COMPARISON OF NCBI PRIMER-BLAST AND PRIMER3PLUS FOR DETERMINATION DNA PRIMER DESIGN OF TYPHONIUM FLAGELLIFORME PLANT

Teddy Siswanto[1,3,*], Nesti Fronika Sianipar[2]
Harco Leslie Hendric Spits Warnars[1] and Harjanto Prabowo[1]

[1]Department of Computer Science, BINUS Graduate Program – Doctor of Computer Science
[2]Department of Food Technology, Faculty of Engineering
Bina Nusantara University
Jl. K. H. Syahdan No. 9, Kemanggisan, Palmerah, Jakarta 11480, Indonesia
{ nsianipar; shendric; harprabowo }@binus.edu

[3]Information System Study Program
Trisakti University
Jalan Kyai Tapa No. 1, Grogol, Jakarta Barat 11440, Indonesia
*Corresponding author: teddysiswanto@trisakti.ac.id

ABSTRACT. *Research on the Typhonium Flagelliforme plant, which originally used the In Vitro and In Vivo methods, is now starting to be synergized with the Bioinformatics method. The amplification process in the Polymerase Chain Reaction will determine the discovery of important compounds in herbal medicine. DNA Primer Design is very influential in the success of the Polymerase Chain Reaction because if the results of the DNA Primer Design are not accurate, then the results of further processing of the amplification process must be repeated so that it is inefficient and ineffective. Tools that are available and often used in determining DNA Primer Design are NCBI Primer-Blast and Primer3Plus. The purpose of this study is to compare the results of the two tools in default setting. The research methodology using the method-comparison is to compare the results of the two tools with the theory in the literature review on aspects of the requirements for Start Codon, Stop Codon, GC quality, melting temperature, difference in temperature melting forward and reverse primer and length of DNA Primer base pair. This research resulted in 5 outputs of base pair options produced by Primer3Plus are considered less than 50% of the quality aspect of Guanine and Cytosine, while 4 of 10 DNA Primer Design base pair options produced by NCBI Primer-Blast meet the specified theory of the literature review.*
Keywords: DNA Primer Design, NCBI Primer-Blast, Primer3Plus, Method-comparison, Typhonium Flagelliforme

1. **Introduction.** The Typhonium Flagelliforme plant is being developed into an anti-cancer herbal ingredient. The results of previous studies have found compounds that contain high antioxidants as anti-cancer ingredients in ethanol compounds, namely tannins, flavonoids and steroids [1-4], lectins [5,6] and stigmasterols [7]. DNA sequence design has a vital role in DNA computing to produce useful chemical compounds [19]. The process of obtaining compounds that have potential as drugs is through a process called the Polymerase Chain Reaction which is often abbreviated as PCR, which is a molecular biology technique to increase the number of copies of a particular DNA chain for further research. So the purpose of PCR is to meet certain DNA targets so that researchers can analyze them. However, to be able to obtain DNA strands that can be duplicated, there are certain requirements in order to be used as a DNA Primer Design. Amplification of target

DNA sequences with Guanine and Cytosine content exceeding 60% can be an obstacle in the success of the PCR process. This is further complicated by the high possibility of formation of secondary structures such as hairpins on the DNA template, formation of dimer primers and primary crosslinking [8]. In addition, genome cloning of high Guanine and Cytosine content by PCR using long primers may pose additional challenges for the amplification process due to the high melting temperature due to the extra hydrogen bonds between Guanine and Cytosine bases. There are several DNA Primer Design tools that researchers often use to obtain DNA primer designs from DNA sequence chains, including Primer3Plus and NCBI Primer-Blast. Both of these tools are Open Sources software that are free to use in research. The two tools have different technical characteristics, so this study aims to compare the results of data processing using the two tools which is more suitable for research needs to obtain DNA primer designs for further processing in PCR amplification. However, it is very open to make new findings in the application of the algorithm so as to be able to find the potential for DNA Primer Design which is better than the two tools tested in this study.

2. **Problem Statement and Preliminaries.** The problem faced by the researchers is how to obtain the primary design DNA chain from raw DNA sequence data in FASTA format so that it can meet the specified requirements. Without DNA primers, the PCR reaction will not occur even though the enzymes are already available. The nucleotide base sequence in the DNA Primer Design must be correctly matched to the base sequence of the target DNA to be amplified and not attached to the non-target region. The good DNA Primer Design is a base pair length of 18-25, the temperature difference between the Forward and Reverse Primer melting is not too far max 5°C, and the quality of Guanine and Cytosine is between 50% to 60% [8,9,17]. The stages of preparation carried out are

a) Raw Data, obtained from research on control plant Typhonium Flagelliforme with file name 1st_BASE_4153159_Bogor Control__B1_upper_OPB18 [10], where research continues to develop to date and was obtained from the researcher who breeds the Rodent Tuber or Typhonium Flagelliforme plant which has obtained Plant Variety Protection certification according to the Decree on Breeding Result Varieties No. 583/PVHP/2018 and No. 584/PVHP/2018 from Indonesian Ministry of Agriculture for Tipobio and Typonesiaraga plants.

b) The Primer3Plus tool, https://primer3plus.com/cgi-bin/dev/primer3plus.cgi, is an online software that can be accessed via the Internet. This tool has several parameters that can be set as needed, but in this study the default settings are used.

c) NCBI Primer-Blast tool, https://www.ncbi.nlm.nih.gov/tools/primer-blast/, is an online tool that can be accessed via the Internet. This tool has several parameters that can be set as needed, but in this study the default settings are used.

3. **Main Results.** The research methodology used in this research is a method-comparison of the results of DNA Primer Design from the two tools to the theory obtained from the literature review, as shown in Figure 1.

Raw DNA in FASTA format was obtained from a study entitled embryonic calli induction, proliferation, and regeneration of rodent tuber plant (Typhonium Flagelliforme lodd.) by single node culture [10] which is entered as input data in the Primer3Plus and NCBI Primer-Blast tools. The results of both processes were recorded for Forward Primer and Reverse Primer and then added with DNA non-reverse primer, melting temperature and quality of Guanine and Cytosine. Based on the literature review, the search for DNA Primer Design characteristics for initial reading of DNA sequences or Start Codon is "ATG" [11-15]; Melting Temperature (Tm) in range between 55°C-65°C [16]; Tm difference <= 5°C [16,18]; composition GC 50%-60% [17]; Length DNA Primer 25 bp [17], 18-22 bp [18] will be used to analyze the output of the two tools.
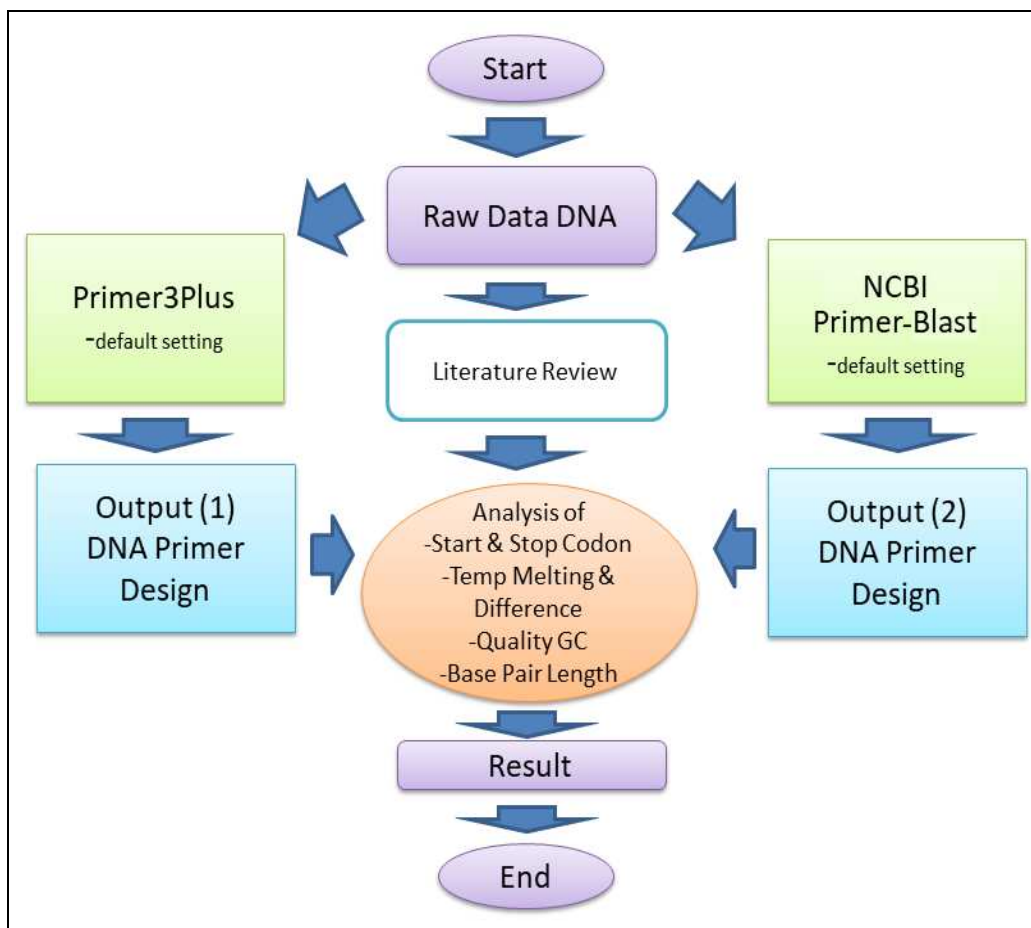
Figure 1. Research methodology

4. Control Design.

4.1. **Raw data.** Raw data FASTA format needs to be identified sequence length, composition of Adenine, Guanine, Thymine and Cytosine as well as non-bases in the raw data, by using a Python script. Characteristics of raw data can be seen in Figure 2.

Figure 2. Characteristics of DNA raw data

DNA raw data has a length of 479 base pairs of which 5 of them are non-base. If it is calculated as a percentage of $5/479 \times 100\% = 1.04\%$, this is relatively small so that it does not interfere with further processes.

4.2. **Primer3Plus.** DNA raw data is entered in the source sequence in Primer3Plus under the default settings as shown in Figure 3.

After selecting Pick Primer, 5 pairs of DNA Primer Design appear and what appears in Figure 4 is Pair-1 of the 5 pairs produced. The first shaded pair is Left Primer-1 or
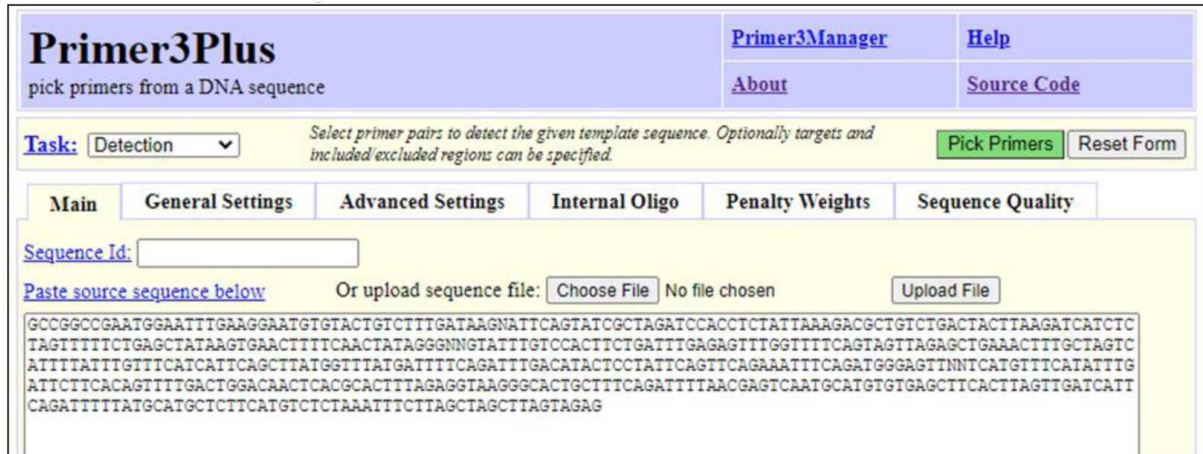
FIGURE 3. DNA raw data in Primer3Plus

FIGURE 4. Forward and Reverse Primer of Pair-1 Primer3Plus

Forward Primer and the second shaded pair is the pair before the reverse, after the reverse it will appear in the Right Primer-1 or Reverse Primer.

4.3. **Output and analysis of DNA Primer Design of Primer3Plus.** The complete Primer3Plus processing results from Pair-1 to Pair-5 are made as shown in Table 1 with information on melting temperature and quality of Guanine and Cytosine as well as additional DNA pair data before being reversed. Based on the quality of GC, from Pair-1 to Pair-5, the two primary pairs are always below 50% so they do not meet the required conditions. Based on the difference in melting temperature, everything is less than 1°C and this is in accordance with the provisions. Based on the ATG Start Codon, there are only Pair-4 and Pair-5, while based on the TAA, TAG and TGA Stop Codon all Pairs meet the requirements.

4.4. **NCBI Primer-Blast.** DNA raw data is entered into the FASTA sequence field in the NCBI Primer-Blast with the default setting conditions as shown in Figure 5.

4.5. **Output and analysis of DNA Primer Design of NCBI Primer-Blast.** After getting primer, 10 DNA Primer Design options appear that can be considered for use in the amplification process for PCR. The result of Pair-1 looks like in Figure 6, which consists of Forward Primer, Reverse Primer, base pair length, Start and Stop Codon starting from the start number position and ending at the stop number position, melting temperature and GC quality.

TABLE 1. Processed results of Pair-1 to Pair-5 from Primer3Plus

Pair	Primer	DNA Primer Design	Length	Tm	GC
1	Forward	CCACTTCTGATTTGAGAGTTTGG	23	60.2	43.5%
	Reverse	GCGTGAGTTGTCCAGTCAAA	20	59.9	50%
		TT TGA CTGGACAACTCACGC			
2	Forward	CCACTTCTGATTTGAGAGTTTGG	23	60.2	43.5%
	Reverse	TGAAAGCAGTGCCCTTACCT	20	59.9	50%
		AGG TAA GGGCACTGCTTTCA			
3	Forward	CCACTTCTGATTTGAGAGTTTGG	23	60.2	43.5%
	Reverse	CTGAAAGCAGTGCCCTTACC	20	59.9	55%
		GG TAA GGGCACTGCTTTCAG			
4	Forward	CCGA ATG GAATTTGAAGGAA	20	59.9	40%
	Reverse	CCAAACTCTCAAATCAGAAGTGG	23	60.2	43.5%
		CCACTTC TGA TT TGA GAGTTTGG			
5	Forward	TTGTTTCATCATTCAGCTT ATG G	23	59.1	34.8%
	Reverse	CTGAAAGCAGTGCCCTTACC	20	59.9	55%
		GG TAA GGGCACTGCTTTCAG			

FIGURE 5. DNA raw data in NCBI Primer-Blast

FIGURE 6. Pair-1 output results of DNA Primer Design of NCBI Primer-Blast

TABLE 2. Pair-1 to Pair-5 DNA Primer Design result of NCBI Primer-Blast process

Pair	Forward & Reverse Primer	Tm	GC%	Raw DNA Data
1	CCGGCCGAATGGAATTTGAA AAGTGCGTGAGTTGTCCAGT *ACTGGACAACTCACGCACTT*	59.19 59.82	50.00 50.00	
2	TTGACTGGACAACTCACGCA GCTCACACATGCATTGACTCG *CGAGTCAATGCATGTGTGAGC*	59.82 60.20	50.00 52.38	
3	CGGCCGAATGGAATTTGAAGG AGTGCGTGAGTTGTCCAGTC *GACTGGACAACTCACGCACT*	60.20 60.25	52.38 55.00	
4	CCGGCCGAATGGAATTTGAAG TGAAAGCAGTGCCCTTACCT *AGGTAAGGGCACTGCTTTCA*	60.20 59.23	52.38 50.00	
5	GACTGGACAACTCACGCACT AGCTCACACATGCATTGACTC *GAGTCAATGCATGTGTGAGCT*	60.25 58.92	55.00 47.62	

10 DNA Primer Design options results of NCBI Primer-Blast process are shown in Table 2 by adding a base pair before being reversed to ensure the position in the DNA raw data. Pair-1 to Pair-5 based on the difference in melting temperature not more than 1.33°C, this meets the requirements. Based on GC quality only Pair-5 does not meet the requirements below 50%. However, if completed based on the ATG Start Codon and TAA or TAG or TGA Stop Codon, only Pair-4 meets the requirements.

Based on Table 3, Pair-6 to Pair-10 based on the difference in melting temperature not more than 2.85°C, this meets the required conditions. Based on GC quality only Pair-7 and Pair-9 do not meet the requirements below 50%. However, if completed based on the Start Codon ATG and Stop Codon TAA or TAG or TGA there are 3 Pair-6, Pair-8 and Pair-10 that meet the requirements.

4.6. **Result analysis.** If it is identified how many "ATG" Start Codons are in the raw data, 11 codons are found as shown in Figure 7, so that the NCBI Primer-Blast looks more suitable to be used as a DNA Primer Design tool; however, out of 10 (10/11 = 91%) obtained only 4 (4/10 = 40%) meet the theoretical requirements. From these results, it is still possible for a new algorithm model to be able to obtain up to a maximum of 11 codons as shown in Figure 7.

5. **Conclusions.** Based on the default settings without parameter changes and with raw data DNA FASTA 1st_BASE_4153159_Bogor Control_B1_upper_OPB18 of Typhonium Flagelliforme plant, it can be concluded:

1) With the Primer3Plus tool, 5 DNA Primer Design options are generated, although the Start and Stop Codon and the melting temperature difference meet the requirements, the quality aspect of the GC is still below 50%.
2) With the NCBI Primer-Blast tool, 10 DNA Primer Design options are generated, 4 of the 10 options meet the requirements of all requirements, namely Pair-4, Pair-6, Pair-8 and Pair-10.

Table 3. Pair-6 to Pair-10 DNA Primer Design result of NCBI Primer-Blast process

Pair	Forward & Reverse Primer	Tm	GC%	Raw DNA Data
6	GCCGGCCGAATGGAATTTGA GCCCTTACCTCTAAAGTGCGT *ACGCCTTTTAGAGGTAAGGGC*	61.66 60.07	55.00 52.38	
7	GGCCGAATGGAATTTGAAGGA TGCCCTTACCTCTAAAGTGCG *CGCACTTTTAGAGGTAAGGGCA*	58.90 60.07	47.62 52.38	
8	CGGCCGAATGGAATTTGAAG AAAGCAGTGCCCTTACCTCT *AGAGGTAAGGGCACTGCTTT*	57.81 58.93	50.00 50.00	
9	GGCCGAATGGAATTTGAAGGAA TGAAAGCAGTGCCCTTACCTC *GAGGTAAGGGCACTGCTTTCA*	59.50 60.27	45.45 52.38	
10	GGCCGAATGGAATTTGAAGG TCTGAAAGCAGTGCCCTTACC *GGTAAGGGCACTGCTTTCAGA*	57.42 60.27	50.00 52.38	

```
In [6]:  # start Codon ATG
         genome = readGenome('C:/Users/Dell/coding/4153159_Bogor.FASTA')
         # print(genome)
         t = 'ATG'
         naive(t, genome)
         posisi_ATG = naive(t, genome)
         print("Position:",posisi_ATG)
         vjumlah=len(posisi_ATG)
         print("Count:",vjumlah)

         Position: [9, 24, 238, 245, 291, 304, 392, 396, 433, 437, 446]
         Count: 11
```

Figure 7. Number of Start Codon "ATG"

3) The potential to find DNA Primer Design through NCBI Primer-Blast was 91%, while with Primer3Plus it was 45%. There are still opportunities for further research to improve accuracy through algorithm improvements.

REFERENCES

[1] B. P. Priosoeryanto, R. Rostantinata, E. Harlina, W. Nurcholis, R. Ridho and L. N. Sutardi, In vitro antiproliferation activity of Typhonium Flagelliforme leaves ethanol extract and its combination with canine interferons on several tumor-derived cell lines, *Veterinary World*, vol.13, no.5, pp.931-939, DOI: 10.14202/vetworld.2020.931-939, 2020.

[2] N. F. Sianipar, Y. E. H. Putri, K. Assidqi, P. Simanjuntak and R. Purnamaningsih, A study of anticancer activity from the fractions of rodent tuber superior mutant extract (Typhonium Flagelliforme) by prestoblue assay method, *Rasayan Journal of Chemistry*, vol.13, no.3, pp.1993-1998, 2020.

[3] N. F. Sianipar, R. Purnamaningsih, I. Darwati and D. Laurent, Gas chromatography-mass spectrometry (GC-MS) analysis of phytochemicals of first generation gamma-irradiated Typhonium Flagelliforme lodd. Mutants, *Jurnal Teknologi*, vol.78, no.10-4, pp.1-7, 2016.

[4] T. Ohi-Toma, S. Wu, S. R. Yadav, H. Murata and J. Murata, Molecular phylogeny of Typhonium sensu lato and its allied genera in the tribe Areae of the subfamily Aroideae (Araceae) based on sequences of six chloroplast regions, *The American Society of Plant Taxonomists, Systematic Botany*, vol.35, no.2, pp.244-251, 2010.

[5] S. Naik and S. Kumar, Applications of plant lectins in biotechnology and therapeutics: Plant lectins application, *Journal of Microbiology, Biotechnology and Food Sciences*, vol.11, no.4, e4224, DOI: 10.55251/jmbfs.4224, 2022.

[6] M. Alfarabi, S. Rosmalawati, M. Bintang, Miftahudin, E. Rofa'ani and Chaidir, Antiproliferation activity of tuber protein from Typhonium Flagelliforme (Lodd.) blume on MCF-7 cell line, *International Journal of Biosciences*, vol.6, no.12, pp.52-60, DOI: 10.12692/ijb/6.12.52-60, 2015.

[7] N. F. Sianipar, Y. E. Hadisaputri, K. Assidqi, S. Salam, M. Yusuf, W. Destiarani, R. Purnamaningsih and I. G. So, Characterization and investigation of stigmasterol isolated from rodent tuber mutant plant (Typhonium Flagelliforme), its molecular docking as anticancer on MF-7 cells, *Preprints*, DOI: 10.20944/preprints202107.0278.v1, 2021.

[8] N. Assal and M. Lin, PCR procedures to amplify GC-rich DNA sequences of Mycobacterium bovis, *Journal of Microbiological Methods*, vol.181, 106121, DOI: 10.1016/j.mimet.2020.106121, 2021.

[9] S. Banoon, *PCR Primer Design Guidelines*, DOI: 10.13140/RG.2.2.15697.58726, 2018.

[10] N. F. Sianipar, Rustikawati, W. Maarisit, A. Wantho and D. N. R. Sidabutar, Embryonic calli induction, proliferation, and regeneration of rodent tuber plant (Typhonium Flagelliforme Lodd.) by single node culture, *International Conference on Biological Science*, Faculty of Biology, Universitas Gadjah Mada, Yogyakarta, Indonesia, pp.80-88, 2011.

[11] T. W. Wan, W. Higuchi, O. E. Khokhlova, W. C. Hung, Y. Iwao, M. Wakayama and T. Yamamoto, Genomic comparison between staphylococcus aureus GN strains clinically isolated from a familial infection case: IS1272 transposition through a novel inverted repeat-replacing mechanism, *PLoS One*, vol.12, no.11, e0187288, DOI: 10.1371/journal.pone.0187288, 2017.

[12] M. Labuschagne, F. Beugnet, S. Rehbein, J. Guillot, J. Fourie and D. Crafford, Analysis of Dipylidium caninum tapeworms from dogs and cats, or their respective fleas, *Parasite*, vol.25, 30, DOI: 10.1051/parasite/2018028, 2018.

[13] J. Yu, J. An, Y. Li and C. B. Boyko, The first complete mitochondrial genome of a parasitic isopod supports Epicaridea Latreille, 1825 as a suborder and reveals the less conservative genome of isopods, *Systematic Parasitology*, vol.95, no.5, pp.465-478, DOI: 10.1007/s11230-018-9792-2, 2018.

[14] B. Cadenas, J. Fita-Torró, M. Bermúdez-Cortés, I. Hernandez-Rodriguez, J. L. Fuster, M. E. Llinares and M. Sanchez, L-ferritin: One gene, five diseases; from hereditary hyperferritinemia to hypoferritinemia – Report of new cases, *Pharmaceuticals*, vol.12, no.1, 17, DOI: 10.3390/ph12010017, 2019.

[15] J. Batki, J. Schnabl, J. Wang, D. Handler, V. I. Andreev, C. E. Stieger and J. Brennecke, The nascent RNA binding complex SFiNX licenses piRNA-guided heterochromatin formation, *Nature Structural and Molecular Biology*, vol.26, no.8, pp.720-731, DOI: 10.1038/s41594-019-0270-6, 2019.

[16] F. D. Wahyuni, Seprianto and H. Saraswati, Detection of cryII gene from Bacillus thuringiensis using Polymerase Chain Reaction (PCR), *IOP Conference Series: Earth and Environmental Science*, East Java, Indonesia, vol.743, no.1, 012093, DOI: 10.1088/1755-1315/743/1/012093, 2021.

[17] H. Y. Wang, J. J. Lu, C. Y. Chang, W. P. Chou, J. C. H. Hsieh, C. R. Lin and M. H. Wu, Development of a high sensitivity TaqMan-based PCR assay for the specific detection of Mycobacterium tuberculosis complex in both pulmonary and extrapulmonary specimens, *Scientific Reports (Nature Publisher Group)*, vol.9, no.1, pp.1-12, DOI: 10.1038/s41598-018-33804-1, 2019.

[18] J. Cannon, R. A. Sanford, L. Connor, W. H. Yang and J. Chee-Sanford, Optimization of PCR primers to detect phylogenetically diverse *nrfA* genes associated with nitrite ammonification, *Journal of Microbiological Methods*, vol.160, pp.49-59, DOI: 10.1016/j.mimet.2019.03.020, 2019.

[19] J. Li, Q. Zhang, R. Li and S. Zhou, Optimization of DNA encoding based on combinatorial constraints, *ICIC Express Letters*, vol.2, no.1, pp.81-88, 2008.