# Cloning and Phylogenetic Analysis of Ryanodine Receptor Orthologues in Bichir (*Polypterus ornatipinnis)*

Siavash Darbandi

Supervisor: Dr. J. Franck

A thesis submitted in partial fulfilment of the Honour Thesis (05.4111/6) Course

Department of Biology

The University of Winnipeg

2008

# Abstract

Ryanodine receptors (RyR) mediate the controlled release of intracellular stores of calcium from the sarcoplasmic reticulum. This release of calcium, triggered by membrane-depolarization, is responsible for initiating muscle contraction. Three isoforms of RyRs have been identified in mammals, RyR1 found predominantly in skeletal muscle, RyR2 in cardiac muscle, and RyR3 which is ubiquitously distributed. The RyR1 and RyR3 isoforms are co-expressed at equal levels in fish skeletal muscle. Additionally, fish express fiber type-specific RyR1 isoforms in fast-twitch and slow-twitch muscles, termed RyR1-fast and RyR1-slow respectively. These isoforms are presumed to be the result of sub-functionalization following a gene duplication event in the teleost lineage. Bichir (*Polypterus ornatipinnis)* is a living representative of the primitive condition before the divergence of teleost fish. The objective of my study is to isolate partial RyR gene sequences from bichir to determine the number of RyR genes in its genome. Using a PCR-based approach with degenerate primer pairs sixteen partial RyR sequences have been cloned from bichir cDNA and genomic DNA. I have putatively identified 8 unique RyR genes from the bichir sequences, which are similar to RyR2 and RyR3. Phylogenetic analysis places these bichir sequences ancestral to more derived vertebrate RyR isoforms.

# Acknowledgements

The task of completing a biology honours thesis project is one that could not have been accomplished without the help of many spectacular people along the way.

First I would like to acknowledge the help and guidance of my supervisor Dr. J. Franck. His unbelievable patience and kindness were invaluable throughout the year and truly made the project a great success. Within the lab there is no doubt that Kelly Reimer was my go-to-guy. He was always there to lend a helping hand and guide me on my research journey.

Also, I would like to thank my committee members Dr. E. Byard and Dr. M. Wiegand who served as voices of reason and inspiration throughout the year. I am very thankful to have had committee members and a supervisor that were always willing to take the time to answer any questions I had. Big thanks to Dr. R. Moodie for organizing the thesis course and doing an excellent job.

Of course no one could survive University life without the support of friends and family. Thanks to friends and family for putting up with my constant rambling about bichir and ryanodine receptors. The support of my whole family has made my life easier and I am certain that they are proud of my accomplishments. I am very fortunate to be surrounded by wonderful people both in and out of school.

# Table of Contents


Abstract ..... ii
Acknowledgements ..... iii
Table of Contents ..... iv
List of Tables ..... v
List of Figures ..... vi
List of Appendices ..... vii
Abbreviations and Nomenclatures ..... viii
1.0 Introduction ..... 1
    1.1 Role of Ryanodine Receptors ..... 1
    1.2 Ryanodine Receptors Structure ..... 2
    1.3 Ryanodine Receptor Isoforms ..... 3
    1.4 RyR Gene Duplication Events ..... 3
    1.5 Regulation of Ryanodine Receptors Activity ..... 5
    1.6 Role of Calcium in Development ..... 6
    1.7 Ornate Bichir as Model Organism ..... 6
    1.8 Objectives ..... 7
2.0 Materials and Methods ..... 8
    2.1 RNA Extraction from the fresh sample ..... 8
    2.2 cDNA Synthesis from Extracted RNA ..... 8
    2.3 Genomic DNA Extraction ..... 9
    2.4 Polymerase Chain Reaction (DNA Amplification) ..... 9
    2.5 Primer Design ..... 11
    2.6 Cloning of PCR Products ..... 12
    2.7 PCR Detection of Recombinants and Plasmid Purification ..... 12
    2.8 Phylogenetic Analysis of Sequenced Genes ..... 13
3.0 Results ..... 15
    3.1 Amplification and Gel Purification ..... 15
    3.2 Detection of recombinants ..... 16
    3.3 Sequencing of Purified Plasmid ..... 16
    3.4 Phylogenetic Analysis of Sequenced Genes ..... 25
4.0 Discussion ..... 27
    4.1 Plasmid Sequencing ..... 27
    4.2 Phylogenetic Analysis of Sequences ..... 28
    4.3 Future Studies ..... 29
5.0 Conclusions ..... 30
6.0 References ..... 31
7.0 Appendices ..... 34

# List of Tables

1. Results of GenBank BLASTx submission for bichir nucleotide sequences ..... 21

2. Percent sequence similarity between bichir sequences and published sequences using BLASTx pairwise comparison ..... 22

3. Characterization of variable amino acid positions between bichir sequences and published sequences from Genbank ..... 24

# List of Figures

1. Three RyR are shown as viewed from the cytoplasmic face ..... 2

2. Hagfish degenerate primer pairs used to amplify RyR genes from bichir ..... 10

3. Gel showing the PCR DETECT products from *E.coli* colonies using M13 forward and M13 reverse primers ..... 13

4. PCR Product amplified using F1/R1, F1/R2, F2/R1, and F2/R2 primers on Bichir brain sample ..... 15

5. Agarose gel showing PCR products of randomly selected colonies using M13 forward and reverse primers ..... 16

6. Alignment of bichir nucleotide sequences ..... 18

7. Alignment of bichir deduced amino acid sequences ..... 20

8. Multiple amino acid alignment of bichir with selected published RyR amino acid sequences ..... 23

9. Amino acid parsimony tree based on multiple amino acid sequence alignments of cloned bichir RyR and published RyR sequences ..... 26

# List of Appendices

1. Nucleotide sequence alignment of cloned Bichir RyR consensus sequence and the published RyR sequences ..... 34

2. pGEM®-T Easy Vector Maps ..... 35

3. Amino acid alignment used for hagfish degenerate primer design..... 36

# Abbreviations and Nomenclatures

BBS: Bichir Brain Sequence

BGS: Bichir Genomic Sequence

BSM: Bichir Skeletal Muscle

cDNA: Complementary DNA (reverse transcribed from RNA)

PCR: Polymerase Chain Reaction

RyR: Ryanodine Receptors

DHPR: Dihydropyridine receptor

$Ca^{2+}$: Calcium cation

$Mg^{2+}$: Magnesium cation

TE: Tris Ethylenediaminetetraacetic acid (Tris EDTA)

dNTP: Deoxyribonucleotide triphosphate

RNasin: Ribonuclease inhibitor

$ddH_2O$: Doubled-distilled water (RNase-free water)

F: Forward primer

R: Reverse primer

HF: Hagfish degenerate primer

# 1. Introduction

## 1.1 Role of Ryanodine Receptors

Ryanodine receptors serve as an important channel in the controlled release of intracellular stores of calcium ($Ca^{2+}$) from the sarcoplasmic reticulum of muscle cells. This release of $Ca^{2+}$, triggered by membrane depolarization, is responsible for initiation of the muscle contraction process.

Excitation – contraction (EC) coupling describes the tight relationship between the excitatory nerve impulse and the resulting muscular contraction (Sherwood *et. al.*, 2005). The cytoplasmic concentration of calcium in the resting cells is much lower relative to extracellular and sarcoplasmic reticulum calcium ($Ca^{2+}$) concentration. $Ca^{2+}$ ion channels, called ryanodine receptors (RyRs), allow the release of calcium ions from the sarcoplasmic reticulum into the cytoplasm when the cell membrane is depolarized. The transverse tubules (t-tubules) of muscle cells allow RyRs to directly contact dihydropyridine receptors (DHPRs) on the cell membrane (Sherwood *et. al.*, 2005). Once a nerve impulse reaches the DHPRs, it induces a conformational change in the DHPR. This conformational change mechanically opens the RyR, a process referred to as depolarization – induced calcium release (DICR). RyR can also be activated by direct binding of $Ca^{2+}$. This is referred to as calcium – induced calcium release (CICR) (Sherwood *et. al.,* 2005).

## 1.2 Ryanodine Receptors Structure

RyRs are large homotetrameric proteins with a total molecular mass of approximately 2.2 million Da (Figure 1) (Sharma and Wagenknecht, 2004). RyRs are expressed in various tissues, but are most abundant in striated muscle (Fill and Copello, 2002). They contain at least two distinct functional domains: regulatory sites are located on the amino-terminal and of the cytoplasmic domain, which accounts for the binding sites for nucleotides, calmodulin and calcium. The cytoplasmic domain also possesses phosphorylation sites, transmembrane and pore-forming regions, which are part of the carboxy-terminal domains of RyRs (Fill and Copello, 2002).

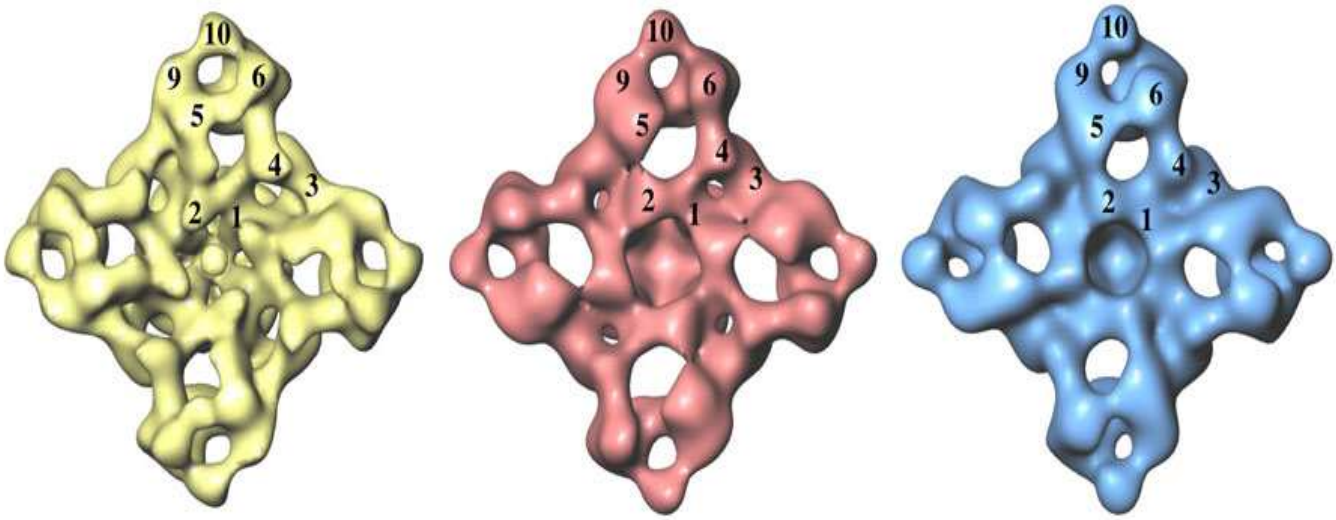



**Figure 1:** Three RyRs are shown as viewed from the cytoplasmic face. RyR1 (yellow), RyR2 (red) and RyR3 (blue). Functional domains are numbered 1-10 in one quarter of the tetramer molecule. *Adapted from Sharma and Wagenknecht, 2004.*

### 1.3 Ryanodine Receptor Isoforms

The three RyR isoforms that have been identified in mammals have an overall amino acid identity of 66-70% (Rossi and Sorrentino, 2002). The RyR isoform that is expressed primarily in skeletal muscle is referred to as RyR1, while RyR2 is predominantly expressed in cardiac muscle. The RyR3 isoform is ubiquitously expressed and has been found in the brain, diaphragm and skeletal muscle (Morrissette *et al.*, 2000). In mammals, the three RyR isoforms are products of three different genes (Fill and Copello, 2002). The RyR isoforms first identified in non-mammals (amphibians such as frog) were designated as α-RyR and β-RyR, because the homology of these isoforms to mammals was uncertain (Franck *et al.*, 1998). It has since been determined that these isoforms are homologous to RyR1 and RyR3 respectively (Morrissette *et al.*, 2000). Two distinct isoforms have been shown to be discretely expressed in slow and fast oxidative muscles of fish. These have been designated RyR1-slow (RyR1a) and RyR1-fast (RyR1b), respectively (Franck *et al.*, 1998). Slow and fast twitch muscles in fish are discretely compartmentalized which makes them a good model organism to study expression of these isoforms.

### 1.4 RyR Gene Duplication Events

During the evolution of the vertebrate lineage there is sufficient evidence that at least two rounds of whole genome duplication have taken place. Supporting evidence comes from studies of gene clusters such as *Hox* genes (Amores *et al.*, 2004; Blomme *et al.*, 2006). Following whole-genome duplication one would expect to see double the number of copies of a gene with its

neighbouring genes in a duplicated cluster maintaining the same order (i.e. same synteny). The first genome duplication may predate the Cambrian explosion (530 million years ago) in diversity in early chordate evolution coinciding with the divergence of cephalochordates (Meyer and Van de Peer, 2005; Xian-guang *et al.*, 2002), and the second duplication during the early Devonian (359 million years ago) during the evolution of vertebrates. Evidence exists for a third round of whole genome duplication in the divergence of teleost fish from sarcopterygians (Finn and Kristoffersen, 2007; Meyer and Van de Peer, 2005).

The duplication of genomes leads to one of several potential outcomes. The extra copy could subsequently become lost through purifying selection, or may be retained as a backup with no change in the function. The additional copy may also diversify to take on completely new roles by “neofunctionalization”, where a gene encodes for a novel function unrelated to its previous function, or undergo “subfunctionalization”, where a gene undertakes a new function similar to its previous function (Blomme *et al.*, 2006). It is believed that this latter mechanism is the main route in the divergence of different RyR isoforms (Sorrentino *et. al.*, 2000). The ancestral RyR gave rise to the RyR2 and RyR3, followed by the duplication and divergence of RyR1 from a RyR3 ancestral form.

The divergence of vertebrate organisms from ancestral invertebrates was accompanied by physiological changes in the EC coupling in muscle tissues (Inoue *et. al.*, 2002). Invertebrates import $Ca^{2+}$ from extracellular fluids into the cell in order for contraction to take place via calcium induced calcium release (CICR) mechanism as discussed earlier in the introduction. Vertebrate organisms

do not require extracellular $Ca^{2+}$ to enter the cell before contraction to occur (Inoue *et. al.*, 2002). During this developmental stage, direct physical association of DHPRs and RyRs appears. Thus this would result in a functional difference between the ancestral and more advanced RyR isoforms (Inoue *et. al.,* 2002). The RyRs in vertebrate skeletal muscle are mechanically gated by the voltage sensor and do not require extracellular $Ca^{2+}$ as a ligand, so called depolarization induced calcium release (DICR) (Inoue *et. al.*, 2002).

**1.5 Regulation of RyR Activity**

Various endogenous molecules can regulate the activity of RyRs (Tripathy *et al.*, 1995). RyRs can be activated or de-activated by ligands that include $Ca^{2+}$, $Mg^{2+}$, ATP and calmodulin (Tripathy *et al.*, 1995). The three RyR isoforms have shown different behaviours when under various concentrations of effector ligands. RyR1 activity is enhanced by calmodulin when the concentration of $Ca^{2+}$ is in the nanomolar range (Tripathy *et al.*, 1995), but inhibited by calmodulin when the $Ca^{2+}$ concentration is in the micromolar range (Tripathy *et. al.*, 1995). RyR2 does not show activation by calmodulin when calcium is not bound, and it is inhibited when $Ca^{2+}$ levels are in the micromolar range (Rossi and Sorrentino, 2002). The RyR2 isoform tends to be more sensitive to activation by micromolar concentrations of $Ca^{2+}$, and less sensitive to inhibition by millimolar concentrations of $Ca^{2+}$ than RyR1. RyR3 is also less sensitive to both activation and inhibition by $Ca^{2+}$. RyR2 and RyR3 are both less sensitive to inhibition by $Mg^{2+}$ than RyR1 (Morrissette *et al.*, 2000).

### 1.6 Role of Calcium in Development

Calcium plays a major role in the regulation of developmental processes. The initiation of development in arrested eggs is activated by a $Ca^{2+}$ wave (Webb and Miller, 2003). Calcium signalling has also been shown to play an important role in the segregation process during the zygote period (Webb and Miller, 2003). During this period, fast $Ca^{2+}$ waves appeared to be generated by the CICR mechanism from the endoplasmic reticulum, presumably via the RyR channels (Webb and Miller, 2003). Intracellular $Ca^{2+}$ signalling is first observed during the blastula period. In this period, localized $Ca^{2+}$ spikes are seen in the cells that eventually develop into the enveloping layer, and $Ca^{2+}$ may play a role in the formation of the yolk "syncytial" layer (Webb and Miller, 2003). Miller and Webb (2003) suggest that these waves of $Ca^{2+}$ may provide positional information to the cells, or may coordinate cell responses.

### 1.7 Ornate Bichir as a Model Organism

Bichir (*Polypterus ornatipinnis)* is a useful vertebrate model organism for the study of the evolution of physiological and developmental processes. Bichir is a representative of the most basal extant species of ray finned (Actinopterygii) fish lineage possessing ancestral morphological features (Chiu *et al.*, 2004). Fossils of *Polypterus* have been found in rocks that date back to the Triassic period (206-248 million years ago) (Budgett, 2001). This organism is also an interesting source for studying the timing of the third round of whole-genome duplication in the vertebrate lineage as there is evidence that bichir contains 5

*Hox* gene clusters in its genome, whereas teleost have seven and other derived organisms such as mammals have four in their genome (Ledje, *et al.* 2002).

**1.8 Objectives**

The objective of this study is to survey the bichir genome for the presence and number of RyR genes. The genes will be annotated by similarity searches against the GenBank database and will be analyzed using phylogenetic methodologies.

# 2. Materials and Methods

## 2.1 RNA Extraction from Fresh Bichir Samples:

A single bichir specimen was obtained from a local tropical fish store (Fish Gallery, Winnipeg) and sacrificed by a protocol approved by the Senate Animal Care Committee issued to Dr. Franck (approved September 28, 2007). The bichir was euthanized immediately after being received in the lab by immersion in MS222 (Tricaine sulfate) followed by pithing. Tissues were dissected from the bichir using a dissecting microscope. Skeletal muscle, cardiac muscle, brain and liver were carefully dissected and instantly frozen in liquid nitrogen. 100 mg of tissue was homogenized in 1 ml of TRIzol reagent (Invitrogen Life Technologies). The tissues were homogenized with a motorized Teflon pestle and glass mortar. TRIzol reagent uses phenol chloroform, to separate RNA to be isolated from other cellular components (primarily protein) (Chomczynski and Sacchi, 1987). Total RNA was precipitated with 95% ethanol, pelleted by centrifugation at maximum speed (14,000 g) and dried in a speed-vacuum evaporation. The dried RNA was resuspended in RNase free water, and quantified by ultraviolet spectrophotometry ($UV_{260}$ was measured) (Birds, 2005).

## 2.2 cDNA Synthesis from Extracted RNA:

10 - 20 μg of total RNA was used as template to synthesize first strand cDNA. A 20 μl solution of the total RNA and 0.005 μg/μl of Oligo $dT_{15}$ primer was heated to 95°C for five minutes and immediately chilled on ice. Once the solution was chilled, 12 μl of 5x Oligo $dT_{15}$ buffer, 2 μl of 10mM dNTPs, 24 μl $ddH_2O$, 1 μl

RNAsin and 1 μl of reverse transcriptase were added and mixed. The solution was incubated at room temperature for 15 minutes, followed by 30 minutes incubation at 37°C. The reaction was stopped by adding 40 μl of 10x TE buffer at room temperature (pH 8.0). The first strand cDNA was precipitated by adding 10 μl of 3M sodium acetate and 250 μl of 95% ethanol. The cDNA was precipitated at -20°C for several hours in this solution and pelleted by centrifugation at 14,000 g for 15 minutes, washed with 95% ethanol and centrifuged again at 6500-7000 rpm for 5 minutes. The final pellet was resuspended in 100 μl of RNase-free water. A speed vacuum centrifugation apparatus was used to dry the pellet. The pellet was dissolved in 100 μl of distilled, RNase and DNase free water.

**2.3 Genomic DNA Extraction:**

50 mg of tissue was homogenized in 1 ml of DNAzol reagent supplemented with 100 μg/ml of proteinase K (Invitrogen Life Technologies). Homogenate was centrifuged for 10 min at 10,000g at room temperature. DNA was precipitated by adding 0.5 ml of 95% ethanol per 1ml of DNAzol. After precipitation, DNA was washed with 0.8 – 1.0ml of 75% ethanol. Finally, DNA was dissolved in 0.2 – 0.3ml of sodium hydroxide (NaOH).

**2.4 Polymerase Chain Reaction (DNA Amplification):**

PCR is a process in which target genes are amplified from the tissue samples using degenerate primers. Degenerate primers were designed with the assistance of the CODEHOP program based on alignments from conserved regions of RyR message from frog, rabbit and fish (Appendix 7.3). The primers were designed to target conserved regions of the amino acid sequence so that

any expressed isoform of RyR would be amplified (Rose *et al,* 1998*).* These primers used are shown in Figure 2.

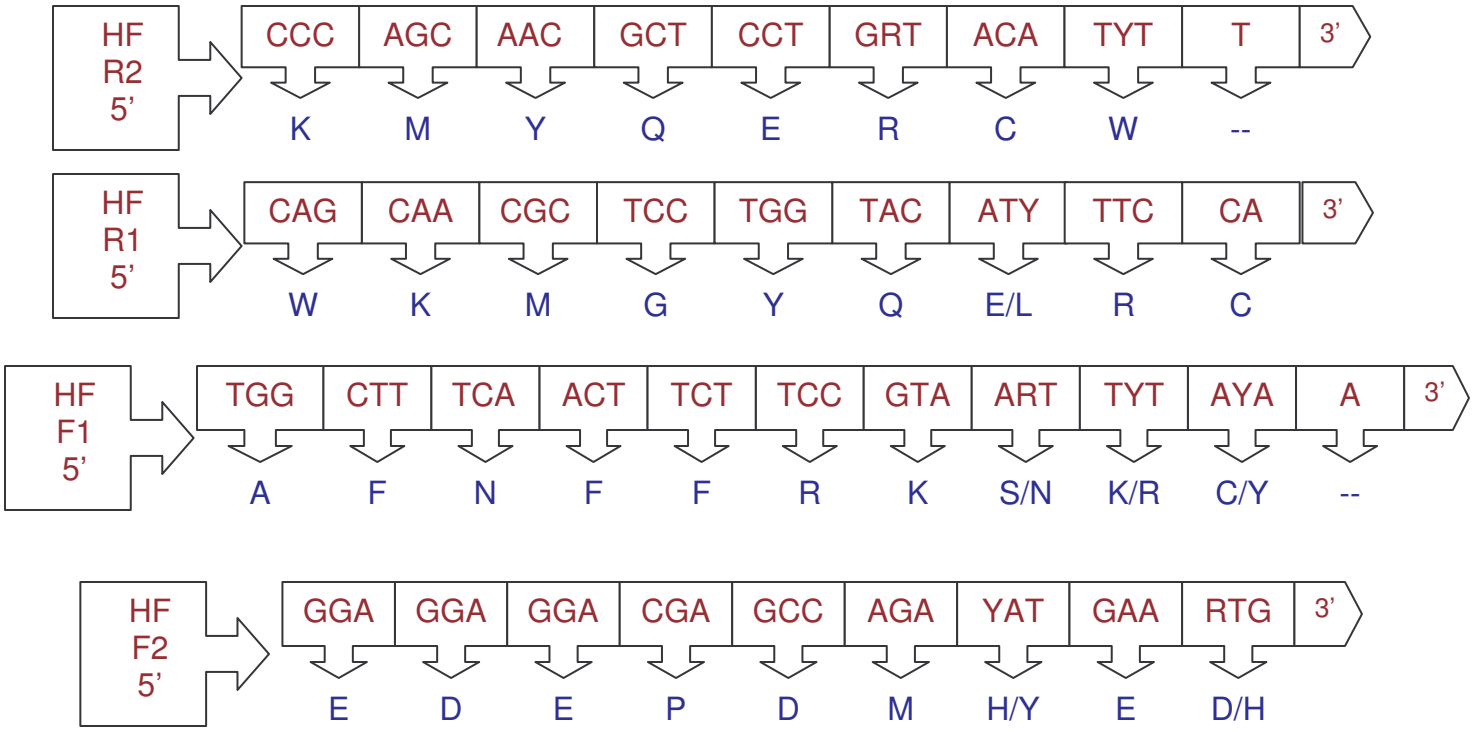



**Figure 2:** Hagfish degenerate primer pairs used to amplify the RyR genes present in cDNA and genomic DNA extracted from bichir tissue. HFF=Hagfish Forward primer, HFR=Hagfish Reverse primer. Nucleotide sequences of primers are shown in red and corresponding amino acids are in blue.

Where R is either G or A, and Y is either C or T. The calculated melting points under standard conditions for HFR1 and HFR2 were between 58℃ and 61℃, whereas HFF1 was between 55℃ and 59℃, and HFF2 was between 61℃ and 64℃. These primers were then used to amplify RyR messages from the first strand cDNA.

45 µl of master mix and 5 µl of template were used for each PCR reaction. The master mix contained 5µl of 10x PCR buffer (Invitrogen Life Technologies), 1.5 µl of 50 mM $MgCl_2$, 1.0 µl of 10 mM dNTPs, 0.5 µM of each primers, 0.5 µl

recombinant Taq DNA Polymerase and 36 µl of $ddH_2O$ (RNase free water). The initial denaturation step was 95°C for 9 minutes. The subsequent denaturation step was 95°C for 1 minute, followed by an annealing step at 60°C for 1 minute and 30 seconds and an extension step at 72°C for 1 minute. These 3 steps were cycled 35 times and a final extension step at 72°C for 7 minutes completed the reaction.

PCR reaction products were fractionated on a 1% low melting point agarose gel and purified by either gel purification (Wizard® PCR Preps DNA Purification System, Promega) or with S.N.A.P. Gel Purification Kit (Invitrogen Life Technologies).

**2.5 Primer Design:**

5 oligonucleotide primers were designed from the aligned bichir sequences. These primers were designed from regions that were unique to that particular sample type. This would provide higher degree of accuracy in terms of which gene (isoform) is expressed in each particular tissue type. The only deviation to the procedure used for the degenerate primers was that the annealing temperature for the PCR was reduced to 50°C. Those primers were as follow:

5′-CGT AAG TTT TAT AAT AAG AGC G-3′ (denoted as BBS F1)

5′-GGA TGA GGA TGA GCC TGA CAT G-3′ (denoted as BBS F2)

5′-GGA CGA GGA CGA GCC AGA CAT G-3′ (denoted as BGS F1)

3′-TCC TGG TAC ATC TTC CAT AC-5′ (denoted as BBS R1)

3′-TCC TGG TAC ATT TTC CAT AC-5′ (denoted as BBS R2)

**2.6 Cloning of PCR Products:**

The purified PCR products were ligated with pGEM® – T Easy vectors (Promega Biotech). The ligated products were transformed into competent *Escherichia coli JM109* strain *(E. coli JM109)* cells and plated on LB media plates containing 0.05 mg/ml ampicillin. Then plates were incubated at 37°C overnight (i.e. 24 hours).

**2.7 PCR Detection of Recombinants and Plasmid Purification:**

Bacterial colonies were randomly selected for screening from each plate. Each colony was touched with a sterile pipette tip and was allowed to incubate in a PCR tube containing 15 µl of PCR master mix for approximately one minute at room temperature. PCR was performed using M13 forward and M13 reverse primer pairs to amplify the cloned insert. The PCR program consist an initial denaturation step at 95°C for five minutes. The second step was 95°C for one minute, followed by an annealing step at 55°C for thirty seconds and an extension step at 72°C for one minute. The second, third and fourth steps were cycled 35 times. The reaction products were fractionated on a 1% agarose gel containing 0.5 µg/ml ethidium bromide and viewed on a UV transilluminator (Figure 3). Bacterial cultures corresponding to PCR products of expected size were grown overnight in LB media containing 0.05mg/ml ampicillin.

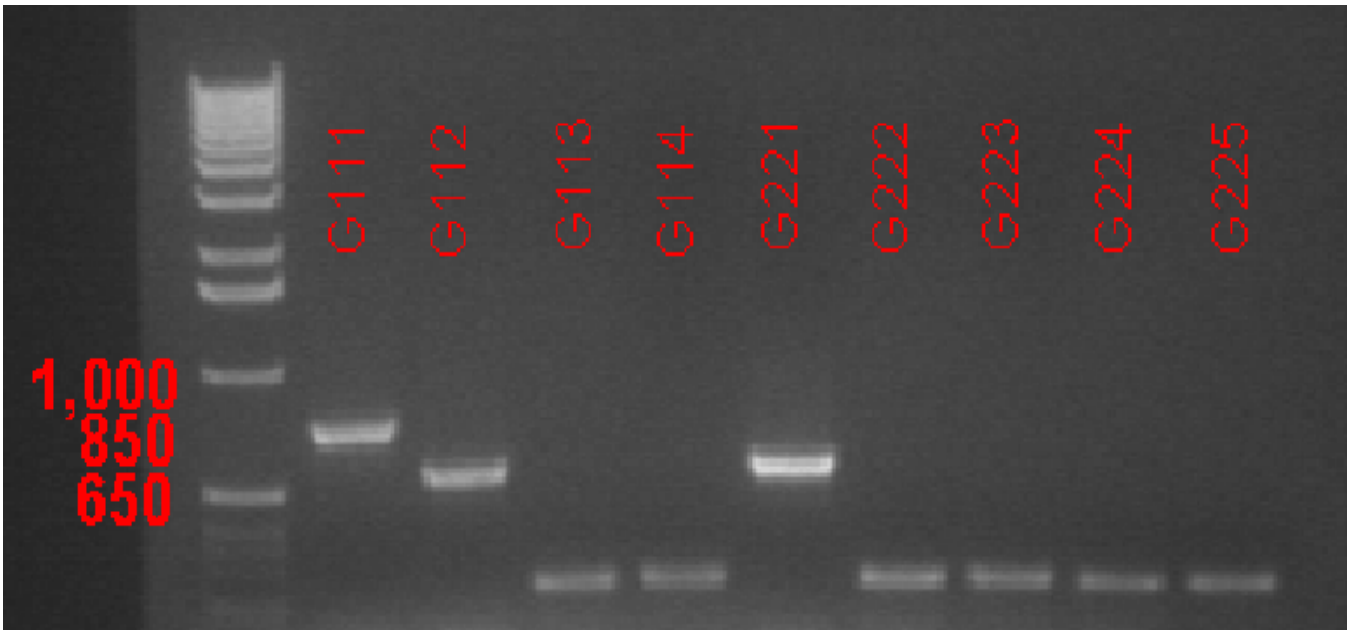



**Figure 3:** Gel showing the PCR DETECT products from *E.coli* colonies using M13 forward and M13 reverse primers. Lines labeled G111, G112 and G221 show the expected size of the insert of amplified RyR gene sequence (G=Genomic). 1kbp ladder is on the left-hand-side.

Plasmids were purified from bacterial cultures using the S.N.A.P. PureLink™ Quick Plasmid Miniprep Kit (Invitrogen Life Technology). Plasmids were sequenced at the Center for Applied Genomics at the University of Toronto. Sequences were edited and compared by BLASTx submission to the GenBank databases. Annotation of RyR sequences was inferred by sequence similarity.

**2.8 Phylogenetic Analysis of Sequenced Genes**

The sequences obtained were aligned with other published RyR isoforms using Clustalx (Thompson *et al.,* 1997*).* Amino acids at variable positions were determined to be ancestral or derived in nature based on their identity to published invertebrate and vertebrate sequences. Bootstrapping was performed on the alignment to generate pseudo-replicates for analysis using Seqboot, a program in the PHYLIP package (Felsenstein, 1992*).* One hundred bootstrap pseudo-replicates were analyzed using the DNApars program of PHYLIP. A

majority rule consensus tree was created using the Consense program from the PHYLIP package.

# 3. Results

## 3.1 Amplification and Gel Purification:

PCR products from first-strand cDNA were amplified using RyR degenerate primers. All lanes showed amplified bands of the expected size and the negative controls corresponding to each primer pairs were clean, which implies absence of any contaminating template (Figure 4). Then the amplified PCR product (band) was purified from the gel. The amplified products from designed primers were disregarded due to very low amplification.

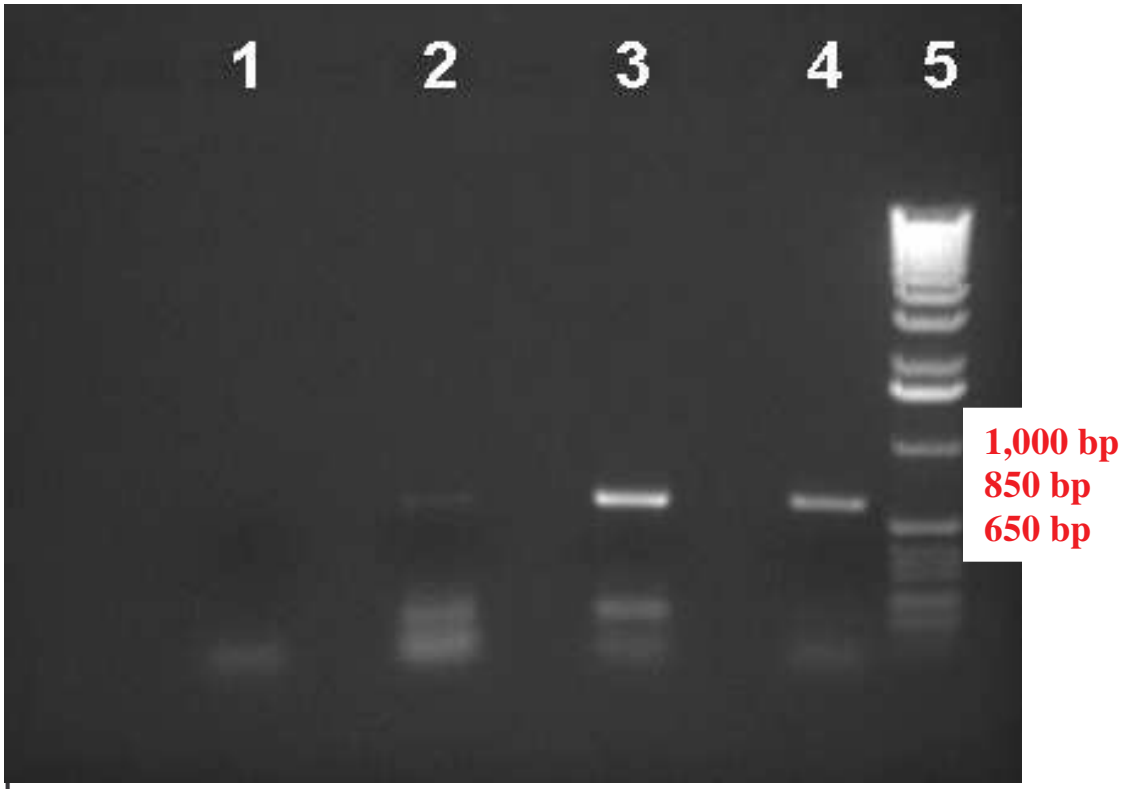



**Figure 4:** PCR Product amplified using F1/R1 (lane 1), F1/R2 (lane 2), F2/R1 (lane 3), and F2/R2 (lane 4) primer pairs. Bichir brain sample cDNA was successfully amplified in lanes 2, 3 and 4. Amplification was seen only with primer pairs F1/R1, F1/R2 and F2/R2. Negative lanes for the above primers show no contamination or presence of any PCR product. Thus the products were purified.

### 3.2 Detection of Recombinants

The PCR products were purified from the gel and cloned into pGEM-T vectors. Recombinants were detected via amplification by PCR. Figure 5 shows the results of gel electrophoresis, with expected sized PCR products present in lanes labeled G111, G112 and G221. White color of colonies showed an insertion in the vector indicating a successful recombination. This was confirmed by the PCR detection method.

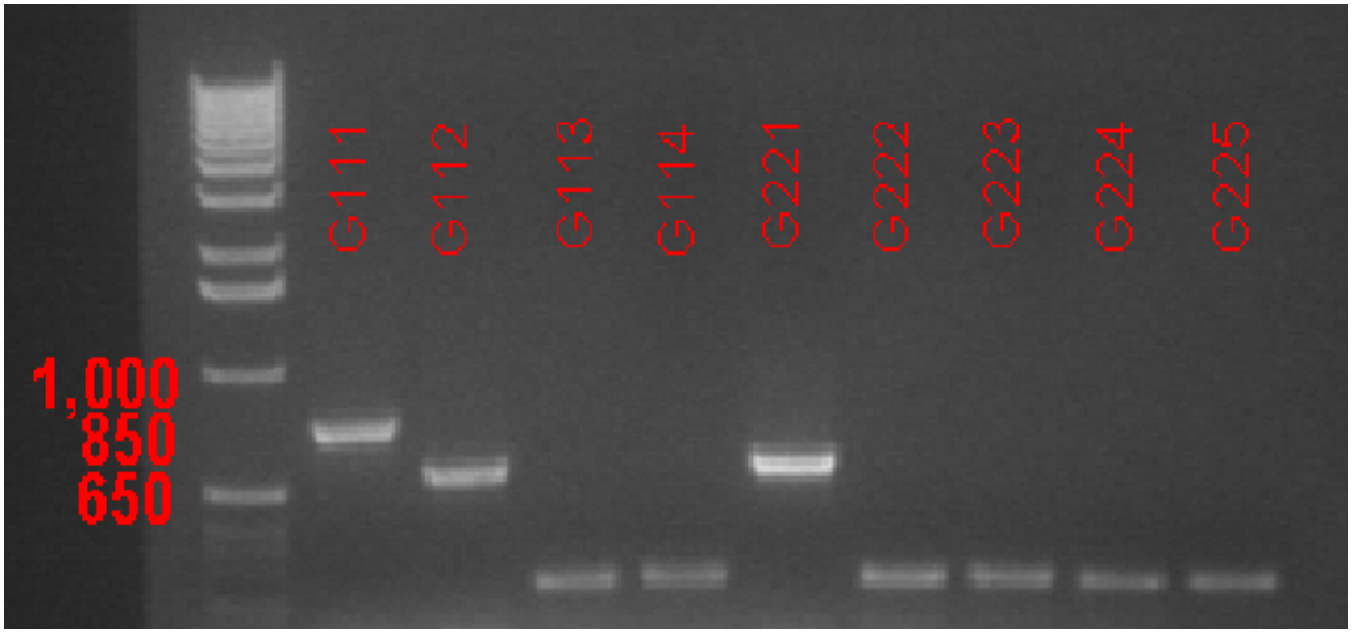



**Figure 5:** Gel showing the PCR DETECT products from *E.coli* colonies using M13 forward and M13 reverse primers. Lines labeled G111, G112 and G221 show the expected size of the insert of amplified RyR gene sequence (G=Genomic). 1kbp ladder is on the left-hand-side.

### 3.3 Sequencing of Purified Plasmids

All clones purified from *E. coli* JM109 were sequenced (Overall 29 plasmids). The aligned nucleotide sequences show such a high degree of identity from nucleotide position ~145 – 460. Brain sequences showed a very distinct pattern at the beginning of each sequence (0-25$^{th}$ nucleotide), which was highly

conserved between all the brain sequences (Figure 6). The translated amino acid sequence were also aligned (Figure 7). BLASTx searches were performed against the GenBank database. The BLASTx algorithm searches all six conceptual translations (3 frames $\times$ 2 strands) against a protein sequence database (Table1). The E-value (Expected value), which is a statistical measurement that represents the likelihood of finding that particular result by chance, is presented in Table 1. Therefore, the lower an E-value, the more significant the similarity is between the sequences. This query revealed *Monodelphis domestica* (gray short-tailed opossum) RyR3 as the top match for all sequences except for BSM 213 (bichir skeletal muscle 213), in which the top hit was *Rana catesbeiana* (North American bullfrog) RyR1. The second top hit for all sequences was *Monodelphis domestica* RyR2. Also by using BLASTx pairwise comparison, percent identity and percent positives are shown in Table 2. Finally, a multiple amino acid sequence alignment was generated with selected published RyR sequences from different species (Figure 8).

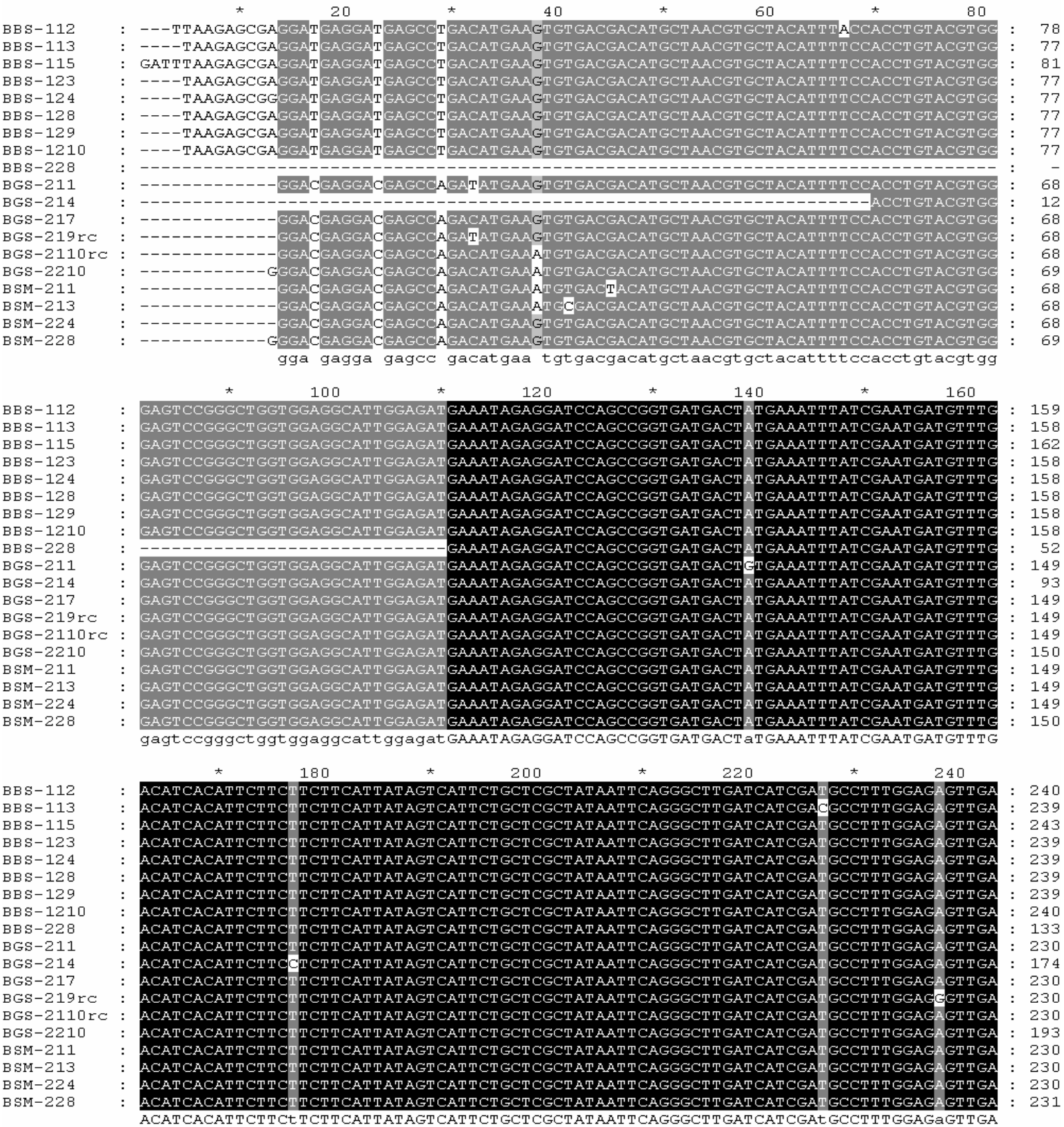



**Figure 6:** Alignment of Bichir nucleotide sequences with identical nucleotides in black. A consensus sequence of all 18 sequences is located at the bottom of the alignment. BBS=Bichir Brain Sequence, BGS=Bichir Genomic Sequence, and BSM=Bichir Skeletal Muscle.

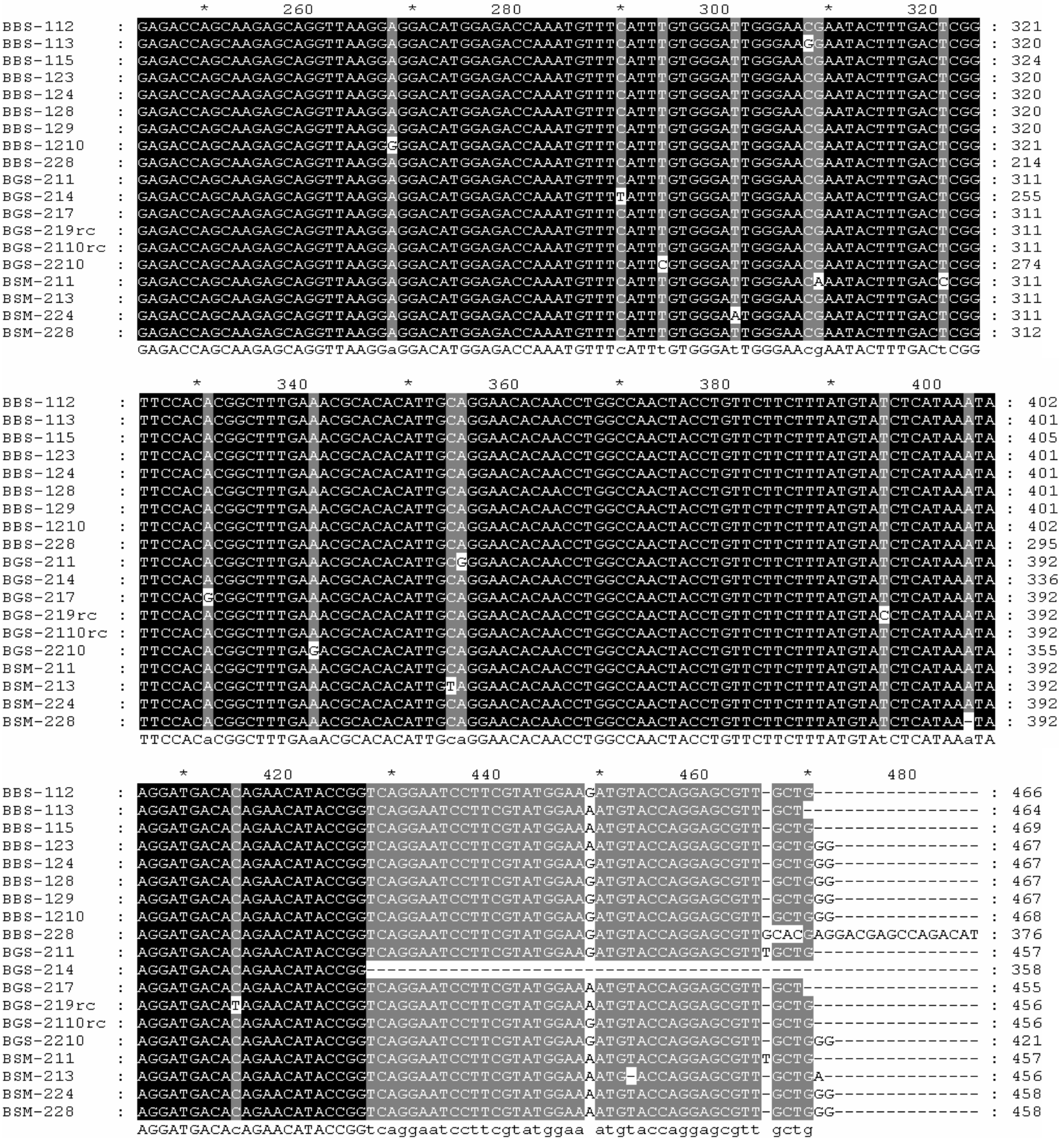



**Figure 6:** Alignment of Bichir nucleotide sequences with identical nucleotides in black. A consensus sequence of all 18 sequences is located at the bottom of the alignment. BBS=Bichir Brain Sequence, BGS=Bichir Genomic Sequence, and BSM=Bichir Skeletal Muscle (Continued).

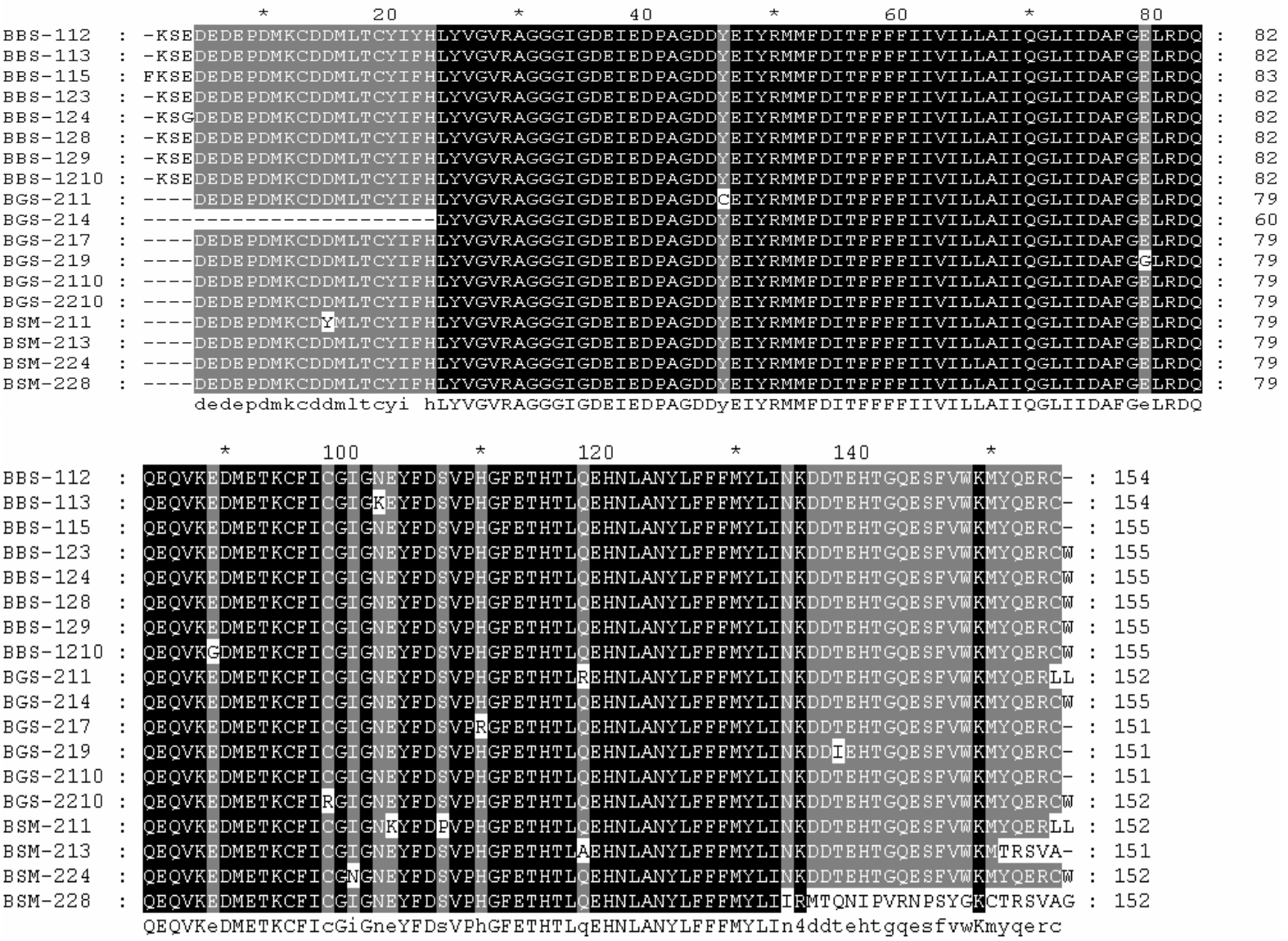



**Figure 7:** Alignment of Bichir deduced amino acid sequences with identical amino acids in black. The consensus sequence is at the bottom of the alignment.

**Table 1:** Top three matches from BLASTx search of GenBank database for the Bichir nucleotide sequences. Accession numbers of the matching hypothetical protein sequences are shown, followed by E values in bracket.

| Sequence | Match #1 (E-value) | Match #2 (E-value) | Match #3 (E-value) |
|---|---|---|---|
| BBS 112 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (5e-75) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (8e-75) | *Mus musculus*<br>RyR2<br>NP_076357 (8e-75) |
| BBS 113 | *Rana catesbeiana*<br>RyR1<br>BAA04646 (6e-75) | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (1e-74) | *Equus caballus*<br>RyR1<br>XP_001496660 (2e-74) |
| BBS 115 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (2e-75) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (4e-75) | *Mus musculus*<br>RyR2<br>NP_076357 (4e-75) |
| BBS 123 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (2e-75) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (4e-75) | *Bos taurus*<br>RyR2<br>XP_001252112 (3e-76) |
| BBS 124 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (1e-75) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (2e-75) | *Mus musculus*<br>RyR2<br>NP_076357 (2e-75) |
| BBS 128 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (4e-76) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (7e-76) | *Mus musculus*<br>RyR2<br>NP_076357 (7e-76) |
| BBS 129 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (4e-76) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (7e-76) | *Mus musculus*<br>RyR2<br>NP_076357 (7e-76) |
| BBS 1210 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (1e-75) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (2e-75) | *Bos taurus*<br>RyR2<br>XP_001252112 (2e-75) |
| BGS 211 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (2e-64) | *Rana catesbeiana*<br>RyR1<br>BAA04646 (2e-64) | *Bos taurus*<br>RyR2<br>XP_001252112 (3e-64) |
| BGS 214 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (2e-64) | *Mus musculus*<br>RyR2<br>NP_076357 (2e-64) | *Bos taurus*<br>RyR2<br>XP_001252112 (2e-64) |
| BGS 217 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (4e-66) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (6e-66) | *Mus musculus*<br>RyR2<br>NP_076357 (6e-66) |
| BGS 219 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (1e-65) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (2e-65) | *Mus musculus*<br>RyR2<br>NP_076357 (2e-65) |
| BGS 2110 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (4e-67) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (7e-67) | *Mus musculus*<br>RyR2<br>NP_076357 (7e-67) |
| BGS 2210 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (6e-67) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (1e-66) | *Mus musculus*<br>RyR2<br>NP_076357 (1e-66) |
| BSM 211 | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (2e-64) | *Pan troglodytes*<br>RyR2<br>XP_514296 (3e-64) | *Homo sapiens*<br>RyR2<br>NP_001026 (3e-64) |
| BSM 213 | *Rana catesbeiana*<br>RyR1<br>BAA04646 (1e-62) | *Monodelphis domestica*<br>RyR3<br>XP_001380803 (2e-62) | *Monodelphis domestica*<br>RyR2<br>XP_001378300 (3e-62) |

**Table 1 cont'd.:**

| Sequence | Match #1 (E-value) | Match #2 (E-value) | Match #3 (E-value) |
|---|---|---|---|
| BSM 224 | *Monodelphis domestica* RyR3 XP_001380803 (1e-67) | *Monodelphis domestica* RyR2 XP_001378300 (2e-76) | *Mus musculus* RyR2 NP_076357 (2e-67) |
| BSM 228 | *Monodelphis domestica* RyR3 XP_001380803 (4e-63) | *Monodelphis domestica* RyR2 XP_001378300 (6e-63) | *Pan troglodytes* RyR2 XP_514296 (6e-63) |

**Table 2:** Percent sequence similarity of bichir sequences with selected published sequences based on BLASTx pairwise comparison. Percentages for identities indicate the proportion of amino acids that match exactly with the published sequence. Percentages for positives indicate the number of amino acids that score a positive value based on the scoring method used, but is not necessarily identical to the query.

| Species \ Sequence | Bichir Sequence Identities | Bichir Sequence Positive |
|---|---|---|
| *Monodelphis domestica* RyR3 | 90% | 97% |
| *Monodelphis domestica* RyR2 | 90% | 97% |
| *Rana catesbeiana* RyR1 | 88% | 97% |
| *Mus musculus* RyR2 | 91% | 97% |
| *Equus caballus* RyR1 | 90% | 97% |
| *Bos taurus* RyR2 | 91% | 98% |
| *Pan troglodytes* RyR2 | 91% | 97% |
| *Homo sapiens* RyR2 | 90% | 96% |

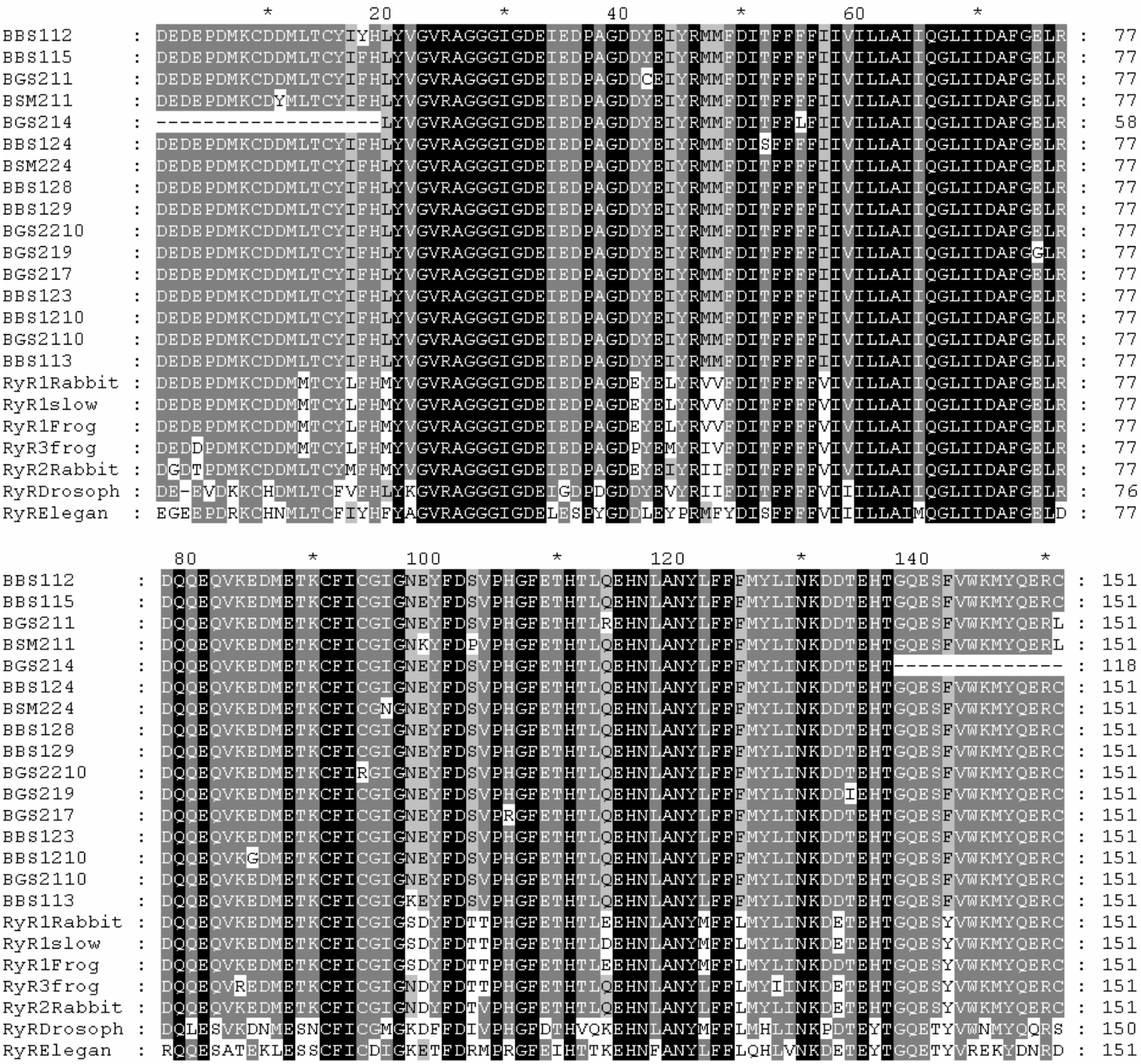



**Figure 8:** Multiple amino acid alignment of Bichir with selected published RyR amino acid sequences from *C. elegans* (BAA08309), Frog RyR3 (D21071), Frog RyR1(D21070), Rabbit RyR1 (X15209), Rabbit RyR2 (U50465). Identical amino acids are shaded black.

**Table 3:** Variant amino acid sites between Bichir and published mammalian RyR amino acid nucleotides. Rabbit RyR1 (X15209), Rabbit RyR2 (U50465) and Mouse RyR3 (NP_08320).

| Bichir amino acid position | Corresponding amino acid in Bichir | Corresponding mammalian RyR1 amino acid (Rabbit) | Corresponding mammalian RyR2 amino acid (Rabbit) | Corresponding mammalian RyR3 amino acid (Mouse) |
|---|---|---|---|---|
| **9** | F | N | N | N |
| **14** | G | E | E | E |
| **16** | E | E | G | D |
| **18** | E | E | T | E |
| **26** | Y | D | D | D |
| **55** | D | E | E | P |
| **56** | C | Y | Y | Y |
| **69** | L | F | F | F |
| **89** | G | E | E | E |
| **100** | G | E | E | E |
| **108** | R | C | C | C |
| **110** | N | I | I | I |
| **112** | K | S | N | N |
| **113** | K | D | D | D |
| **118** | P | T | T | T |
| **119** | V | T | V | T |
| **121** | R | R | H | H |
| **128** | R | E | Q | Q |
| **128** | A | E | Q | Q |

### 3.4 Phylogenetic Analysis of Sequenced Genes

Based on the bichir amino acid alignments (Fig. 8), there are 8 sequences that contain 8 unique amino acids. These are not shared with the published RyR amino acids or with other bichir amino acid sequences. Thus, these 8 sequences were selected for phylogenetic analysis and were aligned with selected published RyR sequences, in order to determine the phylogenetic position of bichir RyRs. The RyR from C*. elegans*, an invertebrate, was used as the designated outgroup, (Fig. 9). The parsimony analysis was performed using Seqboot, DNApars and Consense programs from the PHYLIP package (Felsenstein, 1992*)*.

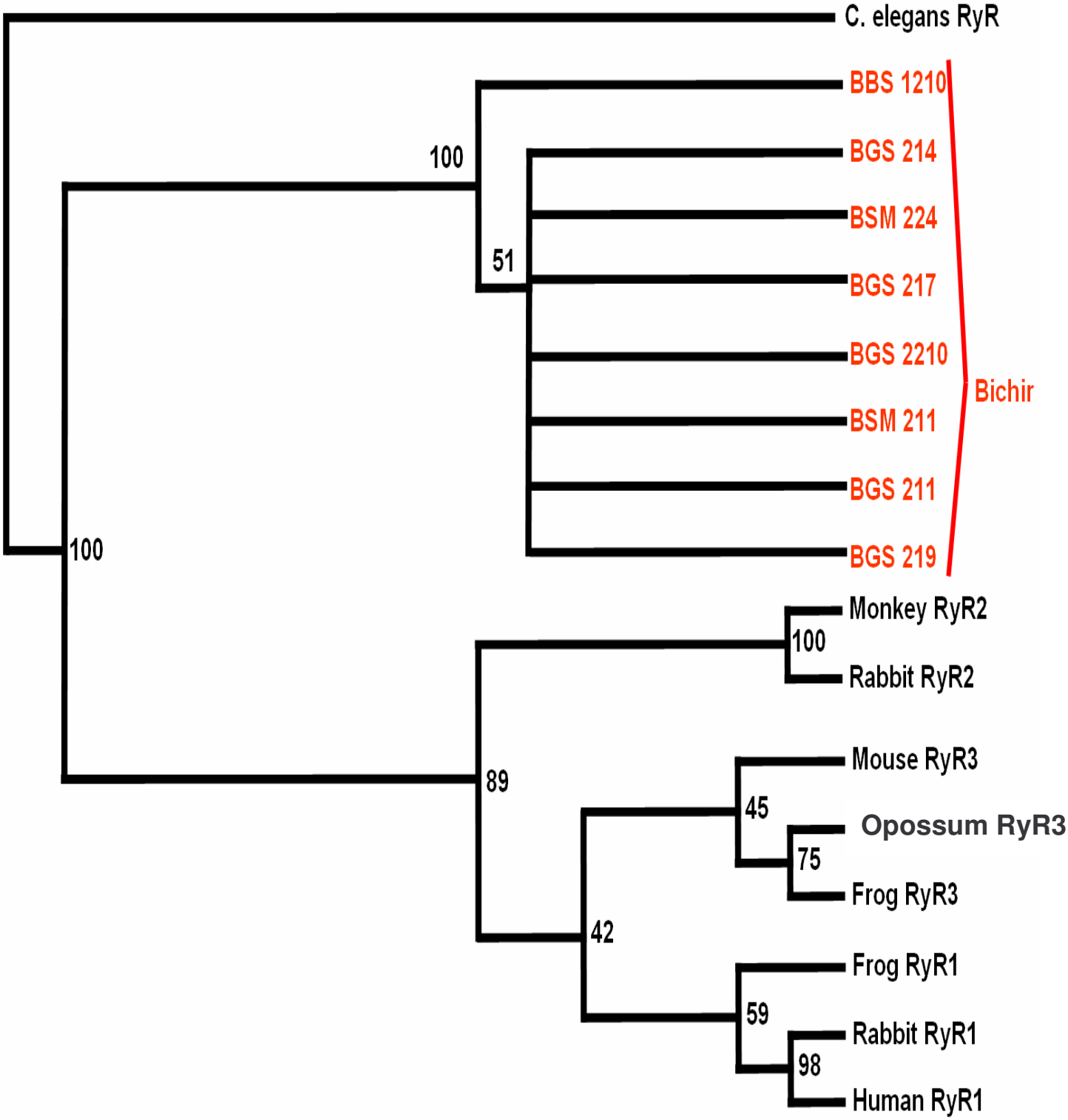



**Figure 9:** Maximum parsimony tree based on the multiple amino acid sequence alignment of cloned Bichir RyR and published RyR sequences including *C. elegans* (BAA08309), *Drosophila* (NM_057646), Fish RyR1-slow (U97329), Frog RyR3 (D21071), Frog RyR1 (D21070), Rabbit RyR1 (X15209), Rabbit RyR2 (U50465) and Mouse (NP_808320). Accession numbers are presented in the brackets.

# 4. Discussion

## 4.1 Sequencing of Bichir RyR Clones

I have identified up to eight RyR genes in the bichir genome, with sequence identity to the major vertebrate RyR isoforms. While 8 unique sequences can be identified in the nucleotide alignment (Fig. 5), the sequences can be divided into 3 main classes that putatively represent orthology to the 3 mammalian RyR isoforms.

Once the sequences were translated, the deduced bichir amino acid sequences were aligned with published representative RyR sequences. According to the amino acid alignment presented in Fig. 7, there are 8 positions in which the amino acids are unique to bichir amino acid sequences; they are not shared with the published RyR sequences or with other bichir amino acid sequences. Presence of 8 unique amino acids would imply existence of 8 unique genes encoded within the genome of bichir. The existence of 8 unique genes would suggest that bichir has undergone a minimum of 3 rounds of whole-genome duplication during its evolution. This theory would support the hypothesis that bichir's genome encodes all 3 major RyR isoforms, which would place bichir ancestral to the more derived vertebrates and teleost fish. The fact that derived genomes like mammals do not encode greater than 3 RyR genes may be explained by gene loss in these lineages.

Deduced amino acid sequences from bichir vary at 14 of 155 positions. BLASTx searches of GenBank were done to identify which isoform is most

similar to each sequenced clone. The top matches from BLASTx searches against the GenBank databases are RyR3 and RyR2 forms (Table 1).

There are several positions (14 positions of 155 representing 9% of the alignment) in which deduced amino acid sequences from bichir varies from published mammalian RyR amino acid sequences. These locations have been identified in Table 3, which suggests that because these sequences differ approximately 9% of the time. Based on the phylogenetic tree presented in Fig. 8 bichir is more primitive to teleost fish and mammals such as rabbit and mouse yet more derived than primitive organisms such as *C. elegans*. The remaining positions are conserved throughout all forms and represent functionally conserved regions (Xian-guang *et al* 2002).

Only two of the 16 bichir sequences queried against the GenBank database returned RyR1 as a top hit. The RyR1 orthologue may be present in the genome of bichir but in very low amounts, lower than 10% of total RyR present (Fill and Copello 2002).

### 4.2 Phylogenetic Analysis of Sequences

Phylogenetic analysis using parsimony analysis positions bichir sequences close to the more ancestral RyR sequences. The tree was rooted using the *C. elegans* nucleotide sequences and indicates that the bichir RyR sequences are more derived than the ancestral sequences present in those organisms.

### 4.3 Future Studies

In order to conclusively determine the presence of 3 different RyR isoforms in bichir much more work needs to be done on surveying the genomic DNA. Many more clones from genomic needs to be obtained in order to determine the number of isoforms present in the genome with much higher degree of confidence.

# 5. Conclusions

1. Up to 8 RyR genes have been identified in bichir, which suggests bichir has undergone a minimum of three rounds of whole-genome duplication during its evolution.
2. The majority of the cloned sequences show greatest similarity to RyR2 and RyR3 based on BLASTx searches.
3. The phylogenetic analysis places this sequence ancestral to the vertebrate RyR sequences.

# 6. References


**Amores A., Suzuki, T., Yan, Y. L., Pomeroy, J., Singer, A., Amemiya, C. and J. H. Postlethwait.** 2004. Developmental roles of puffer fish *hox* clusters and genome evolution in ray-fin fish. *Genome Res*. **14**(1):1 – 10.

**Bird, T.M.** 2005. Extraction of RNA from cells and tissue. *Methods Mol. Med*. **108**:139 – 48.

**Blomme, T., Vandepoele, K., De Bodt, S., Simillion, C., Maere, S., Van de Peer, Y.** 2006. The gain and loss of genes during 600 million years of vertebrate evolution. *Genome Biol*. **7**(5):R43.

**Budgett, J. S.** 2001. In Pursuit of *Polypterus*. *American Institute of Biological Sciences.* **51**(5): 399 – 407.

**Chiu, C.H., Dewar, K., Wagner, G.P., Takahashi, K., Ruddle, F., Ledje, C., Bartsch, P., Scemama, J.L., Stellwag, E., Fried, C., Prohaska, S.J., Stadler, P.F. and C. T. Amemiya**. 2004. Bichir *Hox*A cluster sequence reveals surprising trends in ray-finned fish genomic evolution. *Genome Res*. **14**(1):11-7.

**Chomczunski, P. and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**:156 – 159.

**Felsenstein, J.** 1992. Estimating effective population size from samples of sequences: A bootstrap monte carlo integration method. *Genet Res* **60**(3):209-20.

**Fill, M., and Copello, J.A.** 2002. Ryanodine receptor calcium release channels. *Physiol Rev* **82**(4):893-922.

**Finn, R.N. and Kristoffersen, B.A.** 2007. Vertebrate vitellogenin gene duplication in relation to the "3R hypothesis": Correlation to the pelagic egg and the oceanic radiation of teleosts. *PLoS ONE* **2**:e169.

**Franck, J.P., Morrissette, J., Keen, J.E., Londraville, R.L., Beamsley, M., Block, B.A.** 1998. Cloning and characterization of fiber type-specific ryanodine receptor isoforms in skeletal muscles of fish. *Am J Physiol* **275**:C401-15.

**Inoue, I., Tsutsui, I. and Q. Bone.** 2002. Excitation-contraction coupling in skeletal and caudal heart muscle of the hagfish *eptatretus burgeri girard*. *J Exp Biol* **205**:3535 - 41.

**Jaffe, L. F.** 1999. Organization of early developmental by calcium patterns. *Bioessays* **21**: 657-667.

**Jones, D. D.** 2005. Triplet nucleotide removal at random positions in a target gene: the tolerance of TEM-1 β-lactamase to an amino acid deletion. *Nucleic Acid Research*. **33**: 1 – 8.

**Ledje, C., Kim, C. B., and Francis H. Ruddle.** 2002. Characterization of *Hox* Genes in the Bichir, *Polypterus palmas*. *Journal of experimental Zoology (Mol Dev Evol)* **294**:107-111.

**Meyer, A. and Van de Peer, Y.** 2005. From 2R to 3R: Evidence for fish-specific genome duplication (FSGD). *Bioessays* **27**(9):937-45.

**Morrissette, J., Xu, L., Nelson, A., Meissner, G. and B. A. Block.** 2000. Characterization of RyR-slow, a ryanodine receptor specific to slow-twitch skeletal muscle. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **279**:R1889-98.

**Rossi, D. and V. Sorrentino.** 2002. Molecular genetics of ryanodine receptors $Ca^{2+}$ release channels. *Cell Calcium*. **32**:307 – 319.

**Rose, T. M., Schultz, E. R., Henikoff, J. G , Pietrokovski, S., McCallum, C. M., Henikoff, S.** 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res* **26**(7):1628-35.

**Sharma, M. R. and T. Wagenknecht.** 2004. Cryo-Electron Microscopy and 3D Reconstruction of Ryanodine Receptors and Their Interactions with E-C coupling Proteins. *Basic Appl Myol*. **14**(5): 299-306.

**Sherwood, L., Klandorf, H., and P.H. Yancey.** 2005. Animal Physiology: From Genes to Organisms. *Thompson Brooks/Cole*.

**Sorrentino, V., Barone, V., Rossi, D.** 2000. Intracellular $Ca^{2+}$ release channels in evolution. *Curr Opin Genet Dev* **10**(6):662-7.

**Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and D. G. Higgins.** 1997. The Clustalx windows interface:flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleis Acids Res.* **25**: 4876 – 4882.

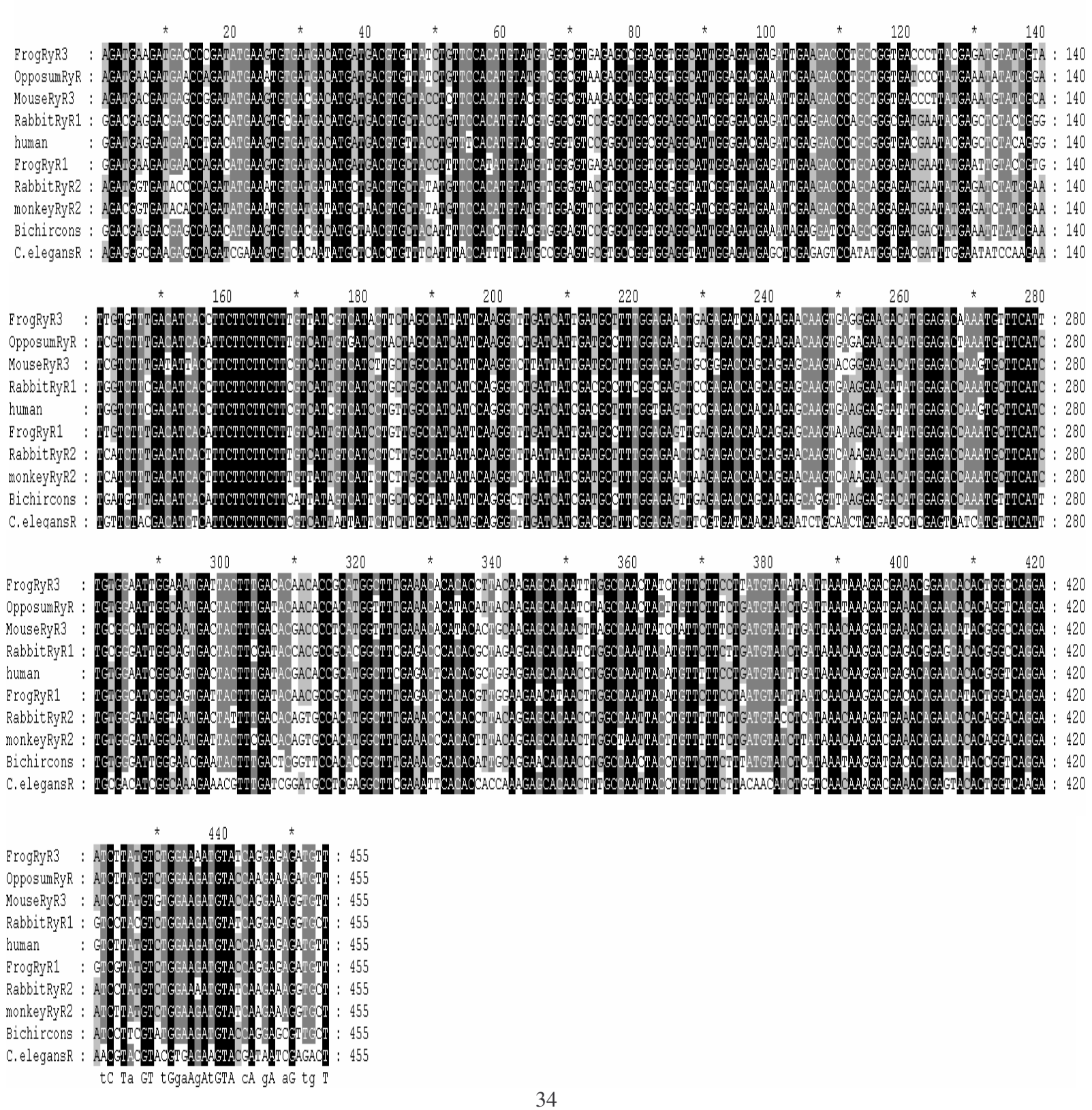
**Tripathy, A., Xu, L., Mann, G. and G. Meissner.** 1995. Calmodulin activation and inhibition of skeletal muscle $Ca^{2+}$ release channels (ryanodine receptor). *Biophys. J.***69**:106 – 119.

**Webb, S. E. and A. L. Miller.** 2003. Calcium signalling during embryonic development. *Nat. rev. Mol. Cell Biol.* **4**:539 – 551.

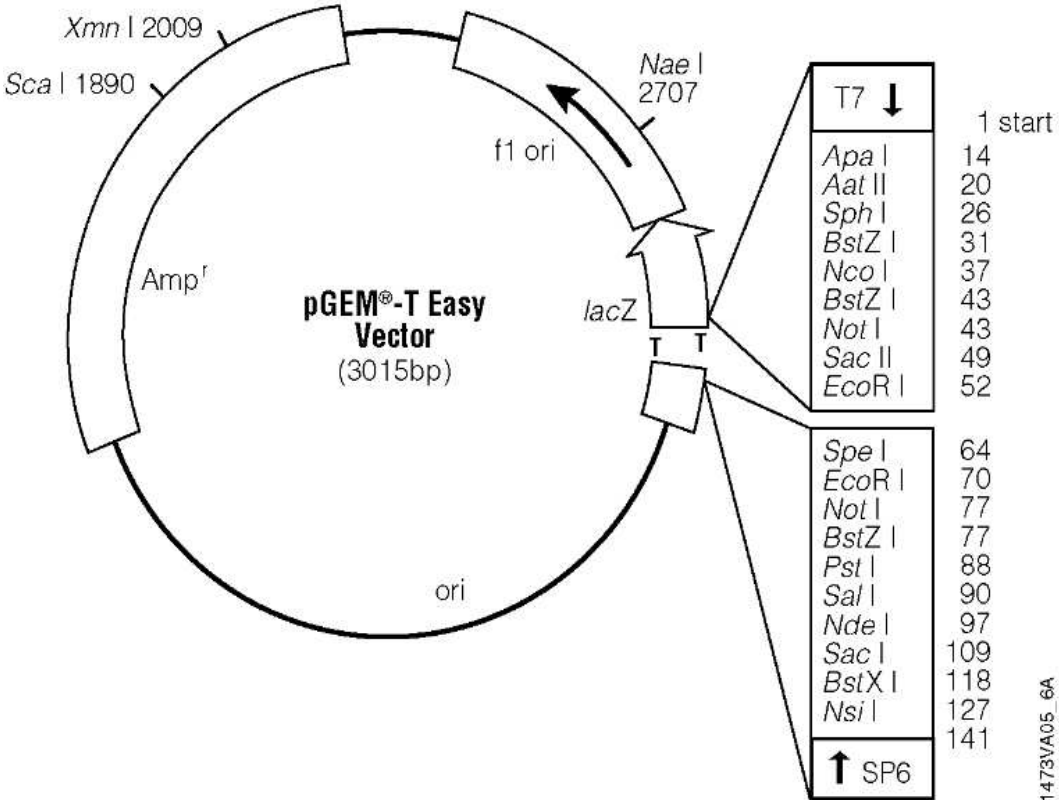
**Xian-guang, H., Aldridge, R.J., Siveter, D.J., Siveter, D.J., Xiang-hong, F.** 2002. New evidence on the anatomy and phylogeny of the earliest vertebrates. *Proc Biol Sci* **269**(1503):1865 – 9.

# 7. Appendices

**7.1** Nucleotide sequence alignment of cloned Bichir RyR consensus sequence and the published RyR sequences of *C. elegans* (D45899), Human RyR1 (P21817), Frog RyR3 (D21071), Frog RyR1 (D21070), Rabbit RyR1 (X15209), Rabbit RyR2 (U50465), Monkey RyR2 (XR_013555) and Mouse RyR3 (NP_808320).



## 7.2 pGEM®-T Easy Vector Maps




**pGEM®-T Easy Vector**

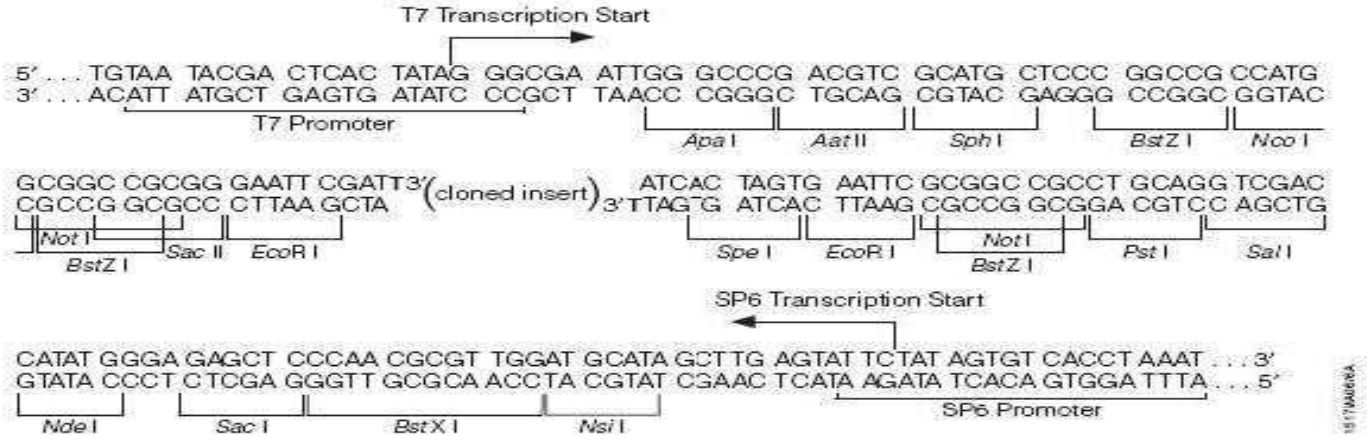



**Figure 1. The promoter and multiple cloning sequence of the pGEM®-T and pGEM®-T Easy Vectors.** The top strand of the sequence shown corresponds to the RNA synthesized by T7 RNA Polymerase. The bottom strand corresponds to the RNA synthesized by SP6 RNA Polymerase.

## 7.3 Amino acid alignment used for hagfish degenerate primer design

```
FrogRyR1          ILSSVTHNGKQLMMTVGLLAVVVYLYTVVAFNFFRKFYNKSEDEDEPDMKCDDMMTCYLF
FishRyR1slow      ILSSVTHNGKQLMMTVGLLAVVVYLYTVVAFNFFRKFYNKSEDEDEPDMKCDDMMTCYLF
RabRyR1           ILSSVTHNGKQLVMTVGLLAVVVYLYTVVAFNFFRKFYNKSEDEDEPDMKCDDMMTCYLF
FrogRyR3          ILSSVTHNGKQLVLTVGLLAVVVYLYTVVAFNFFRKFYNKSEDEDDPDMKCDDMMTCYLF
RabRyR2           ILSSVTHNGKQLVLTVGLLAVVVYLYTVVAFNFFRKFYNKSEDGDTPDMKCDDMLTCYMF
                  ************:::****************************** * ********:***:*

FrogRyR1          HMYVGVRAGGGIGDEIEDPAGDEYELYRVVFDITFFFFVIVILLAIIQGLIIDAFGELRD
FishRyR1slow      HMYVGVRAGGGIGDEIEDPAGDEYELYRVVFDITFFFFVIVILLAIIQGLIIDAFGELRD
RabRyR1           HMYVGVRAGGGIGDEIEDPAGDEYELYRVVFDITFFFFVIVILLAIIQGLIIDAFGELRD
FrogRyR3          HMYVGVRAGGGIGDEIEDPAGDPYEMYRIVFDITFFFFVIVILLAIIQGLIIDAFGELRD
RabRyR2           HMYVGVRAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVIVILLAIIQGLIIDAFGELRD
                  ********************** **:**::****************************

FrogRyR1          QQEQVKEDMETKCFICGIGSDYFDTTPHGFETHTLEEHNLANYMFFLMYLINKDDTEHTG
FishRyR1slow      QQEQVKEDMETKCFICGIGSDYFDTTPHGFETHTLDEHNLANYMFFLMYLINKDETEHTG
RabRyR1           QQEQVKEDMETKCFICGIGSDYFDTTPHGFETHTLEEHNLANYMFFLMYLINKDETEHTG
FrogRyR3          QQEQVREDMETKCFICGIGNDYFDTTPHGFETHTLQEHNLANYLFFLMYIINKDETEHTG
RabRyR2           QQEQVKEDMETKCFICGIGNDYFDTVPHGFETHTLQEHNLANYLFFLMYLINKDETEHTG
                  *****:*************.*****.*********:*******:*****:****:*****

FrogRyR1          QESYVWKMYQERCWDFFPAGDCFRKTYEDQLG
FishRyR1slow      QESYVWKMYQERCWDFFPAGDCFRKQYEDQLS
RabRyR1           QESYVWKMYQERCWDFFPAGDCFRKQYEDQLS
FrogRyR3          QESYVWKMYQERCWDFFPAGDCFRKQYEDQLG
RabRyR2           QESYVWKMYQERCWEFFPAGDCFRKQYEDQLN
                  **************:********** *****.
```