The Use of Degenerate Primer to Isolation and Designing Housekeeping Gene of Eel Fish (*Anguilla bicolor*)

Diah Kusumawaty[1], Tina Yulianti[2], Hernawati[3]
{diah.kusumawaty@upi.edu[1]}

Departemen Pendidikan Biologi, Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam, Universitas Pendidikan Indonesia[1,2,3]

Abstract. This study aims to sequence the isolation results of the housekeeping gene in Anguilla bicolor to get specific sequences by designing degenerate primer on 18s ribosomal RNA (18S rRNA), Beta Actin (ACTB), Elongation Factor 1-Alpha (ef1a) and Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH). The degenerate primer consists of two primer sets (outer and inner) that were designed by the Primaclade method. The result of housekeeping gene amplification was then sequenced directly to determine the nucleotide sequences of the gene. The results of BLAST analysis showed that the three of four genes sequences were corresponding to the target gene, which are ribosome RNA 18s, (ACTB) and Ef1a. The general grouping of dendrograms showed that the three gene sequences were in groups with a gene target. The conclusion is the discovery of three specific sequences in three housekeeping gene targets which can be used as a source for designing specific primers in subsequent studies.

Keywords: Anguilla bicolor, Housekeeping gene, Degenerate Primer, Nested PCR, Specific Primer

1 Introduction

In Indonesia, the use of eel resources in the cultivation business to produce eel fish seeds is still very rare to do. In addition to cultivation efforts which are still rare, genetic information such as the housekeeping gene sequences database from eel fish from Indonesia especially Anguilla bicolor is also still small, not as much as genetic information on other Anguilla species. Preliminary survey on the National Center for Biotechnology Information (NCBI) page on October 7, 2018, the housekeeping gene sequence database from *Anguilla bicolor*, there are only two types, 16s rRNA and 12s rRNA.

The housekeeping gene is a constitutive gene that is needed in maintaining cell function and is expressed in all cell organisms under normal conditions. Housekeeping genes that are widely used in various studies include 18s Ribosomal RNA (18S rRNA), 1-Alpha Elongation Factor (EF1A) [1], Glyceraldehyde-3- Phosphate Dehydrogenase (GAPDH) and Beta Actin (ACTB) [2],18s Ribosomal RNA (18S rRNA) and Elongation Factor 1-Alpha (EF1A) genes are the most stable genes in all samples analyzed in the study of Politis [3] Beta Actin (ACTB) and Glyceraldehyde-3- Phosphate Dehydrogenase (GAPDH) genes are also the most stable genes expressed so they are used to normalize gene expression in research [4].

MSCEIS 2019, October 12, Bandung, Indonesia

DOI 10.4108/eai.12-10-2019.2296368

Genetic information such as the sequential database of the four housekeeping genes is unknown in eel (*Anguilla bicolor*). The housekeeping gene sequence information can be used to understand transcriptomes [5]. Sequences from the housekeeping gene can be obtained using the Nested PCR method, which is a PCR modification designed to increase sensitivity and specificity [6] In Nested PCR two primary sets are used Hoy [7]. The primers used in the amplification process are degenerate because the target gene is not specific or unknown[13]. According to Linhart and Shamir [8] primary degenerate is commonly used to identify genes that have minimal or unknown information.

Based on the description above, the research carried out is intended to sequencing the results of the isolation of the housekeeping gene in eel fish (*Anguilla bicolor*) by first doing a degenerate primary design. The housekeeping gene sequence obtained can be used as a source for designing specific primers in subsequent studies. The data obtained is also expected to complement and enrich existing databases.

2 Method

Housekeeping gene. The initial stage for the primary degenerate design is the collection of housekeeping gene 18s ribosomal protein (18S rRNA), Beta Actin (ACTB), Glyceraldehyde-3-Phosphate Dehydrogenase (GADPH) and Elongation Factor 1-alpha (EF1A) class Actinopterygii on the GenBankank page (http://ncbi.nlm.niv.gov). The sequences obtained are stored in the FASTA format and then the housekeeping gene alignment is performed on the clustalX software. Then the primary design is done online on the Primaclade page (Http: // primaclade/org). At the time of selection of primary candidates, several criteria are considered so that the resulting primer can amplify the target DNA well, while the primary design criteria in this study refer to the Premier Biosoft Design guidelines.

The selected primary is then checked for secondary structure on the NCBI Primary-BLAST page (https://www.ncbi.nlm.nih.gov/tools/primerblast/) and on the Beacon Designer page (Http: //www.premier biosoft.com). After obtaining a good primary candidate BLAST is carried out on the NCBI Nucleotide BLAST page (https://blast.ncbi.nlm.nih.gov/Blast.cgi?P AGE_TYPE = BlastSearch).

Identification of Eel (Anguilla bicolor). Eel fish used in this study were morphologically identified to ensure that eel fish species obtained from aquaculture are Anguilla bicolor species. Then compare it with the key determination that refers to the pedomana of Vertebrate Zoology [9] for the level of classes, orders, and families. While at the species level the morphological features compared were based on the research of Hakim [10]. Samples of eel fish (Anguilla bicolor) that have been identified are surgically removed and then some organs are then stored in a filled falcon tube RNA solution later.

DNA Isolation and amplification. DNA Isolation uses the DNA Isolation method used by Sambrook protocol with several modifications [11]. Amplification of Housekeeping gene using Primary degenerate with Nested PCR (nPCR) Method based on Kusumawaty method. DNA was amplified using tools in this study GeneAmp PCR System 9700 and Eppendorf Mastercycler Personal. the mix composition of the amplification reaction refers to Rahma's research with a total reaction of 10 µl. The PCR method used is a Nested PCR using two pre-designed degenerate primers (outer and inner) and an amplification program used by the GoTaq

Green Master Mix 2x protocol with some modifications. When optimizing the temperature in this study, the PCR touchdown method was also conducted, which refers to the GeneAmp® PCR System 9700 Base Module Biosystem21. Furthermore, the results of the amplification are sequenced by the direct sequencing method performed using a sequencer engine at Macrogen, Inc. South Korea

3 Data analysis

The results of the peak of DNA sequence sequences that unstable were cut off in the upstream and downstream parts. Then do contig (merging). Contig results obtained are stored in FASTA format. Then the BLAST homology test was carried out on the NCBI BLAST page (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

The dendrogram construction that was made in this study used Parsimony (PAUP) V.40 software. As for the use of PAUP software in this study, it refers to the window-based PAUP. Bootstrapping analysis used was 1000 replications to estimate the degree of trust of each group (klad) that was formed. The dendrogram that is formed is stored, and seen the construction that is formed using the treev32 program.

4 Results and discussion

4.1 Primary Candidate Degenerate Housekeeping gene Eel (Anguilla bicolor)

In the selection of primary candidates, the length of the primary base and the number of degenerate bases also need to be considered. In this study, the length of the selected primary bases ranged from 18 bp to 22 bp (Table 1). Based on Biosoft's Primary Design guidelines[22]. The primary length in this range is the optimal length measurement to facilitate the primer in binding the template at the annealing temperature. Then the number of degenerate bases in the selected primer is no more than 2 degenerate bases because based on previous research [12] a good degenerate primer is a primer that has a degenerate base ≤ 2.

4.2 DNA Isolation

A qualitative test of the results of DNA isolation in this study can be seen in Figure 1. The electrophoresis results show that DNA bands are formed in samples of the heart, pancreas, kidney, intestine, stomach, and liver. While in the flesh organ the DNA band is not formed. The formation of DNA bands does not result in electrophoresis due to low DNA concentrations and the occurrence of DNA degradation.

In the quantitative test, the purity and concentration values are known to range between 1.66 -1.813. DNA purity values obtained in the six organ samples ranged from 1,632 to 1,813. According to Fatchiyah et al24 good DNA purity values typically range between 1.8 and 2.0. However, in Rahma's research 20 DNA samples that had purity values of 1.681 and 1.640 could still be used for the amplification process. Based on qualitative test isolates that showed the best results were samples of kidney organs, intestinal organs, and stomach organs. But of the three organ samples that have the best purity is the kidney organ, so that in the process of amplification the kidney organ is selected as a DNA template.

Amplification Results of Eel (Anguilla bicolor) Housekeeping Gene Using Primary Degenerate. The amplification results of the four housekeeping genes using degenerate primers that have been designed show positive results. Positive amplification results using inner Ribosomal RNA (18S rRNA) primers were obtained at annealing temperature (Ta) 53oC to 63oC using the PCR touchdown method. Annealing temperature (Ta) is obtained from the lowering and raising of the temperature of 5oC from the melting temperature (Tm) of 58.4oC. Primary annealing (Ta) temperature is generally 5oC below the primary Tm temperature. The DNA band formed from the results of amplification has a size that corresponds to the estimated size of the amplicon, where the estimated size of the target amplicon from the primary 18S Ribosomal RNA (18S rRNA) inner is 586 bp and on the amplification results the DNA band results are above 500 bp (Figure 3).

Table 1. Size, starting post, and estimated product size of primary degenerate outer and inner housekeeping gene candidates.

Primer	**Siquence (5'- 3')**	**primer size**	**prediction size of amplicon (bp)**
18S_RNA_F_Outer	CCTTTGATCGCT C(Y)(M)ACG	18	1284
18S_RNA_R_Outer	TAAGAAGTTGG ACGC(S)GA	18	
18S_RNA_F_Inner	AAATTACCCACT CCCG(R)CTC	20	586
18S_RNA_R_Inner	AT(Y)GTTTATGG TCGGAACTACG	22	
B_Aktin_F_Outer	AGAGGTATCCTG AC (Y)CTGAAGT	22	888
B_Aktin_R_Outer	ACTC(R)TC(R)TA CTCCTGCTTGCT	22	
B_Aktin_F_Inner	CCTGAC(Y)CTGA AGTACCCCAT	21	745
B_Aktin_R_Inner	GC(W)GTGAT(Y) TCCTTCTGCAT	20	
EF1A_F_Outer	ACAT(Y)AACATC GT(R)GTCATTGG	22	1121
EF1A_R_*Outer*	CTTCTT(S)CC(R) GAACGACG	18	
EF1A_F_*Inner*	CAAGAGAACCA T(Y)GA(R)AAGTTC	22	646
EF1A_R_Inner	CC(R)ATTTTGTA GAC(R)TCCTG	20	
GADPH_F_*Outer*	CCAAGAAGGTGGAGAT(Y)GT	19	883
GADPH_R_Outer	GCAGACACG(K) T(K)GCTGTAG	19	
GADPH_F_Inner	AGTA(Y)GACTC(Y)ACCCACGG	19	704
GADPH_R_Inner	GCC(R)TTGAAGT CTG(W)GGA	18	

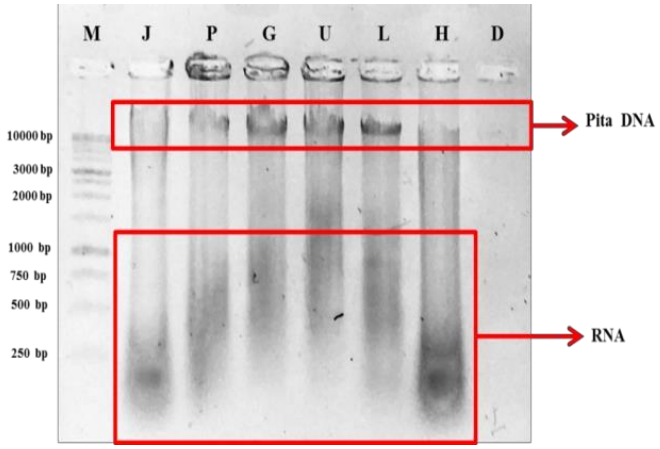

Fig. 1. Elektroforegram DNA isolation of Anguilla bicolor. M = Ladder 1 kb; J = DNA isolated from the heart organ; P = DNA isolated from pancreatic organs; DNA isolates from kidney organs; U = Isolate DNA from intestinal organs; L = DNA isolated from the stomach organ; H = DNA isolated from the liver; D = Isolate DNA from flesh tissue.

In the Beta Actin primer (ACTB), Elongation Factor 1-Alpha (EF1A) and Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) positive results obtained from the usual PCR thermocycler program at temperatures at 50oC. Beta Actin inner primers have a melting temperature (Tm) of 58.85oC so that the optimal annealing temperature is obtained from a decrease in melting temperature (Tm) of 8.85oC. The target size of the Beta Actin primer (ACTB) is 745 bp and the results of the DNA band amplification obtained do not match the estimated size of the amplicon but are greater than the estimate of 1250 bp.

Then for the Elongation Factor 1-Alpha (EF1A) primer the positive results obtained at 45oC. The temperature is obtained from the melting temperature (Tm) 55.45 of 10.45oC. However, the size of the amplification results from the primary Elongation Factor 1-Alpha (EF1A) does not match the estimated amplicon size where the results obtained are 1479 bp while the initial estimate is 646 bp (Figure 5). In the Glyceraldehyde-3-Phosphate Dehydration (GAPDH) inner primer which has a melting temperature (Tm) of 55.6oC a positive result is obtained from a decrease in melting temperature (Tm) of 10.6oC. The estimated amplicon size of this primer is 704 bp but when amplified it is only 300 bp (Figure 2).

4.3 BLAST Analysis

BLAST analysis of the four sequences of housekeeping gene sequences showed that three of the four target genes, namely 18s ribosomal RNA (18S rRNA), Beta Actin (ACTB) and Elongation Factor 1-Alpha (EF1A) correspond to the target gene and attach to class Actinopterygii fish. Meanwhile, BLAST homology test results on Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) sequences did not match the target wherein the sequence of the results of sequencing was not homologous to the target gene and instead attached to other genes, one of them in the assembly genome. This is caused by the incompatibility of the primer designed with the expected target. Primary incompatibility with the target gene can be caused by the secondary structure of the designed primer. In this study, the secondary structure of the designed primer has been minimized and even avoided by analyzing the secondary structure,

while the results are following the tolerance limits. Thus it can be seen that the GAPDH target gene may have a tolerance range of secondary structure that is lower than the guidelines used. Primary incompatibility with the target gene can be caused by too low GC content as well as secondary structure formation

.

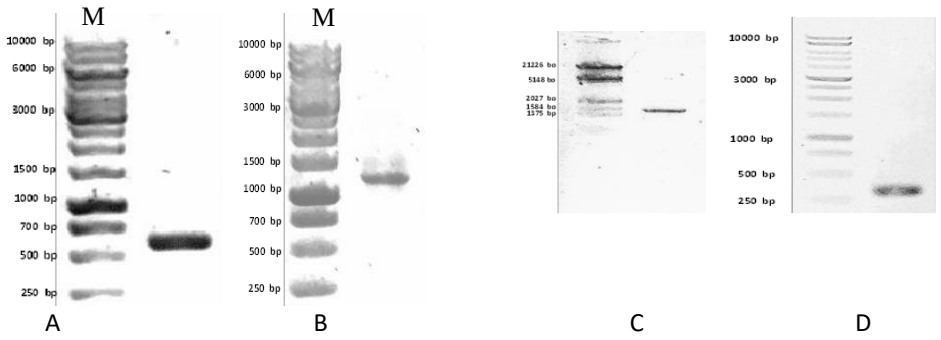

Fig. 2. Amplification Results of Eel (Anguilla bicolor) Housekeeping Gene Using Primary Degenerate.

4.4 Dendrogram Construction of Eel (Anguilla bicolor) Based on Nucleotide Sequences

In the dendrogram construction of the four housekeeping gene obtained only three gene sequences which were grouped with the target sequence were 18S_*Anguilla_bicolor*, EF1A_ *Anguilla_bicolor*, Beta_Actin_*Anguilla_bicolor*. Meanwhile, the gene sequence which was originally expected to be another Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) (GADPH_*Anguilla bicolor*) does not group with the Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) gene (**Figure 3**). These results are consistent with the BLAST results obtained previously where the gene sequences expected of the Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) gene do not match the target gene.

Thus it can be seen that the gene sequence obtained is not the Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) gene. In dendrogram construction, an outgroup of another gene is used, namely the Protein Binding Box (TBP) as a basis for comparison.

Bootstrap value with 1000 replications on the dendrogram construction shows that the Anguilla bicolor gene 18s Ribosomal RNA (18S rRNA) gene (18S_Anguilla_bicolor) obtained from the results of sequencing has a similarity to the Anguilla bengalensis 18s Ribosomal RNA (18S rRNA) gene, which has a similarity to the Anguilla bengalensis 18 75%. While the Anguilla bicolor beta Actin (ACTB) gene (Beta_Actin_^{Anguilla_bicolor}) has similarities with other Beta-actin gene sequences with 100% solubility. Then the Elongation Factor 1- Alpha (EF1A) (EF1A_*Anguilla_bicolor*) sequence has similarities to the Anguilla australis and *Anguilla Anguilla* Elongation Factor 1-Alpha (EF1A) genes with 100% accuracy.

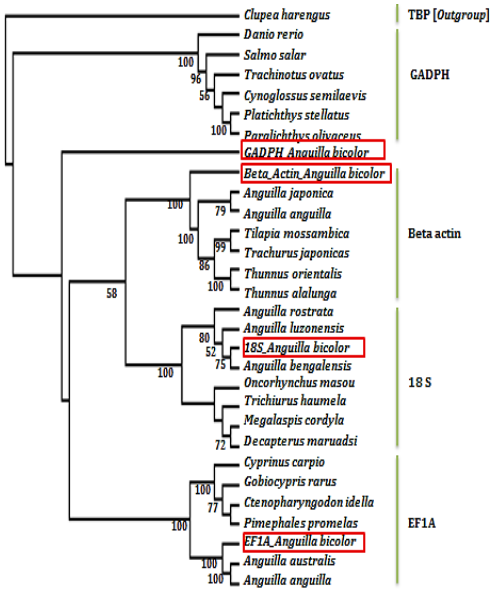

Fig. 3. Dendrogram housekeeping DNA gene18s_Anguilla _bicolor, Beta_Actin_Anguilla _bicolor, EF1A Anguilla _bicolo, and GAPDH_Anguilla_bicolor results from the elbow sequences and some Actinopterygii class fish gene sequences from the NCB1 page.

5 Conclusion

A specific sequence is obtained if the degenerate primer designed is capable of amplifying the target gene and showing homologous results with the target gene. The primary criteria for a good degenerate to be able to amplify the target gene are, having an estimated amplicon size> 800bp, having a secondary structure in the tolerance range or better if it does not have a secondary structure and has an optimal GC percentage and Melting Temperature (Tm) value.

References

[1] Setiawan, A.N & P Mark L.: The use of reference gene selection programs to study the silvering transformation in a freshwater eel Anguilla australis: a cautionary tale. [Online] Retrieved by http://www.biomedcentral.com/147 1-2199/11/75 (2010)

[2] McCurley, A. T. & Gloria V. C.: Characterization Of Housekeeping Genes In Zebrafish: Male-Female Differences And Effects Of Tissue Type, Developmental Stage And Chemical Treatment. BMC Molecular Biology. Vol. 9, No. 102, pp. 1-12 (2008)

[3] Politis, S. N., David M., Ariana S., Jose- Luis Z. I., Joanna J. M., Ian A. E. B.: Temperature Effects on Gene Expression and Morphological Development of European Eel, Anguilla anguilla Larvae. PLoS ONE, Vol. 12, No. 8, pp. 1 - 23 (2017)

[4] Filby, A. L. & Tyler, C. R.: Appropriate 'Housekeeping' Genes for Use in Expression Profiling the Effects of Environmental Estrogens in Fish.BioMed Central Molecular Biology. Vol. 8, pp. 1-13 (2007)

[5] Lin, Y., Shila G, Dario S, Andy Y. W., Ellis P., Trence P. S., ..., Pengyi W.: Housekeeping genes revisited at the single-cell level (2017)
[6] Carr, J., Donna G. W. & Randall T. H.: Molecular Detection of Multiple Respiratory Viruses. Maryland: Elsevier Inc (2010)
[7] Hoy, M. A.: Insect Molecular Genetics. 3 Edition. England: Academic Press (2013).
[8]. Linhart C., & Shamir R.: The Degenerate Primer Design Problem: Theory and Application. School of Computer Sciences: Tel Aviv University (2004)
[9]. Sudargo, F. dan Hernawati.: Zoology Vertebrata (Perkuliahan Berbasis Praktikum Superkelas Pisces). Bandung: Jurusan Pendidikan Biologi UPI (2017)
[10]. Hakim, A. H., Mukhlis K., Nurlisa A. B. dan Ridwan Affandi.: Komposisi Species Ikan Sidat (Anguilla Spp.) di Delapan Sungai yang Bermuara ke Teluk Palabuhanratu,Sukabumi, Indonesia. Jurnal Ilmu dan Teknologi Kelautan Tropis. Vol. 7, No. 2 pp. 573-585 (2015)
[11]. Sambrook, J., Fritsch, E. F., Maniatis, T.: Molecular Cloning: A Laboratory Manual. 2nd ed. New York: Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989)
[12] Hellman, L. M. and Fried M. G.: Electrophoretic Mobility Shift Assay (EMSA) for Detecting Protein-Nucleic Acid Interactions. Nature Protocols, Vol. 2, No. 8, pp. 1849-1861 (2007)