Conservation Genet Resour (2015) 7:53–56
DOI 10.1007/s12686-014-0358-6

TECHNICAL NOTE

Endonuclease V digestion for SNP discovery and marker development in South African white rhinoceros (*Ceratotherium simum*)

Christiaan Labuschagne · Antoinette Kotzé · J. Paul Grobler · Desiré L. Dalton

Received: 7 August 2014 / Accepted: 14 October 2014 / Published online: 22 October 2014

Abstract Single nucleotide polymorphism (SNP) markers are a promising new tool that can be used to study evolutionary processes, population genetic parameters, forensic cases and parentage. However, application of SNP marker analysis to wildlife has been limited, due to the lack of available sequence data in non-model organisms. Here, we describe a simple, rapid and cost effective method to isolate candidate SNPs in non-model organisms using the commercially available Endonuclease V enzyme. In a first application of this method, this SNP isolation strategy resulted in the identification of 12 new SNPs for white rhinoceros (*Ceratotherium simum*). This species has low reported genetic variability and has suffered severe bottlenecks over the last 150 years. Developed SNP markers in white rhinoceros could be used to define the genetic mating system of this species, for forensic applications and to determine population structure and variability when other markers prove problematic.

C. Labuschagne · A. Kotzé · J. Paul Grobler · D. L. Dalton (✉)
Department of Genetics, University of the Free State,
P.O. Box 339, Bloemfontein 9300, South Africa
e-mail: desire@nzg.ac.za

C. Labuschagne
e-mail: Christiaan.Labuschagne@inqababiotec.co.za

C. Labuschagne
Inqaba Biotechnical Industries (Pty) Ltd,
P.O. Box 14356, Hatfield 0028, South Africa

A. Kotzé · D. L. Dalton
National Zoological Gardens of South Africa,
P.O. Box 754, Pretoria 0001, South Africa

Keywords Single nucleotide polymorphism · SNPs · White rhinoceros · *Ceratotherium simum* · Endonuclease V · Non-model organism

Introduction

The white rhinoceros is a species that is affected the most at the hand of illegal trade in rhinoceros horn (Florescu et al. 2003). Highly informative molecular markers are important tools in successful management (Abadia-Cardoso et al. 2011). Single-strand specific nucleases such as CEL I, have been proposed as a simple and rapid method to assay mutations and Single nucleotide polymorphism (SNPs) (Rungis et al. 2005). The enzyme works with a variety of co-factors to digest heteroduplex DNA immediately 3′ of a mismatch site (Oleykowski et al. 1998). CEL I has been widely used in reverse genetics in plants and animals as well as disease diagnostics in human cancers (Colbert et al. 2001; Coghill et al. 2002; Perry et al. 2003; Comai et al. 2004). Xu et al. (2009) applied CEL I in a method to isolate fragments containing SNPs from background DNA in the half-smooth tongue sole (*Cynoglossus semilaevis*). However, CEL I needs to be isolated from celery stems, a time consuming procedure. The aim of the present study was to evaluate the use of commercially available Endonuclease V for the isolation of SNP containing fragments from white rhinoceros DNA.

Materials and methods

Blood samples were collected from 26 rhinoceros. DNA was extracted using the ZR Genomic DNA™-Tissue MiniPrep kit following the manufacturer's instructions.

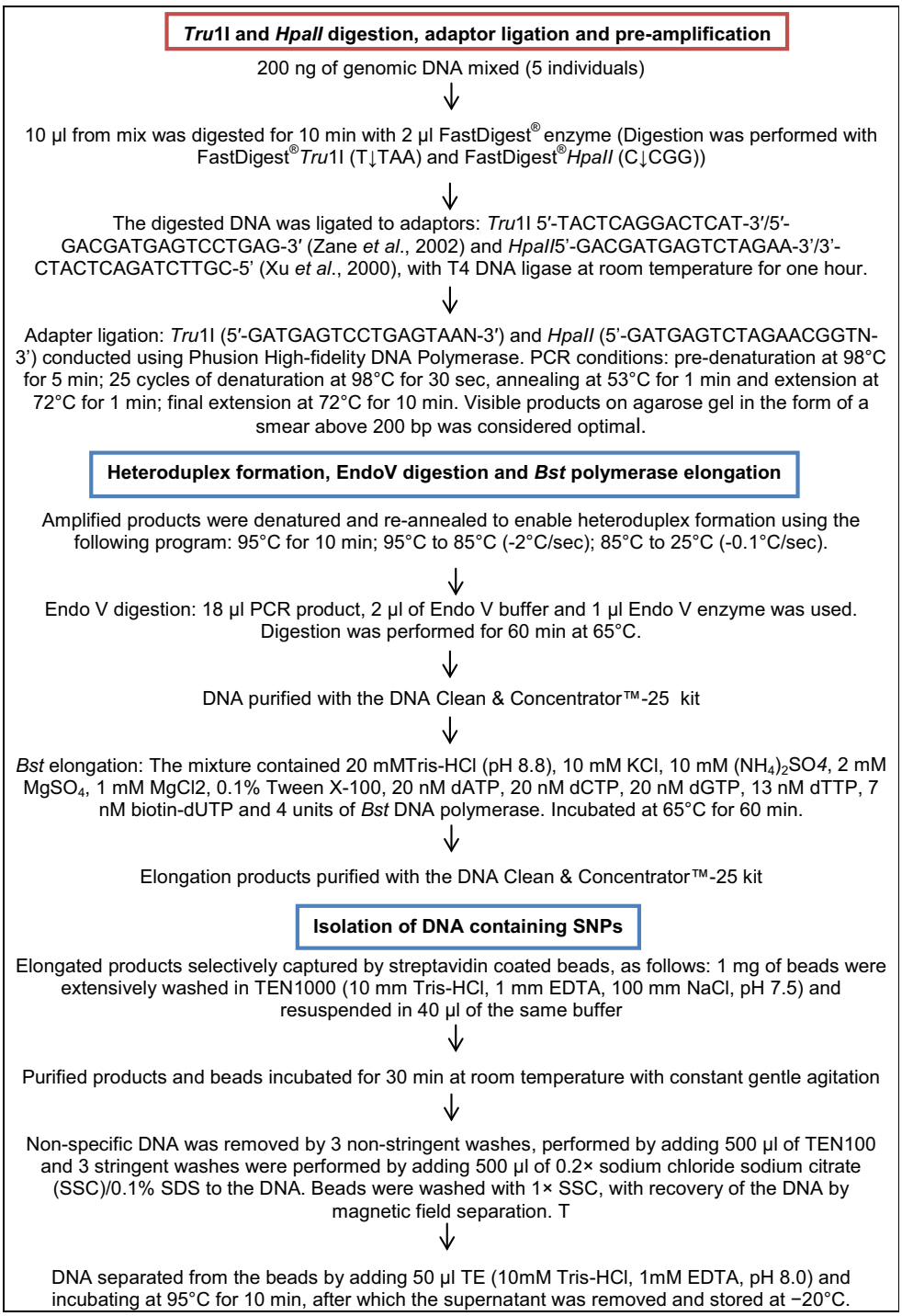

Fig. 1 Method used to isolate DNA containing SNPs

Blood samples were first washed by mixing 100 µl blood with 1,000 µl nuclease free water followed by centrifugation at 4,000 rpm for 2 min to reduce the number of red blood cells and improve DNA yields. The resulting pellet was further extracted using the above mentioned kit. The method used to isolate DNA containing SNPs was performed as shown in Fig. 1. Thereafter, Cloning and Sequencing was performed. The isolated DNA was amplified with either *Tru*1I or *HpaII* primers as described above. Subsequent amplicons were cloned into pJET using the CloneJET PCR Cloning Kit (Thermo Scientific) and Z-Competent™ JM109 *E. Coli* cells (Zymo Research). Ten clones containing fragments ranging from 300 to 800 bp were selected from the libraries. Cloned fragments were purified utilizing the Zymoclean™ Gel DNA Recovery Kit (Zymo Research) and sequenced utilising a

Table 1 Characterization of 12 SNPs in White Rhinoceros (*Ceratotherium simum*): GenBank accession numbers are 825690401-825690412

Locus	SNP name	Sequence length (bp)	Primer sequences (5′-3′)	Minor allele frequency	Heterozygosity		HWE
					He	Ho	
Hpa-1	Hpa-1-K	605	F-GGGATCATTCATTCATTCAGCTG	0.260	0.385	0.280	0.173 (ns)
			R-GGAACTCCAGAAGCCACG				
Hpa-10	Hpa-2-W	449	F-CCTTGTGTGGATTAAATGAGC	0.385	0.473	0.462	0.899 (ns)
			R-CTCAGCGGGTGGTTTCTC				
Tru-1	Tru-1-K	380	F-GAGAGCTTTCTCTCCTGAT	0.058	0.109	0.115	0.755 (ns)
			R-GAACTGGAAGTGTGTCAAC				
Tru-2	Tru-2-1-S	345	F-CCAGCATGGCTAGCATGC	0.423	0.488	0.462	0.781 (ns)
	Tru-2-2-R		R-CAGCCCTATCCGTGACTTTC	0.442	0.493	0.500	0.945 (ns)
	Tru-2-3-Y			0.077	0.142	0.154	0.671 (ns)
	Tru-2-4-Y			0.442	0.493	0.500	0.945 (ns)
	Tru-2-5-R			0.442	0.493	0.500	0.945 (ns)
	Tru-2-6-M			0.077	0.142	0.154	0.671 (ns)
Tru-3	Tru-3-R	335	F-GGCTCTGTTTGCTTGTCTG	0.250	0.375	0.346	0.695 (ns)
			R-CTTAGTGCTAGATTCTGCATG				
Tru-4	Tru-4-K	362	F-GTAGAACCTTCATCTCTGC	0.231	0.355	0.462	0.126 (ns)
			R-GCAGCTGCATTATATCCAC				
Tru-5	Tru-5-W	193	F-CTTGTGCTATTCTTCACTGTC	0.280	0.403	0.480	0.341 (ns)
			R-CAAGACGTCCACTGCAC				

F forward primer, *R* reverse primer, *bp* base pairs, *He* expected heterozygosity, *Ho* observed heterozygosity, *HWE* Hardy-Weinburg equilibrium, *ns* non-significant

Big Dye V3.1 Terminator Kit and an ABI 3500XL genetic analyser.

SNP validation: Primers were designed according to the obtained sequences. The potential SNP loci were amplified in the 5 isolates used for the initial DNA pool. Amplification reactions were done in a final volume of 25 µl containing 30 ng DNA, 25 pM of each primer and 2X DreamTaq® Green Master Mix using a standard PCR protocol. Resulting amplicons were inspected on 1 % agarose gels followed by purification and sequencing as described above. Sequences were inspected and aligned in CLC Bio Genomics work bench 5.0. GENEPOP version 4.0.10 (Raymond and Rousset 1995) was used to test for deviations from expected Hardy–Weinberg proportions, to evaluate loci for gametic disequilibrium and to determine allelic richness. Differences in mean observed heterozygosity, mean expected heterozygosity and mean number of alleles was determined using GenAlEx6 (Peakall and Smouse 2006).

Results

DNA from five rhinoceros was used for SNP discovery. DNA enriched for mismatches through EndoV digestion of *Tru*1I and *HpaII* AFLP libraries were cloned and sequenced. Ten clones were selected for each enzyme. Six out of the 10 *HpaII* clones contained a total of 13 polymorphic sites and were typed in a further 10 isolates. Four of the clones were heterozygous at all polymorphic sites for all individuals typed. These polymorphisms were artefacts from paralogous sequence differences. The remaining two clones both possessed one polymorphic site that presented as both homozygous and heterozygous. These polymorphisms were considered to be true SNPs and were typed through sequencing in the remaining individuals. Nine out of the 10 *Tru*1I clones contained a total of 30 polymorphic sites and were typed in a further 10 isolates. Four of the clones were heterozygous in all and were considered to be artefacts. The remaining 5 clones all showed polymorphic sites that presented as both homozygous and heterozygous alleles in the 10 individuals tested. These were considered to be true SNPs and were typed through sequencing in the remaining individuals. The 12 SNPs, primer sequences and allele frequencies for the 26 individuals are listed in Table 1. SNPs Tru2-1, Tru2-2, Tru2-4 and Tru2-5 are in Linkage disequilibrium ($P = 0.0000$), while Tru2-3 and Tru2-6 are in also linked ($P = 0.0000$). None of the loci deviated significantly from Hardy–Weinberg equilibrium.

Discussion

The preparation of native enzymes from plant material usually involves several steps of purification (Mon et al. 2012). Adopting the SNP enrichment method proposed by Xu et al. (2009), but replacing CEL I with commercially available Endonuclease V, we isolated 12 new SNPs in the white rhinoceros. Endo V is active towards basic sites and urea sites, base pair mismatches, flap and pseudo Y structures, and small insertions/deletions in DNA molecules with the cleavage site generated at the second phosphodiester bond 3′ to a lesion (Huang et al. 2001). Endo V recognised the mismatches formed in the heteroduplex library between different isolates creating a nick, which was subsequently recognised by *Bst* DNA polymerase and extended through strand displacement incorporating biotin-dUTP. The DNA strands containing biotin-dUTPs could then be captured with streptavidin-coated beads and separated from background DNA. Xu et al. (2009) reported that 9 out of 10 fragments contained SNPs through CEL I digestion. In the present study only 2 fragments from the *HpaII* library and 5 fragments from the Tru1I library contained true SNPs. Since paralogous regions form mismatches in the heteroduplex in the same way as a polymorphism between isolates would, the *HpaII* library resulted in 6 out of 10 enriched fragments while the Tru1I library resulted in 9 out of 10 enriched fragments, confirming that EndoV can be used to substitute CEL I. The apparent high proportion of paralogs may be a result of the low genetic diversity observed in white rhinoceros.

References

Abadia-Cardoso A, Clemento AJ, Garza JC (2011) Discovery and characterization of single-nucleotide polymorphisms in steelhead/rainbow trout, *Oncorhynchus mykiss*. Mol Ecol Resour 11(supp. 1):31–49

Butler JM, Coble MD, Vollone PM (2007) STRs vs. SNPs: thoughts on the future of forensic DNA testing. Forensic Sci Med Pathol 3:200–205

Coghill EL, Hugill A, Parkinson N, Davison C, Glenister P, Clements S, Hunter J, Cox RD, Brown SD (2002) A gene-driven approach to the identification of ENU mutants in the mouse. Nat Genet 30:255–256

Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai L, Henikoff S (2001) High-throughput screening for induced point mutations. Plant Physiol 126:480–484

Comai L, Young K, Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR, Henikoff S (2004) Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. Plant J 37:778–786

Florescu A, Davila JA, Scott C, Fernando P, Kellners K, Morales JC, Melnick D, Boag PT, van de Coeverden Groot P (2003) Polymorphic microsatellites in white rhinoceros. Mol Ecol Notes 3:344–345

Huang J, Lu J, Barany F, Cao W (2001) Multiple cleavage activities of endonuclease V from *Thermotoga maritima*: recognition and strand nicking mechanism. Biochemistry 40:8738–8748

Mon H, Lee J, Fukushima M, Nagata Y, Fujii M, Xu J, Nishi O, Iiyama K, Kusakabe T (2012) Production and characterization of the celery mismatch endonuclease CEL II using baculovirus/silkworm expression system. Appl Microbiol Biotechnol 97(15):6813–6822

Oleykowski CA, Bronson Mullins CR, Godwin AK, Yeung AT (1998) Mutation detection using a novel plant endonuclease. Nucleic Acids Res 26:4597–4602

Peakall R, Smouse PE (2006) GENALEX 6: geneticanalysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Perry JA, Wang TL, Welham TJ, Gardner S, Pike JM, Yoshida S, Parniske M (2003) A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume Lotus japonicus. Plant Physiol 131:866–871

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Rungis D, Hamberger B, Berube Y, Wilkin J, Bohlmann J, Ritland K (2005) Efficient genetic mapping if single nucleotide polymorphisms based upon DNA mismatch digestion. Mol Breed 16:261–270

Xu ML, Li XQ, Korban SS (2000) AFLP-based detection of DNA methylation. Plant Mol Biol Report 18:361–368

Xu J-Y, Xu G-B, Chen S-L (2009) A new method for SNP discovery. Biotechniques 46:201–208

Zane L, Baegelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Mol Ecol 11:1–16