Development of polymorphic chloroplast DNA markers for the endangered tree Eusideroxylon zwageri…

Article *in* Conservation Genetics Resources · August 2015

DOI: 10.1007/s12686-015-0485-8

Conservation Genet Resour (2015) 7:845–850
DOI 10.1007/s12686-015-0485-8

METHODS AND RESOURCES ARTICLE

Development of polymorphic chloroplast DNA markers for the endangered tree *Eusideroxylon zwageri* through chloroplast isolation and next-generation sequencing

Hiroyuki Kurokochi[1] · I. L. G. Nurtjahjaningsih[2] · Sukartiningsih[3] · Engkong Tan[2] · Shuichi Asakawa[2] · Yoko Saito[2] · Yuji Ide[2]

Received: 5 August 2015 / Accepted: 12 August 2015 / Published online: 22 August 2015

Abstract Borneo ironwood (*Eusideroxylon zwageri*) is a tall tropical tree that is threatened by illegal logging and exploitation due to its economic value. In the present study, approximately 10,618 bp of *E. zwageri* chloroplast DNA were sequenced genome-wide and 16 polymorphic markers of chloroplast DNA were developed, using next-generation sequencing technology. In total, 26 nucleotide substitutions, 3 indels, 1 sequence inversion, and 1 mononucleotide repeat variation were detected in samples of 10 trees collected from 5 populations in Kalimantan, Indonesia. These markers are a powerful tool that may be used to describe the phylogeographical genetic structure of *E. zwageri* and will be essential for the conservation and management of this endangered tree.

Hiroyuki Kurokochi and I. L. G. Nurtjahjaningsih have contributed equally to this research.

Electronic supplementary material The online version of this article (doi:10.1007/s12686-015-0485-8) contains supplementary material, which is available to authorized users.

✉ Hiroyuki Kurokochi
kurokochi@anesc.u-tokyo.ac.jp; h.kurokochi@gmail.com

[1] Asian Natural Environmental Science Center, The University of Tokyo, 1-1-8 Midori-cho, Nishitokyo, Tokyo 188-0002, Japan

[2] Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1, Yayoi, Bunkyo, Tokyo 113-8657, Japan

[3] Faculty of Forestry, Mulawarman University, Campus Gunung Kelua, Jl. Ki Hajar Dewantara, Samarinda, East Kalimantan 75119, Indonesia

Keywords *Eusideroxylon zwageri* · Next-generation sequencing · Chloroplast DNA · Polymorphic markers · Phylogeography

Introduction

Borneo ironwood, *Eusideroxylon zwageri* (Lauraceae), is a tropical rainforest tree with a long lifespan that is distributed naturally in Brunei, Indonesia, Malaysia, and the Philippines. The wood has high specific gravity and is extremely hard, making it durable for decades or even a century (Whitten et al. 1987). The wood is economically valuable mainly for bridge construction and furniture manufacture. It uses insect-mediated pollination and gravity dispersal of seeds. Clearing of land for agriculture and overharvesting may threaten this species (Partomihardjo 1987). Thus, the conservation of *E. zwageri* is an urgent matter. In order to form an effective conservation strategy, it is essential to understand the underlying genetic structure of the species, because the spatial genetic structure of wild populations is often linked to specific local ecosystems.

Evaluation of the intraspecific mutations of chloroplast DNA is a powerful tool for describing the phylogeographical genetic structure. Chloroplast DNA is generally maternally inherited in angiosperms, thus chloroplast DNA markers record gene flow from seed movement only, whereas nuclear markers record the gene flow of both pollen and seeds. In addition, polymorphisms in haploid genomes are more affected by genetic drift than those from the nuclear genome (Petit et al. 1993). Patterns in the spatial distribution of chloroplast DNA polymorphism established by seed dispersal during range expansion should therefore be more slowly eroded by subsequent

gene flow, compared with spatial patterns in nuclear genetic markers, especially in trees (Petit and Hampe 2006).

However, until now, not enough markers have been available to detect intraspecific variations in chloroplast DNA in *E. zwageri*. In the present study, in order to understand the phylogeographical genetic structure of *E. zwageri*, we developed a sufficient number of chloroplast DNA markers, including intraspecific mutations.

Materials and methods

Preparation of reference chloroplast DNA sequences for *E. zwageri* with a next-generation sequencer

In order to increase the ratio of chloroplast DNA, chloroplasts of *E. zwageri* leaves sampled from Kutai in Indonesia (N0°33′54″, E117°35′27.6″) were isolated according to Sato et al. (1993). Chloroplast DNA was then extracted from the isolated chloroplasts using a DNeasy Blood & Tissue Kit (Qiagen). Fragmented DNA was produced from approximately 50 ng of chloroplast DNA with Ion Shear Plus Reagents. After the adapter was ligated to the fragmented DNA with the Ion Plus Fragment Library Kit, approximately 550 bp of DNA fragments were collected using the E-GelR SizeSelect™ 2 % Agarose Gel and iBase™ unit (Kurokochi et al. 2014; Kurokochi et al. 2015).

The collected DNA was amplified by PCR with an Ion Plus Fragment Library Kit, under the following PCR conditions: a 5-min hold at 95 °C, 12 cycles of 95 °C for 15 s, 58 °C for 15 s, and 70 °C for 1 min, then a final hold at 4 °C. After the concentration of the amplified DNA was determined with Agilent 2200 TapeStation (Agilent Technologies), 1 μL of the amplified DNA was diluted to 2 pM and sequenced using an IonPGM sequencer. All sequencing data were assembled and contigs were generated using a CLC Genomics Workbench (CLC Bio). The contigs with more than 2000 bp were checked to determine whether they were derived from chloroplast DNA, using the BLAST (National Center for Biotechnology Information, NCBI). All contigs judged to be related to chloroplast DNA were treated as reference chloroplast DNA sequences for *E. zwageri*.

Design and characterization of chloroplast DNA markers for *E. zwageri*

Chloroplast DNA markers were designed from the developed reference chloroplast DNA sequences using Primer3Plus (Untergasser et al. 2007). In order to detect chloroplast DNA mutations among individuals, a total of ten *E. zwageri* trees were selected, two each from five geographically separated populations: Kapuas Hulu (N1°5′16.8″, E111°57′7.2″), Berau (N0°35′6″, E117°11′52.8″), Gunung Palung (N0°56′42″, E109°59′9.6″), Kutai (N0°33′54″, E117°35′27.6″), and Kintap (S3°42′10.8″, E115°9′18″). One leaf was collected from each of the selected trees and dried in a separate plastic bag containing silica gel. DNA was extracted from the dried leaves using the CTAB method (Kurokochi et al. 2013).

Table 1 Basic information on contigs with more than 2000 bp after assembly. (Color table online)

Contig	Average coverage[a]	Length of contig (bp)	Accession number (sort by "Max Score" in NCBI BLAST)[b]				
Contig I	84.2	32,767	AJ428413.1	KF753638.1	KJ408574.1	JN867582.1	JN867580.1
Contig II	153.6	22,793	DQ899947.1	JN867577.1	JN227740.1	KF753638.1	KJ408574.1
Contig III	82.4	18,444	DQ899947.1	JN867579.1	KJ408574.1	HM775382.1	JN867582.1
Contig IV	82.0	15,469	DQ899947.1	KJ408574.1	JN867582.1	HM775382.1	KF753638.1
Contig V	81.1	10,413	DQ899947.1	AJ428413.1	JN867584.1	HM775382.1	JN867582.1
Contig VI	84.0	9931	AJ428413.1	JN867587.1	JN867584.1	HM775382.1	KF753638.1
Contig VII	73.5	9452	AJ428413.1	JN227740.1	KF753638.1	JN867577.1	HM775382.1
Contig VIII	82.6	8011	KJ408574.1	JN867580.1	JN867581.1	JN867582.1	DQ899947.1
Contig IX	23.9	3483	AF389256.1	AF479244.1	DQ008624.1	AY095466.1	AY095462.1
Contig X	10.6	3045	AF506028.1	AM487639.2	AM482469.2	AM425773.2	AM459848.2
Contig XI	4.3	2891	AP004894.1	HG975451.1	HG975521.1	CP002684.1	CP002688.1
Contig XII	78.5	2430	DQ899947.1	AJ428413.1	KF753638.1	HM775382.1	KJ408574.1
Contig XIII	4.6	2112	Not hit				
Contig XIV	3.6	2021	Not hit				

[a] Average coverage was calculated on the basis of the reads that were obtained from the next-generation sequencer

[b] Yellow, green, blue, and red indicate that the Query cover in the NCBI BLAST had more than 90, 50–90, 10–50, and fewer than 10 hits, respectively. "Not hit" indicates that the E value in the NCBI BLAST was larger than e^{-10}

Table 2 The 19 primers developed for chloroplast DNA of *Eusideroxylon zwageri*

Locus	Primer F (5′–3′)	Primer R (5′–3′)	Contig[a]	Expected length (bp)[b]	Compared length (bp)[c]	Number of mutation sites
Ez_Cp_01	TCTAGGAATCAACAACACGAAAA	TCCGTTGCTTGTGTGGATAA	Contig I	575	530	2
Ez_Cp_02	AGTAATCCACGCCACTCTCA	GTCGCTGTCAGGGCAACTA	Contig I	586	488	2
Ez_Cp_03	TCGAATCATATCCAATGGTCA	GCAGTGCCGGGATATCTTAT	Contig I	568	482	1
Ez_Cp_04	TTCGGAGATTGGATGCAGTT	GCCCCACCCATGAGTAAATA	Contig II	594	498	1
Ez_Cp_05	TCCCTCTATCCCCAATAAAAAG	AATGAAGTATTAACATTCTTTTTGAGC	Contig III	692	648	3
Ez_Cp_06	CGGAAATTAGAAATGCCAAAA	TCGAGATGGAAATCCCTAGAAA	Contig III	594	497	3
Ez_Cp_07	TTTCAAGAGTTCGCTGAGCTT	CTATCCTTGGGTTGGACGAA	Contig III	569	490	1
Ez_Cp_08	CGGAAAGGAATTGGGAAAA	CACATGTCGTACCATATTCCAC	Contig III	573	501	3
Ez_Cp_09	TGATTTCGTTCTCTCGCTATCA	GGGTTTACTTTTTCAACCATAGG	Contig IV	548	467	1
Ez_Cp_10	GCGGGCTGTTCAAGATCATA	TCGCTTGGATTTACGAAAGG	Contig V	599	515	2
Ez_Cp_11	TTTTGCAGCTTTCGTTGTTG	TAGCACCATGCCAAATGTGT	Contig V	677	573	0
Ez_Cp_12	CCACGGGTTGCTTACTGAAT	GGACAGTAACGAGGGGTCAA	Contig V	983	845	0
Ez_Cp_13	CGGGTGCCTCATAAAAAGAA	GTGCCACCAGATGTGCATAC	Contig V	789	675	1
Ez_Cp_14	TGGAGAAGACAAAATTGTTTGAA	GACCGATTTACCCTCTTTTGC	Contig VIII	595	505	2
Ez_Cp_15	TGCCAGAAGTTGACAAGGTG	TTCCCTAGAAGAAAAAGAAATAGCA	Contig VIII	579	497	2
Ez_Cp_16	CTCTTCGGGACCGAATATCA	AGATGCTCTTGGCTCGACAT	Contig VIII	782	687	4
Ez_Cp_17	GATTTATGATATGCTTCCTGCTGA	TCAAAGCCACCTCTTCTGCT	Contig XII	497	471	0
Ez_Cp_18	TTCACGCATGTCCAGTTGAT	CCCACGATCCTCTGTTGATT	Contig XII	800	703	1
Ez_Cp_19	AGCAGAAGAGGTGGCTTTGA	ATGATGGGATCGTAGCTTGG	Contig XII	648	546	1

[a] The contigs shown are the same as those in Table 1

[b] Expected length is the length from Primer F to Primer R, based on each of the assembled contigs

[c] Compared length is the length that was actually read using the ABI 3130 capillary sequencer in this study. The two lengths were different because the direct sequencing technique was used, a method that sometimes makes it difficult to detect both edges of the sequencing regions. The sequencing data on each compared length are described in Supporting Information

Table 3 Mutation sites and types detected in the present study using five populations of *Eusideroxylon zwageri*

Contig	Contig I					Contig II	Contig III									
Locus	Ez_Cp_01		Ez_Cp_02		Ez_Cp_03	Ez_Cp_04	Ez_Cp_05			Ez_Cp_06			Ez_Cp_07	Ez_Cp_08		
Location	63	264	61–68	257	353	87	56	164–166	560	114–119	268	368	468	53	300	357
Kapuas Hulu	C	A	AAAGGGTT	A	G	A	A	ACA	G	–	T	T	C	A	A	A
Berau	T	C	AAAGGGTT	G	G	A	A	GCC	A	–	T	G	A	T	G	A
Gunung Palung	C	A	AAAGGGTT	A	G	A	A	ACA	G	–	T	T	A	A	A	A
Kutai	T	C	–	G	T	C	A	GCA	G	ACTGTA	A	G	A	T	G	A
Kintap	T	C	–	G	G	A	G	GCA	G	ACTGTA	T	G	C	T	G	C

Contig	Contig IV	Contig V			Contig VIII								Contig XII	
Region	Ez_Cp_09	Ez_Cp_10		Ez_Cp_13	Ez_Cp_14		Ez_Cp_15			Ez_Cp_16			Ez_Cp_18	Ez_Cp_19
Location	171–175	45	374	482	45	137	28	431	20	161–170	256	411	45	69–70
Kapuas Hulu	AGTAG	T	G	C	G	G	G	A	G	A10	C	A	A	TA
Berau	AGTAG	T	A	C	T	G	T	A	A	A9	–	G	G	C
Gunung Palung	CTACT	T	G	C	T	A	G	A	G	A10	C	A	A	C
Kutai	AGTAG	T	G	A	T	G	T	A	A	A9	–	G	G	C
Kintap	CTACT	G	A	C	T	G	T	C	A	A9	–	G	G	C

All the compared sequences of each locus are described in Supporting Information

Each region was amplified by PCR using designed markers. Each reaction mixture contained 1.0 μL of 5× Colorless GoTaq® Reaction Buffer' (Promega), 0.4 μL of MgCl2 (25 mM), 2.775 μL of RNase-free water, 0.1 μL of dNTPs (2.5 mM each), 0.1 μL of primers (forward and reverse, both 5 μM), 0.025 μL of GoTaq (5 U/μL), and 0.5 μL of the extracted DNA. The total volume of each mixture was 5.0 μL. The reaction program had the following thermal profile: 5 min initial denaturation at 95 °C, followed by 40 cycles of 30 s denaturation at 95 °C, annealing for 30 s at 57 °C, and extension for 60 s at 72 °C, followed by a final extension at 72 °C for 7 min. PCR products were purified using an ExoSAP-IT Kit (Affymetrix). The conditions imposed were 40 min of enzymatic treatment at 37 °C, followed by 20 min of inactivation at 80 °C.

Sequencing reactions were performed on the PCR products using a Big Dye Terminator 3.1 Cycle Sequencing Kit (Applied Biosystems). The thermal profile was as follows: 3 min of preheating at 95 °C, followed by 25 cycles each consisting of 10 s denaturation at 95 °C, 5 s annealing at 50 °C, and 60 s extension at 60 °C. Products were purified by adding a mixture of 1.5 μL acetic acid (99 %), 7.5 μL water, and 31.25 μL ethanol (95 %), and allowing the sequencing products to stand for 15 min at room temperature. The mixture was then centrifuged at 1500*g* for 30 min and the supernatant was decanted and mixed with an additional 1.5 μL acetic acid, 7.5 μL DDW, and 75 μL ethanol. After centrifugation for 15 min at 1500*g*, the supernatant was decanted and all ethanol was removed by evaporation. Purified products were separated on an ABI3100 Genetic Analyzer (Applied Biosystems).

To assemble sequence data, the program BioEdit Sequence Alignment Editor (Hall 2004) was used. Sequences were aligned using the software package MEGA5 (Tamura et al. 2007). Then, the chloroplast DNA mutations were determined.

Results and discussion

In total, 200,445 reads (54,470,111 bp) were obtained from the Ion PGM sequencer. After assembly, 3275 contigs (1,193,362 bp) were detected, 14 of which consisted of more than 2000 bp. Those 14 contigs were denoted Contigs I–XIV (Table 1) and included a total 143,262 bp ranging from 2021 bp (Contig XIV) to 32,767 bp (Contig I), with a mean length of 10,233 bp. The mean coverage of Contigs I to XIV in the assembly process was 60.6, ranging from 3.6 (Contig XIV) to 153.6 (Contig II) (Table 1). The results of NCBI BLAST indicated that nine Contigs (I, II, III, IV, V, VI, VII, VIII, and XII) were the most likely to contain chloroplast DNA (Table 1). Contig IX was likely to be a part of the 26S ribosomal RNA gene with a moderate query cover in the NCBI BLAST, but four other Contigs (X, XI, XIII, and XIV) were difficult to identify due to the low query cover (Table 1).

As for the nine Contigs (I, II, III, IV, V, VI, VII, VIII, and XII) containing chloroplast DNA, the coverage of eight Contigs (I, III, IV, V, VI, VII, VIII, and XII) was approximately 80 in the assembly process, but that of Contig II was about double that of the others (153.6) (Table 1). Chloroplast DNA is well known to have a genome size of approximately 100–150 kbp and an inverted repeat sequence (Shinozaki et al. 1986; Hiratsuka et al. 1989; Maier et al. 1995). Thus, Contig II might be the inverted repeat sequence of chloroplast DNA in *E. zwageri*. If this were the case, the total length of the nine Contigs would be 152,503 bp, indicating that we have detected almost all of the chloroplast DNA in *E. zwageri*.

We designed a total of 27 markers in the nine Contigs related to chloroplast DNA. Of these markers, 19 (Ez_Cp_01 to Ez_Cp_19) were well amplified and succeeded in sequencing 3, 1, 4, 1, 4, 3, and 3 markers for Contigs I, II, III, IV, V, VIII, and XII, respectively (Table 2). Using those 19 markers, approximately 10,618 bp of chloroplast DNA was sequenced for each sample (Table 2). All except three of the markers (Ez_Cp_11, Ez_Cp_12, and Ez_Cp_17) included chloroplast DNA mutations among the populations (Table 3). In total, 26 nucleotide substitutions, 3 indels, 1 sequence inversion, and 1 mononucleotide repeat variation were detected. On the basis of these mutations, five haplotypes were detected from ten samples, but there was no variation within populations.

In the present study, we successfully developed chloroplast DNA markers that included intraspecific mutations. These markers will be a powerful tool for describing the phylogeographical genetic structure of *E. zwageri* and will be essential for the conservation and management of this endangered tree.

Acknowledgments This study was supported in part by a Grant-in-Aid for Scientific Research (B) from the JSPS to YI (No. 25304025).

References

Hall T (2004) BioEdit version 7.0.0. http://www.mbio.ncsu.edu/bioedit/biodoc.pdf

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY et al (1989) The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Mol Genet Genom 217(2–3):185–194

Kurokochi H, Saito Y, Chuman M, Ide Y (2013) Low chloroplast diversity despite phylogenetically divergent haplotypes in

Japanese populations of *Ailanthus altissima* (Simaroubaceae). Botany 91(3):148–154

Kurokochi H, Tan E, Asakawa S, Sukartiningsih, Saito Y, Ide Y (2014) Development of 16 microsatellite markers in *Eusideroxylon zwageri* by next-generation sequencing. Conserv Genet Resour 6:593–595

Kurokochi H, Tan E, Asakawa S, Lian C (2015) Development of 18 microsatellite markers in *Pieris japonica*, a poisonous tree insulated from the browsing pressure of herbivores, using a next-generation sequencer. J For Res 20:244–247

Maier RM, Neckermann K, Igloi GL, Kössel H (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. J Mol Biol 25:614–628

Partomihardjo T (1987) The ulin wood which is threatened to extinction. Duta Rimba 13(87–88):10–15

Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. Annu Rev Ecol Evol Syst 37:187–214

Petit RJ, Kremer A, Wagner DB (1993) Geographic structure of chloroplast DNA polymorphisms in European oaks. Theor Appl Genet 87:122–128

Sato N, Albrieux C, Joyard J, Douce R, Kuroiwa T (1993) Detection and characterization of a plastid envelope DNA-binding protein which may anchor plastid nucleoids. EMBO J 12(2):555–561

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5(9):2043–2049

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24(8):1596–1599

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JA (2007) Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res 35:W71–W74

Whitten AJ, Damanik SJ, Anwar J, Hisyam N (1987) The ecology of Sumatra, 2nd edn. Gadjah Mada University Press, Yogyakarta