Agr. Nat. Resour. 54 (2020) 55–63

AGRICULTURE AND NATURAL RESOURCES

Journal homepage: http://anres.kasetsart.org

Research article

Nucleotide sequencing of specific genes in nuclear and plastid genomes of good performance hybrids derived from interspecific hybridization in the genus *Jatropha*

Nutchanun Sakthanutpat[a], Chatuporn Kuleung[a], Vipa Hongtrakul[a.b.c,*]

[a] *Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand.*

[b] *Center of Excellence for Jatropha, Kasetsart University, Bangkok 10900, Thailand.*

[c] *Center for Advanced Studies in Tropical Natural Resources, NRU-KU, Kasetsart University, Bangkok 10900, Thailand.*

Article Info

Article history:
Received 4 August 2018
Revised 8 October 2018
Accepted 19 October 2018
Available online 28 February 2020

Keywords:
DNA sequencing,
Interspecific hybridization,
Jatropha curcas

Abstract

Biodiesel, produced from plant oil, has emerged as an attractive renewable energy source. Physic nut (*Jatropha curcas* L.) has gained widespread attention from both biodiesel producers and the research community due to its advantages of high biodiesel production, rapid growth, tolerance and convenient propagation. The current study used interspecific hybridization between *J. curcas* and *Jatropha* species and good performance hybrids DNA to study the relationship. Standardized DNA in the organism was used for DNA sequencing of the good performance hybrids as well as other *Jatropha* species. Sequencing was performed at four gene regions of the *rbc*L and *trn*H–*psb*A intergenic spacers in the plastid genome and the internal transcribed spacer of *rDNA* and *stearoyl-acyl carrier protein desaturase* (*SAD*) in the nuclear genome. Alignment of the *SAD* gene sequence showed high polymorphism among the five *Jatropha* species and hybrids. Phylogenetic trees based on the nucleotide sequences of regions could show that *J. integerrima* was closely related to *J. curcas* better than either gene.

Introduction

Biodiesel is a fast-developing alternative fuel in many countries around the world and is produced from biological processes using oil extracted from various organisms such as algae, fungi and plants (Nigam and Singh, 2011). Biofuel production from vegetable oils has been suggested as one of the possible ways to reduce greenhouse gas emissions that cause global warming which has a negative environmental effect (Nigam and Singh, 2011). *Jatropha curcas* L is one of the most popular of the *Jatropha* species for biodiesel production; however, usage of biodiesel from *J. curcas* is still relatively limited, largely due to the lack of genetic diversity within the species that can breed for agronomic traits, such as increased yield and tolerance to stresses (Asseleih et al., 1989). Hence, genetic improvement using interspecific hybridization can be applied to broaden its genetic variation and to transfer desirable traits from one species to another (Prabakaran and Sujatha, 1999). Five species of genus *Jatropha*, belonging to the family Euphorbiaceae, are commonly found in Thailand: *J. curcas* L., *J. multifida* L., *J. podagrica* Hook, *J. gossypifolia* L. and *J. integerrima* Jacq. And each species has special traits (Asseleih et al., 1989). *J. curcas* or physic nut is a multipurpose semi-evergreen large shrub or small tree, which is widely distributed in the wild or semi-cultivated areas in Central and South America, Africa, India and South East Asia (Asseleih et al., 1989; Asif et al., 2010). *J. curcas* has gained attention because of its high oil content (30–50% overall, 45–60% in the kernel), hardiness, easy propagation, drought tolerance, rapid growth and adaptation to wide agro-climate conditions (Jones and Miller, 1992). *J. multifida*

* Corresponding author.
E-mail address: fscivph@ku.ac.th (V. Hongtrakul)

https://doi.org/10.34044/j.anres.2020.54.1.08

is a shrub 2–3 m high that is smaller than *J. curcas*, having big fruit, non-dehiscent capsules and a high quantity of oil (Chopra et al., 1965). It originated in the region from Mexico to Brazil (Soontornchainaksaeng and Jenjittikul, 2003. *J. podagrica* is a multipurpose evergreen shrub that is good at adapting to drought and is commonly found in Africa, Asia and Latin America (Aiyelaagbe et al., 2007). *J. gossypifolia* is a bushy shrub about 1.8 m high, with abundant fruit, easy propagation, a short gestation period and drought tolerance (*Oduola et al.,* 2005). *J. integerrima* is one of the most important ornamental species and is claimed to connect the genera *Jatropha* and *Ricinus* (Sujatha, 1996; Sujatha et al., 2005). Seeds of *J. integerrima* have a high quantity of linoleic acid (Rao and Lakshminarayana, 1987). The *Jatropha* genome is approximately 416 Mbp, and chromosome numbers are 22 chromosomes with a base number of x = 11 (Soontornchainaksaeng and Jenjittikul, 2003). *J. curcas* is a cross-pollinated crop that can cross with closely related species to produce *Jatropha* hybrids with good performance (Parthiban et al., 2009).

Specific gene sequencing is a form of molecular technology for identification that uses short standardized DNA within the genome regions to identify plant species (Chen et al., 2014). The standardized DNA regions, used for nucleotide sequencing, should be about 500–800 bp in length, capable of amplification using universal primers and have sufficient nucleotide polymorphism to distinguish among different species. In the current study, DNA sequencing at four gene regions was performed: the *ribulose-1,5-bisphosphate carboxylase* large subunit (*rbc*L) and *trn*H–*psb*A intergenic spacer in the plastid genome and the internal transcribed spacer (ITS) of *rDNA* and *stearoyl-acyl carrier protein* (*ACP*) *desaturase* (*SAD*) in the nuclear genome. The *rbc*L subunit is suitable for inference of phylogenetic relationships at higher taxonomic levels because of its slow rate of evolution compared to nuclear genes, but it has a fast rate of evolution in annual angiosperms (Wolfe et al., 1987; Chase et al., 1993). The *trn*H–*psb*A gene consists of *trn*H codes for $tRNA^{His}$ (GUG) and *psb*A codes for the 32 kD herbicide-binding protein component of the thylakoid membrane (UImasov et al., 1990) which has been successfully used to identify species (Kress et al., 2005; Shaw et al., 2007; Newmaster et al., 2008; Yao et al., 2010). This *trn*H–*psb*A intergenic spacer region can be easily amplified using universal primers, and insertions/deletions (indels) are common and lead to various lengths among different plant groups. Therefore, the *trn*H–*psb*A intergenic spacer region has a high percentage of nucleotide differences and inversions (Kress et al., 2005; Shaw et al., 2007). The eukaryotic *ribosomal DNA* (*rDNA*) gene consists of a small-subunit (18S), 5.8S, and a large-subunit (28S) that are separated by internal transcribed spacer (ITS) regions (Coleman and Mai, 1997). The sequence variation of ITS regions has led to phylogenetic studies of many different organisms (White et al., 1990). SAD is an important enzyme of fatty acid biosynthesis in higher plants (Tong et al., 2006), that catalyzes the desaturation of stearoyl-ACP to oleoyl-ACP. SAD plays a key role in determining the ratio of saturated fatty acids to unsaturated fatty acids in plants. Information from GenBank indicates that the *SAD* gene contains three exons and two introns. The current study investigated specific gene sequences of five *Jatropha* species and the good performance hybrids derived from interspecific hybridization and phylogenetic tree construction based on all gene sequences for specific identification and genetic relationship analysis among *Jatropha* species and hybrids.

Materials and Methods

Plant materials

Five *Jatropha* species (*J. curcas*, *J. multifida*, *J. podagrica*, *J. gossypifolia* and *J. integerrima*) were collected from five locations of Thailand—Ban Paew (BP), Samut Sakhon province; Pakchong (PC), Nakhon Ratchasima province; Klongsamwa (KS), Bangkok; Muang, Chachoengsao (CS) province; and Lamlukka (LK), Pathum Thani province. Four *J. curcas* samples were used: two toxic varieties—KUBP78-9 and Prayapichai (dwarf plant types) —and two non-toxic varieties (very low levels of toxic phorbal ester)—Suksan and M5. Three samples of each species of *J. multifida*, *J. podagrica* and *J. gossypifolia* were collected from three locations. Four *J. integerrima* samples were collected—four varieties of red flower and smooth edge narrow leaf (R-S), red flower and round leaf with a smooth edge (R-R), red flower and three-loped leaf (R-3L) and pink flower (P). Five good performance hybrids (KUBP201–202, KUBP209–211) were obtained from the *Jatropha* Improvement Project under the Kasetsart University Biodiesel Project (KUBP) using interspecific hybridization. All 22 samples are shown in Fig.1 and Table 1.

DNA extraction

Genomic DNA was extracted from leaves using the modified cetyl-trimethylammonium bromide (CTAB) procedure (Doyle and Doyle, 1990). The quality of genomic DNA was analyzed in 1% agarose gel and the concentration was measured using spectrophotometry. Total genomic DNA samples were stored at 4°C.

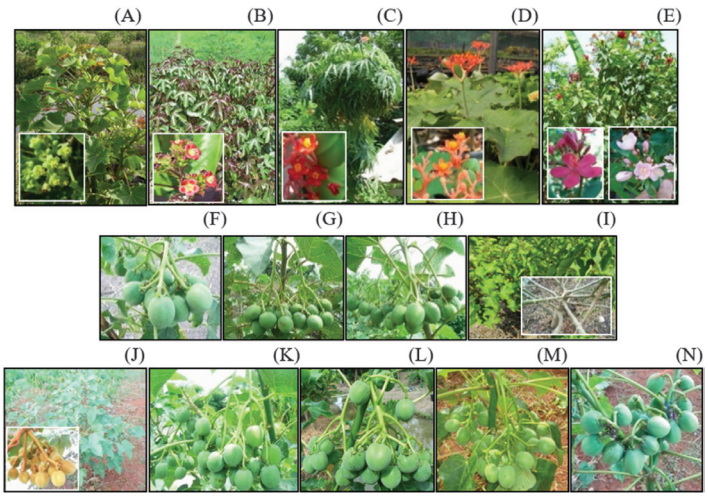

Fig. 1 *Jatropha* species: (A) *J. curcas*; (B) *J. gossypifolia*; (C) *J. multifida*; (D) *J. podagrica*; (E) *J. integerrima*, *J. curcas*: (F) KUBP78-9 (toxic); (G) Suksan (non-toxic); (H) M5 (non-toxic); (I) Prayapichai (toxic/dwarf); and good performance hybrids: (J) KUBP201; (K) KUBP202; (L) KUBP209; (M) KUBP210; (N) KUBP211.

Table 1 Characteristic of samples and sampling location

Sample	Characteristic	Pedigree	Location
J. curcas (JC)			
1. KUBP78-9	toxic and high yield seed yield 1,178.75 kg/ ha 100 seed weight 69.1 g 49.12% oil in kernel		BP
2. Prayapichai	toxic and dwarf plant type		BP
3. M5	high yield and non-toxic		BP
4. Suksan	high yield and non-toxic		BP
5.-7. *J. multifida*			BP/PC/KS
8.-10. *J. podagrica*			BP/PC/CS
11.-13. *J. gossypifolia*			BP/PC/LK
14.-17. *J. integerrima*	- red flower and smooth edge narrow leaf (R-S) - red flower and round leaf with a smooth edge (R-R) - red flower and three-loped leaf (R-3L) - pink flower (P)		BP
18. KUBP201	thick leaf and resistance 50.08% oil in kernel	JC (toxic154) x JI (R-S) ↓ $F_1$5 x ↓ 267 ($OPF_1$5) x ↓ KUBP201 (OP267)	BP
19. KUBP202	big fruit and high yield seed yield 310.3 kg/rai 100 seed weight 78.0 g 51.35% oil in kernel phorbol ester 0.64 mg/g	JC (non-toxic) x JI (R-S) ↓ $F_1$137 x JC (non-toxic) ↓ $BC_1$235 x ↓ KUBP202 ($OPBC_1$235)	BP
20. KUBP209	high yield seed yield 278.7 kg/rai 100 seed weight 60.5 g 54.20% oil in kernel	JC (non-toxic) x JI (R-S) ↓ $F_1$137 x JC (toxic-Nakhon Ratchasima) ↓ $BC_1$202 x ↓ 494 ($OPBC_1$202) x toxic-KUBP78-9 ↓ KUBP209	BP
21. KUBP210	high yield seed yield 252.3 kg/rai 100 seed weight 69.1 g 58.79% oil in kernel phorbol ester 1.39 mg/g	JC (non-toxic) x JI (R-S) ↓ $F_1$137 x JC (toxic-Nakhon Ratchasima) ↓ $BC_1$202 x JC (toxic-Mookdahan 10kr) ↓ $BC_2$561 x ↓ KUBP210 ($OPBC_2$561)	BP
22. KUBP211	high yield seed yield 202.2 kg/rai 100 seed weight 65.2 g 52.06% oil in kernel	JC (non-toxic) x JI (R-S) ↓ $F_1$137 x JC (toxic-Nakhon Ratchasima) ↓ $BC_1$202 x ↓ 313 ($OPBC_1$202) x toxic-KUBP78-9 ↓ KUBP211	BP

BP = Ban Paew; PK = Pak Chong; KS = Klongsamwa; CS = Chachoengsao; LK= Lam Luk Ka

Primer design

Primers used for the *rbc*L, *trn*H–*psb*A genes and for the ITS of *rDNA* were obtained from previous reports (Table 2). The *SAD* primers were designed from alignment of *J. curcas* complete genome and the *SAD* gene of *J. curcas* available in GenBank. The first SAD primer pair was designed at 5′ UTR before the SAD start codon and at 3′ UTR after the stop codon. After sequencing, all sequences from all *Jatropha* species were aligned and used to design the next primer pair for polymerase chain reaction (PCR) amplification of all *Jatropha* species. In total, six primer pairs were designed to obtain the complete SAD gene sequence (Table 2). All designed primers were checked for suitability using the Fast PCR 6.5.5 program (Helsinki, Finland).

PCR amplification

Each PCR amplification reaction was in a total volume of 25 µl containing 100 ng of genomic DNA, 5X Phusion buffer (rbcL, trnH–psbA and SAD = 5X Phusion HF buffer, ITS of rDNA = 5X Phusion GC buffer), 0.2 mM of dNTPs (Thermo Scientific; Waltham, MA, USA), 5 pmol of each primer and 2 U of Phusion DNA polymerase (Thermo Scientific; Waltham, MA, USA). The primer sequences are shown in Table 2. The thermal cycle consisted of an initial denaturation of 94°C for 1 min, followed by 35 cycles of 94°C for 30 s, 52.0–56.5°C for 45 s and 72°C for 1 min and then a final elongation of 72°C 10 min. The PCR products were stored at 4°C for analysis.

Table 2 Primer sequences for polymerase chain reaction amplification at four gene regions of *rbc*L, *trn*H–*psb*A, ITS of *rDNA* and *SAD* in *Jatropha* species and hybrids

Primer	Nucleotide sequence (5′ – 3′)	DNA size (~bp)	Annealing temperature (°C)	Reference
rbcLa	F: ATGTCACCACAAACA GAGACTAAAGC	600	56.5	Levin et al. (2003); Soltis et al. (1992);
	R: GTAAAATCAAGTCCACCRCG			Kress and Erickson (2007); Fofana et al. (1997)
trnH2 psbAF	F: CGCGCATGGTGGATT CACAATCC	600	56.5	Tate and Simpson (2003); Sang et al. (1997)
	R: GTTATGCATGAACGT AATGCTC			
ITS-F ITS4-R	F: TCGCTCCTACCGATTGAATG R: TCCTCCGCTTATTGATATGC	900	56.5	Songpanich and Hongtrakul (2010); White et al. (1990)
SAD1	F: TGTAGCTTAATTTTGATTCC R: CTAACAATAGACTAACGAC	769	56.5	
SAD2	F: TTTGTCAGTTATTCATGTCGG R: GGTATACACATTAATAGAGGG	1,051	52.0	
SAD3	F: TGATGGATGGGTTTAAATGC R: AGATACATCAACCATGTTCC	723	56.5	
SAD4	F: TGGCAGTTGGTTGAGTTTGG R: CTTCATGTCCACTCTTCCAG	755	56.5	
SAD5	F: GTTGTGCTGGTTGGAGATATG R: TCAATCTCAAAGAGCTTCTCC	1,105	56.5	
SAD6	F: GGCAACATTCATCTCCCATG R: GCACACCTATTTCGGTCTAAC	693	56.5	

ITS = internal transcribed spacer; *SAD = stearoyl-acyl carrier protein desaturase* gene.

DNA sequencing and analysis

The PCR products were purified using a FavorPrep™ Gel/PCR Purification Mini Kit (Favorgen; Ping-Tung, Taiwan) and were sent to 1st BASE (Seri Kembangan, Selangor, Malaysia) for sequencing using the forward primers in Table 2. Sequence comparisons of all samples were performed using the ClustalW alignment in biological sequence alignment editor (BioEdit) program (Hall, 1999) for cutting the early and end regions of the sequences. DNA sequence data were also analyzed using the maximum likelihood method and a phylogenetic tree was constructed using the MEGA 5.0 software (Tamura et al., 2011).

Results and Discussion

Nucleotide sequence analysis of rbcL

The *rbc*L sequence of 543 bp was amplified from *Jatropha* species and hybrids. The obtained sequences were submitted to GenBank (*J. curcas*, MH229901; *J. multifida*, MH229902; *J. podagrica*, MH229903; *J. gossypifolia*, MH229904; *J. integerrima*, MH229905). The *rbc*L sequences of *J. multifida*, *J. podagrica* and *J. gossypifolia* differed from the sequences of *J. curcas*, *J. integerrima* and hybrids by two base substitutions (one transition and one transversion) as shown in Table 3. All 22 sequences of *Jatropha* species and good performance hybrids were aligned, and a phylogenetic tree was constructed (Fig. 2). Phylogenetic analysis of *rbc*L sequences demonstrated the similarity between the *Jatropha* species and hybrids. The *rbc*L sequence was reported to be not variable enough in some plants for separating some species in the same genus (Lahaye et al., 2008). With hybridization between closely related species, a moderately conserved gene like *rbc*L is likely to have little or no variation (Chase et al., 1993).

Nucleotide sequence analysis of trnH–psbA intergenic spacer

The *trn*H–*psb*A intergenic spacer is the best plastid option for sequencing that has good priming sites, length and interspecific variation (Kress et al., 2005). The *trn*H–*psb*A intergenic spacer region is located between the *trn*H and *psb*A genes and has high polymorphism that is more suitable for species identification than *rbc*L (Kress et al., 2009). The *trn*H–*psb*A intergenic spacer sequences were submitted to GenBank (*J. curcas*, MH229906; *J. multifida*, MH229907; *J. podagrica*, MH229908; *J. gossypifolia*, MH229909; *J. integerrima* (R-S), MH229910; *J. integerrima* (R-R), MH229911; *J. integerrima* (R-3L), MH229912; *J. integerrima* (P), MH229913). All 22 sequences of *Jatropha* species and good performance hybrids were aligned, and a phylogenetic tree was constructed (Fig. 2). The number of nucleotides and nucleotide substitutions of *trn*H–*psb*A sequences are shown in Table 3 with 144 positions of indels found. The *trn*H–*psb*A intergenic spacer sequences could clearly classify the five *Jatropha* species. The phylogenetic tree based on *trn*H–*psb*A sequences demonstrated five groups of *Jatropha* species and separated *J. integerrima* into two groups (R-S/R-3L and R-R/P) but could not separate hybrids from *J. curcas* (Fig. 2).

Nucleotide sequence analysis of internal transcribed spacer of rDNA

The nuclear ribosomal ITS region is used extensively in phylogenetic analysis and can successfully discriminate species as it provides the highest resolving power for discriminating closely related species of Fungi (Schoch et al., 2012). It exists in multi-copies and exhibits high levels of within-species sequence differentiation and as such ITS1 was earlier suggested as a region for the identification of flowering plants (Kress et al., 2005). This gene inhabits the nucleus, so hybrids could receive all alleles from parents carrying more than one allele. The ITS1–ITS2 of *rDNA* sequences were submitted to GenBank (*J. curcas*, toxic KUBP78-9: MH229886; *J. curcas*, non-toxic M5-a1 & a2: MH229887–MH229888, *J. curcas*, non-toxic Suksan: MH229889; *J. multifida*: MH229890; *J. podagrica*: MH229891; *J. gossypifolia*: MH229892; *J. integerrima*: MH229893; KUBP202: MH229894–MH229895; KUBP210: MH229896; KUBP201–a1 & a2: MH229897–MH229898; KUBP209: MH229899; KUBP211: MH229900). All 22 sequences of the *Jatropha* species and good performance hybrids were aligned, and a phylogenetic tree was constructed (Fig. 2). Two alleles (a1& a2) were found in non-toxic *J. curcas* M5 and resistant KUBP201 (Fig. 3). The ITS of *rDNA* could be used to distinguish toxic *J. curcas* from non-toxic *J. curcas*. Following the alignment of ITS sequences of all 22 Jatropha species and hybrids, 22 positions of indels were observed. The number of nucleotide polymorphisms is shown in Table 3. The phylogenic tree based on ITS of *rDNA* sequences differentiated each species of *Jatropha* into five groups and within *J. curcas* into two groups (Fig. 2).

Nucleotide sequence analysis of stearoyl-acyl carrier protein desaturase

The stearoyl-ACP desaturase (SAD) enzyme converts stearoyl (18:0)-ACP to oleoyl (18:1)-ACP in the unsaturated fatty acid synthesis of plants and plays a role in the regulation of fatty acid composition of seed oil plants (Harwood, 1988).The *SAD* gene has high divergence especially at the intron region and can be used to study the genetic relationship (Table 4). The *SAD* gene consists of three exons and two introns. Six specific primer pairs were developed and used to PCR amplify the *SAD* gene sequence from the start to stop codons. The length of each partial gene sequence was variable in each *Jatropha* species. The CAP3 program was used to combine the six overlapping gene sequences. The complete *SAD* sequences were submitted to GenBank (*J. curcas*, non-toxic M5: MH234457; *J. curcas*, non-toxic Suksan: MH234458; KUBP211: MH234459; *J. multifida*: MH234460; *J. podagrica*: MH234461; *J. gossypifolia*: MH234462; *J. integerrima* (R-S): MH234463; *J. integerrima* (R-R): MH234464; *J. integerrima* (R-3L): MH234465; *J. integerrima* (P): MH234466, toxic KUBP78-9: MH250055; *J. curcas*, toxic Prayapichai: MH250056; KUBP209: MH250057; KUBP202-a1 & a2: MH250058–MH250061; KUBP210: MH250062; KUBP201-a1 & a2: MH250063–MH250064). The number of polymorphisms in the *SAD* gene sequence is shown in Table 3 and Table 4. Two alleles were found in the hybrids KUBP202 and KUBP201. High variation in the number of nucleotides in intron 1 and intron 2, (shown as indels and nucleotide substitutions) was found more in exon 1 than in exon 2. The phylogenetic tree based on the complete *SAD* gene sequences could classify species of the genus *Jatropha*, toxic and non-toxic *J. curcas*, hybrids and the four varieties of *J. integerrima* (Fig. 2).

Table 3 Number of nucleotides and nucleotide substitutions of four gene regions in *Jatropha* species and hybrids

Sample	Number of nucleotides (bp)				Number of nucleotide substitutions** (bp)			
	*rbc*L	*trn*H-*psb*A	ITS*	*SAD*	*rbc*L	*trn*H–*psb*A	ITS*	*SAD*
J. curcas								
KUBP78-9 (toxic)	543	660	802	3,872	-	-	-	-
Prayapichai (toxic)	543	660	802	3,872	-	-	-	-
M5 (non-toxic)	543	660	802	3,872	-	-	1	13
Suksan (non-toxic)	543	660	802	3,872	-	-	1	13
J. multifida	543	668	779	3,796	2	49	154	189
J. podagrica	543	668	794	3,805	2	50	48	189
J. gossypifolia	543	686	803	3,846	2	101	33	191
J. integerrima								
R-S	543	718	801	3,877	-	12	27	109
R-R	543	718	801	3,859	-	27	27	121
R-3L	543	718	801	3,876	-	12	27	106
P	543	718	801	3,876	-	27	27	105
Hybrids								
KUBP202	543	660	802	3,872	-	-	-	151(a1)/1(a2)
KUBP210	543	660	802	3,872	-	-	1	1
KUBP209	543	660	802	3,872	-	-	-	1
KUBP211	543	660	802	3,872	-	-	-	1
KUBP201	543	660	802	3,872	-	-	1	151(a1)/1(a2)

ITS = internal transcribed spacer; *SAD = stearoyl-acyl carrier protein desaturase* gene.

* ITS of *rDNA* **number of nucleotide substitutions compared to KUBP78-9 (high yield *J. curcas* obtained from selection under the KU Biodiesel Project (KUBP) before the improvement project by interspecific hybridization).

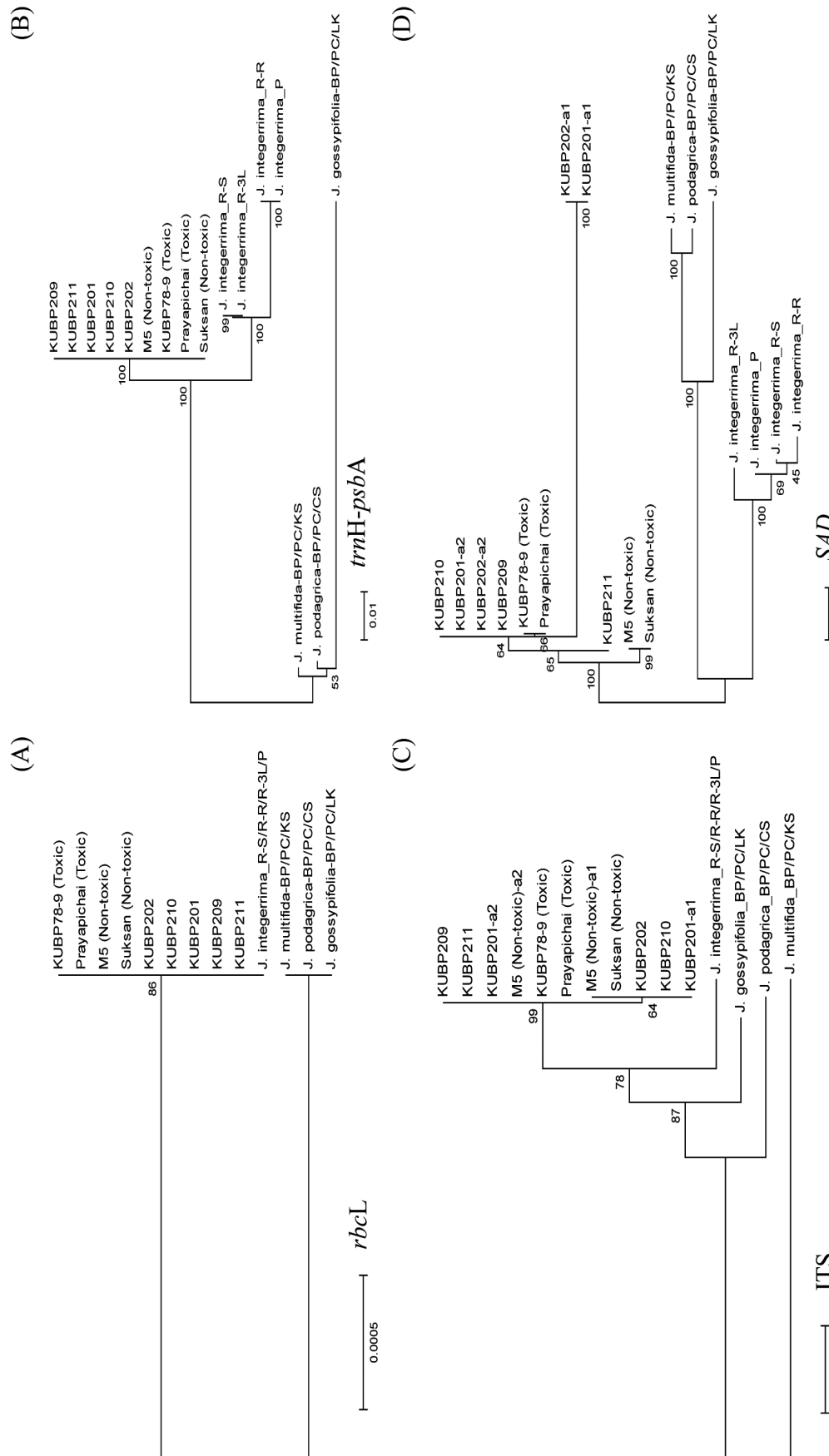

Fig. 2 Phylogenetic tree of 22 samples of *Jatropha* species and hybrids based on maximum-likelihood analysis of specific gene sequences: (A) *rbc*L; (B) *trn*H–*psb*A; (C) ITS of *rDNA*; (D) *SAD*. (a = allele), where ITS = internal transcribed spacer and *SAD* = *stearoyl-acyl carrier protein desaturase* gene

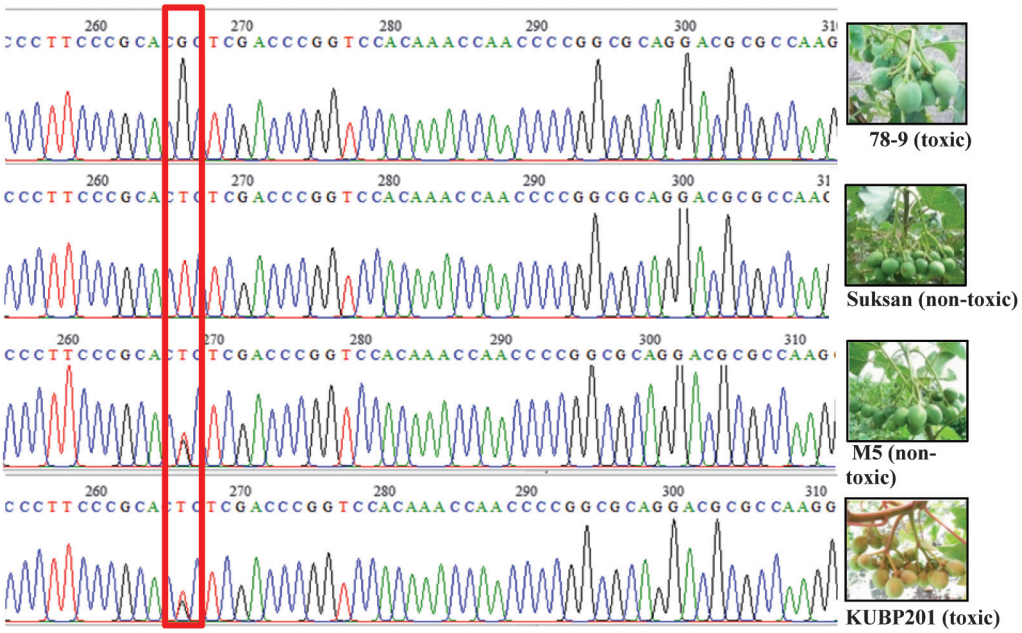

Fig. 3 Position 256 of internal transcribed spacer of *rDNA* indicating different nucleotide within *J. curcas*, where the red box indicates the two peaks (alleles) in M5 and KUBP201 compared to the one peak in 78-9 and Suksan.

Table 4 Number of nucleotides and nucleotide substitutions in *stearoyl-acyl carrier protein desaturase* gene regions of *Jatropha* species and hybrids

Sample	Number of nucleotides/nucleotide substitutions*					
	Exon1	Intron1	Exon2	Intron2	Exon3	Total
KUBP78-9 and Prayapichai (toxic)	122 / -	2,037 / -	505 / -	647 / -	561 / -	3,872 / -
M5 and Suksan (non-toxic)	122 / -	2,037 / 8	505 / -	647 / -	561 / 5	3,872 / 13
J. multifida	122 / -	2,048 / 137	505 / 12	560 / 27	561 / 13	3,796 / 189
J. podagrica	122 / -	2,031 / 135	505 / 10	586 / 31	561 / 13	3,805 / 189
J. gossypifolia	122 / -	2,013 / 150	505 / 8	645 / 21	561 / 12	3,846 / 191
J. integerrima (R-S)	122 / 4	2,040 / 67	505 / 4	649 / 22	561 / 12	3,877 / 109
J. integerrima (R-R)	122 / 4	2,022 / 77	505 / 4	649 / 25	561 / 11	3,859 / 121
J. integerrima (R-3L)	122 / 4	2,040 / 71	505 / 4	648 / 17	561 / 10	3,876 / 106
J. integerrima (P)	122 / 4	2,040 / 67	505 / 4	648 / 19	561 / 11	3,876 / 105
KUBP201/202-a1	122 / -	2,037 / 151	505 / -	647 / -	561/ -	3,872 / 151
KUBP201/202-a2	122 / -	2,037 / 1	505 / -	647 / -	561 / -	3,872 / 1
KUBP209/210/211	122 / -	2,037 / 1	505 / -	647 / -	561 / -	3,872 / 1

*Number of nucleotide substitutions compared to KUBP78-9.

Genetic relationship of genus Jatropha

The genetic relationships among the *Jatropha* species and good performance hybrids were developed based on the combined four gene regions of *rbc*L, the *trn*H–*psb*A intergenic spacer, ITS of *rDNA* and *SAD* using the maximum likelihood procedure. The phylogenetic tree could classify all the *Jatropha* species, toxic and non-toxic *J. curcas* and the selected hybrids better than using a single gene sequence information (Fig. 4).

Considering the four gene sequences studied, the ability to use gene sequences for species discrimination was best based on the *SAD* sequences, followed by the ITS of *rDNA*, *trn*H–*psb*A and *rbc*L, respectively. Both *rbc*L and *trn*H–*psb*A in the plastid genome performed well in terms of sequence quality but had low discriminatory power; whereas the ITS of *rDNA* and *SAD* in the nuclear genome exhibited good discriminatory power. The *rbc*L region had the lowest polymorphisms, with only two substitutions in 543 bp DNA sequences. The rates of substitution by codon position deviate slightly for chloroplast DNA gene sequences (Ritland and Clegg, 1987). The *trn*H–*psb*A demonstrated good amplification, a high rate of polymorphism and multiple indels and high levels of species discrimination that were capable of classifying all 5 *Jatropha* species (Fig. 2). Among plastid regions, *trn*H–*psb*A is an efficient characterized gene due to substantial species-level genetic variability and divergence (Blaxter and Floyd, 2003). The *rbc*L and *trn*H–*psb*A genes are in the chloroplast genome that receives genetic material

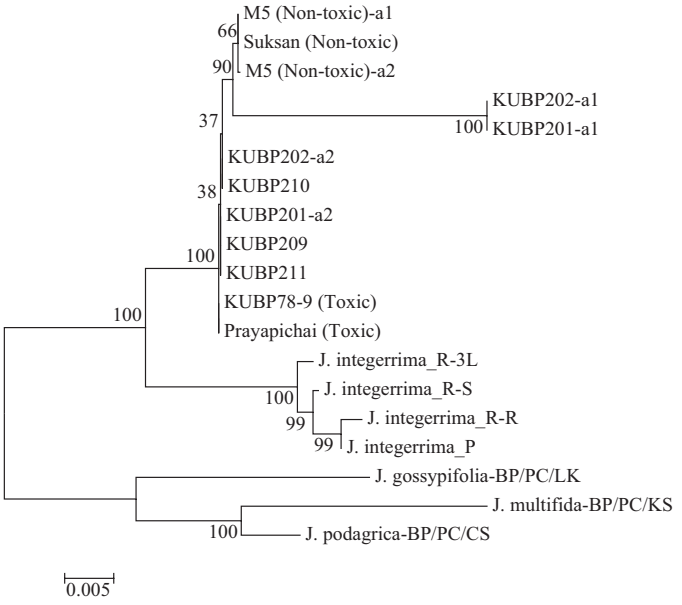

Fig. 4 Phylogenetic tree of 22 samples of *Jatropha* species and hybrids based on maximum-likelihood analysis of combined four regions of *rbc*L, *trn*H–*psb*A, internal transcribed spacer of *rDNA* and *stearoyl-acyl carrier protein desaturase* gene sequences. (a = allele)

via maternal inheritance. Therefore, they could separate *Jatropha* species, but could not classify hybrids from back crossing. However, prior reports demonstrated that both *rbc*L (Kellogg and Juliano, 1997) and *trn*H–*psb*A (Cowan et al., 2006) had significantly lower levels of genetic divergence within species than between species. By comparison, the ITS of *rDNA* had much higher divergence than any of the plastid regions, and *rbc*L was by far the lowest in divergence (Kress et al., 2005). The presence of multiple alleles of the ITS of *rDNA* help the variation of polymorphism to be fast and efficient (Newmaster et al., 2008). The *SAD* gene was appropriate for discrimination among the five *Jatropha* species and hybrids because of the large gene region and high genetic diversity in introns. It was found the 2-allele pattern was the same as the ITS of *rDNA* that could discriminate among the five *Jatropha* species, toxic and non-toxic *J.curcas*, some hybrids and all varieties of *J. integerrima*. The nuclear inheritance of the ITS of *rDNA* and *SAD* regions may have been a contributing factor due to its increased resolving potency compared to the maternally inherited plastid DNA of *rbc*L and *trn*H–*psb*A (Hollingsworth, 2011). The benefits from using nuclear DNA will improve the ability to distinguish among plant species. However, the various combinations of specific genes were all more efficient at differentiating between species than either gene individually (Kress and Erickson, 2007). The phylogenetic tree based on four gene regions confirmed *J. integerrima* has a close relationship with *J. curcas*, followed by *J. gossypifolia*, *J. multifida* and *J. podagrica* (Fig. 4). The intergenic hybridization of *J. curcas* and *J. integerrima* provides a high opportunity to successfully obtain more hybrids than with other species. According to Prabakaran and Sujatha (1999), the breeding of *J. curcas* L. and *J. gossypifolia* L. involved high incompatibility in metaphase I and anaphase of meiosis. The current study analyzed gene sequencing and the results implied genetic evolution in the *Jatropha* species. The information can be applied and used as background information for genetic improvement in *Jatropha* species and for the production of future desirable hybrids. Moreover, the gene sequence can be used for species-specific identification and hybrid detection.

Conflict of Interest

The authors declare that there are no conflicts of interest.

Acknowledgement

This research was supported by a scholarship from the Graduate School of Kasetsart University, Bangkok, Thailand; the Kasetsart University Research and Development Institute (KURDI); and the Kasetsart University Biodiesel Project and partially supported by the Advanced Studies in Tropical Natural Resources, National Research University - Kasetsart University, Higher Education Research Promotion and the National Research University Project of Thailand, Office of the Higher Education Commission.

References

Aiyelaagbe, O.O., Adesogan, K., Ekujndayo, O., Gloer J.B. 2007. Antibacteria diterpenoids from *Jatropha podagrica* Hook. Phytochemistry 68: 2420–2425.

Asif, M.H., Mantri, S.S., Sharma, A., et al. 2010. Complete sequence and organization of the *Jatropha curcas* (Euphorbiaceae) chloroplast genome. Tree Genet. Genomes. 6: 941–952.

Asseleih, L.M.C., Plumbley, R.A., Hylands, P.J. 1989. Purification and partial characterization of a hemagglutinin from seeds of *Jatropha curcas*. J. Food Biochem. 13: 1–20.

Blaxter, M., Floyd, R. 2003. Molecular taxonomics for biodiversity surveys: Already a reality. Trends Ecol. Evol. 18: 268–269.

Chase, M.W., Soltis, D.E., Olmstead, R.G. Morgan, D., Les, D.H., Mishler, B.D., Kron K.A. 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbc*L. Ann. Mo. Bot. Gard. 8: 528–580.

Chen, S., Pang, X., Song, J., Shi, L., Yao, H., Han, J., Leon, C. 2014. A renaissance in herbal medicine identification: From morphology to DNA. Biotechnol. Adv. 32: 1237–1244.

Chopra, R.N., Badhawar, R.L., Ghosh. S. 1965. Poisonous Plants of India Vol. II. New Delhi, India.

Coleman A.W., Mai, J.C. 1997. Ribosomal DNA and ITS-2 sequence comparisons as a tool for predicting genetic relatedness. J. Mol. Evol. 45: 168–177.

Cowan, R.S., Chase, M.W., Kress, W.J., Savolainen, V. 2006. 300,000 species to identify: problems, progress, and prospects in DNA barcoding of land plants. Taxon 55: 611–616.

Doyle, J.J., Doyle, J.L. 1990. Isolation of plant DNA from fresh tissue. Focus 12: 13–15.

Fofana, B., Harvengt, L., Baudoin, J.P., Du Jardin, P. 1997. New primers for the polymerase chain amplification of cpDNA intergenic spacers in Phaseolus phylogeny. Belg. J. Bot. 129: 118–122.

Hall, T.A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Window 95/98/NT. In: Nucleic Acids Symposium Series. Vol. 41. No. 41. [London]: Information Retrieval Ltd., c1979–c2000.

Harwood, J.L. 1988. Fatty acid metabolism. Annu. Rev. Plant Phys. 39: 101–138

Hollingsworth, P.M., Graham S.W., Little D.P. 2011. Choosing and using a plant DNA barcode. Plos one. 6: e19254. doi:10.1371/journal.pone.0019254

Jones, N., Miller. J.H. 1992. *Jatropha curcas*: A multipurpose species for problematic sites. Land Resources Series No. 1. Asia Technical Department. The World Bank. Washington, DC, USA.

Kellogg, E., Juliano, N.D. 1997. The structure and function of RuBisCo and their implications for systematic studies. Am. J. Bot. 84: 413–428.

Kress, W.J. Erickson, D.L. 2007. A two-locus global DNA barcode for land plans: the coding *rbc*L gene complements the non-coding *trn*H-*psb*A spacer region. Plos one. 2: e508. doi:10.1371/journal.pone.0000508

Kress, W.J., Erickson, D.L., Jones, F.A., Swenson. N.G., Perez, R., Sanjur, O., Bermingham, E. 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. P. Natl. Acad. Sci. USA. 106: 18621–18626.

Kress, W.J., Wurdack, K.J., Zimmer, E.A., Weigt, L.A., Jazen, D.H. 2005. Use of DNA barcodes to identify flowering plants. P. Natl. Acad. Sci. USA. 102: 8369–8374.

Lahaye, R., Michelle, B., Bogarin, D., et al. 2008. DNA barcoding the floras of biodiversity hotspots. P. Natl. Acad. Sci. USA. 105: 2923–2928.

Levin, R.A., Wagner, W.L., Hoch, P.C., Nepokroeff, M., Pires, J.C., Zimmer, E.A., Sytsma, K.J. 2003. Family level relationships of *Onagraceae* based on chloroplast *rbc*L and *ndh*F data. Am. J. Bot. 90: 107–115.

Newmaster, S.G., Fazekas, A.J., Steeves, R.A.D., Janovec, J. 2008. Testing candidate plant barcode regions in the Myristicaceae. Mol. Ecol. Resour. 8: 480–490.

Nigam, P.S., Singh, A. 2011. Production of liquid biofuels from renewable resources. Prog. Energ. Combust. 37: 52–68.

Oduola, T., Adeosun, G.O., Oduola, T.A., Avwioro, G.O., Oyeniyi, M.A. 2005. Mechanism of action of *Jatropha gossypifolia* stem latex as a haemostattic agent. Eur. J. Gen. Med. 2: 140–143.

Parthiban, K.T., Kumar, R.S., Thiyagarajan, P., Subbulakshmi, V., Vennila S. Rao, M.G. 2009. Hybrid progenies in *Jatropha* - a new development. Curr. Sci. India. 96: 815–823.

Prabakaran, A.J., Sujatha, M. 1999. *Jatropha tanjorensis* Eills and Saroja: a natural interspecific hybrid occurring in Tamilnadu. Gen. Resour. Crop Evol. 46: 213–218.

Rao, K.S., Lakshminarayana, G. 1987. Characteristics and composition of six newer seeds and the oils. Eur. J. Lipid Sci. Tech. 89: 324–326.

Ritland, K., Clegg, M.T. 1987. Evolutionary analysis of plant DNA sequences. Am. Nat. 130: S74–S100.

Sang, T., Crawford, D., Stuessy, T. 1997. Chloroplast DNA phylogeny, reticulate evolution and biogeography of *Paeonia* (Paeoniaceae). Am. J. Bot. 84: 1120–1120.

Schoch, C.L., Seifert, K.A., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.A., Chen, W., Fungal Barcoding Consortium. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. PNAS. 109: 6241–6246.

Shaw, J., Lickey, E.B., Schilling, E.E. Small, R.L. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am. J. Bot. 94: 275–288.

Soltis, P.S., Soltis, D.E., Smiley. C.J. 1992. An *rbc*L sequence from a Miocene Taxodium (bald cypress). Natl. Acad. Sci. USA. 89: 449–451.

Songpanich, P., Hongtrakul, V. 2010. Intersubgeneric cross in *Nymphaea* spp. L. to develop a blue hardy waterlily. Sci. Hortic-Amsterdam. 124: 475–481.

Soontornchainaksaeng, P., Jenjittikul, T. 2003. Karyology of *Jatropha* (Euphorbiaceae) in Thailand. Thai For. Bull. (Bot.) 31: 105–112.

Sujatha, M. 1996. Genetic and tissue culture studies in castor (*Ricinus communis* L.) and related genera. Ph.D. Thesis, Osmania University. Hyderabad, India.

Sujatha, M., Makkar, H.P.S., Becker, K. 2005. Shoot bud proliferation from axillary nodes and leaf sections of non-toxic *Jatropha curcas* L. Plant Growth Regul. 47: 83–90.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. 2011. MEGA5: molecular evolution distance, and maximum parsimony methods. Mol. Biol. Evol. 28: 2731–2739.

Tate, J.A. Simpson, B.B. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. Syst. Bot. 28: 723–737.

Tong, L., Shu-Ming, P., Wu-Yuan, D., Dan-Wei, M., Ying, X., Meng, X., Fang, C. 2006. Characterization of a new *stearoyl-acyl carrier protein desaturase* gene from *Jatropha curcas*. Biotechnol. Lett. 28: 657–662.

UImasov, T.N., Gulov, M.K., Aliev, K.A., Andrianov, V.M., Piruzian, E.S. 1990. Nucleotide sequence of the chloroplast *psb*A and *trn*H genes from cotton *Gossypium hirsutum*. Nucleic Acids Res. 18: 186–186.

White, T.J., Bruns, T., Lee, S. Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: PCR Protocols: A Guide to Methods and Applications. Academic Press, Inc., New York, pp. 315–322.

Wolfe, K.H., Li, W.H., Sharp, P.M. 1987. Rates of nucleotide substitution vary greatly among plant mitochondria, chloroplast, and nuclear DNAs. PNAS. 84: 9054–9058.

Yao, H., Song, J., Lui, C., et al. 2010. Use of ITS2 region as the universal DNA barcode for plants and animals. Plos one. 5: e13102. doi.org/10.1371/journal.pone.0013102.