Evolution of Vertebrate Ryanodine Receptors Family in Relation to Functional Divergence and Conservation

Zhiwen Ding,1,2,* MSc, Juan Peng,1 MD, Yanyan Liang,1,3,* MD, Chunjie Yang,1 BSc, Guoliang Jiang,1 MD, Jun Ren,1,4 PhD and Yunzeng Zou,1,2 MD

Summary

Ryanodine receptors (RyRs), the large homotetrameric protein complexes, regulate the release of calcium from intracellular stores into the cytosol and play vital roles in the excitation-contraction coupling of cells. However, the evolutionary relationship of RyRs in vertebrates has yet to be elucidated. We identified 22 RyRs from Homo sapiens, Mus musculus, Rattus norvegicus, Gallus gallus, Anolis carolinensis, Rana catesbeiana, and Danio rerio. The phylogenetic relationship, motifs analysis and reconstruction of ancestral RyRs showed that the members of RyR family in vertebrates were grouped into three clades: the RyR1 clade, the RyR2 clade, and the RyR3 clade. Positive selection existed in RyR gene evolution, which is consistent in three site models, and gene ontology (GO) analysis showed that the evolution of RyR family in vertebrates promotes RyRs function differentiation. At last, we predicted 140 mutation sites which may be involved in diseases and 57 phosphorylation sites among RyR1 sequence in human, as well as 61 mutation sites and 70 phosphorylation sites in human RyR2 sequences. Most of these potential sites are arranged in clusters. Our work provides insight into the origin and evolutionary process of RyRs in vertebrates, facilitating their functional investigations in the future.

Key words: Phylogenetic relationship, Motifs analysis, Ancestral reconstruction, Positive selection, Mutation site, Phosphorylation site

Cardiac hypertrophy is a complex process and the abnormality of intracellular Ca2+ homeostasis is critical one of major aspects.1-3) Ryanodine receptor (RyR) is one family of intracellular Ca2+ release channels which regulates the release of Ca2+ from the endoplasmic reticulum or sarcoplasmic reticulum into the cytosol.4) RyR is the largest known ion channel, in form of homotetramers (approximately 2,200 kDa), each subunit of which is comprised of a large N-terminal cytoplasmic domain, accounting for 4/5 of the protein, that modulates the gating of the channel, and 1/5 proportion being luminal and transmembrane spanning (TM) domains.5) RyR is an efficient gateway for Ca2+ release, being involved in the process of excitation-contraction (EC) coupling in contractile tissues, which process links sarcocalmern depolarization with the rise of cellular calcium transients and subsequent contraction.6) RyR also plays a vital role in controlling multiple developmental processes in vertebrates.7)

RyRs exist in both animals and unicellular eukaryotes. There are several isofoms in vertebrates, for example, three mammalian isofoms and five ones in fish. Not all of these isofoms are tissue specific, RyR1 is mainly expressed in skeletal muscle and in cerebellar Purkinje neurons,8,9) instead RyR2 in cardiac myocytes,10) which is also considered to be the most abundant isofom in the brain,11) and RyR3 is detected in smooth muscle, brain, and testis.12,13,14) RyR1 and RyR2 are responsible for myocyte contraction in their respective tissues, moreover, one recent study found that RyR2 is required for the development of pressure overload-induced cardiac hypertrophy.15) However, the role of RyR3 is unclear, although this isofom is discovered earlier in skeletal muscle development, and a role for RyR3 in learning and memory has been proposed based on murine knockout studies.16)

In contrast to expression profiles and functions of RyRs in vertebrates, little is known about the evolution of RyR family, except three main members in mammals sharing 65% amino acid identity.

In this study, we investigated the phylogenetic relationship among RyRs, and reconstructed the ancestral

From the 1Shanghai Institute of Cardiovascular Diseases, Zhongshan Hospital, Fudan University, Shanghai, China, 2Institute of Biomedical Sciences, Fudan University, Shanghai, China, 3Department of Cardiology, The First People’s Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China and 4University of Wyoming College of Health Sciences, Laramie, USA.

*These authors contributed equally to this work.

This work was supported by National Natural Science Foundation of China (81220108003, 31430039, and 81200085).

Address for correspondence: Yunzeng Zou, MD, Shanghai Institute of Cardiovascular Diseases, Zhongshan Hospital, Fudan University, 180 Feng Lin Road, Shanghai 200032, China. E-mail: zou.yunzeng@zs-hospital.sh.cn

Received for publication November 3, 2016. Revised and accepted February 14, 2017.

Released in advance online on J-STAGE November 17, 2017.

doi: 10.1536/ihj.16-558

All rights reserved by the International Heart Journal Association.
RyRs. Subsequently, results of positive selection detection and GO analysis in RyRs revealed that evolution of RyR family to their functions. In addition, we analyzed the mutation sites and phosphorylation of RyR in human and discussed their possible roles with disease.

**Methods**

**Search members and multiple sequence alignments:** We obtained the putative orthologous and paralogous RyR sequences of vertebrates from The National Center for Biotechnology Information, Mouse Genome Informatics, and The Zebrafish Model Organism Database using BLASTP. Based on the results of BLASTP searches, at the same time we obtained the coding sequence (CDS) of these RyR genes. In order to confirm the CDS, we examined the domains which translated nucleotide sequences into amino acid sequences using the SMART Database.

The obtained RyR sequences of vertebrates were aligned by the program MUSCLE3.6 with the default parameter setting. We also used ClustalX software to align sequences and recheck the result as a secondary method. In order to achieve a better result, the alignment result was then adjusted by manually based on the location of the corresponding amino acids in the RyR motif using GeneDoc (version 2.7.000; http://www.psc.edu/index.php/user-resources/software/genedoc).

**Construction of phylogenetic tree:** The RyR phylogenetic tree was constructed with the aligned RyR sequences by neighbor-joining (NJ) method with 1000 bootstrap replicates, poisson correction methods, and pairwise deletion of gaps as implemented in MEGA 6. Meanwhile, the PhyML software (http://atgc.lirmm.fr/phyml/) was used to construct maximum likelihood (ML) tree with the WAG model and an estimated proportion of invariant sites plus 8 categories of gamma distribution of substitution rates and 100 nonparametric bootstrap replicates as a secondary method to validate the results. In addition, the MrBayes software was used to construct Bayesian trees after running for 10⁶ generations, with four Markov chains, and sampled every 1000 generations. The WAG model was used for amino acid substitutions and variable plus 8 categories of gamma-distributed substitution rates were utilized to correct the among-site substitution rate heterogeneity. Tree files were viewed by TreeView software. ML tree was shown with bootstrap values from NJ, ML, and Bayesian analyses.

**Estimation of positively selected sites and ancestral RyR sequence reconstruction:** To test whether some sites in RyRs are under positive selection, we used three pair models from the CODEML program: M0 (one ratio) versus M3 (discrete), M1a (nearly neutral) versus M2a (positive selection), M7 (beta) versus M8 (beta & omega). We also used CODEML program to compute ancestral sequences reconstruction of RyR in vertebrates. The obtained RyR ancestral sequences of vertebrates were aligned with human, mouse, and rat RyR family sequences by the program MUSCLE3.6 with the default parameter setting.

**GO analysis:** We obtained human and mouse RyR family proteins uniprot ID (http://www.uniprot.org/). Subsequently, we analyzed GO of human and mouse RyR family using STRAP 1.5 with the default parameter setting.

**Mutation sites and phosphorylation sites analyzes:** Mutation sites data of human RyR family were obtained from ExPASy Database (http://www.expasy.org/) and phosphorylation sites data of human RyR family were obtained from PSP Database (http://www.phosphosite.org/home/Action).

**Results**

**Classification of RyR sequences in vertebrates:** Here we chose Homo sapiens (Human), Mus musculus (Mouse), Rattus norvegicus (Rat), Gallus gallus (Chicken), Anolis carolinensis (Lizard), Rana catesbeiana (Frog) and Danio rerio (Zebrafish) to represent vertebrate class. To obtain orthologous and paralogous RyR members in vertebrates, we used the full length sequence of RYR1 protein from Human as the TBLASTN query. Candidate amino acid sequences of RyRs were obtained from databases of National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/), MGI (Mouse Genome Informatics; http://www.informatics.jax.org/) and ZFIN (The Zebrafish Model Organism Database; http://zfin.org/). Subsequently, redundant members were removed according to SMART (http://smart.embl-heidelberg.de/) and UniProt (http://www.uniprot.org/). Eventually, 22 vertebrate RyR proteins were harvested, among which Human, Mouse, and Rat containing 3 ones each, Chicken and Frog each having two members, 4 RyRs in Lizard, and five ones in Zebrafish. In addition, we selected the RyR protein sequence of C. elegans (Caenorhabditis elegans) as the outgroup.

**Phylogenetic analysis of RyRs:** We conducted multiple sequence alignment among RyR proteins using Muscle software to determine the RyRs phylogeny. To examine the evolutionary process of RyRs, we constructed the rooted phylogenetic tree by means of three methods (NJ, ML and Bayesian inference) using the full length sequences of 23 RyRs. Results showed that the three phylogenetic trees generated had the pretty similar topology structures, which was depicted in the tree with the NJ, ML and Bayesian inference bootstrap values on it (Figure 1). Overall, based on the topology structure and bootstrap values, the phylogeny is divided into three clades, and the first clade contained 8 RyR1 proteins, the second clade contained 7 RyR3 sequences and the third clade contained 7 RyR2 sequences. We named the three clades RyR1 clade, RyR3 clade, and RyR2 clade. The phylogenetic tree also manifested that Clade RyR2 is the most ancient subfamily in vertebrates group compared with the other two clades based on the evolutionary distance, and RyR members may undergo expansion by gene duplication in vertebrates.

Interestingly, we also found some independent duplication events from the phylogenetic analysis. For example, both RyR1 clade and RyR2 clade in Zebrafish have two copies each, however, there was only one single RyR3 gene, which is distinct from other teleost fishes appear having two ryr3 genes in vivo. Therefore our results suggested in Zebrafish an independent duplication event.
once occurred before vertebrate duplication, and one of two putative RyR3 genes might get lost during evolution.

To further investigate the evolution of the RyR family and the dynamic change in the protein sequences of vertebrates, we also analyzed the conservation of sequence domains. A total of eleven domains were detected in RyR family of vertebrates by the SMART Database (Figure 1), and most members of the RyR family in the same clade shared the common domains, indicating the classification of RyR proteins into these three subfamilies is reliable24). In addition, RyR family has a highly conserved N-terminus, being composed of 4 MIRs, 2 RYDR_TTPRs, 3 SPRYs, and 2 RYRs, all of which exist in each RyR member of the phylogenetic tree. Compared with those in *C. elegans*, most sequences of vertebrate RyR families own only one PDB domain called calcium binding protein (2BCX), except Zebrafish RyR2a and RyR1-like from Lizard, suggesting RyR2 clade may be the ancient one among this family. In addition, a SCOP domain exists in all members of the RyR3 clade, which contains the EF-hand_8 domain showing that these RyR3s might gain the new divergent domain to displace the original one during the evolution process from the ancestor. In the RyR1 clade, all members only have one EF-hand_8 domain, except RyR1s in Lizard, which interestingly still hold two other additional EFh domains, indicating that similar with the RyR3 clade, RyR1 clade might also earn the EF-hand_8 domain instead of the two old EFh domains after the divergence from the ancestor, except the Lizard RyR1 undergoing the strong purifying selection.

**Ancestral RyRs:** Firstly, we reconstructed the ancestral sequence for RyR families in vertebrates. Then *Human*, *Mouse*, and *Rat* RyR sequences were pairwise aligned with the ancestral one. The alignment result revealed that 870 sites belonging to different clades among the amino acids had been changed, among which 512 site mutation only specially happened in one single clade, composing of 230 in RyR1 clade, 170 in RyR3 clade and 112 in RyR2 clade (Figure 2), 297 occurred in two clades and 61 sites were altered in all three clades. This result further demonstrated that among RyR1, RyR2, and RyR3, RyR2 clade is the gene group closest to the ancient RyR.

**Detection of positive selection on RyRs:** RyR family in vertebrates is divided into three clades, but whether they went through positive selection during evolution is still unknown. Here we screened the positive selection sites by using the site model implemented in PAML. The three
Parameter Estimates under Different Site Pair Models for the RyR Family

<table>
<thead>
<tr>
<th>Model</th>
<th>Estimates of parameters</th>
<th>LRT (pairs)</th>
<th>df</th>
<th>2Δl</th>
<th>P</th>
<th>Positively selected sites BEB</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0: one ratio</td>
<td>ω = 0.04649</td>
<td>M0/M3</td>
<td>4</td>
<td>9112.794</td>
<td>P &lt; 0.001</td>
<td>None</td>
</tr>
<tr>
<td>M3: discrete</td>
<td>p_0 = 0.40584, p_1 = 0.42380, p_2 = 0.17036;</td>
<td>M1/M2</td>
<td>2</td>
<td>6.664892</td>
<td>P &lt; 0.05</td>
<td>517 (T),2166 (E),4063 (K),4064 (R),4065 (V),4068 (G),4071 (C),4072 (L),4787 (L),4998 (V),5043 (E),5044 (P),5047 (P),5051 (A),5063 (K),5149 (E),5154 (F),5155 (E),5157 (S),5167 (S),5168 (G),5169 (S)</td>
</tr>
<tr>
<td>M1a: neutral</td>
<td>p_0 = 0.94095, p_1 = 0.05905;</td>
<td>M1/M2</td>
<td>2</td>
<td>6.664892</td>
<td>P &lt; 0.05</td>
<td>517 (T),2166 (E),2241 (T),4063 (K),4064 (R),4065 (V),4068 (G),4071 (C),4072 (L),4787 (L),4998 (V),5043 (E),5044 (P),5047 (P),5051 (A),5063 (K),5149 (E),5154 (F),5155 (E),5157 (S),5167 (S),5168 (G),5169 (S)</td>
</tr>
<tr>
<td>M2a: selection</td>
<td>p_0 = 0.94435, p_1 = 0.00111;</td>
<td>M2/M3</td>
<td>2</td>
<td>6.664892</td>
<td>P &lt; 0.05</td>
<td>517 (T),2166 (E),2241 (T),4063 (K),4064 (R),4065 (V),4068 (G),4071 (C),4072 (L),4787 (L),4998 (V),5043 (E),5044 (P),5047 (P),5051 (A),5063 (K),5149 (E),5154 (F),5155 (E),5157 (S),5167 (S),5168 (G),5169 (S)</td>
</tr>
<tr>
<td>M7: beta</td>
<td>p = 0.54391, q = 8.65065</td>
<td>M7/M8</td>
<td>2</td>
<td>60.18138</td>
<td>P &lt; 0.001</td>
<td>517 (T),2166 (E),2241 (T),4063 (K),4064 (R),4065 (V),4068 (G),4071 (C),4072 (L),4784 (D),4787 (L),4998 (V),5043 (E),5044 (P),5045 (D),5047 (P),5049 (G),5051 (A),5063 (K),5149 (E),5154 (F),5155 (E),5157 (S),5167 (S),5168 (G),5169 (S)</td>
</tr>
<tr>
<td>M8: beta &amp; ω</td>
<td>p_0 = 0.99680, p_1 = 0.55105, q = 9.00980, (p_1 = 0.00320), ω = 92.66927</td>
<td>M8/M9</td>
<td>2</td>
<td>60.18138</td>
<td>P &lt; 0.001</td>
<td>517 (T),2166 (E),2241 (T),4063 (K),4064 (R),4065 (V),4068 (G),4071 (C),4072 (L),4784 (D),4787 (L),4998 (V),5043 (E),5044 (P),5045 (D),5047 (P),5049 (G),5051 (A),5063 (K),5149 (E),5154 (F),5155 (E),5157 (S),5167 (S),5168 (G),5169 (S)</td>
</tr>
</tbody>
</table>

models applied were as follows: M0 (one ratio) versus M3 (discrete), M1a (nearly neutral) versus M2a (positive selection) and M7 (beta) versus M8 (beta & ω). Table I showed the statistics of the likelihood ratio (2Δl) for the hypotheses tested. For pair models of M0 versus M3 and M7 versus M8, the p-values by LRT of them are both less than 0.01; and for the pair model of M1a versus M2a, the p-value are less than 0.05. These results proved that certain sites in RyR family of vertebrates were indeed under positive selection during evolution.

**Go analysis:** To understand the relationship between RyR function and evolutionary events, GO analyses of the three RyR clades in *Human* and *Mouse* were separately performed (Figure 3). In aspects of biological process and cellular component, Clade RyR1 and RyR2 had the similar class structure; however, RyR3 clade was short of developmental process (0% in RyR3, while 16.7% and 17.6% in RyR1 and RyR2, respectively) and response to stimulus (0% in RyR3, while 16.7% and 17.6% in RyR1 and RyR2, respectively) and lacked extracellular, plasma membrane and other intracellular organelles (41.2% and 38.9% in RyR1 and RyR2, respectively), but gains additional component-nucleus in cellular component (15.4%), compared with other clades. In the molecular function aspect, functions of these three clades had a comparable classification, especially for those of RyR2 and RyR3 clade, which were completely the same. In total, different clades of RyR family have their own diversified functions, which may be attributed to the role of evolution.

**Mutation sites:** To further investigate the profound effect of evolutionary events on RyR function, we examined the distribution of mutated sites in *Human* RyRs. RyR1 was found as a candidate pathogenic protein for Malignant hyperthermia 1 (MHS1), Central core disease (CCD), Mul-
timinicore disease with external ophthalmoplegia (MMDO) and Congenital fiber type disproportion (CFTD) in Human. In this study, 140 mutations were detected in RyR1 in which the mutation sites clustered into six major regions (Figure 4A). These six regions together almost covered all domains, except the PDB domain. Interestingly, some mutations reported associate with kinds of diseases, for example, the mutation M23 leading to either CCD or MMDO, and some other ones causing amino acid substitution, such as R163L/C, were all seen (Figure 4A). Figure 4B illustrated that Arginine (R) is the amino acid most easily being mutated among the whole RyR1 sequence in Human, which accounts for over 1/3 of the altered ones, besides the second one most frequently changed is Glycine (G), occupying more than 10%. In addition, we also conducted a statistical analysis on the amino acids after mutation. The generated amino acids were mainly composed of Cysteine (C) and Histidine (H), each having more than 10%, respectively (Figure 4B). RyR2 was discovered as a candidate for Catecholaminergic polymorphic ventricular tachycardia 1 (CPVT1) and Arrhythmogenic right ventricular dysplasia 2 (ARVD2) in Human. We also found 61 mutations in RyR2 in which the mutation sites clustered into four major regions, and no mutations were discovered in regions of divergent sequences (DR1, DR2, and DR3). All of these mutations were associated with CPVT1 (Figure 4A). We also found that different mutation could lead to different diseases. For instance, there were four mutation sites not only associated with CPVT1 but also with ARVD2, and some mutation sites could cause the encoded amino acid change, such as R414L/C. In addition, we explored the amino acid composition before and after mutation events (Figure 4B). The result showed that Arginine (R), Alanine...
and Asparagine (N) are the main amino acids before mutation, which having proportion of over 10% each. Moreover, we analyzed the new generated amino acids after mutation. Amino acids after change were composed of Serine (S), Isoleucine (I) and Proline (P), having proportion of over 10% each (Figure 4B). Interestingly, some mutation sites of RyR1 overlapped well with the mutation sites of RyR2, for example, R163L/C, V2280T, R2355C, A2383Q,
Figure 5. Schematic diagram of RyRs phosphorylation sites in human. The phosphorylation sites are marked by the amino acid’s name and position number.

and V4771I in RyR1 corresponding to R169Q, V2306I, R2359Q, A2387P/T and V4771I in RyR2, respectively.

**Phosphorylation sites:** As is reported, protein phosphorylation is closely associated with the functions of RyRs in human.33,34) So we searched for the phosphorylation sites among Human RyR amino acid sequences. We found 57, 70, and 29 phosphorylation sites in Human RyR1, RyR2 and RyR3, respectively (Figure 5). These sites were distributed in clusters, and also grouped in most domains of RyR1 and RyR2. There was no phosphorylation found in RyR1 PDB domain and EF-hand_8, and PDB domain, RIH_assoc domain and Ion_trans domain of RyR2. However, compared with that in RyR1 and RyR2, the phosphorylation sites in RyR3 are quite fewer, suggesting RyR3 might be the most distinctive one among these three clades.

**Discussion**

RyRs are homotetrameric proteins that regulate the release of calcium from intracellular stores. Here the evolutionary relationship of RyR family in vertebrate was investigated. Phylogenetic analysis shows that RyR family can be divided into three clades evolutionarily, and RyR2 clade is the most conserved group compared with the other two RyRs in vertebrates. Motif distribution and the ancestral RyR sequence reconstruction further confirm the above conclusion. Positive selection and GO analyses shows that the evolution of RyR family in vertebrates promotes function divergence of the RyR family in biological processes and cellular components. Mutation sites and phosphorylation sites of RyRs in human are arranged in clusters.

**RyRs may originate in RyR2 clade:** Previous analyses of RyR family show that RyR family only has one gene in invertebrate and multiple ones in vertebrates.32,35) Our results show that RyR family in vertebrates is divided into three clades. Phylogenetic analysis suggests that RyR2 clade is the most ancient one in vertebrates. Protein domain distribution plays a key role in the evolution process. Domain analysis reveals that vertebrate RyR2 clade shares highly similar domain assignment with that in C. elegans except the PDB domain which only exists in vertebrate RyR family. This result also suggests that vertebrate RyR2 clade members are much closer to RyRs in invertebrates than the other two clade ones of vertebrates. Results of ancestral RyR sequence reconstruction show that RyR sequences have the fewest amino acids mutated. All these results confirm that RyR2 clade is more like RyR of invertebrates in evolution, which also reveal that vertebrate
RyR2s may undergo the strongest purifying selection. Evolution of RyR promotes its function differentiation: Protein function is closely associated with evolution. Positive selection analysis suggests that positive selection on RyR peptides in vertebrates is involved in their function differentiation. In addition, GO analysis results imply that function differentiation of RyR family in vertebrates mainly focuses on their functions in biological process and cellular component. As we all known, RyR genes have differential expression patterns. RyR-J is mainly expressed in skeletal muscle, RyR-2 in cardiac muscle, and RyR-3 in brain. Proteins in the same clade display the similar domain organization, indicating the results of our phylogenetic analysis are reliable. For the RyR protein general structure, 4/5 of it is located in cytoplasm and the remaining part is responsible for lumen localization and membrane spanning. However, RyR2a in Zebrafish and RyR1-like in Lizard both lack the entire C-terminus, the former of which is exclusively expressed in the developing central nervous system of Zebrafish embryos instead of the developing Zebrafish heart where RyR2b is expressed, suggesting that RyR2a function is specialized in the evolution.

The domain classes and distribution in RyR sequences are quite conservative, including 4 MIR domains, 2 RYDR_ITPR domains, 3 SPRY domains, 4 RYR domains, 1 RIH_assoc domain, 1 TM regions and 1 Ion_trans domain. This conservation relationship between family members is closely associated with the evolution of RyRs. MIR is the abbreviation name for mannosyltransferase. Although most of these domains are conserved in both lack the entire C-terminus, the former of which is exclusively expressed in the developing central nervous system of Zebrafish embryos instead of the developing Zebrafish heart where RyR2b is expressed, the mutations within are associated with CPVT and ARVD2. Previous results showed that the disease-associated mutations of RyR1 cluster were mainly located in three regions, and four highly enriched ones for those in RyR2 clade, however, we found that there were six related zones in RyR1 instead probably due to a large sample of mutated sites being discovered in this study. The mutation hot-spot regions within RyR3 have not been reported yet. These results implied that the evolution of RyR family may be driven by relaxed selection. Interestingly, just like the mutation, the phosphorylation sites of RyRs also tended to gather together in some special spots, and phosphorylation status of RyRs can influence the open probability of homotetrameric complex formed between each other, finally leading to heart failure or arrhythmia.

Conclusion

In sum, we have analyzed the phylogeny of the RyR family in vertebrates and found that the RyRs was evolutionarily grouped into three clades. In addition, functions of RyRs are closely related to their evolution pattern. At last, both the human disease-associated mutation and phosphorylation sites clustered together in RyR members, and there are six and four specific regions in RyR1 and RyR2, respectively. Our results lay a solid theoretical foundation for the RyR function study; at the same time provide new clues for the future research of disease susceptibility.

Disclosures

Conflicts of interest: No potential conflicts of interest relevant to this paper.

References