Sequences of Trypsin Inhibitor Gene on *Jatropha curcas* L.

P R Primandiri¹, M Amin², S Zubaidah³, Maftuchah³, A M Santoso¹

¹ Biology Education, University of Nusantra PGRI Kediri, Kediri, Indonesia
² Faculty of Mathematic and Natural Science, State Malang University, Malang, Indonesia
³ Faculty of Agriculture, University of Muhammadiyah Malang, Malang, Indonesia

Corresponding author: poppyprimandiri@unpkediri.ac.id

**Abstract.** Trypsin Inhibitor (*TI*) is one of unique gene which involved in plant protection from biotic and abiotic defense. The exploration variation of *TI* gene sequence is crucially needed as an abiotic reference, especially in *J. curcas* to support national biodiesel production. *TI* gene in *J. curcas* has been successfully isolated and sequenced form several numbers (5 (SP-8 X SP-16), 6 (SP-8 X SP-38), 7 (SP-33 X HS-49), and 18 (SM-35 X SP-38) of *J. curcas*). Fourth sequence samples compared with two prime accessions (IP3A and IP3P). *TI* gene was successfully isolated, sequenced (598-604 bp), and aligned using *BioEdit Sequence Alignment Editor* 7.2.5 program. There were insertion and substitution that occur in all samples that lead to differences *TI* gene sequences. There are 11 sites was inserted. Sample 18 and 7 only inserted at one site only. Substitution has been occurred at 14 sites, 35.71% (5 sites) as transition and 64.29% as transversion. Both of inversion and substitution in some sites leads to change in the proportion of nucleic acid of *TI* genes. Further analysis is needed to determine the impact of *TI* gene variation on *J. curcas* physiological response in environmental defense.

**1. Introduction**

Seeds of *J. curcas* has been potential as one of the renewable biodiesel feedstock [1,2,3,4] and *J. curcas* oils include non-edible oils that do not compete with human consumptions needs such as both of palm and corn oils. The cultivation of *J. curcas* is very easy by stem cuttings [5] and easy to grow on any land condition because it can adapt to marginal environmental condition such as under drought and lack nutrients land [3,6,7,8,9,10] and under saline condition [9,11]. Therefore, government policy of Republic of Indonesia is *J. curcas* cultivation is focused on marginal and dry land.

On the other hand, several field findings indicate that some *J. curcas* are superior to drought stress but are vulnerable to several insects. They were *Valanga nigricornis*, *Selenophthrips rubrocinctus*, *Nipaecoccus viridis* Newstead, *Exopalis hypoleuca* Wied, *Luecopolis rorida*, *Chyschosoris javanus* Westw., *Tetanychidae* [12], *Planococcus* sp., *Polyphagotarsonemus latus* Banks, *Selenophthrips rubrocinctus* Giard [13]. Furthermore, screening of *J. curcas* from local areas is still being carried out, especially those that are tolerant of abiotic and biotic stress.

The molecular approach for the identification of specific gene encoding tolerant to biotic and abiotic stress on *J. curcas* has not been performed. Identification of specific genes is limited. For examples most of gene which involved in regulation responses under drought stress in the other plant such as Δ1-Pyrroline-5-carboxylate synthetase (*P5CS*) [14], *JcRD29B*, *JcNCED3*, *JcDREB*, *JcDI19*, *JcDRS1*, *JcRAP2*, *JcSAM*, *JcPIP2*, *JcERF*, *JcBD1*, and *JcLEA-5* genes [15], gene of *JcDREBA* [16],...
JcMYB2 gene [17], and JcCBF2 which involved in cold stress [18]. However, the identification of Trypsin Inhibitor (TI) gene on J. curcas has not been reported.

Identification of TI genes in J. curcas is important as TI genes are able to regulate plant resistance from both of biotic and abiotic stress [19,20]. TI is a member of the Protease Inhibitor (PI) that plays an important role in plant response to stress and development [21]. In some plants, TI is detected in plants that are tolerant to lack of salt minerals [20,22,23], under drought stress [23,24,25], and also detected in some plants that were tolerant to insect [20,26,27,28,29].

Currently, there were four numbers of J. curcas form several local J. curcas in Indonesia, namely number 5 (SP-8 x SP-16), 6 (SP-8 x SP-38), 7 (SP-33 x HS-49), and 18 (SM-35 x SP-38), that the four crossing number had high productivity [30]. For example, number 6 (SP-8 x SP-38) able to have highest dry seed and number 18 (SM-35 x SP-38) has highest oil content up to 32.035%. This study was aimed to detect variation of TI genes on the four J. curcas cross number.

2. Materials and Method
2.1 Plant Material
There were four numbers of J. curcas which are developed by Maftuchah [30]. The number 5 (SP-8 X SP-16), 6 (SP-8 X SP-38), 7 (SP-33 X HS-49), and 18 (SM-35 X SP-38), accession IP3A and IP3P which has been released as a prime accession by Agriculture Department of Republic of Indonesia was used as a comparison. Mature, fresh, and healthy leaf samples were collected from Pasuruan East Java (112030’ – 113030’ BT and 7030’ – 8030’LS) and used to further steps.

2.2 Isolation and Sequence of TI
DNA isolation was done by using protocol from Doyle & Doyle modified by Maftuchah & Zainudin [31] with CTAB extraction buffer. Isolated DNA was used as a template for DNA amplification. A pair of primer, CpTI-F (5’-ATG AAG AGC ACC ATC TTC TTT GCT C-3’) and CpTI-R (5’-CTT ACT CAT CAT CTT CAT CCC TGG-3’) [32] were used to amplify of TI gene. The total volume of PCR reaction was conducted as following condition: 25 μl total volume, consist of a mixture solution including taq DNA polymerase and 10X buffer TaqPolimerase (100 mMTris-Cl, pH 8.3; 500 mM KCI; 15 mM MgCl2; 0.01% gelatin); dNTPs mix (dGTP, dATP, dTTP and dCTP) (Recho); dH2O. The conditions for the PCR reaction was conducted following condition: pre-denaturation 94 °C (5 min), denaturation 94 °C (1 min), annealing 40 °C (1 min), an elongation of 72 °C (2 min) and post-PCR 4 °C (2 min), for 35 cycles. Electrophoresis on a 1.5% agarose and visualization under a UV-transilluminator were used to check the PCR result. The process of sequencing was performed according to the ABI PRISM DNA Analyzer procedure.

2.3 Alignment and Analysis
The results of sequencing were aligned using BioEdit Sequence Alignment Editor 7.2.5 program. Proportion of amino acids was analyzed using the MEGA 7 program.

3. Results and Discussion
PCR analysis was carried out successfully to amplify TI gene and the visualization is presented in Figure 1. Based on the figure, TI gene was detected on all plant samples including IP3A and IP3P (prime accession) as the species ancestor in this research. In this research, TI gene was detected in 450 bp to all samples. In this experiment, has been obtained for optimization of PCR condition to amplify TI gene from J. curcas.

Figure 1. TI gene was detected in J. curcas with accession number of IP3A (lane 1), IP3P (lane 2), number 5 (lane 3), 6 (lane 4), 7 (lane 5), 18 (lane 6), marker (lane 7) in 1.5% agarose.
An analysis of the TI gene alignment has been performed on four crossed numbers (number 5 (SP-8 X SP-16), 6 (SP-8 X SP-38), 7 (SP-33 X HS-49), and 18 (SM-35 X SP-38) with the IP3A and IP3P as a prime accession. There were differences sequences on TI gene on some sites. That difference can be presented in Table 1. The differences was grouped according to the pattern of nucleotide changes i.e. insertion and substitution. There were two kinds of substitution, transition and transversions. In number 6 (SP-8 X SP-38), there were 6 inserted sites at 37, 39, 85, 111, 259, and 578 and there were 3 inserted sites in number 5 (SP-8 X SP-16) i.e. sites at 35, 43, dan 578. In number 18 (SM-35 X SP-38) and number 7 (SP-33 X HS-49), there was only one inserted site, and they were sites of 575 and 37 respectively.

Table 1. Insertion and substitution (transition and transversion) in TI gene of the four J. curcas samples

| Sample | Sites |
|--------|-------|
| IP3A   | C     A  T    -    G T A C A A A -   A A - C A -  |
| IP3P   | T     -   -    -   A - - - -   -   -   -   -   |
| 5      | T     C   -   -   T C A -   C -   -   -   -   |
| 6      | C     T   A  T    -   A -   -   C T T T G C T T - |

Notes: * insertion

There were 3 patterns of transition nucleotide changes that have been found. First, the transition of T↔C can be found at site 36. Second, the transition of C ↔ T at sites 31, 71, and 292 and the transition of A↔G can be found at site 138 only. On the other hand, it has been found 3 patterns of transversion. They were A↔T at sites 34, 42, 57, 80, 132, and 323; A↔C at sites 54 and 75, then G↔T at site 49 in number 6 (SP-8 X SP-38).

The multiple transitions on J. curcas sample were C → T. The higher C → T transition frequency can be caused by nucleotide methylation [33]. Cytosine is known to be the only base naturally methylated in the eukaryotic genome [34]. Methylation in plant DNA sequences can be up to 30% while in fish and amphibians it can be up to 10% [35]. Cytosine DNA methylation is important in regulating gene expression [36,37]. The main function of cytosine in plants is as a defense system to protect the genome against endogenous DNA damage and viral invasion [38]. Cytosine deamination is a source of the C → T transition under in vivo conditions [39] and deamination of methylated cytosines can cause very high transition rates at some sites [40].

This case, the insertions that occurs above causes a shift in amino acid reading and proportion of amino acid in samples. Proportion amino acid changes of all samples were presented in Table 2.

Table 2. Proportion of amino acids in all of the observed J. curcas sequences.

| Sample | Proportion amino acid (%) |
|--------|---------------------------|
|        | Ala | Cys | Asp | Glu | Phe | Gly | His | Ile | Lys | Leu |
| IP3A   | 2.82| 3.39| 3.95| 3.95| 7.91| 2.26| 1.13| 7.91| 4.52| 15.25|
| IP3P   | 2.82| 3.39| 4.52| 3.95| 7.91| 2.26| 1.13| 7.91| 4.52| 15.25|
| 18     | 2.81| 3.37| 4.49| 3.93| 8.43| 2.25| 1.12| 7.87| 4.49| 15.17|
| 5      | 2.78| 3.33| 4.44| 3.89| 7.78| 2.22| 2.22| 7.78| 4.44| 15.00|
| 7      | 2.82| 3.39| 5.65| 2.82| 7.91| 2.26| 1.13| 7.91| 4.52| 15.25|
| 6      | 2.76| 3.31| 4.42| 3.31| 7.74| 2.21| 1.11| 9.39| 4.42| 15.47|
| Avg.   | 2.80| 3.36| 4.58| 3.65| 7.94| 2.24| 1.31| 8.13| 4.49| 15.23|

| Met | Asn | Pro | Gln | Arg | Ser | Thr | Val | Trp | Tyr |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| IP3A| 2.83| 4.52| 7.35| 3.96| 3.39| 11.30| 3.96| 5.65| 2.83| 1.13|
The highest proportion of amino acids was leucine at 15.23\%, followed by serine at 11.5\%. The amino acid in small amounts is tyrosine, 1.22\%, followed by histidine, 1.31\%. The height of leucine is a manifestation of the codons rich in C and T bases, besides that leucine is encoded by 6 different codons with variations in the position of the third base at different codons. The role of leucine is to modulate protein synthesis, primarily by stimulating the activity of proteins involved in the translation process.

4. Conclusion

Based on the exposure and data analysis, it can be concluded that the TI gene band size on J. curcas is 450 bp. The highest average proportion of amino acids was leucine, 15.23\%, the amino acid in the least amount was tyrosine 1.22\%, followed by histidine, 1.31%.

Acknowledgment

All author thanks to University of Nusantara PGRI Kediri, Kediri, Indonesia for supporting and facilitate this study.

References

[1] Achten WMJ and Mathijs E 2007 Biofuels, Bioprod. Bioref 1 283-291
[2] van Eijck J and Romijn H 2008 Energy Policy 36 311-325
[3] Sharma DK Pandey A K Lata 2009 Biomass and Energy 33 159-162
[4] Syakir M 2010 Perspektif 9 (2) 55-65
[5] Charloq 2008 Jurnal Penelitian Rekayasa 1 (1) 51-56
[6] Kheira AAA and Atta NMM 2009 Biomass & Energy 33 1343-1350.
[7] Maes WH Achten WMJ Reubens B Raes D Samson R Muys B 2009 Journal of Arid Environments 73 877-884
[8] Achten et al, 2010
[9] Fujimaki H and Kikuchi N 2010 Agrophysics 24 121-127.
[10] Sapeta H Costa J M Lourenco T Maroco J van der Linde P Oliveira M M 2013 Environmental and Experimental Botany 85 76–84
[11] Patel A D Panchal N S Pandey I B Pandey A N 2010 Anales de Biología 32 59-71
[12] Rumini W dan Karmawati E 2006 Prosiding Lokakarya II Status Teknologi Tanaman Jarak Pagar Bogor 29 November
[13] Rumini W dan Karmawati E 2007 Lokakarya Nasional III Inovasi Teknologi Jarak Pagar untuk Mendukung Program Desa Mandiri Energi Malang 5 November
[14] Zhuang G Q Li B Guo H Y Liu J L Chen F 2011 African Journal of Biotechnology 10 (66) 14803-14811
[15] Cartagena J A Seki M Tanaka M Yamauchi T Sato S Hirakawa H Tsuge T 2014 Plant Mol Biol Rep
[16] Tang Y Liu Kun L Zhang J Li X Xu K Zhang Y Yu D Wang J Li C Qi J 2017 Frontiers in Plant Science 8 1-12.
[17] Peng X Liu H Wang D Shen S 2016 BMC Genomics 17 251-262.
[18] Wang L Luo L Hu M Gao J Zhang G Liu Q Qin X Yu H Xu Y Shi X Li C Shen F 2014 Mol Biol Rep
[19] Sanchez-HC Martinez-GN Guerrero-RA Valdes-RS Delano-FJ 2004 Physiol Plantarum 122 254-264
[20] Srinivasan T Kumar KRR Kirti PB 2009 Plant Cell Physiol 50 (3) 541-553
[21] Singh A Sahi C Grover A 2009 Gene 428 9-19
[22] Shan L Li C Chen F Zhao S Xia G 2008 Plant, Cell and Environment 31 1128–1137
[23] Kidric M Kos J Sabotic J 2014 Botanica Serbica 38 (1) 139-158
[24] Lawrence PK and Koundal KR 2002 EJB Electronic Journal of Biotechnology 5 (1) 93-109
[25] Anandhan S Insaf A Qureshi and KR Koundal 2010 Physiol. Mol. Biol. Plants 16 (1) 31-37
[26] Maftuchah Zainudin A Sudarmo H 2013 Agricultural Sciences 4 (1) 48-56
[27] Zhang Q Zou Y Lin S Wang Z Lin Y Zhang Z Liu H 2004 Forestry Studies in China, 6 (3): 15-19.
[28] Coulondre C Miller JH Farabaugh PJ Philip J Walter G 1978 Nature 274: 775–780.
[29] Adams RLP 1996 DNA Methylation in Principles of Medical Biology. (Bittar, E.E., ed.). New York: JAI Press Inc. 5: 33-66.
[30] Cokus SJ Feng S Zhang X Chen Z Merriman B Haudenschild CD Pradhan S Nelson SF Pellegrini M Jacobsen SE 2008 Nature 452: 215-219
[31] Kreutzer DA Essigmann JM 1998 Proc. Natl. Acad. Sci 95: 3578–3582.
[32] Keller I Bensasson D Nichols RA 2007 PLOS Genetics 3 (2): 185-191.