EDINBURGH JOURNAL OF BOTANY 62 (3): 119–126 (2005)
doi:10.1017/S0960428606000126
Issued 12 May 2006

A PRIMER SET FOR SPECIFIC AMPLIFICATION OF TWO CYCLOIDEA-LIKE GENES IN THE GENISTOID CLADE OF *LEGUMINOSAE* SUBFAM. *PAPILIONOIDEAE*

H. L. CITERNE

Locus-specific primers were designed to amplify the homologues LEGCYC1A and LEGCYC1B of the snapdragon (*Antirrhinum majus*) floral symmetry gene CYCLOIDEA in the genistoid legumes *Lupinus nanus* and *Cadia purpurea*. These primers were tested successfully in a range of genistoid taxa. Sequence comparison between *L. nanus* and *L. angustifolius* revealed that these CYCLOIDEA-like genes can provide up to four times as many variable characters as the ribosomal internal transcribed spacers (ITS) at this taxonomic level. These genes are therefore potentially useful for studying evolutionary relationships between closely related species within the genistoid clade of legumes, which contains many large and diverse genera. Comparisons between *L. nanus*, *Cadia purpurea* and *Calpurnia aurea* LEGCYC sequences show that these genes may also be useful for phylogenetic studies between genera in this large clade of 1800 species.

Keywords. CYCLOIDEA, *Fabaceae*, genistoid clade, *Leguminosae*, nuclear gene, phylogenetics.

INTRODUCTION

The most commonly used nuclear regions for phylogenetic analysis at low taxonomic levels are the internal transcribed spacers (ITS) of the nuclear 18S–5.8S–26S ribosomal DNA gene family (Alvarez & Wendel, 2003). There is a need, however, for rapidly evolving low-copy nuclear genes in systematic studies (e.g. Mort & Crawford, 2004). Multiple sources of informative molecular data are required for testing the congruence of topologies of different gene trees in order to have more reliable estimates of taxic relationships, or to investigate hybridization events (Doyle, 1992). In addition, the nature of the ITS region has certain limitations, being part of a multigene family that is homogenized through concerted evolution. Reports of incomplete concerted evolution or pseudogene evolution in this gene family suggest that sequencing of ITS may be subject to complicating factors (Alvarez & Wendel, 2003; Bailey *et al.*, 2003). ITS divergence between closely related taxa may also be too small to resolve relationships, in part due to the short length of the ITS region (c.450 bp), and to the homogenizing effect of concerted evolution (reviewed in Hershkovitz *et al.*, 1999). For instance, species relationships within the

Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, UK.

genistoid genus *Lupinus* L. could not be well resolved using ITS (Ainouche & Bayer, 1999). Molecular data from single-copy nuclear genes providing more variable characters are therefore needed to resolve rapid radiations at the species level.

CYCLOIDEA (CYC) is a rapidly evolving transcription factor that controls floral symmetry in *Antirrhinum majus* L. (*Veronicaceae*, *Lamiales*; Luo *et al.*, 1996; Gübitz *et al.*, 2003). Homologues of CYC have been isolated in the *Leguminosae* (Citerne *et al.*, 2003), a family that has evolved highly specialized zygomorphic (bilaterally symmetric) flowers independently of the *Lamiales*. Phylogenetic analyses suggest that legume CYC (LEGCYC) genes have duplicated at least three times in the evolution of subfamily *Papilionoideae* (Citerne *et al.*, 2003). These genes were found to be evolving rapidly by substitution, insertions and deletions so that unambiguous alignment of sequences in the variable regions outside the conserved TCP and R domains was possible only between closely related taxa (Citerne *et al.*, 2003).

As part of a study to investigate whether CYC-like genes are involved in the control of floral symmetry in legumes, locus-specific primers were designed to amplify two paralogous CYC-like genes, LEGCYC1A and LEGCYC1B (Citerne *et al.*, 2003), in closely related species belonging to the core genistoid clade (*sensu* Wojciechowski, 2003) with contrasting floral morphology: *Lupinus nanus* Doug. ex Benth., which has zygomorphic flowers characteristic of the *Papilionoideae*, and *Cadia purpurea* (Picc.) Aiton, a species with atypical near-actinomorphic flowers.

This paper describes these primers, which amplify CYC-like genes, and their possible use in phylogenetic studies within the core genistoid clade. This large clade, defined from analyses of molecular data as comprising seven different tribes, some of which were previously thought to be unrelated, is reviewed in Wojciechowski (2003). Despite considerable work on members of this group, relationships between certain genera or between species which have undergone rapid diversification are still unclear (Wojciechowski, 2003). These primers should be useful for molecular systematic studies in this clade, which contains many large genera such as *Crotalaria* L. (c.600 species), *Aspalathus* L. (c.250 species), *Genista* L. (c.90 species) and economically important genera such as *Lupinus* (c.250 species), *Sophora* L. and *Ulex* L.

Materials and methods

Genomic DNA of *Cadia purpurea* and *Lupinus nanus* was extracted as described in Citerne *et al.* (2003). The open reading frames (ORF) of LEGCYC1A and LEGCYC1B in *C. purpurea* and *L. nanus* were isolated by genome walking following a modified protocol of Siebert *et al.* (1995) [GenBank accession numbers: *L. nanus* LEGCYC1A: AY382156; *L. nanus* LEGCYC1B: AY382155; *C. purpurea* LEGCYC1A: AY225826; *C. purpurea* LEGCYC1B: AY225825]. An array of primers, described here, was then designed based on these sequences specifically to amplify part of each locus (Table 1, Fig. 1). PCR amplifications were carried out on *L. nanus* and *C. purpurea* DNA using Bioline *Taq* and reagents (Bioline, London,

TABLE 1. Primers for amplification of LEGCYC1A and LEGCYC1B in genistoid legumes, with predicted annealing temperature in °C (T_{an}). Location of priming sites is given in Fig. 1

Primer name	Primer sequence (5′–3′)	T_{an}	Note
LEGCYC1-F1	CTT CTA CTT ACA YWT CYT CAG GC	58.9	General (LEGCYC1)
LEGCYC1-F2	CTT TCY TTA ACC CTG AAA ATG CTT C	58.9	General (LEGCYC1)
LEGCYC1-R1	CAC TCY TCC CAR GAY TTT CC	58.3	General (LEGCYC1)
LEGCYC1-R2	YAT TSG CAT CCC AAT TTG GAG	56.9	General (LEGCYC1)
LEGCYC1A-F1	CCA GAA GGG GTA GTR GTA G	57.7	Specific to LEGCYC1A
LEGCYC1A-R1	CTA CYA CTA CCC CTT CTG G	57.7	Specific to LEGCYC1A
LEGCYC1B-F1	CAC ARA AGG AAC CWG CTT G	55.6	Specific to LEGCYC1B
LEGCYC1B-R1	CAA GCS GGT TCC TTY TGT G	57.7	Specific to LEGCYC1B
LEGCYC1A-R2	GGT TTC TTW GYA AGA AAA TTG GAG	56.7	Specific to LEGCYC1A
LEGCYC1B-R2	AGC ARA CAA GAA AGS CCA TAG TG	59.8	Specific to LEGCYC1B

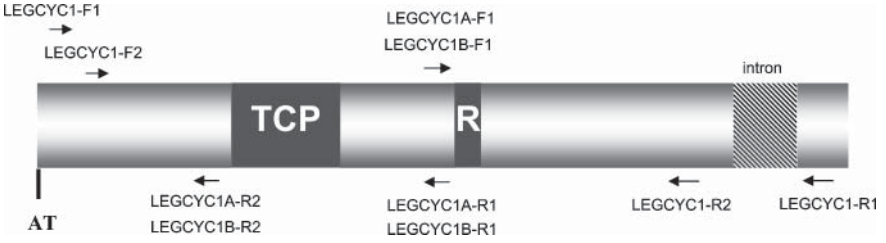

FIG. 1. Schematic representation of LEGCYC1 (A and B) showing the approximate location of primers described in Table 1. Black areas represent conserved domains (TCP and R) characteristic of CYC-like genes, and the hypervariable regions are paler. Forward and reverse locus-specific primers binding to a region between the TCP and R domains are complementary.

UK). The reaction mix contained 2.5μl of 50mM $MgCl_2$, 5μl of a 2mM dNTP mix, 10μM made up of 2.5μl of each primer (MWG Biotech, Ebersberg, Germany), 1 unit of BIOTAQ, 10–20ng of genomic DNA, and dH_2O to make up 50μl. Cycling conditions consisted of an initial denaturation step at 94°C (3 min), followed by 30 cycles of denaturation at 94°C (1 min), annealing at 50–55°C (30 s), and extension at 72°C (1 min), followed by a final extension step at 72°C (5 min), using a DNA Engine thermal cycler (MJ Research, Inc., Waltham, MA, USA). PCR products were purified using the QIAquick PCR Purification Kit (Qiagen Ltd, Dorking, Surrey, UK) and sequenced directly using the primers described here. Dye-terminator cycle sequencing was carried out using Thermosequenase II (Amersham Pharmacia, Buckinghamshire, UK). Samples were analysed on an ABI 377 Prism Automatic DNA Sequencer (Perkin Elmer, Applied Biosystems Division, Foster City, CA, USA).

To estimate the sequence divergence of LEGCYC genes within a large genus such as *Lupinus*, parts of the two loci were amplified and sequenced in a cultivar of another *Lupinus* species, *L. angustifolius* L. 'Merrit'. The genes were amplified in two parts using a combination of the general LEGCYC1-F1 and LEGCYC1-R1 primers and complementary locus-specific primers (LEGCYC1A-R1/LEGCYC1B-R1 and LEGCYC1A-F1/LEGCYC1B-F1) [GenBank accession numbers: DQ333950, LEGCYC1A 5′; DQ333951, LEGCYC1A 3′; DQ333952, LEGCYC1B 5′; DQ333953, LEGCYC1B 3′]. Genomic DNA of *L. angustifolius* 'Merrit' was provided by Susan Barker (University of Western Australia). Pairwise divergence of LEGCYC1A and LEGCYC1B between the two *Lupinus* species was compared with the nuclear ribosomal DNA ITS1 and ITS2 [GenBank accession numbers (submitted by Kaess in 1996): *L. nanus*: Z72176, ITS1; Z72177, ITS2; *L. angustifolius*: Z72202, ITS1; Z72203, ITS2]. All alignments were carried out manually and are available from the author.

In addition, the primers described here were tested in a range of taxa, including a putative sister taxon to *Cadia*, *Calpurnia aurea* (Aiton) Benth., which has typical zygomorphic papilionoid flowers, and in taxa belonging to the sister lineage of the

core genistoid clade: *Ormosia amazonica* Ducke, *Acosmium subelegans* (Mohl.) Yakovlev and *Bowdichia virgiloides* Kunth (Pennington *et al.*, 2001; Wojciechowski, 2003). *Calpurnia aurea* LEGCYC1A and LEGCYC1B were sequenced directly in two parts using primers LEGCYC1-F2 and LEGCYC1A-R1/LEGCYC1B-R1, and LEGCYC1A-F1/LEGCYC1B-F1 and LEGCYC1-R1 [GenBank accession numbers: DQ333954, LEGCYC1A 5′; DQ333955, LEGCYC1A 3′; DQ333956, LEGCYC1B 5′; DQ333957, LEGCYC1B 3′]. Genomic DNA of *C. aurea* was provided by Matt Lavin (Montana State University) and of *O. amazonica*, *A. subelegans* and *B. virgiloides* by Toby Pennington (Royal Botanic Garden Edinburgh).

Results and discussion

Potential of LEGCYC genes for phylogenetics

Pairwise sequence comparison of the coding regions of LEGCYC homologues in *L. nanus* and *L. angustifolius* revealed that the two copies are diverging rapidly: 8.49% nucleotide sequence divergence between LEGCYC1A orthologues and 5.34% nucleotide sequence divergence between LEGCYC1B orthologues. Sequence comparisons including total number of substitutions are summarized in Table 2. Comparison of pairwise divergence from the nuclear ribosomal DNA ITS (3.70% sequence divergence between *L. nanus* and *L. angustifolius* ITS1 + ITS2; Table 2) suggests that LEGCYC1A and LEGCYC1B sequences can provide up to four times as many variable characters as ITS at this taxonomic level. Although the level of nucleotide substitutions suggests that LEGCYC1A is evolving more rapidly than LEGCYC1B, sequence alignment between *L. nanus* and *L. angustifolius* homologues required the insertion of more gaps between LEGCYC1B than LEGCYC1A orthologues (7 and 4 gaps, respectively; Table 2). Most of these are trinucleotide repeats between LEGCYC1B orthologues, whereas there is a large indel (30 bp) between LEGCYC1A orthologues. The short intron region at the 3′ end of the gene was alignable between the two *Lupinus* species, providing further characters potentially useful for phylogenetic analyses. The intron region of LEGCYC1A and LEGCYC1B had a similar number of substitutions (4 substitutions over 79 bp versus 3 substitutions over 92 bp, respectively), but only LEGCYC1A had indels (3 indels of 1–31 bp).

Sequence comparison between closely related genera of the genistoid clade such as *Cadia* and *Lupinus* suggested that the level of variation between LEGCYC orthologues is also phylogenetically informative at this taxonomic level. Nucleotide pairwise distances in *C. purpurea* and *L. nanus* were greater between LEGCYC1A than LEGCYC1B orthologues (17.53% versus 13.29% nucleotide sequence divergence, respectively; see Table 2). Proportionally four times more gaps (17 for LEGCYC1A and 30 for LEGCYC1B; Table 2) were required for alignment of *C. purpurea* and *L. nanus* sequence pairs than in the comparison of *Lupinus* species.

TABLE 2. Pairwise comparison of *Lupinus nanus*, *L. angustifolius*, *Cadia purpurea* and *Calpurnia aurea* LEGCYC1A and LEGCYC1B coding regions. ITS values are shown for comparison between *Lupinus* species. **C. aurea* sequences have some missing data

Parameter	*Lupinus nanus*/ *L. angustifolius*	*L. nanus*/ *Cadia purpurea*	*L. nanus*/ *Calpurnia aurea**
LEGCYC1A			
Aligned length	1058 bp	1085 bp	971 bp
Aligned length excluding gaps	1013 bp	941 bp	873 bp
Number of indels	4	17	14
Length of indels	3–30 bp	3–30 bp	3–30 bp
Number of substitutions	86	165	132
Sequence divergence	8.49%	17.53%	15.12%
LEGCYC1B			
Aligned length	1187 bp	1280 bp	1204 bp
Aligned length excluding gaps	1160 bp	1038 bp	940 bp
Number of indels	7	30	30
Length of indels	3–6 bp	3–45 bp	3–36 bp
Number of substitutions	62	138	113
Sequence divergence	5.34%	13.29%	12.02%
ITS			
Aligned length	460 bp	–	–
Aligned length excluding gaps	459 bp	–	–
Number of indels	1	–	–
Length of indels	1 bp	–	–
Number of substitutions	17	–	–
Sequence divergence	3.70%	–	–

Despite these numerous insertion–deletion events, sequence alignment was for the most part unambiguous (details are available from the author). That these results are unrelated to the difference in floral symmetry between *L. nanus* (zygomorphic) and *C. purpurea* (actinomorphic) is borne out by comparison between *L. nanus* LEGCYC sequences and their orthologues in the zygomorphic putative sister species of *Cadia purpurea*, *Calpurnia aurea*, providing similar values to those described above (15.12% and 12.02% nucleotide sequence divergence between LEGCYC1A and LEGCYC1B orthologues, respectively; Table 2).

To evaluate the taxonomic range of the primers described here, some primer combinations were tested on genomic DNA from taxa in the sister group of the core clade: *Ormosia amazonica*, *Acosmium subelegans* and *Bowdichia virgiloides*. A single band of the expected size was amplified, with some exceptions reflecting the rapid rate of evolution of these genes, suggesting that these primers may be particularly useful within the core genistoids and their sister group, but may be less effective in a wider range of papilionoid legumes. In particular, the general primers LEGCYC1-F2, LEGCYC1-R1 were found to be most reliable in *Ormosia*, *Acosmium* and *Bowdichia*. The high level of variation within LEGCYC copies does however imply

that, although some of these primers were found to work in taxa sister to the core genistoid clade, they may not work in certain lineages even within this group. Nevertheless, these primers are likely to be useful for phylogenetic analysis of many large genera, particularly those in the range of tribes containing *Lupinus* (*Genisteae* s.s.) and *Cadia*/*Calpurnia* (*Podalyriae*), as defined by current phylogenies (Wojciechowski, 2003). The phylogenetic potential of LEGCYC1A and LEGCYC1B has been confirmed since the submission of this paper by Ree *et al.* (2004) who successfully used the primers described here to generate a phylogeny of 15 *Lupinus* species. The phylogenetic utility of CYC genes in the genistoid clade mirrors their use as a variable phylogenetic marker in other families such as *Gesneriaceae* (Smith *et al.*, 2004; Wang *et al.*, 2004).

Other applications

Regulatory genes, such as transcription factors, are believed to play an important role in the control of phenotypic diversity that distinguishes species (reviewed in Cronk, 2001). The study of natural variation of regulatory genes within populations or between closely related species is fundamental for understanding the processes of selection acting on these genes (Purugganan, 2000). CYC-like genes are suspected of being involved in the control of floral symmetry in this clade of the *Leguminosae* (Citerne, unpublished). Although most members of the genistoid clade have typical zygomorphic papilionoid flowers, some genera such as *Cadia* and *Acosmium* have more open, actinomorphic (radially symmetric) flowers. This provides a good framework for studying developmental gene evolution in relation to changes in a related morphological trait, in this case floral symmetry.

In addition, these genes may be useful molecular markers at the population level. Single nucleotide polymorphisms (SNPs) in a CYC-like gene were found to be informative for studying population structure in *Ramonda myconi* (L.) Reichb. (*Gesneriaceae*) (Pico *et al.*, 2002). The high rate of evolution described here between *Lupinus* species suggests that SNPs or trinucleotide repeats between CYC alleles could be used as population markers in legumes.

Acknowledgements

I would like to thank Toby Pennington and Quentin Cronk for comments on the manuscript. Laboratory assistance was provided by Michelle Hollingsworth and Alex Clark. This work was part of a project funded by the Carnegie Trust for the Universities of Scotland.

References

Ainouche, A. K. & Bayer, R. J. (1999). Phylogenetic relationships in *Lupinus* (Fabaceae: Papilionoideae) based on internal transcribed spacer sequences (ITS) of nuclear ribosomal DNA. *Amer. J. Bot.* 86: 590–607.

ALVAREZ, I. & WENDEL, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Mol. Phylogenet. Evol.* 29: 417–434.
BAILEY, C. D., CARR, T. G., HARRIS, S. A. & HUGHES, C. E. (2003). Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. *Mol. Phylogenet. Evol.* 29: 435–455.
CITERNE, H. L., LUO, D., PENNINGTON, R. T., COEN, E. & CRONK, Q. C. B. (2003). A phylogenomic investigation of *CYCLOIDEA*-like TCP genes in the Leguminosae. *Plant Physiol.* 131: 1042–1053.
CRONK, Q. C. B. (2001). Plant evolution and development in a post-genomic context. *Nat. Rev. Genet.* 2: 607–619.
DOYLE, J. J. (1992). Gene trees and species trees: molecular systematics as one-character taxonomy. *Syst. Bot.* 17: 144–163.
GÜBITZ, T., CALDWELL, A. & HUDSON, A. (2003). Rapid molecular evolution of CYCLOIDEA-like genes in *Antirrhinum* and its relatives. *Mol. Biol. Evol.* 20: 1537–1544.
HERSHKOVITZ, M. A., ZIMMER, E. A. & HAHN, W. J. (1999). Ribosomal DNA sequences and angiosperm systematics. In: HOLLINGSWORTH, P. M., BATEMAN, R. M. & GORNALL, R. J. (eds) *Molecular Systematics and Plant Evolution*, pp. 268–326. London: Taylor & Francis.
LUO, D., CARPENTER, R., VINCENT, C., COPSEY, L. & COEN, E. (1996). Origin of floral asymmetry in *Antirrhinum*. *Nature* 383: 794–799.
MORT, M. E. & CRAWFORD, D. J. (2004). The continuing search: low-copy nuclear sequences for lower-level plant molecular phylogenetics studies. *Taxon* 53: 257–261.
PENNINGTON, R. T., LAVIN, M., IRELAND, H., KLITGAARD, B., PRESTON, J. & HU, J-M. (2001). Phylogenetic relationships of basal papilionoid legumes based upon sequences of the chloroplast *trnL* intron. *Syst. Bot.* 26: 537–556.
PICO, F. X., MÖLLER, M., OUBORG, N. J. & CRONK, Q. C. B. (2002). Single nucleotide polymorphisms in the coding region of the developmental gene Gcyc in natural populations of the relict *Ramonda myconi* (Gesneriaceae). *Plant Biol.* 4: 625–629.
PURUGGANAN, M. D. (2000). The molecular population genetics of regulatory genes. *Mol. Ecol.* 9: 1451–1461.
REE, R. H., CITERNE, H. L., LAVIN, M. & CRONK, Q. C. B. (2004). Heterologous selection on LEGCYC paralogs in relation to flower morphology and the phylogeny of *Lupinus* (Leguminosae). *Mol. Biol. Evol.* 21: 321–331.
SIEBERT, P. D., CHENCHICK, A., KELLOGG, D. E., LUKYANOV, K. A. & LUKYANOV, S. A. (1995). An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* 23: 1087–1088.
SMITH, J. F., DRAPER, S. B., HILEMAN, L. C. & BAUM, D. A. (2004). A phylogenetic analysis within tribes Gloxinieae and Gesnerieae (Gesnerioideae: Gesneriaceae). *Syst. Bot.* 29: 947–958.
WANG, C-N., MÖLLER, M. & CRONK, Q. C. B. (2004). Phylogenetic position of *Titanotrichum oldhamii* (Gesneriaceae) inferred from four different gene regions. *Syst. Bot.* 29: 407–418.
WOJCIECHOWSKI, M. F. (2003). Reconstructing the phylogeny of legumes (Leguminosae): an early 21st century perspective. In: KLITGAARD, B. B. & BRUNEAU, A. (eds) *Advances in Legume Systematics, part 10, Higher Level Systematics*, pp. 5–35. Royal Botanic Gardens, Kew.

Received 31 March 2004; accepted after moderate revision 27 December 2005