Journal of Phylogenetics & Evolutionary Biology

Zheng L, et al., J Phylogen Evolution Biol 2015, 3:2
DOI: 10.4172/2329-9002.1000150

Research Open access

The Evolution of Calcium Release Channels: a Story of Expansion and Loss?

Limian Zheng and John James Mackrill*

Department of Physiology, School of Medicine, University College Cork, BioSciences Institute, College Road, Cork, Ireland

***Corresponding author:** John James Mackrill, Department of Physiology, School of Medicine, University College Cork, BioSciences Institute, College Road, Cork, Ireland, Tel: +353 (0)21 490 1400; Fax: +353 (0)21 4205370; Email: j.mackrill@ucc.ie

Received date: March 21, 2015, **Accepted date:** March 30, 2015, **Published date:** April 03, 2015

Abstract

Calcium ions are utilised as a second messenger in all forms of cellular life. In contrast to bacteria and archaea, eukaryotes possess endomembrane systems, exemplified by the endoplasmic reticulum. Such organelles act as intracellular stores in Ca^{2+} signalling processes, with two distantly related calcium channels, the inositol 1,4,5-trisphosphate receptors and ryanodine receptors, acting as Ca^{2+} release mechanisms. Despite their fundamental role, the evolutionary origins of such Ca^{2+} release channels have proven difficult to elucidate. The current study presents updates on the phylogeny of this channel superfamily and analyses of the domain architectures of these proteins. We demonstrate that inositol 1,4,5-trisphosphate receptor homologues are present in every major taxonomic group of eukaryotic life, suggesting that they were utilised early on in the evolution of these organisms. Certain taxonomic groups contain multiple Ca^{2+} release channel homologues, suggesting expansion and diversification. Early diverging fungi and green plants contain a single canonical inositol 1,4,5-trisphosphate receptor, which is absent in later branching members, suggesting loss from at least two distinct lineages. A key difference in the protein architecture of the two channel families is the presence or absence of ryanodine receptor domains. Such ryanodine receptor domains occur in multiple families of proteins present in eukaryotes, bacteriophage viruses, bacteria and archaea. In eukaryotes, canonical ryanodine receptors are first detected in the choanoflagellate/metazoan lineage, but a distinct family of ryanodine receptor domain-containing proteins with the potential to form cation channels, that we name the 'PKD-RR' family, are conserved among oomycetes. Modelling of the tertiary structures of ryanodine receptor domains from viruses, bacteria, fungi and oomycetes indicates that they are likely to closely resemble those from mammalian ryanodine receptor channels. We also present evidence that horizontal gene transfer has occurred during the evolution of ryanodine receptor domain-containing proteins, thereby contributing to calcium release channel structural and functional diversity.

Keywords: Calcium release channel; Ryanodine receptor; Inositol 1,4,5-trisphosphate receptor; Protein domains; Horizontal gene transfer

Introduction

Calcium ions are used as a second messenger within all three domains of cellular life, coupling intrinsic and extrinsic cues to downstream responses. Extensive experimental investigations have firmly established that eukaryotes use Ca^{2+} in this way to control almost every cellular process, from gene expression, to metabolism, to motility (for reviews, see [1, 2]). Although not as thoroughly investigated, evidence also supports a second messenger role for Ca^{2+} in the regulation of diverse functions in bacteria [3] and in methanogenesis in the archaeon Methanothermobacter thermautotrophicus [4]. Several hypotheses have been proposed to explain the evolution of the use of Ca2+ as a second messenger by cellular organisms, all of which are underpinned by the necessity of keeping the cytosolic level of this ion in a sub-micromolar range, owing to its unfavourable interactions with biomolecules [5,6]. In most environments, extracellular Ca^{2+} concentrations are in the low millimolar range, providing a large electrochemical gradient that favour influx of this ion and its potential use in a second messenger role. This function was facilitated by the early evolution of both calcium channels and extrusion mechanisms [6].

One key difference in the cell biology of eukaryotes compared to bacteria and archaea is the presence of endomembrane systems, including the endoplasmic reticulum (ER). Such endomembrane systems permit compartmentalisation of cellular activities and are endowed with unique subsets of proteins that perform these specialised roles. Many of these endomembrane located proteins lack detectable homologues in archaea and bacteria. It has been proposed that endomembrane specific proteins have been generated by duplication and rapid diversification of those present in the prior to the existence of the last eukaryotic common ancestor (LECA). Subsequently, post-LECA expansions resulted in the generation of multi-paralogue protein families, which were diversified or lost to generate lineage-specific complements of endomembrane components [7].

One role of the endomembrane system, in particular the ER, which is conserved in all eukaryotic lineages is that of Ca^{2+} storage, which regulates both protein synthesis and Ca^{2+} signalling. Intracellular Ca^{2+} stores are endowed with channels to increase the cytosolic levels of this ion in response to stimuli and with active transport mechanisms, to re-accumulate it. In terms of signal transduction, this function circumvents problems associated with high levels of buffering of this ion by cellular components that slow its diffusion, limiting the velocity at which Ca^{2+} entering via plasma membrane channels can reach deep within eukaryotic cells, which are typically much larger than bacteria and archaea [8]. In contrast, Ca^{2+} channels located in endomembrane systems can communicate 'locally' with Ca^{2+}-sensing effector mechanisms, without causing large changes of the levels of this cation in the bulk cytoplasm [1]. This has the advantages of increasing the

fidelity of signal transduction and reducing the 'global' levels of this cytotoxic ion within a cell.

Two major families of calcium release channel (CRC) protein were initially identified in the endomembranes of mammalian cells: the inositol 1,4,5-trisphosphate receptors (ITPRs) and the ryanodine receptors (RyRs). These channel proteins are members of a protein superfamily, with mammalian ITPRs sharing about 30% identity with mammalian RyRs [9]. Functional RyR channels are formed by association of four monomers, each of around 5000 amino acid residues. ITPRs are also tetrameric, composed of monomers of approximately 3000 residues. In mammals, each family is comprised of three homologues. Despite their similarities, ITPRs and RyRs display distinct protein domain organisations, pharmacology and electrophysiological properties [10]. In particular, RyRs are gated by rises in cytoplasmic Ca^{2+} concentration in a process termed Ca^{2+}-induced Ca^{2+}-release (CICR) or by allosteric interactions with voltage-gated calcium channels in the plasma membrane; whereas ITPRs are gated by the second messenger inositol 1,4,5-trisphosphate ($InsP_3$), the sensitivity to which is enhanced by CICR. Pharmacologically, RyRs are defined by complex interactions with the neutral plant alkaloid ryanodine and agonism by caffeine; while ITPRs are antagonised by low molecular weight heparin. In terms of structural features that distinguish RyRs from ITPRs, the former contain at least one tandem pair of repeated structures of about 100 residues in length, termed 'RyR domains' [11]. These structures are also present in a range of viral, archaeal, bacterial and oomycete proteins, suggesting that they might act as a module within proteins, conferring a specific function [12]. Three independent elucidations of high resolution structures of type 1 RyR (RyR1) from rabbit (Oryctolagus cuniculus) skeletal muscle indicates that RyR domains play a role in protein-protein interactions, possibly underpinning the assembly of supramolecular arrays of multiple RyR channel complexes [13-15].

Phylogenetic investigations have revealed the presence of ITPR and RyR homologues in multiple branches of the eukaryotic tree of life. Based on the available genomes, RyRs are reported to have first appeared in choanoflagellates [12,16-18], a sister lineage of metazoans that last shared a common ancestor at least 600 million years ago (Ma) [19]. ITPRs evolved earlier: combined phylogenetic, biochemical and physiological approaches have provided strong support for the presence of ITPR homologues in the alveolate *Paramecium tetraurelia* [20], the amoebozoan *Dictyostelium discoideum* [21] and the euglenozoan *Trypanosoma brucei* [22]. Physiological data strongly support the notion that CRCs are also present in both green plants and in fungi [23]. For example, vacuolar membranes from the storage root of beetroot (*Beta vulgaris*), an angiosperm plant, are endowed with two subsets of calcium channel, one of which is sensitive to $InsP_3$ and the other to cyclic ADP ribose, a candidate second messenger that is an agonist of RyRs [24]. Functional and genetic data demonstrate a role for $InsP_3$ in glucose-induced elevations in cytoplasmic Ca^{2+} in the yeast Saccharomyces cerevisiae [25], despite the well-characterised genome of this organism apparently lacking ITPR orthologues. Indeed, no detectable homologues of either ITPRs or RyRs have been identified in the genomes of green plants or fungi, despite the availability of high-quality, well-annotated genomic data for multiple members of each lineage. In parallel, it has been estimated that $InsP_3$ has been employed as a Ca^{2+}-mobilising second messenger within eukaryotes for at least 1000 Ma [26], early during the evolution of these organisms.

In the light of these controversies and the continually increasing number of organisms whose genomes have been sequenced, the current study sought to fulfil the following objectives: 1) to survey the presence of ITPR and RyR homologues among eukaryotes; 2) to gain insights into the evolutionary relationships between CRCs within and between different eukaryotic lineages; 3) to investigate the evolution of the RyR domain, that is involved in protein-protein interactions, in the light of expanding genomic information and increased knowledge of structure-function relationships in mammalian RyR complexes; and 4) to examine the mechanisms by which $InsP_3$ can release Ca^{2+} from the endomembrane systems of organisms that apparently lack homologues of mammalian CRCs. These analyses suggest that CRC families evolved early during eukaryotic evolution, and subsequently expanded, diverged and in some lineages, were lost over evolutionary time. For example, ITPRs orthologues with canonical protein domain organisations are present in both early branching fungi and plants but are undetectable in late-branching counterparts. Phylogenetic and protein-fold modelling analyses suggest the presence of diverse RyR domain containing proteins metazoa, choanoflagellates, fungi, bacteria, archaea and viruses. The presence of the RyR domain in multiple proteins from several lineages suggests that it is a promiscuous domain, perhaps being incorporated into various protein families by partial horizontal gene transfer (HGT) events [12]. Finally, the tertiary structures of candidate non-RyR, non-ITPR CRC proteins and domains were modelled, in order to determine if they share predicted structural similarities with those within canonical family members.

Methods

Identification of Calcium Release Channel Homologues and Prediction of their Protein Domain Architecture

In order to survey the presence of ITPR and RyR homologues throughout the tree of life, BLASTP [27,28] searches were conducted via the National Center for Biotechnology Information (NCBI) website: http://blast.ncbi.nlm.nih.gov/Blast.cgi/. Initially, the query sequences selected were mouse RyR1 (Mus musculus NP_033135.2) and ITPR1 (GenBank: EDK99409.1) proteins, as these are extremely well characterised representives of each CRC family. BLAST searches were conducted with standard algorithm parameters using a 'bottom-up' approach, initially searching within each taxonomic domain (eukaryotes, archaea, bacteria and viruses), then restricting searches to smaller branches of the tree of life. In general, an expected value of 1 x 10^{-3} was taken as the lower limit for a significant relationship between the query sequence and a subject, but in some cases potential homologues of higher E-values were BLASTed back against the mouse genome, or were modelled to examine any structural relationships with the query proteins. In cases in which taxa were over-represented (such as vertebrates, in particular mammals), examples of protein sequences of each CRC type were recorded, then the taxonomic group was then excluded from subsequent searches. In order to detect CRC homologues in under-represented eukaryotic taxa, the query sequences were used to probe genomic data at the 'Origins of Multicellularity' project at the Broad Institute: http://www.broadinstitute.org/annotation/genome/multicellularity_project/Blast.html.

The sequences of all of the CRC proteins present in representatives of major taxonomic groups were stored in FASTA format and the percentage identity, percentage similarity and E-value of each

homologue compared to the query sequence were tabulated. In order to compare the predicted protein domain architectures of the CRCs detected, protein sequences were analysed using the Conserved Domain Database [29] at NCBI: http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi. In some cases, the tertiary structures of domains detected in candidate CRC homologues were modelled using PHYRE2 (Protein Homology/analogY Recognition Engine V2.0) software at: http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index, using the intensive mode [30]. For prediction of transmembrane topology, protein sequences were analysed using two independent algorithms. These were TMPRED, at: http://www.ch.embnet.org/software/TMPRED_form.html and TopPred [31] at: http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::toppred.

Multiple Sequence Alignment (MSA) and Construction of Phylogenetic Relationships

MSA and phylogenetic analyses were performed using MEGA5 software [32]. MSA of either full-length proteins or protein domains was carried out with the MUSCLE programme using standard parameters [33]. Evolutionary relationships between CRC homologues were inferred using the Maximum Likelihood method, employing the Jones-Taylor Thornton (JTT) matrix-based model [34]. For each analysis, the tree topology with the highest log likelihood was retained. The statistical likelihood at each node was determined using the bootstrap method with 500 replicates [35].

Detection of Horizontal Gene Transfer (HGT) Events

HGT is one mechanism that could have contributed to the diversity of CRC protein architectures during evolution [12]. Putative HGT of RyR domains was investigated in two distinct databases of bacterial and archaeal genomes. One of these databases, HGT-DB at: http://genomes.urv.cat/HGT-DB/, detects genes with atypical DNA composition (G+C content and codon usage) to identify candidates that have potentially undergone HGT [36]. The other database, DarkHorse HGT Candidate Resource at: http://darkhorse.ucsd.edu/, employs an algorithm that uses both DNA composition and phylogeny to calculate a lineage probability index (LPI) score for each protein in a genome [37]. Within this database, an LPI score of less than 0.6 is suggestive of HGT.

Results and Discussion

ITPR and RyR Homologues are present in some, but not all, eukaryotic lineages: evidence of early expansion and loss from some taxa.

A survey of the presence of CRC homologues through the eukaryotic tree of life (based on references [7,38]) revealed a 'patchy' distribution, see Figure 1 and Table S1. For example, within the SAR (stramenopile-alveolate-rhizarial) supergroup, diverse stramenopile and alveolate species contain multiple CRC orthologues, displaying protein architectures characteristic of canonical metazoan members of this superfamily. In contrast, those proteins in rhizaria generating significant sequence alignments with mouse RyR1 showed strong identity only at Spore lysis A and RyR (SPRY) domains, which are ubiquitous among eukaryotes [12]. However, these proteins do not contain other domains characteristic of CRCs, suggesting that rhizaria lack RyR and ITPR channels. In contrast, homologues of the mouse sarcoplasmic/endoplasmic reticulum Ca^{2+} ATPase (SERCA) pump, the key mechanism for Ca^{2+}-sequestration into these organelles, were detected in multiple rhizarial genomes [16]. A lack of uniformity in the distribution of CRC homologues was also observed within all of the other major taxonomic assemblages. For example, ITPR homologues were detected in discoseans (Acanthamoeba castellanii) and mycetozoans (Dictyostelium fasciculatum), but were not detected in any other amoebozoans. Despite this, every major eukaryotic taxonomic assemblage contains representatives of the CRC superfamily. Although these are hypothetical proteins in most cases, there is strong genetic, physiological and biochemical experimental evidence to support the expression of CRC channel proteins in opisthokonts (with particularly comprehensive data for animals), excavates (Trypanosoma brucei [22]) and the SAR group (in particular for the alveolate Paramecium tetraurelia [20]). In agreement with previous studies, these observations indicate that CRCs arose early in the evolution of eukaryotes, prior to them branching into major groups [12,17,18,39,40]. Assuming that CRCs were present in LECA, current estimates would suggest that they have existed for between 1007 and 1898 Ma [41], correlating with the era that $InsP_3$ was first predicted to be employed as a Ca^{2+}-mobilising second messenger [26].

By analysing protein domain architectures, all candidate members of the CRC family contained the following canonical arrangement: an amino-terminal $InsP_3$ receptor domain ('I'), an RyR and ITPR homology associated domain (RIHA, 'A') and an ion channel domain ('C'). A possible exception to this is a hypothetical protein in the fungus Absidia idahoensis, which contains I and A domains, but apparently lacks an ion channel domain, see figure 1. However, inspection of the primary structure of this protein indicated that it contains a candidate selectivity filter (that allows it to discriminate between different types of ion), figure 2, and pore-lining transmembrane domain (not shown) that are highly conserved with other CRCs [39]. Canonical ITPRs are characterised by the presence of an additional mannosyltransferase, ITPR and RyR (MIR, 'M') domain and one or two copies of the RyR and ITPR homology (RIH, 'H'). RyRs differ from this architecture in terms of containing SPRY ('S'), ryanodine receptor (RyR, 'R') and EF-hand ('EF') domains. In both protein families, the $InsP_3$ receptor and MIR domains participate in channel gating, assembly and interaction with ligands. The RIH and RIHA domains have recently been shown to contribute large components of the cytoplasmic ultrastructure of rabbit RyR1, called the α-solenoids, which have been postulated to play a role in interactions with protein regulators [13-15]. It is tempting to speculate that the RIH and RIHA domains serve similar roles in the ITPRs and in other CRCs. SPRY domains are common in eukaryotic proteins and participate in intra- and inter-molecular protein-protein interactions, including the allosteric coupling between dihydropyridine receptor/voltage-gated calcium channels and RyR1 in mammalian skeletal muscle [42]. EF-hand domains were first demonstrated in lobster RyRs [43] and are thought to regulate the calcium dependent gating of these channels [13-15].

Certain CRC homologues contain diverse additional domains, generally located either close to the amino-terminus or adjacent to the channel domain (Figures 1 and 3) and (Table S1). The consequences of this are unclear, but it is likely that such domains would increase the functional diversity of CRC paralogues. Analogous patterns of domain insertion and loss have been reported for proteins in the Rab GTPase prenylation complex [44]. Such structural and functional diversity has been experimentally demonstrated in the ciliated alveolate P.tetraurelia, whose genome encodes 34 CRC subtypes, some of which display distinct subcellular locations and differential gating by ligands.

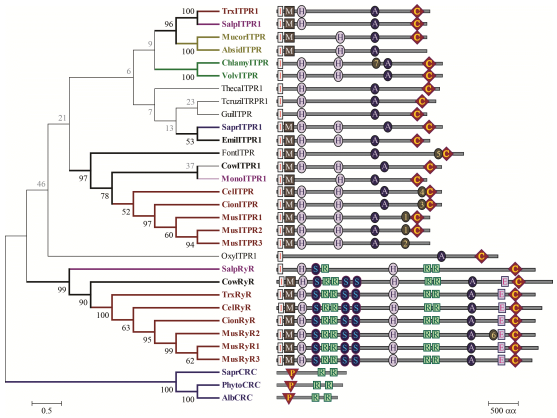


Figure 1: Phylogenetic and Protein Domain Analyses of Calcium Release Channels (CRCs). Homologues of mouse type 1 ryanodine receptor protein were identified by BLAST searching of genomic databases and their sequences aligned using MUSCLE then their phylogenetic relationships were modelled using the Maximum Likelihood method. The unrooted tree shown is the consensus with the highest likelihood (log likelihood of -49294.12). The numbers at each branch represent the boot-strap value from 500 replicates for that grouping, with a value of 50% taken as the threshold for a significant relationship. The scale-bar represents the number of substitutions per residue. Blue branches are oomycetes; maroon, metazoans; purple, choanoflagellates; green, viridiplantae; olive, fungi; black, other eukaryotes. Key to species names: Trx = Trichoplax adhaerens; Salp = Salpingoeca rosetta; Mucor = Mucor circinelloides; Absid = Absidia idahoensis; Chlamy = Chlamydomonas reinhardtii; Volv = Volvox carteri; Theca = Thecamonas trahens; Tcruzi = Trypanosoma cruzi marinkellei; Guil = Guillardia theta; Sapr = Saprolegnia declina; Emil = Emiliania huxleyi; Font = Fonticula alba; Cow = Capsaspora owczarzaki; Mono = Monosiga brevicollis; Cel = Caenorhabditis elegans; Cion = Ciona intestinalis; Mus = Mus musculus; Oxy = Oxytricha trifallax; Phyto = Phytophthora infestans; Alb = Albugo laibachii. For accession numbers, please see Supplemental Table 1(examples used are highlighted in bold font). The right-hand panel represents the corresponding domain architectures of these CRC homologues. Key to domains: I = Ins143_P3_rec domain = $InsP_3$ receptor domain; M = MIR domain = mannosyltransferase, ITPR and RyR domain; H = RIH domain = RyR and ITPR homology domain; A = RIHA domain = RIH associated domain; S = SPRY domain = Spore lysis A and RyR domain; R = RyR domain; C = channel domain; EF = EF-hand domain = calcium binding domain; P = PKD channel domain; 1 = COG1266 = Metal-dependent membrane protease domain; 2 = Bax1I = Inhibitor of apoptosis-promoting Bax1; 3 = SCL5/6 = Solute carrier 5/6-likedomain; 4 = PMT-2 = Dolichyl-phosphate-mannose-protein mannosyltransferase domain; 5 = MdlB = Multidrug transporter domain; 6 = Y = YjfB_motility domain; 7 = Dsh = Segment polarity protein dishevelled.

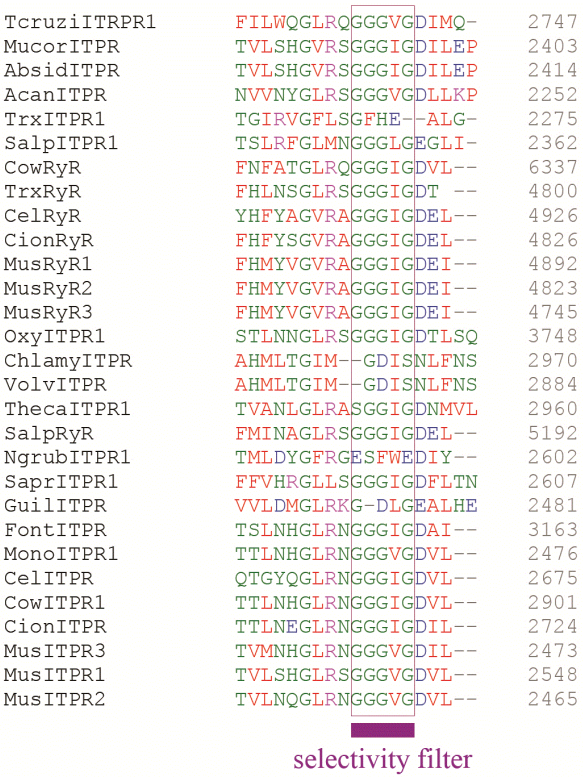


Figure 2: CRCs from Diverse Eukaryotes Share Similar Selectivity Filters. The selectivity filter of an ion channel is the structure that enables it to distinguish between different ions, either allowing them to enter the permeation pathway or excluding them. Canonical ITPRs and RyRs contain highly conserved selectivity filter motifs, of either GGVG or GGIG. In the MSA shown, most but not all, candidate CRCs share similar selectivity filter motifs. The key to the species abbreviations is given in Figure 1. The numbers to the right indicate the position of this alignment within each protein.

It was proposed that this structural and functional diversity could be explained by whole genome duplication and may facilitate local delivery of Ca^{2+} from distinct endocellular compartments [45,46]. In the current analyses, multiplicity of CRC homologues was also detected in heteroloboseans, such as Naegleria gruberi (Figure 3, 5 ITPR-like genes), oomycetes ('water-moulds', organisms that superficially resemble fungi, but are distinct from them) including the fish parasite Saprolegnia declina (Figure 3, 13 members), the cryoptophyte Guillardia theta (5 candidate homologues), other ciliates including Tetrahymena thermophila (23 putative CRCs), and either 6 or 7 subtypes in vertebrates. Compared with the three RyR subtypes present in most vertebrates, fish such as the blue marlin (Makaira nigricans) contain an additional RyR form within their fast-twitch muscles that shows a distinct primary structure and functional properties [47]. These observations indicate that diversification of CRC structures has occurred at multiple times throughout eukaryote evolution. This hypothesis is supported by the analysis of intraspecies evolution of RyRs in vertebrates, which suggests the occurrence of

positive selection at divergent regions within each protein, but purifying selection of each subtype overall [48].

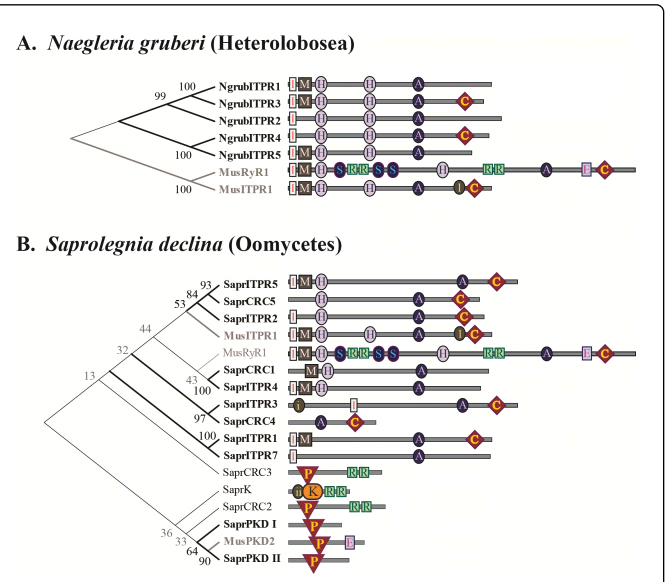


Figure 3: Examples of Expansion and Diversification of CRCs in A. the Heterolobosean Naegleria gruberi and B. the Oomycete Saprolegnia declina. Unrooted Maximum Likelihood trees were constructed as described in Fig. 1. Paralogous proteins sharing particular domains were included for comparative purposes (eg. PKD proteins from M. musculus and S. declina). Key to domains (also see Fig. 1): i = ankyrin repeat superfamily; ii = pleckstrin homology domain; K = calmodulin dependent protein kinase domain.

However, not all eukaryotic taxa possess multiple CRC subtypes. For example, most non-vertebrate animals contain one ITPR and one RyR homologue, compared with the total of 6 or 7 in vertebrates. An exception are porifera, such as the sponge Amphimedon queenslandica, whose genome potentially encodes one RyR and four ITPR subtypes. Early branching opisthokonts, such as Fonticula alba (of the Nucleariidae and Fonticula group) and apusomonads such as Thecamonas trahens, only have one or two CRC homologues, respectively. Within the SAR grouping, S. declina possesses 13 candidate CRC homologues, whereas Albugo laibachii, another oomycete ('water moulds', superficially resembling fungi), only contains only two. More striking examples are found among fungi and plants. Early diverging fungi such as Mucor circinelloides and Absidia idahoensis (both of the Order Mucorales) seem to have only one candidate ITPR homologue (Figure 1) which was not detected in later branching members. Both of these candidate ITPRs possess an RyR-like selectivity filter (Figure 2). Experimental evidence supports roles for inositol $InsP_3$-Ca^{2+} signalling in such fungi: enhanced phosphatidylinositol turnover and perhaps, $InsP_3$ mediated Ca^{2+}-release, regulate morphogenesis in M. racemosus [49]. Similarly, in the current, a candidate ITPR was detected in early diverging green plants, such as the chlorophytes Chlamydomonas reinhardtii and Volvox carteri, but not in later branching forms, such as the angiosperm Arabidopsis thaliana. Experimental evidence supports the presence of functional ITPRs in chlorophytes, since $InsP_3$ mobilises Ca^{2+} from an intracellular store in C. moewusii [50]. However, these chlorophyte CRC homologues did not contain a selectivity filter similar to those of canonical ITPRs (GGVG) or RyRs (GGIG) (Figure 2) suggesting that they might have distinct electrophysiological properties.

In the context of contemporary models of eukaryotic evolution [7,38], the current analysis supports the existence of CRCs very early during this process. Given the time involved, it is difficult to establish the evolutionary relationships between different CRC families, or even between the domains of life in which they operate. For example, there is uncertainty in even the direction of transition between eukaryotic and non-eukaryotic ('akaryotic') cellular life [51]. In some taxa, CRC diverged into a range of subtypes, probably by mechanisms involving whole genome duplication, horizontal gene transfer (HGT) and positive selection. In other taxa, there was a lack of detectable CRC homologues. Evidence from chlorophyte and fungal lineages suggests that this has probably occurred via loss of CRC homologues, rather than their divergence into forms that are no longer recognisable on the basis of sequence identity. An alternative explanation for $InsP_3$ induced Ca^{2+}-release in organisms that apparently lack ITPR homologues is that of homoplasy, or 'convergent evolution', in which unrelated structures evolve to fulfil equivalent roles [8]. An example of this might be the yeast vacuolar cation channel (Yvc1), which is not closely related to ITPRs in structure, but which releases Ca^{2+} from this endomembrane system in response to $InsP_3$ [52]. Furthermore, it is unlikely that all eukaryotic calcium signalling proteins have been identified to date. For example, a previously uncharacterised channel protein called OSCA1 has recently been demonstrated to be essential for hyperosmolality-induced Ca^{2+} signals in the plant A. thaliana [53]. This suggests that multiple homoplastic equivalents of CRCs could be present throughout eukaryotic evolution and raises the possibility that not all calcium channel families have been identified to date. Published evidence also supports the possibility of multiple expansions and losses of lineages during this evolution owing to ecological catastrophes, with the complement of genes in surviving organisms seeding the next expansion of protein structural diversity [54]. Such processes are likely to have contributed the diversity of CRCs in extant eukaryotes.

The RyR domain

One issue that arises from analysis of CRC domain architectures, is the question 'what makes an RyR an RyR, rather than an ITPR?' Pharmacologically, the answer is probably that RyRs are cation channels whose gating and conductance can be modified by the plant alkaloid ryanodine [10]. However, data on ryanodine interactions with potential RyR channels are restricted to organisms from a limited number of eukaryotic taxa. In terms of protein domain architecture, the answer is probably the presence of domains that are shared between RyRs and ITPRs ($InsP_3$ receptor, MIR, RIH, RIHA and channel domains) with additional RyR and SPRY domains. If this is the case, then RyR channels have only been detected in metazoa and choanoflagellates, two sister groups of opisthokonts, Figure. 1. Even among choanoflagellates, one representative, *Salpingoeca rossetta*, possesses an RyR homologue, whereas another, *Monosiga brevicolis*, lacks recognisable RyRs [12,16,18]. These observations suggest that RyRs evolved before the time that metazoans and choanoflagellates diverged.

In addition to being present within canonical RyR channel proteins, RyR domains are also found within diverse eukaryotic, bacterial, archaeal and viral proteins [12]. At least one additional family of RyR-domain containing eukaryotic proteins has the potential to act as CRCs: these are found in oomycetes and we have named them 'PKD-

RR' channels. Although these are hypothetical proteins, we have found that every oomycete genome contains at least one candidate PKD-RR homologue (Table S1) (Figures 1 and 3) and experimental evidence indicates that mRNA encoding one of these proteins is expressed during zoosporogenesis in *Phytophthora infestans*, the causative organism of late potato blight [55]. Strong homology between oomycete PKD-RR channels and choanoflagellate/metazoan RyRs is restricted to their RyR domains. In order to gain further insights into the structures of oomycete PKD-RR, the transmembrane topologies of representatives from *P.infestans* (XP_002909214), *S. declina* (XP_008611403.1) and *Albugo laibachii* (CCA15061.1) were predicted using TopPred and TMPRED programs. Although there were discrepancies in the number of predicted transmembrane segments between species and between algorithms (between 6 and 9), the consensus was a model with 7 transmembrane stretches, figure 4. Six of these transmembrane segments reside within a predicted polycystic kidney disease (PKD) channel domain. PKD channel proteins are distinct from the RyR/ITPR superfamily and belong to the transient receptor potential (TRP) superfamily of cation channels, which characteristically possess 6 transmembrane spanning regions and reside in the cell surface membrane [56]. The additional, N-terminally located predicted transmembrane segment of oomycete PKD-RR channels might be required to allow an appropriate transmembrane orientation within endomembrane systems, as opposed to the cell surface membrane.

To gain further insights into their structural properties, oomycete PKD-RR channels were modelled using the PHYRE^2 server. This can permit detection of tertiary structural features of proteins, unveiling relationships between proteins that are not readily apparent on the basis of identity shared between primary structures [30]. Analyses of full-length oomycete PKD-RR proteins failed to generate any high-confidence models. For example, the best model of full-length *P.infestans* PKD-RR was of only 35.7% confidence. As an alternative, models were generated from the three protein domains predicted in *P.infestans* PKD-RR: the PKD channel domain and the two RyR domains. Although the PKD domain did not generate any high-confidence models, both RyR domains ('RyRd1' and 'RyRd2') generated several matches with 100% confidence, despite sharing limited amino acid identity with the matching protein structures (between 26% and 45%). The highest ranking models were both based on the structure of an RyR domain protein from the bacterium *Bacteroides thetaiotaomicron* ('BthetaR', NP_811160.1) whose high-resolution teriary structure has been deduced by X-ray crystallography, Figure 4. High ranking matches (all 100% confidence) were also found with structures of RyR domains from the RyR1, RyR2 and RyR3 channels from several mammalian species. These predictions support the concept that oomycete PKD-RR channels contain genuine RyR domains.

In order to generate an updated survey of RyR domain-containing proteins throughout the tree of life, homologous domains were detected by BLAST searches using the B. thetaiotaomicron protein NP_811160.1 as the query sequence. As reported previously, RyR proteins domains were detected in archaea, bacteria, eukaryotes and double-stranded DNA bacteriophage [12], (Table S2) (Figure 5). Among eukaryotes, RyR domains could not be detected in rhodophytes, glaucophytes, haptophytes, cryprophytes, viridiplantae, rhizaria, excavates and alveolates. A novel finding in the current study is that among fungi, RyR domains were only detected in members of the family Orbiliaceae: *Dactylellina haptotyla*, *Drechslerella stenobrocha* and *Arthrobotrys oligospora*. These are nematode-trapping fungi, so it is speculated that RyR domains might have been transferred from their prey by HGT, or may have been retained from an ancestral eukaryote if they are essential for this predatory life-style. Although these viral, fungal and bacterial RyR domains show relatively low amino acid identity with domains from mammalian RyR channels, PHYRE2 modelling indicates that they are extremely likely to adopt highly similar tertiary structures (Figure 6).

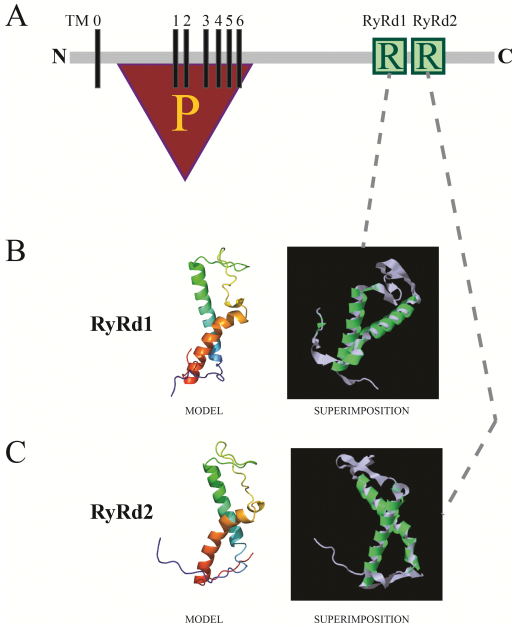


Figure 4: Predictions on the structure of a PKD-RR channel from the Oomycete Phytophthora infestans. A. Consensus model of the transmembrane topology of CRC1 from Phytophthora infestans, a PKD-RR protein. Black vertical bars ('0' to '6') indicate the positions of predicted transmembrane segments. B. Modelling of the tertiary structure of the first RyR domain from this PKD-RR protein, RyRd1, using PHYRE2 software. Approximate model dimensions are X:36 Y:55 Z:36 Å. The right-hand panel represents the model (green) superimposed on the crystal structure of a putative RyR from B. thetaiotaomicron (grey, c3nrtC, TM=1, confidence=100%, id=45%). C. The predicted tertiary structure of RyRd2 from this PKD-RR (X:44 Y:51 Z:29 Å). The right-hand panel represents the model (green) superimposed on the crystal structure of the putative RyR from B. thetaiotaomicron (grey, TM=1, confidence=100%, id=43%). Both models are coloured rainbow from amino to carboxy terminus.

The presence of RyR domains in a restricted range of eukaryotic taxa may have arisen by two mechanisms, which are not mutually exclusive. Firstly, RyR domains may have been lost from certain taxa, as we have suggested might have happened to full-length ITPR proteins in certain lineages, including green plants and fungi. Alternatively, RyR domains might have been incorporated into a range of proteins by HGT. Although the contribution of HGT to the evolution of RyR domain-containing proteins has been dismissed by other authors [20], two lines of evidence to support this possibility, at least among bacterial homologues. Firstly, in the current phylogenetic analyses, a protein from the alkaliphilic gamma-proteobacterium Thioalkalivibrio sp. ARh3 forms a well-supported group (boot-strap

value of 81%) with RyR domain proteins from bacteriophages (Figure 5). This tree incongruency is suggestive of HGT. The presence of cellular calcium signalling components in viruses is not unprecedented: a functional calcium-transporting ATPase pump is encoded within the genomes of chlorella viruses and phylogenetic analysis demonstrated that these group with homologues from chlorophyte plants [57]. Secondly, both the HGT-DB and the DarkHorse HGT Candidate Resource indicate potential HGT in bacterial RyR domain encoding genes. For the HGT-DB, HGT is predicted for a putative RyR domain protein (GenBank:ABD09801.1) in the actinobacterium Frankia sp. CcI3. In the DarkHorse HGT database, an LPI score of less than 0.6 is considered supportive of HGT. Querying the archaeal and bacterial genomes within this resource (1456 genomes) with an RyR domain protein (NP_811160.1) from B. thetaiotamicron (Phylum: Bacteroidetes) revealed potential HGT in an RyR domain protein (WP_004854008.1) from Copprococcus eutactus (Phylum: Firmicutes; LPL score=0.468) and in a kinase domain containing protein (EEB78804.1) in the marine gamma proteobacterium HTCC2148 (LPL score=0.477).

Both phylogenetic analysis and HGT database information support the possibility of HGT of RyR domain-containing proteins, at least in bacteria. On the basis of this, we propose that proteins containing the RyR domain may have arisen by partial HGT of this domain into pre-existing PKD-like (in oomycetes) or ITPR-like (in the choanoflagellate and metazoan lineage) channels in ancestral eukaryotes. In keeping with this hypothesis, recent analyses suggest that about 4.4% of the genes present in the choanoflagellate Monosiga brevicolis have been derived by HGT from bacteria and plankton, prey organisms of this species [58].

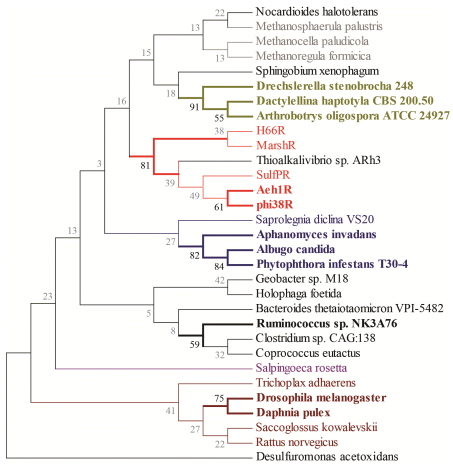
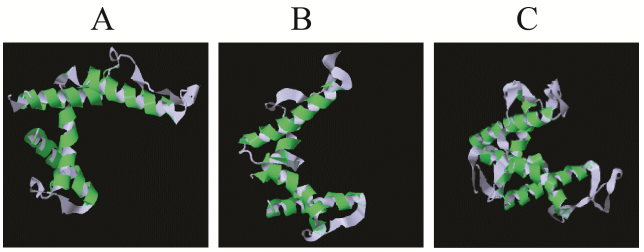


Figure 5: Phylogenetic Analysis of Ryanodine Domain Containing Proteins. Homologues of a putative ryanodine receptor from the bacteroidete B. thetaiotamicron were identified by BLAST searching of genomic databases (Supplemental Table 2), sequences were then aligned using MUSCLE and their phylogenetic relationships were investigated by Maximum Likelihood analyses. The consensus tree with the highest likelihood is shown (log likelihood of -4180.01). Black branches are bacteria; grey, archaea; blue, oomycetes; maroon, metazoans; purple, choanoflagellates; olive, fungi; red, viruses.



Thioalkalivibrio + MusRyR2 (TM = 1, 99.9% confidence, 21% id)

Aeromonas phage Aeh1 + MusRyR2 (TM = 1, 100% confidence, 26% id)

Dactylellina haptotyla + Rabbit RyR1 (TM = 1, 100% confidence, 21% id)

Figure 6: Structural Models of Ryanodine Receptor Domains from Bacteria, Bacteriophage and Fungi. Predicted tertiary structures of ryanodine domains from A. the gamma-proteobacterium Thioalkalivibrio sp. ARh3 (dimensions: X:39 Y:52 Z:30 Å); B. the virus Aeromonas phage Aeh1 (X:45 Y:49 Z:33 Å) and C. the fungus Dactyllina haptotyla (X:60 Y:50 Z:36 Å). Models (green) are superimposed on the best ranking template (grey).

Acknowledgement

This project was generously supported by funding from a private source, administered by Cork University Foundation (CUF: R15645). JJM is grateful to Dr. Marco Mariotti, Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG), Barcelona, Spain, for helpful discussions.

References

1. Berridge MJ, Lipp P, Bootman MD (2000) The versatility and universality of calcium signalling. Nat Rev Mol Cell Biol 1: 11-21.
2. Clapham DE1 (2007) Calcium signaling. Cell 131: 1047-1058.
3. Dominguez DC1 (2004) Calcium signalling in bacteria. Mol Microbiol 54: 291-297.
4. Vancek M, Vidová M, Majerník AI, Smigán P (2006) Methanogenesis is Ca2+ dependent in Methanothermobacter thermautotrophicus strain DeltaH. FEMS Microbiol Lett 258: 269-273.
5. Plattner H, Verkhratsky A2 (2015) The ancient roots of calcium signalling evolutionary tree. Cell Calcium 57: 123-132.
6. Verkhratsky A, Parpura V2 (2014) Calcium signalling and calcium channels: evolution and general principles. Eur J Pharmacol 739: 1-3.
7. Elias M1 (2010) Patterns and processes in the evolution of the eukaryotic endomembrane system. Mol Membr Biol 27: 469-489.
8. Bothwell JH, Ng CK (2005) The evolution of Ca2+ signalling in photosynthetic eukaryotes. New Phytol 166: 21-38.
9. Sorrentino V, Barone V, Rossi D (2000) Intracellular Ca(2+) release channels in evolution. Curr Opin Genet Dev 10: 662-667.
10. Mackrill JJ1 (2010) Ryanodine receptor calcium channels and their partners as drug targets. Biochem Pharmacol 79: 1535-1543.
11. Zorzato F, Fujii J, Otsu K, Phillips M, Green NM, et al. (1990) Molecular cloning of cDNA encoding human and rabbit forms of the Ca2+ release channel (ryanodine receptor) of skeletal muscle sarcoplasmic reticulum. J Biol Chem 265: 2244-2256.
12. Mackrill JJ1 (2012) Ryanodine receptor calcium release channels: an evolutionary perspective. Adv Exp Med Biol 740: 159-182.
13. Efremov RG, Leitner A2, Aebersold R3, Raunser S4 (2015) Architecture and conformational switch mechanism of the ryanodine receptor. Nature 517: 39-43.

14. Yan Z, Bai XC2, Yan C3, Wu J3, Li Z4, et al. (2015) Structure of the rabbit ryanodine receptor RyR1 at near-atomic resolution. Nature 517: 50-55.
15. Zalk R, Clarke OB2, des Georges A2, Grassucci RA3, Reiken S, et al. (2015) Structure of a mammalian ryanodine receptor. Nature 517: 44-49.
16. Cai X, Wang X2, Clapham DE3 (2014) Early evolution of the eukaryotic Ca2+ signaling machinery: conservation of the CatSper channel complex. Mol Biol Evol 31: 2735-2740.
17. Plattner H, Verkhratsky A (2013) Ca2+ signalling early in evolution--all but primitive. J Cell Sci 126: 2141-2150.
18. Cai X, Clapham DE (2012) Ancestral Ca2+ signaling machinery in early animal and fungal evolution. Mol Biol Evol 29: 91-100.
19. Peterson KJ, Butterfield NJ (2005) Origin of the Eumetazoa: testing ecological predictions of molecular clocks against the Proterozoic fossil record. Proc Natl Acad Sci U S A 102: 9547-9552.
20. Plattner H1 (2015) Molecular aspects of calcium signalling at the crossroads of unikont and bikont eukaryote evolution - The ciliated protozoan Paramecium in focus. Cell Calcium 57: 174-185.
21. Lusche DF, Wessels D, Scherer A, Daniels K, Kuhl S, et al. (2012) The IplA Ca2+ channel of Dictyostelium discoideum is necessary for chemotaxis mediated through Ca2+, but not through cAMP, and has a fundamental role in natural aggregation. J Cell Sci 125: 1770-1783.
22. Huang G, Bartlett PJ, Thomas AP, Moreno SN, Docampo R (2013) Acidocalcisomes of Trypanosoma brucei have an inositol ,4,5-trisphosphate receptor that is required for growth and infectivity. Proc Natl Acad Sci U S A 110: 1887-1892.
23. Krinke O, Novotná Z, Valentová O, Martinec J (2007) Inositol trisphosphate receptor in higher plants: is it real? J Exp Bot 58: 361-376.
24. Allen GJ, Muir SR, Sanders D (1995) Release of Ca2+ from individual plant vacuoles by both InsP3 and cyclic ADP-ribose. Science 268: 735-737.
25. Bouillet LE, Cardoso AS, Perovano E, Pereira RR, Ribeiro EM, et al. (2012) The involvement of calcium carriers and of the vacuole in the glucose-induced calcium signaling and activation of the plasma membrane H(+)-ATPase in Saccharomyces cerevisiae cells. Cell Calcium 51: 72-81.
26. Michell RH1 (2007) Evolution of the diverse biological roles of inositols. Biochem Soc Symp : 223-246.
27. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.
28. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, et al. (2005) Protein database searches using compositionally adjusted substitution matrices. FEBS J 272: 5101-5109.
29. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, et al. (2015) CDD: NCBI's conserved domain database. Nucleic Acids Res 43: D222-D226.
30. Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 4: 363-371.
31. Claros MG, von Heijne G (1994) TopPred II: an improved software for membrane protein structure predictions. Comput Appl Biosci 10: 685-686.
32. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-2739.
33. Edgar RC1 (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792-1797.
34. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8: 275-282.
35. Felsenstein J (1985) Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. Evolution 39: 783-791.
36. Garcia-Vallve S, Guzman E, Montero MA, Romeu A (2003) HGT-DB: a database of putative horizontally transferred genes in prokaryotic complete genomes. Nucleic Acids Res 31: 187-189.
37. Podell S, Gaasterland T, Allen EE (2008) A database of phylogenetically atypical genes in archaeal and bacterial genomes, identified using the DarkHorse algorithm. BMC Bioinformatics 9: 419.
38. Burki F1 (2014) The eukaryotic tree of life from a global phylogenomic perspective. Cold Spring Harb Perspect Biol 6: a016147.
39. Prole DL, Taylor CW (2011) Identification of intracellular and plasma membrane calcium channel homologues in pathogenic parasites. PLoS One 6: e26218.
40. Cai X, Wang X2, Patel S3, Clapham DE4 (2015) Insights into the early evolution of animal calcium signaling machinery: A unicellular point of view. Cell Calcium 57: 166-173.
41. Eme L, Sharpe SC, Brown MW, Roger AJ1 (2014) On the age of eukaryotes: evaluating evidence from fossils and molecular clocks. Cold Spring Harb Perspect Biol 6.
42. Tae H, Casarotto MG, Dulhunty AF (2009) Ubiquitous SPRY domains and their role in the skeletal type ryanodine receptor. Eur Biophys J 39: 51-59.
43. Xiong H, Feng X, Gao L, Xu L, Pasek DA, et al. (1998) Identification of a two EF-hand Ca2+ binding domain in lobster skeletal muscle ryanodine receptor/Ca2+ release channel. Biochemistry 37: 4804-4814.
44. Rasteiro R, Pereira-Leal JB (2007) Multiple domain insertions and losses in the evolution of the Rab prenylation complex. BMC Evol Biol 7: 140.
45. Ladenburger EM, Sehring IM, Korn I, Plattner H (2009) Novel types of Ca2+ release channels participate in the secretory cycle of Paramecium cells. Mol Cell Biol 29: 3605-3622.
46. Ladenburger EM, Plattner H (2011) Calcium-release channels in paramecium. Genomic expansion, differential positioning and partial transcriptional elimination. PLoS One 6: e27111.
47. Franck JP, Morrissette J, Keen JE, Londraville RL, Beamsley M, et al. (1998) Cloning and characterization of fiber type-specific ryanodine receptor isoforms in skeletal muscles of fish. Am J Physiol 275: C401-415.
48. McKay PB, Griswold CK (2014) A comparative study indicates both positive and purifying selection within ryanodine receptor (RyR) genes, as well as correlated evolution. J Exp Zool A Ecol Genet Physiol 321: 151-163.
49. Ito ET, Cihlar RL, Inderlied CB (1982) Lipid synthesis during morphogenesis of Mucor racemosus. J Bacteriol 152: 880-887.
50. Kuin H, Koerten H, Ghijsen WE, Munnik T, van den Ende H, et al. (2000) Chlamydomonas contains calcium stores that are mobilized when phospholipase C is activated. Planta 210: 286-294.
51. Penny D, Collins LJ, Daly TK, Cox SJ (2014) The relative ages of eukaryotes and akaryotes. J Mol Evol 79: 228-239.
52. Denis V, Cyert MS (2002) Internal Ca(2+) release in yeast is triggered by hypertonic shock and mediated by a TRP channel homologue. J Cell Biol 156: 29-34.
53. Yuan F, Yang H2, Xue Y2, Kong D3, Ye R3, et al. (2014) OSCA1 mediates osmotic-stress-evoked Ca2+ increases vital for osmosensing in Arabidopsis. Nature 514: 367-371.
54. Harish A, Tunlid A, Kurland CG (2013) Rooted phylogeny of the three superkingdoms. Biochimie 95: 1593-1604.
55. Tani S, Yatzkan E, Judelson HS (2004) Multiple pathways regulate the induction of genes during zoosporogenesis in Phytophthora infestans. Mol Plant Microbe Interact 17: 330-337.
56. Cantiello HF1 (2004) Regulation of calcium signaling by polycystin-2. Am J Physiol Renal Physiol 286: F1012-1029.
57. Bonza MC, Martin H, Kang M, Lewis G, Greiner T, et al. (2010) A functional calcium-transporting ATPase encoded by chlorella viruses. J Gen virol 91: 2620-2629.
58. Yue J, Sun G, Hu X, Huang J1 (2013) The scale and evolutionary significance of horizontal gene transfer in the choanoflagellate Monosiga brevicollis. BMC Genomics 14: 729.