Albanian j. agric. sci. 2013;12 (4): 575-583 *Agricultural University of Tirana*

RESEARCH ARTICLE **(Open Access)**

Phylogenetic analyses with proximal promoter from skeletal and cardiac calsequestrins

RIGERS BAKIU[1*], GIORGIA VALLE[2] AND ALESSANDRA NORI[2]

[1]Department of Aquaculture and Fishing, Agricultural University of Tirana, Albania

2Department of Biomedical Science, University of Padova, Italy

Abstract:

Calsequestrin (CASQ) is the main calcium binding protein of the sarcoplasmic reticulum. In mammalian muscles, it exists as a skeletal isoform (CASQ1) found in fast- and slow-twitch skeletal muscles and a cardiac isoform (CASQ2) expressed in the heart and slow-twitch muscles. Evolutionary biology studies have a great impact to problems in medicine because evolutionary thinking do not displace other approaches to medical science, such as molecular medicine and cell and developmental biology, but that evolutionary insights can combine with and complement established approaches. Thus, we followed evolutionary biology approach by studing *Casq* promoters evolution. Phylogenetic analyses were performed using the minimum promoter of CASQ1 and CASQ2. Emerged topological discordances from the comparison between CASQ promoter based phylogenetic tree and amino acid based phylogenetic tree showed similar, but not identical evolutionary pathways.

Keywords: Sarcoplasmic reticulum; Phylogeny; Molecular clock.

1. Introduction

Calsequestrin is the most abundant calcium (Ca^{2+}) binding protein in the sarcoplasmic reticulum (SR) of cardiac and skeletal muscles with moderate affinity (Kd ~ 1 mM) and high capacity (60–80 mol Ca^{2+}/mol calsequestrin) for Ca^{2+} [1]. Localization in a sub-compartment of the SR (terminal cisternae, TC), is accomplished by anchoring to the membrane through interaction with trans-membrane proteins such as triadin (Trd) and junctin (Jct). CASQ polymerizes in response to rising Ca^{2+} concentrations in the lumen of SR, to form dimers and polymers. Post-traslational modifications such as phosphorylation and glycosylation are important for CASQ localization and function [1]. Beyond its capability of Ca^{2+} buffering, CASQ has additional roles in muscle contraction such as regulator of Ca^{2+} release through the channel Ryanodine Receptor (RyR) either directly or through interactions via triadin/junctin [2]. Finally CASQ is coordinately expressed with mayor SR and contractile protein during development and is important for the structural organization of the SR in adult [3]. Two CASQ isoforms have been described in mammals, each one encoded by a single gene. The "skeletal" isoform (encoded by *Casq1* gene) is found in the fast-twitch skeletal muscle (like soleus), whereas the "cardiac" isoform (encoded by *Casq2* gene) is expressed in the heart and in the slow-twitch skeletal muscle (like EDL). Cardiac muscle expresses exclusively the *Casq2* gene. Both protein isoforms show considerable sequence and structure similarity. For example share a high nucleotide and amino acid homology, 84 and 80%, respectively in humans [4]. At this time there are few studies that point to differences in the physiological role of the different CASQ isoforms [5] thus their roles is sometimes considered equivalent.

CASQ1 interacts with many SR proteins and is proposed that these interactions and the molecular ratio between CASQ1 and Trd/Jct play a role in SR biogenesis. A quantitative link between CASQ1 and Trd is described in different mouse knock out (KO) models and also in physio-pathological conditions: CASQ1-null mice [3] show a decrease in Trd content; pan Trd KO [6] show decrease in CASQ1 expression, but if the equilibrium between CASQ1 and Trd is due to specific gene regulation or is an effect of post-translational mechanisms is not known. Another protein whose content is strictly coordinated with CASQ1 is SERCA1 [7]. Also opposite regulation of CASQ1 and CASQ2 is reported in different models and processes: is well known that CASQ1 and CASQ2 are co-expressed during early myotube formation in vitro, but as myogenesis progresses, CASQ2 is replaced by CASQ1. In mouse and rat embryos, CASQ2 is detected in fetal heart and skeletal muscles, whereas CASQ1 transcripts are only

Correspondence: Rigers Bakiu, Department of Aquaculture and Fishing, Agricultural University of Tirana, Albania; Email: rigers.bakiu@ubt.edu.al
(Accepted for publication 17 September 2013)
***ISSN**: 2218-2020, © Agricultural University of Tirana*

found in fetal skeletal muscles [8]. Similarly, the expression of rabbit CASQ2 in the fast-twitch skeletal muscle declines progressively with development, whereas the expression of CASQ1 increases and totally replaces CASQ2 in adults. In contrast, the cardiac muscle expresses exclusively CASQ2 at all stages [9]. Recent work shows that KO of Six1 and Six4 homeoproteins alterates expression of CASQ1/2 and SERCA1 and leads to difficult in differentiation of fast skeletal muscles [10]. In models of chronic low frequence stimulation, CASQ1 decreases and CASQ2 increases in fast muscle [11]. Also denervation causes a considerable reduction in Ca^{2+}-ATPase (SERCA) and calsequestrin in EDL, making it resemble the control of slow muscle like soleus. A recent work describes different effects on CASQ1 and SERCA1 expression due to different trainings [12].

It appears that both CASQ1 and CASQ2 are up- or down-regulated in many different models, pathologies and during development implying that the two proteins are not interchangeable or are submitted to quite different regulators, but the mechanisms that operate at genic and protein level are still obscure. Evolutionary biology studies have a great impact to problems in medicine because evolutionary thinking do not displace other approaches to medical science, such as molecular medicine and cell and developmental biology, but that evolutionary insights can combine with and complement established approaches to reduce suffering and save lives [13]. Thus, we followed evolutionary biology approach by studing *Casq* promoters evolution, in order to resolve such biomedical problems.

2. Material and Methods

2.1 Phylogenetic Analyses

CASQ amino acid and coding cDNA sequences were found in GenBank (www.ncbi.nlm.nih.gov/genbank/) and ENSEMBL (http://www.ensembl.org) databases, while all *Casq* promoter region sequences were retrieved from ElDorado annotation database of Genomatix Software Suite [14]. All respective sequences were aligned using T-Coffee multiple sequence alignment software package [15]. To find the best-fit model of molecular evolution we used ProtTest3 [16] on CASQ amino acid sequence alignment and jModelTest 0.1.1 [17] on *Casq* coding and promoter region sequence alignments. Phylogenetic analyses were conducted using the maximum likelihood (ML) and Bayesian inference (BI) methods. The ML (including bootstraps) of phylogeny was conducted on promoter region sequence alignment using PhyML [18]. Non parametric bootstrap analysis was performed using PhyML software (1000 replicates). BI of phylogeny was performed on all sequence alignments using BEAST v1.7.0 software package [19] with a tree prior of Yule speciation process under a strict clock model. Four independent runs, each one with four simultaneous Markov Chain Monte Carlo (MCMC) chains, were performed for 10,000,000 generations sampled every 1000 generations. TreeAnnotator v1.7.0 software was used to identify the posterior probabilities of the nodes in the target tree (maximum clade credibility tree) generated by BEAST v1.7.0. after burning 500 trees. FigTree v1.3.1 software was used to display the annotated phylogenetic trees. We used BEAST v1.7.0. also for clock rate estimation. Trace files generated by Bayesian MCMC runs in BEAST v1.7.0. were analyzed using Tracer 1.5.

2.2 Statistical Analyses

In three parameter data statistical comparison, one-way ANOVA was performed (Tukey, Bonferroni and Sceffè tests) using Origin 8.6. **Student's t test** was used for comparisons between two parameter data like substitution rates of mammals and all analyzed species promoter sequences. Statistical significance was set at $P < 0.05$.

3. Results and Discussion

3.1. Molecular Clock Tests

The molecular clock has become an indispensable tool within evolutionary biology, enabling independent timescales to be placed on evolutionary events. Despite these valuable contributions, date estimates derived from molecular data have not been without controversy. In particular, when molecular clocks have been employed to estimate the timing of recent events already tentatively dated on the basis of paleontological, archaeological or biogeographic sources, conflicting dates are frequently obtained. In its most extreme form, the molecular clock hypothesis postulates that homologous stretches of DNA evolve at essentially the same rate along all evolutionary lineages for as long as they maintain their original function [19]. It was shown that the substitution rate of mitochondrial encoded proteins has increased in the order of fishes, amphibians, birds, and mammals and that the rate in mammals is at least six times, probably an order of

magnitude, higher than that in fishes [19]. The higher evolutionary rate in birds and mammals than in amphibians and fishes was attributed to relaxation of selective constraints operating on proteins in warm-blooded vertebrates and to high mutation rate of bird and mammalian mitochondrial DNAs. Since the assumption of rate constancy is violated even within Mammalians, a truly universal molecular clock that applies to all organisms cannot be assumed to exist. In phylogenetics, the unrooted model of phylogeny and the strict molecular clock model are two extremes of a continuum. Despite their dominance in phylogenetic inference, it is evident that both are biologically unrealistic and that the real evolutionary process lies between these two extremes. Local molecular clocks are another alternative to the global molecular clock. A local molecular clock permits different regions in the tree to have different rates, but within each region the rate must be the same. This new method conveniently allows a comparison of the strict molecular clock against a large array of alternative local molecular clock models [19]. A Likelihood Ratio Test (LRT) was performed in order to know which was the best-fit model to analyze CASQ protein sequence evolution. The LRT (**LR=2[lnL(HA) - ln L(HO)]**) was conducted with **n-2** degrees of freedom, where **n** is the number of considered taxa in the phylogeny. Likelihood scores were estimated using in BEAST v1.7.0 on JTT + G + F matrix which was determined previously by ProtTest3 software application [16] as the best model (-lnL= -5306.53) with a gamma shape value (four rate categories) of 1.057. We chose to use the constant rate birth–death process as it is probably the most popular homogeneous model. A birth–death process is a stochastic process which starts with an initial species. A species gives birth to a new species after exponential (rate λ) waiting times and dies after an exponential (rate μ) waiting time. A special case of the birth–death process is the Yule model where μ=0. This birth-death model is implemented in BEAST v1.7.0. The estimated likelihood scores for both the null ($\mathbf{H_O}$: random local molecular clock) and the alternative hypotheses ($\mathbf{H_A}$: strict molecular clock) were -5381.14 and -5427.36, respectively. Following a chi-squared distribution with 18 degrees of freedom, the alternative hypotheses was accepted for $P < 0.05$; analyzed CASQ molecular evolution is based on the strict molecular clock model. It is similar to subunits 4 and 5 of the enzyme NADH-dehydrogenase and three tRNAs evolution, which appear as molecular clock examples in several phylogenetic software releases [20] and find good support for a clock among the anthropoids, but no support is found for random local clock. We used the strict molecular clock model in the phylogenetic tree constructions and substitution rate estimations.

3.2 Phylogenetic Tree Construction

All species promoter sequences present in ElDorado database were used in our bioinformatic analyses. ElDorado is the Genomatix genome annotation. It is based on the publicly available reference genome assemblies of 31 different organisms, from which only 13 organisms *Casq* promoter region sequences are well known (Table 1).

Table 1. All annotated *Casq* promoter sequences found in ElDorado database.

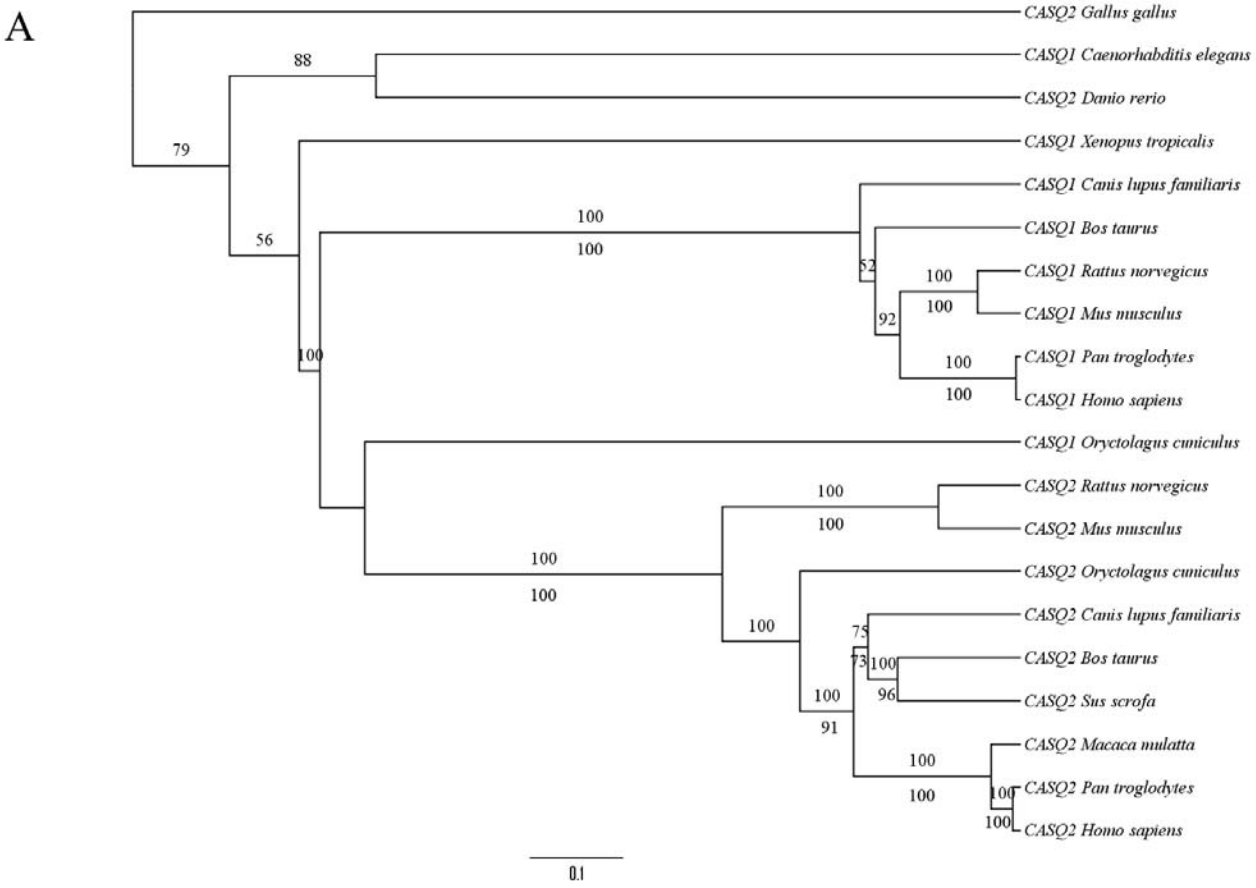
Species	Accession no.	Promoter
Bos taurus (**cow**)	GXP_988745	*Casq1*
	GXP_1320581	*Casq2*
Caenorhabditis elegans (**worm**)	GXP_476696	*Casq1*
Canis lupus familiaris (**dog**)	GXP_558411	*Casq1*
	GXP_3238664	*Casq2*
Danio rerio (**zebrafish**)	GXP_620319	*Casq2*
Gallus gallus (**chicken**)	GXP_1705730	*Casq2*
Homo sapiens (**human**)	GXP_69193	*Casq1*
	GXP_182357	*Casq2*
Macaca mulatta (**monkey**)	GXP_1382073	*Casq2*
Mus musculus (**mouse**)	GXP_285897	*Casq1*
	GXP_3083575	*Casq2*
Oryctolagus cuniculus (**rabbit**)	GXP_3105846	*Casq1*
	GXP_3106758	*Casq2*
Pan troglodytes (**chimp**)	GXP_1203321	*Casq1*
	GXP_1434139	*Casq2*
Rattus norvegicus (**rat**)	GXP_1151130	*Casq1*
	GXP_1412508	*Casq2*
Sus scrofa (**pig**)	GXP_2006898	*Casq2*
Xenopus tropicalis (**frog**)	GXP_3381510	*Casq1*

These promoter sequences were aligned using T-Coffee in combined libraries of local and multiple alignments, which are known to induce high accuracy and performance in sequence alignments [15]. The residue consistency mean score of the all sequence alignment reported by T-Coffee aligner was very low (SCORE=29) demonstrating that *Casq* promoter sequences alignment is a low quality alignment. One possible reason for the low quality of the alignment would be the fact that the transcription in CASQ1 occurs in the opposite direction relative to CASQ2. For this reason *Casq2* promoter were reverse-complemented in order to match the direction *Casq1* transcription, and then were aligned together using T-

coffee. Surprisely the score was identical to the previous alignment score. These results can be linked to the high number of observed nucleotide differences between the two classes of promoters in the comparison between them (*Casq1 vs Casq2*). However, each promoter class alignment (*Casq1* and *Casq2* promoter alignments) showed higher scores than previous alignments (*Casq1*: SCORE=46 and *Casq2*: SCORE=53). jModelTest 0.1.1 software [17] determined the HKY+G model as being the best-fit model of promoter DNA sequence evolution with a gamma shape value (four rate categories) of 4.02, using two statistical criterion, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) (-lnL=12098.567). Phylogenetic relationships of all these different organisms *Casq* promoter sequences were determined using the two most powerful statistical methods: ML and BI. The best phylogeny generated by the BI method is depicted in Figure 1A.

First we used a BEAUti v1.7.0 generated input file with no out-group sequences; then we used as out-group all *non*-Mammalian sequences. The "out-group" maximum clade credibility tree is the best of two trees generated by BEAST v1.7.0, because it is characterized of statistically significant lower values of likelihood (lnL) (-12165.057±0.0996 for out-group tree *vs* -12163.746±0.1131 for non out-group tree; $P < 0.05$) and auto-correlation time (ACT) (4805.239 for out-group tree *vs* 5243.1218 for non out group tree) than the other tree. The BI phylogenetic reconstruction was highly representative of the phylogeny generated by ML. As shown in Figure 1A, the major branches were highly supported both in BI and ML trees, even BI phylogenetic tree was more resolved than ML one. Another difference between BI and ML, was the different position of rabbit (*Oryctolagus cuniculus*) *Casq2* promoter sequence in the two trees; this branch was higher supported in BI phylogenetic tree (100% posterior probability) than in ML one (71% bootstrap value) (not shown).

.



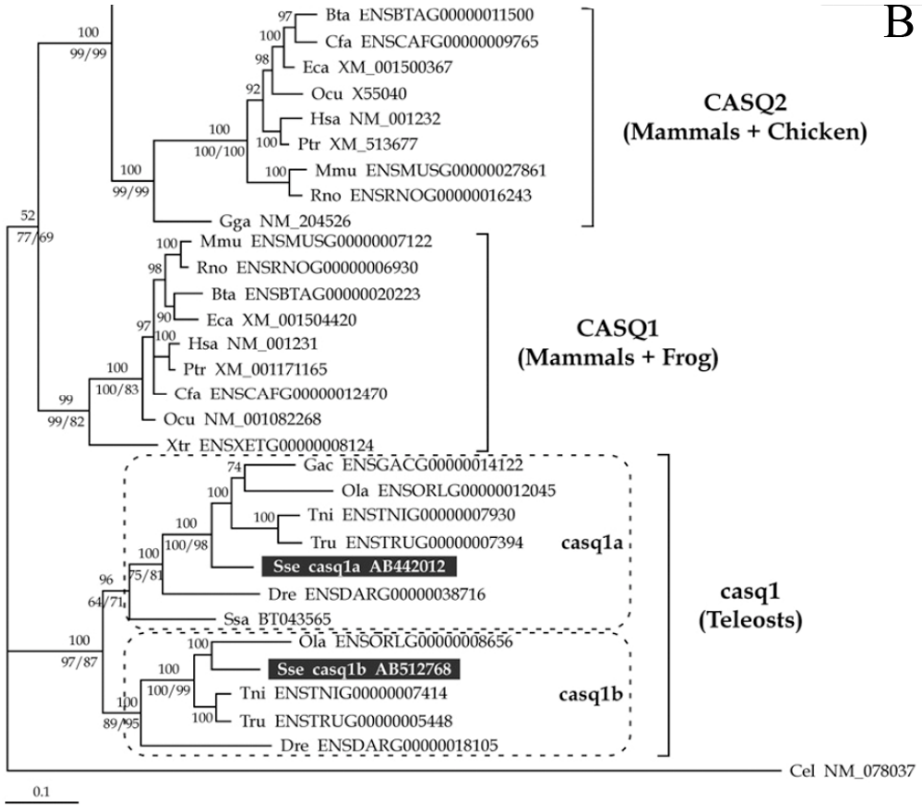


Figure 1. (A) Phylogenetic relationships among mammals and other organisms *Casq* promoter sequences using BI and ML methods. Posterior probability values higher than 50% are indicated on each branch above nodes; below nodes indicate the considered bootstrap support (>50%) of the clades for ML method. The scale for branch length (0.1 substitution/site) is shown below the tree. (B) Phylogenetic relationships among the predicted sequence of *Senegalese sole* calsequestrin isoforms and the corresponding deduced amino acid sequences from other vertebrates [21] using the BI method. GenBank/EMBL/DDBJ or ENSEMBL accession no. is indicated in each case. Bootstrap values higher than 50% are indicated on each branch above nodes; below nodes indicate the bootstrap support of the main clades for ML/NJ methods. The scale for branch length (0.1 substitutions/site) is shown below the tree. Bta: *Bos taurus*; Cfa: *Canis familiaris*; Cel: *Caenorhabditis elegans*; Eca: *Equus caballus*; Gga: *Gallus gallus*; Hsa: *Homo sapiens*; Mmu: *Mus musculus*; Ocu: *Oryctolagus cuniculus*; Ptr: *Pan troglodytes*; Rno: *Rattus norvegicus*; Xtr: *Xenopus tropicalis*.

Both *Casq1* and *Casq2* promoter sequences of mammals grouped into two separate and highly consistent clades (100% posterior probabilities in both cases), even rabbit *Casq1* promoter sequence was an exception, because it's position was in the mammals "big" clade, but it was outside from each of two clades. In order to know if *Casq* promoter and CASQ proteins had followed the same pathway of evolution, a tree topology comparison between CASQ promoter and amino acid sequences [21] phylogenetic trees was conducted. Infante and colleagues [21] performed the phylogenetic relationships among the predicted sequence of *Senegalese sole* CASQ isoforms and the corresponding deduced amino acid sequences from other vertebrates using the BI method (Figure 1B). In that phylogenetic tree, mammals and frog (*Xenopus tropicalis*) CASQ1 were grouped together into a clade, while mammals and chicken (*Gallus gallus*) CASQ2 were grouped into the other one. As it is shown in Figure 1A, in *Casq* promoter based phylogeny, the chicken *Casq2* was positioned far outside the mammals *Casq2* clade; only frog *Casq1* position was identical to that observed in the CASQ amino acid based phylogeny [21]. Probably, this emerged topological discordance from the comparison between CASQ promoter based phylogenetic tree and amino acid based [21] phylogenetic tree show similar, but not identical evolutionary pathways and may be linked to possible different substitution rates present in the comparison between *Casq* promoter and protein sequences. As shown in Figure 1A, the *Casq* promoter sequences positions in each of the two mammals calsequestrin isoform clades were identical to those observed in the amino acid based phylogenetic tree (Figure 1B). To confirm if the pathways of evolution

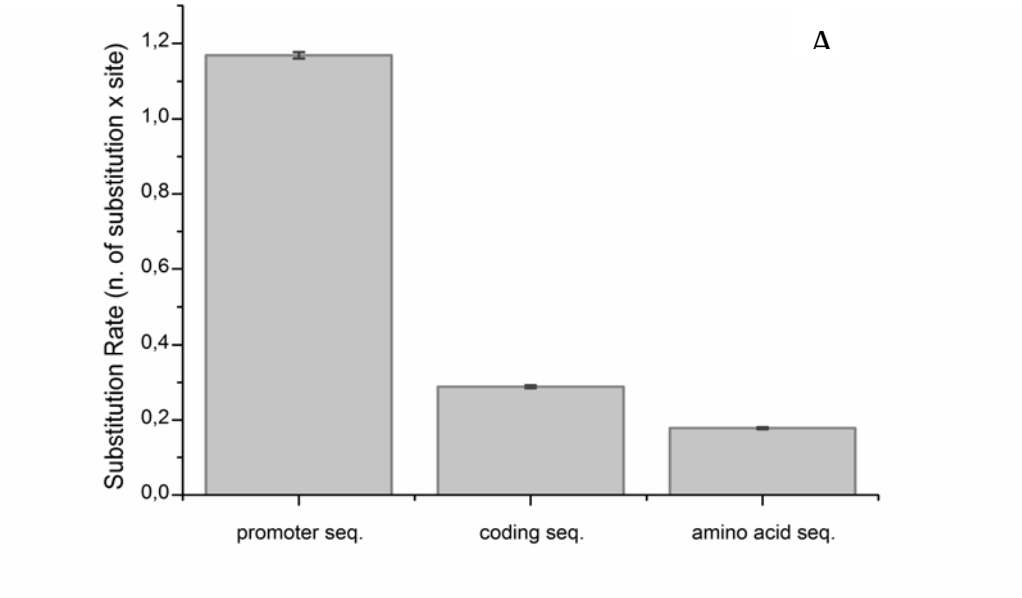
were identical in mammals, we performed a comparison between estimated *Casq* promoter and CASQ protein, substitution rates.

3.3 Substitution Rate Comparison

As it is shown in Figure 1A, the *Casq* promoter positions in each of the two mammals CASQ isoform clade are identical to those observed in the amino acid based phylogenetic tree (Figure 1B). Thus, first we estimated nucleotide substitution rates of mammals *Casq* promoter sequences using BEAST v1.7.0; then we estimated nucleotide and amino acid substitution rates of mammals *Casq* coding cDNA and protein sequences, respectively. Previously, the jModelTest 0.1.1 software determined HKY+G and TrN+G as being the best-fit models (four rate categories) of promoter and coding sequences evolutions (-lnL = 8653.62 and 6713.66) with a gamma shape value of 2.20 and 0.38, respectively. ProtTest3 was used for determination of amino acid sequence evolution best-fit model. The best resulted to be the JTT + G + F model (-lnL= -2777.9) with a gamma shape value (four rate categories) of 0.46. By comparison of estimated substitution rate mean values (Figure 2A), statistical significant ($P < 0.05$, One-Way ANOVA tests) differences among mean rates of the analyzed categories were found.

In fact, as shown in Table 2, the promoter substitution rate mean value (1.169±0.008) was notably higher than the coding and protein sequence substitution rates (0.288±0.004 and 0.179±0.003).

We estimated also all ElDorado *Casq* promoter sequence substitution mean rate (Figure 2B). It resulted to be significantly higher than the mammals promoter mean rate (2.4521±0.012 to 1.169±0.008, $P < 0.05$ by unequal variances student's t test). However this difference was not so elevated like that between mammals and coding (or protein) substitution mean rates (Table 2). These results suggest different evolutionary pressure operating onto the CASQ gene (protein) and promoter sequences. Although substitution rate differences were present between promoter and amino acid sequences, mammals sequence phylogenetic relationships remained the same at the phylogenetic topology of promoter and amino acid sequence based trees (Figure 1A and B, respectively). This suggested that the majority of substitutions didn't change the structure/ function of the *Casq* promoters and *Casq* genes.



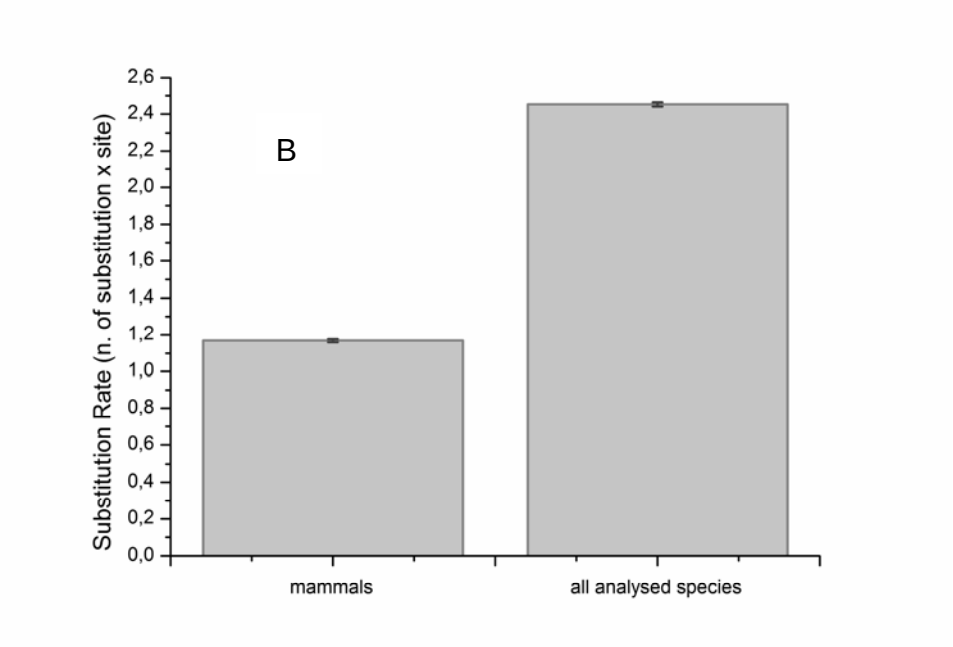


Figure 2. Graphical representations of estimated substitution rates in mammals. (A) The *Casq* promoter and coding DNA sequence rates were calculated as the number of nucleotide substitutions for nucleotide site, while the CASQ amino acid sequence rates were calculated as the number of amino acid substitutions for amino acid site. One-Way ANOVA Tukey, Bonferroni and Sceffé test results showed that estimated substitution rate means are significantly different between them for $P < 0.05$. (B) The all analyzed species and mammals *Casq* promoter rates were calculated as the number of nucleotide substitutions for nucleotide site. Unequal variances student's t test result showed that estimated substitution rate means are significantly different between them for $P < 0.05$.

Table 2. Calculated statistics for the estimated substitution rate means. Mean - the mean value of the sampled substitution rate across the chain (excluding the burn-in); standard error of mean - it takes into account the effective sample size, so a small ESS will give a large standard error of mean; median - the median value of the sampled substitution rate across the chain (excluding the burn-in); 95% HPD Lower - the lower bound of the highest posterior density (HPD) interval. The HPD is a credible set that contains 95% of the sampled values; 95% HPD Upper - the upper bound of the highest posterior density (HPD) interval; Auto-Correlation Time (ACT) - the number of states in the MCMC chain that two samples have to be from each other for them to be uncorrelated. The ACT is estimated from the samples in the substitution rate (excluding the burn-in); Effective Sample Size (ESS) – the number of independent samples that the substitution rate is equivalent to. This is essentially the chain length (excluding the burn-in) divided by the ACT [21]. The ESS are sufficiently large to make these analyses really reliable.

	Mammals promoter seq.	Mammals amino acid seq.	Mammals coding seq.	All analyzed species promoter seq.
Mean	1.1685	0.1786	0.2881	2.4521
Standard error of mean	8.13E-03	2.45E-03	3.76E-03	1.17E-02
Median	1.0514	0.1433	0.2359	2.3044
Geometric mean	1.0169	0.1318	0.216	2.2692
95% HPD lower	0.1983	3.86E-03	8.96E-03	0.8242
95% HDP upper	2.3883	0.4533	0.7108	4.3932
Auto-Correlation Time (ACT)	1541.4658	2736.2685	2660.5513	1301.5448
Effective Sample Size (ESS)	5839.2471	3289.5163	3383.1334	6915.6282

Though the *Casq* promoters mean rate was almost 2 times higher than coding region and protein mean rates, the *Casq* promoter structure /function remained unaltered like *Casq* gene. Thus, it could have been a consequence of the higher frequency of "synonymous" (no altered structure/function) substitutions in the *Casq* promoter than in *Casq* gene. This suggested that negative selection could have operated with higher degree onto the *Casq* promoter region than on *Casq* gene or CASQ protein, responsible for *Casq* promoter sequence conservation. Negative selection prevents deleterious mutation from reaching common frequencies and so should produce

an excess of rare variation. It is acceptable that deleterious mutations are present also in non-coding regions [22]. In conclusion our analysis indicates that mammals *Casq* promoter (non-coding) region could be a perfect platform where the negative or purifying selection might operate to conserve structurally and/or functionally important nucleotide motifs.

4. References

1. Sanchez EJ, Lewis KM, Danna BR, Kang C: **High-capacity Ca^{2+} binding of human skeletal calsequestrin.** *Journal of Biological Chemistry 2012*, **287**: 11592-11601.
2. Boncompagni S, Thomas M, Lopez JR, Allen PD, Yuan Q, Kranias EG: **Triadin/junctin double null mouse reveals a differential role for triadin and junctin in anchoring CASQ to the jSR and regulating Ca^{2+} homeostasis.** *PLoS One 2012*, **7**: e39962.
3. Tomasi M, Canato M, Paolini C, Dainese M, Reggiani C, Volpe P: **Calsequestrin (CASQ1) rescues function and structure of calcium release units in skeletal muscles of CASQ1-null mice.** *American Journal of Physiology - Cell Physiology 2012*, **302**: 575-586.
4. Yang A, Sonin D, Jones L, Barry WH, Liang BT: **A beneficial role of cardiac P2X4 receptors in heart failure: rescue of the calsequestrin overexpression model of cardiomyopathy.** *American Journal of Physiology 2004*, **287**: 1096–1103.
5. Milstein ML, Houle TD, Cala SE: **Calsequestrin isoforms localize to different ER subcompartments: evidence for polymer and heteropolymer-dependent localization.** *Experimental Cell Research 2009*, **315**: 523-534.
6. Chopra N, Yang T, Asghari P, Moore ED, Huke S, Akin B: **Ablation of triadin causes loss of cardiac Ca^{2+} release units, impaired excitation-contraction coupling, and cardiac arrhythmias.** *Proceedings of the National Academy of Sciences 2009*, **106**: 7636-7641.
7. Murphy RM, Larkins NT, Mollica JP, Beard NA, Lamb GD: **Calsequestrin content and SERCA determine normal and maximal Ca^{2+} storage levels in sarcoplasmic reticulum of fast- and slow-twitch fibres of rat.** *Journal of Physiology 2009*, **587**: 443-460.
8. Park KW, Goo JH, Chung HS, Kim H, Kim DH, Park WJ: **Cloning of the genes encoding mouse cardiac and skeletal calsequestrins: expression pattern during embryogenesis.** *Gene 1998*, **217**: 25–30.
9. Sacchetto R, Volpe P, Damiani E, Margreth A: **Postnatal development of rabbit fast-twitch skeletal muscle: accumulation, isoform transition and fibre distribution of calsequestrin.** *Journal of Muscle Research and Cell Motility 1993*, **14**: 646–653.
10. Richard AF, Demignon J, Sakakibara I, Pujol J, Favier M, Strochlic L, Le Grand F: **Genesis of muscle fiber-type diversity during mouse embryogenesis relies on Six1 and Six4 gene expression.** *Developmental Biology 359*, **359**: 303-320.
11. Donoghue P, Doran P, Dowling P, Ohlendieck K: **Differential expression of the fast skeletal muscle proteome following chronic low-frequency stimulation.** *Biochimica et Biophysica Acta 2005*, **1752**: 166-176.
12. Kinnunen S, Mänttäri S: **Specific effects of endurance and sprint training on protein expression of calsequestrin and SERCA in mouse skeletal muscle.** *Journal of Muscle Research and Cell Motility 2012*, **33**: 123-130.
13. Stearns SC: **Evolutionary medicine: its scope, interest and potential.** *Proceedings of the Royal Society B - Biological Sciences 2012*, **280**: 1750.
14. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A: **MatInspector and beyond: promoter analysis based on transcription factor binding sites.** *Bioinformatics 2005*, **21**: 2933-2942.
15. Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for fast and accurate multiple sequence alignment.** *Journal of Molecular Biology 2000*, **302**: 205-217.
16. Darriba D, Taboada GL, Doallo R, Posada D: **ProtTest 3: fast selection of best-fit models of protein evolution.** *Bioinformatics 2011*, **27**: 1164-1165.
17. Posada D: **jModelTest: Phylogenetic Model Averaging.** *Molecular Biology and Evolution 2008*, **25**: 1253-1256.
18. Guindon S, Gascuel O: **A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.** *Systematic Biology 2003*, **53**: 696-704.
19. Drummond AJ, Rambaut A: **BEAST: Bayesian evolutionary analysis by sampling trees.** *BMC Evolutionary Biology 2007*, **7**: 214.
20. Yang Z: **PAML 4: phylogenetic analysis by maximum likelihood.** *Molecular Biology and Evolution 2007*, **24**: 1586-1591.

21. Infante C, Ponce M, Manchado M: **Duplication of calsequestrin genes in teleosts: molecular characterization in the Senegalese sole (*Solea senegalensis*).** *Comparative Biochemistry and Physiology Part B - Biochemistry and Molecular Biology 2011*, **158**: 304-314.

22. Wakeley J: **Substitution rate variation among sites in hypervariable region 1 of human mitochondrial DNA.** *Journal of Molecular Evolution 1993*, **37**: 613-623.