Prospects for Primary Design Study of Protease Encoding Genes for Recombinant Protease Enzymes - A Bibliographic and Journal Review

Rima Agnes Widya Astuti[1], Nanik Rahmani[2], Dewi Seswita Zilda[3], Sri Darmawati[4], Stalis Norma Ethica[5]

[1,3,4,5]Universitas Muhammadiyah Semarang, Semarang, Central Java 50273, Indonesia
[2]National Research and Innovation Agency, Central Jakarta, DKI Jakarta 10340, Indonesia
norma@unimus.ac.id

Abstract. Protease enzymes in the health sector are very important, especially as anticoagulant, antithrombotic agents, as diagnostic tools and plasmin standards. It is not surprising that the need for protease enzymes in the world is very high, including in Indonesia. The supply of protease in Indonesia still depends on imports so its availability is limited. One way to overcome this problem is to produce protease enzymes from Indonesia's biological resources. Previous research reported that the bacteria Bacillus sp. HSFI-3 isolated from fermented digestive organs of Holothurian scabra sand sea cucumbers (West Nusa Tenggara) has proteolytic, fibrinolytic and thrombolytic characteristics. Bacillus sp. HSFI-3 is a wild type bacteria. The challenge with wild type bacteria as protease producers is that the gene promoter is not always on so it does not always express protease, the enzyme activity is low and it is not pure. Strategies need to be implemented to overcome this, one of which is cloning and overexpressing proteases for scale up efforts from wild type Bacillus sp bacteria. HSFI-3. The steps to carry out cloning and overexpression require the full-length gene sequence encoding the protease enzyme from Bacillus sp. HSFI-3. Full-length gene sequences can be obtained from the results of Whole Genome Sequencing (WGS) of Bacillus sp. HSFI-3 amplified from specific primer design. Full-length gene encoding the protease enzyme from Bacillus sp. After obtaining HSFI-3, it was then continued by amplifying the full-length gene using in silico PCR and in vitro PCR with specific primer designs based on the WGS sequence results. This journal review summarizes the types of primer design software specific to protease genes throughout the world in the last decade. This research specifically targets the software used in the design of specific primers for the full-length protease gene. The aim is to obtain the most appropriate software recommendations regarding specific primer design for the full-length protease encoding gene from Bacillus sp. HSFI-3. Based on the results of the journal review, all the software contained in this study only amplifies fragments. Benchling software (https://www.benchling.com/) is capable of designing primers for full-length genes. So we recommend this software for the design of full-length gene specific primers for the protease gene of Bacillus sp. HSFI-3 is an effort to scale up protease enzymes so that the need for protease enzymes in Indonesia is met.

G. Setya Ayu Putri et al. (eds.), *Proceedings of the 2nd Lawang Sewu International Symposium on Health Sciences: Medical Laboratory Technology (LSISHS-MLT 2023)*, Advances in Biological Sciences Research 40,
https://doi.org/10.2991/978-94-6463-457-0_2

Keywords: Specific Primer Design Software, Protease Genes, Recombinant Protease Enzymes, Cloning and Overexpression, Bacillus sp. HSFI-3.

1. Introduction

The importance of protease enzymes in the health sector is that they are widely used in the treatment of cardiovascular disease (CVD), anticoagulants, antithrombotic drugs, diagnostic tools and plasmin [1][2][3][4]. Plasmin is the active form of plasminogen from the serine protease gene which plays a role in breaking down blood clots caused by fibrin [3].

The need for protease enzymes in the world reaches 70% of all enzymes needed in various industrial sectors, including health. The demand for protease enzymes in Indonesia is also increasing and this need still depends on imported production so the availability of this protease enzyme is very important. Data from the Agency for the Assessment and Application of Technology (BPPT) in 2015, almost 99% of the need for protease enzymes in Indonesia was still imported [5][6][7].

One way to overcome the problem of dependence on imported protease enzymes is by producing protease enzymes from Indonesia's biological resources [5][6]. The Holothurian scabra sand sea cucumber (H. scabra) is one of the native biological resources of Indonesian marine waters (West Nusa Tenggara) which produces protease enzymes in the rusip products of its digestive organs producing the bacteria Bacillus sp. HSFI-3, this bacterium has proteolytic, fibrinolytic and thrombolytic characteristics. In the thrombolysis activity test on the lysis of blood clots in vitro using the gravimetric method, Bacillus sp. HSFI-3 is able to lyse blood clots with a percentage of 26.06% [4]. Bacillus sp. HSFI-3 is a wild type bacteria. The challenge with wild type bacteria is that the gene promoter is not always on so it does not always express protease, the enzyme activity is low and it is not pure. So strategies need to be implemented to overcome this, one of which is through cloning and overexpression [8].

Efforts to scale up the anti-thrombosis fibrinolytic protease Bacillus sp. HSFI-3 via cloning and overexpression has not been reported. So it is important to carry out research related to the recombinant protease enzyme Bacillus sp. HSFI-3 with genetic engineering techniques. Genetic engineering is a technology for manipulating microbial genes in order to regulate gene expression levels and the production of certain compounds so that they can produce useful products [9]. The advantages of genetic engineering are that the enzymes produced have good stability and durability, consistent product performance, and increased production capacity (www.scrippslabs.com) [8].

Full-length gene encoding the protease enzyme from Bacillus sp. HSFI-3 needs to know the sequence first before carrying out genetic engineering. Efforts to obtain full-length gene

sequences can be done through WGS on the Bacillus sp. HSFI-3 genome. Full-length gene encoding the protease enzyme from Bacillus sp. HSFI-3 once was obtained, it was then continued by amplifying the full-length gene using in silico PCR and proven by in vitro PCR amplification with a specific primer design based on the WGS sequence results.

Specific primer design plays an important role in PCR, therefore it is important to find specific primers of good quality. Manual design and selection of primers tends to produce poor results due to human limitations. So, primary selection using automatic computing (software) is more practical and easier. In recent years, software for specific primer design has been well developed and applied effectively to the problem of primer selection with good results [10].

There are many and varied types of specific primer design software used and reported on PCR applications, therefore it is important to determine design software Specific primers suitable for application in whole genome sequencing research as well as in-silico and in vitro amplification of the anti-thrombosis protease gene encoding Bacillus sp. HSFI-3 through specific primer design. This journal review summarizes the types of primer design software specific to protease genes throughout the world in the last decade. The research specifically targeted the software used in the design of primers specific to the protease gene. The aim is to obtain the most appropriate software recommendations regarding specific primer design for the protease gene from Bacillus sp. HSFI-3.

2. Methods

Bibliographic was carried out through a database search http://app.dimensions.ai/. The search took into account the term "protease gene primer design" published in the data range from 2014 to 2023 and searched the title and abstract. The bibliographic aims to look at research developments related to protease gene primer design. Network visualization and density visualization maps for co-occurrence using VOSviewer.

The journal review was carried out by selecting sources from the PubMed and Google Scholer databases published between 2009-2022 which discussed specific primer design software for protease genes. Search for articles using Medical Subject Title Headings (MeSH) with several combinations including "protease gene", "protease enzyme", "protease gene characteristics", "protease enzyme characteristics", "cloning and overexpression", "primer design", "genetic engineering", “recombinant enzyme”. The journal review aims to find out the software used in designing protease gene primers and whether the designed primers were continued for in vitro PCR.

2.1. Journal Eligibility Criteria

Journal eligibility screening is based on the inclusion criteria set as follows: (i) cloning and overexpression of protease genes; (ii) protease enzymes; (iii) specific primer design software; (iv)reported in Indonesian or English; (v) search for review journals published in 2009-2022. All journals were extracted using search computerized and manuals from PubMed and Google Scholar. Journals excluded from the study were journals encoding other than protease genes.

2.2. Journal Selections

Based on guidelines from Polanin et al., [11], journal selection was carried out to identify journals that met the inclusion criteria applied in this journal review. Titles and abstracts produced by the search were identified and analyzed carefully to determine which inappropriate sources should be excluded. Journals from the remaining searches were also retrieved and evaluated whether they met the inclusion criteria or not [10].

2.3. Research Bias Control

The following things are included in the risk of bias or quality assessment in this journal review: (i) completeness of reporting information on specific primer design software for protease genes, (ii) selective reporting of results, (iii) in vitro PCR testing. When overall criteria are met, the overall reasonable risk of bias is considered low.

3. Result and Discussion

A search through the database http://app.dimensions.ai/ which was published in the data range from 2014 to 2023 and produced 10 scientific article publications or proceedings. The number of publications on "protease gene primer design" per year is presented in Figure 1. VOSviewer provides a network visualization map for co-occurrence. Network visualization of these 33 terms is presented in Figure 2. VOSviewer also provides a density visualization map for co-occurrence. Density visualization of these 33 terms is presented in Figure 3.

Figure 1 shows that research on protease gene primer designs is still relatively small and from 2019 to 2023 research will not increase. Figures 2 and 3 show research related to proteases connected to genes and primer but does not mention which primer used whether specific, fragment or full- length gene. So research related to protease gene primer design is still relatively small and has quite high novelty.

Fig. 1. Number of publications on "protease gene primer design" from 2014 to 2023 (source: http://app.dimensions.ai/)

Primer design is the first step that determines the success of DNA amplification using the PCR method [12]. Things that need to be considered in selecting a primer include primer length, melting temperature (Tm), GC (Guanine Citocin) content. Good primers range from 18-30 base pairs. Primers that are more than 30 base pairs long will cause the primer attachment to be non-specific. The second characteristic that needs to be considered in primary selection is Tm. A good primer has a Tm difference of around 5°C. This aims to ensure that there is no decrease in the amplification process. The percentage of G and C bases also needs to be considered because the number of G and C bases can affect the Tm of a primer [13]. A good primer has a G and C percentage of around 40-60%. Another criterion for a good primer is that it does not have cross dimers and a hairpin structure [14].

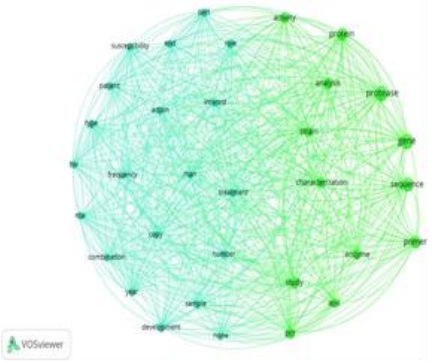

Fig. 2. Network visualization for co-occurrence of "protease gene primer design" (source: VOSviewer and http://app.dimensions.ai/)

Several studies have reported specific primer design software related to protease genes worldwide in the last decade. This journal review specifically summarizes journals that report specific primer design software related to protease genes. This journal review aims to analyze the prospects for designing specific primers related to protease coding genes to

be used as target genes or inserts in recombinant enzymes based on supporting references published in the last decade.

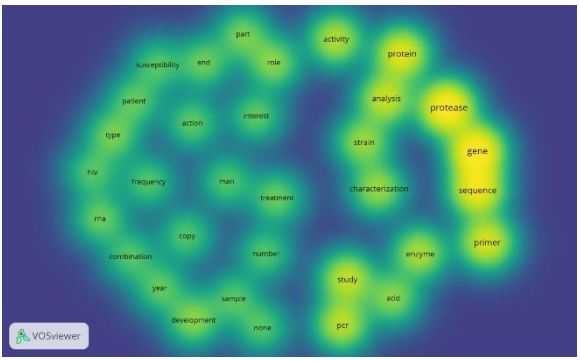

Fig. 3. Density visualization for co-occurrence of "protease gene primer design" (source: VOSviewer and http://app.dimensions.ai/)

Table 1. summarizes specific primer design software related to protease genes reported in the last decade (2009-2022). In table 1 the data shows that the primer design software that is widely used is dominated by NCBI Primer Blast. In Table 1, the data also reveals that all identified journals reporting the process of cloning and overexpression of antithrombotic bacterial protease genes are dominated by Asian countries including China, India, and Iran [15][16][17][18][19][20][21][22][24].

Primer is a component that plays an important role in the PCR process [24]. A good primer is a specific primer. Non- specific primers can cause other regions in the gene that are not targeted to be amplified or, conversely, none of them are amplified. Supporting the findings in table 1, the specificity of the primer must be considered so that it can recognize and attach to the desired target gene. For analysis, BLAST from NCBI can be used (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Using this software, candidate primer sequences will be compared with a collection of sequence databases from various organisms at NCBI. The similarity between primers and genes of a particular organism will be expressed in percentage identity. The higher this percentage value, the better the gene-specific primer [25].

Table 2 summarizes the results of specific primer designs tested in vitro which are reported to all produce bands on agarose gels. The primers designed using the software in table 1 show that the primers can amplify the target region or gene according to what we designed in the software [15][16][17][18][19][20][21][22][24]. All the software in table 1 only amplifies fragments, to obtain the target gene/insert in a recombinant enzyme requires primer design software that is capable of amplifying the full length gene. In research by Mazloum et al., (2018) it was reported that Benchling software (https://www.benchling.com/) was able to design primers for full-length genes [27].

Table 2. Results of specific primer designs tested in vitro PCR

No.	Primary Design Program	Web Address	Species	Country	Reference
1	Oligo 7 software	https://www.oligo.net/	*Bacillus* sp.	Iran	Ariyaei A, *et al*., 2019
2	Oligonucleotide Properties Calculator (Oligo Calculator version 3.27)	http://basic.northwestern.edu/biotools/OligoCalc.html	*Bacillus subtilis* RD7	Nigeria	Suberu Y, *et al*., 2019
3	NCBI Primer Design Tool	https://www.ncbi.nlm.nih.gov/tools/primer-blast/	*Bacillus licheniformis* MK90	Mesir	Ahmed A. Hamed, *et al*., 2019
4	Multiple NCBI Primer Designing Tools	https://www.ncbi.nlm.nih.gov/tools/primer-blast/	*Streptomyces thermovulgaris*	Pakistan	Mushtaq A *et al*., 2020
5	NCBI Primer BLAST	https://www.ncbi.nlm.nih.gov/tools/primer-blast/	*Nectria haematococca*	India	Swati N. Madhu, *et al*., 2020
6	Oligonucleotide Primer	https://www.bio-rad.com/en-rs/applications-technologies/oligonucleotides- design-applications?ID=MOYTCO15	*Bacillus subtilis* 168	Iran	Sadeghi M. Mohammad, *et al*., 2009
7	Primer Premier 5.0 Programs	http://www.premierbiosoft.com /primerdesign/	*Crassostrea virginica*	China	Yu, *et al*., 2011
8	'Oligoperfect' Primer Design Tool	https://www.thermofisher.com/ br/pt/home/life-science/oligonucleotides-primers-probes-genes/custom- dna-oligos/oligo-design-tools/oligoperfect.html	*Myzus persicae*	India	Bhatia, *et al*., 2012
9	Primer 3	https://primer3.ut.ee/	*Tritrichomonas foetus*	Australia	Sun, *et al*., 2012
10	Primer Premier 5.0 Programs	http://www.premierbiosoft.com /primerdesign/	*Agropyron mongolicum Keng*	China	Yang, *et al*., 2012

Table 2. Results of specific primer designs tested in vitro PCR

No.	Primary Design Program	Species	In-Vitro PCR	
			Yes	No
1.	Oligo 7 software	*Bacillus* sp.	✓	
2.	Oligonucleotide Properties Calculator (Oligo Calculator version 3.27)	*subtilis Bacillus* RD7	✓	
3.	NCBI primer design tool	*Bacillus licheniformis* MK90	✓	
4.	Multiple NCBI primer designing tools	*Streptomyces thermovulgaris*	✓	
5.	NCBI Primer BLAST	*Nectria haematococca*	✓	
6.	Oligonucleotide Primer	*Bacillus subtilis* 168	✓	
7.	Primer Premier 5.0	*Crassostrea virginica*	✓	
8.	'Oligoperfect' primer design tool	*Myzus persicae*	✓	
9.	Primer 3	*Tritrichomonas foetus*	✓	
10.	Primer Premier 5.0 programs	*Agropyron mongolicum Keng*	✓	

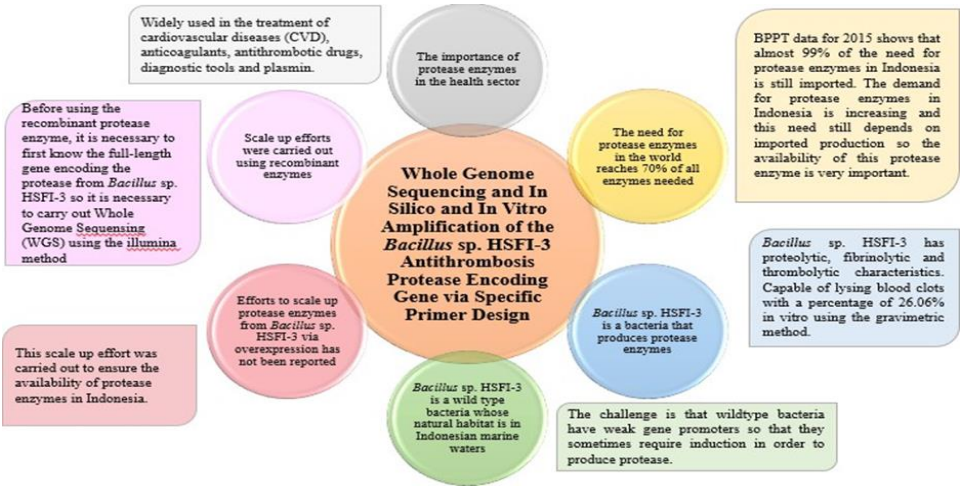

Fig. 4. The importance of target/insert genes in efforts to scale up recombinant protease enzymes from bacteria *Bacillus sp*. HSFI-3

Figure 4 shows the importance of obtaining target/insert genes in efforts to scale up recombinant protease enzymes from Bacillus sp. HSFI-3 bacteria. Based on the results of this journal review, it can be concluded that Benchling software is capable of designing primers for full-length genes. So we recommend this software for the design of specific full-length gene primers for scaling up protease enzymes to meet the needs of protease enzymes in Indonesia is fulfilled. Research like this, because it has not been reported before, provides originality and novelty.

Authors' Contributions. RAWA and SNE designed the entire research, RAWA collected, reviewed, screened and summarized all the journals obtained. SNE and RAWA evaluate the figures, tables and schemes created, also analyzing research bias. The main text was written by RAWA and SNE. The manuscript was originally written by RAWA and proofread by SNE.

Acknowledgments. All authors would like to express their deepest gratitude to the Master of Clinical Laboratory Science, Postgraduate Program, Muhammadiyah University, Semarang, who has supported this research

References

1. D. Gogoi, N. Arora, B. Kalita, R. Sarma, T. Islam, S. S. Ghosh, et al., Anticoagulant mechanism, pharmacological activity, and preclinical safety assessment of novelolitic serine fibrin proteases (ogen) from Leucas indica leaves, Sci Rep, 2018; vol, 8(1), p. 1–17. http://dx.doi.org/10.1038/s4159 8-018-24422-y
2. T.E, Serena, B.M. Cullen, S.W. Bayliff, M.C. Gibson, M.J. Carter MJ, L. Chen, et al., Define new diagnostic assessment parameters for wound care: Increased protease activity, nonhealing indicator, for targeted protease modulation treatment, Wound repair regen, 2016, vol. 24(3), p.589–95.
3. C.S. Craik, M.J. Page, E.L. Madison, Protease as therapy, Biochem J, 2011, vol. 435(1), p.1–16.
4. N. Hidayati, N. Nurrahman, H. Fuad, H. Munandar, D.S. Zilda, A.R. Ernanto, et al., Bacillus tequilensis isolated from intestinal fermentation of Holothuria scabra produces fibrinolytic proteases with thrombolytic activity, IOP Conf Ser Earth Environment Sci, 2021, vol. 707(1).
5. L.A. Utami, A. Suprihadi, Utilization Limbah Tahu sebagai Media Pertumbuhan Aspergillus flavus DUCC-K225 untuk Produksi Enzim Protease, Berk Bioteknol, 2018, vol. 1(1), p.1–6.
6. V. Deavina, A. Rahmawati, E. Yulianti, uji aktivitas enzim protease bakteri termofilik pasca erupsi merapi terhadap tepung jeroan ikan lele, Pend Biol, 2018, vol. 7(6), p.398–402.

7. R. Melliawati, Rohmattusolihat, Nurhayati, N. Rahmani, Yopi, Seleksi Dan identifikasi bakteri endofit potensial penghasil enzim protease dari taman nasional gunung halimun, Biopropal Ind, 2016, vol. 7(2), p.73–82.
8. M. Ananda, M. Ivana, P.P. Tokede, F. Natalia, B. Ginting, S.N. Tifen, R.A. Madao, D.N. Endah, E. Pangesti, et al., Potensi dan efektivitas salmonella typhimurium hasil rekayasa genetika sebagai terapi antikanker glioblastoma, Biotropika J Trop Biol, 2021, vol. 9(2), p.96– 104.
9. N.Z. Jaouadi, B. Jaouadi, N. Aghajari, S. Bejar, The overexpression of the SAPB of Bacillus pumilus CBS and mutated sapB-L31I/T33S/N99Y alkaline proteases in Bacillus subtilis DB430: New attractive properties for the mutant enzyme, Bioresour Technol, 2012, vol. 105 p.142–51
10. Y.H. Cheng, Computational intelligence-based polymerase chain reaction primer selection based on a novel teaching-learning-based optimisation, IET Nanobiotechnology, 2014, vol. 8(4), p. 238–46.
11. J.R. Polanin, T.D. Pigott, D.L. Espelage, J.K. Grotpeter. Best practice guidelines for abstract screening large-evidence systematic reviews and meta-analyses, Res Synth Methods, 2019, vol. 10(3), p.330– 42.
12. S. Suparman, Desain primer PCR secara in silicountuk amplifikasi gen coi pada kupu-kupu papilio ulysses linnaeus dari pulau bacan, J Pendidik Mat dan IPA, 2016, vol.7(1), p.14.
13. E.N. Sari, Desain primer untuk amplifikasi regio promoter gen inha isolat P016 multidrug resistance mycobacterium tuberculosis dengan metode polymerase chain reaction, J Farm Udayana, 2018, vol.7(1), p.34.
14. T.A. Safitri, N.J. Patty, H. Saraswati, Gen L1 Hpv 16 dan 18 sebagai dasar dalam desain primeruntuk deteksi kanker leher rahim dengan in-House Multiplex PCR, Ijobb, 2018, vol. 2(2), p.67
15. S.N. Madhu, S. Sharma, D.U Gajjar, Identification of Proteases: carboxypeptidase and aminopeptidase as putative virulence factors of fusarium solani species complex, Open Microbiol J, 2020, vol. 14(1), p.266–77.
16. T.G.B.Y. Ao, M.L. Lang, Y.Q. Li, Y. Zhao, L.C. Wang, X.J. Yang, Cloning and expression analysis of cysteine protease gene (MwCP) in Agropyron mongolicum Keng, Genet Mol Res, 2016, vol. 15(1), p.1–18.
17. Z. Sun, C. Stack, J. Šlapeta, Sequence differences in the diagnostic region of the cysteine protease 8 gene of Tritrichomonas foetus parasites of cats and cattle, Vet Parasitol, 2012, vol.v186(3– 4), p.445–9.
18. V. Bhatia, R. Bhattacharya, P.L. Uniyal, R. Singh, R.S. Niranjan, Host generated siRNAs attenuate expression of serine protease gene in myzus persicae, PLoS One, 2012, vol. 7(10), p1–9.

19. H. Yu, Y. He, X. Wang, Q. Zhang, Z. Bao, X. Guo X, Polymorphism in a serine protease inhibitor gene and its association with disease resistance in the eastern oyster (Crassostrea virginica Gmelin), Fish Shellfish Immunol, 2011, vol. 30(3), p.757 http://dx.doi.org/10.1016/j.fsi.2010.1 2.015
20. H. Mir, M. Sadeghi, M. Rabbani, M. Naghitorabi, Cloning of alkaline protease gene from Bacillus subtilis, Res Pharm Sci, 2009, vol. 4(1), p. 43–6.
21. A. Mushtaq, T.M. Ansari, G. Mustafa, M.A. Shad, J. Cruz-Reyes, A. Jamil, Isolation and characterization of nprB, a novel protease from Streptomyces thermovulgaris, Pak J Pharm Sci, 2020, vol. 33(5), p.2361–9.
22. A. Hamed, M. Khedr, M. Abdelraof, Molecular characterization of alkaline protease-coding gene from Bacillus licheniformis MK90 mutants with biofilm inhibitory activity, Egypt Pharm J, 2019, vol. 18(4), p:419.
23. Y. Suberu, I. Akande, T. Samuel, A. Lawal, A. Olaniran, Cloning, expression, purification and characterisation of serine alkaline protease from Bacillus subtilis RD7, Biocatal Agric Biotechnol, 2019, vol. 20(July), p. 101264. https://doi.org/10.1016/j.bcab.20 19.101264
24. A. Ariyaei, A. Farhadi, F. Moradian, M.G. Rahimi, Cloning, expression and characterization of a novel alkaline serine protease gene from native Iranian Bacillus sp.; a producer of protease for use in livestock. Gene, 2019, 693(December 2018), p.10-5 https://doi.org/10.1016/j.gene.2019.0 1.020
25. I.G.A.A. Septiari, P.S. Yustiantara, S.C. Yowani, Analisis primer untuk amplifikasi promoter inha multidrug resistance Tuberculosis (MDR-TB) dengan metode Polymerase Chain Reaction (PCR), J Kim, 2015, vol. 9(1), p.117–23.
26. H. Saraswati, Seprianto, F.D. Wahyuni, Desain primer secara in silico untuk amplifikasi gen cryiii dari bacillus thuringiensis isolat lokal. Indonesia, J Biotechnol Biodivers, 2019. Vol/ 3(1), p. 33–8.
27. A. zloum, N.G. Zinyakov, A.S. A.S. Igolkin AS, Vlasova. Cloning of genes encoding transmembrane proteins and proteins responsible for african swine fever virus virulence, Vet Sci Today, 2018; vol. 1(2), p.8–12.