

# Identification and analysis of snps in population of Vietnamese catfish (*Pangasianodon hypophthalmus*), using next generation sequencing and snp validation

## Abstract

Tra catfish (*Pangasianodon hypophthalmus*) is a commercially important aquaculture species in the Mekong Delta, Vietnam; accounting for approximately 60% of Vietnam freshwater fish production. However, the food trade globalization and the growing demand for selected varieties of food have led to the intensification of commercial fraud, especially in the form of substitution species and mixing with cheaper taxa. DNA bar-coding methods have shown a potential well proven molecular approach to assess the authenticity of food items. In this study, we utilized high throughput NextSeq500 sequencing to capture and identified SNP markers on 400 wild and farming catfish individuals from 8 Vietnamese provinces (e.g., An Giang, Dong Thap, Can Tho, Tien Giang, Ben Tre, Hau Giang, Vinh Long, and Ho Chi Minh City) and 10 international locations of Cambodia, Thailand and Bangladesh. Stringent filtering of SNP-calling parameters resulted in 11,009 SNP loci represented across all 4 countries. We utilized the VCF tools to validate two panels of SNPs selected from the NGS data. Selection criteria included SNPs shared between catfish's Vietnam and international populations, and SNPs specific to each population. A total of 780 and 809 SNPs were filtered in Vietnamese and international catfishes, respectively, the final 12 special SNPs were detected. SNP validation by using Sanger sequencing and PCR showed a result of 2 per 12 SNPs was confirmed successfully in 300 individuals from Vietnamese catfishes. The first times our results provide genotype informative marker loci in Vietnamese Tra catfish in order to approach molecular traceability in aquaculture species with minimal previous genetic information.

**Keywords:** vietnamese tra catfish, next generation sequencing (NGS), single nucleotide polymorphism (SNP)

Volume 1 Issue 1 - 2018

Thuy Thi Bich Vo,<sup>1</sup> Hieu Duc Nguyen,<sup>1</sup> Thu Thi Hoai Nguyen,<sup>1</sup> Ha Thi Thuy Tran<sup>2</sup>

<sup>1</sup>Institute of Genome Research, Vietnam Academy of Science and Technology, Vietnam

<sup>2</sup>Centre of Aquaculture Biotechnology, Research Institute for Aquaculture No 1, Vietnam

**Correspondence:** Thuy Thi Bich Vo, Institute of Genome Research, Vietnam Academy of Science and Technology, Hanoi, Vietnam, Tel +84-24-32191174, E-mail thuytbvo@igr.ac.vn

**Received:** November 26, 2017 | **Published:** January 10, 2018

**Abbreviations:** AS, allele specific; NGS, next generation sequencing; SNP, single nucleotide polymorphism; PCR, polymerase chain reaction

## Introduction

The Vietnamese striped catfish (locally name as Tra catfish) has been cultured in the Mekong Delta, from the 1960s of the last century.<sup>1</sup> Average output of Tra catfish reaches 1.5 million tons per year, and has profited 1.6 billion USD in export business in 2016, increasing about 7% compared to 2015. Vietnamese catfish products have exported to 140 markets in the world, and the United States continued to be a major market for Vietnam's pangasius exports, accounting for about 20% of total exports.<sup>2</sup> Vietnam has become the third largest producer of fisheries products in the world.<sup>3</sup> However, in recent years, Vietnamese Tra catfish export industry has been facing many challenges such as illegal fishing, no declaration of origin, failing to comply with regulations and technical standards on food safety conditions and commercial fraud in fisheries. These stories have a huge impact on the economic and sometimes even public health implications. Therefore, the development of an efficient seafood traceability framework is crucial for the management of sustainable fisheries and the monitoring of potential substitution fraud across the food chains. There have been relatively few genetic studies of Tra catfish and information on genome resources in this species also limited.<sup>4-6</sup> Next generation sequencing (NGS) technologies have opened many opportunities to develop molecular resources

in biological and economic interest. In addition, genomic datasets represent may be mined in order to identify suitable candidates for SNP marker development. SNPs potentially can be utilized for developing aquaculture lines<sup>7</sup> or seafood traceability.<sup>8-10</sup> Recently, over one hundred studies of whole genome sequences, transcriptome analyses, gene expression patterns, ESTs, and SNP identification for aquaculture species using NGS have been conducted to date, in particular for Atlantic salmon, catfish, rainbow trout, striped bass, oyster, zebra fish and medaka.<sup>11</sup> The available genetic tools for traceability are SNPs, representing sites in the genome with minute mutations (novel genetic differences) in the DNA sequence. They are very abundant and widespread. Analyses of SNPs reach hitherto levels of population identification, rendering them optimal tools in fundamental biology, conservation and traceability.<sup>12</sup> In the current study, the first times provide a genotype informative in Vietnamese striped catfish *P. hypophthalmus* using the NextSeq500 platform, Sanger sequencing and PCR.

## Materials and methods

### Fish sources and sampling

All procedures involving the handling and collection fish samples during this study were approved by the Research Institute for Aquaculture No.1, Vietnam prior to initiation of the project. A total of 400 individuals catfishes, with 24 locations (i.e., 14 locations in 8 provinces of Vietnam and 10 locations in 3 countries as Cambodia, Thailand and Bangladesh- hereafter referred to as international

samples) were used for this study. These locations were from different geographic areas, which possess different production traits such as growth rate, disease resistance and feed conversion efficiency. The original fish samples were collected from the Hau Giang River, Tien Giang River, the local fish farms (in Vietnam), Mekong river (the region of Cambodia, Thailand, and Bangladesh), and other local fish farms of 3 international countries (i.e., Strung treng, Kratie, and Tonle Sap of Cambodia; Loie, Nakhon Phanom, and Ubon Ratchathani of Thailand; Jessore, Brahmanbaria, Jamalpur, and Badal of Bangladesh). Catfish fins (approximately 300-500 mg) were sampled from different locations, placed in eppendorf micro tubes with absolute ethanol and maintained in a freezer at -20°C.

### DNA extraction and sequencing

Total DNA was isolated using the GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific, CA, USA) following the manufacturer's protocol. Equal amounts of DNA (100mg) from each individual were pooled for sequencing, one pool for each location. Sequencing was conducted at Institute of Genome Research (Hanoi, Vietnam). Genomic libraries were prepared with the TruSeq® DNA PCR-Free (Illumina, San Diego, CA) with 5mg of genomic DNA for all samples, according to the manufacturer's instructions. For each pool sample, the prepared DNA library was sequenced on one lane of the Illumina NextSeq 500 platform for 100bp paired-end reads.

### Reference mapping

Sequence mapping was performed using CLC Genomics Workbench (version 4.0.2; CLC bio, Aarhus, Denmark). Before mapping, adaptor sequences, ambiguous nucleotides (N's), extreme short reads (<30bp) and low quality sequences (Quality score <20) were trimmed using CLC Genomics Workbench. After that, all trimmed reads were then aligned with the Vietnamese catfish genome reference (unpublished data). The mapping settings were: mismatch cost of 2, deletion cost of 3 and insertion cost of 3. The alignment had more than 95% similarity with the reference sequence across 90% of their length was accepted. The mapping output was export as BAM format<sup>13</sup> for further analysis.

### SNP identification and filtering

SNPs were identified from the alignment sequences using the SAM tools (version 0.1.18)<sup>13</sup> and PoPoolation2<sup>14</sup> with the lowest criteria setting to obtain all potential SNPs. Two factors that are important for excluding false SNPs caused by sequencing errors were set: minimum read depth and maximum read depth. SNPs were defined as specific SNPs if the SNP polymorphisms were found in more than 95% organisms from only one location and less than 5% organisms from the other location. SNPs that were polymorphic in all strains were considered as common SNPs.

### Polymerase chain reaction (PCR) amplification using primers for non-informative markers and Sanger sequencing

Genomic DNA was amplified by PCR using the selected primers (Table 1) in a 25µL PCR reaction consist of 12.5µL GoTaq Green Master Mix (Thermo Fisher Scientific, CA, USA), 1.0µL (10 ng) DNA template, 0.5µL of each primer (10 ppm), and 10.5µL sterile deionized water. Amplification was done in a Veriti Thermal Cycler (Applied Biosystems, Foster City, CA, USA) programmed for an initial denaturation at 94°C for 5min, followed by 25 cycles of 94°C for 30s, 45°C to 65°C based on the primer sequences for 30s, and 72°C for 50 min, with a final extension at 72°C for 10 min. The PCR products were visualized by standard 2 % agarose gel electrophoresis and purified by GenJet PCR Purification Kit (Thermo Fisher Scientific, CA, USA). A 10µL sequencing reaction volume was prepared using 2µL of the cleaned PCR product, 0.5µL (10 µM) of either the forward or reverse primer, 1µL BigDye terminator v3.1 (Applied Biosystems), and 1 µl Big Dye Sequencing (Genetix). The sequencing PCR consisted of 25 cycles of 10 seconds at 96°C, 5 sec at 50°C, and 4 min at 60°C. The sequencing reaction product was cleaned through sephadex columns before adding 20µl of Hi-Di formamide (Applied Biosystems) and then denaturation at 98°C for 2 min. The samples were analyzed on an ABI 3500 Genetic Analyzer (Applied Biosystems) with a 36 cm capillary and POP-6 polymer and using the default settings. The resulting parental sequences were visually screened for SNPs in Sequencher 4.2 software (GeneCode).

**Table 1** Primer for PCR amplification

No	Sequencing (5'-3')	Size of PCR product (bp)	Annealing temperature (T°C)	
SNP 1	Forward	AGAGCATAAGCAAATCGCTGC	531	59°C
	Reverse	TTCCAATTGGGGCTTTGTGTG		
SNP 2	Forward	TGGTATTCGGAGGAAAGATGGC	654	57°C
	Reverse	TTGGACCACTAGGAGGAGACA		
SNP 3	Forward	GCTGTTTCGGTTCAGAAGCTG	617	59°C
	Reverse	AGCACAGGTGAAGGGTTGTC		
SNP 4	Forward	CACCCTGTGGACACGAACA	516	59°C
	Reverse	AGTGACTTCTCTAGCGGAC		
SNP 5	Forward	CCAGTGGTCATCTGGGGAAAG	580	63°C
	Reverse	CGTCACTCACACACCACAAG		
SNP 6	Forward	AGTGCCATTGCTGTGGTAAC	626	57°C
	Reverse	CACAGTCATGGCGTTGGATG		

Table Continued....

No	Sequencing (5'-3')		Size of PCR product (bp)	Annealing temperature (T°C)
SNP 7	Forward	GGCCGTGCTGTTATGCGAAA	504	65°C
	Reverse	CATTCAGCAGCAATCAGGAGC		
SNP 8	Forward	CGGGTGTCCACCATGCTTTA	559	63°C
	Reverse	GTAGCCGCACCATTCTCAGT		
SNP 9	Forward	CCCCAGGTTTGACATTGCAC	650	63°C
	Reverse	CATCAGCGCCTGATCTCACT		
SNP 10	Forward	TCTGGGTGGACAAAGGTAGG	528	63°C
	Reverse	CATGCACAACAGCTCACTGC		
SNP 11	Forward	GGCCTACGACGATGGTTACA	532	63°C
	Reverse	CAGTTTAGGGAAGGCCACACA		
SNP 12	Forward	TAGTCAGACGCTCCAAACCG	518	65°C
	Reverse	GAACCGAGCGTAGAAGTCGT		

### SNP validation with allele-specific (AS) markers by Nested-PCR amplification

Two SNP markers were selected from the merged of sequencing results and reference sequencing. The AS primers were designed by a high throughput web tool for picking PCR and sequencing primers as BatchPrimer3 v1.0, containing SNP5- 45:G>A (45 indicates the position of the selective nucleotide) with the primer sequence, forward 5'-GCGAGGAGATCACTCACAA-3', reverse 5'-GAGTGGCTATGAGTTGGCTC-3' (fragment size: 114 bp), and SNP9- 66:T>G with the primer forward 5'-AGCAAGATACAGCATTACCG-3', reverse 5'-GACTCACCTGCATGCGTAAT-3' (fragment size: 145 bp). The first PCR amplified a region of SNP 5 and 9 using the designed primers

in (Table 2). This step was performed to isolate regions of interest containing the relevant SNP 5 and 9 polymorphisms that were later used for the second allele-specific PCR to avoid amplification of similar sequences in the catfish genome that may be located outside the gene. After a successful the first PCR, 2.0 µL of diluted PCR product were used as a template for the detection of mutant in the second PCR. The second PCR was carried out in a reaction mixture and PCR condition with the modifications; higher annealing temperature (3-5°C increase) and reduced MgCl<sub>2</sub> concentration (1.5 mM), which further increased the specificity of amplification. Amplicons were detected by electrophoresis in 2%-agarose gels and 1× TBE at 130 V for 90 minutes. SNP validation was confirmed from genomic DNA of all 400 individuals of the mapping population (including 300 and 100 individuals of Vietnamese and international catfishes, respectively).

Table 2: The sequencing of SNPs location in Vietnamese and international catfishes.

No	Sequencing of SNPs (Trimmed Seq)	SNP position	SNPs	Call rate	Location
SNP 1	TGCAGCTGTAGCTGAATTGTAAAGTGTGCCTGGAGCCCTT CCTTCCAATTAATGCAATGCATG	28	C>T	1	Vietnam
SNP 2	TGCAGTCATGATGTGTTTCTTAAACTGGATGCATAGTTACT GAGCAACACACACACATACACACACA	17	T>C	0.989247	Vietnam
SNP 3	TGCAGATGCATTAAGGATAAATAATTATATTGTTTAGGCCT ACATTACGACATTAATATATGGAACCCT	61	G>A	0.989247	Vietnam
SNP 4	TGCAGCTCTGAAACACAGCACTATACAGCACGTCCTAAT AATAGTAAAAGCAAGCCATGTAATGGAAA	31	C>T	0.989247	International
SNP 5	TGCAGGCTCCACTGCTGTTTCATGTTGGCGAGGAGATTCAC TCTGTACGTGCTGGAGGATGATCTGCTG	45	G>A	0.989247	Vietnam
SNP 6	TGCAGAGTGACATTAGGTTATGTTGTATTTGCTTATCATAG ACCTGTATCAGGGGTTCCCT	62	C>G	0.989247	International

Table Continued...

No	Sequencing of SNPs (Trimmed Seq)	SNP position	SNPs	Call rate	Location
SNP 7	TGCAGTACTACAGCTACAATCCACTCAGCCACAGCACCAC TGGGCAGCAACAGTGTGATTGCTGTGCTG	50	C>A	0.989247	International
SNP 8	TGCAGCTCTCTGTGTCCATGTTTCTCATTITTTTGTGAACTA AAGAGATTCGGCTGACTTGTGTATTCC	50	C>T	0.978495	Vietnam
SNP 9	TGCAGCTCTCGCTCTCAAATTCAAATGAGCCTTACTGGCAT GGCGCTAGCAAGATACAGCATTGGGAA	66	T>G	0.978495	Vietnam
SNP 10	TGCAGCAATTTTCTCATAACATCAGTGAGTACTGCTCCTG GAGTACACATAATGTTTCAAAAATGTGC	58	C>T	0.978495	International
SNP 11	TGCAGGTAAGTATTGGGCTCCACATTCAGTTAAAA TTCAGTTAAAGTTATTAATAAAT	14	C>T	0.967742	International
SNP 12	TGCAGTCATCCACGGCTTCTGGTTGGAGCGTGATGTGATGG TCTTGGAGGCGGTGACGTCATCAATGCA	56	C>T	0.967742	Vietnam

### Analysis of significant SNPs

Statistical analysis was performed using GraphPad Prism5 (GraphPad Software Inc., La Jolla, CA). SNPs with significantly different frequency ratios between Vietnamese catfish and international catfishes population were analyzed by one-way ANOVA. A p value of <0.05 was considered statistically significant.

## Results

### Identification SNPs among populations

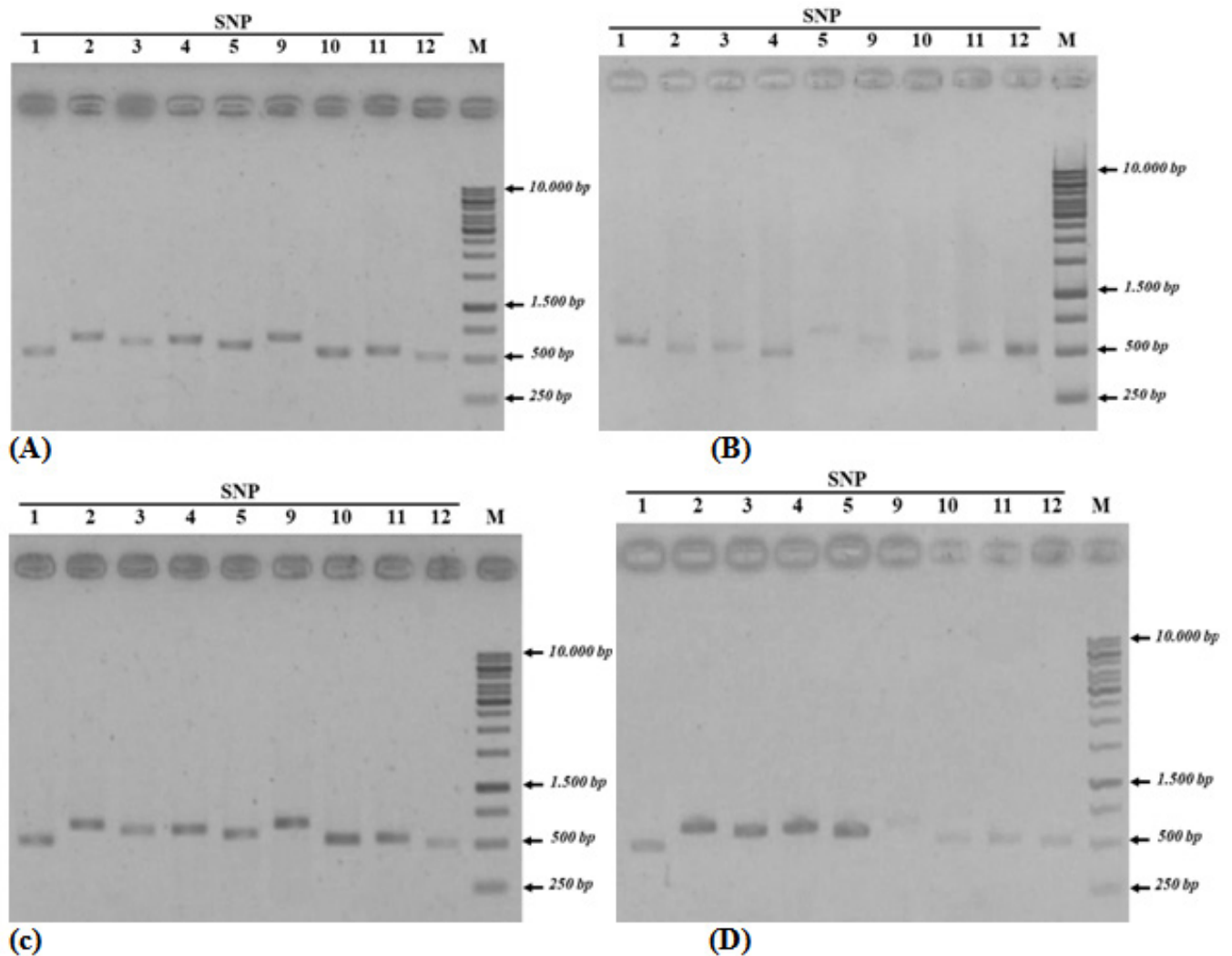
A total of potential 16,689 SNPs were observed at the lowest criteria setting. At the next selected set of criteria, minimum read depth is set as excluding the top 5% of minimum reads for SNP detection and maximum read depth is set as excluding top 5% of maximum reads for SNP detection, a total of 11,009 putative SNPs were identified, in which 780 SNPs observed in Vietnamese catfishes and others in international catfishes. Finally, we compared all putative SNPs between two strains and received twelve specific SNP markers (Table 2).

### SNPs validation in Vietnamese catfish populations

Twelve special SNP markers selected from the preliminary Vietnamese catfish genome assembly map that confirmed by both PCR method using primers designed as (Table 1) and Sanger sequencing. Special SNPs identified from each of the area were shown in (Figure 1) (Figure 2). As shown in Figure 1, the primers designed for the first PCR successfully produced the desired products, except primers for SNP 6, 7, and 8 that failed in primer design and data not showed here. Analysis of the results showed that SNP 1, 2,

3, 4, 10, 11, and 12 appeared in the test regions, were different with the expected analysis. As shown in (Table 2), some SNPs (e.g., SNP 1, 2, and 3) only appeared in Vietnamese catfish samples without the presence of international catfish samples, however, the Sanger sequences showed that they have appeared in the samples of both areas. Similarly, SNP 4 and 11 should only appear in international catfish samples and not appear in Vietnamese catfish, but the results have been shown the opposite data. Other SNPs (e.g., SNP 10 and 12) did not show the mutation position like analysis results in (Table 2). Only two SNPs 5 and 9 were observed a similar data, these mutants appear only in Vietnamese catfish, but not in international samples. Thus, by re-checking the position of the mutant via PCR method and sequencing analysis, the SNPs 1, 2, 3, 4, 10, 11, and 12 were not used to determine the difference between catfish's Vietnam and international catfish (Thailand, Cambodia, and Bangladesh). In this case, the observation of both SNP 5 and SNP 9 have been corrected from both PCR and sequencing data, and may be considered as special SNPs for Vietnamese catfishes.

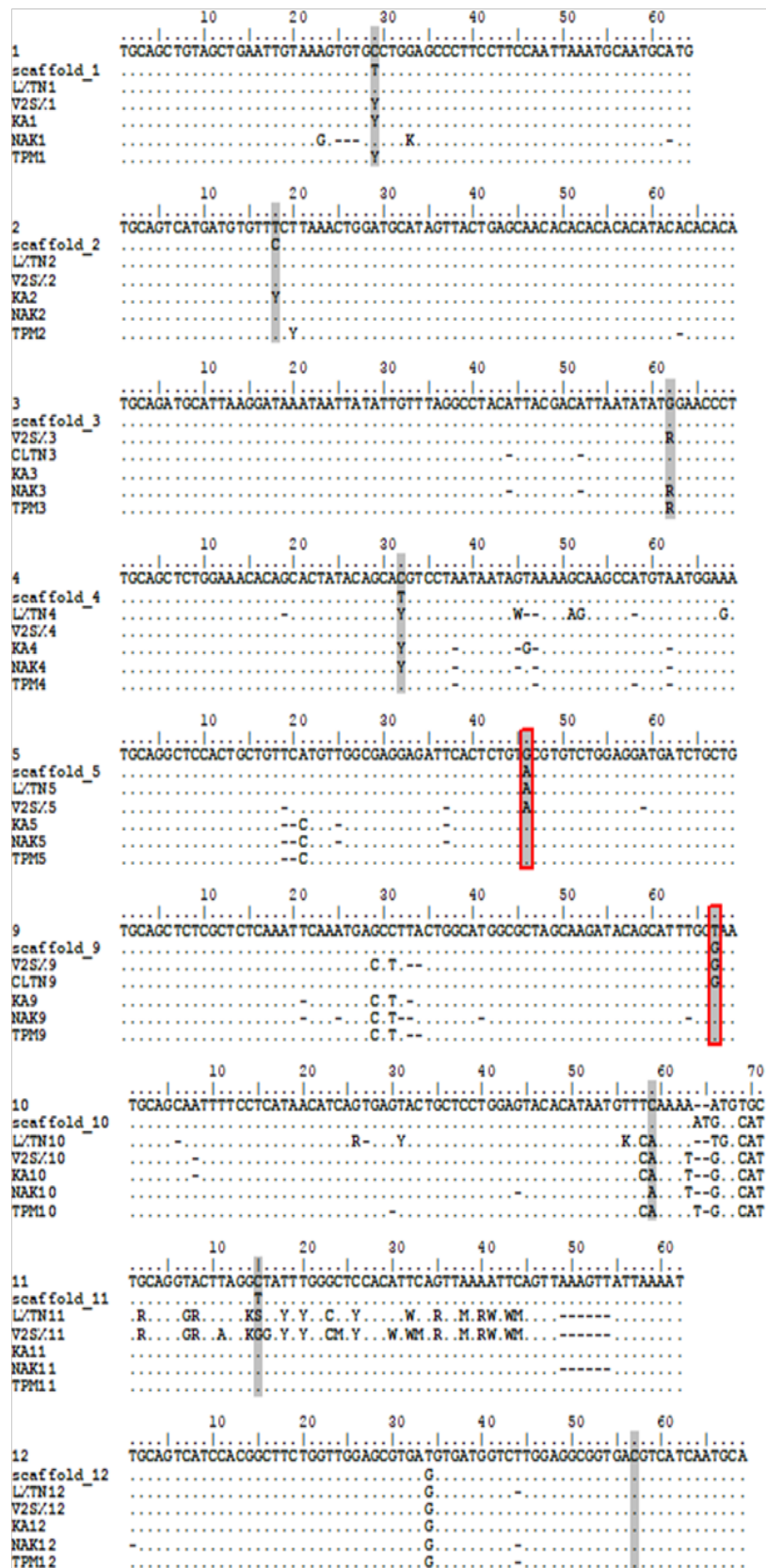
The number of each specific SNPs (e.g., SNP 5 and 9) was confirmed with gel electrophoresis in 300 individual Vietnamese catfishes (including 14 locations of 8 provinces) and 100 other individual international catfishes (e.g., 10 locations in 3 countries as Cambodia, Thailand and Bangladesh). The results were shown in (Table 3). Interestingly, 100% SNP5 observed in Vietnam catfish samples, whereas no expression in international catfish samples. Another SNP 9, approximately 75% - 84.4% of this SNP was identified in wild and farming catfishes, respectively. A significant difference ( $P < 0.01$ ) was appeared when validating the occurrence of SNP5 and 9 in both catfish populations, its mean may be expected that SNP 5 and 9 are specific for Vietnamese catfishes.



**Figure 1** Electrophoresis pattern for the first PCR carried out with the specific primers. (A) PCR amplification of Vietnam regions; (B) PCR amplification of Thailand regions; (C) PCR amplification of Cambodia regions; (D) PCR amplification of Bangladesh regions. M: 1kb-plus DNA ladder.

**Table 3** The occurrence of SNP 5 and SNP 9 in wild and farming catfish samples between Vietnam and the international regions

Location	SNP 5		SNP 9	
	Wild catfish	Farming catfish	Wild catfish	Farming catfish
Vietnam	300/300 (100%)	300/300 (100%)	75/100 (75%)	76/90 (84.4%)
International (Cambodia, Thailand, Bangladesh)	0/30 (0%)	0/60 (0%)	4/30 (13.3%)	6/60 (10%)
<i>P</i> value	0.0001 ( $P < 0.01$ )	0.0001 ( $P < 0.01$ )	0.0001 ( $P < 0.01$ )	0.0001 ( $P < 0.01$ )



**Figure 2** Nucleotide sequence of the PCR product obtained from sequencing was aligned using BLAST. The position of the specific SNPs merged between sequencing results and reference sequence are framed.

## Discussion

In recent years, several studies have successfully used NGS technologies for SNP marker discovery in fish species relevant to fishery and aquaculture, for instance, catfish,<sup>15</sup> lake sturgeon,<sup>16</sup> rainbow trout,<sup>17,18</sup> lake whitefish,<sup>19</sup> Atlantic cod,<sup>20</sup> salmonids,<sup>17,18</sup> hake,<sup>21</sup> turbot,<sup>22</sup> Atlantic herring<sup>23</sup> and Pacific herring.<sup>24</sup> According to the latest statistics on fish species, approximately 7,800 species (25%) have successfully researched at least one characteristic barcode for the species and the original data related to the species was deposited in BOLD.<sup>25</sup> In the study, we have identified 12 specific SNPs (about 0.1 %) in the Vietnam catfish and the international catfish samples. The small numbers of SNPs detection may have been a result of sequencing the small population sizes. Besides, the catfish of four countries have not yet published their reference genome sequencing, therefore, when assembling and discovering SNPs, we used all their assumption genome sequencing of four catfish species. This might lead to the assembly of the same contig due to their high sequence similarity, single nucleotide differences between duplicated loci in the genome but invariant at the population or species level<sup>26</sup> or single nucleotide variants with complex characteristics due to polymorphisms within duplicated regions,<sup>26</sup> and all of situations were indistinguishable from SNPs during filtration. Additionally, the SNPs identified in this study using stringent criteria in the construction of special SNPs.<sup>27</sup> Like Chao Li et al.<sup>28</sup> studied in 190 individual fishes of five wild and domesticated populations (e.g., Mississippi River, Missouri River, D & B, Rio Grande and Texas) scored for 4,275 SNPs across all five populations. The confirmation these SNP in five populations only showed a total of 64 SNPs (1.5 %) that have successfully emerged in 100% individual in all populations.<sup>28</sup> Another study found about 6.6 million, 5.3 million, 4.9 million, 7.1 million and 6.7 million common SNPs in Marion, Thompson, USDA103, Hatchery and wild fish populations, respectively, using the new generation sequencing system. Approximately 6% of all SNPs are identified in individual species.<sup>9</sup>

SNP combined with sequencing is also known as a useful indicator in the traceability of some aquatic species. Identification of specific SNPs of Vietnamese catfishes to avoid trade frauds in aquaculture and export between Vietnamese catfishes and *Pangasius* in the surrounding areas. We identified two SNPs appearing at levels of 75-100% in the catfishes of Vietnam, whereas only 0-13.3% in the international catfish samples. The other studies have identified SNPs in population that showed the similar parameters. Iratxe Montes et al.<sup>10</sup> used the NGS sequencing technique to detect and confirm SNPs in European anchovy species without reference sequence. From the analysis results for the assumed 2,317 SNPs suitable for the design of priming tests in populations. A set of 530 individuals were tested, resulting in 83.2% of individuals presenting these SNPs.<sup>10</sup> Neilsen et al.<sup>29</sup> using genetic polymorphisms have found the origin of four commercial marine fish populations (e.g., Snowy Fish, Trichoderma fish, Trilobites and Coalfish) on European scale. The results were reported that 93-100% of individuals able to accurately retrieve origin via SNPs combined with Bayesian statistical methods.<sup>29</sup> DNA barcodes are also used to detect fraudulent foods, including fish products.<sup>30</sup> Recently, many cases of tuna identification have found misprinting and illegal trade.<sup>31</sup> Yoon et al.<sup>32</sup> developed a DNA chip to identify nine species of salmonids and managed the origin and commercial traceability of salmon.<sup>32</sup> For acacia fishes, traceability of products by high molecular indicators avoids commercial fraud.<sup>33</sup> From our results suggested that SNP5 and 9 can be used to make the bar-coding sequence distinguish between Vietnamese and international catfishes in the future.

## Conclusion

Catfish species are almost impossible to distinguish based on phenotypes, therefore, the number of SNPs identified from this study can be potentially used for species identification, tracing the origin of commercial species, analyzing the genetic difference, and marking fish for other genetic experiments in Vietnamese catfish. The further, it will be useful for the development of high density SNP arrays for genetic and genomic analysis in catfish.

## Acknowledgments

This research project was sponsored by Vietnam Ministry of Agriculture and Rural Development.

## Conflict of interest

The author declares no conflict of interest.

## References

1. Phan TL, Bui MT, Nguyen TTT, et al. Current status of farming practices of striped catfish, *Pangasianodon hypophthalmus* in the Mekong Delta, Vietnam. *Aquaculture*. 2009;296(3-4):227-236.
2. Nguyen THP. Effects of temperature and salinity on growth performance in cultured tra catfish (*Pangasianodon hypophthalmus*) in Vietnam. Environmental and Biological Sciences, Queensland University of Technology Brisbane, Australia; 2015. p. 1-159.
3. FAO. The State of World Fisheries and Aquaculture: Opportunities and Challenges. Nations FaOoTU, Rome; 2014. p. 1-243.
4. Na Nakorn U, Moeikum T. Genetic diversity of domesticated stocks of striped catfish, *Pangasianodon hypophthalmus* (Sauvage 1878), in Thailand: relevance to broodstock management regimes. *Aquaculture*. 2009;297(1-4):70-77.
5. Nguyen TTT. Patterns of use and exchange of genetic resources of the striped catfish *Pangasianodon hypophthalmus* (Sauvage 1878). *Rev Aquaculture*. 2009;1(3-4):224-231.
6. Ha HP, Nguyen TTT, Poempuang S, et al. Microsatellites revealed no genetic differentiation between hatchery and contemporary wild populations of striped catfish, *Pangasianodon hypophthalmus* (Sauvage 1878) in Vietnam. *Aquaculture*. 2009;291(3-4):154-160.
7. Su B, Shang M, Grewe PM, et al. Suppression and restoration of primordial germ cell marker gene expression in channel catfish, *Ictalurus punctatus*, using knockdown constructs regulated by copper transport protein gene promoters: Potential for reversible transgenic sterilization. *Theriogenology*. 2015;84(9):1499-1512.
8. Li C, Gowan S, Anil A, et al. Discovery and validation of gene-linked diagnostic SNP markers for assessing hybridization between Largemouth bass (*Micropterus salmoides*) and Florida bass (*M. floridanus*). *Mol Ecol Resour*. 2015;15(2):395-404.
9. Sun L, Liu S, Wang R, et al. Identification and analysis of genome-wide SNPs provide insight into signatures of selection and domestication in channel catfish (*Ictalurus punctatus*). *PLoS One*. 2014;9(10):e109666.
10. Montes I, Conklin D, Albaina A, et al. SNP discovery in European anchovy (*Engraulis encrasicolus*, L) by high-throughput transcriptome and genome sequencing. *PLoS One*. 2013;8(8):e70051.
11. Zhang J, Sun F, Jiang Y, et al. Functional Genomics in Aquaculture. USA; 2012. p. 41-77.
12. Gary Carvalho, Eoin Mac Aoidh, et al. Traceability of Fish Populations and Fish Products: Advances and Contribution to Sustainable Fisheries. FishPopTrace, EU; 2013. p. 1-40.

13. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–2079.
14. Kofler R, Pandey RV, Schlotterer C. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*. 2011;27(24):3435–3436.
15. Wang S, Sha Z, Sonstegard TS, et al. Quality assessment parameters for EST-derived SNPs from catfish. *BMC Genomics*. 2008;9:450 p.
16. Hale MC, Mc Cormick CR, Jackson JR, et al. Next-generation pyrosequencing of gonad transcriptomes in the polyploid lake sturgeon (*Acipenser fulvescens*): the relative merits of normalization and rarefaction in gene discovery. *BMC Genomics*. 2009;10:203 p.
17. Everett MV, Grau ED, Seeb JE. Short reads and nonmodel species: exploring the complexities of next-generation sequence assembly and SNP discovery in the absence of a reference genome. *Mol Ecol Resour*. 2011;11(S 1):93–108.
18. Seeb JE, Pascal CE, Grau ED, et al. Transcriptome sequencing and high-resolution melt analysis advance single nucleotide polymorphism discovery in duplicated salmonids. *Mol Ecol Resour*. 2011;11(2):335–348.
19. Renaut S, Nolte AW, Rogers SM, et al. SNP signatures of selection on standing genetic variation and their association with adaptive phenotypes along gradients of ecological speciation in lake whitefish species pairs (*Coregonus* spp.). *Mol Ecol*. 2011;20(3):545–559.
20. Hubert S, Higgins B, Borza T, et al. Development of a SNP resource and a genetic linkage map for Atlantic cod (*Gadus morhua*). *BMC Genomics*. 2010;11:191 p.
21. Milano I, Babbucci M, Panitz F, et al. Novel tools for conservation genomics: comparing two high-throughput approaches for SNP discovery in the transcriptome of the European hake. *PLoS One*. 2011;6(11):e28008.
22. Millan A, Gomez Tato A, Pardo BG, et al. Gene expression profiles of the spleen, liver, and head kidney in turbot (*Scophthalmus maximus*) along the infection process with *Aeromonas salmonicida* using an immune-enriched oligo-microarray. *Mar Biotechnol (NY)*. 2011;13(6):1099–1114.
23. Helyar SJ, Hemmer Hansen J, Bekkevold D, et al. Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Mol Ecol Resour*. 2011;11(S1):123–136.
24. Roberts SB, Hauser L, Seeb LW, et al. Development of genomic resources for Pacific Herring through targeted transcriptome pyrosequencing. *PLoS One*. 2012;7(2):e30908.
25. Becker S, Hanner R, Steinke D. Five years of FISH-BOL: brief status report. *Mitochondrial DNA*. 2011;22(S1):3–9.
26. Gut IG, Lathrop GM. Duplicating SNPs. *Nat Genet*. 2004;36(8):789–790.
27. Langmead B, Schatz MC, Lin J, et al. Searching for SNPs with cloud computing. *Genome Biol*. 2009;10(11):R134.
28. Li C, Waldbieser G, Bosworth B, et al. SNP discovery in wild and domesticated populations of blue catfish, *Ictalurus furcatus*, using genotyping-by-sequencing and subsequent SNP validation. *Mol Ecol Resour*. 2014;14(6):1261–1270.
29. Nielsen R, Korneliussen T, Albrechtsen A, et al. SNP calling, genotype calling, and sample allele frequency estimation from New-Generation Sequencing data. *PLoS One*. 2012;7(7):e37558.
30. Yancy HF, Zemlak TS, Mason JA, et al. Potential use of DNA barcodes in regulatory science: applications of the Regulatory Fish Encyclopedia. *J Food Prot*. 2008;71(1):210–217.
31. Lowenstein JH, Burger J, Jeitner CW, et al. DNA barcodes reveal species-specific mercury levels in tuna sushi that pose a health risk to consumers. *Biol Lett*. 2010;6(5):692–695.
32. Yoon M, Choi YS, Jin HJ, et al. Heteropolymorphism of mitochondrial NADH dehydrogenase subunit 3 gene for the population analysis of chum salmon, *Oncorhynchus keta*. *J Environ Biol*. 2008;29(4):567–570.
33. Engler M, Knapp A. Briefing on the evolution of the caviar trade and range state implementation of Resolution Conf. 12.7 (Rev. Cop 14). European Commission2008, Belgium; 2008. p. 1–109.