Genetics: Published Articles Ahead of Print, published on February 3, 2008 as 10.1534/genetics.107.086421

Research article

Genomic organization, rapid evolution and meiotic instability of NBS-encoding genes in a new fruit crop "chestnut rose"

Qiang Xu,[*] Xiaopeng Wen [†] and Xiuxin Deng[*, 1]

[*]National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, Hubei, P. R. China and [†]Guizhou Key Laboratory of Agricultural Bioengineering, Guizhou Univeristy, Guiyang 550025, Guizhou, P. R. China

Running head: Fruit crop, resistance gene, evolution

Keywords: chestnut rose, genomic organization, meiotic instability, rapid evolution, resistance gene

[1]*Corresponding author*: Prof. Xiuxin Deng, National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Lion Street No.1, Wuhan 430070, P.R. China. Tel: +86-27-87286906; Fax +86-27-87280016; E-mail xxdeng@mail.hzau.edu.cn

ABSTRACT

From chestnut rose, a promising fruit crop of Rosa species, powdery mildew disease resistant and susceptible genotypes and their F_1 progenies were used to isolate NBS-encoding genes using 19 degenerate primer pairs and an additional cloning method called overlapping extension amplification. A total of 126 genes were harvested; of these 38 were from resistant parent, 37 from susceptible, and 51 from F_1 progenies. Phylogenetic tree was constructed and revealed that NBS sequences from parents and F_1 progenies tend to form mixture and are well distributed among the branches of the tree. Mapping of these NBS genes suggested their organization in the genome as "tandem duplicated cluster" and to a less extent "heterogeneous cluster". Intra-specific polymorphisms and inter-specific divergence were detected by Southern blotting with NBS-encoding genes as probes. Sequencing on nucleotide level revealed even more intra-specific variation: for R4 gene 9.81% of the nucleotides are polymorphic. Amino acid sites under positive selection were detected in the NBS region. Some NBS-encoding genes were meiotically unstable, which may due to recombination and deletion events. Moreover, a transponson-like element was isolated in the flanking region of NBS gene, implying a possible role for transponson in the evolution history of resistance genes.

Plants utilize two distinct defense systems to recognize and respond to pathogen challenges (CHISHOLM *et al.* 2006; JONES and DANGL 2006; DE YOUNG and INNES 2006). The first line of defense against pathogens is basal defense which uses pattern recognition receptors (PRRs) to recognize pathogen-associated molecular patterns (PAMPs), the common features of pathogens such as bacterial flagellin. The second line of defense acts strongly to specifically recognize pathogen effectors, using resistance (*R*) genes. Many *R* genes have now been cloned from a wide variety of plant species (HAMMOND-KOSACK and JONES 1997; DANGL and JONES 2001), of which the largest class encodes a nucleotide-binding site (NBS) and a series of carboxy-terminal leucine-rich repeats (LRRs). The NBS domain contains several strictly ordered motifs, i.e. P-loop, kinase-2, kinase-3a, and GLPL motifs, that are highly conserved across plant species, and therefore they have been used extensively to identify and classify NBS-encoding genes. By using PCR amplification with degenerate primers targeted to these conserved motifs, a large number of NBS sequences with homology to *R* genes, so-called "R gene homologues (RGHs)", have been isolated from various plant species (KANAZIN *et al.* 1996; LEISTER *et al*. 1998; SHEN *et al.* 1998; YU *et al.* 1996; COLLINS *et al.* 1998; ZHU *et al.* 2002; ASHFIELD *et al.* 2003). RGHs are abundant in plant genomes. Approximately 160, for example, are present in *Arabidopsis* genome (MEYERS *et al.* 2003). Genetic analysis has shown that RGHs tend to occur in clusters, and often map to major resistance genes or quantitative trait loci (KANAZIN *et al.* 1996; COLLINS *et al.* 1998; YOUNG 2000; DONALD *et al.* 2002; ASHFIELD *et al*. 2003; CALENGE *et al.* 2005; WELTER *et al.* 2007).

The large number of RGH sequences provides a facile system to study the evolution of this gene family across plant taxa (GRUBE *et al.* 2000; PAN *et al.* 2000b; BAI *et al.* 2002; MEYERS

et al. 2003; see review in McHale *et al.* 2006). After analysis of more than 400 NBS-encoding homologs of 26 genera, Meyers and colleges (1999) demonstrated that on genome these RGHs frequently reside in "mega-clusters" consisting of smaller clusters with several members each, and they phylogenetically fall into two distinct groups, i.e. TIR and Non-TIR subfamilies. Cannon *et al.* (2002) investigated even larger number of sequences, 800 RGHs from 30 genera, and suggested that evolutionary histories are different for the TIR and Non-TIR subfamilies. With the increasing information in database and growing power of computational biology, researchers shift their focus to the comparative genetics of RGHs on/within family-level. Comparative analysis of tomato and Arabidopsis revealed remarkably rapid evolution of RGHs during radiation of plant families (Pan *et al.* 2000a). In a deeper step, Plocik *et al.* (2004) comparatively analyzed the RGHs within a specific family (Asteraceae), and indicated that gene duplication and loss events occur and change the composition of these gene subfamilies over time. Comparative analysis continued to be further applied in RGHs within a specific species, most of which concentrate on wild and cultivated plants such as apple (Lee *et al.* 2003), strawberry (Martinez Zamora *et al.* 2004), and potato (Kuang *et al.* 2005). However, relatively few studies have comparatively analyzed RGHs patterns between resistant and susceptible lines within a fruit tree species, and the behaviors of these RGHs during meiosis have been ignored.

Chestnut rose, a rare fruit crop in Southwest China, belongs to Rosaceae family *Rosa* genus (Figure 1). It has recently been labeled as one of the three promising new fruit crops in China (Wen and Deng 2005) due to its fruits having high content of vitamin C (2000-3000 mg/100g FW), displaying high levels of superoxide dismutase (SOD) activity, and therefore is believed

to have senescence-retarding and cancer-preventing effects (MA *et al.* 1997). Unfortunately, chestnut rose crops are suffering from powdery mildew disease caused by fungi *Sphaerotheca pannosa*. An indigenous cultivar from Guizhou province of China, Guinong No. 6, however, is highly resistant to powdery mildew. Within this species the other cultivar Guinong No. 5 is highly susceptible to powdery midew (supplementary Figure 1). Classical genetic analyses have showed that Guinong No. 6 and Guinong No. 5 are very similar in many evolutionarily characters such as growth habit, tree and leaf shape, and they share near identical genetic background. The relatively close phylogenetic relationship between Guinong No. 6 and Guinong No. 5 makes comparative analysis of R gene candidates a particularly attractive approach to understanding how the resistance specificity generated.

In this study, we aim to clone RGHs from resistant and susceptible genotypes of chestnut rose (*Rosa roxaburghii*) and their F_1 progenies for comparative analysis of polymorphisms of the RGHs to identify genetic parameters relevant to the resistance specificity properties. Genetic mapping and evolutionary analysis provide us insights into the genomic organization and evolutionary pattern of RGHs in chestnut rose genome, and allow us to detect amino acid sites under positive selection in NBS-coding region. A transponson-like element was isolated in the flanking region of NBS domain. Moreover some RGHs meiotic stability was investigated.

MATERIALS AND METHODS

Plant materials: *Rosa roxburghii* cv. Guinong No. 6 (powdery mildew disease resistant genotype) and *R. roxburghii* cv. Guinong No. 5 (susceptible genotype), and an F_1 population

of 109 plants derived from an intraspecific cross between them were used for molecular cloning, genetic mapping and powdery mildew disease phenotyping. Disease phenotypes were evaluated the same as previously reported (XU *et al.* 2005).

Isolation of NBS-encoding genes from resistant genotype, susceptible genotype and their F_1 progenies: For resistant and susceptible parents, genomic DNA were extracted and used for PCR amplification. For F_1 progenies bulk DNA was constructed for molecular cloning. RGHs were broadly obtained as much as possible by employing two cloning method: one is the PCR amplification with 19 degenerate primer pairs targeted to P-loop and GLPL motifs; the other is the overlapping extension approach which using a target strategy to capture TIR-type RGHs from genome (XU *et al.* 2005). PCR products were cloned into pMD18-T vector (Takara Bio Inc.). Recombinant plasmid DNA was extracted by alkaline lysis (SAMBROOK *et al.* 1989). Each clone was re-amplified with M13 universal primers and then subject to restriction analysis using three restriction endonucleases (*Taq*I, *Hae*III, and *Hin*fI). Based on the restriction patterns, representative clones of each type were used for sequencing on both directions on an ABI 3730 sequencer.

For the purpose of comparative analysis, several sets of previously published RGHs from other plant species were used. Sixteen Arabidopsis RGHs were retrieved form At-Rgenes (http://www.niblrrs.edu/At_RGenes) in March 2006. Fifteen NBS sequences was selected from Cereal genomes described by PAN *et al.* (2000b) and 10 sequences from a coniferous species (*Pinus monticola*) were used to determine whether chestnut rose NBS gene subfamily risen before or after the monocot/dicot and angiosperm/gymnosperm separation. The other sequences were from *Medicago truncatula* (ZHU *et al.* 2002), *Cicer arietinum, Pennisetum*

glaucum, Cajanus cajan, Sorghum bicolor (MEYERS *et al.* 1999) and unpublished data from *Poncirus trifoliate* in our Lab.

Isolation of flanking sequences of NBS-encoding genes: To isolate flanking sequences of NBS-encoding genes, a method similar to GenomeWalker™ Universal Kit (Clontech) was used except that asymmetric PCR (TAIL-PCR) was introduced to eliminate artificial PCR products. An adaptor was designed according to CALENGE *et al.* (2005). The adapter had a short and a long arm, leaving the 5' part of the adapter single stranded, and 3' part containing *Hpa*II endonuclease restriction site which sticks to digested DNA fragment. 3 μg genomic DNA was digested with 50 U *Hpa*II endonuclease enzymes for 8-16 hrs, and de-activated at 65 °C for 10 min. Then the *Hpa*II adapter was ligated to the ends of the digested DNA fragments by adding 40 U T4 ligase overnight at 16°C. The ligated products were used as templates for PCR amplification with an adapter primer and a specific primer in the NBS region. The adapter primer was identical to the single-stranded 5' part of the *Hpa*II adaptor, and extension of the 3' end of the short arm in the PCR reaction was effectively blocked by an amino linker. To increase the specificity of PCR products, thermal asymmetric interlaced PCR (TAIL-PCR) was used (supplementary Table 3)

Gene-specific PCR amplification, southern blotting, and genetic mapping: Forty two gene-specific primers were designed with primer3 software (http://www-genome.wi.mit.edu/cgi-bm/primer/primer3-www.cgi). The PCR cycling conditions were as follows: 94 °C 3 min followed by 32 cycles of 94 °C 30 s, annealing temperature (supplementary Table 2) 1 min, and 72 °C 1 min. PCR products were separated by electrophoresis on 1.5% agarose gel. Genomic DNA (10 μg) of both parents and a DNA

bulk constituted from F_1 progenies were digested with six restriction endonucleases (i.e. *Eco*RI, *Hin*dIII, *Bam*HI, *Eco*RV, *Dra*I, and *Taq*I). Southern hybridization was performed according to XU *et al.* (2005).

Molecular markers were scored in the 109 F_1 individuals according to RITTER *et al.* (1990). Linkage analysis and QTL detection were performed using Map Manager QTXb20 (MEER *et al.* 2004), with the linkage criterion being set at P=0.05.

Sequence analyses: Phylogenetic and evolutionary analyses were carried out according to Xu *et al.* (2007). Sequences were aligned using CLUSTALX (THOMPSON *et al.* 1997) and manually edited in GENEDOC (http://www.psc.edu/biomed/genedoc/). All the sequence statistics were output from GENEDOC, Duncan's multiple range tests were performed by SPSS11.0 soft (SPSS inc., Chicago, IL). Neighbor-joining trees using Kimura's two-parameter model and maximum parsimony phylogenetic trees were constructed and bootstrap number were calculated by heuristic search in PAUP* 4.0 (Sinauer Associates, Sunderland, MA, USA). The trees were visualized using the program TREWVIEW (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).

Evidence for positive selection was evaluated by using a maximum likelihood (ML) method (YANG 1997), which allows searches for individual amino acid sites under selection. Site-specific model M_7 and M_8 in program CODEML from PAML soft were run to detect positive selection at individual sites. Model M_7 is a special case of model M_8 that assumes no selection, whereas model M_8 allows for positively selected sites (YANG *et al.* 2000). Likelihood ratio tests (LRTs) were used to determine which model fit the observed data. The significance of the LRTs was evaluated by comparing it to a χ^2 statistic. When the LRTs are

significant, Bayesian approach was used to calculate posterior probabilities of each positively selected site.

Three-dimensional structures for RGHs were constructed by homology modeling running MODELLER software 9v2 (MARTI-RENOM *et al.* 2000). The X-ray diffraction structural data from APAF-1 (PDB code 1z6t) was used as template to model the NBS-encoding genes. The sequence alignments of APAF-1 and NBS gene was obtained from PHYRE, a threading service soft. Amino acids identified as possible targets of positive selection were mapped onto the 3D structure using VMD software (HUMPHREY *et al.* 1996; TIFFIN 2004)

RESULTS

NBS-encoding genes in resistant parent, susceptible parent, and F_1 progenies: Fragments of predicted sizes (500-700 bp) were cloned and sequenced from resistant line (Guinong No. 6), susceptible line (Guinong No. 5) and their F_1 progenies. 126 of the 150 clones showed significant similarity (E-values $<e\times10^{-5}$) to the known R-genes or published RGHs as revealed by BlastX analysis in GenBank. For the remaining 24 sequences, 15 encompassed repetitive elements, while the rest 9 sequences gave no or weak blast hits in the database (data not shown).

Conceptual translations of the above 126 sequences (38 resistant parent, 37 susceptible parent, and 51 F_1 progenies) revealed the presence of premature stop codons in four clones from resistant line, seven clones from susceptible line, and 19 clones from F_1 progenies. Interestingly, Bk174 pseudogene carried a six direct repeat oligos of 18 nucleotides (CAGCCTTCCTTTAGCCCT). The remained 96 sequences (34 resistant parent, 30 susceptible parent, and 32 F_1 progenies) were identified as RGHs by the presence of continue

open reading frame (ORF) and by characteristic motifs such as GVGKTT (P-loop), LLVLDDVW/D (Kinase-2) and GLPL motif (MEYERS *et al.* 1999; PAN *et al.* 2000b). Moreover, differences were observed in these motifs among genotypes; Sequnece Logos were made from the P-loop, Kinas-2 and GLPL motifs to illustrate the difference among resistant parent, susceptible parent and F_1 progenies (supplementary Figure 2). Sequence diversity was analyzed on nucleotide and amino acid levels (supplementary Table 1). Within genotypes, for resistant parent, sequence nucleotide identity averaged at 52%; and for susceptible parent, was 57%. Unexpected, for inter-genotype between resistant and susceptible, average nucleotide identity was 54%, higher than that within resistant genotype (supplementary Figure 1).

A transponson-like element is neighbor to NBS region: After the *Hpa*II adaptor was ligated to digested DNA fragments, ligated products were used as templates to isolate genes neighbor to the NBS-encoding region. Thirteen NBS-specific primers in combination with the adaptor primer were tested in PCR amplification. One primer combination produced a specific product near 500 bp after three round of PCR. Sequencing revealed that fragment contains the adapter sequence at one end and an NBS primer sequence on the other end, indicating that this procedure efficiently excludes adapter-adapter fragments. The sequence was further characterized by sequence similarity searches in the GenBank databases using BLASTX program. The sequence was highly homologous to a published retrotranspon from *Populus trichocarpa*, with 66% identity. Southern analysis with this sequence as probe produced a classical pattern of transposon, i.e. large number of copies and pattern look like smear, further confirming the existence of a transponson gene in the neighbor region of NBS-encoding gene (Figure 2).

Phylogenetical analysis of the NBS genes from resistant, susceptible parents and F_1 generation: The 96 sequences from chestnut rose and 87 previously published sequences and mammalian apoptosis-related protein Apaf1 (outgroup) were used in phylogenetic analysis. Trees were created by neighbor-joining method and the reliability of the trees was established by conducting 1000 bootstrap re-sampling steps (Figure 3). The overall impression of this tree is that NBS genes from resistant/susceptible genotypes and F_1 progenies are well distributed among branches of this tree. This suggests that these genes probably arose from common ancestors that existed before variety split within species. Two major branches designated as TIR and non-TIR subfamilies were obtained and well supported by high bootstrap value. The two branches were further confirmed by the presence or absence of characteristic subfamily-specific motifs, for example RNBS-A-nonTIR (FDLxKxWVSVSDDF) and RNBS-A-TIR (LQxQLLSxxL) motifs are mutually exclusive in the two groups, consistent with previous results from MEYERS *et al.* (1999) and PAN *et al.* (2000b).

Five NBS genes from susceptible line, eleven from resistant line, two from F_1 progenies, three from *Arabidopsis*, and two from *Pinus monticola* constitute the TIR branch. Branch of TIR group showed rather distantly clustered nodes, reflecting relative longevity of diversification within species. Some branches contained sequences from *Medicago truncatula*, implying that the ancient ancestor of chestnut rose and legume contained multiple TIR sequences that have since diverged. The phylogenies of TIR group showed some chestnut rose TIR-type NBS sequences is phylogenetically closed to pinus TIR-type sequences, indicating the antiquity of these sequences and the origin of this gene family may predates the evolutionary split of the plant lineages leading to Angiosperms and Gymnosperms.

The nonTIR branche contains 62 members, with 19 from susceptible line, 18 from resistant line, 22 from F_1 progenies, one from *Arabidopsis*, and two from *Oryza sativa*. Long branch lengths and closely clustered nodes were observed in this branch, indicating ancient divergence into separate lineages followed by more recent duplication and diversification.

Describetrees command was used in the PAUP soft to output branch tree length statistics. The resistant line has a total length of 2001 units for 34 sequences, or 58.8 per sequence; the susceptible line has a total length of 1590 units for 30 sequences, or 53 per sequence; the F_1 progenies have a total length of 2018 unit for 32 sequences, or 63.5 per sequence. Multiple comparisons LSD test reveals that tree distance for them are significantly different from each other, with the difference 3.2± 0.7 unit between resistant/susceptible genotypes.

Genetic mapping reveals three clusters of NBS genes in chestnut rose genome: Sequence tagged site (STS) marker was used to mapping the NBS genes because it can be easily applied by designing gene-specific primers. Forty two gene-specific primer pairs (12 from resistant genotype, 17 from susceptible and 13 from F_1 generation) were designed and used for scoring 109 F_1 progenies (supplementary Table 2). Conventional RFLP was also employed to map the genes as previously described (Xu *et al.* 2005). Eleven representative clones in combination with 5 restriction endonucleases were used to detect polymorphisms between resistant/susceptible lines. Six clones yielded distinct polymorphisms (Figure 4) and were then used to survey Southern blots of the F_1 progenies.

Three clusters of NBS genes were obtained (Figure 3). The largest one, designated CR1, includes 23 STS markers developed from resistant genotype and F_1 progenies. NBS genes from susceptible form a single cluster CR2, which composed of 12 loci. The third cluster CR3,

with 6 loci, contains a major QTL that located between RGA22C marker and RAPDAL09 marker for powdery mildew resistance. Noticeably, among the 109 F_1 progenies, 26 individuals tend to be resistant, and 25 of them were genotyped with RGA22C marker; while 33 individuals are susceptible, all of them are absent from this marker. Therefore RGA22C could be used as a molecular marker for efficient selection of resistant progeny (Xu et al. 2005).

When comparing the mapping data with phylogenic clades, RGHs from the same phylogenetic clade tend to cluster together. S101, S93, S133 and S107 were closely related in the same clade and sequence similarity between them more than 90%, they clustered closely with genetic distance ~ 10 CM in cluster CR2. Same phenomenon could be seen in cluster CR1. This phenomenon suggests the tandem duplication followed by divergence occur recently. This kind of cluster was named as “tandem duplicated cluster” according to LEISTER (2004). However, sometimes NBS genes from distinguished clades could cluster together. For example in cluster CR1, R5 gene, phylogenetically belonging to TIR group, flanked by Bk89 and Bk66 genes, which are nonTIR type gene. This kind of cluster was named as “heterogeneous cluster”, as suggested by LEISTER (2004).

Rapid evolution of NBS-encoding genes in chestnut rose: The RFLP patterns of NBS-encoding genes were quite different between resistant and susceptible genotypes (Figure 4). Less number of hybridizing fragments was detected for five probes R22, S5, R5, Bk71 and S2 in susceptible genotype in comparison to resistant genotype. While for R4 probe, the RFLP pattern is different between the two genotypes although they both have two copy numbers. Further sequencing on nucleotide level revealed a considerable sequence

polymorphism between the two genotypes (The intra-specific variation, Figure 4). Within NBS-encoding region 42 polymorphic sites were observed; of these, 15 result in silent substitutions, 27 involve non-conservative amino acid changes. Approximately 9.81% of the total 428 nucleotides are polymorphic for this gene, higher than that observed in Arabidopsis *RPS2* gene (CAICEDO *et al.* 1999). The intra-specific copy number and sequence nucleotide variation strongly suggest the rapid evolution and re-arrangement of NBS-encoding genes in the chestnut rose genome.

Inter-specific polymorphisms were also revealed from RFLP pattern among *Rosa roxburghii*, *R. sterilis*, and *R. chinensis* (Figure 4). On nucleotide level, a total of 23 non-conservative amino acid changes were observed in *R. sterilis*, and a polymorphic site in *R. chinesis* result in a change from Valine to a stop condon. Further inter-genus RFLP survey in *Prunus* and *Malus* also revealed presence/absence polymorphism. The absence of inter-specific and inter-genus cross hybridization signals and the considerable nucleotide changes confirmed the rapid evolution of NBS-encoding genes.

Positively selected sites were detected in NBS region: Types of forces that drive natural selection can been inferred from the ratio of nonsynonymous substitution (d_N) that cause an amino acid change to the synonymous substitution (d_S) that dose not, also known as ω. A site-specific model M7 and M8 (described in materials and method) were used to determine the sites which are under purifying selection ($\omega<1$), or neutral selection ($\omega\approx1$), or diversifying selection ($\omega>1$). Pairwise values of d_S and d_N indicated an overall scarcity of amino acid substitutions. The average d_S for resistant parent, susceptible parent, and F_1 generation ranged from 0.8 to 1.5 (supplementary Table 1), meaning a high sequence divergence according to

the criteria described by YANG *et al.* (2000). It is therefore the ratios (ω) are reliable. Among the 146 sites, most ω are below 1, indicating that most of the sites are under purifying selection. However, one residue 125 H (Histidine) was detected to be under strong diversifying selection (ω>1) with a posterior probability 92.4%. Other sites including 47 P (Proline), 57 E (Glutamic acid), 107 K (Lysine), 127 L (Leucine) are likely to be under diversifying selection with probabilities 60-80% (Table 1; supplementary Figure 3).

Meiotic instability of NBS genes: Fragment deletion event was observed during the meiosis. Bk111 gene, cloned from F_1 progenies that showed sequence affiliation to R26 gene from resistant genotype, lost a 300 bp fragment during meiosis (Figure 5). Estimates of meiotic instability were further investigated by RFLP analysis. The Southern blotting results revealed that the hybridization pattern for F_1 individuals can be the combined pattern of both parents, but to some extent, new hybridizing fragment appeared, forming "recombination pattern" of the two parents (Figure 5). This indicated that during meiosis NBS-encoding genes are not stable, and recombination occurred at these loci and consequently resulted in the appearance of new alleles in the F_1 progenies.

DISCUSSION

Previous studies of *R* genes or RGHs on model plants have accumulated knowledge on the generation of resistance specificities and evolutionary dynamics of *R* genes (BERGELSON *et al.* 2001; BAI *et al.* 2002; RICHLY *et al.* 2002; CAICEDO and SCHAAL 2004; KUANG *et al.* 2004; XIAO *et al.* 2004; MONDRAGON-PALOMINO and GAUT 2005; BAKKER *et al.* 2006; FRIEDMAN and BAKER 2007). This study, however, focused on a fruit crop chestnut rose. Two cloning methods were used in this study; one is direct PCR amplification with degenerate primers,

while this method was believed with risk of biased sampling due to preferential amplification. To broaden the sequences, another method called overlapping extension approach was used to capture NBS genes from genome as much as possible (Xu *et al.* 2005). Under this strategy, a total of 126 NBS-encoding genes were isolated, less than those reported in Arabidopsis and Rice. One possible reason for the less number of NBS genes would be the preferential amplification caused by degenerate primers though we tried best to avoid this as mentioned above; the second reason could be due to the scarce genome information from this fruit tree, no sequence information could be utilized from public database; another reason would be the slower diversification rate of NBS genes in wild chestnut rose than those in tamed plants such as Rice because of the selective pressure. However, based on the cloned 126 genes, this study still provides some interesting knowledge for the characteristics of NBS-encoding genes in chestnut rose, such as genomic organization as tandem cluster, neighbor to transponson element, rapid evolution, meiotic instable, and etc.

Evolutionary complexities of NBS-encoding gene in chestnut rose: NBS-encoding genes in chestnut rose exhibited high level of intra-specific polymorphisms: e.g. for R4 gene 9.81% of nucleotides were polymorphic, of which 64% were non-conservative amino acid changes, suggesting that this gene is maintained for short time periods. What are the evolutionary forces that shaping the polymorphisms in NBS-encoding genes in chestnut rose genome?

Intra-specific variations in NBS gene copy number and size were observed between resistant and susceptible genotypes (Figure 4). The two chestnut rose cultivars Guinong No. 6 and No. 5 were differentiated and selected at recent 30 years. Seen from Figure 4, it is obviously the signature of unequal crossing-over events that can produce gene copy number

variation and size difference (HAMMOND-KOSACK and JONES 1997; MCDOWELL and SIMON 2006). Such process is almost certainly leading to rapid gene divergence between different genotypes (gene amplification and reduction respectively); compatible with the opinion that NBS genes evolve rapidly. Moreover, the process of unequal crossing-over could be facilitated by the repetitive elements, which were also found in this study, within the RGH clusters on chestnut rose genome, as suggested by MCDOWELL and SIMON (2006) and FRIEDMAN and BAKER (2007). Unequal crossing-over events are believed to homogenize genes within a genotype, resulting paralogs being more similar than orthologs (MICHELMORE and MEYERS 1998). However, in contrast to the above expectations, inter-genotype comparison of nucleotide identity between resistant and susceptible turned out to be higher than that within resistant genotype, a surprising result with orthologs being more similar than paralogs. The contradiction implies that there have been other processes that shaping the NBS gene polymorphisms as well as unequal crossing-over events, and the unequal crossing-over events are not predominantly in generation gene variations.

Evolutionary analysis and genetic mapping revealed the existence of "tandem duplicated cluster", where point mutations were observed among the tandem direct repeats. The accumulated mutations are another evolutionary way to increase the complexity of RGHs and the opportunity of producing new homolog. Moreover, some RGHs share high homology with pinus RGHs, and produced uniform hybridization band among Rosaceae plants (Figure 4), indicating that these genes are ancient and maybe evolutionarily maintained by some form of balancing selection (TIAN *et al.* 2002).

Detecting adaptive evolution by comparing amino acid substitution rates (d_N) to

synonymous substitution rates (d_s) has indicated that positive selection has contributed to the evolution of NBS genes in chestnut rose. For many reported *R* genes, positive selection has been detected primarily in LRR regions. However, positively selected sites were also detected in NBS-encoding region, e.g. MONDRAGON-PALOMINO *et al.* (2002) found that in Arabidopsis five positive selected sites were positioned in NBS region, implying that NBS domain may function in determining resistance specificity. This could be evidenced from domain swapping experiments by replacement of the NBS-encoding region of *L10* gene with the equivalent region of *L2* or *L9* generated new recombinant alleles with novel specificity (LUCK *et al.* 2000). In chestnut rose, NBS domain may also functioned in determining the resistance specificity, further research are required to investigate.

It seems that transposable-like elements were involved in the evolutionary dynamics of NBS-encoding genes in chestnut rose. A transposable element was isolated and neighbor to NBS gene. For most of Rosaceae genomes, a large number of gene copies were detected by RFLP analysis. However, on cultivar Guinong No. 1 (intra-specific genotype), Peach, and Pear, one to three copies were detected. The markedly difference of copy number implied that such element could play active roles in the evolutionary history of resistance genes in chestnut rose. In rice, eleven different families of transposable elements were identified at the *Xa21* cluster, and the elements appeared to be a major source of variation in this cluster (SONG *et al.* 1998). In addition, in plants it is common that transposable elements activated by environmental stresses such as pathogen infection (GRANDBASTIEN 1998), and it is believed that such activation can increase genomic flexibility with a possible selective advantage.

Together, positive selection, balancing selection, recombination, point mutation, and even

transposable elements may constitute the driving forces that shaping the complexity, rapid evolution, and even the generation new resistance specificity of *R*-gene sequences in chestnut rose.

Meiotic instability of RGH genes: Sequence pairwise comparison revealed that within resistant parent nucleotide identity averaged at 52%, and 57% within susceptible parent, while 45% within F_1 generation. Statistical analysis showed that the latter was significantly lower than the former two at 0.05 level, meaning a higher sequence variation in F_1 progenies than that in parents. Moreover, some sequences from F_1 progenies carried repetitive elements.

Evidence from RFLP markers demonstrated that RGH genes are meiotically instable. New alleles were observed in F_1 plants. To confirm this, we used the RGH sequences from F_1 plants to design specific PCR primers and to determine the gene status in two parents. Three types were detected; what merits attention is type III where gene was only detected in some F_1 plants but not in either parent (supplementary Figure 4), suggesting that this allele was newly produced during meiosis. However, this is a preliminary data for the estimation of meiotic instability. One classical example was *Rp1* complex loci in maize, where homozygous line for *Rp1* locus was used to generate a large number of testcross progenies. Surprisingly, high frequency of susceptibles was found in the progenies, indicating the occurrence of meiotically instable gene (SUDUPAK *et al.* 1993). Further research on *Rp1* locus demonstrated that recombination is the primary mechanism of meiotic instability, and such recombination can result in new race specificity (SMITH and HULBERT 2005). But for this study, fruit tree is believed to be highly heterozygous for most gene loci; it is difficult to create homozygous line to get a depth insight into the mechanism for meiotic instability. However, further research on

sequencing the flanked reigon around NBS domain may help us understand the types of recombination and the genetic mechanism for meiotic instability.

The authors are grateful to Dr. S. Xiao from Center for Biosystems Research of University of Maryland Biotechnology institute for his valuable comments. Thanks also to the anonymous reviewer for critical reviews. This project was supported by the Ministry of Education of China (No. IRT0548), 973 project (2006CB708202) and the National Natural Science Foundation of China (30660115).

LITERATURE CITED

ASHFIELD, T., A. BOCIAN, D. HELD, A. D. HENK, L. F. MAREK *et al.*, 2003 Genetic and Physical Localization of the Soybean Rpg1-b Disease Resistance Gene Reveals a Complex Locus Containing Several Tightly Linked Families of NBS-LRR Genes. Mol. Plant-Microbe Interact. **16**: 817-826.

BAI, J., L. A. PENNILL, J. NING, S. W. LEE, J. RAMALINGAM *et al.*, 2002 Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. Genome Res. **12**: 1871-1884.

BAKKER, E. G., C. TOOMAJIAN, M. KREITMAN and J. BERGELSON, 2006 A genome-wide survey of R gene polymorphisms in Arabidopsis. Plant Cell **18**: 1803–1818.

BERGELSON, J., M. KREITMAN, E. A. STAHL and D. TIAN, 2001 Evolutionary dynamics of plant R-genes. Science **292**: 2281-2284.

CAICEDO, A. L., B. A. SCHAAL and B. N. KUNKEL, 1999 Diversity and molecular evolution of the RPS2 resistance gene in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **96**: 302-306.

CAICEDO, A. L., and B. A. SCHAAL, 2004 Heterogeneous evolutionary processes affect R gene diversity in natural populations of *Solanum pimpinellifolium*. Proc. Natl. Acad. Sci. USA **101**: 17444-17449.

CALENGE, F., C. G. VAN DE LINDEN, E. VAN DE WEG, H. J. SCHOUTEN, G. VAN ARKEL *et al.*, 2005 Reisstance gene analouges identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. Theor. Appl. Genet. **110**: 660-668.

CANNON, S. B., H. Y. ZHU, A. M. BAUMGARTEN, R. SPANGLER, G. MAY *et al.*, 2002 Diversity, distribution, and ancient taxonomic relationships within the TIR and Non-TIR NBS-LRR resistance gene families. J. Mol. Evol. **54**: 548-562.

CHISHOLM, S. T., G. COAKER, B. DAY and B. J. STASKAWICZ. 2006 Host-microbe interactions: shaping the evolution of the plant immune response. Cell **124**: 803-814.

COLLINS, N. C., C. A. WEBB, S. SEAH, J. G. ELLIS, S. H. HULBERT *et al.*, 1998 The isolation and mapping of disease resistance gene analogs in maize. Mol Plant-Microbe Interact. **11**: 968-978.

DANGL, J. L., and J. D. G. JONES, 2001 Plant pathogens and integrated defence responses to infection. Nature **411**: 826-833.

DEYOUNG, B. J., and R.W. INNES, 2006 Plant NBS-LRR proteins in pathogen sensing and host defense. Nature Immunol. **7**: 1243-1249.

DONALD, T. M., F. PELLERONE, A. F. ADAM-BLONDON, A. Bouquet, M. R. Thomas *et al.*, 2002 Identification of resistance gene analogs linked to a powdery mildew resistance locus in grapevine. Theor Appl Genet **101**: 301-308.

FRIEDMAN, A. R., and B. J. BAKER, 2007 The evolution of resistance genes in multi-protein plant resistance systems. Curr. Opin. Genet. Dev. **17**: 1-7.

GRANDBASTIEN, M. A., 1998 Activation of plant retrotransposons under stress conditions. Trends in Plant Sci. **3**: 181-187.

GRUBE, R. C., RADWANSKI E. R. and M. JAHN, 2000 Comparative Genetics of Disease Resistance Within the Solanaceae. Genetics **155**: 873-887.

HAMMOND-KOSACK, K. E., and J. D. G. JONES, 1997 Plant disease resistance genes. Annu. Rev. Plant Physiol. Plant Mol. Biol. **48**: 575-607.

HUMPHREY, W., A. DALKE and K. SCHULTEN, 1996 VMD-Visual Molecular Dynamics. J. Molec. Graphics **14**: 33-38.

JONES, J. D. G., and J. L. DANGL, 2006 The plant immune system. Nature **444**: 323-329.

KANAZIN, V., L. F. MAREK and R. C. SHOEMAKER, 1996 Resistance gene analogs are conserved and clustered in soybean. Proc. Natl. Acad. Sci. USA **93**: 11746-11750.

KU, H. M., T. VISION, J. LIU and S. D. TANKSLEY, 2000 Comparing sequenced segments of the tomato and Arabidopsis genomes: large-scale duplication followed by selective gene loss

creates a network of synteny. Proc. Natl. Acad. Sci. USA **97**: 9121-9126.

KUANG, H., S. S. WOO, B. C. MEYERS, E. NEVO, and R. W. MICHELMORE, 2004 Multiple genetic processes result in heterogeneous rates of evolution within the major cluster disease resistnace genes in lettuce. Plant Cell **16**: 2870-2894.

KUANG, H., F. WEI, M. R. MARANO, U. WIRTZ, X. WANG *et al.*, 2005 The R1 resistance gene cluster contains three groups of independently evolving, type I R1 homologues and shows substantial structural variation among haplotypes of *Solanum demissum*. Plant J. **44**: 37-51.

LEE, S. Y., J. S. SEO, M. RODRIGUEZ-LANETTY and D. H. LEE, 2003 Comparative analysis of superfamilies of NBS-encoding disease resistance gene analogs in cultivated and wild apple species. Mol. Genet. Genomics **269**: 101-108.

LEISTER, D., J. KURTH, D. A. LAURIE, M. YANO, T. SASAKI *et al.*, 1998 Rapid re-organization of resistance gene homologues in cereal genomes. Proc. Natl. Acad. Sci. USA **95**: 370-375.

LEISTER, D., 2004 Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. Trends Genet. **20**: 116-122.

LUCK, J. E., G. J. LAWRENCE, P. N. DODDS, K. W. SHEPHERD and J. G. ELLIS, 2000 Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. Plant Cell **12**: 1367-1378.

MA, Y.X., Y. ZHU and C. F. WANG, 1997 The aging retarding effect of 'Long-Life CiLi'. Mech. Ageing Dev. **96**: 171-189.

MARTINEZ-ZAMORA, M. G., A. P. CASTAGNARO and J. C. DIAZ-RICCI, 2004 Isolation and diversity analysis of resistance gene analogues (RGAs) from cultivated and wild strawberries. Mol. Genet. Genomics **272**: 480-487.

MARTI-RENOM, M. A., A. STUART, A. FISER, R. SÁNCHEZ, F. MELO *et al.*, 2000 Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. **29**: 291-325.

MCDOWELL, J. M., and A. S. A. SIMON, 2006 Recent insights into *R* gene evolution. Mol. Plant Pathol. **7**: 437-448.

MCHALE, L., X. TAN, P. KOEHL and R. W. MICHELMORE, 2006 Plant NBS-LRR proteins: adaptable guards. Genome Biol. **7**: 212.

MEER, J., R. CUDMORE and K. F. MANLY, 2004 Map Manager QTX.

http://www.mapmanager.org/mmQTX.html

MEYERS, B. C., A. W. DICKERMAN, R. W. MICHELMORE, S. SIVARAMAKRISHNAN, B. W. SOBRAL *et al.*, 1999 Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. Plant J. **20**: 317-332.

MEYERS, B. C., A. KOZIK, A. GRIEGO, H. KUANG and R. W. MICHELMORE, 2003 Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell **15**: 809-834.

MICHELMORE, R. W., and B. C. MEYERS, 1998 Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res. **8**: 1113-1130.

MONDRAGON-PALOMINO, M., B. C. MEYERS, R. W. MICHELMORE and B. S. GAUT, 2002 Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. Genome Res. **12**: 1305-1315.

MONDRAGON-PALOMINO, M., and B. S. GAUT, 2005 Gene conversion and the evolution of three leucine-rich repeat gene families in *Arabidopsis thaliana*. Mol. Biol. Evol. **22**: 2444-2456.

PAN, Q., Y. S. LIU, O. BUDAI-HADRIAN, M. SELA, L. CARMEL-GOREN *et al.*, 2000a Comparative genetics of nucleotide binding site-leucine rich repeat resistance gene homologues in the genomes of two dicotyledons: tomato and arabidopsis. Genetics, **155**: 309-322.

PAN, Q., J. WENDEL and R. FLUHR, 2000b Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. J. Mol. Evol. **50**: 203-213.

PLOCIK, A., J. LAYDEN and R. KESSELI, 2004 Comparative analysis of NBS domain sequences of NBS-LRR disease resistance genes from sunflower, lettuce, and chicory. Mol. Phylogenet. Evol. **31**: 153-163.

RICHLY, E., J. KURTH and D. LEISTER, 2002 Mode of amplification and reorganization of resistance genes during recent Arabidopsis thaliana evolution. Mol. Biol. Evol. **19**: 76-84.

RITTER, E., C. GEBHARDT and F. SALAMINI, 1990 Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. Genetics, **125**: 645-654.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 Molecular cloning: A laboratory manual, 2rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

SHEN, K. A., B. S. MEYERS, M. N. ISLAM-FARIDI, D. B. CHIN, D. M. STELLY *et al.*, 1998

Resistance gene candidates identified by PCR with degenerate oligonucleotide primers map to clusters of resistance genes in lettuce. Mol. Plant-Microbe Interact. **11**: 815-823.

SMITH, S. M., and S. H. HULBERT, 2005 Recombination events generating a novel *Rp1* race specificity. Mol. Plant Microbe Interact. **18**: 220-228.

SONG, W. Y., L.Y. PI, T. E. BUREAU and P. C. RONALD, 1998 Identification and characterization of 14 transponson-like elements in the noncoding regions of members of the Xa21 family of disease resistance genes in rice. Mol Gen Genet. **258**: 449-456.

SUDUPAK, M. A., J. L. BENNETZEN and S. H. HULBERT, 1993 Unequal exchange and meiotic instability of disease–resistance genes in the Rp1 region of Maize. Genetics **133**: 119-125.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN and D. G. HIGGINS, 1997 The ClustalX windows interface:flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**: 4876-4882.

TIAN, D**.**, H. ARAKI, E. STAHL, J. BERGELSON and M. KREITMAN, 2002 Signature of balancing selection in Arabidopsis. Proc. Natl. Acad. Sci. USA **99**: 11525-11530.

TIFFIN, P., 2004 Comparative evolutionary histories of chitinase genes in the genus zea and family poaceae. Genetics **167**: 1331-1340.

WELTER, L. J., N. GOKTURK-BAYDAR, M. AKKURT, E. MAUL, R. EIBACH *et al.*, 2007 Genetic mapping and localization of quantitative trait lociaffecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L). Mol Breeding **20**: 359–374.

WEN, X. P., and X. X. DENG, 2005 Micropropagation of chestnut rose (*Rosa roxburghii* Tratt) and genetic stability assessment of the in vitro plants using RAPD and AFLP markers. J. Hortic. Sci. Biotechnol. **80**: 54-60.

XIAO, S., B. EMERSON, K. RATANASUT, E. PATRICK, C. O'NEILL *et al.*, 2004 Origin and maintenance of a broad-spectrum disease resistance locus in arabidopsis. Mol. Biol. Evol. **21**: 1661-1672.

XU, Q., X. P. WEN and X. X. DENG, 2005 Isolation of TIR and NonTIR NBS-LRR resistance gene analogues and identification of molecular markers linked to a powdery mildew resistance locus in chestnut rose (*Rosa roxburghii* Tratt). Theor. Appl. Genet. **111**: 819-830.

XU, Q., X. P. WEN and X. X. DENG. 2007 Phylogenetic and evolutionary analysis of

NBS-encoding genes in Rosaceae fruit crops. Mol. Phylogenet. Evol. **44**: 315-324.

YOUNG, N. D., 2000 The genetic architecture of resistance. Curr. Opin. Plant Biol. **3**: 285-290.

YANG, Z., 1997 PAML: A program package for phylogenetic analysis by maximum likehood. Comput. Appl. Biosci. **13**: 555-556.

YANG, Z., R. NIELSEN, N. GOLDMAN and A. M. K. PEDERSEN, 2000 Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics **155**: 431-449.

YU, Y. G., G. R. BUSS and M. A. MAROOF, 1996 Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. Proc. Nal. Acad. Sci. USA, **93**: 11751-11756.

ZHU, H., S. B. CANNON, N. D. YOUNG and D. R. COOK, 2002 Phylogeny and genomic organization of the TIR and non-TIR NBS-LRR resistance gene family in *Medicago truncatula*. Mol. Plant-Microbe Interact. **15**: 529-539.

TABLE 1

Parameter estimates for positively selected sites in NBS-encoding genes in chestnut rose

Model	Estimates of Parameters [a]	Positively Selected Sites [b]
M0: one-ratio	$\omega=0.301$	None
M7: β	p=0.456, q=1.062	Not allowed
M8: β & ω	p0=0.956, p=0.647, q=1.363	125 H (at $P>0.9$)
	p1=0.044, $\omega=1.138$	47 P, 127 L ($0.7<P<0.8$)
		57 E, 107 K ($0.6<P<0.7$)

[a] ω is the ratio of nonsynonymous substitution (d_N) that cause an amino acid change to the synonymous substitution (d_S) that dose not.

[b] P is the probability produced by Naive Empirical Bayes (NEB) analysis.

Figure legends:

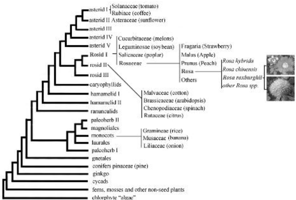
FIGURE 1.-Phylogenetic position of chestnut rose in higher plant taxa. Flower and fruit of chestnut rose was showed. Picture was modified from figure 1 of Ku et al. (2000).

FIGURE 2.-A transposon-like element was isolated neighbor to NBS domain. (A) The strategy to isolate flanking sequences, an adaptor was designed. Arrows shows the location of the PCR primers. (B) PCR products of three round TAIL-PCR: I, the first round; II, the second round; III, the third round. (C) The sequence was highly homologous to a transponson. (D) RFLP pattern confirmed the gene as a transposon. And polymorphism was detected among crops in Rosaceae. I, negative control; II, resistnat parent; III, susceptible parent; IV, *R. roxburghii* cv. Guinong No.1; V, *R. Sterilis*; VI, *Prunus persica*; VII, *Pyrus communis*; VIII, *Malus baccata*.

FIGURE 3.- Phylogenetic analysis and genetic mapping of chestnut rose NBS genes. The tree was constructed by the neighbor-joining method with human APAF1 as outgroup by PAUP* 4.0 software. Different colors denoted different sources. Bootstrap values (1000 replicates) with only values >50% were shown on the branches. On the right, three clusters were obtained by MAPMAKER and Map Manager QTXb20 softwares. Some RGH genes' phylogenetic position was connected with their locations in the genetic map by lines, illustrating two forms of cluster in chestnut rose genome, "tandem duplicated cluster" that genes within a cluster occupy the same phylogenetic lineage, and in converse named as "heterogeneous cluster".

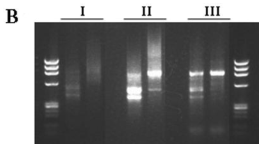
FIGURE 4.-Rapid evolution of NBS-encoding genes in chestnut rose. (A) Intra-specific variation of copy number and size between resistant parent (R) and susceptible (S). (B) Inter-specific polymorphisms in Rosaceae family. The upper showed a rapid evolving gene; the lower showed a rather old gene which show uniform band among different species in Rosaceae. (C) Nucleotide polymorphism detected among different species. G6, G5, G1 denote different cultivars in *Rosa roxburghii*; WZ, *R. sterilis*; and Rose, *R. chinensis*.

FIGURE 5.-Meiotic instability of NBS-encoding genes. (A) New homolog was produced in F_1 progenies. (B) Homologs with different size were produced in F_1 individuals. (C) Deletion was detected in gene from F_1 progenies.


asterid I
Solanaceae (tomato)
Rubiace (coffee)
asterid II Asteraceae (sunflower)
asterid III
asterid IV
asterid V
Rosid I
rosid II
rosid III
Cucurbitaceae (melons)
Leguminosae (soybean)
Salicaceae (poplar)
Rosaceae
Fragaria (Strawberry)
Malus (Apple)
Prunus (Peach)
Rosa
Others
Rosa hybrida
Rosa chinensis
Rosa roxburghii
other Rosa spp.
caryophyllids
hamamelid I
hamamelid II
ranunculids
Malvaceae (cotton)
Brassicaceae (arabidopsis)
Chenopodiaceae (spinach)
Rutaceae (citrus)
paleoherb II
magnoliales
monocots
laurales
paleoherb I
Gramineae (rice)
Musaceae (banana)
Liliaceae (onion)
gnetales
conifers pinaceae (pine)
ginkgo
cycads
ferns, mosses and other non-seed plants
chlorphyte "algae"

A

B



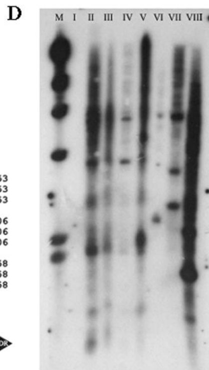
C

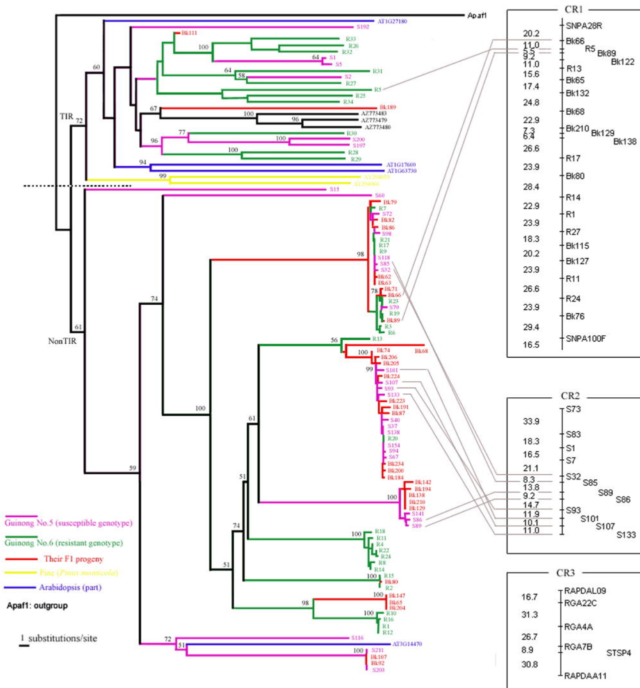
PT TCDKCQRTGNLGPRDQMPLSNIITVEIFDVWGIDFMGPFPSSFGFLYILLAVD 53
ABG37657 TCENCQKLGSISKLHMMPLNPILVIEIFECWGIDFMGPFPPSFGFLYILVAVD 53
ABG37668 TCENCQKLGSISKLHMMPLNPILVIEIFECWGIDFMGPFPPSFGFLYILVAVD 53

PT YVSKWVEAKATRTNDSKVVADFIKSNIFSRFGMPRVLISDGGSHFCNRTIEAL 106
ABG37657 YVSKWIEAIPSRTNDHKTVIKFLKENILSRFGIPRAMISDGGTHFCNKPFESL 106
ABG37668 YVSKWIEAIPSRTNDHKTVIKFLKENILSRFGIPRAMISDGGTHFCNKPFESL 106

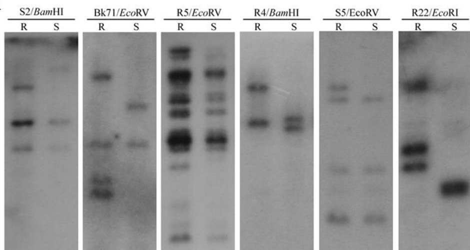
PT LKRYDVTHKVSTPYHSQTSGQAEVSNRQIKGILEKTVNPNRKDWSLRVQDAV 158
ABG37657 MKKYGITHKVATPYHPQTSGQVELANREIKQILEKTVNPNRKDWSLRLNDAL 158
ABG37668 MKKYGITHKVATPYHPQTSGQVELANREIKQILEKTVNPNRKDWSLRLNDAL 158

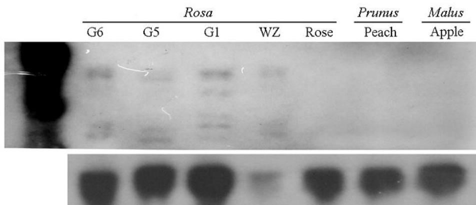
D





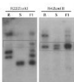
Apaf1
AT1G27180
S192
Bk111
R33
R26
R32
S1
S5
R31
S2
R27
R5
R25
R34
Bk189
AZ773483
AZ773479
AZ773480
R30
S200
S197
R28
R29
AT1G17600
AT1G63730
TIR
NonTIR
S15
S60
Bk79
R7
S72
Bk82
Bk86
S98
R21
R17
R9
S118
S85
S32
Bk62
Bk63
Bk71
Bk66
R23
S79
R19
Bk89
R3
R6
R13
Bk68
Bk74
Bk206
Bk205
S101
Bk224
S107
S93
S133
Bk223
Bk191
Bk87
S40
S37
S138
R20
S154
S94
S67
Bk234
Bk200
Bk184
Bk142
Bk194
Bk138
Bk210
Bk129
S141
S86
S89
R18
R11
R4
R22
R24
R8
R14
R15
Bk80
R2
Bk147
Bk65
Bk204
R10
R16
R1
R12
S116
AT3G14470
S211
Bk107
Bk92
S203
Guinong No.5 (susceptible genotype)
Guinong No.6 (resistant genotype)
Their F1 progeny
Pine (Pinus monticola)
Arabidopsis (part)
Apaf1: outgroup
1 substitutions/site
CR1
SNPA28R
20.2
Bk66
11.0
R5
Bk89
Bk122
9.2
11.0
R13
15.6
Bk65
17.4
Bk132
24.8
Bk68
22.9
7.3
Bk210
Bk129
6.4
Bk138
26.6
R17
23.9
Bk80
28.4
R14
22.9
R1
23.9
R27
18.3
Bk115
20.2
Bk127
23.9
R11
26.6
R24
23.9
Bk76
29.4
SNPA100F
16.5
CR2
S73
33.9
S83
18.3
S1
16.5
S7
21.1
S32
8.3
S85
13.8
S89
9.2
S86
14.7
S93
11.9
S101
10.1
S107
11.0
S133
CR3
RAPDAL09
16.7
RGA22C
31.3
RGA4A
26.7
RGA7B
8.9
STSP4
30.8
RAPDAA11



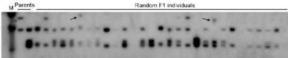


C

```
                *        20         *        40         *        60         *        80      
Rose : ...............................--------............................G.................G................T..C.. :  78
WZ   : ..........................................................C...T.................T..... :  86
G5   : ..........................---CA...AT..C....A.G........................................ :  83
G1   : .......................................................G.........................C.... :  86
G6   : .......................................................G.........................C.... :  86

          *       100         *       120         *       140         *       160         *  
Rose : ....G.....C...................A......G................T..............G................ : 164
WZ   : ....G.....C..........................G................T..............G................ : 172
G5   : .........................A.......................G.............T.............A........ : 169
G1   : ......C.......A..........A......G.GG......C.......A.........................G......... : 172
G6   : ......C.......A.................G.CG......C.......A.........................G......... : 172

            180         *       200         *       220         *       240         *       2
Rose : ...............A.............................................A.......G................ : 250
WZ   : ...............A...................C..........G......................G................ : 258
G5   : ..G..............C.G.A................................................................. : 255
G1   : ....A.....TG.T......C......T....................................A.....A...CT.C........ : 258
G6   : ....A.....TG.T......C......T....................................A.....A...CT.C........ : 258

       60         *       280         *       300         *       320         *       340    
Rose : ...........C...C.C...G.C.....C.......................G................................ : 336
WZ   : ...............C.C...G.C.....C............................................A.....A..... : 344
G5   : ..........................................................................A.....A..... : 341
G1   : .C.T...............T..A.....................C...................C...........C....G.... : 344
G6   : .C.T...............T..A.....................C...................C...........C....G.... : 344

            *       360         *       380         *       400         *       420        
Rose : ...........G..A....T................................................................ : 420
WZ   : ...T.G.............................................................................- : 427
G5   : ...T.G..........................GA............T.....C.A............................. : 425
G1   : ......................AG......T................A.................................... : 428
G6   : ......................AG......T................A.................................... : 428
```


A
R
S
F1
R
S
F1

C


B
M
Parents
Random F1 individuals