Genet. Res., Camb. (2006), **88**, *pp.* 165–175. © 2006 Cambridge University Press
doi:10.1017/S0016672307008518 Printed in the United Kingdom

Revealing constitutively expressed resistance genes in *Agrostis* species using PCR-based motif-directed RNA fingerprinting

HIKMET BUDAK*, SENEM SU AND NESLIHAN ERGEN
Biological Science and Bioengineering Program, Faculty of Engineering and Natural Sciences, Sabanci University, Orhanlı, 34956, Tuzla/Istanbul, Turkey

(*Received 2 November 2006 and in revised form 11 January 2007*)

Summary

Agrostis species are mainly used in athletic fields and golf courses. Their integrity is maintained by fungicides, which makes the development of disease-resistance varieties a high priority. However, there is a lack of knowledge about resistance (*R*) genes and their use for genetic improvement in *Agrostis* species. The objective of this study was to identify and clone constitutively expressed cDNAs encoding *R* gene-like (RGL) sequences from three *Agrostis* species (colonial bentgrass (*A. capillaris* L.), creeping bentgrass (*A. stolonifera* L.) and velvet bentgrass (*A. canina* L.)) by PCR-based motif-directed RNA fingerprinting towards relatively conserved nucleotide binding site (NBS) domains. Sixty-one constitutively expressed cDNA sequences were identified and characterized. Sequence analysis of ESTs and probable translation products revealed that RGLs are highly conserved among these three *Agrostis* species. Fifteen of them were shown to share conserved motifs found in other plant disease resistance genes such as *MLA13*, *Xa1*, *YR6*, *YR23* and *RPP5*. The molecular evolutionary forces, analysed using the Ka/Ks ratio, reflected purifying selection both on NBS and leucine-rich repeat (LRR) intervening regions of discovered RGL sequences in these species. This study presents, for the first time, isolation and characterization of constitutively expressed RGL sequences from *Agrostis* species revealing the presence of TNL (TIR-NBS-LRR) type *R* genes in monocot plants. The characterized RGLs will further enhance knowledge on the molecular evolution of the *R* gene family in grasses.

1. Introduction

Bentgrass (*Agrostis*), a genus of the Poaceae family, consists of more than 200 species. However, there are only three species that are commonly used commercially, especially in athletic fields, golf courses, home lawns, parks and roadsides to enhance natural beauty (Hitchcock, 1971; Scheef *et al.*, 2003; Bonos *et al.*, 2006). Creeping bentgrass (*A. stolonifera* L.) is the most commonly used species in maintaining golf courses, followed by colonial bentgrass (*A. capillaris* L.) and velvet bentgrass (*A. canina* L.) (Belanger *et al.*, 2003; Dai *et al.*, 2003). Each of these species has particular traits such as vulnerability to certain diseases and environmental stress conditions (Vincelli & Doney, 1997; Belanger *et al.*, 2003).

* Corresponding author. Telephone: +90 216 4839575. Fax: +90 216 4839550. e-mail: budak@sabanciuniv.edu

The most serious problems of bentgrasses are their susceptibility to necrotrophic pathogens, especially dollar spot (caused by *Sclerotinia homoeocarpa*), red thread (*Laestisaria fuciformis*) and brown patch (*Rhizoctonia solani*), and to biotrophic pathogens such as powdery mildew (Deacon, 1973; Jørgensen, 1994; Vargas, 1994; Couch, 1995; Bliffeld *et al.*, 1999; Zhou *et al.*, 2001; Bonos *et al.*, 2006). It is known that necrotrophic plant pathogens kill the host tissue by means of toxic compounds and feed on the remains, whereas biotrophic pathogens produce effector molecules which are introduced into the plant cytoplasm by as yet unknown mechanisms (Glazebrook, 2005; Jones & Dangl, 2006). The current method for disease control is the heavy use of fungicides, which creates environmental and economic concerns (Dai *et al.*, 2003; Guo *et al.*, 2003). Improvement of the disease resistance is, therefore, a high priority for bentgrass

https://doi.org/10.1017/S0016672307008518 Published online by Cambridge University Press

breeding and genetics programmes (Boulter *et al.*, 2002; Belanger *et al.*, 2004). One of the most convenient ways of controlling plant diseases is to develop resistant genotypes.

Naturally occurring plant disease resistance often results from the presence of a specific resistance (*R*) gene in the plant for the effector molecule coded by the corresponding avirulence (*avr*) gene of the pathogen (Flor, 1971). The recognition of an *avr* gene product by a receptor in the plant initiates a signal transduction cascade leading to the activation of plant defence mechanisms through *R* gene products (Bent, 1996; Tameling *et al.*, 2002; Hammond-Kosack & Parker, 2003; Innes, 2004). Although many studies have been conducted to understand plant disease resistance in molecular means, this process has been slower than expected. The main reason for this is that plant cells have a very low abundance of *R* genes, and thus their isolation and characterization is rather difficult. The cloning of the initial plant *R* genes by map-based methods or transposon tagging revealed the presence of conserved domains in the protein structure (Bent, 1996). This finding enabled the isolation and characterization of various resistance gene-like sequences (RGLs) using PCR-based degenerate primers (Yu *et al.*, 1996; Kanazin *et al.*, 1996; Aarts *et al.*, 1998; Seah *et al.*, 1998; Donald *et al.*, 2002; Shen *et al.*, 2002; Budak *et al.*, 2006).

Sequence analysis of the plant RGLs revealed that common motifs occur in resistance genes from diverse origins and pathogen specificity (Hulbert *et al.*, 2001). There are at least five different classes of R proteins, the major one carrying a nucleotide-binding site (NBS) followed by a leucine-rich repeat (LRR) region at the C-terminal end of the protein (Richter & Roland, 2000; Meyers *et al.*, 2003; Michelmore & Meyers, 1998). NBS contains three peptide motifs (kinase-1a, kinase-2 and kinase-3) that are critical for ATP/GTP binding and the presence of these domains in R proteins suggests that nucleotide binding is essential for proper functioning (Bent, 1996; Hammond-Kosack & Jones, 1997; Tameling *et al.*, 2002; Belkhadir *et al.*, 2004). The LRR domain contains leucines or other hydrophobic residues at regular intervals and is functional in protein–protein interaction and peptide-ligand binding (Bent, 1996; Jones & Jones, 1997; Kajava, 1998; Tameling *et al.*, 2002). The NBS-LRR gene family is further subdivided into two, based on motifs at the N-terminal to the NBS (Cannon *et al.*, 2002). The *Toll-Interleukin* receptor-like regions (TIR) subfamily (TIR-NBS-LRR, TNL) contains approximately 200 amino acid residues at the N-terminal with high similarity to *Drosophila Toll* and mammalian *Interleukin* receptor-like regions (Meyers *et al.*, 1999; Pan *et al.*, 2000*a*; Young, 2000; Cannon *et al.*, 2002). Plant TIRs, in contrast to their animal counterparts functioning in signal transduction, appear to have a role in pathogen recognition (Young, 2000). In the non-TIR subfamily (non-TIR-NBS-LRR), the N-terminal region consists of either a coiled-coil (CC) or leucine zipper (LZ) structure, both of which facilitate protein–protein interactions (Baker *et al.*, 1997; Pan *et al.*, 2000*a*; Cannon *et al.*, 2002). The CC-NBS-LRR (CNL) subfamily is further divided into the four subgroups CNL-N1, -N2, -N3 and -N4 where CNL-N4 was shown to have conserved motifs like TVS and PKAE domains (Meyers *et al.*, 2003; Chini & Loake, 2005). Other RGL subfamilies include those with an LRR domain and a serine–threonine protein kinase domain with no apparent NBS domain, with a protein kinase domain without an NBS or LRR domain or a CC domain with a putative transmembrane region at the N-terminus (Hulbert *et al.*, 2001; Liu & Ekramoddoullah, 2004).

There is a clear distinction between monocotyledons and dicotyledons with respect to the presence of TNL-type *R* gene families. Neither database searches nor genomic amplification using universal degenerate primers has revealed the presence of any TNL-type *R* genes in monocotyledons (Meyers *et al.*, 1999; Pan *et al.*, 2000*b*; Martin *et al.*, 2003). It was speculated that these type of genes are either completely missing from grass genomes or are present below detectable levels (Pan *et al.*, 2000*b*).

Molecular analyses have indicated a high level of polymorphism of plant *R* genes and enhanced our understanding of the possible mechanisms involved in their evolution (Song *et al.*, 1997; Hulbert *et al.*, 2001; Hammond-Kosack & Parker, 2003; McDowell & Simon, 2006). Most *R* genes are arranged in the genome forming an array of complex loci which can evolve through frequent recombination events (Baker *et al.*, 1997). It is thought that a high frequency of recombination increases the selective advantage of plant *R* gene evolution in the face of rapidly changing pathogen populations (Baker *et al.*, 1997; Baumgarten *et al.*, 2003). Genome sequencing studies have indicated the presence of some RGLs that have no apparent expression and are therefore clearly pseudogenes, supporting the presence of a high rate of evolution in plant *R* genes (Michelmore & Meyers, 1998). In contrast, not all *R* genes are maintained by the same evolutionary forces. For instance, the *Rpm1* locus of *Arabidopsis* was reported to be maintained by balanced selection, indicating the presence of different evolutionary forces on *R* gene families (Stahl *et al.*, 1999).

The availability of *R* genes conferring resistance to diseases opens up an opportunity to improve commercially important plant species. To date, most of the studies have focused on the isolation and characterization of RGLs from genomic DNA

https://doi.org/10.1017/S0016672307008518 Published online by Cambridge University Press

Table 1. *Degenerate primer sequences used for amplification of RGLs from three* Agrostis *species*

Primers	Domain sequence	Primer sequence[a]	Degeneracy
Forward-1	GGVGKTT	5′ ggg ggg gTg ggg AAR Acg Ac 3′	2
Forward-2	GLPLAL	5′ ccH Acg ccR ATg gAW gAc cc 3′	12
Forward-3	GSGGKTT	5′ ggR AcT ggN AAR AcN Acc c 3′	64

[a] R = A, G; N = A, C, G, T; W = A, T; H = A, C, T.

Table 2. *Oligo*(*dT*) '*T*' *primers used for RNA fingerprinting*

Primers[a]	Sequence (5′→3′)
T1	CATTATGCTGAGTGATATCTTTTTTTTT**AA**
T2	CATTATGCTGAGTGATATCTTTTTTTTT**AC**
T3	CATTATGCTGAGTGATATCTTTTTTTTT**AG**
T4	CATTATGCTGAGTGATATCTTTTTTTTT**CA**
T5	CATTATGCTGAGTGATATCTTTTTTTTT**CC**
T6	CATTATGCTGAGTGATATCTTTTTTTTT**CG**
T7	CATTATGCTGAGTGATATCTTTTTTTTT**GA**
T8	CATTATGCTGAGTGATATCTTTTTTTTT**GC**
T9	CATTATGCTGAGTGATATCTTTTTTTTT**GG**

[a] Each 'T' primer contains two variant nucleotides at the 3′-end (bold underlined).

directly, with the aim of understanding the evolution of plant *R* genes. Here we report a PCR-based motif targeted RNA fingerprinting for the cloning and sequencing of differentially expressed cDNAs from *Agrostis*.

2. Materials and methods

(i) *Plant materials*

Approximately 1000 seeds each from creeping bentgrass (*A. stolonifera* L., Southshore), velvet bentgrass (*A. canina* L., SR7200) and colonial bentgrass (*A. capillaris* L., Heriot) species were planted in the greenhouse to ensure that the resulting plant populations represented the diversity of each species. The genotypes were planted in 15 cm diameter pots containing a mixture of peat, vermiculite, soil and sand (35 : 32 : 9 : 24, by volume). Soil was saturated biweekly with nutrient solution (21N–1·5P–12·5K) supplemented with 200 mg/l nitrogen. The greenhouse was maintained at 25 ± 1 °C with supplemental light on a 15 h day/9 h night photoperiod (Budak *et al*., 2005). Individual seedlings of each *Agrostis* species were used for cDNA amplification of RGL sequences.

(ii) *Total RNA isolation*

Total RNA was isolated from 0·2 g frozen tissue using Trizol reagent (Invitrogen) according to the manufacturer's instructions. The concentration was quantified spectrophotometrically. DNase treatment was performed in 50 μl reaction mixture containing 1× reaction buffer with $MgCl_2$, 5 μg of total RNA and 5 U of RNase-free DNase I (Fermantas). The reaction was maintained at 37 °C for 30 min and terminated by adding 5 μl of 25 mM EDTA followed by an incubation at 65 °C for 10 min. Samples were then stored at −20 °C.

(iii) *Degenerate primer design*

Three oligonucleotides (Table 1) were designed based on the highly conserved motifs in the kinase-1a domain of NBS-LRR type *R* genes using the Vector NTI program (version 9.0, Invitrogen), using known *R* sequences from rice (*Oryza sativa* L., AB002266 and AB019186), diploid wheat (*Aegilops tauschii*, AAC05834), barley (*Hordeum vulgare* L., AJ302292) and buffalograss (*Buchloe dactyloides* (Nutt.) Engelm, AY966893 and AY970294). Each of the designed primers was combined with the 'T' primers (Table 2) for PCR amplification of RGL sequences from bentgrass species.

(iv) *cDNA synthesis and PCR amplifications*

Total RNA from bulk leaf samples was reverse transcribed using an Omniscript RT kit (Qiagen). Amplification of cDNA fragments was performed in 20 μl PCR reactions. Each reaction mixture contained 2 μl first-strand cDNA, 2 μl 10× PCR buffer without $MgCl_2$, 2·5 mM $MgCl_2$, 200 μM dNTP mix, 2 μCi [α-^{33}P]dATP (Institute of Isotopes Co., Hungary), 1 μM of forward primer (Table 1), 1 μM of reverse primer (Table 2) and 1·25 units *Taq* DNA polymerase. Amplifications were carried out using an MJ Research PTC-100 thermocycler programmed for 32 cycles of 1 min at 94 °C, 1 min at 50 °C, 1 min at 72 °C; followed by a final extension at 72 °C for 5 min. The PCR products obtained from first-strand cDNAs synthesized from total RNA were size-fractionated in 6% polyacrylamide/8 M urea gel. The gel was blotted onto a Whatman 3M paper and dried using a gel dryer (E-C355, E-C Apparatus Corporation) at 76 °C for 50 min. A Kodak X-ray film was placed on the dried gel and exposed for 14 h at −80 °C.

https://doi.org/10.1017/S0016672307008518 Published online by Cambridge University Press

(v) *cDNA cloning and sequence analysis*

Amplified fragments were excised from gels, extracted by a QiaexII gel extraction kit (Qiagen), inserted into pGEMT-Easy vector (Promega) and propagated in *E. coli* grown in Luria Bertani (LB) medium supplemented with 100 μg/ml ampicillin.

The subcloned fragments were sequenced and exposed to the VecScreen algorithm (http://www.ncbi.nlm.nih.gov/) to remove vector contamination. Sequences were then grouped at 90% nucleotide identity level using ContigExpress (Vector NTI, version 9.0, Invitrogen) and the longest sequence from each group was used for further analysis. The BLAST algorithm (Altschul *et al*., 1997) was used to analyse ESTs. The identified RGL sequences were compared with known *R* genes in order to show their structural resemblance. The ClustalW algorithm (Chenna *et al*., 2003) was performed to identify the motifs. Exon prediction was conducted using the gene-finding programs GENESCAN (Burge & Karlin, 1997) and GENEMARK (Lomsadze *et al*., 2005) as outlined by Budak *et al*. (2005). All sequences were run against rice.

(vi) *Data analysis*

A phylogenetic tree was constructed based on the neighbour joining (NJ) method with a Dayhoff matrix (Schwartz & Dayhoff, 1978) using the Vector NTI program (version 9.0, Invitrogen).

(vii) *Rapid amplification of cDNA ends (5′RACE)*

To amplify the corresponding full-length cDNA, RACE was performed using the GeneRacer kit (Invitrogen) according to the manufacturer's protocol. Gene-specific primers (i.e. 5′RACE-ccagcggggagcccatccttgatgg) and (3′-aagtcaagcatactcctgtgccgtt) were designed based on the sequence information obtained from cDNAs. Each amplification reaction was cloned into pBlueScript II KS (+) (Stratagene) for sequence analyses.

3. Results

The combinations of each kinase-1a domain specific degenerate (Table 1) and 'T' primer (Table 2) pairs were used for cDNA amplification of RGL sequences from three *Agrostis* species. Amplification products that had been size-separated on 6% polyacrylamide/8 M urea gel generated 40–90 bands (Fig. 1). The sizes of the PCR products ranged from 150 to 1100 base pairs. Of the total 180, only about 80 fragments were found to be unique. Sixty-one clones (29 from creeping bentgrass, 21 from velvet bentgrass and 11 from colonial bentgrass) were successfully isolated and

Fig. 1. ^{33}P-labelled PCR amplification products of three *Agrostis* species separated on a 6% polyacrylamide/8 M urea gel. Arrows point to the fragment size.

sequenced from the three *Agrostis* species. After removal of the primer and vector sequences, 39 unique contig tags were identified using ContigExpress (Vector NTI, version 9.0) and the longest sequence of each contig was used for further analysis.

The RGL sequences cloned from all three bentgrass species were found to be highly divergent both at nucleotide and amino acid levels. Putative protein sequences of 15 RGLs were compared with sequences available in the GenBank database by BLASTX searches (Altschul *et al*., 1997). These sequences showed either partial or overall homology to the known or putatively annotated genes with different E values (Table 3). The remaining 24 *Agrostis* sequences proved similarity to pathogenesis-related proteins (proteases, kinases and phosphatases), ATP- and GTP-binding motif carrying proteins and stress-related kinases, indicating these species contain a large and diverse family of genes for proteins carrying this domain. Since our main objective was to isolate and characterize constitutively expressed RGLs, these

https://doi.org/10.1017/S0016672307008518 Published online by Cambridge University Press

Table 3. *BLASTX results of RGLs obtained from* Agrostis *species*

Accession no.	Homologous GenBank accessions	Species *R* gene isolated	% identity	Probability	ORF		
					Frame	From	To
EE284243	CC-NBS-LRR resistance protein MLA13 *Hordeum vulgare* (AAO16014.1)	Velvet bentgrass	75	4e-28	−3	144	446
EE284244	RGA: NBS-LRR-like resistance protein *Oryza sativa* (AAM69508.1)	Velvet bentgrass	97	2e-34	+3	114	290
EE284245	Similar to NBS-LRR-type resistance gene *Oryza sativa* (NP_910483.1)	Velvet bentgrass	57	2e-07	+3	159	531
EE284246	NBS-LRR resistance protein *Lolium perenne* (AAY56467.1)	Velvet bentgrass	91	1e-76	+2	449	589
EE284247	CC-NBS-LRR resistance protein MLA13 mRNA, complete cds; alternatively spliced *Hordeum vulgare* (AF523683.1)	Creeping bentgrass	97	3e-146	−2	596	877
EE284248	NBS-LRR-like protein (YR6) gene, partial cds *Oryza sativa* (AF220732.1)	Creeping bentgrass	93	3e-73	+1	346	540
EE284249	NBS-LRR-like resistance protein *Oryza sativa* (AAM69514.1)	Creeping bentgrass	98	5e-39	+2	128	466
EE284250	Downy mildew resistance protein RPP5 *Arabidopsis thaliana* (AAF08790.1)	Creeping bentgrass	5851	3e-40, 3e-33	+3	318	530
EE284251	Putative disease resistance complex protein *Oryza sativa* (AAP05798.1)	Creeping bentgrass	78	3e-72	+2	299	533
EE284252	RGA: *Zizania latifolia* clone Zi-FL10 disease partial cds (DQ239431.1)	Creeping bentgrass	89	2e-157	+3	108	536
EE284253	Disease resistance gene analogue *Zizania latifolia*	Creeping bentgrass	86	6e-84	+3	108	517
EE284254	NBS-LRR-like resistance protein *Oryza sativa* (AAM69508.1)	Colonial bentgrass	95	2e-80	+1	181	530
EE284255	NBS-LRR-type resistance protein *Hordeum vulgare* (AAB96979.1)	Colonial bentgrass	77	1e-71	+3	108	539
EE284256	Putative NBS-LRR-type resistance *Oryza sativa* (AAT69649.1)	Colonial bentgrass	80	3e-72	−3	1	154
EE284257	Mla-like protein *Triticum aestivum* (AAO65985.1)	Colonial bentgrass	75	2e-71	+1	1	465

sequences were not analysed further. Some of the sequences isolated did not match with any of the plant NBS/LRR resistance genes even in relatively conserved regions. The same behaviour was also observed in buffalograss species and might indicate the high variability of the RGL sequences (Budak *et al*., 2006).

Open reading frame analysis (ORF) of the 39 different bentgrass RGLs was performed using GENESCAN (Burge & Karlin, 1997) and GENEMARK (Lomsadze *et al*., 2005) programs as outlined by Budak *et al*. (2005). Of the 39 *Agrostis* sequences, ORFs of 15 RGL sequences were characterized and are depicted in Table 3. However, 10 cDNA clones (4 from creeping bentgrass, 3 from colonial bentgrass and 3 from velvet bentgrass) did not have any ORFs as indicated by the presence of multiple stop codons and were not analysed further. This response might indicate that some isolated *Agrostis* RGL sequences could be pseudogenes.

BLASTX analysis of EE284250 showed high homology with downy mildew resistance protein RPP5 from *Arabidopsis thaliana* (NM 179067·1) with an E value of 3e-40. Two other RGL sequences, EE284247

https://doi.org/10.1017/S0016672307008518 Published online by Cambridge University Press

(creeping bentgrass) and EE284243 (velvet bentgrass), have high levels of homology (75% and 97% identity, respectively) to MLA13, which confers resistance to the fungal pathogen *B. graminis*, isolated from the Poaceae family. Because *Agrostis* belongs to the Poaceae family, it is possible to assume that these RGL sequences may be involved in resistance to *B. graminis*.

The C-terminal amino acid of the kinase-2 motif indicates whether the RGL is of the CNL (Coiled-Coil-NBS-LRR) or TNL (*Toll-Interleukin* receptor-like-NBS-LRR) type: the CNL kinase-2 domain ends with a tryptophan (W) residue, whereas TNL types end with an aspartic acid (D) residue (Meyers *et al*., 1999). Of the predicted kinase-2 motifs among RGL sequences, EE284246 from velvet bentgrass and EE284254 from creeping bentgrass end with a tryptophan (W) residue (Fig. 2), indicating that these sequences belong to CNL-type *R* genes. For the CNL-type RGLs, NBS and GLPL domains are also depicted in Fig. 2. Comparison of the deduced amino acid sequences of these RGLs with known CNL and TNL types of NBS-LRR proteins also revealed a high level of similarity to CNL-type *R* genes (Table 3). On the other hand, EE284250 from creeping bentgrass was shown to possess a RNBS-A motif and kinase-2 domain of EE284257 from colonial bentgrass ends with an aspartic acid (D) residue, implying these RGLs are of TNL type (Fig. 2).

A neighbour joining tree of the deduced amino acid sequences of 15 RGLs from *Agrostis* species and comparison with known TIR- and non-TIR-NBS-LRR type *R* genes from different plant species revealed that constitutively transcribed RGLs of *Agrostis* species are highly diverse (Fig. 3). Translated RGLs formed two main groups; in the first group, RGLs of all three *Agrostis* species – EE284243, EE284245 (both from velvet bentgrass), EE284247, EE284250 (both from creeping bentgrass) and EE284257 (from colonial bentgrass) – were found to be in close proximity with the known TNL-type *R* genes RPP5 (NM179067, *A. thaliana*), Lu-M (AAB47618, *L. usitassimum*, flax), L6 (AAD25965, *L. usitassimum*, flax) and Lu-L5 (AFO93645, *L. usitassimum*, flax). Although not all RGLs from this group could be shown to carry TNL-type motifs like RNBS-A or a kinase-2 domain ending with a D residue, it can be assumed on the basis of the neighbour joining tree that they are also TNL-type *R* genes. The RGLs in the second group showed similarities only to known non-TIR type R proteins, namely YR16 (AF220739·1, *O. sativa*), Xa1 (BAA25068, *O. sativa*), MLA6 (CAC29241, *H. vulgare*), MLA13 (AAO16014·1, *H. vulgare*). YR23 (AF220745·1, *O. sativa*), AtRPS2 (AF487823, *A. thaliana*), AtRPS5 (AFO74916, *A. thaliana*) and BoRPS2 (AAF19803, *B. oleracea*) (Fig. 3).

Evolutionary forces were also examined on NBS and LRR domains of constitutively expressed RGL sequences by calculating the ratio of non-synonymous nucleotide substitutions (Ka) to synonymous nucleotide substitutions (Ks) in *Agrostis* species (Nei & Gojobori, 1986). The majority of nucleotide variations in NBS of RGL sequences represent the non-synonymous substitutions. Ka/Ks values for pairwise comparisons ranged from 0·23 to 0·34 in creeping bentgrass, 0·13 to 0·21 in velvet bentgrass, and 0·21 to 0·32 in colonial bentgrass. These results (i.e. Ka/Ks ratio smaller than 1) indicate purifying selection in *Agrostis* species. Similar results were also found in the LRR domain. The Ka/Ks values for pairwise comparison of LRR domain sequences ranged from 0·45 and 0·42 in creeping bentgrass, 0·25 to 0·43 in velvet bentgrass and 0·34 to 0·42 in colonial bentgrass, supporting purifying selection in constitutively expressed RGL sequences.

The 5′-RACE was performed to obtain the untranslated region (5′-UTR) of the identified *R* gene-like sequences in all three *Agrostis* species. This analysis helped to locate the position of the translation start codon more precisely. Two distinct 5′-RACE products from both creeping and colonial bentgrass and one distinct 5′-RACE product from velvet bentgrass were isolated. Sequence analysis of the identified 5′-RACE products in these species showed a high level of similarity, ranging from 75% to 98% in *Agrostis* species. The 5′-UTRs were found to be approximately 80 bp in length. In colonial bentgrass RGL sequences, the upstream sequences showed 78–99% similarity with an average 5′-UTR length of 70 bp (data not shown).

The existence of common RGL sequences was tested in three *Agrostis* and buffalograss species. The sequence analysis of the resulting PCR products of expected size (500 bp) from buffalograss and three *Agrostis* species provided the evidence for common RGL sequences among the species compared (Fig. 4). The existence of common RGL sequences suggests that RNA fingerprinting combined with PCR-based motif-directed profiling can potentially be used for comparative genomics analysis in grasses.

4. Discussion

To our knowledge, this is the first study on the isolation and characterization of RGL sequences based on PCR-based motif-directed RNA fingerprinting from three *Agrostis* species. There are a number of studies in the literature based on the degenerate primer design on conserved motifs and genomic DNA amplification to identify and characterize RGL sequences from a variety of plant species (Yu *et al*., 1996; Kanazin *et al*., 1996; Aarts *et al*., 1998;

https://doi.org/10.1017/S0016672307008518 Published online by Cambridge University Press

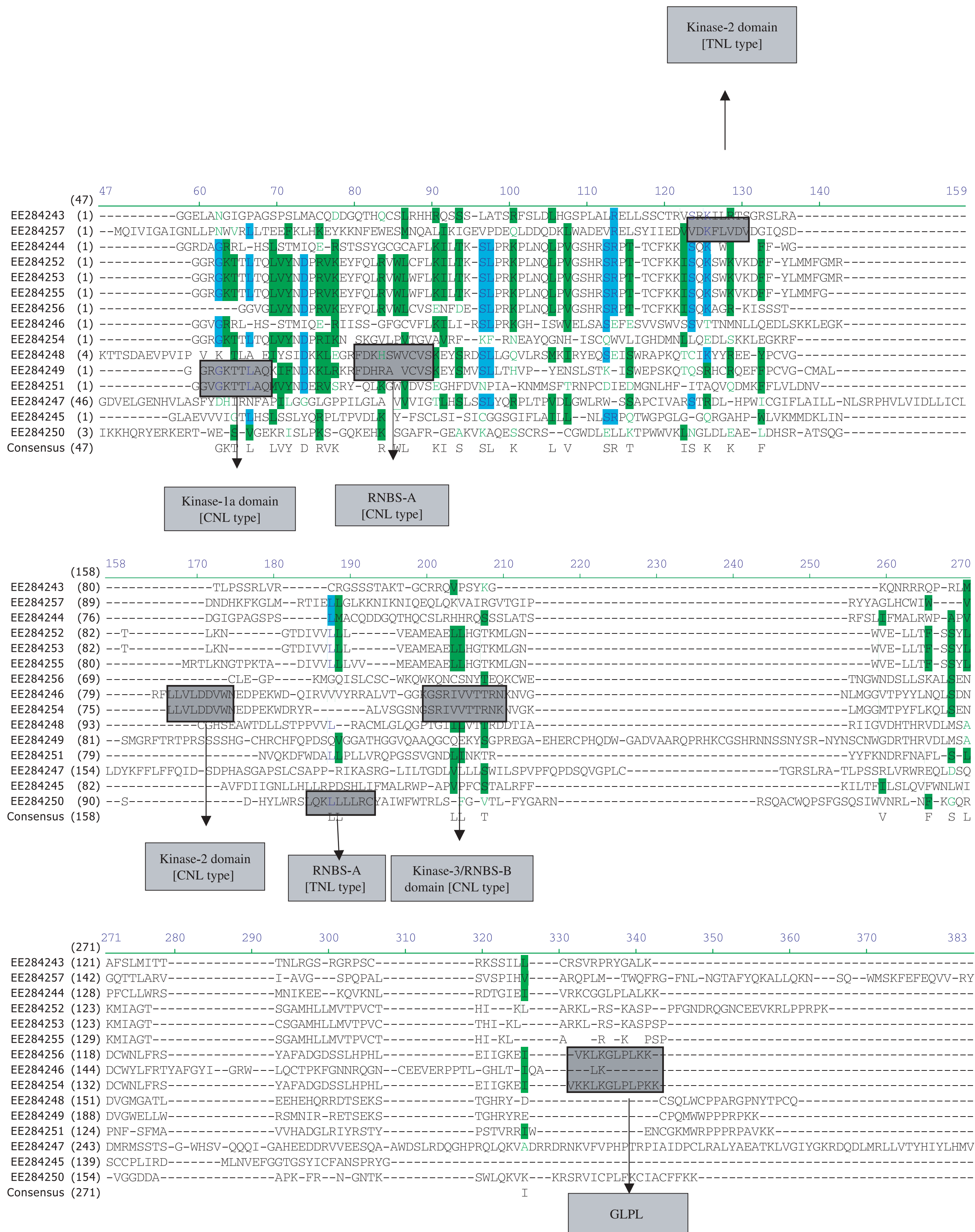

Fig. 2. ClustalW analysis of identified RGLs. Boxes indicate the CNL-type or TNL-type motifs found in these sequences.

Seah *et al.*, 1998; Donald *et al.*, 2002; Shen *et al.*, 2002). This study differs from previous studies by isolating cDNA (not genomic DNA) in the absence of a pathogen attack.

A PCR-based motif-directed RNA fingerprinting can be useful for systematically searching for RGLs in the *Agrostis* genome. Thus the present study demonstrates the isolation of constitutively expressed RGLs

https://doi.org/10.1017/S0016672307008518 Published online by Cambridge University Press

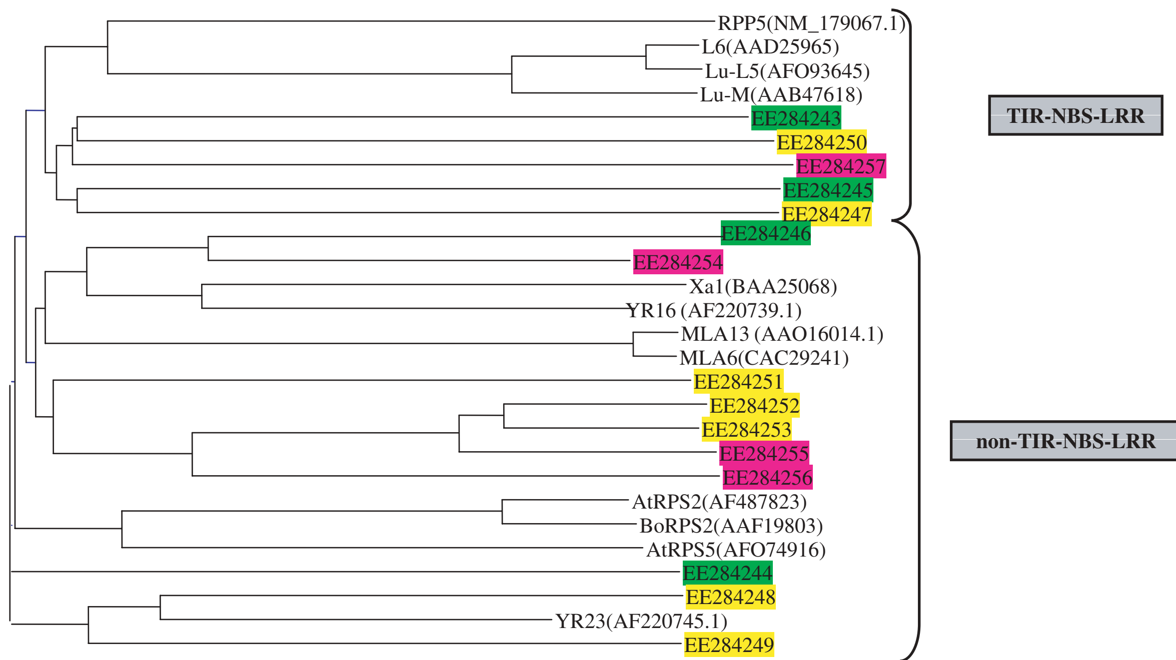

Fig. 3. Neighbour joining tree constructed on the basis of a Dayhoff matrix (Schwartz & Dayhoff, 1978) on 15 ESTs from *Agrostis* species and comparison with known *R* genes from different plant species. GenBank accession numbers between EE284243 and EE284246 are for velvet bentgrass (green colour), EE284247 and EE284253 for creeping bentgrass (yellow colour), and EE284254 and EE284257 for colonial bentgrass (pink colour).

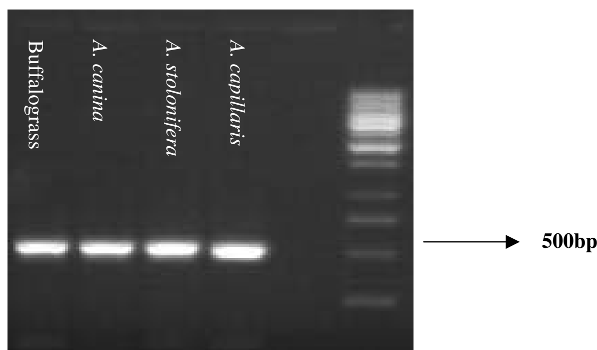

Fig. 4. cDNA amplification products derived from buffalograss and three *Agrostis* species.

and structural and genetic diversity that exist among *R* genes at both the nucleotide and amino acid level (Figs. 2 and 3, Table 3). Although some RGL sequences might be not functional due to stop codons or disruption of open reading frames, the *Agrostis* RGLs cloned in the present study resemble *R* genes (Table 3). Ploidy level and genome organization are important factors that might affect the accumulation of pseudogenes in *Agrostis* species. The pseudogene ratio was 10% in *Arabidopsis* (Meyers *et al*., 2003), 11% in soybean (Kanazin *et al*., 1996), 9·3% in tomato (Pan *et al*., 2000*a*) and 29% in buffalograss (Budak *et al*., 2006). Several researchers have demonstrated that sequencing of paralogues of *Xa21*, *Cf9* and *Dm3* indicated the presence of pseudogenes (Parniske *et al*., 1997; Song *et al*., 1997; Meyers *et al*., 1998). Shen *et al*. (2002) identified only one functional resistance gene in 800 resistance gene analogues from 20 plant species.

Deduced amino acid sequence results indicated EE284243 and EE284245 of velvet bentgrass, EE284247 and EE284250 of creeping bentgrass, and EE284257 of colonial bentgrass RGLs closely match with downy mildew resistance protein RPP5 from *Arabidopsis* (AAFO8790), which confers resistance against *Hyaloperonospora parasitica*, an oomycete. Therefore, it is possible to assume that the RGLs indicated can be involved in this type of resistance.

A neighbour joining tree constructed on the basis of a Dayhoff matrix (Schwartz & Dayhoff, 1978) indicated that RGLs isolated from *Agrostis* species are highly diverse (Fig. 3). The obtained *Agrostis* RGLs sequences mainly grouped into two classes (TIR and non-TIR-NBS-LRR). Five of the bentgrass RGL sequences were in the same cluster with known TNL-type R proteins such as RPP5, AtRPS2 and BoRPS2. However, two RGL sequences (EE284250 and EE284257) were shown to contain clear TNL-type motifs (Fig. 2). The majority of constitutively expressed RGL sequences in *Agrostis* grouped with known *R* genes and proteins of *O. sativa*, common rice. Since both species are of the same family, Poaceae, this may support the hypothesis that the divergence of *R* genes is also controlled by family- or

https://doi.org/10.1017/S0016672307008518 Published online by Cambridge University Press

species-specific pathogens. The branching of R proteins of *H. vulgare* (Poaceae) with that of *L. usitassimum* (Linaceae), and the presence of two of the RGLs identified from *Agrostis* species, may indicate that the same evolutionary forces were effective on this group of proteins.

The Ka/Ks ratio, an indicator of the evolutionary pressures on a gene, revealed that both NBS and LRR domains seem to be under purifying selection; thus, divergence mechanisms may be acting very slowly within the constitutively expressed RGL genes of these three *Agrostis* species. Our results are in good agreement with comments made by Michelmore & Meyers (1998) who reported that *R* genes might not be evolving rapidly but rather slowly to provide resistance against pathogens that are heterogeneous in space and time. This finding might be the result of not having enough RGL sequences from *Agrostis* species to clearly identify the selection methods on evolution.

In the literature, all the *R* genes amplified appears to belong to the non-TIR-NBS-LRR subfamily and there are no reports revealing TNL-type *R* gene families in monocotyledons (Meyers *et al*., 1999; Pan *et al*., 2000*b*; Martin *et al*., 2003). It was speculated that these type of genes are either completely missing from the grass genomes or are present beyond detectable levels (Pan *et al*., 2000*b*). In the present study, for the first time, two TNL-type RGLs were isolated from *Agrostis*, a monocotyledonous plant. This striking finding might indicate the presence of both CNL- and TNL-type *R* genes in monocots and can enhance knowledge on the evolution and utilization of RGLs in grass species. Furthermore, the RGL sequences obtained in the present study could facilitate the design of new molecular markers to trace *R* genes and could also be placed onto the *Agrostis* linkage map to identify R loci to improve the disease resistance of this economically important species.

The authors thank Drs Stacy Bonos and William Meyer for providing us with the *Agrostis* species used in this study. This project was partially supported by Sabanci University Research Funding and the European Commission's Marie Curie International Reintegration Grant (FP6).

References

Aarts, M. G. M., Hekkert, B. L., Holub, E. B., Beynon, J. L., Stiekema, W. J. & Pereira, A. (1998). Identification of R-gene homologous DNA fragments genetically linked to disease resistance loci in *Arabidopsis thaliana*. *Molecular Plant–Microbe Interactions* **11**, 251–258.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–3402.

Baker, B., Zambryski, P., Staskawicz, B. & Dinesh-Kumar, S. P. (1997). Signaling in plant–microbe interactions. *Science* **276**, 726–733.

Baumgarten, A., Cannon, S., Spangler, R. & May, G. (2003). Genome-level evolution of resistance genes in *Arabidopsis thaliana*. *Genetics* **165**, 309–319.

Belanger, F. C., Plumley, K. A., Dat, P. R. & Meyer, W. A. (2003). Interspecific hybridization as a potential method for improvement of *Agrostis* species. *Crop Science* **43**, 2172–2176.

Belanger, F. C., Bonos, S. & Meyer, W. A. (2004). Dollar spot resistant hybrids between creeping bentgrass and colonial bentgrass. *Crop Science* **44**, 581–586.

Belkhadir, Y., Sumramaniam, R. & Dangl, J. L. (2004). Plant disease resistance protein signaling: NBS-LRR proteins and their partners. *Current Opinion in Plant Biology* **7**, 391–399.

Bent, A. F. (1996). Plant disease resistance genes: Function meets structure. *Plant Cell* **8**, 1757–1771.

Bliffeld, M., Mundy, J., Potrykus, I. & Fütterer, J. (1999). Genetic engineering of wheat for increased resistance to powdery mildew resistance. *Theoretical and Applied Genetics* **98**, 1079–1086.

Bonos, S. A., Clarke, B. B. & Meyer, W. A. (2006). Breeding for disease resistance in the major cool-season turfgrasses. *Annual Review of Phytopathology* **45**, 213–234.

Boulter, J. I., Boland, G. J. & Trevors, J. T. (2002). Evaluation of composts for suppression of dollar spot (*Sclerotinia homoeocarpa*) of turfgrass. *Plant Disease* **86**, 405–410.

Budak, H., Shearman, R. C., Gulsen, O. & Dweikat, I. (2005). Understanding ploidy complex and geographic origin of the *Buchloe dactyloides* genome using cytoplasmic and nuclear marker systems. *Theoretical and Applied Genetics* **111**, 1545–1552.

Budak, H., Kasap, Z., Shearman, R. C., Dweikat, I., Sezerman, U. & Mahmood, A. (2006). Molecular characterization of cDNA encoding resistance gene like sequences in *Buchloe dactyloides*. *Molecular Biotechnology* **34**, 293–301.

Burge, C. & Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. *Journal of Molecular Biology* **268**, 78–94.

Cannon, S. B., Zhu, H., Baumgarten, A. M., Spangler, R., May, G., Cook, D. R. & Young, N. D. (2002). Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. *Journal of Molecular Evolution* **54**, 548–562.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. & Thompson, J. D. (2003). Multiple sequence alignment with the clustal series of programs. *Nucleic Acids Research* **31**, 3497–3500.

Chini, A. & Loake, G. J. (2005). Motifs specific for the ADR1 NBS-LRR protein family in *Arabidopsis* conserved among NBS-LRR sequences from both dicotyledonous and monocotyledonous plants. *Planta* **21**, 597–601.

Couch, H. B. (1995). *Disease of Turfgrasses*. London: Chapman & Hall.

Dai, W. D., Bonos, S., Guo, Z., Meyer, W. A., Day, P. R. & Belanger, F. C. (2003). Expression of pokeweed antiviral proteins in creeping bentgrass. *Plant Cell Reports* **21**, 497–502.

Deacon, J. W. (1973). Factors affecting occurrence of the *Ophiobolus* patch disease of turf and its control by *Phialophora radicicola*. *Plant Pathology* **22**,

https://doi.org/10.1017/S0016672307008518 Published online by Cambridge University Press

149–153. [Note: *Ophiobolus graminis* is now named *Gaeumannomyces graminis*; *Phialophora radicicola* is now named *P. graminicola*.]

Donald, T. M., Pellerone, F., Adam-Blondon, A. F., Boquet, A., Thomas, M. R. & Dry, I. B. (2002). Identification of resistance gene analogs linked to a powdery mildew resistance locus in grapevine. *Theoretical and Applied Genetics* **104**, 610–618.

Flor, H. H. (1971). Current status of the gene for gene concept. *Annual Reviews of Phytopathology* **9**, 275–296.

Glazebrook, J. (2005). Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology* **43**, 205–227.

Guo, Z., Bones, S., Meyer, W. A., Day, P. R. & Belanger, F. C. (2003). Transgenic creeping bentgrass with delayed dollar spot symptoms. *Molecular Breeding* **11**, 95–101.

Hammond-Kosack, K. E. & Jones, J. (1997). Plant disease resistance genes. *Annual Review of Plant Physiology and Plant Molecular Biology* **48**, 575–608.

Hammond-Kosack, K. E. & Parker, J. E. (2003). Deciphering plant pathogens communications: fresh perspectives for molecular resistance breeding. *Current Opinion in Biotechnology* **14**, 177–193.

Hitchcock, A. S. (1971). *Manual of the Grasses of the United States*. New York: Dover Publications.

Hulbert, S. H., Webb, C. A., Smith, S. M. & Sun, Q. (2001). Resistance gene complexes: evolution and utilization. *Annual Reviews of Phytopathology* **39**, 285–312.

Innes, R. W. (2004). Guarding the goods. New insights into the central alarm system of plants. *Plant Physiology* **135**, 695–701.

Jones, D. A. & Jones, J. D. G. (1997). The role of leucine-rich repeats proteins in plant defenses. *Advanced Botanical Research* **24**, 89–167.

Jones, J. D. G. & Dangl, J. L. (2006). The plant immune system. *Nature* **444**, 323–329.

Jørgensen, J. H. (1994). Genetics of powdery mildew resistance in barley. *Critical Reviews in Plant Sciences* **13**, 97–119.

Kajava, A. V. (1998). Structural diversity of leucine-rich repeat proteins. *Journal of Molecular Biology* **277**, 519–527.

Kanazin, V., Marek, L. F. & Shoemaker, R. C. (1996). Resistance gene analogs are conserved and clustered in soybean. *Proceedings of the National Academy of Sciences of the USA* **93**, 11746–11750.

Liu, J. J. & Ekramoddoullah, A. K. M. (2004). Isolation, genetic variation and expression of TIR-NBS-LRR resistance gene analogs from western white pine (*Pinus monticola* Dougl. ex. D. Don.). *Molecular Genetics and Genomics* **270**, 432–441.

Lomsadze, A., Ter-Houhannisyan, V., Chernoff, Y. & Borodovsky, M. (2005). Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Research* **33**, 6494–6506.

Martin, G. B., Bagdanove, A. J. & Sessa, G. (2003). Understanding the function of plant disease resistance proteins. *Annual Review of Plant Biology* **54**, 23–61.

McDowell, J. M. & Simon, S. A. (2006). Recent insight into *R* gene evolution. *Molecular Plant Pathology* **7**, 437–448.

Meyers, B. C., Shen, K. A., Rohani, P., Gaut, B. & Michelmore, R. W. (1998). Receptor like genes in the major resistance locus of lettuce are subject to divergent selection. *Plant Cell* **10**, 1833–1846.

Meyers, B. C., Dickerman, A. W., Michelmore, R. W., Sivaramakrishnan, S., Sobral, B. W. & Young, N. D. (1999). Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant Journal* **20**, 317–332.

Meyers, B. C., Kozik, A., Griego, A., Kuang, H. & Michelmore, R. W. (2003). Genome wide analysis of NBS-LRR encoding genes in *Arabidopsis*. *Plant Cell* **15**, 809–834.

Michelmore, R. W. & Meyers, B. C. (1998). Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Research* **8**, 1113–1130.

Nei, M. & Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and non-synonymous nucleotide substitutions. *Molecular Biology and Evolution* **3**, 418–426.

Pan, Q., Liu, Y. S., Budai-Hadrian, O., Sela, M., Carmel-Goren, L., Zamir, D. & Fluhr, R. (2000*a*). 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., Wendel, J. & Fluhr, R. (2000*b*). Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. *Journal of Molecular Evolution* **50**, 203–213.

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

Richter, T. E. & Ronald, P. C. (2000). The evolution of disease resistance genes. *Plant Molecular Biology* **42**, 195–204.

Scheef, E. A., Casler, M. D. & Jung, G. (2003). Development of species-specific SCAR markers in bentgrass. *Crop Science* **43**, 345–349.

Schwartz, R. M. & Dayhoff, M. O. (1978). *Atlas of Protein Sequences and Structure*. Washington: National Biomedical Research Foundation.

Seah, S., Bariana, H., Jahier, J., Sivasithamparam, J. & Lagudah, E. S. (1998). Cloning and characterization a family of disease resistance gene analogs from wheat and barley. *Theoretical and Applied Genetics* **97**, 937–945.

Shen, K. A., Chin, D. B., Arroyo-Garcia, R., Ochoa, O. E. & Lavelle, D. O. (2002). Dm3 is one member of a large constitutively expressed family of nucleotide binding site leucine-rich repeat encoding genes. *Molecular Plant Microbe Interactions* **15**, 251–261.

Song, W. Y., Pi, L. Y., Wang, G. L., Gardner, J., Holsten, T. & Ronal, P. C. (1997). Evolution of the rice Xa21 disease resistance gene family. *Plant Cell* **9**, 1279–1287.

Stahl, E. A., Dwyer, G., Mauricio, R., Kreitman, M. & Bergelson, J. (1999). Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. *Nature* **400**, 667–671.

Tameling, W. I., Elzinga, S. D., Darmin, P. S., Vossen, J. H., Takken, F. L., Haring, M. A. & Cornelissen, B. J. (2002). The tomato *R* gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity. *Plant Cell* **14**, 2929–2939.

Vargas, J. M. Jr (1994). *Management of Turfgrass Diseases*. Florida: CRC Press.

Vincelli, P. & Doney, J. C. (1997). Variation among creeping bentgrass cultivars in recovery from epidemics of dollar spot. *Plant Disease* **81**, 99–102.

Young, N. D. (2000). The genetic architecture of resistance. *Current Opinion in Plant Biology* **3**, 285–290.

https://doi.org/10.1017/S0016672307008518 Published online by Cambridge University Press

Yu, Y. G., Bussand, G. F. & Maroof, M. A. (1996). Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. *Proceedings of the National Academy of Sciences of the USA* **93**, 11751–11756.

Zhou, F., Kurth, J., Wei, F., Elliott, C., Vale, G., Yahiaoui, N., Keller, B., Somerwille, S., Wise, R. & Schulze-Lefert, P. (2001). Cell-autonomous expression of barley *Mla1* confers race-specific resistance to the powdery mildew fungus via a *Rar1*-independent signaling pathway. *Plant Cell* **13**, 337–350.

https://doi.org/10.1017/S0016672307008518 Published online by Cambridge University Press