Analysis of flowering gene in palm oil (*Elaeis guineensis*)

IRVAN FAIZAL[1,2,♥], AXEL EMDI[2]

[1]Agency for the Assessment and Application of Technology (BPPT). Gedung 2 BPPT Lantai 10, Jl. Thamrin No. 8, Jakarta 10340, Indonesia. ♥email:
irvan.faizal@bppt.go.id
[2]Faculty of Biotechnology, Universitas Katholik Atma Jaya. Jl. Jenderal Sudirman 51, Jakarta 12930, Indonesia

Abstract. *Faizal I, Emdi A. 2017. Analysis of flowering gene in palm oil (*Elaeis guineensis*). Asian J Agric 1: 53-58.* Palm oil has always been an important commodity in Indonesia. The most common species of palm oil is *Elaeis guineensis*. Palm oil is a monoecious plant with a tendency to be a temporal dioecious. The female flower will be the one that produces palm oil fruit, which later is treated with palm oil while the male flower only takes part in the fertilization process. To know the ratio between female and male flower trees in a plantation, this study was performed to detect a distinction between female and male flowering gene sequences from DNA samples of *E. guineensis*. Based on a previous study that managed to characterize *MADS*-box gene of palm oil, a primer was designed and named GmG (Globosa-male-Gaps). The result shows that the primer can differentiate DNA sequence female and male flower of *E.guineensis*, Palm oil. However, further studies with full sequence and more samples are needed to find distinctive results between female and male flower sequences as the GmG primer could be used to design a specific marker or primer to detect the presence of female or male flower within a tree.

Keywords: *Elaeis guineensis,* dioecious, male-female ratio, *MADS*-Box, flowering gene

INTRODUCTION

Palm oil is one of Indonesia's main commodities. It is used for many daily aspects of Indonesian life as frying oil, chocolate butter, or even biodiesel; and these days there is an increasing demand for Crude Palm Oil (CPO) in international market. Indonesian palm oil sources are mainly from the species *Elaeis guineensis*. The other species that are commonly available around the world is *Elaeis oleifera,* which has smaller palm oil yield (Durand-Gasellin et al. 2005).

Palm oil has proved itself as an important commodity, in 2012 Indonesia became the largest consumer and producer of palm oil (Obidzinski et al. 2012). Indonesian palm oil production rose to 35% of global production in 2012, and a further increase to more than 50% in 2013, moreover there was an increase in global import trend of palm oil up to 42,7 billion USD and Indonesia managed to produce approximately 17,3 billion USD which is 43% of global needs. The Indonesian government saw this trend as a beneficial opportunity, which led to a plan to increase productivity up to 33 million tons in 2015 (McClanahan 2013; Sandi 2014; Armindya 2014).

For years, the Indonesian government has tried many ways to increase palm oil production. Various studies are available and have been used to try to increase the production rate of palm oil plantations such as importing *Elaidobious kamerunicus*, a certain type of insect which was mainly used in Southern Africa to help palm oil trees to pollinate. This method has been used since 1983 in Indonesia when the insect was first time introduced (Tuo et al. 2011).

However, there are no studies to differentiate male and female flower composition in one area of a plantation, even though it is a crucial aspect of producing palm oil. Palm oil trees are a monoecious plant with a tendency to become a temporal dioecious. As a monoecious plant, a palm oil tree has two flowers: female and male type. The female flower will be the one that produces palm oil fruit which later is treated with palm oil, while the male flower only takes part in the fertilization process. Moreover, as a temporal dioecious plant, each tree has mostly just one kind of flower and to increase productivity we cannot waste our time on the trees which do not produce palm oil, so we need to have more trees with mostly female flowers rather than the male flowers. Also, to differentiate between both flowers visually the flowers need to be fully grown. However, it will often be too late since a lot of expense and time have been spent in maintaining the plantation and it's impossible to just start over again. Farmers are then forced to accept the current harvest rate (Iswanto et al. 2010; Adam et al. 2011).

A study about genes taking part in controlling and developing the flowering organ had been conducted by Adam et al. (2007), which showed that the flowering process was controlled by *MADS-box* gene. In their study, Adam et al. used cDNA samples to characterize *MADS*-box gene of palm oil. Based on their study, a primer was designed and utilized to look for the correlation between DNA sequences of female and male flowering gene samples of palm oil at Indonesia's Agency for Assessment and Application of Technology (BPPT).

The aim of this study is to detect the distinction between female and male flowering gene sequences from DNA sample of *Elaeis guineensis* and then analyze it using bioinformatics software.

MATERIALS AND METHODS

Plant materials. The plant materials included staminate flowers that were collected from one oil palm plantation in Central Kalimantan, Indonesia and Puspiptek Serpong Botanic/Experimental Gardens, Banten, Indonesia. The flowers were produced within tenera, derived from dura and pisifera.

DNA isolation. A female and male flower from two different trees were gathered as a start for DNA isolation. Cetyl trimethyl ammonium bromide (CTAB) 1%, EDTA 0.5M, Tris HCl 1M pH8, and NaCl 5 M were mixed as extraction buffer. The other materials were Chloroform: Isoamyl 24: 1, Na-acetate, isopropanol, ddH$_2$O, and RNAse. First, the samples were weighed and crushed with liquid nitrogen and polyvinylpyrrolidone (PVP), then the samples were moved into falcon tube and extraction buffer was added. After that, the tube was soaked in the water bath at 65°C for 30 minutes and shook every 5 minutes. Next, the samples were put at room temperature, then C: I was added as much as samples volume. Then samples were centrifuged at 15572 G, 20 minutes 4°C, the supernatant was collected, and then Chloroform: Isoamyl solution was added. After that, the samples were re-centrifuged, repeated for 3 times. After that, the supernatant was removed. The pellets were dried under vacuum pump for 90 minutes. After that, the pellets were added with TE Buffer, NaCH$_3$COO, and chilled with absolute ethanol. Then the samples were kept at -20°C overnight. The next day, the samples were centrifuged at 1846 G for 5 minutes at 4 °C and re-dried at room temperature. Last ddH$_2$O and RNAse were added to samples and incubated at 37°C for 1 hour.

PCR amplification. The PCR amplification was conducted using GmG primer (F RCACTAAYAGCRCA / R TCACTTARTTCPCA), PCR modified master mix, and the PCR process is prepared as shown in Tables 1 and 2.

Visualization. The DNA visualization was conducted using electrophoresis instrument. The 1% agarose gel was made with a mixture of 0.1 g agarose and 100 mL TAE buffer. The samples were mixed with Tritrack 6x Loading dye and then were injected into the well. A 1 Kb DNA ladder marker was also injected into another well. The process was run at 100 volts for 30 minutes.

Sequencing. The amplified samples were sequenced using sequencer analysis by Biotech Centre, BPPT Puspiptek Serpong.

Table 1. PCR Master Mix.

Composition	Volume for 1X Reaction (µL)
2mM dNTP mix	1
Taq polymerase	0.1
Buffer Taq 10X	1
Primer GmG F	0.5
Primer GmG R	0.5
ddH$_2$O	5.9
DNA Sample	1
Total volume	10

Table 2. Thermocycler setting.

Stage	Temp. (°C)	Time (sec.)	
Pre-denaturation	95	300	-
Denaturation	95	30	35 cycles
Annealing	65 (±11)	30	35 cycles
Elongation	72	60	35 cycles
Final elongation	72	300	35 cycles
Hold	4	~	35 cycles

RESULTS AND DISCUSSION

The first part was DNA isolation of both samples, and the results of DNA isolation (Figure 1) show that the total DNA from both samples was successfully isolated. All bands are shown above the highest band of marker 1Kb DNA ladder, which is 10kb. It is proven that the whole DNA had been isolated from the total genome size of palm oil tree is 1.8 GB (Singh et al. 2013).

To determine the position of Globosa gene within the genome region, analyses were conducted using PerlPrimer 1.1.21. They used Spidey algorithm to predict the gene position with mRNA sequenced, using mRNA data with accession number AF411848.1 from the study by Adam et al. (2007), and whole DNA sequence, using a genome region data with accession number NW_011550962.1 from the study of Singh et al. (2013). According to *PerlPrimer* (application for designs primers, ORF and CpG island detection algorithms) analysis of globosa gene was separated into seven exons, between each exon are introns (Figure 2.). According to next analysis using Ugene (http://ugene.net/)(application for multiple sequence alignment and genome sequencing data analysis), the amplified region of GmG primer was stretched from the 4th exon to the 7th exon as shown in Figure 3.

The amplified male and female samples (#3 and #5 of genomic DNA, respectively) size were shown in an electrophoresis result as bands between 4th and 5th band of 1 Kb DNA ladder, which is between 1000 bp to 1500 bp (Figure S1.A), when analyzed using Ugene software, the size of the amplified region is 1125 bp (Figure 3). The amplified region of GmG primer is longer than the mRNA data, which is 897 bp because in the previous study they used RNA samples to analyze the characteristic of *MADS*-box mRNA sequence (Adam et al. 2006). However, in this study, the samples were all genome or DNA samples which explains the longer size of the amplified region. The mRNA sequences were divided into several exons within the DNA genome sequences. The primer is expected to attach at a certain position within the exons, in this study the forward primer was attached at 4th exon and the reverse primer was attached at 7th exon. Thus, the amplified region included introns (Figure 3). Optimization processes were also conducted and found that the most optimal temperature for GmG primer to anneal was 58.4°C as shown in Figure S1.B.

Since the result of samples, sequencing was not fully assembled, it is presented as forward and reverse of the primer sequence. Male forward and reverse sequences, also

female forward and reverse sequences are presented in Table S1. All sequences were aligned to the amplified genomic region of GmG primer. Male and Female forward sequences were aligned using pairwise alignment against the amplified region, and the results were compared between male forward (Figure 4) and Female forward (Figure 5).

The same process was done for male and female reverse (Figures 6 and 7). Even though the samples sequences were not aligned from the first sequence of the amplified region for forwarding sequence and the last part for reverse sequence, the alignment analysis proof that the samples were indeed part of globosa sequence by Adam et al. (2006). The samples sequencing result was not shown as fully sequenced, compared to the amplified region from the database caused by the presence of a repeated sequence, which caused problems to occur within sequencing process. The problem usually occurs two-fold. First, since the repeated sequence tends to be quite long, it may result in further incorporation for further sequence runs. This problem often leads to sequence petering out in the middle of the process. Second, the enzyme used in the process will often slip for certain number of times and the enzyme will re-join in a random position that would result in the complete sequence. Moreover, palm oil is a high GC content sample, thus it poses greater difficulty in sequence process if repeated sequences occur (Low et al. 2008; Treangen and Salzberg 2011).

From the result, a few gaps were found in both male-female forward sequences, and the gaps were shown in a different position. The male forward sequence's gap was found at the 269th base, while female forward sequence's gaps were found at the 270th and 311th base (Figures 4 and

5). Gaps were also found in both male and female reverse sequences (Figures 6 and 7). However, both samples showed the same gaps position, which was at the 1065th base. It shows that the differences between male-female sequences are all in the front part of the sequence, or at 5'end since the sequencing coding starts from 5' to 3'. This base position difference is a significant marker that shows the difference between male and female sequences. Even in just one base different, at least there would be one amino acid different; and in one amino acid different, there could be a different protein conformation which results in a significant expression difference (Spindel et al. 2013).
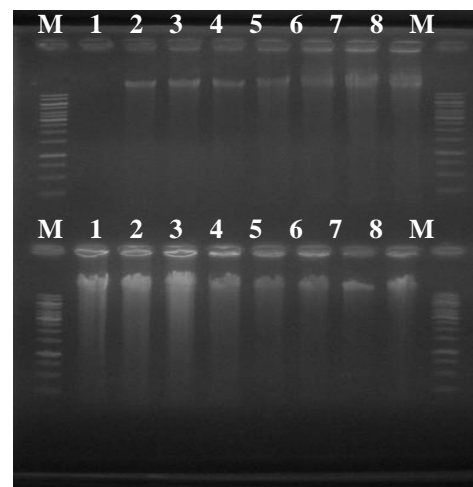

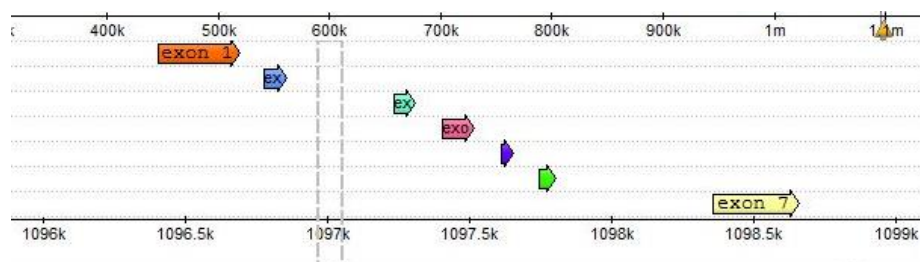**Figure 1.** DNA Isolation of male (*top*) and female (*bottom*) flower, M=DNA ladder, 1-8: samples.

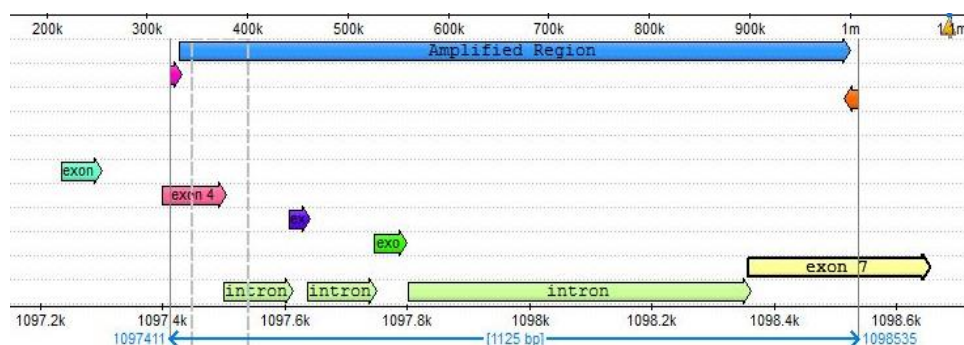**Figure 2.** The exons of globosa according to PerlPrimer analysis (taken from Ugene).

Figure 3. Amplified region of GmG primer within the genome region (taken from Ugene).
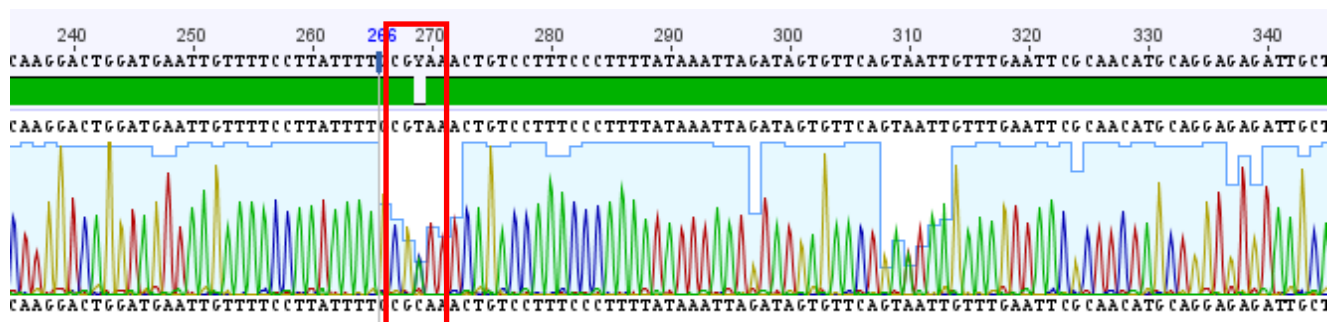
Figure 4. Male forward sequence compared to the amplified part of genome region NW_011220962 from NCBI database (red box: gap).
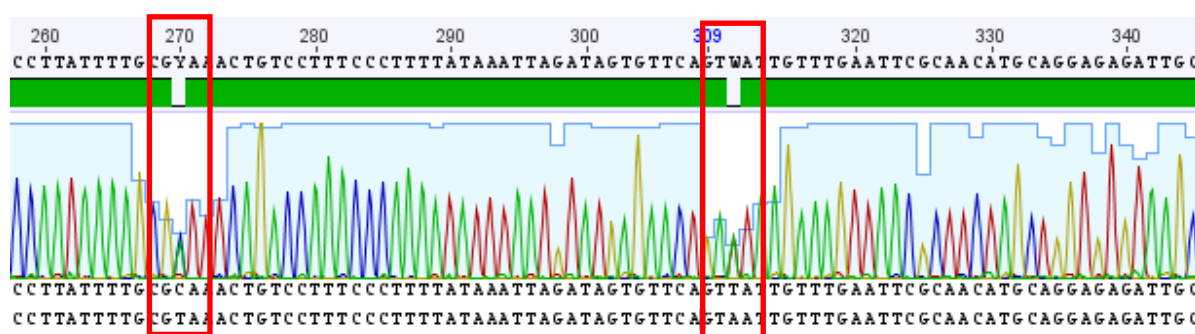

Figure 5. Female forward sequence compared to the amplified part of genome region NW_011220962 from NCBI database (red box: gaps)
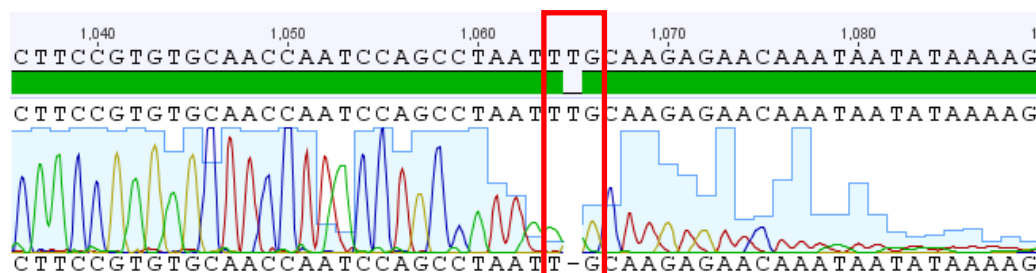

Figure 6. Male reverse sequence compared to the amplified part of genome region NW_011220962 from NCBI database (red box: gap)
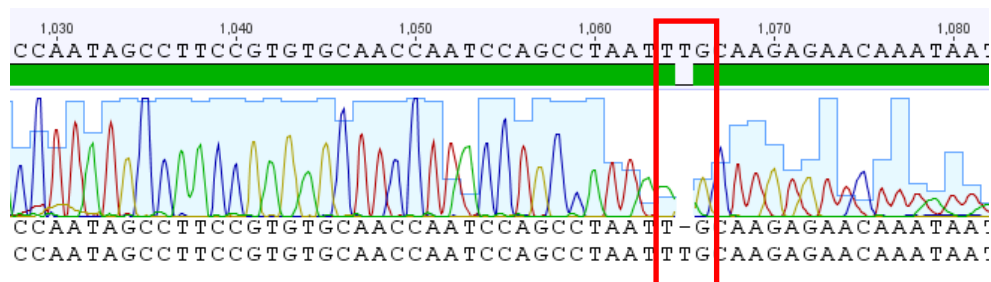

Figure 7. Female reverse sequence compared to the amplified part of genome region NW_011220962 from NCBI database (red box: gap).

These differences can be used as an initial study in determining the difference between a female and male flower. Though gaps sometimes posed as the effect of SNPs, a single nucleotide difference that usually appears in a different individual, a recent study showed that gaps could act as a marker and possess a great value for numerous applications in plant breeding and mapping. However, it would be beneficial in further studies to complete the full sequence and more samples to understand the difference between male and female flower sequences as whole, since the GmG primer could be used to design a specific marker or primer to detect the presence of female or male flower within a tree (Spindel et al. 2013).

Overall, this study shows the GmG primer can differentiate DNA sequence female and male flower of *Elaeis guineensis*, palm oil. As shown in the result of male and female forward sequences, there are gaps in both samples. However, the position of those gaps are in different positions, the male forward sequence at the 269th base and female forward sequence at the 270th and 311th base . This shows that the GmG primer has the potential to differentiate the DNA sequence of the female and male flowers. However, further studies with full sequence and more samples are needed to find distinctive results between female and male flower sequences as the GmG primer can be used to design a specific marker or primer to detect the presence of female or male flower within a tree, to increase the yield of Indonesian palm oil production.

ACKNOWLEDGEMENTS

REFERENCES

Adam H, Collin M, Richaud F, Beule T, Cros D, Omore A, Nodichao L, Nouy B, Tragear JW. 2011. Environmental regulation of sex determination in oil palm: current knowledge and insight from other species.Annals of Botany: 1-9 doi: 10.1093/aob/mcr151

Adam H, Jouannic S, Orieux Y, Morcillo F, Richaud F, Duval Y, Tregear JW. 2007. Functional characterisation of MADS box genes involved in the determination of oil palm flower structure. J Exp Bot 58 (6): 1245-1259.

Adam H, Jouannic S, Orieux Y, Morcillo F, Richaurd F, Duval Y, Tregear JW. 2006. MADS Box genes in oil palm (*Elaeis guineensis*): Patterns in the evolution of the SQUAMOSA, DEFICIENS, GLOBOSA, AGAMOUS, and SEPALLATA Subfamilies. J Mol Evol 1 (62): 15.

Armindya YR. 2014. Ini Prediksi Pasar Minyak Sawit Tahun Depan. Tempo. Analisa Bisnis. [BLTD] Biomatter LTD. 2015. Geneious 8.1 manual. http://assets.geneious.com/documentation/geneious/GeneiousManual.pdf [16 Apr 2015].

Durand-Gaselin T, Asmady H, Flori A, Jacquemard JC, Hayun Z, Breton F, de Franqueville H. 2005. Possible sources of genetic resistance in oil palm (*Elaeis guineensis* Jacq.) to basal stem rot caused by *Ganoderma boninense*--prospects for future breeding.Mycopathologia 159 (1): 93-100.

Iswanto AH, Sucipto T, Azhar I, Coto Z, Febrianto F. 2010. Physical and mechanical properties of palm oil trunk from Aek Pancur farming-North Sumatra. Jurnal Ilmu dan Teknologi Hasil Hutan 3 (1): 1-7.

Low EL, Alias H, Boon SH, Shariff EM, Tan CA, Ooi LCL, Cheah SC, Raha AR, Wan KL, Singh R. 2008. Oil palm (*Elaeis guineensis* Jacq.) tissue culture ESTs: Identifying genes associated with callogenesis and embryogenesis. BMC Plant Biol 8: 62. DOI: 10.1186/1471-2229-8-62

McClanahan P. 2013. Can Indonesia increase palm oil output without destroying its forest? Environmentalists doubt the world's biggest palm oil producer can implement ambitious plans without damaging woodland. The Guardian. 11 September 2013. https://www.theguardian.com/global-development/2013/sep/11/indonesia-palm-oil-destroy-forests

Obidzinski K, Adriani R, Komarudin H, Adrainto A. 2012. Environmental and social impacts of oil palm plantations and their implications for biofuel production in Indonesia. Ecol Society 17 (1): 25. DOI: 10.5751/ES-04775-170125

Sandi AP. 2014. Pemerintah Optimistis Pasar CPO Makin Baik. Tempo. 5 September 2014. https://bisnis.tempo.co/read/604663/pemerintah-optimistis-pasar-cpo-makin-baik [Indonesian]

Singh, R., M. Ong-Abdullah, E.-T. Low, M. A. A. Manaf, R. Rosli. 2013. Oil palm genome sequence reveals divergence of interfertile species in Old and New Worlds. Nature 500: 335-339

Spindel J, Wright M, Chen C, Cobb J, Gage J, Harrington S, Lorieux M, Ahmadi N, McCouch S. 2013. Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. Theor Appl Genet 126 (11): 2699-2716.

Treangen TJ, Salzberg SL. 2011. Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat Rev Genet 13 (1): 36-46.

Tuo Y, Koua HK, Hala N. 2011. Biology of *Elaeidobius kamerunicus* and *Elaeidobius plagiatus* (Coleoptera: Curculionidae) main pollinators of oil palm in West Africa. Eur J Sci Res 49 (3): 426-432.

Table S1. Samples sequences

The following is Male forward sequence.
>JG_F
GAATCTTTCTATTGAGGTGTCCTCCAGAATGGTCTCACCAGTGT
TAGGGGCAAGCAGGTTGGTATATGTTTGTGCCTTCTTTCTGTTT
GATTGATCATGTAAATCCTGTGATATGTTTCCTTTGTGCATGGG
AATGTTACTGATTAGTTCTTTTACTATTCCAGATGGATTTCTTG
AAGAAGCTCAAGAAGAATGTATGATTCAAGGACTGGATGAATTG
TTTTCCTTATTTTGCGCAAACTGTCCTTTCCCTTTTATAAATTA
GATAGTGTTCAGTAATTGTTTGAATTCGCAACATGCAGGAGAGA
TTGCTGGAAGAGGAGAACAAGCATCTGACTTTTCTATTGGTAAT
CATCTAATCGGATGAGCTTTTGAGTTTCAAGTTATAGCAGAGGA
TGCACTTTTTATGCATGGATGGCAAAATTATAAGCATCATGATA
AACTCAACGCAGTGTGATATGTAGATAGTGATAGTAAGGAGCTA
CTTGGAGCTTTTACTTATTTGATATTTGGCACTCCACTTTACCA
TTTTTCATCTGGGACCATGCTTTTGTTGAGCCAGTTGGAGTGTT
CATTGTGATTGAAATTTTAAAGCT

The following is male reverse sequence.
>JG_R
AAAGGGGCCTTTTATATTATTTGTTCTCTTGCAATTAGGCTGGA
TTGGTTGCACACGGAAGGCTATTGGAATCTGGCATGCAAAATCC
CTAGCTTTCAGAGGATCCCCAAGCTCCAATTCCCTTACATTTCC
ATTCATTGCCAATTCCTGGTGATGCTGTGAAAAGAAAGATGATT
ATTTCAGTTAAGAAAGAGAGAAAGACTGTTGAGTTTTTGGTAAT
GTCCACAATAAGCACTAACTGCACTATCTCACATTTGGCAAATT
CTTACATGGTATAGTAACCAAAATGATGCTAATATCATGACTTC
TGAGTGATACCATTACATGGAGCAACCAAAGTTATGCGTGGTGT
GAGGAAGGCAGATGCAGCAGAAACAAAGATGATGCCGCTATAGT
ATGCACATGACCTTGATGCATCCATAGTCTAAAACTTTACAAGA
CAGTCAAATCATGGGAAAGAGTTGATAGAGCTTTAAAATTTCAA
TCACAATGAACACTCCAACTGGCTCAACAAAAGCATGGTCCCAG
ATGAAAAATGGTAAAGTGGAGTGCCAAATATCAAATAAGTAAAA
GCTCCAAGTAGCTCCTTACTATCACTAT

The following is female forward sequence
>BG_F
CTGGAGGGATGTCCTCCAGAATGGTCTCACCAGTGTTAGGGGCA
AGCAGGTTGGTATATGTTTGTGCCTTCTTTCTGTTTGATTGATC
ATGTAAATCCTGTGATATGTTTCCTTTGTGCATGGGAATGTTAC
TGATTAGTTCTTTTACTATTCCAGATGGATTTCTTGAAGAAGCT
CAAGAAGAATGTATGATTCAAGGACTGGATGAATTGTTTTCCTT
ATTTTGCGCAAACTGTCCTTTCCCTTTTATAAATTAGATAGTGT
TCAGTTATTGTTTGAATTCGCAACATGCAGGAGAGATTGCTGGA
AGAGGAGAACAAGCATCTGACTTTTCTATTGGTAATCATCTAAT
CGGATGAGCTTTTGAGTTTCAAGTTATAGCAGAGGATGCACTTT
TTATGCATGGATGGCAAAATTATAAGCATCATGATAAACTCAAC
GCAGTGTGATATGTAGATAGTGATAGTAAGGAGCTACTTGGAGC
TTTTACTTATTTGATATTTGGCACTCCACTTTACCATTTTTCAT
CTGGGACCATGCTTTTGTTGAGCCAGTTGGAGTGTTCATTGTGA
TTGAAATTTTAAAGCTC

The following is female reverse sequence.
>BG_R
TTGGGGACTTTTATATTATTTGTTCTCTTGCAATTAGGCTGGAT
TGGTTGCACACGGAAGGCTATTGGAATCTGGCATGCAAAATCCC
TAGCTTTCAGAGGATCCCCAAGCTCCAATTCCCTTACATTTCCA
TTCATTGCCAATTCCTGGTGATGCTGTGAAAAGAAAGATGATTA
TTTCAGTTAAGAAAGAGAGAAAGACTGTTGAGTTTTTGGTAATG
TCCACAATAAGCACTAACTGCACTATCTCACATTTGGCAAATTC
TTACATGGTATAGTAACCAAAATGATGCTAATATCATGACTTCT
GAGTGATACCATTACATGGAGCAACCAAAGTTATGCGTGGTGTG
AGGAAGGCAGATGCAGCAGAAACAAAGATGATGCCGCTATAGTA
TGCACATGACCTTGATGCATCCATAGTCTAAAACTTTACAAGAC
AGTCAAATCATGGGAAAGAGTTGATAGAGCTTTAAAATTTCAAT
CACAATGAACACTCCAACTGGCTCAACAAAAGCATGGTCCCAGA
TGAAAAATGGTAAAGTGGAGTGCCAAATATCAAATAAGTAAAAG
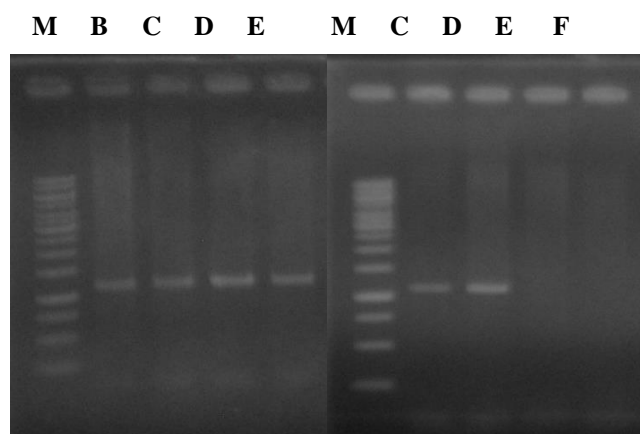CTCCAAGTAGCTCCTTACTATCACTATCTAC

Figure S1.A. Results of PCR amplification and GmG primer temperature optimization. Male (*left*) and female (*right*) samples amplification results, the samples were labeled to represent different temperature gradient B: 55.8 °C, C: 57.1 °C, D: , 58.4 °C, and E: 59.7 °C
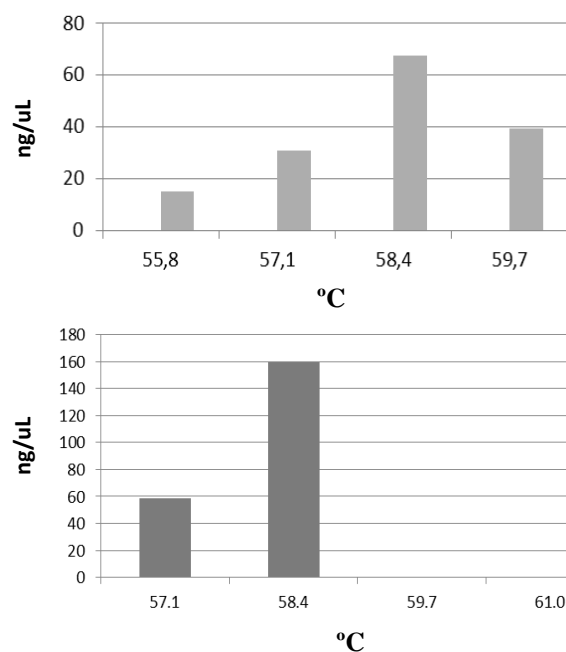

Figure S1.B. Male (*top*) and female (*bottom*) flower samples temperature optimization result.