trait loci and metabolic pathways: genetic control of the concentration of maysin, a corn earworm resistance factor, in maize silks. *Proc Natl Acad Sci USA* 93:8820-8825

Cai L., Taylor J., Wing R., Gallagher D., Woo S. and Davis S. (1995). Construction and characterization of a Bovine Artificial Chromosome library. *Genomics* 29:413-425

Clarke L. and Carbon J. (1976). A colony bank containing synthetic ColE1 hybrid plasmids representative of the entire *E. coli* genome. *Cell* 9: 91-100

Collins A., Milbourne D., Ramsay L., Meyer R., Chalot-Balandras C., Oberhagemann P., De Jong W., Gebhardt C., Bonnel E. and Waugh R. (1999). QTL for field resistance to late blight in potato are strongly correlated with maturity and vigour. *Mol Breeding* 5: 387-398

Cornelissen B., Van Huijsduijnen R. and Bol J. (1986). A tobacco mosaic virus-induced tobacco protein is homologous to the sweet tasting protein thaumatin. *Nature* 321:531-532

Cutt J. and Klessing D. (1992). Pathogenesis-related proteins. In: Boller, Meins F (eds) Genes involved in plant defence. Springer, Berlin Heidelberg New York, pp209-243

Chen M., SanMiguel P., De Oliviera A., Woo S., Zhang H., Wing R. and Bennetzen J. (1997). Microcolinearity in sh2-homologuos regions of the maize, rice and sorghum genomes. *Proc Natl Acad Sci USA* 94:3431-3435

Chen M., SanMiguel P. and Bennetzen J. (1998). Sequence organization and conservation in sh2/a1-homologous regions of sorghum and rice. *Genetics* 148:435-443

Chen W., Chen P., Liu D., Kynast R., Friebe B., Velazhahan R., Muthukrishnan S. and Gill B. (1999). Development of wheat scab symptoms is delayed in transgenic wheat plants that constitutively express a rice thaumatin-like protein gene. *Theor Appl Genet* 99:755-760

Cheong N., Choi Y., Kim W., Kim S., Cho J. and Lee S. (1997). Purification of an antifungal PR-5 protein from flower buds of *Brassica campestris* and cloning of its gene. *Physiol Plant* 101:583-590

Choi S., Creelman R., Mullet J. and Wing R. (1995). Construction and characterization of a bacterial artifitial chromosome library of *Arabidopsis thaliana*. *Plant Mol Bio Rep* 13:124-128

Darby R., Firek S., Mur L. and Draper J. (2000). A thaumatin-like gene from *Asparagus officinalis* (AoPRT-L) exhibits slow activation following tissue maceration or salicylic acid treatment, suggesting convergent defence-related signalling in monocots. *Mol Plant Pathol* 1(6), 357-366

Danesh D., Penuela S., Mudge J., Denny R., Nordstrom H., Martinez J. and Young N. (1998). A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. *Theor Appl Genet* 96:196-202

Datta K., Velazhahan R., Oliva N., Ona I., Mew T., Khush G., Muthukrishnan S. and Datta S. (1999). Over-expression of the cloned rice thaumatin-like protein (PR-5) gene in transgenic rice plants enhances environmental friendly resistance to *Rhizoctonia solani* causing sheath blight disease. *Theor Appl Genet* 98:1138-1145

Deahl K., Inglis D. and DeMuth S. (1993). Testing for resistance to metalaxyl in *Phytophthora infestans* isolates from Northwestern Washington. *Am Potato J* 70:779-795

Deng Z., Tao Q., Chang Y., Huang S., Ling P., Yu C., Chen C., Gmitter F. and Zhang H. (2001). Construction of a bacterial artificial chromosome (BAC) library for citrus and identification of BAC containing resistance gene candidates. *Theor Appl Genet* 102:1177-1184

Ding Y., Johnson M., Colayco R., Chen Y., Melnyk J., Schmitt H. and Shizuya H. (1999). Cloning assembly of Bacterial Artificial Chromosome clones through multiplexed flourescence-labeled fingerprinting. *Genomics* 56:237-246

Ditta G., Stanfield S., Corbin D. and Helinski D. (1980). Broad host range DNA cloning system for Gram-negative bacteria:construction of a gene bank of *Rhizobium meliloti*. *Proc Natl Acad Sci USA* 77:347-351

Dong F., Song J., Naess S., Helgeson J., Gebhardt C., and Jiang J. (2000). Development and applications of a set of chromsome-specific cytogenetic DNA markers in potato. *Theor Appl Genet* 101:1001-1007

Drenth A., Turkensteen L. and Govers F. (1993). The occurrence of the A2 mating type of *Phytophthora infestans* in the Netherlands: significance and consequences. *Neth J Plant Path* 99 (suppl.) 3:57-67

Edwards K., Thompson H., Edwards D., De Saizieu A., Sparks C., Thompson J., Greenland A., Eyers M. and Schuch W. (1992). Construction and characterization of a yeast artifitial chromosome library containing three haploid maize genome equivalents. *Plant Mol Biol* 19:299-308

El-Kharbotly A., Leonards-Schippers C., Huigen D., Jacobsen E., Pereira A., Stiekema W., Salamini F. and Gebhardt C. (1994). Segregation analysis and RFLP mapping of the *R1* and *R3* alleles conferring race specific resistance to *Phytophthora infestans* in progenies of dihaploid potato parents. *Mol Gen Genet* 242:749-754

El-Kharbotly A., Palomino Sanchez C., Salamini F., Jacobsen E. and Gebhardt C. (1996). *R6* and *R7* alleles of potato conferring race-specific resistance to *Phytophthora infestans* (Mont.) de Bary identified genetic loci clustering with the *R3* locus on chromosome XI. *Theor Appl Genet* 92:880-884

Ewing E., Simko I., Smart C., Bonierbale M., Mizubuti E., May G. and Fry W. (2000). Genetic mapping from field tests of qualitative and quantitative resistance to Phytophthora infestans in a population derived from *Solanum tuberosum* and *Solanum berthaultii*. *Mol Breeding* 6:25-36

Faris J., Li W., Liu D., Chen P. and Gill B. (1999). Candidate gene analysis of quantitative disease resistance in wheat. *Theor Appl Genet* 98:219-225

Feinberg A. and Vogelstein B. (1984). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity (addendum). *Anal Biochem* 137: 266-267

Fritjers A., Zhang Z., van Damme M., Wang G., Ronald P. and Michelmore R. (1997). Construction of a bacterial artificial chromosome library containing large *Eco*RI and *Hin*dIII genomic fragments of lettuce. *Theor Appl Genet* 94:390-399

Folkertsma R., Spassova M., Prins M., Stevens M., Hille J. and Goldbach R. (1999). Construction of a bacterial artificial chromosome (BAC) library of *Lycopersicon esculentum* cv. Stevens and its application to physically map the Sw-5 locus. *Mol Breeding* 5:197-207

Frendo P., Didierjean L., Passelgue E. and Burkard G. (1992). Abiotic stresses induce a thaumatin-like protein in maize-cDNA isolation and sequence analysis. *Plant Science* 85:61-69

Fritzemeier K., Cretin C., Kombrink E., Rohwer F., Taylor J., Scheel D. and Hahlbrock K. (1987). Transient induction of Phenylalanine Ammonia-Lyase and 4-Coumarate: CoA Ligase mRNAs in potato leaves infected with virulent races of *Phytophthora infestans*. *Plant Physiol.* 85:34-41

Fry W., Goodwin S., Dyer A., Matuszak J., Drenth A., Tooley P., Suijkowski L., Koh Y., Cohen B., Spielman L., Deahl K., Inglis D. and Sandlan K. (1993). Historical and recent migrations of *P. infestans*: chronology and implications. *Plant Disease* 77:653-661

Gebhardt C., Ritter E., Debener T., Schachtschabel U., Walkemeier B., Uhrig H. and Salamini F. (1989). RFLP analysis and linkage mapping in *Solanum tuberosum*. *Theor Appl Genet* 78:65-75

Gebhardt C., Ritter E., Barone A., Debener T., Walkemeier B., Schachtschabel U., Kaufmann H., Thompson R., Bonierbale M., Ganal M., Tanksley S. and Salamini F. (1991). RFLP maps of potato and their alignment with the homoeologous tomato genome. *Theor Appl Genet* 83:49-57

Gebhardt C. and Valkonen J. (2001). Organization of genes controlling disease resistance in the potato genome. *Annu Rev Phytopathol* 39:79-102

Gebhardt C., Ritter E. and Salamini F. (2001). RFLP map of the potato in Vasil K. and Phillips R. (eds) DNA based markers in plants. Kluwer academic press, The Netherlands.

Geffroy V. Sicard D. de Oliveira J., Fouilloux G., Skroch P., Thoquet P., Gepts P., Langin T. and Dron M. (2000). Inheritance of partial resistance against *Colletotrichum lindemuthianum* in *Phaseolus vulgaris* and co-localization of quantitative trait loci with genes involved in specific resistance. *Mol Plant Microb Interact* 13:287-296

Gentzbittel L., Abbott A., Galaud J., Georgi L., Fabre F., Liboz T. and Alibert G. (2002). A bacterial artificial chromosome (BAC) library for sunflower, and identification of clones containing genes for putative transmembrane receptors. *Mol Gen Genomics* 266: 979-987

Ghislain M., Trognitz B., Herrera M., Solis J. Casallo G., Vasquez C., Hurtado O., Castillo R. Portal L and Orillo M. (2001). Genetic loci associated with field resistance to late blight in offspring of *Solanum phureja* and *S. tuberosum* grown under short-day conditions. *Theor Appl Genet* 103:433-442

Goodwin S., Cohen B. and Fry W. (1994). Panglobal distribution of a single clonal lineage of the Irish potato famine fungus. *Proc Nat Acad Sci USA* 91:11591-11595

Goodwin S., Sujkowski L. and Fry W. (1996). Widespread distribution and probable origin of resistance to metalaxyl clonal genotypes of *Phytophthora infestans* in the United States and Western Canada. *Phytopathology* 86:793-800

Goovers F. (2001). Misclassification of pest as 'fungus' puts vital research on wrong track. *Nature* 411:633

Grenier J., Potvin C., Trudel J. and Asselin A. (1999). Some thaumatin-like proteins hydrolyse polymeric β-1,3-glucans. *The Plant Journal* 19(4), 473-480

Grill E. and Somerville C. (1991). Construction and characterization of a yeast artificial chromosome library of *Arabidopsis* which is suitable for chromosome walking. *Theor Appl Genet* 226:484-490

Grube R., Radwanski E. and Jahn M. (2000). Comparative genetics of disease resistance within the *Solanaceae. Genetics* 155:873-887

Hamilton C., Frary A., Lewis C. and Tanksley S. (1996). Stable transfer of intact high molecular weight DNA into plant chromosomes. *Proc Natl Acad Sci USA* 93:9975-9979

Hamilton C, Frary A, Xu Y, Tanksley S. and Zhang H. (1999). Construction of tomato genomic DNA libraries in a binary BAC (BIBAC) vector. *The Plant Journal* 18(2): 223-229

Hawkes J. (1991). The Potato: Evolution, Biodiversity and Genetic Resources. Belhaven Press, UK.

Hawkes J. (1994). Origins of cultivated potato and species relationships. In Bradshaw J. and Mackay G (eds). Potato Genetics. CAB International University Press, UK.

Hu X. and Reddy A. (1997). Cloning and expression of a PR-5 like protein from Arabidopsis: inhibition of fungal growth by bacterially expressed protein. *Plant Mol Biol* 34:949-959

Hunt M., Eannetta N., Yu H., Newman S. and Steffens J. (1993). cDNA cloning and expression of a potato polyphenol oxidase. *Plant Mol Biol* 21: 59-68

Iyengar R., Smits P., Van der Ouderaa F., Van der Wel H., Brouwershaven J., Ravestein P., Richters G. and Van Wassenaar P. (1979). The complete amino acid sequence of the sweet protein thaumatin I. *Eur J Biochem* 96:193

Jacobs J., van Eck H., Arens P., Verkerk-Bakker B., Hekkert B., Bastiaanssen H., El-Kharbotly A., Pereira A., Jacobsen E. and Stiekema W. (1995). A genetic map of potato (*Solanum tuberosum*) integrating molecular markers, including transposons, and classical markers. *Theor Appl Genet* 91:289-300

Jiang J., Gill B., Wang G., Ronald P. and Ward D. (1995). Metaphase and interphase fluorescence in situ hybridisation mapping of the rice genome with bacterial artificial chromosome. *Proc Natl Acad Sci USA* 92:4487-4491

Jones J., Schlumukov L., Carland F., English J., Scofield S., Bishop G. and Harrison K. (1992). Effective vectors for transformation, expression of heterologous genes, and assaying transposon excision in transgenic plants. *Transgenic Research* 1:285-297

Joos H. and Hahlbrock K. (1992). Phenylalanine ammonia- lyase in potato (*Solanum tuberosum* L.) Genomic complexity, structural comparison of two selected genes and modes of expression. *Eur J Biochem* 204:621-629

Jung J., Maurel S., Fritig B. and Hahne G. (1995). Different pathogenesis-related –proteins are expressed in sunflower (*Helianthus-annuus*) in response to physical, chemical and stress factors. *J Plant Physiol* 145:153-160

Kamoun S. (2001). Nonhost resistance to *P. infestans*: novel prospects for a classical problem. *Current Opinion in Plant Biology* 4:295-300

Kleine M., Michalek W., Graner A., Herrmann R. and Jung C. (1993). Construction of a barley (*Hordeum vulgare* L.) YAC library and isolation of a Hor1-specific clone. *Mol Gen Genet* 240:265-272

Koiwa H., Sato F. and Yamada Y. (1994). Characterization of accumulation of tobacco PR-5 proteins by IEF immunoblot analysis. *Plant Cell Physiol* 35(5):821-827

Koiwa H., Kato H., Kato H., Nakatsu T., Oda J., Yamada Y. and Sato F. (1997). Purification and characterization of tobacco pathogenesis related protein PR-5d, an antifungal thaumatin-like protein. *Plant Cell Physiol* 38(7):783-791

Koiwa H., Kato H., Nakatsu T., Oda J., Yamada Y. and Sato F. (1999). Crystal structure of tobacco PR-5d protein at 1.8 Å resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like proteins. *J Mol Biol* 286:1137-1145

Kombrink E. and Sommsich I. (1995). Defense responses of plants to pathogens. *Advances in Botanical Research* Vol 21:1-34

Kowalski S., Eanneta N., Hirzel A. and Steffens J. (1992). Purification and characterization of polyphenol oxidase from glandular trichomes of *Solanum berthaultii*. *Plant Physiol* 100:677-684

Kuhl J., Hanneman R. and Havey M. (2001). Characterization and mapping of *Rpi1*, a late blight resistance locus from diploid (1EBN) Mexican *Solanum pinnatisectum*. *Mol Gen Genomics* 265(6):977-985

Leister D., Ballvora A., Salamini F. and Gebhardt C. (1996). A PCR-based approach for isolating resistance genes from potato with potential for wide application in plants. *Nature Genetics* 14:421-429

Leister D., Berger A., Thelen H., Lehmann W., Salamini F. and Gebhardt C. (1997). Construction of a potato YAC library and identification of clones linked to the disease resistance loci *R1* and *Gro1*. *Theor Appl Genet* 95:954-960

Leonards-Schippers C., Gieffers W., Salamini F. and Gebhardt C. (1992). The *R1* gene conferring race-specific resistance to *Phythopthora infestans* in potato is located on potato chromosome V. *Mol Gen Genet* 233:278-283

Leonards-Schippers C., Gieffers W., Schäfer-Pregl R., Ritter E., Knapp S., Salamini F. and Gebhardt C. (1994). Quantitative resistance to *Phythopthora infestans* in potato: a case-study for QTL mapping in an allogamous plant species. *Genetics* 137: 67-77

Li X., van Eck H., Rouppe van der Voort J., Huigen D., Stam P., and Jacobsen E. (1998). Autotetraploids and genetic mapping using common AFLP markers: the *R2* allele conferring resistance to *P. infestans* mapped on potato chromosome 4. *Theor Appl Genet* 96:1-8

Li W., Faris J., Chittoor J., Leach J., Hulbert S., Liu D., Chen P. and Gill B. (1999). Genomic mapping of defense response genes in wheat. *Theor Appl Genet* 98:226-233

Lin Y., Zhu L., Ren S., Yang J., Schertz K. and Paterson A. (1999). A *Sorghum propinquum* BAC library, suitable for cloning genes associated with loss-of-function mutations during crop domestication. *Mol Breed* 5:511-520

Liu D., Raghothama K., Hasegawa P. and Bressan R. (1994). Osmotin overexpression in potato delays development of disease symptoms. *Proc Natl Acad Sci* Vol. 91 1888-1892

Liu D., Narasimhan M., Xu Y., Raghothama K., Hasegawa P. and Bressan R. (1995). Fine structure and function of the osmotin gene promoter. *Plant Mol Biol* 29: 1015-1026

Liu D., Rhodes D., Paino D' Urzo M., Xu Y., Narasimhan M., Hasegawa P., Bressan R. and Abad L. (1996). *In vivo* and *in vitro* activity of truncated osmotin that is secreted into the extracellular matrix. *Plant Science* 121:123-131

Liu Y., Shirano Y., Fukaki H., Yanai Y., Tasaka M., Tabata S. anad Shibata D. (1999). Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artifitial chromosome vector accelerates positional cloning. *Proc Natl Acad Sci USA* 96:6535-6540

Livingstone K., Lackney V., Blauth J., Wijk R. and Jahn M. (1999). Genome mapping in *Capsicum* and the evolution of genome structure in the *Solananeae*. *Genetics* 152:1183-1202

Malcolmson J. and Black W. (1966). New R genes in *Solanum demissum* Lindl. and their complementary races of *Phytophthora infestans* (Mont.) de Bary. *Euphytica* 15:199-203

Marek L. and Shoemaker R. (1997). BAC contig development by fingerprint analysis in soybean. *Genome* 40:420-427

Martin G., Ganal M. and Tanksley S. (1992). Construction of a yeast artificial chromosome library of tomato and identification of cloned segments linked to two disease resistance loci. *Mol Gen Genet* 233:25-32

Meksem K., Zobrist K., Ruben E., Hyten K., Quanzhuo T., Zhang H. and Lightfoot D. (2000). Two large-insert soybean genomic libraries constructed in a binary vector: applications in chromosome walking and genome wide physical mapping. *Theor Appl Genet* 101:747-755

Melchers L., Sela-Buurlage M., Vloemans M., Woloshuk C., Van Roeckel, Pen J., van den Elzen P. and Cornelissen B. (1993). Extracellular targeting of the vacuolar tobacco proteins AP24, chitinase and β-1,3-glucanase in transgenic plants. *Plant Mol Biol* 21:583

Meyer R., Milbourne D., Hackett C., Bradshaw J., McNichol J. and Waugh R. (1998). Linkage analysis in tetraplooid potato and association of markers with quantitative resistance to late blight (*Phytophthora infestans*). *Mol Gen Genet* 259:150-160

Ming R., Moore P., Zee F., Abbey C., Ma H. and Paterson A. (2001). Construction and characterization of a papaya BAC library as a foundation for molecular dissection of a tree-fruit genome. *Theor Appl Genet* 102:892-899

Moullet O., Zhang H. and Lagudah E. (1999). Construction and characterization of a large DNA insert library from the D genome of wheat. *Theor Appl Genet* 99:305-313

Naess S., Bradeen J., Wielgus S., Haberlach G., McGrath J. and Helgeson J. (2000). Resistance to late blight in *Solanum bulbocastanum* is mapped to chromosome 8. *Theor Appl Genet* 101:697-704

Nam Y., Penmetsa R., Endre G., Uribe P, Kim D. and Cook D. (1999). Construction of a bacterial artificial chromosome library of *Medicago truncatula* and identification of clones containing ethylene-response genes. *Theor Appl Genet* 98:638-646

Nakamura S., Asakawa S., Ohmido N., Fukui K., Shimizu N. and Kawasaki S. (1997). Construction of an 800-kb contig in the near centromeric region of the rice blast resistance gene *Pi-ta*^{2} using a highly representative rice BAC library. *Mol Gen Genet* 254:611-620

Newton S. and Duman J. (2000). An osmotin-like cryoprotective protein from the bittersweet nightshade *Solanum dulcamara*. *Plant Mol Biol* 44:581-589

Oberhagemann P., Chalot-Balandras C., Schäfer-Pregl R., Wegener D., Palomino C., Salamini F., Bonnel E. and Gebhardt C. (1999). A genetic analysis of quantitative resistance to late blight in potato: towards marker –assisted selection. *Mol Breed* 5:399-415

O'Sullivan D., Ripoll P., Rodgers M. and Edwards K. (2001). A maize bacterial artificial chromosome (BAC) library from the European flint inbred line F2. *Theor Appl Genet* 103:425-432

Payne G. Middlesteadt W., Williams S., Desai N., Parks D., Dincher S., Carnes M. and Ryals J. (1988). Isolation and nucleotide sequence of a novel cDNA clone encoding the major form of pathogenesis – related protein R. *Plant Mol Biol* 11: 223-224

Perez E., Menendez A., Dehal P. and Quiros C. (1999). Genomic structural differentiation in *Solanum*:comparative mapping of the A-and E-genomes. *Theor Appl Genet* 98:1183-1193

Pierpoint W., Tathan A. and Pappin D. (1987). Identification of the virus-induced protein of tobacco leaves that resembles the sweet-tasting protein thaumatin. *Physiol Mol Plant Path* 31:291-298

Pierpoint W., Jackson P. and Evans R. (1990). The presence of a thaumatin-like protein, a chitinase and a glucanase among the pathogenesis-related proteins of potao (*Solanum tuberosum*). *Physiological and Molecular Plant Pathology* 36:325-338

Pierpoint W., Gordon-Weeks R. and Jackson P. (1992). The occurrence of the thaumatin-like, pathogenesis-related protein, PR-5, in intercellular fluids from *Nicotiana* species and from an interspecific *Nicotiana* hybrid. *Physiological and Molecular Plant Pathology* 41:1-10

Pflieger S., Lefebvre V., Caranta C., Blattes A., Goffinet B. and Palloix A. (1999). Disease resistance gene analogs as candidates for QTLs involved in pepper/pathogen interactions. *Genome* 42:1100-1110

Pflieger S., Lefebvre V. and Causse M. (2001a). The candidate gene approach in plant genetics: a review *Mol Breeding* 7:275-291

Pflieger S., Palloix A., Caranta C., Blattes A., Lefebvre V. (2001b). Defense response genes co-localize with quantitative disease resistance loci in pepper. *Theor Appl Genet* 103:920-929

Quanzhou T. and Zhang H. (1998). Cloning and stable maintenance of DNA fragments over 300 kb in *Escherichia coli* with conventional plasmid-based vectors. *Nucleic Acids Res* Vol. 26 N°21, 4901-4909

Rebmann G., Mauch F., and Dudler R. (1991). Sequence of a wheat cDNA encoding a pathogen-induced thaumatin-like protein. *Plant Mol Biol* 17:283

Ristaino J., Groves C. and Parra G. (2001). PCR amplification of the Irish potato famine pathogen from historic specimens. *Nature* 411:965-697

Roberts K. and Selitrennikoff C. (1990). Zeamatin, an antifungal protein from maize with membrane-permeabilizing activity. *J Gen Microbiol* 136:1771-1778

Salimath S. and Bhattacharyya M. (1999). Generation of a soybean BAC library, and identification of DNA sequences tightly to the *Rps1*-k disease resistance gene. *Theor Appl Genet* 98:712-720

Sambrook J., Fritsch E. and Maniatis T. (1989). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Sandbrink J., Colon L., Wolters P. and Stiekema W. (2000). Two related genotypes of *Solanum microdontum* carry different segregating alleles for field resistance to *Phytophthora infestans*. *Mol Breeding* 6: 215-225

Sanguinetti C., Neto E., Simpson A. (1994). Rapid silver staining and recovery of PCR products separated on polyacrilamide gels. *Biotechniques* 17:915-919

Sato F., Kitajima S., Koyama T. and Yamada Y. (1996). Ethylene-induced gene expression of osmotin-like protein, a neutral isoform of tobacco PR-5, is mediated by the AGCCGCC cis-sequence. *Plant Cell Physiol* 37:249

Scheweizer P., Buchala A., Dudler R. and Metraux J. (1998). Induced systemic resistance in wounded rice plants. *The Plant Journal* 14:475-481

Shih C., Wu J., Jia S., Khan A., Ting K. and Shih D. (2001). Purification of an osmotin-like protein from the seeds of *Benincasa hispida* and cloning of the gene encoding this protein. *Plant Science* 160:817-826

Shizuya H., Birren B., Kim U., Mancino V., Slepak T., Tachiiri Y. and Simon M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor based vector. *Proc Natl Acad Sci USA* 89:8794-8797

Singh N., Handa A., Hasegawa P. and Bressan R. (1985). Proteins associated with adaptation of cultured tobacco cells to NaCl. *Plant Physiol* 79:126-137

Skadsen R., Sathish P. and Kaeppler H. (2000). Expression of thaumatin-like permatin PR-5 genes switches from the ovary wall to the aleurone in developing barley and oat seeds. *Plant Science* 156:11-22

Somssich I. (1994). Regulatory elements governing pathogenesis-related (PR) gene expression. In: L. Nover (ed.). Plant promoters and transcription factors. Springer-Verlag.

Song W., Wang G., Chen L., Kim H., Pi L., Gardner J., Wang B., Holsten T., Zhai W., Zhu L., Fauquet C. and Ronald P. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene Xa21. *Science* 270:1804-1806

Staskawicz B., Ausubel F., Baker B., Ellis J. and Jones J. (1995). Molecular genetics of plant disease resistance. *Science* 268: 661-667

Strong S., Ohta Y., Litman and Amemiya T. (1997). Marked improvement of PAC and BAC cloning is achived using electroelution of pulsed-field gel- separated partial digests of genomic DNA. *Nucleic Acids Research* 25(19): 3959-3961

Tanksley S., Ganal M., Prince J., De Vicente M., Bonierbale M., Broun P., Fulton T., Giovannoni J., Grandillo S., Martin G., Messeguer R., Miller J., Miller L., Paterson A., Pineda O., Röder M., Wing R., Wu W. and Young N. (1992). High density molecular linkage map of tomato and potato genomes. *Genetics* 132:1141-1160

Tanksley S., Ganal M. and Martin G. (1995). Chromosome landing: a paradigm for map-based cloning in plants with large genomes. *Trends Genetics* 2:63-68

Taylor J., Fritzemeier I., Hauser E., Kombrink F., Rohwer F., Schroder M., Strittmatter G. and Hahlbrock K. (1990). Structural analysis and activation by fungal infection of a gene encoding a pathogenesis-related protein in potato. *Mol Plant Microbe Interact* 3:72-77

Thygesen P., Dry I. and Robinson S. (1995). Polyphenoloxidase in Potato: A multigene family that exhibits differential expression patterns. *Plant Physiol* 109: 525-531

Tikhonov A., SanMiguel P., Nakajima Y., Gorenstein N., Bennetzen J. and Avramova Z. (1999). Colinearity and its expectations in orthologous adh regions of maize and sorghum. *Proc Natl Acad Sci USA* 96:7409-7414

Tomkins J., Mahalingam R., Smith H., Goicoechea J., Knap H. and Wing R. (1999a). A bacterial artificial chromosome library for soybean PI 437654 and identification of clones associated with cyst nematode resistance. *Plant Mol Biol* 41:25-32

Tomkins J., Yu Y., Miller-Smith H., Frisch D., Woo S. and Wing R. (1999b). A bacterial artificial chromosome library for sugarcane. *Theor Appl Genet* 99:419-424

Trognitz F., Manosalva P., Niño-Liu D., Herrera M., Ghislain M., Trognitz B., and Nelson R. (2002). Plant defense genes associated with quantitative resistance to potato late blight in *Solanum phureja* x *Solanum tuberosum* hybrids. *Mol Plant Microb Interact* (in press)

Vanhouten W. and MacKenzie S. (1999). Construction and characterization of a common bean artificial chromosome library. *Plant Mol Biol* 40:977-983

Van Huijsduijen H., Kaufmann S., Brederode F., Legrand M., Fritig B. and Bol J. (1987). Homology between chitinases that are induced by TMV infection of tobacco. *Plant Mol Biol* 9:411-420

Van Kan J., Van de Rhee M., Zuidema D., Cornelissen B. and Bol J. (1989). Structure of tobacco genes encoding thaumatin-like proteins. *Plant Mol Biol* 12:153-155

Van Loon L. and van Kammen (1970). Polyacrylamide disc electrophoresis of the soluble leaf proteins from Nicotiana tabacum var. Samsun NN II. Change in protein constitution after infection with tobacco mosaic virus. *Virology* 40:199-221

Van Loon L. (1985). Pathogenesis-related proteins. *Plant Mol Biol* 4:111-116

Van Loon L., Pierpoint W., Boller T. and Conejero V. (1994). Recommendations for naming plant pathogenesis proteins. *Plant Mol Biol Rep* 12:245-264

Van Loon L. (1999). Ocurrence and Properties of Plant Pathogenesis Related Proteins. In : Datta and Muthukrishnan (eds). Pathogenesis Related proteins in Plants. CRC Press, USA

Velzhahan R., Datta S. and Muthukrishnan S. (1999). The PR-5 family: Thaumatin-like proteins. In Datta and Muthukrishnan (eds). Pathogenesis Related proteins in Plants. CRC Press Printed in USA

Vigers A., Wiedermann S., Roberts W., Legrand M., Selitrennikoff C. and Fritig B. (1992). Thaumatin-like pathogenesis-related proteins are antifungal. *Plant Science* 83:155-161

Vinatzer B., Zhang H. and Sansavini (1998). Construction and characterization of a bacterial artifitial chromosome library of apple. *Theor Appl Genet* 97:1183-1190

Vleeshouwers V., van Dooijeweert W., Keizer P., Sijpkes L., Govers F. and Colon L. (1999). A laboratory assay for *Phytophthora infestans* resistance in various *Solanum* species reflects the field situation. *European Journal of Plant Pathology* 105:241-250

Wang Y. and Schwartz D. (1993). Chopped inserts: a convinient alternative to agarose inserts or beads. *Nucleic Acids Research* 21(10)

Wang G., Holsten T., Song W., Wang H. and Ronald P. (1995). Construction of a rice bacterial artificial chromosome library and identification of clones linked to the *Xa-21* disease resistance locus. *The Plant Journal* 7(3):525-533

Wang G., Warren R., Innes R., Osborne B., Baker B. and Ronald P. (1996). Construction of an *Arabidopsis* BAC library and isolation of clones hybridizing with disease-resistance, gene-like sequences. *Plant Mol Biol Rep* 14:107-114

Wang X., Zafian P., Chudharx M., and Lawton M. (1996). The PR5K receptor kinase from *Arabidopsis thaliana* is structurally related to a family of plant defense proteins. *Proc Natl Acad Sci* USA 93:2598-2602

Wang Z., Taramino G., Yang D., Liu G., Tingey S., Miao G. and Wang G. (2001). Rice ESTs with disease-resistance gene- or defense-response gene-like sequences mapped to regions containing major resistance genes or QTLs. *Mol Genet Genomics* 265:302-310

Wang Q., Zhang K., Qu X., Jia J., Shi J., Jin D. and Wang B. (2001). Construction and characterization of a bacterial artificial chromosome library of peach. *Theor Appl Genet* 103:1174-1179

Wang W., Zhai W., Luo M., Jiang G., Chen X., Li X., Wing R. and Zhu L. (2001). Chromosome landing at the bacterial blight resistance gene *Xa4* locus using a deep coverage rice BAC library. *Mol Genet Genomics*. 265:118-125

Woo S., Jiang J., Gill B., Paterson A. and Wing R. (1994). Construction and characterization of a bacterial artificial chromosome library of *Sorghum bicolor*. *Nucleic Acids Research* 22, 4922-4931

Woloshuk C., Meulenhoff J., Sela-Buurlage M., van den Elzen P. and Cornelissen B. (1991). Pathogen-Induced proteins with inhibitory activity toward *Phytophthora infestans*. *The Plant Cell* 3:619-628

Yang D., Parco Y., Nandi S., Subundhi P., Zhu Y., Wang G. and Huang N. (1997). Construction of a bacterial artificial chromosome (BAC) library and identification of overlapping BAC clones with chromosome specific RFLP markers in rice. *Theor Appl Genet* 95: 1147-1154

Yang D., Sanchez A., Khush G., Zhu Y. and Huang N. (1998). Construction of a BAC contig containing the *xa5* locus in rice. *Theor Appl Genet* 97: 1120-1124

Yu Y., Tomkins J., Waugh R., Frish D., Kudrna D., Kleinhofs A., Brueggeman R., Muehlbauer G., Wise and Wing R. (2000). A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. *Theor Appl Genet* 101:1093-1099

Yun D., Zhao Y., Pardo J., Narashimhan M., Damsz B., Lee H., Abad L., D'Urzo M., Hasegawa P., and Bressan R. (1997). Stress proteins on the yeast cell surface determine resistance to osmotin, a plant antifungal protein. *Proc Natl Acad Sci USA* 94, 7082-7087

Zhang H., Zhao X., Ding D., Paterson A. and Wing R. (1995). Preparation of megabase-size DNA from plant nuclei. *The Plant Journal* 7:175-184.

Zhang H., Choi S., Woo S. Li Z. and Wing R. (1996). Construction and characterization of two rice bacterial artificial chromosome libraries from the parents of a permanent recombinant inbred mapping population. *Mol Breeding* 2:11-24

Zhu B., Chen T. and Li P. (1995a). Expression of three osmotin-like protein genes in response to osmotic stress and fungal infection in potato. *Plant Mol Biol* 28: 17-26

Zhu B., Chen T. and Li P. (1995b). Activation of two osmotin-like proteins genes by abiotic stimuli and fungal pathogen in transgenic potato plants. *Plant Physiol* 108: 929-937

Zhu B., Chen T. and Li P. (1996). Analysis of late-blight disease resistance and freezing tolerance in transgenic potato plants expressing sense and antisense genes for an osmotin-like protein. *Planta* 198:70-77

VII Appendices

7.1 Osmotin-like BAC end primers

Primer name	Primer sequence	Size (bp)	Tm
18h19T3for	5'-gct tca cca tga aac ggg tct aca t-3'	800	58
18h19T3rev	5'-cct ttc gac tta cat gac cat tac g-3'		
18h19T7for	5'-cat tcc ata cca aga cca gtc gtc a-3'	700	55
18h19T7rev	5'-agc ttc taa gaa atg tga cat cat c-3'		
41d14T3for	5'-ggt cca aca aat gat ctg tta cat c-3'	650	56
41d14T3rev	5'-cct tgg agc att gat gac cca tgt-3'		
41d14T7for	5'-cca cta tag ata cag ggt tgg a-3'	500	58
41d14T7rev	5'-tga aac gta ggc ttg agt ctc ca-3'		
48h15T3for	5'-gac cgt gtg cac gcg gag gag-3'	1000	56
48h15T3rev	5'-tga att cgc atg tta atg gtt agc-3'		
48h15T7for	5'-aga cca atg atg gag atg tat acg-3'	650	56
48h15T7rev	5'-cac acc tgg taa gga ttt aca tca-3'		
82a11T3for	5'-ctg tga ccc tat ttg tta ctc gat-3'	350	56
82a11T3rev	5'-cat gtt agg aga gtg ttc caa ttc-3'		
82a11T7for	5'-ggt atg tga tgc cag ctt aac cac-3'	1200	58
82a11T7rev	5'-agt cat ggg aat tgg aac gat ggt-3'		
68311	5'-tga tgg tgc tgg tag agg ttc ttg-3'	700	58
68321	5'-aag acc gat tgc ctg aag cat tag-3'		
86e10T3for	5'-gct cta aac tgt cga tga acc-3'	900	56
86e10T3rev	5'-gat gcg gac gaa ttt gag gac-3'		
86e10T7for	5'-tgg tat atg tac tga gcc tcg ata-3'	800	56
86e10T7rev	5'-gga aat aac gac acg caa tta acg		
99n12T7for	5'-cac act tgc tgt aaa tga gat agg-3'	750	56
99n12T7rev	5'-gtc cga ata tgg aca tag acg-3'		

7.2 *PR-5* CAPS primers

Primer name	Primer sequence	Size (bp)	Tm	Restriction enzyme
319prf	5'-gga cac gaa cct aca att gca gat g-3'	400	56	*Ava*II
319prr	5'-ggt aat gtc aaa agt ggc agc atg agt-3'			
321E22f	5'-cac aca aca att acg ata ctc ctt c-3'	500	56	*Apo*I
321E22r	5'-gag atg tcc tta aac aga gag agt ag-3'			

7.3 Nucleotide sequence of *PR-5*-like genes of potato

7.3.1. Genomic sequence of 21a

```
   1  gttattatta tcacatttat tcttaaagca aac
                                 21a ATGCGCTTCATCAAATC
                                      M  R  F  I  K  S
  51  CTTCTTCATATCCCTCTATATTATAACTTCTCATATCTCAATACAAGCAA
        F  F  I  S  L  Y  I  I  T  S  H  I  S  I  Q  A
 101  ACGCAATAATCGACATACAAAATAATTGTCCCTTCACGGTATGGGCAGCA
      N  A  I  I  D  I  Q  N  N  C  P  F  T  V  W  A  A
 151  GCCGTCCCCGGTGGGGGACGACAACTCGAGTATGGTGATACATGGAAAAT
      A  V  P  G  G  G  R  Q  L  E  Y  G  D  T  W  K  I
 201  CGAGCCAGACATGACAACGAGCTCAACTAAAAAGGGTCGTATTTGGGGTC
        E  P  D  M  T  T  S  S  T  K  K  G  R  I  W  G
 251  GAACCAATTGTAATTTTGACTCATTAGGCCACAGTCAATGCCAAACGGGG
      R  T  N  C  N  F  D  S  L  G  H  S  Q  C  Q  T  G
 301  GATTGCAATGGACTGCTCGAATGTCAAACTTTTAGTAATACACCCCCAAA
      D  C  N  G  L  L  E  C  Q  T  F  S  N  T  P  P  N
 351  CACTTTGGCTGAAAACGCTTTAAACCAATACAACAATAATGATTTCTTCG
        T  L  A  E  N  A  L  N  Q  Y  N  N  N  D  F  F
 401  ACTTATCTCTTGTTGAGGGTTTCAACGTCCCAATGGAATTTAGTCCTGTG
      D  L  S  L  V  E  G  F  N  V  P  M  E  F  S  P  V
 451  TTCGCGGATGAGTGCTCTACTAGGATCAGATGCACCGCGGACATCATAGG
      F  A  D  E  C  S  T  R  I  R  C  T  A  D  I  I  G
 501  ACAATGTCCAAATGAATTAAGGACACCTGGAGGATGCAACAATCCTTGTA
        Q  C  P  N  E  L  R  T  P  G  G  C  N  N  P  C
 551  CGATTTTCAAGACGGACGAATATTGTTGCACGTCTGGAAATTGTGGTCCG
      T  I  F  K  T  D  E  Y  C  C  T  S  G  N  C  G  P
 601  ACCAACTATTCGAGGTTCTTTAAGGATAGATGTTCAACTTCTTTTAGTTA
      T  N  Y  S  R  F  F  K  D  R  C  S  T  S  F  S  Y
 651  TCCTGCAGATGATTTAAGTAGTCTCTTCACTTGTCCTAGTGGAACTACCT
        P  A  D  D  L  S  S  L  F  T  C  P  S  G  T  T
 701  ATAGAGTTGTCTTTTGTCCGTAA
      Y  R  V  V  F  C  P  *
                           cgtacat catacataaa tatgtttttt
```

7.3.2. Genomic sequence of 319b

```
   1  aataaacaca ttctcaagtt taacaacaca aaaaaaa
                                 319b ATGCATTTCCTCA
                                      M  H  F  L
  51  AGTTTTTCCCCCTTTTTGTCTTCCTTTATTTTGGTCAATACTATTTATAT
      K  F  F  P  L  F  V  F  L  Y  F  G  Q  Y  Y  L  Y
 101  GTCACTCATGCTGCCACTTTTGACATTACCAATCGGTGCACCTACCCGGT
      V  T  H  A  A  T  F  D  I  T  N  R  C  T  Y  P  V
 151  CTGGGCCGCTGCCTCTCCTGGGGGAGGCAGAAGACTCGACTCGGGCCAAA
        W  A  A  A  S  P  G  G  G  R  R  L  D  S  G  Q
 201  CTTGGAACCTTAATGTGAACCCAGGAACAATCCAGGCTCGCATTTGGGGC
      T  W  N  L  N  V  N  P  G  T  I  Q  A  R  I  W  G
 251  CGAACCAATTGTAACTTTGATGGTAGTGGCCGAGGCAAATGCGAGACGGG
      R  T  N  C  N  F  D  G  S  G  R  G  K  C  E  T  G
 301  AGATTGTAATGGGCTGTTAGAATGTCAAGGTTATGGTAGTCCACCCAACA
        D  C  N  G  L  L  E  C  Q  G  Y  G  S  P  P  N
```

```
351  CTTTAGCCGAATTTGCTCTAAATCAGCCCAATAACTTGGATTTTGTCGAT
      T L  A  E  F  A  L  N  Q  P  N  N  L  D  F  V  D
401  ATTTCTCTAGTCGATGGATTCAACATTCCTATGGAATTCAGCCCAATCAA
      I  S  L  V  D  G  F  N  I  P  M  E  F  S  P  I  N
451  CGGTGGGTGTCGTAATCTCTTATGCAACGCACCTATTAACGACCAATGTC
        G  G  C  R  N  L  L  C  N  A  P  I  N  D  Q  C
501  CAAACGAATTACGGACACCGGGTGGGTGTAACAACCCGTGCACGGTTTTC
      P  N  E  L  R  T  P  G  G  C  N  N  P  C  T  V  F
551  AAGACGAATGAATTTTGTTGTACAAATGGGCCAGGGTCATGTGGTCCTAC
      K  T  N  E  F  C  C  T  N  G  P  G  S  C  G  P  T
601  TGATTTATCGAGATTTTTCAAACAAAGATGCCCAGATGCTTATAGTTATC
        D  L  S  R  F  F  K  Q  R  C  P  D  A  Y  S  Y
651  CACAAGATGATCCAACAAGTTTGTTTACATGTCCTGCTGGAACAAATTAC
      P Q  D  D  P  T  S  L  F  T  C  P  A  G  T  N  Y
701  AAGGTTGTCTTTTGCCCTTGA
      K  V  V  F  C  P  *
                           aggctacaa aattagaact atatagtact
```

7.3.3 Genomic sequence of 321c

```
  1  cataaaagta at
           321c ATGCATTTCCTCAAGTTTTCCCCCCTTTTTGTCTTCCT
                    M  H  F  L  K  F  S  P  L  F  V  F  L
 51  TTACTTTTTATCTGTTACTCATGCTGCCACTTTTGACATTACCAATCGAT
       Y  F  L  S  V  T  H  A  A  T  F  D  I  T  N  R
101  GCACCTACCCGGTCTGGGCCGCTGCCTCTCCTGGGGGAGGCAGACGACTC
     C  T  Y  P  V  W  A  A  A  S  P  G  G  G  R  R  L
151  GACTTGGGCCAAACTTGGAACATTAATGTGAATCCAGGAACAACCATGGG
      D  L  G  Q  T  W  N  I  N  V  N  P  G  T  T  M  G
201  TCGCATTTGGGGTAGAACCAATTGTAACTTTGATGGTAGTGGTCGAGGCA
       R  I  W  G  R  T  N  C  N  F  D  G  S  G  R  G
251  AATGCCAGACTGGAGATTGTAATGGGCGGCTAGAATGTCAAGGTTTTGGT
     K  C  Q  T  G  D  C  N  G  R  L  E  C  Q  G  F  G
301  ACTGTACCCAATACTTTAGCTGAATTTGCACTTAACCAGCCCAATAACTT
      T  V  P  N  T  L  A  E  F  A  L  N  Q  P  N  N  L
351  GGACTTTGTCGATATTTCTCTAGTCGATGGATTCAATATCCCTATGGAAT
       D  F  V  D  I  S  L  V  D  G  F  N  I  P  M  E
401  TCAGCCCAATCAACGGTGGGTGTCGCAAAATCCGATGCTCAGCAGATATT
      F  S  P  I  N  G  G  C  R  K  I  R  C  S  A  D  I
451  AATGGGCAATGCCCAAGCGAATTACGGGCACCTGGTGGATGTAACAACCC
      N  G  Q  C  P  S  E  L  R  A  P  G  G  C  N  N  P
501  GTGCACAGTTTTCAAGATGAAGGAATTTTGTTGTACAAATGGACCAGGGT
       C  T  V  F  K  M  K  E  F  C  C  T  N  G  P  G
551  CATGTGGTCCTACTGATTATTCGAAATTTTTCAAACAAAGATGCCCAGAT
     S  C  G  P  T  D  Y  S  K  F  F  K  Q  R  C  P  D
601  GCTTATAGTTATCCACTGGATGATCCAACAAGTATGTTTACATGTCATGC
      A  Y  S  Y  P  L  D  D  P  T  S  M  F  T  C  H  A
651  TGGTACAAATTACAAGGTTGTCTTCTGCCCAGGTAATACATTTACTCCGT
       G  T  N  Y  K  V  V  F  C  P  G  N  T  F  T  P
701  TCAATTTCTTTTACTTGTCACGTTTAGATTTGACAAATCTATTAAGAAAA
      F N  F  F  Y  L  S  R  L  D  L  T  N  L  L  R  K
751  TAA
      *
        ataattt gagaaataag taattaatat ttgaatattt gtcctcaaac
```

7.3.4. Genomic sequence of 534a

```
   1  aaataatttg tccacatata aatttcaaca aac
                              534a  ATGGGCTACTTGAGATC
                                      M  G  Y  L  R  S
  51  TTCTTTTGTTTTCTCCCTTCTTGCTTTTGTGACTTTCACTTATGCTGTCA
        S  F  V  F  S  L  L  A  F  V  T  F  T  Y  A  V
 101  CTATCGAGGTACGCAACAACTGTCCATACATCGTCTGGGCGGCATCGACC
       T  I  E  V  R  N  N  C  P  Y  I  V  W  A  A  S  T
 151  CCGATAGGCGGTGGCCGACGTCTCGATCGAGGCCAAACGTGGGTTATTGA
       P  I  G  G  G  R  R  L  D  R  G  Q  T  W  V  I  D
 201  TGCGCCTAGGGGCACTACGATGGCACGTATATGGGGTCGTACTGGATGTA
        A  P  R  G  T  T  M  A  R  I  W  G  R  T  G  C
 251  ACTTCGATGGTGCTGGTAGAGGTTCGTGCATGACTGGTGATTGTGGTGGT
       N  F  D  G  A  G  R  G  S  C  M  T  G  D  C  G  G
 301  GTCCTGCAGTGCACCGGGTGGGGCAAACCACCAAACACCCTAGCTGAGTA
       V  L  Q  C  T  G  W  G  K  P  P  N  T  L  A  E  Y
 351  CGCCTTGGACCAGTTCAGCAACTTAGATTTCTGGGACATTTCTTTAGTTG
        A  L  D  Q  F  S  N  L  D  F  W  D  I  S  L  V
 401  ATGGATTTAATATACCGATGACTTTTGCCCCGACTAATCCAAGTGGAGGG
       D  G  F  N  I  P  M  T  F  A  P  T  N  P  S  G  G
 451  AAATGCCATGCAATTCAATGCACGGCAAATATAAACGGTGAATGTCCTGG
       K  C  H  A  I  Q  C  T  A  N  I  N  G  E  C  P  G
 501  TCAACTTAAAGTACCCGGAGGATGTAACAATCCTTGTACCACGTTTGGAG
        Q  L  K  V  P  G  G  C  N  N  P  C  T  T  F  G
 551  GACAACAATATTGTTGCACCCAAGGTCCATGTGGTCCTACGTTCTTTTCC
       G  Q  Q  Y  C  C  T  Q  G  P  C  G  P  T  F  F  S
 601  AATTTTTTCAAACAAAGATGTCCTGATGCGTATAGCTACCCACAAGATGA
       N  F  F  K  Q  R  C  P  D  A  Y  S  Y  P  Q  D  D
 651  TCCTACTAGCACATTTACTTGTCCTAGTGGTAGTACAAACTATAGGGTTG
        P  T  S  T  F  T  C  P  S  G  S  T  N  Y  R  V
 701  TCTTTTGTCCTAATGGTGTTGCTGGCCCAAATTTCCCCTTGGAGATGCGT
       V  F  C  P  N  G  V  A  G  P  N  F  P  L  E  M  R
 751  GCAAGTACTGATGAAGTGGCCAAGTAA
       A  S  T  D  E  V  A  K  *
                                    atg agaaattttg ctctcttata
```

7.3.5 Genomic sequence of 517b

```
   1  caaattaac
        517b  ATGGGCTACTTGAGATCTTCTTTTATTTTCTCCCTTCTTGC
                 M  G  Y  L  R  S  S  F  I  F  S  L  L  A
  51  TTTTGTGACTTACACTTATGCTGCCACTATCGAGGTACGCAACAACTGTC
        F  V  T  Y  T  Y  A  A  T  I  E  V  R  N  N  C
 101  CATACACCGTGTGGGCAGCATCGACCCCGATAGGCGGTGGTCGACGTCTC
       P  Y  T  V  W  A  A  S  T  P  I  G  G  G  R  R  L
 151  AATAAGGGCCAAACATGGGTCATCAATGCTCCAAGGGGAACAAAGATGGC
       N  K  G  Q  T  W  V  I  N  A  P  R  G  T  K  M  A
 201  ACGTATATGGGGTCGTACTGGTTGTAACTTCAATGCTGCAGGCAGGGGTT
        R  I  W  G  R  T  G  C  N  F  N  A  A  G  R  G
 251  CGTGTCAGACTGGTGATTGTGGTGGAGTCTTGCAGTGTACCGGGTGGGGC
       S  C  Q  T  G  D  C  G  G  V  L  Q  C  T  G  W  G
 301  AAGCCCCCAAACACCTTGGCTGAATATGCCTTGGATCAGTTTAGCAACCT
       K  P  P  N  T  L  A  E  Y  A  L  D  Q  F  S  N  L
```

```
351  AGATTTCTGGGATATTTCTTTAGTTGACGGATTCAATATTCCAATGACTT
       D  F  W  D  I  S  L  V  D  G  F  N  I  P  M  T
401  TTGCCCCTACCAAACCTAGTGCTGGGAAATGCCATGCAATCCATTGCACG
       F  A  P  T  K  P  S  A  G  K  C  H  A  I  H  C  T
451  GCCAATATAAATGGTGAATGTCCTCGCGCCCTTAAGGTACCTGGAGGATG
       A  N  I  N  G  E  C  P  R  A  L  K  V  P  G  G  C
501  TAACAATCCTTGTACCACGTTTGGAGGACAACAATATTGTTGCACCCAAG
       N  N  P  C  T  T  F  G  G  Q  Q  Y  C  C  T  Q
551  GTCCATGTGGTCCTACAGAGTTGTCCAAATTTTTCAAGAAAAGATGCCCC
       G  P  C  G  P  T  E  L  S  K  F  F  K  K  R  C  P
601  GATGCTTATAGCTACCCACAAGATGATCCTACTAGCACATTTACTTGTCC
       D  A  Y  S  Y  P  Q  D  D  P  T  S  T  F  T  C  P
651  TGGAGGTAGTACAAACTATAAGGTTGTCTTTTGTCCCAATGGCGTTGCTG
       G  G  S  T  N  Y  K  V  V  F  C  P  N  G  V  A
701  ATCCAAATTTCCCCTTGGAGATGCCTGCAAGTACTGATGAAGTGGCCAAG
       D  P  N  F  P  L  E  M  P  A  S  T  D  E  V  A  K
751  TAA
       *
         atttgag tctctttctt taaaattact tcaagtggtc gagtgatctc
```

7.3.6. Genomic sequence of 517c

```
  1  tatatccaaa caacttctta tactaaa
                        517c  ATGAGTCATTTGACAACTTGTTT
                                 M  S  H  L  T  T  C  L
 51  AGTGTTCTTCCTTCTTGCTTTTGTGACTTACACTTATGCTTCCGGTGTGT
       V  F  F  L  L  A  F  V  T  Y  T  Y  A  S  G  V
101  TTGAGGTCCATAACAACTGCCCCTACACCGTTTGGGCGGCGGCGGTCCCC
       F  E  V  H  N  N  C  P  Y  T  V  W  A  A  A  V  P
151  GTAGGAGGTGGCCGACGTCTCGAGAGAGGTCAGAGTTGGTGGTTTTGGGC
       V  G  G  G  R  R  L  E  R  G  Q  S  W  W  F  W  A
201  CCCACCGGGCACTAAAATGGCACGTATATGGGGTCGTACTAATTGCAACT
       P  P  G  T  K  M  A  R  I  W  G  R  T  N  C  N
251  TTGATGGTGCTGGTAGAGGTGGGTGCCAGACCGGTGATTGTGGTGGAGTC
       F  D  G  A  G  R  G  G  C  Q  T  G  D  C  G  G  V
301  CTAGATTGCAAAGGATGGGGTAAACCGCCAAACACCTTAGCTGAATACGC
       L  D  C  K  G  W  G  K  P  P  N  T  L  A  E  Y  A
351  TTTGAACCAGTTTGGTAACCTAGATTTCTGGGATATTTCTGTAATTGATG
       L  N  Q  F  G  N  L  D  F  W  D  I  S  V  I  D
401  GATTTAACATCCCTATGTCTTTTGGGCCAACTAAGCCTGGACCTGGAAAA
       G  F  N  I  P  M  S  F  G  P  T  K  P  G  P  G  K
451  TGTCATGGAATTCAATGCACAGCCAATATAAACGGTGAATGCCCTGGTTC
       C  H  G  I  Q  C  T  A  N  I  N  G  E  C  P  G  S
501  ACTTAGGGTACCTGGAGGATGTAACAACCCTTGTACCACATTCGGAGGAC
       L  R  V  P  G  G  C  N  N  P  C  T  T  F  G  G
551  AACAATACTGTTGTAATCACGGTCCATGTGGTCCTACTGATTTGTCAAGA
       Q  Q  Y  C  C  N  H  G  P  C  G  P  T  D  L  S  R
601  TTTTTCAAACAAAGATGTCCTGATGCCTATAGTTACCCTCAAGACGATCC
       F  F  K  Q  R  C  P  D  A  Y  S  Y  P  Q  D  D  P
651  AACAAGTACTTTTACTTGTCAGAGTTGGACTACGGACTACAAGATTATGT
       T  S  T  F  T  C  Q  S  W  T  T  D  Y  K  I  M
701  TCTGTCCTTATGGCTCTACTCACAATGAAACAACAAATTTCCCATTGGAG
       F  C  P  Y  G  S  T  H  N  E  T  T  N  F  P  L  E
751  ATGCCTACAAGTATTCTTGAAGTTGCTTAA
       M  P  T  S  I  L  E  V  A  *
```

gtagattcat catgagtgta

7.3.7. Genomic sequence of 57d

```
   1  taaatatatc caacaaataa atattcttct tattctaaa
                                      57d ATGAGTCACTT
                                           M  S  H  L
  51  GACAACTTGTTTAGTGTTCTTCCTCCTTGCCTTTGTGACTTACACTAATG
        T  T  C  L  V  F  F  L  L  A  F  V  T  Y  T  N
 101  CTTCCGGCGTATTTGAGGTCCATAACAACTGTCCATACACCGTATGGGCG
       A  S  G  V  F  E  V  H  N  N  C  P  Y  T  V  W  A
 151  GCGGCAACCCCCATAGGAGGTGGCCGACGTCTCGAGAGAGGTCAAAGTTG
       A  A  T  P  I  G  G  G  R  R  L  E  R  G  Q  S  W
 201  GTGGTTTTGGGCCCCACCGGGCACTAAAATGGCACGTATTTGGGGTCGTA
        W  F  W  A  P  P  G  T  K  M  A  R  I  W  G  R
 251  CTAATTGCAACTTCGATGGTGCTGGTAGAGGTTGGTGCCAGACTGGTGAT
       T  N  C  N  F  D  G  A  G  R  G  W  C  Q  T  G  D
 301  TGTGGTGGAGTCCTAGAATGCAAAGGATGGGGTAAACCACCAAACACCTT
       C  G  G  V  L  E  C  K  G  W  G  K  P  P  N  T  L
 351  AGCTGAGTATGCTTTGAATCAATTTAGCAACTTAGATTTCTGGGACATTT
        A  E  Y  A  L  N  Q  F  S  N  L  D  F  W  D  I
 401  CTGTTATTGATGGATTCAACATCCCTATGTCTTTCGGCCCAACTAACCCT
       S  V  I  D  G  F  N  I  P  M  S  F  G  P  T  N  P
 451  GGGCCGGGAAAATGTCATCCAATTCAATGTGTTGCCAATATAAATGGTGA
       G  P  G  K  C  H  P  I  Q  C  V  A  N  I  N  G  E
 501  ATGCCCTGGTTCACTTAGGGTACCCGGAGGATGTAACAACCCTTGTACCA
        C  P  G  S  L  R  V  P  G  G  C  N  N  P  C  T
 551  CATTCGGAGGACAACAATATTGTTGCACCCAAGGTCCATGTGGTCCTACC
       T  F  G  G  Q  Q  Y  C  C  T  Q  G  P  C  G  P  T
 601  GATTTGTCAAGATTTTTCAAACAAAGATGTCCCGATGCCTATAGTTACCC
       D  L  S  R  F  F  K  Q  R  C  P  D  A  Y  S  Y  P
 651  TCAAGATGATCCAACAAGTACATTTACTTGCCAAAGTTGGACTACAGACT
        Q  D  D  P  T  S  T  F  T  C  Q  S  W  T  T  D
 701  ACAAGGTTATGTTTTGTCCTTATGGCTCTACTCACAATGAAACAACAAAT
       Y  K  V  M  F  C  P  Y  G  S  T  H  N  E  T  T  N
 751  TTCCCATTGGAGATGCCTACAAGTACTCTTGAAGTGGCTTAA
       F  P  L  E  M  P  T  S  T  L  E  V  A  *
                                                 ttaagtat
```

7.3.8. Genomic sequence of 137d'

```
   1  caaataaata ttcttcttat tctaaa
                          137d' ATGAGTCACTTGACAACTTGTTTA
                                     M  S  H  L  T  T  C  L
  51  GTGTTCTTCCTCCTTGCCTTTGTGACTTACACTAATGCTTCCGGTGTATT
       V  F  F  L  L  A  F  V  T  Y  T  N  A  S  G  V  F
 101  TGAGGTCCATAACAACTGTCCATACACCGTATGGGCGGCGGCCACCCCCA
        E  V  H  N  N  C  P  Y  T  V  W  A  A  A  T  P
 151  TAGGAGGTGGCCGACGTCTCGAGAGAGGTCAAAGTTGGTGGTTTTGGGCT
       I  G  G  G  R  R  L  E  R  G  Q  S  W  W  F  W  A
 201  CCACCGGGTACTAAAATGGCACGTATTTGGGGTCGTACTAATTGCAACTT
        P  P  G  T  K  M  A  R  I  W  G  R  T  N  C  N  F
 251  TGATGGTGCTGGTAGAGGTTGGTGCGAGACTGGTGATTGTGGTGGAGTCC
        D  G  A  G  R  G  W  C  E  T  G  D  C  G  G  V
```

```
301  TAGAATGCAAAGGATGGGGTAAACCACCAAACACCTTAGCTGAGTACGCT
      L E  C  K  G  W  G  K  P  P  N  T  L  A  E  Y  A
351  TTGAATCAATTTAGCAACTTAGATTTCTGGGATATTTCTGTTATTGATGG
      L  N  Q  F  S  N  L  D  F  W  D  I  S  V  I  D  G
401  ATTCAACATCCCTATGTCTTTCGGCCCAACTAACCCTGGGCCGGGAAAAT
        F  N  I  P  M  S  F  G  P  T  N  P  G  P  G  K
451  GTCATCCAATTCAATGTGTTGCCAATATAAACGGTGAATGCCCTGGTTCA
      C H  P  I  Q  C  V  A  N  I  N  G  E  C  P  G  S
501  CTTAGGGTACCCGGAGGATGTAACAACCCATGCACCACATTCGGAGGACA
      L  R  V  P  G  G  C  N  N  P  C  T  T  F  G  G  Q
551  ACAATATTGTTGCACTCAAGGTCCATGTGGTCCTACCGATTTATCAAGGT
        Q  Y  C  C  T  Q  G  P  C  G  P  T  D  L  S  R
601  TTTTCAAACAAAGATGTCCCGATGCCTATAGTTACCCTCAAGATGATCCA
      F F  K  Q  R  C  P  D  A  Y  S  Y  P  Q  D  D  P
651  ACAAGTACATTTACTTGTCAAAGTTGGACTACAGACTACAAGGTTATGTT
      T  S  T  F  T  C  Q  S  W  T  T  D  Y  K  V  M  F
701  TTGTCCTTATGGGTCTACTCACAATGAAACAACAAATTTCCCATTGGAGA
        C  P  Y  G  S  T  H  N  E  T  T  N  F  P  L  E
751  TGCCTACAAGTACTCTTGAATTGGCTTAA
      M P  T  S  T  L  E  L  A  *
                                  t taagtagaat cttgagttga
```

7.3.9. Genomic sequence of 68e

```
  1  aaaaaaaaag
          68e ATGTCCACAAACATGGACTATTTGATCAGATCTTCTTCTG
                 M  S  T  N  M  D  Y  L  I  R  S  S  S
 51  TTTTCTTCCTCCTTTCTTTTGTGACTTATACTTATTCTACCTCTTTTGAA
      V F  F  L  L  S  F  V  T  Y  T  Y  S  T  S  F  E
101  GTCCGAAACAACTGTCCATACACCATCTGGGCGGCATCGACCCCGATAGG
      V  R  N  N  C  P  Y  T  I  W  A  A  S  T  P  I  G
151  CGGTGGTCGACGCCTCGATTCAGGCCAGACCTGGGTCATCGATGCGCCGA
        G  G  R  R  L  D  S  G  Q  T  W  V  I  D  A  P
201  GGGGCACTAAGATGGCACGTATATGGGGTCGTACTAATTGCAACTTTGAT
      R G  T  K  M  A  R  I  W  G  R  T  N  C  N  F  D
251  GGTGCTGGTAGAGGTTCTTGCCAGACCGGTGATTGTGGTGGAGTCTTGCA
      G  A  G  R  G  S  C  Q  T  G  D  C  G  G  V  L  Q
301  GTGTACCGGGTGGGGCAAATCGCCAAACACCCTAGCCGAATATGCCTTGA
        C  T  G  W  G  K  S  P  N  T  L  A  E  Y  A  L
351  ACCAATTTAGCAACCTAGATTTCTGGGACATTTCTTTAATCGAAGGATTC
      N Q  F  S  N  L  D  F  W  D  I  S  L  I  E  G  F
401  AATATACCAATGACTTTCGCCCCGACCAATCCTAGTGGAGGAAAATGCCA
      N  I  P  M  T  F  A  P  T  N  P  S  G  G  K  C  H
451  CGCAGTTCAATGCACAGCCAATATAAATGGTGAATGCCCTAATCAACTTA
        A  V  Q  C  T  A  N  I  N  G  E  C  P  N  Q  L
501  AAGTATCCGGAGGATGTAACAATCCTTGTACCATAGGTCAATGTGGTCCT
      K V  S  G  G  C  N  N  P  C  T  I  G  Q  C  G  P
551  AATGACTTGTCCAGATTTTTCAAACAAAGATGCCCTGATGCATATAGCTA
      N  D  L  S  R  F  F  K  Q  R  C  P  D  A  Y  S  Y
601  CCCACAAGATGATCCTACTAGCTTATTTGCTTGCCCTAGTGATAGTACAA
        P  Q  D  D  P  T  S  L  F  A  C  P  S  D  S  T
651  ATTATAGGGTTGTTTTTTGCTAA
      N Y  R  V  V  F  C  *
                             tgccaac aataaagcat attttgctac
```

Eidesstattliche Erklärung

Ich versichere, daß ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jeder Einzelfall als Entlehnung kenntlich gemacht habe; daß diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; das Sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, daß ich eine solche Veröffentlichung vor Abschluß des Promotionsverfahrens nicht vornehmen werde.

Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Priv.- Doz. Dr.Christiane Gebhardt (MPIZ, Köln) betreut worden.

Köln, den 26.03.02

Rosa Castillo

Teiler dieser Arbeit sind in folgenden Veröffentlichungen enthalten:

Castillo R., Gebhardt C. and Ghislain M. Screening for late blight resistance using a BAC library from a high field resistant potato hybrid (Poster 366). 10th Molecular Plant Microbe Interactions Meeting, July 2001. Madison-USA.

Acknowledgements

This work has been done at the Max Planck Institute for Plant Breeding Research (Cologne), in the Department of Plant Breeding and Yield Physiology. Thanks to Prof. Francesco Salamini for the facilities in lab. equipment and reagents provided during my training. I would like to thank Dr. Marc Ghislain (CIP-Lima, Peru) who gave me the chance to realize this work. I am especially thankful to my supervisor Dr. Christiane Gebhardt for her advice and guidance during these more than three years.

Thanks a lot to Agim Ballvora for his help to solve practical and theoretical problems. Thanks for being so patient!. I would like to say thanks to all my past and present labmates for their help and kindness: Cristina Menéndez, Jost Muth, Birgit Walkemeier, Heike Henselewski, Monique Winkler, Iris Schmidt, Christina Bormann, Andreas Heibges, Jürgen Paal, Li Li, Jeong He Kim, Ralf Schäfer-Pregl and Julia Weiss. Thanks to Andreas Rickert for having spent time on a precise german translation of the summary. Many thanks to Werner Lehmann for his invaluable technical assistance.

Thanks to all the friends for the nice times we have had together. Thanks for their support and help.

Tambien agradezco a mis papas Victor y Edith, a Victor E., Sandra, Chuli, abuelita, Leonardo; Felipe, Manuel, Josefina, Matilde y Mercedes, Balki y Basi por su cariño y apoyo permanentes desde donde estén.

Y no me olvido de ti, Wim. Gracias por tu cariño, paciencia y apoyo.

Lebenslauf

Name	Rosa Angela Castillo Ruiz
Geburtsdatum und Ort	22.09.73 Lima-Peru
Nationalität	Peruanerin
Eltern	Victor Castillo Calderón Edith Ruiz de Castillo
Familienstand	Ledig
Schule	1981-1985 Primary School 1986-1990 Secondary School
Studium	1991-1996 Bachellor Biology Universidad Nacional Agraria La Molina (UNALM) Lima-Peru 1997-1998 Thesis of Professional degree Biologist International Potato Center/ UNALM, Lima-Peru
Promotion	1998-2002 Ph.D. Genetics Universität zu Köln/Max-Planck- Insitut für Züchtungsforschung