Isolation and Characterization Of Rice *R* Genes: Evidence for Distinct Evolutionary Paths in Rice and Maize

Jianping Hu, Beth Anderson and Susan R. Wessler

Department of Botany and Genetics, University of Georgia, Athens, Georgia 30602

Manuscript received September 27, 1995
Accepted for publication November 16, 1995

ABSTRACT

R and *B* genes and their homologues encode basic helix-loop-helix (bHLH) transcriptional activators that regulate the anthocyanin biosynthetic pathway in flowering plants. In maize, *R/B* genes comprise a very small gene family whose organization reflects the unique evolutionary history and genome architecture of maize. To know whether the organization of the *R* gene family could provide information about the origins of the distantly related grass rice, we characterized members of the *R* gene family from rice *Oryza sativa*. Despite being a true diploid, *O. sativa* has at least two *R* genes. An active homologue (*Ra*) with extensive homology with other *R* genes is located at a position on chromosome *4* previously shown to be in synteny with regions of maize chromosomes *2* and *10* that contain the *B* and *R* loci, respectively. A second rice *R* gene (*Rb*) of undetermined function was identified on chromosome *1* and found to be present only in rice species with AA genomes. All non-AA species have but one *R* gene that is *Ra*-like. These data suggest that the common ancestor shared by maize and rice had a single *R* gene and that the small *R* gene families of grasses have arisen recently and independently.

THE *R* and *B* genes of maize comprise a very small gene family that regulates anthocyanin biosynthesis by activating transcription of some of the structural genes in the pathway. This gene family can have as few as two members with a single *B* gene on chromosome *2* and a single *R* gene on chromosome *10*. However, *R* complexes with multiple *R* genes and gene fragments have also been described (STADLER and NUFFER 1953; DOONER and KERMICLE 1971, 1976; ROBBINS *et al.* 1989, 1991; WALKER *et al.* 1995).

Consistent with their regulatory role was the finding that *R* and *B* genes encode homologous proteins with features characteristic of transcriptional activators including a bHLH domain (LUDWIG *et al.* 1989; CHANDLER *et al.* 1989) and nuclear localization signals (SHIEH *et al.* 1993). Furthermore, *R* and *B* gene products can activate transcription of chimeric genes containing structural gene promoters (isolated from the *Bz1* and *A1* genes) fused to reporter genes (KLEIN *et al.* 1989; GOFF *et al.* 1990, 1992; ROTH *et al.* 1991). Activation of the structural genes in the pathway is also dependent on the presence of a second transcriptional activator containing a *myb* domain that is encoded by the duplicate *C1* and *Pl* genes (PAZ-ARES *et al.* 1987; GOFF *et al.* 1992; CONE *et al.* 1993).

Although activation of the anthocyanin biosynthetic pathway requires the action of two classes of regulatory proteins (*R* or *B* and *C1* or *Pl*), the diverse pigmentation patterns displayed by maize strains primarily reflect the allelic diversity of only the *R* and *B* loci. Genetic studies indicated that *R* and *B* alleles condition distinct patterns of pigmentation in the plant (STADLER and FOGEL 1943, 1945; STYLES *et al.* 1973). Molecular studies, in turn, have shown that *R* and *B* alleles differ in their pattern of gene expression rather than in their protein product. Thus, it is the presence or absence of the *R/B* protein that determines whether a particular cell type will be pigmented (LUDWIG *et al.* 1989; RADICELLA *et al.* 1992). Furthermore, the overall pattern of pigmentation in the plant represents the additive effects of the particular allele at the *B* locus and the composition of the *R* locus or *R* complex (LUDWIG and WESSLER 1990).

Molecular characterization of *B* and *R* alleles has indicated that the proliferation and diversification of this small gene family has occurred very recently (PURUGGANAN and WESSLER 1994). The duplicate *B* and *R* loci are located on regions of chromosomes *2* and *10* shown to be in synteny and thought to be derived from an ancient polyploidization event (ANDERSON 1945; HELENTJARIS *et al.* 1988; ROBBINS *et al.* 1989; WHITKUS *et al.* 1992; AHN and TANKSLEY 1993). In contrast, the duplication of *R* genes on chromosome *10* may have arisen as a byproduct of transposable element activity (ROBBINS *et al.* 1991; WALKER *et al.* 1995). Thus the evolutionary history and unique genomic architecture of maize appear to have contributed to the diversification of the *R/B* gene family.

Although anthocyanin regulation has been well characterized only in maize, recent experiments suggest that aspects of the regulatory network have been conserved in all flowering plants. First, maize *R* and *B* genes can activate pigmentation in other plants including the

Corresponding author: Susan R. Wessler, Department of Genetics, Life Sciences Bldg., University of Georgia, Athens, GA 30602. E-mail: sue@dogwood.botany.uga.edu

grasses sorghum (CASAS *et al.* 1993) and wheat (BILANG *et al.* 1993), and the dicots tobacco, Arabidopsis and petunia (LLOYD *et al.* 1992; QUATTROCCHIO *et al.* 1993; GALWAY *et al.* 1994). Second, *R* homologues have been isolated from *Antirrhinum majus* (*delila*) and petunia (*jaf13*) and shown to be involved in regulation of the anthocyanin pathway. Interestingly, these *R* homologues also appear to be members of small gene families (GOODRICH *et al.* 1992; QUATTROCCHIO 1994).

Domesticated rice is another member of the grass family that, like maize, is the product of intensive human selection. In addition, rice and its wild relatives display diverse pigmentation patterns that may also reflect proliferation of *R* alleles. Despite the similarities between these two important crop plants, their genome architecture and evolutionary history differ dramatically. Unlike maize, domesticated rice is a true diploid with no cytological or molecular evidence for an ancient polyploidization event (OKA 1988). In addition, rice has the smallest genome of all of the members of the grass family analyzed to date (ARUMUGANATHAN and EARLE 1991). The approximate sixfold difference in genome size between maize and rice is thought to result from, in part, a higher level of mobile element activity in maize (FLAVELL 1984).

Given the role played by polyploidization and transposable elements in shaping the *R/B* gene family of maize, we were interested in determining if *R/B* genes existed as a gene family in rice. If so, we could use the available domesticated and wild rice species to understand how the gene family may have evolved. As a first step, we took advantage of the cloned maize *R (Lc)* gene to isolate its homologue from rice. Here we describe the identification of two *R* genes in *Oryza sativa,* designated *Ra* and *Rb,* that are mapped to chromosome *4* and *1,* respectively. The active *Ra* gene shows similarity along its entire length with the maize *R (Lc)* gene and the Antirrhinum *delila* gene and is able to induce anthocyanin expression in maize. Our data indicates that the common ancester of maize and rice may have had only a single *R* gene and that the small *R* gene families of grasses have arisen recently and independently.

MATERIALS AND METHODS

Plant materials: The *O. sativa* strains Purple522, Green412, IR36, BS125, the wild species *O. rufipogon, O. longistaminata, O. officinalis, O. australiensis, O. eichingeri,* and the rice mapping population were kindly provided by GARY KOCHERT (University of Georgia). SH88-2 was obtained from MARC COHN (Louisiana State University). IR29, Calmochi 201, Star Bonnet, Susono Mochi, Stg 772199, and Malagkit Songsong were from the International Rice Research Institute. The maize inbred line W22 (*r-g, A1, A2, Bz1, C1, C2, pl, B-b*), provided by JERRY KERMICLE (University of Wisconsin) was used for the bombardment assay.

DNA extraction and DNA gel blot analysis: DNA was isolated from 2 g of rice leaf tissue using the DNA miniprep procedure of DELLAPORTA *et al.* (1983), except that phenol/chloroform extraction was done before the final isopropanol precipitation. Five micrograms of DNA that had been pretreated with RNase was digested with restriction enzymes (BRL) at 37° for ≥5 hr or overnight. Conditions for electrophoresis, blotting, prehybridization and hybridization were modified from SAMBROOK *et al.* (1989). DNA was transferred to GeneScreen (NEN), and blots were prehybridized for 3–5 hr in 1 M NaCl, 1% SDS, 50 mM Tris HCl pH 7.5 and 200 μg/ml ssDNA. Probes were made by the Random Primer DNA Labeling System (BRL) and added directly to the prehybridization solution. Filters were hybridized overnight at 65°, washed in 0.1× SSC/0.5% SDS at 65°, and exposed to X-ray film (Kodak). When necessary, the probes were stripped off the filters by incubating for 30 min in 0.4 N NaOH at 42°, then for 30 min in 0.5 M Tris pH 8.0/2× SSC/0.1% SDS at 42°.

RNA extraction and gel blot analysis: RNA was isolated from 20 g of leaf tissue from either Purple522 or Green412 using the procedure of LUDWIG *et al.* (1989). Five micrograms of polyA$^+$ RNA from either Purple522 or Green412 were fractionated in a 1.2% agarose/formaldehyde gel, transferred to a Magnagraph membrane (Fisher) and prehybridized for 2 hr at 65° in 5% SDS/0.33 M sodium phosphate pH 7.0/0.1 M EDTA/0.15 mg/ml heparin. The filter was probed with a 2.4-kb *XbaI/ApaI* fragment containing the full-length *Ra* cDNA that was labeled using the random primer protocol (BRL), washed under stringent conditions (0.1× SSC/0.5% SDS at 65° twice for 30 min) and exposed to X-ray film (Kodak) overnight. To reprobe the filter, labeled *Ra* was first removed by incubating the filter in 5 mM Tris pH 8.0/0.02% $Na_4P_2O_7 \cdot 10H_2O$/0.1% SDS at 65° for 2 hr and then reprobed with maize actin cDNA under conditions of reduced stringency (hybridized as above but at 55° and washed in 2× SSC/0.5% SDS).

***Ra* cDNA isolation:** PolyA$^+$ RNA isolated from Purple522 leaf tissue and analyzed on Northern blots was used for cDNA library construction with the ZAP cDNA Synthesis Kit (Stratagene). Screening of 130,000 plaques with a 1.5-kb *SalI* genomic fragment containing part of the *Ra* gene resulted in the identification of 64 positive signals. The clone designated *Ra* cDNA was chosen for further analysis because it contained the longest insert.

Isolation of genomic clones: A rice (*O. sativa*) genomic DNA library containing *Hin*dIII fragment inserts in the lambda 2001 vector (OKAGAKI and WESSLER 1988) was probed with an 800-bp *SstI* fragment of the maize *Lc* cDNA containing the HLH domain (LUDWIG *et al.* 1989). A phage containing a 12.5-kb *Hin*dIII fragment was isolated from which the 3′ half of the *Ra* gene was subcloned on a 1.5-kb *SalI* fragment (used in the *Ra* cDNA isolation).

The *Ra* gene was isolated from a genomic library that was constructed following partial *Sau*3A digestion of DNA from Purple522 (SAMBROOK *et al.* 1989), ligation into the Lambda FIX II Vector (Stratagene), packaging (Gigapack II Packaging Extract protocol, Stratagene) and screening 300,000 plaques with the cloned *Ra* cDNA. A phage containing an 11.5-kb insert was isolated, and *Ra* cDNA-homologous regions were identified. The *Rb* gene that was in a phage containing a 10-kb insert was also isolated from this library using the same methods except using the *Rb* specific probe derived mostly from *Rb* intron 9 sequence.

5′ end determination: Primer extension assays (SAMBROOK *et al.* 1989) were used in conjunction with 50 pmol of a *Ra* gene-specific oligonucleotide derived from exon 2 sequences (5′-GCTCGCAGACCAAGAACTCG-3′) and labeled by T4 DNA kinase with 80 pmol (6000 Ci/mmol) of r-ATP at 37° for 1 hr. Extension was done at 43° with the labeled primer, 10 μg of rice leaf polyA$^+$ RNA (from Purple522 or IR36) and AMV reverse transcriptase. Extension products were run in

an 8% PAGE-urea sequencing gel, along with the sequencing reaction products from a control DNA of bacteriophage M13mp18 (USB).

5′ RACE was performed with the 5′ RACE System for Rapid Amplification of cDNA Ends (GIBCO BRL). The two gene-specific primers used were as follows: external, 5′-TGGGCG-GCTGGTTGAAATGG-3′ from exon 4, and internal, 5′-GCT-CGCAGACCAAGAACTCG-3′ from exon 2.

PCR amplification: Fragments of the *O. sativa Ra* and *Rb,* and their homologues were amplified from genomic DNA of Purple522 and other wild species with the degenerate primers P1 (5′-GGGAACGGCAARAANCAYGTNATG-3′) and P2 (5′-AGGTGCRTCRAANACNCKNGTCAT-3′) (see Figure 3) and the conditions used previously to isolate *R* gene sequences from the genomes of other grasses (PURUGGANAN and WESSLER 1994). Intron 7 of *Rb* was amplified from the *Rb* genomic clone using primers P3 (5′-AGTGCATCTATTAAG-3′) on exon 7 and P4 (5′-GCATGATAAATGGCACGC-3′) on exon 8 of the *Ra* gene with the same conditions as described above. Amplified products were cloned using the TA Cloning Kit (Invitrogen). At least two clones from each PCR product were sequenced in this study.

DNA sequencing and alignments: DNA was sequenced using the dideoxy method (SANGER *et al.* 1977) and a sequencing kit (USB) or by the University of Georgia Molecular Genetics Facility. Alignments were made by PILEUP of the Genetics Computer Group (GCG) package (Wisconsin Package, version 8, 1994). The sequences of maize *Lc,* maize *Rs,* maize *Sn,* maize *B-Peru,* and Antirrhinum *delila* are from Genbank under the accession numbers of MZ6227, X15806, X60706, X57276 and M84913. The sequences of *Stow1,* the *Stowaway* element from the *O. sativa* heat shock protein 82A, and *Stow2,* the Tnr1 element from *O. glaberrima,* are from BUREAU and WESSLER (1994) and TENZEN *et al.* (1994). The rice sequences used in this study have been deposited into GenBank under the accession numbers of U39860–U39872.

Bombardment assays: An expression vector (pRICERA) containing the *Ra* cDNA was constructed by inserting a 2.4-kb *Eco*RI/*Xho*I fragment [containing the entire *Ra* open reading frame (ORF) and 333 bp of the 5′ untranslated region] between the CaMV 35S promoter and the nopaline synthase terminator in the pBI221 vector (Clontech). The first 140 bp of the 5′ leader containing the sole upstream AUG was not included in the construct. Bombardment of this construct, the maize control construct (pLcm1) (DAMIANI and WESSLER 1993) and the negative control construct (containing the 5′ and 3′ of the CaMV 35S only) into imbibed maize kernels was carried out as described previously (DAMIANI and WESSLER 1993) except that microprojectiles were prepared by precipitation of a total of 1 μg plasmid DNA (pRICERA or pLcm1). The kernels used were from the maize inbred line W22 (*r-g, A1, A2, Bz1, C1, C2, pl, B-b*). Pigmented cells were scored after incubation of bombarded kernels (12 with pLcm1, eight with pRiceRa, and three with the negative control construct) for 48 hr.

Restriction fragment length mapping: Segregation analysis was carried out using a backcross population of 113 plants derived from the cross BS125 (*O. sativa*)/WL02 (*O. longistaminata*)//BS125. To score for *Ra* segregation, a labeled *Ra*-specific fragment was used to probe two sets of DNA blots containing *Hin*dIII or *Bam*HI digests of the mapping population. Similarly, *Rb* segregation was scored using the labeled *Rb*-specific fragment to probe three sets of DNA blots containing *Hin*dIII, *Bam*HI or *Msp*I digests of the mapping population. The segregation patterns from different digestions were consistent with each other. The maps were constructed using the MAPMAKER program version 2.0 (DuPont and Whitehead Institute for Biomedical Research 1992), and the genetic distances were calculated with two-points estimates. The markers were considered to be linked if LOD > 3.0. The map positions of *B* and *Lc* in Figure 9 are derived from E. COE (1993).

Determining *Ra* and *Rb* gene phylogeny: The amino acid sequences of the genes were first aligned by Pileup of the GCG package. The gene phylogeny was then constructed using the PAUP program of parsimony analysis (Paup 3.1.1, Smithsonian Institution 1993) with the heuristic search algorithm using the Antirrhinum *delila* gene as an outgroup. All insertions and deletions were treated as informative in the phylogeny construction.

RESULTS

Isolation and characterization of a rice *R* homologue: A 12.5-kb *Hin*dIII fragment was isolated from a rice genomic library that had been probed with a fragment of the maize *R (Lc)* cDNA (LUDWIG *et al.* 1989). DNA sequence analysis revealed that part of this clone (1.5 kb) had significant similarity with the coding region of the 3′ half of the maize *Lc* cDNA, which includes the bHLH domain. This 1.5-kb *Sal*I rice fragment was, in turn, used to probe a RNA blot that contained polyA$^+$ RNA isolated from leaf tissue of a rice strain (Purple522) that has purple leaves and stigmas. Detection of a 2.5-kb transcript led to the construction of a cDNA library that was generated from Purple522 polyA$^+$ RNA. Screening of this library with the 1.5-kb genomic fragment allowed the identification of 64 cDNA-containing clones, of which the longest (2.4 kb) was sequenced and designated *Ra.* Comparison of the derived amino acid sequence of *Ra* with the maize *Lc* and Antirrhinum *delila* cDNAs revealed extensive similarity along their entire length (Figure 1). Furthermore, when *Ra* was used to probe an RNA blot containing polyA$^+$ RNA isolated from leaf tissues of Purple522 or an unpigmented strain (Green412), a 2.5-kb transcript was only detected among the purple rice sample (Figure 2A), suggesting the involvement of *Ra* expression in leaf pigmentation.

To isolate the *Ra* gene, a genomic library was generated from Purple522 DNA and probed with the *Ra* cDNA. One genomic clone was determined to encode the *Ra* mRNA because it contained sequences identical to the *Ra* cDNA. The *Ra* gene was further defined by using primer extension (Figure 2B) and 5′ RACE analysis to position the transcription start site. The 11 exons that comprise the rice *Ra* gene are shown diagrammatically in Figure 3 along with the positions of the start and stop of translation that defines the limits of the 1761-bp *Ra* ORF. The *Ra* ORF contains seven introns that interrupt the coding region at precisely the same positions as the maize *Lc* introns (Figure 3, solid lines connect the regions of homology). In contrast to the coding region similarity, the 5′ leaders of the *Ra* and *Lc* genes differ dramatically. The *Lc* 5′ leader (235 bp), encoded by the first exon and 7 bp of the second exon, has a 38-codon upstream ORF that is not found in the rice gene. Instead, the *Ra* 5′ leader is unusually long at

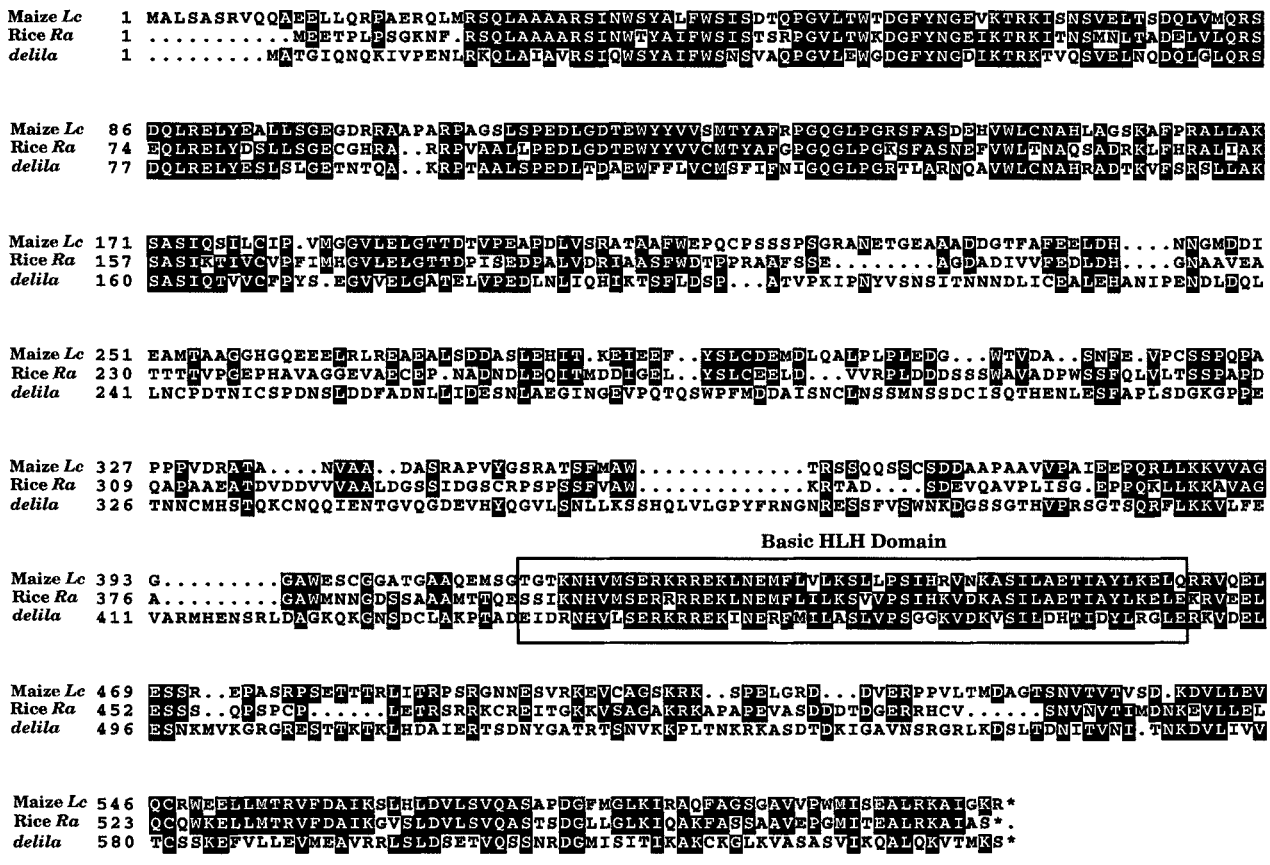

FIGURE 1.—Alignment of derived amino acid sequences from maize *Lc*, rice *Ra* and Antirrhinum *delila*. Identical amino acids are blocked in black. The sequence of *Ra* has been deposited in GenBank under the accession number of U39860.

470 bp and is encoded by the first three exons and the first 58 bp of the fourth exon. The 470-bp *Ra* leader contains only a single AUG codon (in exon 1) that initiates a three-codon ORF.

***Ra* can activate the maize anthocyanin pathway:** The identification of the *Ra* transcript among RNA isolated from purple, but not green, rice leaves suggested that the *Ra* protein functions in the activation of the anthocyanin pathway. To test this notion directly, a transcriptional fusion of the *Ra* cDNA and the 35S promoter was constructed and bombarded into maize aleurone cells (see MATERIALS AND METHODS). Observation of pigmented cells 48 hr after bombardment suggested that *Ra* protein is able to interact with the maize *C1* protein to activate the maize structural gene promoters. However, activation appeared extremely inefficient since the number of red spots was ~10% of the maize *Lc* control (Figure 4), and the pigmentation was weak (not shown). The inefficiency could reflect the fact that *Ra* differs from *Lc* at almost 50% of the amino acid positions. Given this level of divergence, it is likely that the interaction of *Ra* with the maize C1 protein and the maize genes encoding enzymes in the anthocyanin pathway will not be optimal.

Isolation of a second *R* gene from the rice genome: As in maize, Antirrhinum and petunia (DELLAPORTA *et al.* 1988; GOODRICH *et al.* 1992; QUATTROCCHIO 1994), some rice strains appear to have more than one *R* gene. DNA gel blots of *O. sativa* genomic DNA digested with different restriction enzymes and probed with *Ra* cDNA fragments suggested the existence of a second *R* gene (data not shown). To identify other *R* genes in rice, we took advantage of degenerate oligonucleotides that had been utilized in a previous study to amplify *R* genes from the genomes of several grasses (PURUGGANAN and WESSLER 1994). The positions of these oligos with respect to the *Ra* gene are shown in Figure 3, as P1 and P2. When used with genomic DNA from Purple522, two products of 720 and 550 bp were obtained. The sequence of the 720-bp product was identical to *Ra* and corresponded to parts of exons 9 and 10 and all of intron 9.

The sequence of the 550-bp product revealed that it was related to, but clearly different than, *Ra* with greatest similarity in the highly conserved bHLH and C-terminal domains (Figure 5A). There is very little similarity in the region corresponding to intron 9 or in the exon sequences between the two domains. To isolate the rest of this putative second *R* gene, the intron 9 region of the 550-bp fragment was used to probe the Purple522 genomic library. A genomic clone containing sequences identical to the 550-bp fragment was isolated and shown to hybridize with the *Ra* cDNA along its entire length (data not shown). This second *R* gene was designated *Rb*.

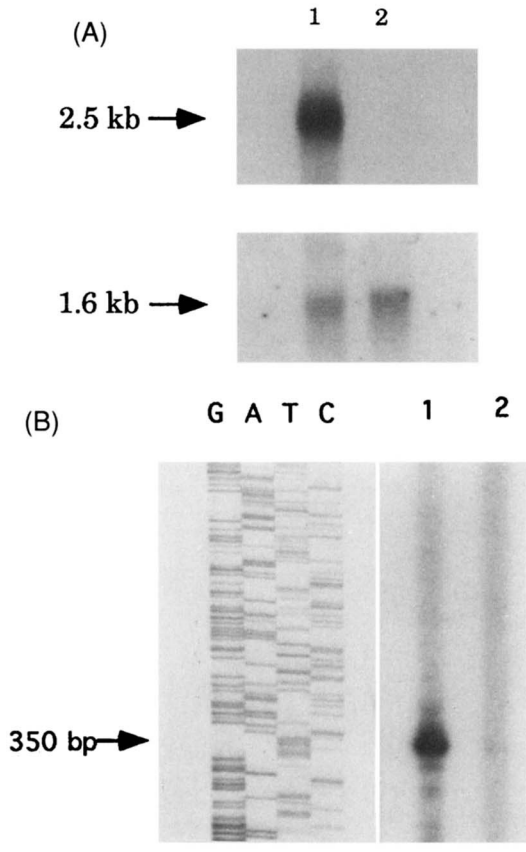

FIGURE 2.—Northern blot analysis of *Ra* transcript and primer extension analysis of polyA$^+$ RNA from purple and green leaf tissues. (A) Northern blot analysis of *Ra* transcripts in purple and green leaf tissues. PolyA$^+$ RNA from leaves of Purple522 (lane 1) and Green412 (lane 2) were probed with *Ra* cDNA (top). After removing the probe, the same blot was hybridized with the maize actin cDNA (bottom) under less stringent conditions (see MATERIALS AND METHODS). (B) Primer extension analysis of polyA$^+$ RNA from purple and green leaf tissues. PolyA$^+$ RNA from leaves of Purple522 (lane 1) or IR36 (lane 2) were subjected to primer extension analysis using a gene-specific primer derived from the *Ra* cDNA. The sequencing reaction products from a control single-stranded DNA (bacteriophage M13mp18) serve as size markers on the left.

***Stowaway* elements in *Ra* but not *Rb*:** A search of Genbank with the intron sequences of *Ra* as queries suggested the presence of members of the *Stowaway* family of inverted repeat elements in intron 7 and intron 9 (Figure 3 and Figure 5B) (BUREAU and WESSLER 1994; TENZEN *et al.* 1994). *Stow1* in intron 7 is 227 bp and *Stow2* in intron 9 is 171 bp in length. Both elements share the conserved terminal inverted repeats and are flanked by a direct repeat of TA. By comparing the regions corresponding to introns 7 and 9 in *Ra* and *Rb*, it was determined that *Rb* lacked both insertions and contained only one copy of the TA sequence. TA is the preferred insertion site of the *Stowaway* family; a direct repeat of TA flanks most *Stowaway* elements. These data are consistent with the view that both elements inserted into the *Ra* gene and were never present in *Rb*.

***R* copy number in rice genomes:** Intron 9 sequences from *Ra* and *Rb* were used as gene specific probes (designated the *Ra*- and *Rb*-specific probes) to determine the copy number of the *R* gene family in *O. sativa*. Although intron 9 of *Ra* contains *Stow2*, this element is substantially diverged and can no longer hybridize with other *Stowaway* elements if stringent conditions are employed. Genomic DNA from Purple522 and green (IR36) strains were digested with *Sph*I (Figure 3) and probed with either the 3′ half of the *Ra* cDNA (which contains the highly conserved bHLH domain), the *Rb*-specific probe, or the *Ra*-specific probe (Figure 6). Two bands were detected by the conserved (cDNA) probe, while the gene-specific probes hybridized to each band individually. These data suggest that the *R* gene family in these strains consists of at least two members, *Ra* and *Rb*.

Intron 9 of *Ra* is located on a polymorphic fragment. A 2.3-kb fragment is derived from the highly expressed *Ra* gene in Purple522, whereas an apparently inactive *Ra* gene from green rice resides on a 4.0-kb *Sph*I fragment. To address the question of whether a gene rearrangement of *Ra* could have been involved in its activation in pigmented rice or its inactivation in green rice, we assayed additional strains of *O. sativa* that lacked leaf pigmentation (Figure 7A). Although two of the unpigmented strains contained the 4.0-kb *Ra* fragment (lanes 2 and 4), the rest of the green strains contained a 2.3-kb fragment that comigrates with *Ra* frag.nent in the pigmented rice. Thus, the presence of the larger fragment does not correlate with the presence of leaf pigmentation. However, unlike the situation in maize where *R* copy number can vary, this preliminary analysis suggests that most or all strains of *O. sativa* have one copy of *Ra* and one of *Rb*. We have noticed however, that in some strains the *Rb*-specific probe also hybridized to a fragment slightly smaller than the *Rb* band, thus giving a doublet in these lanes (Figure 6, lane 4, and Figure 7A, lanes 2–4 and 6–10). In contrast, the conserved cDNA probe does not detect two bands in this region. These observations suggest that in some strains there appears to be a duplication in the genome of part of the sequence in *Rb* intron 9.

The chromosomal locations of *Ra* and *Rb*: The map positions of *Ra* and *Rb* were of interest for two reasons. First, to understand the evolution of this gene family it is important to know if these genes are linked, as is the case for the maize *R-r* complex, or unlinked, as is the case for the maize *R* and *B* genes. Second, several genetically defined loci have been shown to influence the distribution of rice pigmentation. Knowledge of the map positions of *Ra* and *Rb* might permit a correlation between their chromosomal location and the position of genetically defined loci. To this end, we utilized an interspecific backcross population and carried out segregation analysis of 113 plants derived from the cross BS125 (*O. sativa*)/WL02 (*O. longistaminata*)//BS125.

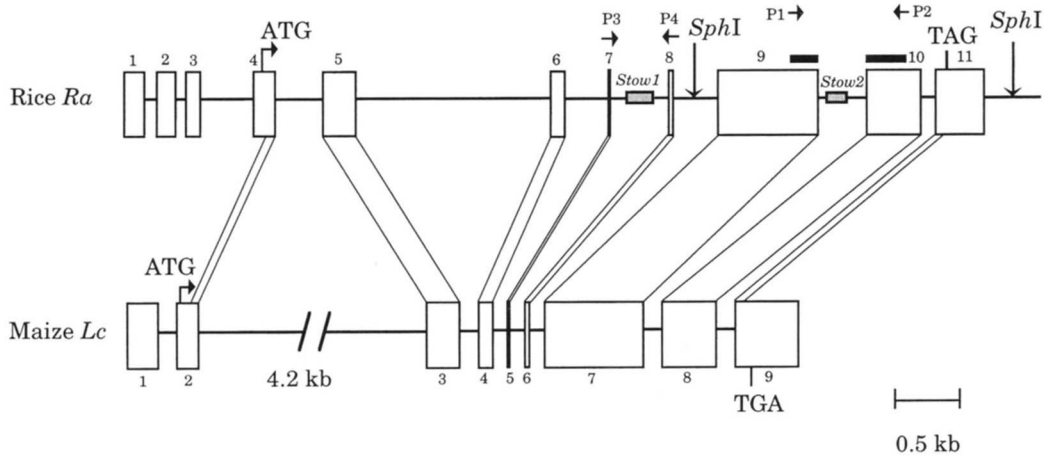

FIGURE 3.—Gene structure comparison of rice *Ra* and maize *Lc*. The exons are shown as □, and — between them represent introns. Homologous exon regions in the two genes are connected with thin solid lines. ■ at the end of exon 9 and the beginning of exon 10 indicates the 3′ *Ra* cDNA probe (see Figure 6 and 7). P1 and P2 are the degenerate primers used to amplify *R* sequences from domesticated and wild rice species. P3 and P4 indicate the positions of the primers used to amplify intron 7 from the *Rb* genomic clone. *Stow1* and *Stow2* are the *Stowaway* elements found in *Ra* introns (see details in text and Figure 5B).

DNA blots were probed with *Ra*- and *Rb*-specific probes, and the map was constructed using the Mapmaker program. The data indicates that *Ra* and *Rb* are unlinked and located on chromosomes *4* and *1*, respectively (Figure 8). The position of the *Ra* gene is of some significance since this region of rice chromosome *4* has been reported to show synteny with the regions on maize chromosomes *2* and *10* where the *B* and *R* genes, respectively, are located (Figure 8) (AHN and TANKSLEY 1993). Thus, there exists a direct ancestral relationship between the chromosomal location of rice *Ra* and the locations of maize *B* and *R* genes.

R genes in wild species: Given that *Ra* is located at the ancestral position, we were interested in knowing more about the organization of *Rb*. Two methodologies were used to address this issue. First, the presence of *Ra* and *Rb* genes in several wild species of rice was assessed by probing genomic Southern blots with conserved and gene-specific probes (Figure 7B). The genus Oryza is comprised of 22 species, with AA, BB, CC, EE, FF, BBCC, CCDD genomes and a few unidentified genome types (VAUGHAN 1989). There are two domesticated rice species, one of which is *O. sativa*. The three AA genome-containing species analyzed in this study are the domesticated *O. sativa* and two wild species, the closely related *O. rufipogon* and the more distantly related *O. longistaminata*. Both wild species appear to contain two copies of *R* (Figure 7B, lanes 11 and 12). In contrast, *O. australiensis* (CC genome, lane 13), *O. eichingeri* (EE genome, lane 14) and *O. officinalis* (CC genome, lane 15) all contain a single band that hybridizes with the conserved probe. Taken together, these data suggest that only AA genomes have two *R* genes.

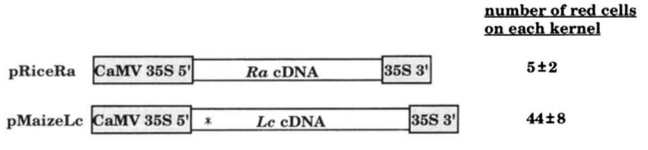

FIGURE 4.—Schematic representation of the *Ra* and *Lc* constructs and the number of red cells observed on kernels after bombardment. The *Ra* cDNA lacks exon 1 (which contains the sole upstream AUG) and part of exon 2. The *Lc* cDNA is full length, except that the first AUG (denoted with *) in the 5′ leader was mutagenized to relieve translational control (DAMIANI and WESSLER 1993).

Due to the divergence between the domesticated and the wild species, the gene-specific probes (derived mainly from intron 9 of the *Ra* and *Rb* genes) could not be used to determine whether the *R* genes in these species were more closely related to *Ra* or *Rb*. Instead, the degenerate PCR primers P1 and P2 were used to amplify *R* sequences from these wild species. Amplified products were sequenced and aligned with *Ra* and *Rb* from *O. sativa*. As expected from the Southern blot results, two fragments were amplified from each of the two wild species with AA genomes (Orra and Orrb from *O. rufipogon*, Olra and Olrb from *O. longistaminata*). The sequence alignments (Figure 9) reveal that one product from each species is more closely related to *Ra* (Figure 9A) and one to *Rb* (Figure 9B). Only a single PCR product was obtained from each non-AA genome species, and these appear to be more closely related to *Ra* (Figure 9A) (Oar from *O. australiensis*, Oer from *O. eichingeri*, and Oor from *O. officinalis*). The identification of an *Ra*-like gene in all rice strains examined is consistent with the notion that *Ra* is at the ancestral locus.

Phylogeny of *R* genes in rice and maize: To further understand the origin of *Rb* in *O. sativa* and the relationship of *Ra* and *Rb* to the single *R* gene in non-AA genomes, we constructed a gene phylogeny of the avail-

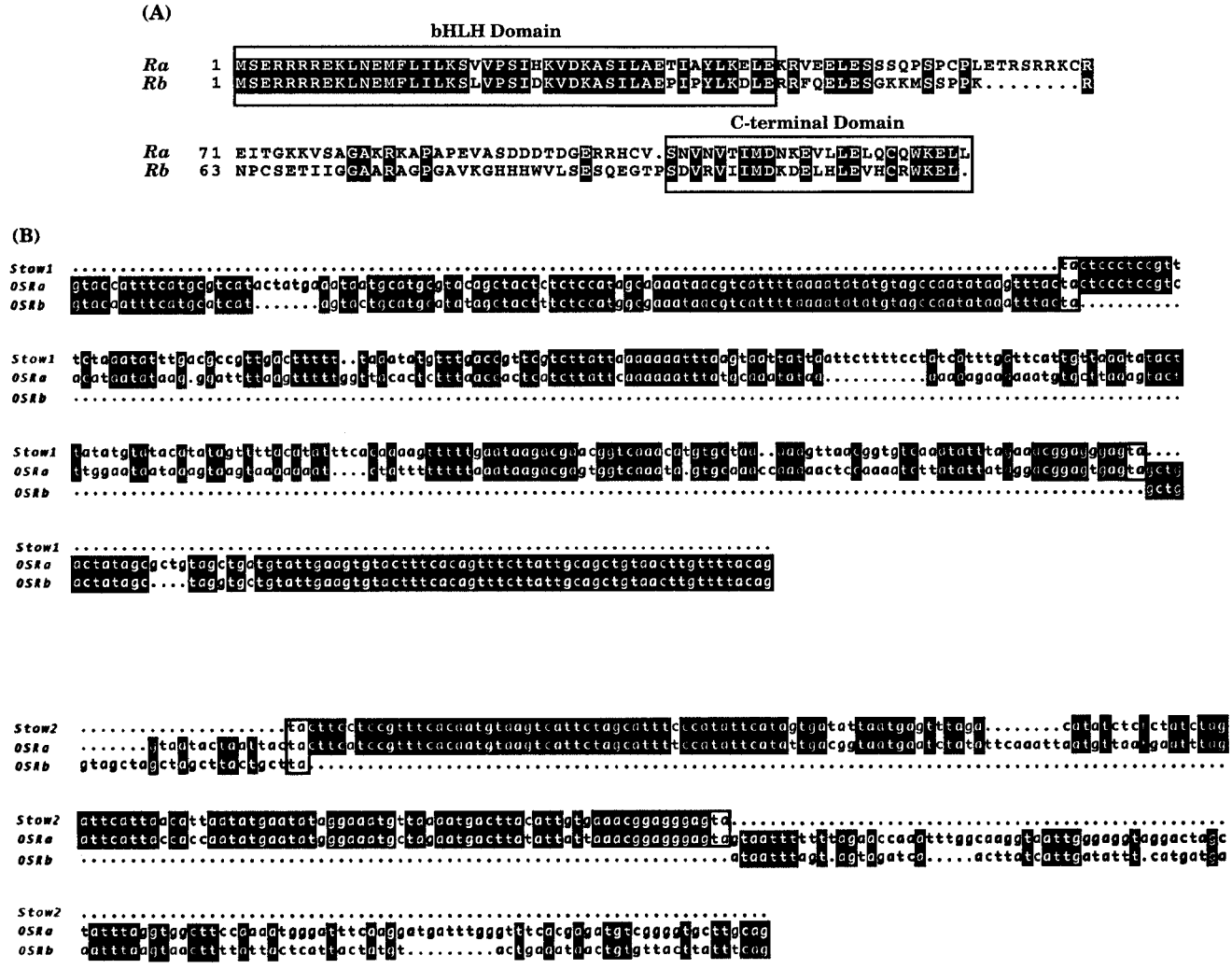

FIGURE 5.—Comparison of the *Ra* and *Rb* genes. (A) Comparison of predicted amino acid sequeces from regions of *Ra* and *Rb*. Identical amino acids are blocked in black. The open boxes delimit the basic helix-loop-helix and the C-terminal domains (defined by PURUGGANAN and WESSLER 1994) that are conserved among the grasses. The *Rb* sequence has been deposited in GenBank under the accession number of U39866. (B) Comparison of introns 7 and 9 of *Ra* and *Rb* with rice *Stowaway* elements. The top figure aligns the *Stow1* element (located in intron 2 of the *O. sativa* heat shock protein 82A) with intron 7 of *Ra* and *Rb* (accession numbers, U39871 and U39872). The bottom figure aligns the *Stow2* element from the genome of *O. glaberrima* (defined as the *Tnr1* element by TENZEN *et al.* 1994) with intron 9 of *Ra* and *Rb* (accession numbers, U39869 and U39870). The black blocks indicate identical nucleotides, and the TA sequences in the open boxes is the presumed target site for *Stowaway* element insertion.

able *R* genes from rice and maize. Using parsimony analysis, a phylogenetic tree was constructed from the derived sequences of 143 amino acids found in the 3′ half of *R*. The Antirrhinum *delila* gene served as the outgroup. The inferred tree reveals that rice *R* genes form two major groups: we call these the *Ra* clade and the *Rb* clade (Figure 10). The three AA species (*O. sativa, O. rufipogon, O. longistaminata*) each contain two genes with one member in the *Ra* clade and one in the *Rb* clade. The three non-AA species (*O. australiensis, O. eichingeri* and *O. officinalis*) each contain one *R* gene that is a member of the *Ra* clade. Although the rice genes fall into two clusters, it is clear from the data that all rice genes are more closely related to each other than they are to the maize genes.

DISCUSSION

We have identified the *Ra* gene as a rice homologue of the maize *R/B* genes based upon the following criteria: (1) the putative amino acid sequence of *Ra* shows extensive homology with maize *R* and *B* genes, (2) the positions of introns are precisely conserved within the coding regions of *Ra* and the maize *R (Lc)* gene, (3) the location of *Ra* on chromosome *4* is in synteny with the regions of maize chromosomes *2* and *10* containing the *B* and *R* loci, respectively, and (4) *Ra* can complement maize *R* gene function in a bombardment assay.

Although the coding regions of *Ra* and *Lc* are homologous along their entire length, the 5′ untranslated regions are dramatically different. The maize *Lc* gene

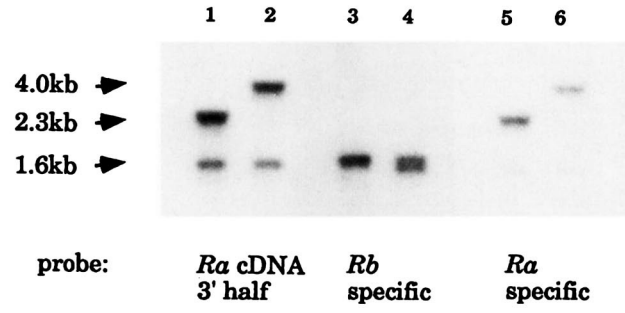

FIGURE 6.—Southern blot analysis of Purple522 and IR36. Genomic DNA from Purple522 (lanes 1, 3, and 5) or IR36 (lanes 2, 4, and 6) were digested with *Sph*I (see Figure 3), fractionated on an agarose gel and blotted. The blot was first probed with part of the *Ra* cDNA sequence from the bHLH to C-terminal region (see Figure 3) (lanes 1 and 2). After removing the first probe, the same blot was hybridized with a *Rb*-specific probe (intron 9) (lanes 3 and 4), and finally with *Ra*-specific probe (intron 9) (lanes 5 and 6) after the second probe was removed. All the probes have no internal *Sph*I restriction site.

has a long (235 bp) 5′ leader that contains an uORF of 38 codons. Translation of the *Lc* uORF has been shown to repress translation of the downstream ORF both *in vitro* and *in vivo* (DAMIANI and WESSLER 1993). Translational repression of *R* gene expression is thought to be necessary because overexpression of *R* protein may interfere with normal development, possibly through aberrant HLH interactions (DAMIANI and WESSLER 1993). In contrast, the 5′ leader of *Ra* is unusually long (470 nt) and is encoded by multiple exons. Despite its length and complexity, it contains only a single AUG codon that probably could not initiate translation because it is in a very poor sequence context (GALLIE 1993). However, long leaders are frequently associated with translational repression, whether it be due to the formation of inhibitory secondary structures or the binding of regulatory proteins (KOZAK 1989). If such a mechanism is found to regulate *Ra*, it would suggest that the *R* genes of grasses may be a unique resource for studies on translational repression of plant gene expression.

One reason for initiating a study of *R* genes in rice was to compare the molecular basis of allelic diversity in two related but distinct plants. In maize, the pattern of anthocyanin deposition primarily reflects the expression pattern of *R* and *B* genes (reviewed in LUDWIG and WESSLER 1990). This was demonstrated in two ways. First, by showing that an *R* cDNA fused to the constitutive 35S promoter could induce pigmentation in most cell types (LUDWIG *et al.* 1990). Second, by swapping the promoters and 5′ leaders of two *B* alleles (*B-I* and *B-Peru*) and demonstrating that the phenotype of the chimeric gene mirrored the allele that donated the promoter and leader (RADICELLA *et al.* 1992). The fact that *Ra* can activate the anthocyanin pathway in maize aleurone cells and that *Ra* mRNA is found in pigmented rice leaves but not in unpigmented leaves is consistent with the view that *Ra* regulates pigmentation pattern in rice just as *R* and *B* genes do in maize.

The pigmentation pattern in maize results from the additive effects of the *B* allele on chromosome *2* and the *R* allele(s) on chromosome *10* (LUDWIG and WESSLER 1990). Therefore, to fully understand what regulates rice pigmentation patterns, we must also know whether there are multiple *R* genes in the genome. Two *R* genes have been found in rice species with AA genomes. Whether *Rb* encodes a functional gene that contributes to the overall pattern of pigmentation is unknown at this time. However, the fact that *Rb* and *Ra* are highly conserved in the bHLH and C-terminal domains and

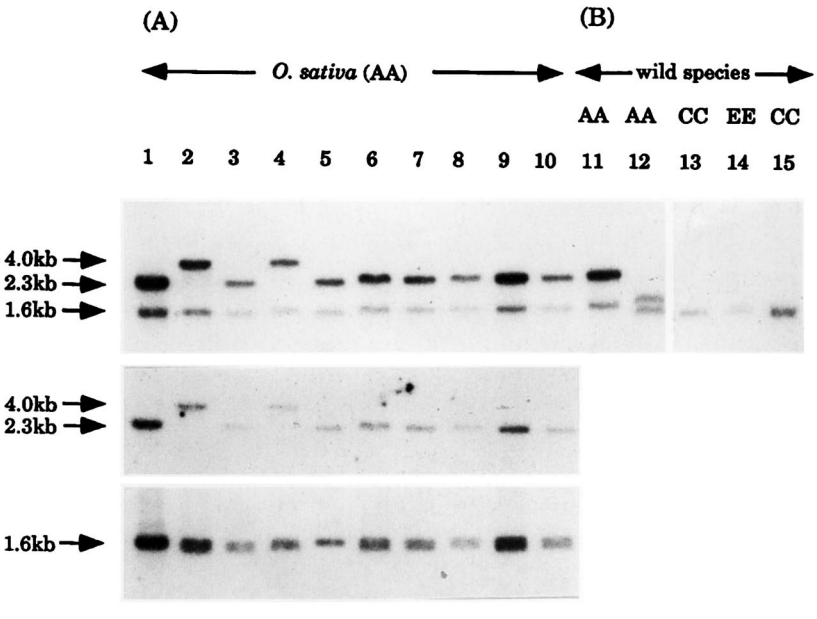

FIGURE 7.—Southern blot analysis of domesticated and wild rice. Genomic DNA from different *O. sativa* strains and some wild rice species were digested with *Sph*I and probed with part of the *Ra* cDNA containing the bHLH to C-terminal region (top). After removing this probe, the *O. sativa* DNAs were hybridized with the *Ra*-specific probe (middle) and then with the *Rb*-specific probe (bottom) after the second probe was removed. All the probes used are the same as those in Figure 6. Lanes 1–10: *O. sativa* strains Purple522, IR36, SH88-2, IR29, BS125, Calmochi 201, Star Bonnet, Susono Mochi, Stg 772199, and Malagkit Songsong. Lanes 11–15: wild rice species *O. rufipogon* (AA), *O. longistaminata* (AA), *O. officinalis* (CC), *O. australiensis* (EE) and *O. eichingeri* (CC).

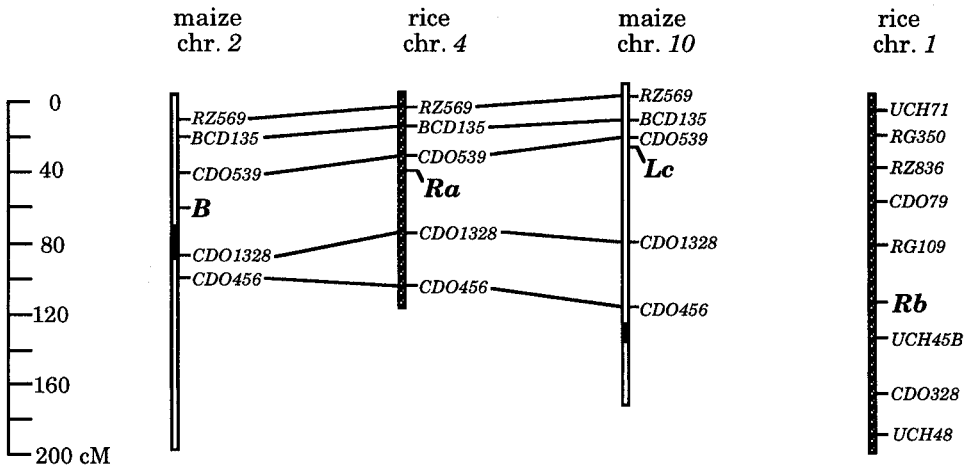

FIGURE 8.—Chromosomal locations of *Ra* and *Rb* and the approximate region of synteny between rice chromosome *4* and maize chromosomes *2* and *10* (AHN and TANKSLEY 1993). The *Ra* and *Rb* genes were mapped using MAPMAKER 2.0. The genetic distances were calculated with two-point estimates. Maize chromosome *10* is shown in reversed order to better demonstrate its homology with other chromosomes. The approximate positions of centromeres are indicated by black bars on maize chromosomes *2* and *10*. The scale on the left represents the genetic distance in cM.

that the deduced protein sequence from the 3′ half of the *Rb* coding region is in the correct ORF argues that *Rb* may be functional.

To more rigorously determine the contribution of *Ra* and *Rb* to the establishment of pigmentation patterns in rice, it will be necessary to analyze their expression in strains that display anthocyanin in a variety of tissues. Fortunately, anthocyanin distribution in rice has been the subject of genetic analysis for decades (KINOSHITA 1984). Three genes, *C* (chromogen for anthocyanin, chromosome *6*), *A* (anthocyanin activator, chromosome *1*) and *P* (purple apiculus, chromosome *4*) are thought to be fundamental for pigmentation of the plant and the distinct coloration of the apiculus (NAGAO 1951; NAGAO and TAKAHASHI 1956; TAKAHASHI 1957). When the *Pl* gene (purple leaf, chromosome *4*) was added to these three genes, the entire leaf turned purple (KINOSHITA 1984). The genetic behavior of *Pl* suggests that it might be synonymous with *Ra*. However, although both *Ra* and *Pl* map to chromosome *4*, they are not tightly linked. This descrepancy may indicate that there are two loci on chromosome *4* that influence pigment distribution. Alternatively, *Pl* and *Ra* may be synonymous and the mapping differences may be due to the use of different mapping populations. *Rb*, which has not as yet been shown to be functional, maps near (a few map units away from) the *Pn* locus on chromosome *1* (KINOSHITA 1993). Addition of *Pn* to the *P, C* and *A* genes results in coloration of the node, auricle and ligule (KINOSHITA 1984). Careful examination of the green *O. sativa* plants used in this study failed to detect anthocyanin in these tissues. Thus, if *Rb* and *Pn* are the same locus, the plants are green either because the *Pn* allele is null or because there is a mutation in another gene of the anthocyanin pathway. With the isolation of gene-specific probes for *Ra* and *Rb* and the availability of isogenic strains with and without alleles of *Pn* or *Pl*, it should now be possible to correlate the presence or absence of these loci with the expression of *Ra* and *Rb*.

The second rationale for initiating a study of rice *R* genes was to determine whether there was more than one rice *R* gene, and if so, to contrast the evolution of gene families in rice and maize. The data presented in this study have allowed us to add more details to the model of *R* gene evolution in rice and maize and to

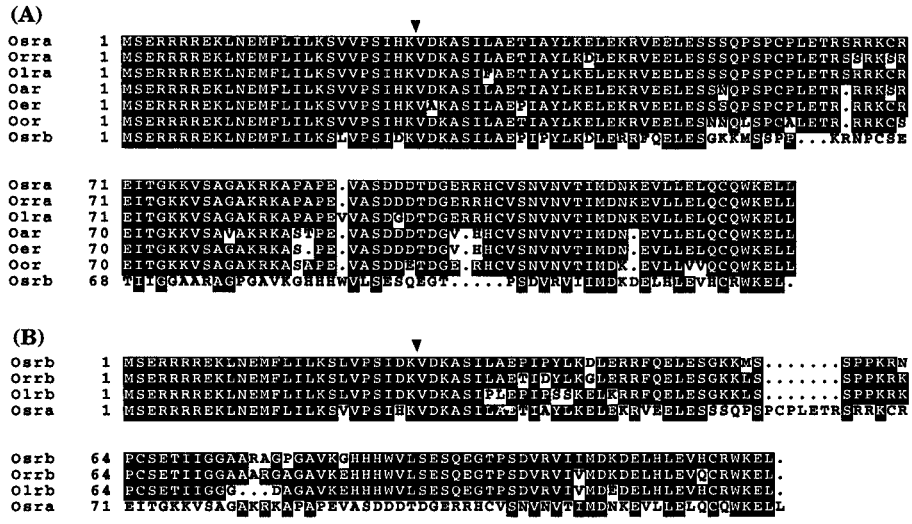

FIGURE 9.—Alignment of amino acid sequences from exon 9 and exon 10 of *O. sativa Ra, Rb* and their homologues from wild rice species. Identical sequences are blocked in black. Osra, Orra, Olra, Oar, Oer and Oor (accession numbers U39860–U39865) are the *Ra* gene and its homologues from *O. sativa, O. rufipogon, O. longistaminata, O. australiensis, O. eichingeri,* and *O. officinalis,* respectively. Osrb, Orrb and Olrb (U39866–U39868) are the *Rb* gene and its homologues from *O. sativa, O. rufipogon* and *O. longistaminata,* respectively. The arrow indicates the junction of exon 9 and 10.

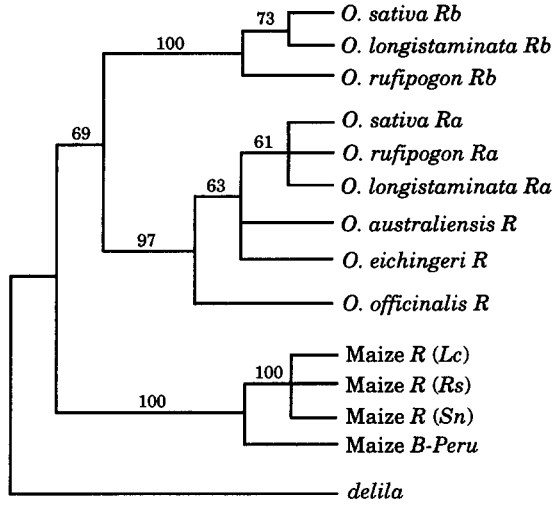

FIGURE 10.—Phylogenetic tree of *R* genes in rice and maize. The gene phylogeny was constructed using the PAUP program of maximum parsimony analysis. A total of 143 amino acids from the bHLH through C-terminal regions were used in the analysis. The numbers by the nodes give bootstrap values from 100 replicates. The tree length is 232 steps, with a consistency index of 0.921, a homoplasy index of 0.079, and a retention index of 0.939.

demonstrate an independent origin for *R* gene families in these grasses. The maize *R/B* gene family arose from a genome duplication event followed by further duplications and rearrangements at the *R* locus (ANDERSON 1945; HELENTJARIS *et al.* 1988; ROBBINS *et al.* 1989; 1991; WHITKUS *et al.* 1992; WALKER *et al.* 1995). Rice, as a true diploid, would not be expected to have a small gene family that resulted, in part, from an ancient polyploidization event. Consistent with this notion is our identification of an active *R* homologue (*Ra*) in a region of the *O. sativa* genome shown previously to be in synteny with duplicate regions of the maize genome harboring the *B* and *R* loci (AHN and TANKSLEY 1993), suggesting that there was a single *R* gene in the progenitor of rice and maize. Our phylogenetic analysis also suggests the independent origins of *R* gene families in maize and rice after the separation of their ancestors. The sequence divergence between *Ra* and *Rb* suggests that the duplication of the ancestral *R* into *Ra* and *Rb* may have taken place very early after the division of maize and rice progenitors. Given the fact that both *Ra* and *Rb* homologues are found in the AA genome, whereas only an *Ra* homologue is found in other types of genomes, *Rb* may have been lost in the lineage leading to some non-AA genome species. Alternatively, *Rb* may have been introgressed into chromosome *1* of the ancestor of some AA genome-containing species from an as yet undetermined wild species. By studying more wild rice species with all types of genomes, we will be able to know more about the distribution of *R* genes in the genus of Oryza. An interesting question is whether selection for a particular pigmentation pattern was the basis for the propagation of this hypothesized duplication or introgression event. Such questions must await a more rigorous analysis of *Rb* structure and function to ascertain its influence, if any, on the pigmentation pattern of rice species with the AA genome.

We are grateful to GARY KOCHERT, MARC ALAN COHN, LENORE LAMPERT and QIANG ZHANG for providing the rice strains and mapping materials, and for helping with the mapping program. We thank RON DAMIANI for performing the bombardment experiments. We also thank SHAWN WHITE and LANE ARTHUR for critical reading of the manuscript. This research was supported by grants from the Department of Energy and the Rockefeller Foundation to S.R.W.

LITERATURE CITED

AHN, S., and S. D. TANKSLEY, 1993 Comparative linkage maps of the rice and maize genomes. Proc. Natl. Acad. Sci. USA **90:** 7980–7984.

ANDERSON, E., 1945 What is *Zea mays*? A report of progress. Chron. Bot. **9:** 88–92.

ARUMUGANATHAN, K., and E. D. EARLE, 1991 Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep. **9(3):** 208–218.

BILANG, R., S. ZHANG, N. LEDUC, V. A. IGLESIAS, A. GISEL *et al.*, 1993 Transient gene expression in vegetative shoot apical meristems of wheat after ballistic microtargeting. Plant J. **4(4):** 735–744.

BUREAU, T. E., and S. R. WESSLER, 1994 *Stowaway:* a family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. Plant Cell **6:** 907–916.

CASAS, A. M., A. K. KONONOWICZ, U. B. ZEHR, D. T. TOMES, J. D. AXTELL *et al.*, 1993 Transgenic sorghum plants via microprojectile bombardment. Proc. Natl. Acad. Sci. USA **90:** 11212–11216.

CHANDLER, V. L., J. P. RADICELLA, T. P. ROBBINS, J. CHEN and D. TURKS, 1989 Two regulatory genes of the maize anthocyanin pathway are homologous: isolation of *B* utilizing *R* genomic sequences. Plant Cell **1:** 1175–1183.

COE, E. H., 1993 Gene list and working maps. Maize Genet. Co. Newsl. **67:** 133-166.

CONE, K. C., S. M. COCCIOLONE, F. A. BURR and B. BURR, 1993 Maize anthocyanin regulatory gene *pl* is a duplicate of *c1* that functions in the plant. Plant Cell **5:** 1795–1805.

DAMIANI JR., R. D., and S. R. WESSLER, 1993 An upstream open reading frame represses expression of *Lc*, a member of the *R/B* family of maize transcriptional activators. Proc. Natl. Acad. Sci. USA **90:** 8244–8248.

DELLAPORTA, S. L., J. WOOD and J. B. HICKS, 1983 A plant DNA minipreparation: version 2.0. Plant Mol. Biol. Rep. **1:** 19–22.

DELLAPORTA, S. L., I. GREENBLATT, J. L. KERMICLE, J. B. HICKS and S. R. WESSLER, 1988 Molecular cloning of the maize *R-nj* allele by transposon tagging with *Ac*, pp. 263–282 in *Chromosome Structure and Function: Impact of New Concepts, 18th Stadler Genetics Symposium,* edited by J. P. GUSTAFSON and R. APPLES. Plenum, New York.

DOONER, H. K., and J. L. KERMICLE, 1971 Structure of the *Rr* tandem duplication in maize. Genetics **67:** 427–436.

DOONER, H. K., and J. L. KERMICLE, 1976 Displaced and tandem duplication in the long arm of chromosome *10* in maize. Genetics **82:** 309–322.

FLAVELL, R. B., 1984 DNA transposition-a major contributor to plant chromosome structure BioEssays **1:** 21–22.

GALLIE, D. R., 1993 Posttranscriptional regulation of gene expression in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. **44:** 77–105.

GALWAY, M. E., J. D. MASUCCI, A. M. LLOYD, V. WALBOT, R. W. DAVIS *et al.*, 1994 The *TTG* gene is required to specify epidermal cell fate and cell patterning in the *Arabidopsis* root. Dev. Biol. **166:** 740–754.

GOFF, S. A., T. M. KLEIN, B. A. ROTH, M. E. FROMM, K. C. CONE *et al.*, 1990 Transactivation of anthocyanin biosynthetic genes following transfer of *B* regulatory genes into maize tissues. EMBO J. **9:** 2517–2522.

GOFF, S. A., K. C. CONE and V. L. CHANDLER, 1992 Functional analysis of the transcriptional activator encoded by the maize *B* gene:

evidence for a direct functional interaction between two classes of regulatory proteins. Genes Dev. **6:** 864–875.

GOODRICH, J., R. CARPENTER and E. S. COEN, 1992 A common gene regulates pigmentation pattern in diverse plant species. Cell **68:** 955–964.

HELENTJARIS, T., D. WEBER and S. WRIGHT, 1988 Identification of the genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. Genetics **118:** 353–363.

KINOSHITA, T., 1984 Gene analysis and linkage map, pp. 187–274 in *Biology of Rice,* edited by S. TSUNODA and N. TAKAHASHI. Jpn. Sci. Soc. Press, Elsevier, Tokyo.

KINOSHITA, T., 1993 Report of the committee on gene symbolization. Rice Genet. Newsl. **10:** 18–32.

KLEIN, T. M., B. A. ROTH and M. E. FROMM, 1989 Regulation of anthocyanin biosynthetic genes introduced into intact maize tissues by microprojectiles. Proc. Natl. Acad. Sci. USA **86:** 6681–6685.

KOZAK, M.,1989 The scanning model for translation: an update. J. Cell Biol. **108:** 229–241.

LLOYD, A. M., V. WALBOT and R. W. DAVIS, 1992 *Arabidopsis* and *Nicotiana* anthocyanin production activated by maize regulators *R* and *C1*. Science **258:** 1773–1775.

LUDWIG, S. R., and S. R. WESSLER, 1990 Maize *R* gene family: tissue specific helix-loop-helix proteins. Cell **62:** 849–851.

LUDWIG, S. R., L. F. HABERA, S. L. DELLAPORTA and S. R. WESSLER, 1989 *Lc,* a member of the maize *R* gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the *myc*-homology region. Proc. Natl. Acad. Sci. USA **86:** 7092–7096.

LUDWIG, S. R., B. BOWEN, L. BEACH and S. R. WESSLER, 1990 A regulatory gene as a novel visible marker for maize transformation. Science **247:** 449–450.

NAGAO, S., 1951 Genetic analysis and linkage relationship of characters in rice. Adv. Genet. **4:** 181–212.

NAGAO, S., and M. TAKAHASHI, 1956 The third gene in apiculus coloration. Genetic studies on rice plant, XIX. Jpn. J. Bot. **15:** 141–151.

OKA, H. I., 1988 *Origin of Cultivated Rice.* Jpn Sci. Soc. Press, Elsevier, Amsterdam.

OKAGAKI, R. J., and S. R. WESSLER, 1988 Comparison of non-mutant and mutant *waxy* genes in rice and maize. Genetics **120:** 1137–1143.

PAZ-ARES, J., D. GHOSAL, U. WIENAND, P. A. PETERSON and H. SAEDLER, 1987 The regulatory *C1* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators. EMBO J. **6:** 3553–3558.

PURUGGANAN, M. D., and S. R. WESSLER, 1994 Molecular evolution of the plant *R* regulatory gene family. Genetics **138:** 849–854.

QUATTROCCHIO, F. M., 1994 Regulatory genes controlling flower pigmentation in *Petunia hybrida.* Ph.D. thesis, Vrije University, Amsterdam.

QUATTROCCHIO, F., J. F. WING, H. T. C. LEPPEN, J. N. M. MOL and R. E. KOES, 1993 Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. Plant Cell **5:** 1497–1512.

RADICELLA, J. P., D. BROWN, L. A. TOLAR and V. L. CHANDLER, 1992 Allelic diversity of the maize *B* regulatory gene: different leader and promoter sequences of two *B* alleles determine distinct tissue specificities of anthocyanin production. Genes Dev. **6:** 2152–2164.

ROBBINS, T. P., J. CHEN, M. A. NORELL and S. L. DELLAPORTA, 1989 Molecular and genetic analysis of the *R* and *B* loci in maize, pp. 105–114 in *The Genetics of Flavanoids—Proceedings of a Post Congress Meeting of the XVI International Congress of Genetics,* edited by D. E. STYLES, G. A. GAVAZZI and M. L. RACCHI. Edizioni Unicopli, Milan.

ROBBINS, T. P., E. L. WALKER, J. L. KERMICLE, M. ALLEMAN and S. L. DELLAPORTA, 1991 Meiotic instability of the *R-r* complex arising from displaced intragenic exchange and intrachromosomal rearrangement. Genetics **129:** 271–283.

ROTH, B. A., S. A. GOFF, T. M. KLEIN and M. E. FROMM, 1991 *C1*- and *R*-dependent expression of the maize *Bz1* gene requires sequences with homology to mammalian *myb* and *myc* binding sites. Plant Cell **3:** 317–325.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Mannual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

SANGER, E., S. NICKLEN and A. R. COULSON, 1977 DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74(12):** 5463–5467.

SHIEH, M. W., S. R. WESSLER and N. V. RAIKHEL, 1993 Nuclear targeting of the maize *R* protein requires two nuclear localization sequences. Plant Physiol. **101:** 353–361.

STADLER, L. J., and S. FOGEL, 1943 Gene variability in maize I. Some alleles of *R* (*R-r* series). Genetics **28:** 90–91.

STADLER, L. J., and S. FOGEL, 1945 Gene variability in maize II. The action of certain *R* alleles. Genetics **30:** 23–24.

STADLER, L. J., and M. G. NUFFER, 1953 Problems of gene structure. II. Separation of *Rr* elements (*S*) and (*P*) by unequal crossing over. Science **117:** 471–472.

STYLES, E. D., O. CESKA and K. T. SEAH, 1973 Developmental differences in action of *R* and *B* alleles in maize. Can. J. Genet. Cytol. **15:** 59–72.

TAKAHASHI, M., 1957 Analysis of apiculus color genes essential to anthocyanin coloration in rice. J. Fac. Agr. Hokkaido Univ. **50(3):** 266–362.

TENZEN, T., Y. MATSUDA, H. OHTSUBO and E. OHTSUBO, 1994 Transposition of Tnr*1* in rice genomes to 5′-PuTAPy-3′ sites, duplicating the TA sequence. Mol. Gen. Genet. **245:** 441–448.

VAUGHAN, D. A., 1989 The genus Oryza L.: current status of taxonomy. Int. Rice Res. Inst. Res. Pap. Ser. No. 138: 2–21.

WALKER, E. L., T. P. ROBBINS, T. E. BUREAU, J. KERMICLE and S. L. DELLAPORTA, 1995 Transposon-mediated chromosomal rearrangements and gene duplications in the formation of the maize *R-r* complex. EMBO J. **14(10):** 2350–2363.

WHITKUS, R., J. DOEBLEY and M. LEE, 1992 Comparative genome mapping of sorghum and maize. Genetics **132:** 1119–1130.

Communicating editor: V. SUNDARESAN