TROPICAL GENETICS

Volume 3, No 2, November 2022
https://ojs.genetikawan-muda.com/index.php/tg

Original Research

Analysis of rbcL gene sequence genetic variation in *Dermatophyllum gypsophilum* subsp. *guadalupense* NCBI Popset 2440747772 using RFLP in silico

Cici Mustika, Afifatul Achyar*
Department of Biology, Faculty of Science and Mathematics, Universitas Negeri Padang, 25131, Kota Padang

***Corresponding author(s):** afifatul.achyar@fmipa.unp.ac.id

Article Info

Article history:
Received 25 Juni 2023
Accepted 10 Oktober 2023
Available online 30 November 2023

Keywords:
Dermatophyllum gypsophilum subsp. *guadalupense*, RFLP, *rbcL*, *in silico*

How to cite:
Mustika and Achyar. 2023. Analysis of rbcL gene sequence genetic variation in Dermatophyllum gypsophilum subsp. guadalupense NCBI Popset 2440747772 using RFLP in silico. *Tropical Genetics,* 3(2): 46-52.

Abstract

Dermatophyllum gypsophilum subsp. *guadalupense* is a rare species of flowering plant in the legume family known by the common name *Guadalupe Mountain necklacepod.* This plant is native to New Mexico and Texas in the United States. The *rbcL* gene encoding the RuBisCO protein is thought to cause this gene sequence to have a low mutation rate compared to other barcode genes in cpDNA so that the level of similarity between species is quite high. This study used restriction enzyme *Pci*I. This study aimed to analyze genetic variation in *rbcL* gene sequence in *Dermatophyllum gypsophilum* subsp. *guadalupense* NCBI with PopSet 2440747772 using RFLP *in silico.* The results of this study showed genetic variations in the *rbcL* gene sequence of *Dermatophyllum gypsophilum* subsp. *guadalupense* and 8 allele variations contained in 73 gene sequences using the restriction enzyme *Pci*I.

Introduction

Dermatophyllum gypsophilum is a rare species of flowering plant in the legume family known by the common names Guadalupe *Mountain necklacepod, Guadalupe mescalbean* (var. *guadalupensis*), and *gypsum necklace*. This plant is native to New Mexico and Texas in the United States, and is known from one location in Chihuahua in Mexico (Turner, 2012).

Ammodendron, *Ammothamnus*, and *Echinosophora* are embedded within *Sophora* s.s. and that nine well supported clades can be recognized within comprise *Sophora* s.l. Ancestral character state estimation revealed that the most recent common ancestor of *Sophora* s.l. was a deciduous shrub that lacks rhizome spines and has unwinged legumes. Divergence times estimation and ancestral area reconstruction showed that Sophora s.l. originated in Central Asia and/or adjacent Southeast China in the early Oligocene (ca. 31 Mya) and dispersed from these regions into East and South Asia's adjacent areas and North America via the Bering land bridge (Liao et al., 2023).

Data character sources to identify molecular genetic diversity can be obtained from nuclei (nDNA), chloroplasts (cpDNA) and mitochondria (mtDNA). The source of genetic diversity data that is widely chosen for the object of plant research is cpDNA. Pierce (2002) explains that one of the key proteins encoded by cpDNA is ribulose-1,5-biphosphate carboxylase-oxygenase (abbreviated as

RuBisCO), which participates in carbon fixation in the process of photosynthesis. RuBisCO makes about 50% of the protein found in green plants which is thought to be the most abundant protein on earth. The RuBisCO protein complex is composed of eight identical large subunit proteins and eight identical small subunit proteins. Large subunit proteins are encoded by cpDNA, while small subunit proteins are encoded by nuclear DNA. The *rbcL* gene measures about 1400 bp in length (Syamsurizal et al., 2021), thus providing many characters for phylogenetic studies (CBOL, 2009). The role of the *rbcL* gene that codes for the RuBisCO protein is thought to cause this gene sequence to have a low mutation rate compared to other barcode genes in cpDNA so that the level of similarity between species is quite high (Kellgog & Juliani, 1997). This low mutation rate provides advantages for in-depth study of intraspecies genetic and phylogenetic variation.

Genetic variation is a change that occurs in nucleotides, genes, chromosomes, and genomes in an organism. Genetic variation in a population will affect the survival ability of an individual (Frankham et al., 2002). According to Dunham (2004), the higher the genetic variation found in a population, the better the individual's ability to adapt to environmental changes (Achyar et al., 2021).

RFLP (Restriction Fragment Length Polymorphism) is a commonly used technique for genotyping through cutting RNA sequences with restriction enzymes. The restricted gene fragments were separated using electrophoresis and visualized using the Southern Blotting technique (Dai and Long, 2015). According to Siti et al., (2013) the RFLP method as one method to determine polymorphism in studying the evolutionary history of human populations (lineage / genealogy) and to determine mutations. The RFLP method is an analytical method using restriction enzymes that cut typical nucleotide sequences at different specific locations so that fragments of different lengths are produced (Theodore, 2000).

Test *in silico* is a term for experiments or tests conducted with computer simulation methods. The use of *in silico* tests is to predict, hypothesize, provide new discoveries or new advances in medicine and therapy (Hardjono, 2013). According to Suhanna, 2012; Johan, 2016 simply *in silico* can be translated as a method to strive for a real-condition approach into computer-based simulations using certain application programs or software (Achyar et al, 2021).

Materials and Methods

1. Materials

   In this study, a sequence of *Dermatophyllum gypsophilum* subsp*. guadalupense plant genes* was used, namely the ribulose-1,5- biphosphate carboxylase-oxygenase (*rbcL*) gene with PopSet identity number 2440747772 downloaded in FASTA format from the NCBI website. In this Popset there are 73 gene sequences with accession numbers (Table 1). This PopSet originally submitted by Liao,M., Shepherd,L.D., Zhang,J.Y., Feng,Y., Mattapha,S., Zhang,L.B., Gao,X.F. and Xu,B. in their study entitled "Phylogeny, biogeography, and character evolution of the genus Sophora s.l. (Fabaceae, Papilionoideae)" (https://www.ncbi.nlm.nih.gov/popset/2440747772).

2. Methods

   Manuscript is divided using the numbered sections. Authors should divide the manuscript into clearly defined and numbered sections. Second level section numbering is done automatically; following the upper level's number. Use this numbering also for internal

cross-referencing: do not just refer to the text. Any subsection should be given a brief heading.

Table 1. *rbcL* gene sequences in *Dermatophyllum* sp. / *Sophora* sp. NCBI

No	Acc Number	No	Acc Number	No	Acc Number	No	Acc Number
1	ON996767.1	21	ON996747.1	41	ON996727.1	61	ON996707.1
2	ON996766.1	22	ON996746.1	42	ON996726.1	62	ON996706.1
3	ON996765.1	23	ON996745.1	43	ON996725.1	63	ON996705.1
4	ON996764.1	24	ON996744.1	44	ON996724.1	64	ON996704.1
5	ON996763.1	25	ON996743.1	45	ON996723.1	65	ON996703.1
6	ON996762.1	26	ON996742.1	46	ON996722.1	66	ON996702.1
7	ON996761.1	27	ON996741.1	47	ON996721.1	67	ON996701.1
8	ON996760.1	28	ON996740.1	48	ON996720.1	68	ON996700.1
9	ON996759.1	29	ON996739.1	49	ON996719.1	69	ON996699.1
10	ON996758.1	30	ON996738.1	50	ON996718.1	70	ON996698.1
11	ON996757.1	31	ON996737.1	51	ON996717.1	71	ON996697.1
12	ON996756.1	32	ON996736.1	52	ON996716.1	72	ON996696.1
13	ON996755.1	33	ON996735.1	53	ON996715.1	73	ON996695.1
14	ON996754.1	34	ON996734.1	54	ON996714.1		
15	ON996753.1	35	ON996733.1	55	ON996713.1		
16	ON996752.1	36	ON996732.1	56	ON996712.1		
17	ON996751.1	37	ON996731.1	57	ON996711.1		
18	ON996750.1	38	ON996730.1	58	ON996710.1		
19	ON996749.1	39	ON996729.1	59	ON996709.1		
20	ON996748.1	40	ON996728.1	60	ON996708.1		

Restriction Enzyme Candidate Screening Method

In candidate screening, restriction enzymes are performed at the site of the http://insilico.ehu.es/restriction/ . For this method, use the tool 'compare restriction pattern of many sequences'. Then, enter the sequence file that has been downloaded in the form of FASTA. Then click 'Go to next' to process the file that has been inputted. After a while, a list of sequences appears and ticks the 'only restriction enzymes with known bases (no N,R,Y,...)' so that only definite bases will be processed. The next step, click 'Get list of restriction enzymes' so that several restriction enzyme candidates will be used in the next stage.

RFLP Method *in silico*

RFLP is done *in silico* or restriction virtually using https://www.benchling.com/ site. At the initial stage, you have to register using email. If you have registered, continue with importing the RNA sequence that has been downloaded from NCBI into the project folder on the Benchling site. Then click the scissors mark in the right corner. Then the "find enzyme" tool is selected and the names of the restriction enzymes that are predetermined at the time of screening, are typed in the column. Next, click the "run digest" menu for restrictions. With the "virtual digest" tool, you can see electropherogram images (Yeriska et al., 2021; Syamsurizal, S., Kadri, 2018).

Results and Discussion

Restriction Enzyme Candidate Screening

At the restriction enzyme screening stage, various restriction enzymes were obtained and *Pci*I restriction enzymes were selected that cut on the 5'-A'CATGT-3' introduction side. This enzyme is used because there is a variation in the *rbcL* gene sequence in *Dermatophyllum gypsophilum* subsp. *guadalupense* with Popset 2440747772.

RFLP *in silico*

RFLP (Restriction Fragment Length Polymorphism) is a commonly used technique for genotyping determination through cutting DNA sequences with restriction enzymes (Achyar et al, 2021). According to Siti et al., (2013) the RFLP method is one method to determine polymorphism in studying the evolutionary history of human populations (lineage / genealogy) and to determine mutations. The RFLP method is an analytical method using restriction enzymes that cut typical nucleotide sequences at different specific locations so that fragments of different lengths are produced (Theodore, 2000).

RFLP analysis which is a codominant marker has been widely used to achieve various goals. Given that restriction sites have certain DNA sequences, it means that variations in the existence of restriction sites reflect variations in DNA sequences. In other words, RFLPs can serve as predictors of DNA variation. Variation is detected in the form of cutting polymorphic (double) long circuits where the assessment time of the variation series allows from the fragment data itself, long series of variations in a section can be assessed from nucleotide substitution (Fatchiyah et al., 2011)

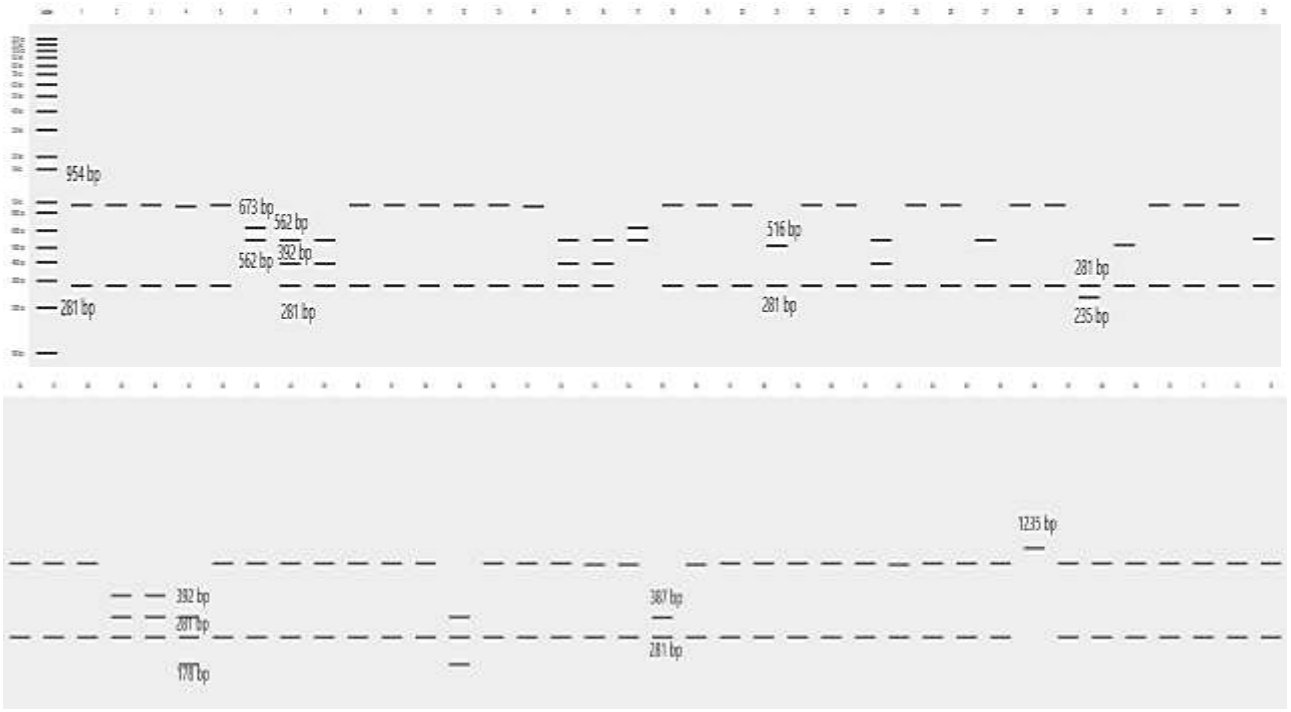

Figure 1. Electropherogram of *rbcL* gene sequence restriction results *in Dermatophyllum gypsophilum* subsp*. guadalupense* using *Pci*I restriction enzyme *in silico*. Left : (Ladder Life 1 kb Plus), (1) ON996767.1, (2) ON996766.1, (3) ON996765.1, (4) ON996764.1, (5) ON996763.1, (6) ON996762.1, (7) ON996761.1, (8) ON996760.1, (9) ON996759.1, (10) ON996758.1, (11) ON996757.1, (12) ON996756.1, (13) ON996755.1, (14) ON996754.1, (15) ON996753.1, (16) ON996752.1, (17) ON996751.1, (18) ON996750.1, (19) ON996749.1, (20) ON996748.1 (21) ON996747.1, (22) ON996746.1, (23) ON996745.1, (24) ON996744.1, (25) ON996743.1, (26) ON996742.1, (27) ON996741.1, (28) ON996740.1, (29) ON996739.1, (30) ON996738.1, (31) ON996737.1, (32) ON996736.1, (33) ON996735.1, (34) ON996734.1, (35) ON996733.1, (36) ON996732.1, (37) ON996731.1, (38) ON996730.1, (39) ON996729.1, (40) ON996728.1, (41) ON996727.1, (42) ON996726.1, (43) ON996725.1, (44) ON996724.1, (45) ON996723.1, (46) ON996722.1, (47) ON996721.1, (48) ON996720.1, (49) ON996719.1, (50) ON996718.1, (51) ON996717.1, (52) ON996716.1, (53) ON996715.1, (54) ON996714.1, (55) ON996713.1, (56) ON996712.1, (57) ON996711.1, (58) ON996710.1, (59) ON996709.1, (60) ON996708.1, (61) ON996707.1, (62) ON996706.1, (63) ON996705.1, (64) ON996704.1, (65) ON996703.1, (66) ON996702.1, (67) ON996701.1, (68) ON996700.1, (69) ON996699.1, (70) ON996698.1, (71) ON996697.1, (72) ON996696.1, (73) ON996695.1

RFLP on 73 *rbcL* gene sequences *of Dermatophyllum gypsophilum* subsp*. guadalupense* was cut with restriction enzyme *Pci*I*.* The results of the RFLP electropherogram from the restriction enzyme *Pci*I can be seen in the picture above. Restriction using the *Pci*I enzyme produces 8 allele variations, the first allele produces 2 DNA bands measuring 954 bp and 281 bp contained in the sample (1-5, 9-14, 18-20, 22-23, 25, 26, 28, 29, 32-34, 36-38, 42-48, 50-54, 56-65, 67-73), the second allele produces 2 DNA bands measuring 673 bp and 562 bp contained in the sample (6 and 17), the

third allele produces 3 DNA bands measuring 562 bp, 392 bp, and 281 bp contained in the sample (7, 8, 15, 16, 24, 39, 40), the fourth allele produces 2 DNA bands measuring 516 bp and 281 bp contained in the sample (21, 27, 31, 35), the fifth allele produces 2 DNA bands measuring 281 bp and 235 bp contained in the 30th sample, the sixth allele produces 3 DNA bands measuring 392 bp, 281 bp, and 178 bp in the sample (41 and 49), the seventh allele produces 2 DNA bands measuring 387 bp and 281 bp in the 55th sample, and the eighth allele producing 1 DNA band measuring 1235 bp in the 66th sample. Allele frequency in *Dermatophyllum gypsophilum* subsp*. guadalupense* NCBI *rbcL* gene sequence with restriction enzyme *Pci*I can be seen in table 2.

Table 2. Gene Allele Frequency *rbcL Dermatophyllum gypsophilum* subsp. *guadalupense* NCBI Based on RFLP *In silico* Results

Restriction Enzyme	Recognition Site	Fragment Size (bp)	Allele	Number of Fragment Presence (N=73)	Fragment Presence Percentage (%)	Allele Frequency
*Pci*I	A'CATGT	954 and 281	1	55	75,34	0,75
		673 and 562	2	2	2,73	0,03
		562, 392, 281	3	7	9,58	0,096
		516 and 281	4	4	5,47	0,055
		281 dan 235	5	1	1,36	0,013
		392, 281, 178	6	2	2,73	0,03
		387 dan 281	7	1	1,36	0,013
		1235	8	1	1.36	0,013

Based on table 2, it is known that the total amplified DNA bands at 954 bp and 281 bp fragment sizes are 110 DNA bands with each amplified fragment size as many as 55 DNA bands, besides that this fragment has a percentage of 75.34% presence. In fragments measuring 673 bp and 562 bp have a total of 4 DNA bands with each fragment size amplified as many as 2 DNA bands, and in this fragment has a percentage of presence of 2.73%. In fragments measuring 562 bp, 392 bp and 281 bp have a total of 21 DNA bands with each fragment size amplified as many as 7 DNA bands, besides that this fragment has a percentage of presence of 9.58%. In fragments measuring 516 bp and 281 bp have a total of 8 DNA bands with each fragment size amplified as many as 4 DNA bands, and in this fragment has a percentage of presence of 5.47%. In fragments measuring 281 bp and 235 bp have a total of 2 DNA bands with each fragment size amplified as much as 1 DNA band, and in this fragment has a percentage of presence of 1.36%. In fragments measuring 392 bp, 281 bp and 178 bp have a total of 6 DNA bands with each fragment size amplified as many as 2 DNA bands, besides that this fragment has a percentage of presence of 2.73%. In fragments measuring 387 bp and 281 bp have a total of 2 DNA bands with each fragment size amplified as much as 1 DNA band, and in this fragment has a percentage of presence of 1.36%. In a fragment measuring 1235 bp has a total DNA band of 1 DNA band with each fragment size amplified as much as 1 DNA band, and in this fragment has a percentage of presence of 1.36%.

Conclusion

The results of RFLP *in silico* in this study showed genetic variation, by giving rise to 8 allele variations in 73 DNA sequences of the *rbcL* gene in plants *Dermatophyllum gypsophilum* subsp. *guadalupense* with Popset 2440747772.

References

Achyar, A., Hindayageni, A., Humaira, F., Wijaya, N.N., Aqsha, N., Zultsatunni'mah. (2021). Analysis of Genetic Variations in POLY Gene Sequences in Dengue Virus 2 Using In-Silico RFLP. Jurnal Bioscience. Vol. 05. No. 1: 80-86.

Consortium Barcode of Life (CBOL). (2009). A DNA Barcode for Land Plants. *PNAS*, 106 (31)

Dai S, Long Y. (2015). Genotyping analysis using an RFLP assay. *Methods Mol Biol*. 1245:91-9.

Dunham, RA. (2004). *Aquaculture and Fisheries Biotechnology Genetic Approaches*. USA: Department of Fisheries and Allied Aquacultures Auburn University Alabama. 86

Frankham, R., Ballou, J.D., Briscoe, D.A. (2002). *Introduction to Conservation Genetics.* Cambridge: Cambridge University

Hardjono S. (2013). Sintesis dan Uji Aktivitas Antikanker Senyawa 1-(2- Klorobenzoiloksi) urea dan 1-(4-klorobenzoiloksi) urea. *Berkala Ilmiah Kimia Farmasi*. 2 (1), 1.

Johan, A. K. (2016). Uji *In silico* Senyawa Genistein sebagai Ligan pada Reseptor Estrogen Beta. *Skripsi*. Yogyakarta: Universitas Sanata Dharma.

Liao, M., Shepherd, L. D., Zhang, J. Y., Feng, Y., Mattapha, S., Zhang, L. B., Gao, X. F., & Xu, B. (2023). Phylogeny, biogeography, and character evolution of the genus Sophora s.l. (Fabaceae, Papilionoideae). *Molecular phylogenetics and evolution*, 181, 107713. https://doi.org/10.1016/j.ympev.2023.107713

Pierce, B. A. (2002). *Genetiks: A Conceptual Approach*. New York: W. H. Freeman and Company.

Siti, Heli H.M., G. Gun G., Desy N., Achmad S. Noer. (2013). *Variasi Urutan Nukleotida Daerah D-Loop DNA Mitokondria Manusia pada Dua Populasi Asli Indonesia Tenggara.* P. 440- 446.

Suharna. (2012). Studi *In silico* Senyawa Turunan Flavonoid terhadap Penghambatan Enzim Tirosinase. *Skripsi*. Makassar: UIN Alauddin

Syamsurizal, S., Kadri, H. (2018). Genotyping SNP Rs12255372 TCF7L2 Gene Using Three-Primer ARMS-PCR for Detection T2DM n Indonesian Batak Ethnic. *Journal of Physics: Conference Series*, *1040*(1), 12003. IOP Publishing.

Syamsurizal, S., Ardi, A., Des, M., Fevria, R., Atifah, Y., Badriyya, E., & Achyar, A. (2021). Design of primer Ipomoea batatas chloroplast gene matK. *Tropical Genetics*, *1*(1), 12–16.

Turner, B.L. (2012). New names in Dermatophyllum (Fabaceae). *Phytoneuron*. 3: 1–4.