CYTOCHROME B GENE SEQUENCES AND MITOCHONDRIAL CONTROL REGION SEQUENCES ANALYSIS IN *ALBURNUS ALBURNUS* AND *VIMBA VIMBA* INDIVIDUALS

LAURA BUBURUZAN[1]*, LUCIAN D. GORGAN[1], VLAD ARTENIE[1]

Key words: mtDNA, *Alburnus alburnus*, *Vimba vimba*, cytochrome b gene, mitochondrial control region.

Abstract: The present study presents and compares two types of gene sequences belonging to the mitochondrial DNA, cytochrome b gene and the mitochondrial control region. For the extraction of DNA two species of fishes were used: *Alburnus alburnus* and *Vimba vimba*, both belonging to the family *Cyprinidae*. Comparing the gene sequences we were able to observe the mutational changes and differences that appeared in the gene structure of individuals coming from the same population.

INTRODUCTION

The genetical structure of a population and its evolutionary path can be studied using the molecular investigations. It has been proved that mitochondrial DNA has a greater efficiency in obtaining interesting phylogenetical or phylogeographical results. In the phylogeny studies the most used is mitochondrial DNA because of its great number of nucleotide sequences polymorphisms (Birki, 2001; Yamada et al., 2001; Yhang et al., 2006; Murgia et al., 2002; Nelson et al., 2000; Oleinik et al., 2004; Parson et al., 2000; Pavlov, 2004).

Due to the high reproduction rate and prolificity, essential qualities for the populational genetics studies, fishes represent an important and interesting study material. Family *Cyprinidae* has many economically important species, that is why it has been chosen for investigations.

The main objective of this study is the sequencing of the cytochrome b gene and the mitochondrial control region, the description of the haplotypes and the comparison between the obtained haplotypes. The genes belong to individuals collected from two biogeographic areas that are little scientifically investigated from the genetical point of view: Suceava and Siret rivers

MATERIALS AND METHODS

The species used for the research study are *Alburnus alburnus* Linnaeus 1758 and *Vimba vimba* Linnaeus 1758, belonging to the family *Cyprinidae*. The *Alburnus alburnus* individuals were collected from Suceava river and the *Vimba vimba* individuals belong to Siret river.

The biological material is represented by a dorsal muscle fragment for each individual, fragment which was kept in ethylic alcohol until the DNA extraction procedure started. The mitochondrial DNA sequences investigation started with total DNA extraction from tissues that were kept in ethanol (Ausubel et al., 1995). The desired sequences were amplified by PCR reaction with its three steps: denaturation of the sequences at high temperatures, binding of the primers at lower temperature and finally the extension with the help of Taq polymerase. The primers used for amplification were universal primers for the mitochondrial control region and for the gene that determines cytochrome b syntesis. The resulting PCR products were tested through agarose gel electrophoresis (for verifying if the amplification was correct), than purified in QIAGEN colomns, quantified through electrophoresis and sequenced.

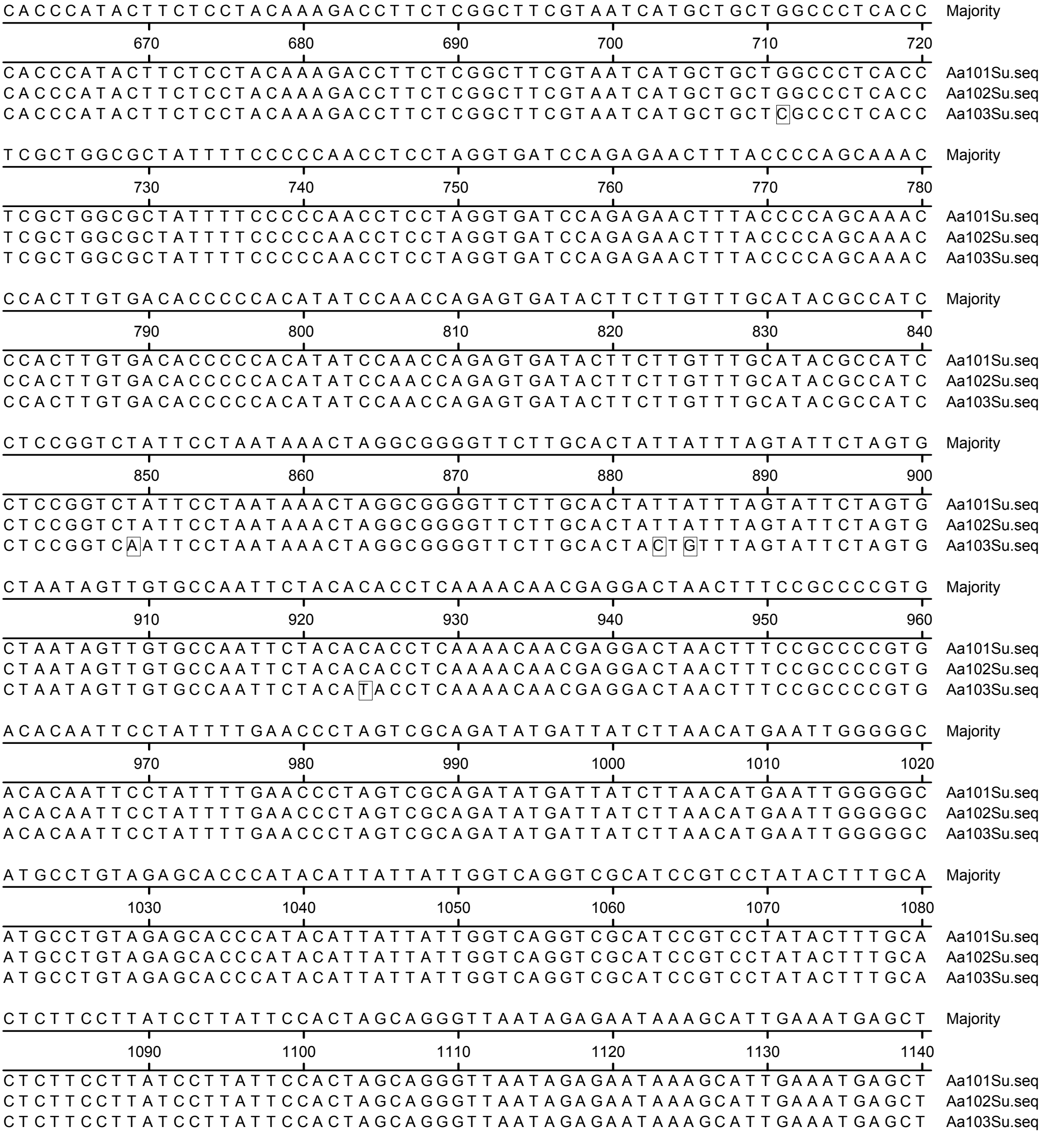
The sequencing process was implemented with an eight capillars Beckman-Coulter sequenator. There have been sequenced three nucleotidic fragments of cytochrome b gene for *Alburnus alburnus* and 2 nucleotidic fragments of the mitochondrial control region for *Vimba vimba*.

RESULTS AND DISSCUSIONS

The cytochrome b gene sequences analysis in *Alburnus alburnus* individuals

The three nucleotidic fragments sequenced were aligned using Clustal V method (Higgins and Sharp, 1989) and can be observed in figure 1.

```
CTTCCAACACCATCTAACATTTCAGCAATATGAAATTTCGGATCCCTTCTAGGGTTATGT   Majority
        70        80        90       100       110       120
CTTCCAACACCATCTAACATTTCAGCAATATGAAATTTCGGATCCCTTCTAGGGTTATGT   Aa101Su.seq
CTTCCAACACCATCTAACATTTCAGCAATATGAAATTTCGGATCCCTTCTAGGGTTATGT   Aa102Su.seq
CTTCCAACACCATCTAACATTTCAGCAATATGAAATTTCGGATCCCTTCTAGGGTTATGT   Aa103Su.seq

TTAATTACCCAAATTCTAACAGGGTTATTCCTAGCCATACACTACACCTCCGATATCTCA   Majority
       130       140       150       160       170       180
TTAATTACCCAAATTCTAACAGGGTTATTCCTAGCCATACACTACACCTCCGATATCTCA   Aa101Su.seq
TTAATTACCCAAATTCTAACAGGGTTATTCCTAGCCATACACTACACCTCCGATATCTCA   Aa102Su.seq
TTAATTACCCAAATCCTAACAGGGTTATTCCTAGCCATACACTACACCTCCGATATCTCA   Aa103Su.seq

ACCGCATTCTCATCAGTCACCCATATTTGCCGGGACGTTAACTACGGCTGGCTCATTCGA   Majority
       190       200       210       220       230       240
ACCGCATTCTCATCAGTCACCCATATTTGCCGGGACGTTAACTACGGCTGGCTCATTCGA   Aa101Su.seq
ACCGCATTCTCATCAGTCACCCATATTTGCCGGGACGTTAACTACGGCTGGCTCATTCGA   Aa102Su.seq
ACCGCATTCTCATCAGTCACCCATATTTGCCGGGACGTTAACTACGGCTGGCTCATTCGA   Aa103Su.seq

AACCTACATGCCAACGGAGCATCCTTCTTCTTCATCTGCCTATATATGCATATCGCACGA   Majority
       250       260       270       280       290       300
AACCTACATGCCAACGGAGCATCCTTCTTCTTCATCTGCCTATATATGCATATCGCACGA   Aa101Su.seq
AACCTACATGCCAACGGAGCATCCTTCTTCTTCATCTGCCTATATATGCATATCGCACGA   Aa102Su.seq
AACCTACATGCCAACGGAGCATCCTTCTTCTTCATCTGCCTATATATGCATATCGCACGA   Aa103Su.seq

GGTCTATATTACGGCTCATATCTTTATAAAGAGACCTGAAACATTGGGGTAGTACTATTT   Majority
       310       320       330       340       350       360
GGTCTATATTACGGCTCATATCTTTATAAAGAGACCTGAAACATTGGGGTAGTACTATTT   Aa101Su.seq
GGTCTATATTACGGCTCATATCTTTATAAAGAGACCTGAAACATTGGGGTAGTACTATTT   Aa102Su.seq
GGTCTATATTACGGCTCATATCTTTATAAAGAGACCTGAAACATTGGGGTAGTACTATTT   Aa103Su.seq

CTTCTGGTTATGATAACAGCCTTCGTGGGCTATGTACTCCCATGAGGACAAATATCCTTT   Majority
       370       380       390       400       410       420
CTTCTGGTTATGATAACAGCCTTCGTGGGCTATGTACTCCCATGAGGACAAATATCCTTT   Aa101Su.seq
CTTCTGGTTATGATAACAGCCTTCGTGGGCTATGTACTCCCATGAGGACAAATATCCTTT   Aa102Su.seq
CTTCTGGTTATGATAACAGCCTTCGTGGGCTATGTACTCCCATGGGGACAAATATCCTTT   Aa103Su.seq

TGAGGTGCTACCGTAATCACGAACCTCCTCTCAGCAGTTCCCTACATGGGAGATACCCTT   Majority
       430       440       450       460       470       480
TGAGGTGCTACCGTAATCACGAACCTCCTCTCAGCAGTTCCCTACATGGGAGATACCCTT   Aa101Su.seq
TGAGGTGCTACCGTAATCACGAACCTCCTCTCAGCAGTTCCCTACATGGGAGATACCCTT   Aa102Su.seq
TGAGGCGCTACCGTAATCACGAACCTCCTCTCAGCAGTTCCCTACATGGGAGATACCCTT   Aa103Su.seq

GTTCAATGAATTTGGGGCGGTTTCTCAGTAGATAACGCGACTCTTACGCGATTCTTCGCG   Majority
       490       500       510       520       530       540
GTTCAATGAATTTGGGGCGGTTTCTCAGTAGATAACGCGACTCTTACGCGATTCTTCGCG   Aa101Su.seq
GTTCAATGAATTTGGGGCGGTTTCTCAGTAGATAACGCGACTCTTACGCGATTCTTCGCG   Aa102Su.seq
GTTCAATGAATTTGGGGCGGTTTCTCAGTAGATAACGCGACTCTTACCCGATTCTTCGCG   Aa103Su.seq

TTCCACTTCCTCCTGCCGTTCGTCGTTGCAGGCGCAACCGTCCTCCACTTACTATTCCTA   Majority
       550       560       570       580       590       600
TTCCACTTCCTCCTGCCGTTCGTCGTTGCAGGCGCAACCGTCCTCCACTTACTATTCCTA   Aa101Su.seq
TTCCACTTCCTCCTGCCGTTCGTCGTTGCAGGCGCAACCGTCCTCCACTTACTATTCCTA   Aa102Su.seq
TTCCACTTCCTCCTGCCGTTCGTCGTTGCAGGCGCAACCGTCCTCCACTTACTATTCCTA   Aa103Su.seq

CACGAGACAGGATCAAACAACCCAGCCGGGTTAAACTCTGACGCGGATAAAATTTCTTTC   Majority
       610       620       630       640       650       660
CACGAGACAGGATCAAACAACCCAGCCGGGTTAAACTCTGACGCGGATAAAATTTCTTTC   Aa101Su.seq
CACGAGACAGGATCAAACAACCCAGCCGGGTTAAACTCTGACGCGGATAAAATTTCTTTC   Aa102Su.seq
CACGAGACAGGATCAAACAACCCAGCCGGGTTAAACTCTGACGCGGATAAAATTTCTTTC   Aa103Su.seq
```



```
CACCCATACTTCTCCTACAAAGACCTTCTCGGCTTCGTAATCATGCTGCTGGCCCTCACC  Majority
         670       680       690       700       710       720
CACCCATACTTCTCCTACAAAGACCTTCTCGGCTTCGTAATCATGCTGCTGGCCCTCACC  Aa101Su.seq
CACCCATACTTCTCCTACAAAGACCTTCTCGGCTTCGTAATCATGCTGCTGGCCCTCACC  Aa102Su.seq
CACCCATACTTCTCCTACAAAGACCTTCTCGGCTTCGTAATCATGCTGCTCGCCCTCACC  Aa103Su.seq

TCGCTGGCGCTATTTTCCCCCAACCTCCTAGGTGATCCAGAGAACTTTACCCCAGCAAAC  Majority
         730       740       750       760       770       780
TCGCTGGCGCTATTTTCCCCCAACCTCCTAGGTGATCCAGAGAACTTTACCCCAGCAAAC  Aa101Su.seq
TCGCTGGCGCTATTTTCCCCCAACCTCCTAGGTGATCCAGAGAACTTTACCCCAGCAAAC  Aa102Su.seq
TCGCTGGCGCTATTTTCCCCCAACCTCCTAGGTGATCCAGAGAACTTTACCCCAGCAAAC  Aa103Su.seq

CCACTTGTGACACCCCCACATATCCAACCAGAGTGATACTTCTTGTTTGCATACGCCATC  Majority
         790       800       810       820       830       840
CCACTTGTGACACCCCCACATATCCAACCAGAGTGATACTTCTTGTTTGCATACGCCATC  Aa101Su.seq
CCACTTGTGACACCCCCACATATCCAACCAGAGTGATACTTCTTGTTTGCATACGCCATC  Aa102Su.seq
CCACTTGTGACACCCCCACATATCCAACCAGAGTGATACTTCTTGTTTGCATACGCCATC  Aa103Su.seq

CTCCGGTCTATTCCTAATAAACTAGGCGGGGTTCTTGCACTATTATTTAGTATTCTAGTG  Majority
         850       860       870       880       890       900
CTCCGGTCTATTCCTAATAAACTAGGCGGGGTTCTTGCACTATTATTTAGTATTCTAGTG  Aa101Su.seq
CTCCGGTCTATTCCTAATAAACTAGGCGGGGTTCTTGCACTATTATTTAGTATTCTAGTG  Aa102Su.seq
CTCCGGTCAATTCCTAATAAACTAGGCGGGGTTCTTGCACTACTGTTTAGTATTCTAGTG  Aa103Su.seq

CTAATAGTTGTGCCAATTCTACACACCTCAAAACAACGAGGACTAACTTTCCGCCCCGTG  Majority
         910       920       930       940       950       960
CTAATAGTTGTGCCAATTCTACACACCTCAAAACAACGAGGACTAACTTTCCGCCCCGTG  Aa101Su.seq
CTAATAGTTGTGCCAATTCTACACACCTCAAAACAACGAGGACTAACTTTCCGCCCCGTG  Aa102Su.seq
CTAATAGTTGTGCCAATTCTACATACCTCAAAACAACGAGGACTAACTTTCCGCCCCGTG  Aa103Su.seq

ACACAATTCCTATTTTGAACCCTAGTCGCAGATATGATTATCTTAACATGAATTGGGGGC  Majority
         970       980       990      1000      1010      1020
ACACAATTCCTATTTTGAACCCTAGTCGCAGATATGATTATCTTAACATGAATTGGGGGC  Aa101Su.seq
ACACAATTCCTATTTTGAACCCTAGTCGCAGATATGATTATCTTAACATGAATTGGGGGC  Aa102Su.seq
ACACAATTCCTATTTTGAACCCTAGTCGCAGATATGATTATCTTAACATGAATTGGGGGC  Aa103Su.seq

ATGCCTGTAGAGCACCCATACATTATTATTGGTCAGGTCGCATCCGTCCTATACTTTGCA  Majority
        1030      1040      1050      1060      1070      1080
ATGCCTGTAGAGCACCCATACATTATTATTGGTCAGGTCGCATCCGTCCTATACTTTGCA  Aa101Su.seq
ATGCCTGTAGAGCACCCATACATTATTATTGGTCAGGTCGCATCCGTCCTATACTTTGCA  Aa102Su.seq
ATGCCTGTAGAGCACCCATACATTATTATTGGTCAGGTCGCATCCGTCCTATACTTTGCA  Aa103Su.seq

CTCTTCCTTATCCTTATTCCACTAGCAGGGTTAATAGAGAATAAAGCATTGAAATGAGCT  Majority
        1090      1100      1110      1120      1130      1140
CTCTTCCTTATCCTTATTCCACTAGCAGGGTTAATAGAGAATAAAGCATTGAAATGAGCT  Aa101Su.seq
CTCTTCCTTATCCTTATTCCACTAGCAGGGTTAATAGAGAATAAAGCATTGAAATGAGCT  Aa102Su.seq
CTCTTCCTTATCCTTATTCCACTAGCAGGGTTAATAGAGAATAAAGCATTGAAATGAGCT  Aa103Su.seq
```

Figure 1. The alignment of cytochrome b gene sequences in *Alburnus alburnus* individuals

Each sequence has a number of 1140 nucleotides. The haplotype Aa101Su and Aa102Su are identical and do not have any mutational modifications.

The third sequence, Aa103Su has a number of nine mutational modifications, all being substitutions of a purinic base with another purine or of a pyrimidinic base with another pyrimidine (transitions), or substitutions of a purine with a pyrimidine or of a pyrimidine with a purine (transversions).

From these nine substitutions six are transitions and three are transvesions. Purine transtions took place in locus 135, 426, 883 (where thymine was replaced by cytosine), 405, 885 (where adenine was replaced by guanine) and 924 (where cytosine was replaced by thymine). In two of the three transvesions cytosine replaced guanine (in locuses 528 and 711) and in locus 849 adenine replaced thymine.

Comparing the number and the percentage of nucleobases in the three different haplotypes (table 1), it can be noticed that the most frequent base is thymine in the first two haplotypes (28.75%) and cytosine for haplotype Aa103Si. It is interesting the fact that the percentage of pyrimidines is higher than the one of purines. The most rare base is guanine for all the three haplotypes. Regarding the number of complementary bases the percentage of A+T is higher than that of the G+C. This fact demonstrates that the bounds between the complementary strains are not very strong and the DNA fragment is not very stabile.

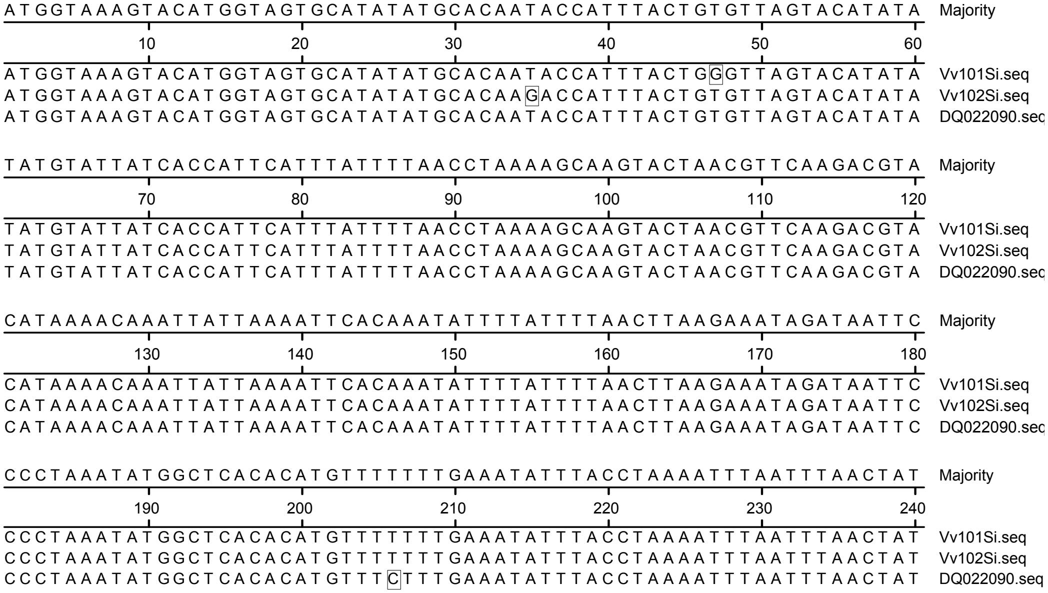
Table 1. The number and percentage of nucleobases for the studied haplotypes

Analysed sequences	Nucleobases											
	A		G		T		C		A+T		C+G	
	Nr.	%	Nr.	%	Nr.	%	Nr.	%	Nr.	%	Nr.	%
Aa101Su	297	26.03	195	17.09	328	28.75	321	28.13	625	54.78	516	45.22
Aa102Su	297	26.03	195	17.09	328	28.75	321	28.13	625	54.78	516	45.22
Aa103Su	296	25.96	195	17.11	324	28.42	325	28.51	620	54.39	520	45.61

The mitochondrial control region sequences analysis in*Vimba vimba* individuals

For the mitochondrial control region two haplotypes have been sequences (Vv101Si and Vv102Si) and compared with a sequence from GenBank (with the code number Dq022090), used as control.

The three nucleotidic sequences were aligned using Clustal V method (Higgins and Sharp, 1989) and can be observed in figure 2.



```
ATGGTAAAGTACATGGTAGTGCATATATGCACAATACCATTTACTGTGTTAGTACATATA  Majority
         10        20        30        40        50        60
ATGGTAAAGTACATGGTAGTGCATATATGCACAATACCATTTACTGGGTTAGTACATATA  Vv101Si.seq
ATGGTAAAGTACATGGTAGTGCATATATGCACAAGACCATTTACTGTGTTAGTACATATA  Vv102Si.seq
ATGGTAAAGTACATGGTAGTGCATATATGCACAATACCATTTACTGTGTTAGTACATATA  DQ022090.sec

TATGTATTATCACCATTCATTTATTTTAACCTAAAAGCAAGTACTAACGTTCAAGACGTA  Majority
         70        80        90       100       110       120
TATGTATTATCACCATTCATTTATTTTAACCTAAAAGCAAGTACTAACGTTCAAGACGTA  Vv101Si.seq
TATGTATTATCACCATTCATTTATTTTAACCTAAAAGCAAGTACTAACGTTCAAGACGTA  Vv102Si.seq
TATGTATTATCACCATTCATTTATTTTAACCTAAAAGCAAGTACTAACGTTCAAGACGTA  DQ022090.sec

CATAAAACAAATTATTAAAATTCACAAATATTTTATTTTAACTTAAGAAATAGATAATTC  Majority
        130       140       150       160       170       180
CATAAAACAAATTATTAAAATTCACAAATATTTTATTTTAACTTAAGAAATAGATAATTC  Vv101Si.seq
CATAAAACAAATTATTAAAATTCACAAATATTTTATTTTAACTTAAGAAATAGATAATTC  Vv102Si.seq
CATAAAACAAATTATTAAAATTCACAAATATTTTATTTTAACTTAAGAAATAGATAATTC  DQ022090.sec

CCCTAAATATGGCTCACACATGTTTTTTTGAAATATTTACCTAAAATTTAATTTAACTAT  Majority
        190       200       210       220       230       240
CCCTAAATATGGCTCACACATGTTTTTTTGAAATATTTACCTAAAATTTAATTTAACTAT  Vv101Si.seq
CCCTAAATATGGCTCACACATGTTTTTTTGAAATATTTACCTAAAATTTAATTTAACTAT  Vv102Si.seq
CCCTAAATATGGCTCACACATGTTTCTTTGAAATATTTACCTAAAATTTAATTTAACTAT  DQ022090.seq
```

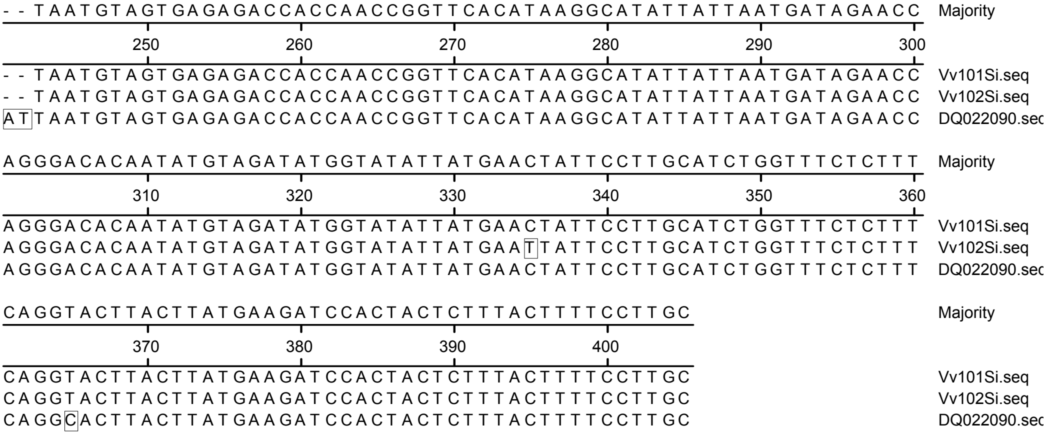


Figure 2. The alignment of mitochondrial control region sequences in *Vimba vimba* individuals

Each fragment has 405 nucleotides. The first haplotype Vv101Si has three mutational modifications; one of these is a transvesion in locus 47, where thymine was replaced by guanine. The other two substitutions are transitions and occured in the same locuses as in the other sequenced fragment Vv102Si: 206 and 365. In these locuses cytosine was replaced by thymine. It can be observed that both sequences belonging to *Vimba vimba* individuals from Siret river have these two locuses with identical mutations compared with the control sequence. This is very important because it means that these two mutations could be fixed inside the genome of Siret river *Vimba vimba* populaton.

The second haplotype Vv102Si has four substitutions, two in the same locuses as the first one: 206 and 365, one transition in locus 335 where thymine replaced cytosine and a transversion in locus 35 where guanine replaced thymine.

An interesting thing can be noticed in locuses 241 and 242 for both sequences. In these places another type of mutation occured: deletion of nucleotides. In the control sequence these locuses were populated with a purine, adenine (locus 241) and a pyrimidine, thymine (locus 242), but both sequenced fragments lack these nucleotides.

Comparing the number and the percentage of nucleobases in the three different haplotypes (table 2), it can be noticed that the most frequent bases are adenine and thymine in almost equal proportions, between 35,24 and 35,82%. The most rare base is guanine for all the three haplotypes. The number of purines is higher then the one of pyrimidines but not with a very large difference.

Regarding the number of complementary bases the percentage of A+T is much more higher than that of the G+C, comparing with the proportion found in cytochrome b genes. This fact demonstrates that the bounds between the complementary strains are weaker than the ones in analyzed cytochrome b genes.

Table 2. The number and percentage of nucleobases for the studied haplotypes

Analysed sequences	Nucleobases											
	A		G		T		C		A+T		C+G	
	Nr.	%	Nr.	%	Nr.	%	Nr.	%	Nr.	%	Nr.	%
Vv101Si	142	35,24	51	12,66	143	35,48	67	16,63	285	70,72	118	29,28
Vv102Si	142	35,32	51	12,69	144	35,82	65	16,17	286	71,14	116	28,86
DQ022090	143	35,31	50	12,35	143	35,31	69	17,04	286	70,62	119	29,38

CONCLUSIONS

Analyzing the cytochrome b gene in three individuals of *Alburnus alburnus* it has been noticed that only one has mutational modifications and this is not a certain proof that these mutations could be fixed in the Suceava river population. This is not the situation with the two haplotypes of mitochondrial control region in *Vimba vimba* population, where the same mutations occurred for both analyzed individuals. This could be a clue that these mutation can fix inside the genome of *Vimba vimba* population.

Both types of genes, cytochrome b gene and mitochondrial control region contain a larger number of A+T nucleotides than G+C, which proves that the DNA fragments is not very stabile.

REFERENCES

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., 1995. *Current protocols in molecular biology*, vol. 1, cap. 2 – *Preparation and analysis of DNA. Phenol extraction and ethanol precipitation of DNA*, Ed by John Wiley & Sons, Inc., 211 – 213.

Birky, W.C., 2001. *The inheritance of genes in mitochondria and chloroplasts: Laws, Mechanisms and models*, *Annu. Rev. Genet.* 35, 125 – 148.

Higgins,D.G., Sharp, P.M., 1989. *Fast and sensitive multiple sequence alignments on a microcomputer*, CABIOS 5(2), 151-153.

Murgia, R., Tola, G., Archer, S.N., Vallerga, S., Hirano, J., 2002. *Genetic Identification of Grey Mullet Species (Mugilidae) by Analysis of Mitochondrial DNA Sequence: Application to Identify the Origin of Processed Ovary Products (Bottarga)*, Marine Biotechnology 4(2), 119-126.

Nelson, J.S., Hoddell, R.J., Chou, L.M., Chan, W.K., Phang, V.P.E., 2000. *Phylogeographic structure of false clownfish, Amphiprion ocellaris, explained by sea level changes on the Sunda shelf, Marine Biology*137(4), 727-736.

Oleinik, A.G., Skurikhina, L.A., Frolov, S.V., Brykov, V.A., Chereshnev, I.A., 2004. *Differences Between two Subspecies of Dolly Varden, Salvelinus malma, Revealed by RFLP—PCR Analysis of Mitochondrial DNA, Environmental Biology of Fishes* 69(1-4), 449-459.

Parson, W., Pegoraro, K., Niederstätter, H., Föger, M., Steinlechner, M., 2000. *Species identification by means of the cytochrome b gene, International Journal of Legal Medicine* 114(1-2), 23-28.

Pavlov, S.D., Kolesnikov, A.A., Melnikova, M.N., Ushakova, M.V., 2004. *Genetic divergence of mykizha (Parasalmo (Oncorhynchus) mykiss) from Kamchatka inferred from restriction analysis and sequencing of mtDNA cytochrome b gene, Russian Journal of Genetics* 40(12), 1407-1412.

Yamada, M., Higuchi, M., Goto, A., 2001. *Extensive Introgression of Mitochondrial DNA Found between two Genetically Divergent Forms of Threespine Stickleback, Gasterosteus Aculeatus, around Japan, Environmental Biology of Fishes* 61(3), 269-284.

Yang, J., He, S., Freyhof, J., Witte, K., Liu, H., 2006. *The Phylogenetic Relationships of the Gobioninae (Teleostei: Cyprinidae) Inferred from Mitochondrial Cytochrome b Gene Sequences. Hydrobiologia* 553(1), 255-266.

1 – University „Alexandru Ioan Cuza" Iasi
* laura_sv2002@yahoo.com