Available online at www.sciencedirect.com
ScienceDirect
Plant Science 172 (2007) 441–451
PLANT SCIENCE
www.elsevier.com/locate/plantsci

Isolation and characterization of microsatellite markers for analysis of molecular variation in the medicinal plant Madagascar periwinkle (*Catharanthus roseus* (L.) G. Don)

Bhumika Shokeen, Niroj Kumar Sethy, Sushil Kumar, Sabhyata Bhatia *

National Center for Plant Genome Research, Aruna Asaf Ali Marg, Post Box No. 10531, Jawaharlal Nehru University Campus, New Delhi 110067, India

Received 23 July 2006; received in revised form 17 October 2006; accepted 17 October 2006
Available online 27 November 2006

Abstract

Catharanthus roseus (L.) G. Don, commonly known as Madagascar periwinkle, is a plant of great medicinal value. In this study microsatellite markers were developed and utilized for analysis of genetic diversity in *C. roseus*. In order to isolate microsatellite sequences, two small insert genomic libraries of *C. roseus* cv. Nirmal were constructed and screened with CA, CT, GC and GCG oligonucleotide repeats. Sixty-five microsatellite motifs were identified, from which 38 functional STMS primer pairs were designed and validated. Out of these, 24 STMS markers were used to evaluate the genetic polymorphism in 37 genotypes, which comprised of 32 accessions of *C. roseus*, a single accession each of two related species (*C. trichophyllus* and *C. pusillus*) and one accession each of three related genera (*Vinca minor*, *Thevetia peruviana* and *Nerium indicum*). The 24 STMS markers detected 26 loci with two markers amplifying more than one locus. A total of 124 alleles were amplified in the 37 genotypes ranging from 2 to 10 alleles with an average of 4.76 alleles per locus. The high average expected heterozygosity (H_E) value of 0.56 and observed heterozygosity (H_O) value of 0.52 established the efficiency of the STMS markers for discriminating the *C. roseus* genotypes. Nei and Li's similarity coefficients were calculated and a UPGMA-based dendrogram was constructed which clearly distinguished all genotypes except two pairs. Sequence analysis of the length variant alleles at three STMS loci revealed that the variation in the copy number of repeat motifs was the major source of length polymorphism within *C. roseus*. However, isolated point mutations and indels in the microsatellite flanking regions (MFRs) of homologus loci from other species also contributed to size homoplasy and allelic size variation. This study is the first report of microsatellite development and utilization in *C. roseus*, providing significant insights into its genome structure and organization.

Keywords: *Catharanthus roseus*; Microsatellites; Genetic diversity; Madagascar periwinkle; STMS

1. Introduction

Catharanthus roseus (L.) G. Don (Madagascar periwinkle) of the Apocynaceae family is widely grown both for its ornamental and more recently for its medicinal importance. The plant possesses an unsurpassed spectrum of chemodiversity in the form of 130 mono-terpenoid indole alkaloids including the antileukemic vincristine and vinblastine, and antihypertensive ajmalicine and serpentine [1]. However, the major factor limiting the extensive use of these alkaloids for medicinal applications is their extremely low yields *in vivo* coupled with the difficulties in their isolation and purification. Different conventional and biotechnological approaches are being used for increasing the pharmaceutically important alkaloids [2]. In this direction major success has been achieved in dissecting the highly complex alkaloid biosynthetic pathway. A number of enzymes along with several genes have been isolated and are available for construction of transgenic cells and plants with modified alkaloid profiles [3]. Large-scale efforts are on for the metabolic engineering of the alkaloid biosynthetic pathway. Hence, due to the large research efforts targeted towards achieving better yields of the pharmaceutically important alkaloids, *C. roseus* has become an important model medicinal plant system for biotechnology and secondary metabolism studies [4].

The genus *Catharanthus* consists of eight species, of which seven are endemic to Madagascar and one species (*C. pusillus*) is endemic to India [5]. *C. roseus* ($2n = 2x = 16$) is the most

* Corresponding author. Tel.: +91 11 26735159; fax: +91 11 26716658.
E-mail address: sabhyatab@yahoo.com (S. Bhatia).

0168-9452/$ – see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.plantsci.2006.10.010

important species in the genus with a genome size of 1500 Mbp (van Iren, unpublished). It is an ever blooming, perennial, tropical plant with reports of frequent outcrossing [6,7]. Even though many aspects of the alkaloid biosynthesis have been investigated, attempts need to be made to identify high alkaloid yielding accessions out of the vast resources of naturally occurring *C. roseus* germplasm. The genetic resources remain largely uncharacterized and unexploited and so does the structure and organization of the *C. roseus* genome [1]. Therefore, in order to use the available germplasm effectively, a thorough understanding of its genetic structure and characterization of genetic variability is required. This may best be done by utilization of DNA-based molecular markers for genome analysis, germplasm characterization and elucidation of genetic relationships. Among the available molecular markers, simple sequence repeats (SSRs), also known as microsatellites, are the most promising for genomic applications [8]. Microsatellites consist of short (1–6 bp) tandemly repeated DNA sequences [9], dispersed randomly and ubiquitously throughout the genome [10]. They represent hypervariable regions of the genome that arise due to replication slippage or unequal crossing over resulting in differences in the copy number of the repeat motifs [11]. However the regions flanking the microsatellites are conserved and can be used to design locus specific STMS (Sequence Tagged Microsatellite Site) markers. These markers are co-dominant, hypervariable, reproducible, having high discriminatory power and therefore suitable for a wide range of applications in genetic mapping, fingerprinting, germplasm characterization and marker assisted breeding [8,12,13]. Their validation as effective markers for genetic diversity analysis is demonstrated by their widespread use in different plant systems such as barley [14], wheat [15,16], maize [17], soybean [18], rice [19], tomato [20], *Arabidopsis* [21], *Cannabis* [22], sorghum [23] and chickpea [24].

This is one of the first structural genomics investigation conducted in *C. roseus.* Microsatellite-based STMS markers were developed from genomic libraries and used for assessing genetic diversity in a collection of *C. roseus* germplasm with the long-term objective of constructing a genetic linkage map of *C. roseus*. The utility of these markers in cross-species and cross-genera transferability studies was investigated. The hypervariability occurring in microsatellite motifs within the genus was also analyzed at the sequence level.

2. Materials and methods

2.1. Germplasm and DNA extraction

In this study a set of 37 *C. roseus* and related genotypes were analyzed which included 32 accessions of *C. roseus* and one accession each of *C. trichophyllus*, *C. pusillus*, *Vinca minor*, *Thevetia peruviana* and *Nerium indicum* (Table 1). DNA extraction using several protocols was attempted from leaf tissue, however, high polysaccharide content in the leaf hindered efforts to obtain high quality DNA. Eventually the best quality DNA was obtained by utilizing the protocol described by Khanuja et al. [25] which uses high salt. For each accession of *C. roseus*, genomic DNA was extracted by pooling leaves of 5 individual plants. Briefly, 3 g of leaf tissue was ground in liquid nitrogen, suspended in 3 ml of extraction buffer [100 mM Tris–Cl (pH 8.0), 25 mM EDTA, 1.5 M NaCl, 2.5% CTAB, 0.2% ß-mercaptoethanol (v/v fresh) and 1% PVP] and incubated at 60 °C for 1 h. Following an extraction with chloroform–isoamylalcohol (24:1), the aqueous layer was collected in a separate tube and DNA was precipitated by adding 1.5 ml of 5 M NaCl and 0.6 volumes of isopropanol. The DNA was pelleted down at 10,000 rpm for 15 min, washed with 80% ethanol, dried and dissolved in 500 μl of high salt T.E (1 M NaCl, 10 mM Tris–Cl, 1 mM EDTA) buffer. This was treated with RNase A (final concentration of 10 μg/ml) by incubating at 37 °C for 30 min. The enzyme was removed using an equal amount of chloroform–isoamylalcohol (24:1) and the DNA was precipitated by adding two volumes of ice-cold

Table 1
List of *C. roseus* and related genotypes analyzed using STMS markers

S. no.	Genotype (name)	Source/location
1	CrN1 (Nirmal)	Uttar Pradesh, India [68]
2	CrN2 (*lli/lli* mutant of CrN1)	Uttar Pradesh, India [69]
3	CrN35 (Prabal)	Uttar Pradesh, India [70]
4	CrN57	Uttar Pradesh, India
5	CrN13	New Delhi, India
6	CrN8	New Delhi, India
7	CrN48	New Delhi, India
8	CrN47	New Delhi, India
9	CrN76	Uttaranchal, India
10	CrN77	Uttaranchal, India
11	CrN80	Uttar Pradesh, India
12	CrN83	Karnataka, India
13	CrN43	Orissa, India
14	EC 120837	NBPGR, India
15	IC 49560	NBPGR, India
16	IC 49581	NBPGR, India
17	IC 210607	NBPGR, India
18	IC 49595	NBPGR, India
19	CrN78 (Vinca ice brocade)	IAHS, India
20	CrN38 (Pacifica light pink)	Namdhari Seed Co.
21	CrN39 (Apricot eye)	Namdhari Seed Co.
2	CrN40 (Pacifica blush)	Namdhari Seed Co.
23	CrN42 (Pacifica burgundy)	Namdhari Seed Co.
24	CrN62 (Pacifica mix)	Namdhari Seed Co.
25	CrN65 (Pacifica coral)	Namdhari Seed Co.
26	CrN72	Sri Lanka
27	CrN73	Mozambique
28	CrN74	Mozambique
29	CrN81	RBG, Kew, U.K.
30	CrN82	RBG, Kew, U.K
31	CrN52 (Grape cooler)	Sweden Seed Co.
32	CrN55 (Pretty mix)	Sweden Seed Co.

S. no.	Related genotypes	Source/location
33	CtN58 (*C. trichophyllus*)	[3]
34	CpN59 (*C. pusillus*)	[3]
35	A 4 (*Vinca minor*)	Uttranchal, India
36	A 5 (*Thevetia peruviana*)	New Delhi, India
37	A 6 (*Nerium indicum*)	New Delhi, India

The source/location of the accessions is mentioned (NBPGR-National Bureau of Plant Genetic Resources, IAHS-Indo-American Hybrid Seeds, RBG-Royal Botanic Gardens).