16. Romano, A., Raemakers, K., Visser, R. et al., Transformation of potato (*Solanum tuberosum*) using particle bombardment, Plant Cell Rep., 2001, 20: 198－204.
17. Romano, A., Raemakers, K., Bernardi, J. et al., Transgene organization in potato after particle bombardment-mediated (co-) transformation using plasmids and gene cassettes, Transgenic Res., 2003, 12: 461－473.
18. Romano, A., Vreugdenhil, D., Jamar, D. et al., Evidence of medium-chain-length polyhydroxyoctanoate accumulation in transgenic potato lines expressing the *Pseudomonas oleovorans* Pha-C1 polymerase in the cytoplasm, Biochem. Eng. J., 2003, 16: 135－143.
19. Wróbel, M., Zebrowski, J., Szopa, J., Polyhydroxybutyrate synthesis in transgenic flax, Journal of Biotechnology, 2004, 107: 41－45.
20. Menzel, G., Harloff, H. J., Jung, C., Expression of bacterial poly (3-hydroxybutyrate) synthesis genes in hairy roots of sugar beet (*Beta vulgaris* L.), Appl. Micro. Bioech., 2003, 60: 571－576.
21. LÖssl, A., Eibl, C., Harloff, J. et al., Polyester synthesis in transplastomic tobacco (*Nicotiana tabacum* L.): Significant contents of polyhydroxybutyrate are associated with growth reduction, Plant Cell. Rep., 2003, 21: 891－899.
22. Romano, A., van der Plas, L. H., Witholt, B. et al., Expression of poly-3-Hydrozyalkanoate (PHA) polymerase and acyl-CoA-transacylase in plastids of transgenic potato leads to the synthesis of a hydrophobic polymer, presumably medium-chain-length PHAs, Planta, 2005, 220(3): 455－464.
23. Gong, X. S., Yan, L. F., The method of isolation chloroplast DNA from higher plants, Chinese Science Bulletin, 1991, 36(6): 467－469.
24. Hiratsuka, J., Shimada, H., Whittier, R. et al., The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals, Mol. Gen. Genet., 1989, 217: 185－194.
25. Shinozaki, K., Ohme, M., Tanaka, M. et al., The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression, EMBO J., 1986, 5: 2043－2049.
26. Wakasugi, T., Sugita, M., Tsudzuki, T. et al., Updated gene map of tobacco chloroplast DNA, Plant Mol. Biol. Rep., 1998, 16: 231－241.
27. Svab, Z., Maliga, P., High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene, Proc. Natl. Acad. Sci. USA, 1993, 90: 913－917.
28. Lin, J. A., An, J. L., Hong, J., Applied biological electron microscopy technique, Liao Ning Scientific Publishing Company, 1998, 48－52.
29. Huijberts, G., Eggink, G., Waard, de. P. et al., Pseudomonas putida KT2442 cultivated on glucose accumulates poly (3-hydroxyalkanoates) consisting of saturated and unsaturated monomers, Appl. Environ. Microbiol., 1992, 58: 536－544.
30. Rehm, B., Kruger, N., Steinbuchel, A., A new metabolic link between fatty acid de novo synthesis and polyhydroxyalkanoic acid synthesis, J. Biol. Chem., 1998, 273: 24044－24051.

(Received January 21, 2005; accepted April 4, 2005)

Chinese Science Bulletin 2005 Vol. 50 No. 11 1120—1125

Genome-wide identification of R genes and exploitation of candidate RGA markers in rice

WANG Xusheng[1], WU Weiren[1,2], JIN Gulei[2] & ZHU Jun[1]

1. Institute of Bioinformatics, Zhejiang University, Hangzhou 310029, China;
2. College of Crop Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China

Correspondence should be addressed Wu Weiren (email: wuwr@zju.edu.cn)

Abstract **By scanning the whole genomic sequence of *japonica* rice using 45 known plant disease resistance (R) genes, we identified 2119 resistance gene homologs or analogs (RGAs) and verified that RGAs are not randomly distributed but tend to cluster in the rice genome. The RGAs were classified into 21 families according to their functional domain based on Hidden Markov model (HMM). By comparing the RGAs of *japonica* rice with the whole genomic sequence of *indica* rice, we found 702 RGAs allelic between the two subspecies and revealed that 671 (95.6%) of them have length difference (InDels) in their genomic sequences (including coding and non-coding regions) between the two subspecies, suggesting that RGAs are highly polymorphic between the two subspecies in rice. We also exploited 402 PCR-based and co-dominant candidate RGA markers by designing primer pairs on the regions flanking the InDels and validating them via e-PCR. The length differences of the candidate RGA markers between the two subspecies are from 1 to 742 bp, with an average of 10.26 bp. All related information of the RGAs is available from our web site (http://ibi.zju.edu.cn/RGAs/index.html).**

Keywords: rice, resistance gene, resistance gene analog (RGA), polymorphism, molecular marker.

DOI: 10.1360/982005-213

Plant disease resistance (R) genes are the sort of genes that determine the specific recognition and activate the defense response of host plants to pathogens. They interact with pathogen avirulence genes. Classical genetics suggests that the plant-pathogen interaction follows a "gene by gene" mechanism, and a receptor-ligand model has been proposed to explain the hypothesis[1,2]. Since 1992, more than 40 R genes have been cloned from various plant species such as *Oryza sativa*, *Arabidopsis thaliana*, *Nicotiana tabacum* and *Linum usitatissimum* by means of map-based cloning or transposon-tagging[3]. Studies have shown that R genes have some highly conserved amino acid sequences. According to their protein structures and locations in cells, R genes could be roughly

divided into 5 groups[4,5]: (i) NBS-LRR, encoding intracellular acceptor proteins containing nucleotide binding site (NBS) and leucine-rich repeat (LRR) domains. It consists of two sub-classes. One is TIR-NBS-LRR, represented by *Arabidopsis thaliana RPP5* gene, *Nicotiana tabacum N* gene, and *Linum usitatissimum L6* and *M* genes; the other is CC-NBS-LRR, represented by *Arabidopsis thaliana RPS2* and *RPM1* genes, *Lycopersicon esculentum I2* gene, and *Hordeum vulgare Mla1* gene. (ii) Intercellular serine/threonine protein kinase (PK) genes, such as *Lycopersicon esculentum pto* gene and *Hordeum vulgare Rpg1* gene. (iii) LRR-TM, encoding extracellular acceptor proteins with LRR domain, single transmembrane region and short cytoplasmic carboxyl terminus, such as *Cf-2*, *Cf-4*, *Cf-5* and *Cf-9* genes of *Lycopersicon pimpinellifolium*. (iv) PK-LRR-TM, encoding extracellular proteins with LRR domain, single transmembrane region and cytoplasmic kinase domain, including *Oryza sativa Xa-21* gene and *Arabidopsis thaliana FLS2* gene. (v) SA-CC, encoding intracellular proteins with signalling anchor and a coiled-coil domain, such as *RPW8.2* and *RPW8.1* genes of *Arabidopsis thaliana.* In addition, there are some R genes with other domains such as *Zea mays Hm1* and Hsl^{Pro-1} genes and *Lycopersicon esculentum Asc* gene.

R genes are a huge multigene family. Knowledge on R genes is still very limited although more than 40 R genes have been cloned. Nowadays, the work of genome sequencing has been basically fulfilled in *Arabidopsis* and rice. The drafts of genomes of two rice subspecies (*indica* and *japonica*) have been completed[6,7], and precisely assembled sequences of chromosomes 1, 4 and 10 of *japonica* have been published[8−10]. These achievements provide a unique opportunity for genome-scale investigation of R genes. Meyers et al. analyzed the distribution of NBS-LRR-type genes in *Arabidopsis* genome[11]. Preliminary analysis of rice genome sequence has revealed that there are a large number of R genes in rice genome[6,7] and they tend to cluster[12,13]. Monosi et al. found that there are about 500 R genes of NBS-LRR type in rice, but none with TIR domain[14]. Using 18 known R genes to analyze the genome sequences of *Arabidopsis* and *japon- ica* rice, Chelkowski and Koczyk[15,16] identified 549 and 1744 R genes in *Arabidopsis* and rice, respectively; 597 of the R genes in rice are of NBS-LRR type. It is seen that all of these studies based on genome sequences are mainly focused on the analysis of NBS-LRR-type R genes.

Molecular markers are powerful tools for current genetic research. Since the concept of molecular marker was proposed in 1980[17], molecular markers have been applied to various research fields including construction of genetic maps, gene mapping, gene cloning, genome comparison, analysis of genetic diversity, marker-assistant breeding and so on[18−20]. Using R gene analogs (RGAs) as DNA probes, the restriction fragment length polymorphisms (RFLPs) at these loci can be detected and therefore can be used as RFLP markers, usually called RGA markers. RGA markers are particularly useful for the cloning and marker-assistant selection of R genes because RGAs themselves are potential R genes and often distributed as clusters in genomes. However, since RGA markers are based on the complicated technology of RFLP analysis, they are not widely utilized at present.

In this study, we performed a more detailed investigation on the number, distribution and classification of RGAs in the rice genome using bioinformatics approaches and published genomic sequences of *indica* and *japonica* rice so as to better understand rice R genes at the whole genome level. In addition, we also investigated the genetic polymorphisms (SNPs and InDels) of RGA loci between the two subspecies of rice, designed PCR primers for the polymorphic RGAs and validated the primers via e-PCR. The work will facilitate the exploitation of convenient and practical new PCR-based RGA markers.

1 Materials and methods

1.1 Sources of genomic and protein sequences

The data of genomic and protein sequences of *japonica* rice (Nipponbare) and *indica* rice (93-11) used were downloaded from TIGR web site (http://www.tigr.org/) and Beijing Genomics Institute web site (http://rise.genomic.org.cn/). Both sets of data were updated in April 2004. Data analysis was conducted on IBM P650 server using IBM AIX Unix operation system.

1.2 Search of rice RGAs

RGA sequences of *japonica* rice were searched by comparing 45 known R genes to the *japonica* rice protein database using BLASTP[21] (with parameters $E < 10^{-10}$ and length > 80% of query gene). A *japonica* rice RGA protein database was then established with the RGAs identified after removing duplicate clones in the *japonica* rice database. Meanwhile, the DNA sequences of the RGAs were acquired in term of the best hits of BLASTP search. To further validate the identified RGAs, the RGA sequences were compared with CDS sequences published by TIGR and mismatched sequences were removed.

1.3 Classification and chromosomal distribution of the rice RGAs

According to the domain list established above, domains possessed by each RGA sequence were searched in the established database using the "hmmsearch" function of program Hmmer[22]. The possibility for a RGA sequence to contain the coiled-coil (CC) structure was analyzed with program Pepcoil[23] and an RGA was thought to have the structure when the probability was greater than 90%. Transmembrane (TM) domains were analyzed with program TM-HMM[24]. The RGAs were classified according to their domains and the distribution of each type

of RGA on different chromosomes was counted.

1.4 Identification and exploitation of RGA polymorphisms between two subspecies

All RGAs found from *japonica* rice were compared with the *indica* genome sequence database using program TBLASTN[21] to identify their alleles in *indica* rice. To avoid non-allelic alignment, a strict expectation value ($E < 10^{-20}$) was used for TBLASTN search. For each pair of candidate *indica* and *japonica* RGA alleles obtained, alignment analysis was further conducted with program SIM4[25] to remove any pair with identity $\leqslant$ 85% and a gap bigger than 200 bp in both sequences. Then, program diffseq from EMBOSS suits[26] was used to identify SNPs and InDels in the genomic sequences of the RGAs, and those with InDels were taken as candidate RGA markers. For each candidate RGA marker, two 100-bp sequences from *japonica* rice flanking the InDel site on both sides were extracted and joined to make a 200-bp template sequence for designing PCR primers using program ePrime3[27]. In general, we selected the optimal pair of primers of the five resultant pairs provided by ePrimer3 for each RGA. In addition, we required that the paired primers should be located on each side of the target InDel site and generate the target PCR product not longer than 1000 bp. Finally, the primers were evaluated by electronic PCR (e-PCR)[28]. The RGA markers of rice were named with an abbreviation OSR followed by a unique number, for example, OsRGA0255. The whole process described above was implemented automatically by a program written in Perl script.

2 Results

2.1 Number, density and distribution of RGAs on chromosomes in *japonica* rice

2119 RGAs were identified from *japonica* by searching the *japonica* rice protein database (Table 1). The number of RGAs on each chromosome ranged from 113 (chromosome 3) to 333 (chromosome 1), with an average of 176 and the most on chromosomes 1, 2 and 11. The density of RGAs on each chromosome ranged 0.66－2.42 RGAs/cM or 2.68－9.44 RGAs/Mb. Analysis of RGA's distribution in the rice genome according to the published assembled rice genomic sequence TIGR revealed that RGAs tend to exist as clusters (with 2－12 RGAs in a cluster in most cases). For example, 10 RGAs were found in clone AP003209 on chromosome 1.

2.2 Structural classification of rice RGAs

R genes are generally classified into 5 groups, among which the NBS-LRR group is the largest[4,5]. In this study, however, we classified rice RGAs more detailedly (into 21 classes in total) according to the structures and domains of R genes (Fig. 1). The PK class (*Pto*, *Fen*, *Lr10*) has the largest number of RGAs, accounting for 26.7% (566/2119) of the total, and followed by TM-LRR, accounting for 20.5% (435/2119). It is necessary to point out that in this study the RGAs possessing NBS or LRR domains are divided into 9 classes, namely TM-LRR, PK-LRR, NBS-LRR, CC-NBS-LRR, CC-LRR, CC-NBS, PK-NBS-LRR, PK-NBS and CC-PK-LRR. Therefore, none of the class is the largest. However, if we take them all as the NBS-LRR type, then the number of RGAs in this class will account for more than half of the total number (1091/2119). There were 77 RGAs belonging to the toxin reductase type represented by maize *Hm1* gene, the first R gene to be cloned. This type domain can also be combined with CC structure to form a CC-Hm1 type. Three RGAs of this type were found. Groups PK-NBS, CC-PK-LRR, TIR, Hs1 and Pad4 only contain a single member in rice. Structural analysis showed that most of them are pseudogenes or functionless genes. Pan et al.[29] suggest that

Table 1 Number and density of RGAs on each chromosome in *japonica* rice

Chromosome	Chromosome length		No. of RGAs	Density of RGAs	
	cM	Mb		cM^{-1}	Mb^{-1}
1	181.8	44.3	333	1.83	7.51
2	157.9	39.9	217	1.37	5.44
3	166.4	41.1	110	0.66	2.68
4	129.6	38.2	195	1.50	5.10
5	122.3	33.2	134	1.09	4.04
6	126.3	31.7	190	1.50	5.99
7	118.6	35.0	125	1.05	3.57
8	121.2	27.6	158	1.30	5.72
9	93.5	21.6	133	1.42	6.16
10	83.8	25.7	113	1.35	4.40
11	117.9	30.2	285	2.42	9.44
12	109.5	30.6	126	1.15	4.12
Whole genome	1528.8	399.1	2119	1.39	5.31

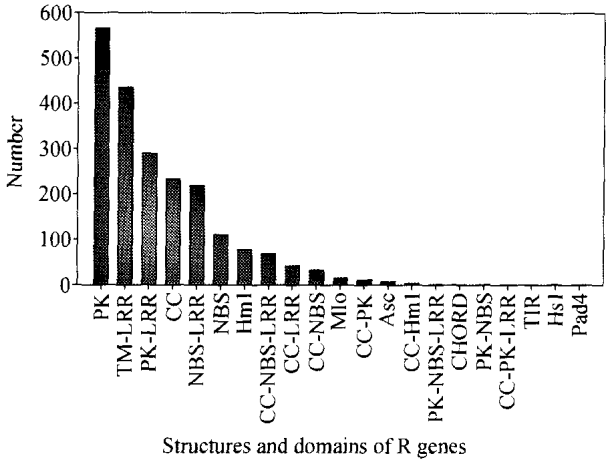

Fig. 1. Types and number distribution of RGAs in rice. PK, serine-threonine protein kinase; TM, transmembrane protein; LRR, leucine-rich repeat; CC, coiled-coil; NBS, nucleotide binding site; Hm1, *Hm1* gene of *Zea mays*; CHORD, cysteine- and histidine-rich domain; TIR, Toll/interleukin-1 receptor; Mlo, Asc, Hs1 and Pad4 represent the characteristic domains of corresponding genes, respectively. Hyphens indicate the junction of two domains.

the NBS-LRR group was subdivided into two subgroups, TIR-NBS-LRR and CC-NBS-LRR, during the process of differentiation between dicot and monocot plants. Our research shows that there is no RGAs of TIR-NBS-LRR type in rice. This is similar to the case of beet[30], but 117 RGAs of this type have been found in *Arabidopsis*[18]. Related data of the RGAs identified in this study are obtainable from our web site (http://ibi.zju.edu.cn/RGAs/index.html).

2.3 RGA polymorphisms between rice subspecies

A total of 1860 *indica* R-gene homologs were identified by aligning the *japonica* RGA sequences to the *indica* genomic sequence using program TBLASTN, and 861 remained after removing those that were duplicated or not well matched (matched sequence length < 80 bp and identity < 40%) or not consistent with the CDS in the TIGR database. Some *indica* homologs appearing nonallelic to *japonica* because of being located on different chromosomes from the expected ones were further discarded. Finally, 702 RGAs showing allelism between *indica* and *japonica* were acquired. Analysis of these RGAs with program sim4 revealed that 671 (95.6%) of them contained InDels between *indica* and *japonica*, indicating that there are very high polymorphisms in RGAs between the two subspecies. By designing primers on both sides of the InDels and selecting those that could generate unique products as expected from e-PCR, a total of 402 candidate rice RGA markers were obtained. There were 269 polymorphic RGAs that could not be exploited as candidate markers. The reasons might be: (i) the specificity of some primers was not very high due to structural similarity among RGAs so that they could not generate unique products; (ii) some primers might generate too large (exceeding the limit of 1000 bp) PCR products and therefore were discarded; and (iii) some primers could not completely match the genomic sequence of *indica* rice 93-11 because the primers were designed according to the genomic sequence of *japonica* rice Nipponbare. Information about these candidate RGA markers (including their primers, sequences, located BAC clones of Nipponbare or scaffolds of 93-11, etc.) has been made public in our web site (http://ibi.zju.edu.cn/RGAs/index.html). Physical mapping of the candidate RGA markers was conducted according to the assembled rice genomic sequence put out by TIGR. Result shows that the candidate RGA markers are also non-randomly distributed and tend to exist as clusters in the rice genome, similar to the situation of all RGAs (Fig. 2). There are large blanks on some chromosomal regions (e.g., the long arm of chromosome 1).

The length difference (LD) of candidate RGA markers varies from 1 to 742 bp with an average of 10.15 bp and appears to follow an exponential distribution. Most (68.16%) of the LDs are < 5 bp; 24.88% of them fall in the range of 5—30 bp; only 6.96% are > 30 bp (Fig. 3). It is worth pointing out that we have found 14 RGAs of which the length differences are longer than 1 kb between the two subspecies and the inserted sequences all contain an independent and complete gene structure. Homology analyses have revealed that these 14 inserted sequences are all closely related to acceptor proteins. As R genes themselves are a kind of acceptor proteins, it would be an interesting question whether the insertion of acceptor protein-related genes in R genes implies some important biological mechanisms.

3 Discussion

We have identified 2119 RGAs in the rice genome by sequence homology comparison and functional domain analysis. The results indicate that R genes are abundant in the rice genome, being a very large gene family. Certainly, apart from R genes, there might be some functionless genes or pseudogenes among the RGAs. To determine what of the RGAs are real R genes, we compared all RGAs with 32127 published rice full-length cDNAs using BLAST program[31]. We found that 1851 RGAs could well match the cDNAs. Thereby, they might be real R genes (certainly, there might possibly be a few expressed pseudogenes). The remaining 268 RGAs could be: (i) genes not included in the cDNA database because there are about 50000 genes in the rice genome; (ii) pseudogenes not expressed; (iii) genes induced by specific pathogens. With the enrichment and improvement of the rice full-cDNA database, the identities of these RGAs will be clarified. The future challenge will be to elucidate the functions of all R genes. Acquiring genome-wide information of RGAs using bioinformatics approaches will greatly promote the functional research of R genes.

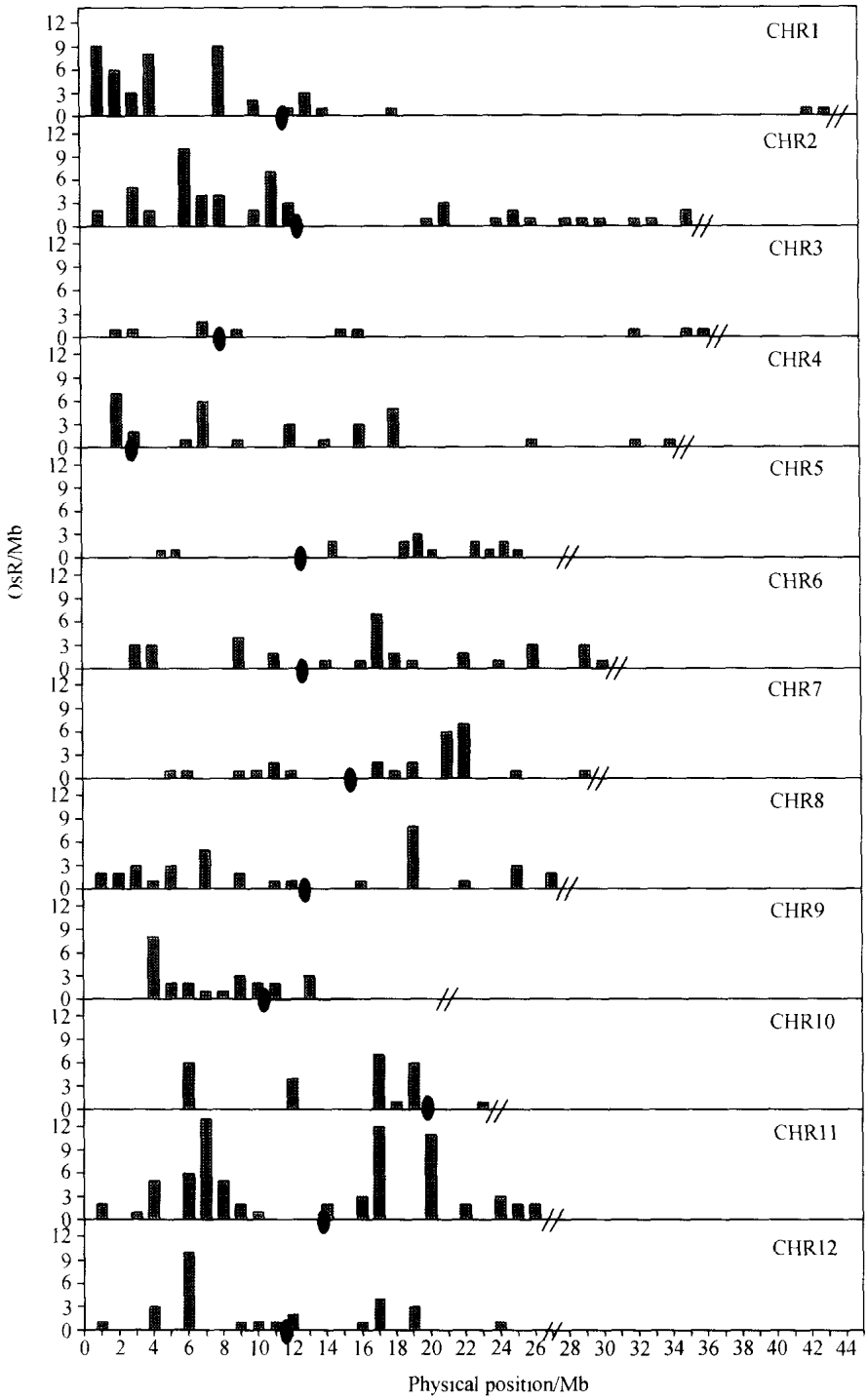

Fig. 2. Distribution of the candidate RGA markers in the rice genome. The double slashes indicate the terminal of chromosome; the black ovals indicate the position of centromere.

The traditional RGA marker is a kind of RFLP marker. It is not convenient to use and has a very limited number exploited because of its technical complexity and high cost. In this study, we have developed PCR-based candidate RGA markers using rice genomic sequence data and bioinformatics approaches. This will make the RGA length polymorphism to be a practical molecular marker. We have also successfully developed markers of intron length polymorphism (ILP) in rice using the similar methods (to be published). We have found by experiments that rice ILP markers have obvious subspecies specificity. This enable us to infer that the candidate RGA markers exploited on the basis of sequence comparison between *indica* and *japonica* rice would be highly subspecies-spe-

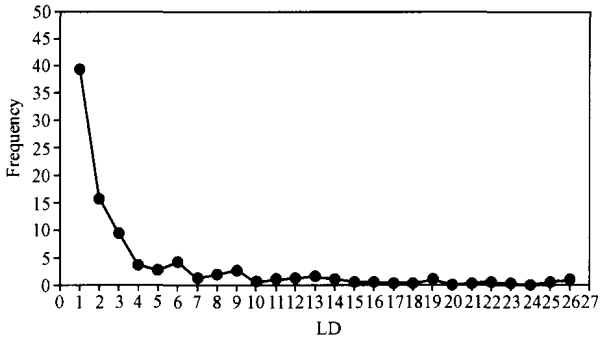

Fig. 3. Frequency distribution of length difference (LD) of candidate RGA markers between two subspecies. 30 markers with LD > 26 are not presented.

cific, too. The feature would hopefully enable RGA markers to be useful for between-subspecies hybridization breeding and utilization of between-subspecies heterosis in rice. Moreover, it has been known that RGAs are non-randomly distributed as clusters in the rice genome; and the distribution of the candidate RGA markers developed in this study is well representative for that of all RGAs in the rice genome (Fig. 3). In addition, RGA markers themselves are candidate R genes. Therefore, RGA markers are helpful for fast mapping of R genes and therefore speed up the process of mapping and positional cloning of R genes.

Aknowledgements This work was partially supported by the National High-Tech Research and Development Program of China (Grant Nos. 2003AA207160 & 2002AA234031) and the Natural Science Foundation of Fujian Province (Grant No. B9910011).

References

1. Flor, H. H., The complementary genic systems in flax and flax rust, Adv. Genet., 1956, 8: 29－54.
2. Flor, H. H., Current status of the gene-for-gene concept, Annu. Rev. Phytopathol., 1971, 9: 275－296.
3. Dangl, J. L., Jones, J. D., Plant pathogens and integrated defense responses to infection, Nature, 2001, 411: 826－833.
4. Hammond-Kosack, K. E., Jones, J. D. G., Plant disease resistance genes, Annu. Rev. Plant Physiol. Plant Mol. Biol., 1997, 48: 575－607.
5. Jones, J. D., Putting knowledge of plant disease resistance genes to work, Curr. Opin. Plant Biol., 2001, 4: 281－287.
6. Goff, S. A., Ricke, D., Lan, T. H. et al., A draft seqeunce of the rice genome (*Oryza sativa* L. ssp. *japonica*), Science, 2002, 296: 92－100.
7. Yu, J., Hu, S., Wang, J. et al., A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*), Science, 2002, 296: 79－92.
8. Sasaki, T., Matsumoto, T., Yamamoto, K. et al., The genome sequence and structure of rice chromosome 1, Nature, 2002, 420: 312－316.
9. Feng, Q., Zhang, Y., Hao, P. et al., Sequence and analysis of rice chromosome 4, Nature, 420: 316－320.
10. The Rice Chromosome 10 Sequencing Consortium, In-depth view of structure, activity, and evolution of rice chromosome 10, Science, 2003, 300: 1566－1569.
11. Meyers, B. C., Kozik, A., Griego, A. et al., Genome-wide analysis of NBS-LRR encoding genes in *Arabidopsis*, Plant Cell, 2003, 15: 809－834.
12. Meyers, B. C., Dickerman, A. W., Michelmore, R. W. et al., Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily, Plant J., 1999, 20: 317－332.
13. Bai, J., Pennill, L. A., Ning, J., Lee, S. W. et al., Diversity in nucleotide binding site-leucine rich repeat genes in cereals, Genome Res., 2002, 12: 1871－1884.
14. Monosi, B., Wisser, R. J., Pennill, L. et al., Full-genome analysis of resistance gene homologues in rice, Theor. Appl. Genet., 2004, 109: 1434－1447.
15. Chelkowski, J., Koczyk, G., Resistance gene analogues of Arabidopsis thaliana: Recognition by structure, J. Appl. Gen., 2003, 44: 311－321.
16. Koczyk, G., Chelkowski, J., An assessment of the resistance gene analogues of *Oryza sativa* ssp. *japonica*: Their presence and structure, Cell Mol. Biol. Lett., 2003, 8: 963－972.
17. Botstein, D., White, R. L., Skolnick, M. et al., Construction of a genetic linkage map in the man using restriction fragment length polymorphisms, Am. J. Hum. Genet., 1980, 32: 314－331.
18. McCouch, S. R., Chen, X., Panaud, O., Microsatellite mapping and applications of SSLP's in rice genetics and breeding, Plant Mol. Biol., 1997, 35: 89－99.
19. Coburn, J., Temnykh, S., Paul, E. et al., Design and Application of Microsatellite Marker Panels for Semi-automated Genotyping of Rice (*Oryza sativa* L.), Crop Sci., 2002, 42: 2092－2099.
20. Ponce, M. R., Robles, P., Micol, J. L., High throughput genetic mapping in *Arabidopsis thaliana*, Mol. Gen. Genet., 1999, 261: 408－415.
21. Altschul, S., Madden, T., Schaffer, A. et al., Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, Nucleic Acids Res., 1997, 25: 3389－3402.
22. Eddy, S. R., Profile hidden Markov models, Bioinformatics, 1998, 14: 755－763.
23. Lupas, A., van Dyke, M., Stock, J., Predicting coiled coils from protein sequences, Science, 1991, 252: 1162－1164.
24. Krogh, A., Larsson, B., von Heijne, G, et al., Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes, Journal of Molecular Biology, 2001, 305: 567－580.
25. Florea, L., Hartzell, G., Zhang, Z. et al., A computer program for aligning a cDNA sequence with a genomic DNA sequence, Genome Res., 1998, 8: 967－974.
26. Rice, P., Longden, I., Bleasby, A. J. Internet resources: EMBOSS, Trends in Genetics, 2000, 16: 276－277.
27. http://www.hgmp.mrc.ac.uk/.
28. Schuler, G. D., Sequence mapping by electronic PCR, Genome Res., 1997, 7: 541－550.
29. Pan, Q. L., Wendel, J., Fluhr, R., Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes, J. Mol. Evol., 2000, 50: 203－213.
30. Tian, Y., Fan, L., Thurau, T. et al., The absence of TIR-type resistance gene analogues in the sugar beet (*Beta vulgaris* L.) genome, J. Mol. Evol., 2004, 58: 40－53.
31. Kikuchi, S., Satoh, K., Nagata, T. et al., Collection, mapping, and annotation of over 28000 cDNA clones from *japonica* rice, Science, 2003, 301: 376－379.

(Received February 5, 2005; accepted April 5, 2005)