





Rel/NF- κ B Transcription Factors Emerged at the Onset of Opisthokonts

Michelle M. Leger ^{1,*†}, Núria Ros-Rocher ^{1,†}, Sebastián R. Najle ^{1,4}, and Iñaki Ruiz-Trillo ^{1,2,3}

¹Institute of Evolutionary Biology (Consejo Superior de Investigaciones Científicas-Universitat Pompeu Fabra), Barcelona, Catalonia, Spain

²Department of Genetics, Microbiology and Statistics, Institute for Research on Biodiversity, University of Barcelona, Catalonia, Spain

³Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Catalonia, Spain

⁴Present address: Center for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain

[†]These authors contributed equally to this work.

*Corresponding author: E-mail: michelle.leger@ibe.upf-csic.es.

Accepted: 27 December 2021

Abstract

The Rel/NF- κ B transcription factor family has myriad roles in immunity, development, and differentiation in animals, and was considered a key innovation for animal multicellularity. Rel homology domain-containing proteins were previously hypothesized to have originated in a last common ancestor of animals and some of their closest unicellular relatives. However, key taxa were missing from previous analyses, necessitating a systematic investigation into the distribution and evolution of these proteins. Here, we address this knowledge gap by surveying taxonomically broad data from eukaryotes, with a special emphasis on lineages closely related to animals. We report an earlier origin for Rel/NF- κ B proteins than previously described, in the last common ancestor of animals and fungi, and show that even in the sister group to fungi, these proteins contain elements that in animals are necessary for the subcellular regulation of Rel/NF- κ B.

Key words: NF- κ B, Rel homology domain, transcription factors, opisthokonts, NF-kappa B.

Significance

The Rel/NF- κ B transcription factor family plays a fundamental role in animal innate immunity, programmed cell death, intercellular signaling, and transcriptional regulation. Here, we show that proteins sharing the characteristic features of animal Rel/NF- κ B transcription factors originate much earlier than previously described: not in a relatively recent common ancestor of animals and some of their unicellular relatives, but prior to the divergence of animals and fungi. Intriguingly, we show that even the earliest-diverging nonmetazoan Rel/NF- κ B-like protein has sequence features consistent with an animal-like mode of regulation.

The Rel homology region (RHR) is an evolutionarily conserved N-terminal DNA-binding region present in two major paralogous families of animal transcription factors with crucial roles in immunity and development: the Rel/Nuclear Factor- κ B (NF- κ B) and the Nuclear Factor of Activated T-cells (NFAT) families. Members of the Rel/NF- κ B family were originally described in the late 1980s as oncogenes (Gilmore and Temin 1986) and immunoglobulin κ light chain enhancer-binding proteins in vertebrates (Sen and Baltimore 1986). Members of this family have since been implicated in a wide range of cellular

processes in animals (Ghosh and Hayden 2012), including innate and adaptive immunity (Hayden and Ghosh 2011), cell cycle regulation (Ledoux and Perkins 2014), apoptosis (Kucharczak et al. 2003), autophagy (Salminen et al. 2012), and regulation of oxidative stress responses (Lingappan 2018). Despite these crucial roles in animals, their functions and domain architectures in other taxa remain to be fully explored. Pinpointing the origins and taxonomic distribution of Rel/NF- κ B proteins is a fundamental first step to tackle these questions and understand the evolutionary history of

© The Author(s) 2022. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Rel/NF- κ B proteins. Rel/NF- κ B proteins are characterized by the presence of an RHR, consisting of an N-terminal DNA-binding and dimerization domain that facilitates homo- or heterodimerization, and a short stretch of arginine and/or lysine residues forming a nuclear localization signal (NLS) that mediates the translocation of these proteins into the nucleus (fig. 1A) (reviewed in Napetschnig and Wu [2013] and Williams and Gilmore [2020]).

The Rel/NF- κ B family can be further subdivided into two classes according to the transactivation potential of its members. The first, collectively termed Rel proteins, include the vertebrate RelA (p65), RelB, and c-Rel and their orthologs (fig. 1A). Besides the RHR, Rel proteins also contain a poorly conserved C-terminal transactivation domain (TAD) that is acidic, and may be rich in proline, serine, glutamine, and/or hydrophobic residues (Bull et al. 1990; Blair et al. 1994; Gross et al. 1999), that allows them to activate target gene expression. RelB proteins additionally possess an N-terminal leucine zipper domain critical for their activity (Dobrzanski et al. 1993). Rel proteins can homodimerize, or heterodimerize with other Rel/NF- κ B family members lacking a classical TAD (reviewed in Napetschnig and Wu [2013]).

The second class includes the vertebrate p50 and p52 and their orthologs (fig. 1A). p50 and p52 are synthesized as larger precursors termed NF- κ B1/p105 and NF- κ B2/p100, respectively. These precursors include an RHR followed by a glycine-rich region (GRR) and a variable number of C-terminal ankyrin repeats. In their inactive state, ankyrin repeats inhibit nuclear localization and transcriptional activity and keep the NF- κ B dimers sequestered in the cytosol. They additionally possess a death domain that mediates interaction with other death domain-containing signaling proteins (Hayden and Ghosh 2008). The p105 and p100 precursors are classified as inhibitors of NF- κ B ($I\kappa$ Bs). This category also includes a family of separate C-terminal ankyrin repeat-containing proteins that carry out the same inhibitory function for Rel proteins, by sequestering them in the cytosol (Kanarek et al. 2010) (fig. 1A). Upon upstream activation, C-terminal serine residues in ankyrin-repeat-rich regions of $I\kappa$ Bs are phosphorylated by an $I\kappa$ B kinase complex (IKK) (Karin 1999). The IKK complex is also involved in the processing of the p105 and p100 precursors, leading to the ubiquitination and proteasomal degradation of their C-terminal regions (fig. 1A). In this case, the proteasome falls off at the GRR located between the RHR and the C-terminal ankyrin repeats (Lin and Ghosh 1996; Williams and Gilmore 2020). This process releases an intact N-terminal part of the NF- κ B protein, including the GRR (Moorthy et al. 2006), and leads to its nuclear translocation for gene expression activation (reviewed in Napetschnig and Wu [2013]). A key scaffolding component of the IKK complex, the IKK γ /NF- κ B essential modulator (NEMO), is also required for IKK recruitment and NF- κ B activation (fig. 1A) (reviewed in Napetschnig and Wu [2013]).

The NFAT family constitutes the paralogous group of RHR-containing proteins. NFAT proteins were first described almost three decades ago as calcium-dependent transcription factors implicated in T-cell activation (Shaw et al. 1988), cell proliferation, migration, and angiogenesis (reviewed in Müller and Rao [2010]). NFAT proteins contain a more centrally located RHR flanked by longer N- and C-terminal regions, and lack ankyrin repeats (fig. 1A). The NFAT NLS is contained not within the RHR, but within the N-terminal regulatory region known as the NFAT homology region (NHR) (fig. 1A). This region also contains calcineurin-binding sites required for nuclear translocation (Park et al. 2000). NFAT1-4 also possess N-terminal TADs (Serfling et al. 2004). NFAT5, the only noncalcium regulated NFAT protein in humans, lacks an NHR (fig. 1A) but is generally located in the nucleus, and plays a role in osmotic stress response and immune cell development (Lee et al. 2019).

Animals and their closest unicellular relatives together comprise the eukaryotic group Holozoa; the larger eukaryotic group comprising Holozoa, fungi, and their closest unicellular relatives, is known as Opisthokonta (fig. 1B). Rel homology domain-containing proteins were initially believed to be specific to animals (Metazoa), in which these proteins have been extensively studied. However, they were later reported in two lineages of unicellular holozoans: the filastereans (Mikhailov et al. 2009; Seb e-Pedr s et al. 2011) and choanoflagellates (Richter et al. 2018; Williams and Gilmore 2020). These sequences branched sister to all metazoan NF- κ B and NFAT sequences, but contained RHRs, GRR, NLS, and/or ankyrin repeats more characteristic of the Rel/NF- κ B family (Seb e-Pedr s et al. 2011; Williams and Gilmore 2020). Rel homology DNA-binding domains had also been reported as being present in more distantly related taxa, including a nucleariid and possibly one or more ichthyosporeans and/or pluriformeans ([de Mendoza et al. 2013; de Mendoza and Seb e-Pedr s 2019]; fig. 1B and [supplementary table S1, Supplementary Material online](#)). However, these studies focused solely on the Rel homology DNA-binding domain, and obtained conflicting results that cast doubt on the domain's true taxonomic distribution (de Mendoza et al. 2013; Richter et al. 2018; de Mendoza and Seb e-Pedr s 2019). As a result, the precise origin, early evolution, and molecular context of Rel homology domain-containing transcription factors remained unknown.

To resolve these questions, we performed a taxonomically broad survey of genomic and transcriptomic sequence data representing all major eukaryotic supergroups, including newly sequenced protistan lineages within Holozoa (Grau-Bov e et al. 2017; Hehenberger et al. 2017; Tikhonenkov et al. 2020; Urrutia et al. 2021). We surveyed data from 180 species for key Rel homology DNA-binding and dimerization domain-containing proteins, examined their domain architecture, and constructed phylogenies of the proteins identified. We additionally extended the search to homologs of IKK components.



FIG. 1.—Rel/NF- κ B proteins emerged at the onset of Opisthokonta (A) Domain architecture representation of members of the Rel/NF- κ B, and NFAT protein families, I κ B, and the IKK complex. Details of the represented features are shown in the domains and regions key. The Rel Homology Region, characteristic of Rel/NF- κ B and NFAT proteins, contains conserved Rel homology DNA-binding (RHD DB) and dimerization domains (Dim) and, in the case of Rel/NF- κ B proteins, an NLS (orange bar). Rel proteins also contain a C-terminal, serine-rich Transactivation Domain (Ct TAD) or a RelB TAD; and RelB proteins additionally possess an N-terminal leucine zipper domain (RelB LZ). The NF- κ B1 precursor (p105) and NF- κ B2 precursor (p100) contain a more centrally located GRR and C-terminal Death domain (Death). These precursors share with I κ B proteins C-terminal ankyrin repeats (light green bars). Other domains present in I κ B families include proline-, glutamic acid-, serine-, and threonine-rich regions (PEST). Key domains specific to NFAT proteins, include an N-terminal TAD (Nt TAD) inside an NHR, an NLS (orange bar) and C-terminal TAD (Ct TAD). PxlIT and LxVP Calcineurin-binding motifs in the NHR are depicted with a magenta and a green bar, respectively. Domains for IKK α and β include kinase domains (Kinase); ubiquitin-like domains (ULD); ubiquitin-binding domains (UBD); IQBAL scaffold dimerization domain (SDD); NEMO-binding domains (NBD). Domains for IKK γ /NF- κ B essential modulator (NEMO) include NEMO domain (NEMO); UBD; leucine-zipper domains (LZ); and zinc finger domains (ZF). (B) Presence or absence of key Pfam domains analyzed in this study are represented in columns and color-coded according to genome or transcriptome data source (indicated in the Domain presence/absence key). The phylogenetic relationships of selected taxa are based on several recent phylogenomic studies (Torruella et al. 2015; Grau-Bové et al. 2017; Hehenberger et al. 2017; Tikhonenkov et al. 2020; Urrutia et al. 2021). Taxa are color-coded according to the Taxonomic Group key.

We identified RHRs in all but one of the holozoan groups examined, including the filastereans *Pigoraptor* spp., the recently sequenced *Txikispora philomaios* and *Tunicaraptor unikontum*, and several ichthyosporean species (fig. 1B-2; supplementary fig. S1, tables S1 and S2, Supplementary Material online). A candidate protein was also present in *Parvularia atlantis* (formerly referred to as *Nuclearia* sp. ATCC 50694 [López-Escardó et al. 2018]), a member of the sister group to Fungi (Nucleariida), suggesting that Rel homology proteins were present in the last common ancestor of Opisthokonta, and were secondarily lost in Fungi (de Mendoza et al. 2013; de Mendoza and Sebé-Pedrós 2019). Strikingly, at least one sequence from each of these lineages included both DNA-binding and dimerization domains (fig. 1B-3; supplementary fig. S1 and table S2, Supplementary Material online), and C-terminal ankyrin repeat-rich regions preceded by a GRR were found in at least one sequence from Filasterea, from Ichthyosporea, and from *P. atlantis* (fig. 2; supplementary table S2, Supplementary Material online). The finding of these traits in the nucleariid *P. atlantis*, in particular, not only confirms the origin of Rel/NF- κ B proteins in the opisthokont stem lineage, but also is consistent with an ancient conserved mechanism of cytosolic sequestration for these proteins. Similarly, nonmetazoan Rel/NF- κ B-like sequences from all of these groups share with their animal homologs key domains for DNA-binding specificity, including a highly conserved specific recognition loop (RL) within the RHR, as well as a dimerization domain and a highly conserved monopartite NLS (figs. 3 and 4; supplementary table S2, Supplementary Material online). In contrast, the linker region between the DNA-binding and dimerization domains (Ghosh et al. 1995; Müller 1995) appears to be animal-specific (fig. 3).

Metazoan Rel/NF- κ B proteins and metazoan NFAT proteins each formed a well-supported clade, within a larger well-supported clade of metazoan sequences. Sequences from choanoflagellates, filastereans, and ichthyosporeans formed a clade sister to all metazoan sequences (fig. 2). The most parsimonious explanation suggested by the phylogeny is a duplication of ancestral Rel/NF- κ B proteins in the metazoan stem lineage followed by loss of the GRR and ankyrin repeats from NFAT proteins (fig. 4; Sebé-Pedrós et al. [2011] and Gilmore and Wolenski [2012]), and additional duplications in individual metazoan and choanoflagellate lineages (Williams and Gilmore 2020). None of the nonmetazoan Rel/NF- κ B sequences contained death domains. The animal NFAT sequences recovered in our survey contain the calcineurin-binding motifs LxVP, and, in the case of chordates, PxlIT (Wigington et al. 2020) near the N-terminus. Seven of the nonmetazoan Rel/NF- κ B-like proteins contained a LxVP motif. In the choanoflagellates *Codosiga hollandica* and *Savillea parva* (two out of three proteins), and in the filasterean *Pigoraptor vietnamica*, this motif was found near the N-terminus, before the RHR (fig. 2; supplementary table S2,

Supplementary Material online). This raises the possibility that calcineurin may be an additional regulator of some nonmetazoan Rel/NF- κ B-like proteins, as it is in NFAT proteins.

Despite the presence of C-terminal ankyrin-rich repeats preceded by a GRR in some sequences from nonmetazoan opisthokonts, and consistent with earlier reports (Williams and Gilmore 2020), we were unable to retrieve any apparent orthologs of NEMO (fig. 1B; supplementary fig. S1, Supplementary Material online) or the other IKK subunits (IKK α , IKK β , and IKK ϵ) outside Metazoa (data not shown). If the C-terminal region of nonmetazoan Rel/NF- κ B-like proteins is processed, it may be phosphorylated by another kinase, or targeted for degradation by a different mechanism.

The diversity of nonmetazoan Rel/NF- κ B-like proteins likely reflect the variety of lifestyles of the organisms in which they are found. Choanoflagellates are free-living, mostly marine or freshwater bacterivores, some of which form clonal multicellular structures in response to specific bacterial molecules (Alegado et al. 2012; Leadbeater 2015); filastereans include both free-living freshwater bacterivores and endobiotic species (Stibbs et al. 1979; Tong 1997; Hehenberger et al. 2017; Tikhonenkov et al. 2020; Urrutia et al. 2021), at least some of which can form multicellular aggregates (Sebé-Pedrós et al. 2013; Hehenberger et al. 2017; Mylnikov et al. 2019); ichthyosporeans include free-living species and parasites of invertebrates or fish, with diverse life cycles and cell states including multinucleate coenocytic stages (reviewed in Mendoza et al. [2002]); and *Parvularia* is a free-living freshwater bacterivorous amoeboid (López-Escardó et al. 2018). Rel/NF- κ B-like proteins may play similar or very different roles in how these organisms interact with a variety of prey or host organisms, and/or environmental factors. Interactions with newly evolved partners and gene duplications may have been key to increasing their combinatorial regulatory capabilities in different lineages, including along the animal stem.

Overall, we provide an updated evolutionary reconstruction of Rel/NF- κ B and NFAT transcription factor families, based on a broad taxon sampling including representatives of all major eukaryotic lineages. We show that Rel/NF- κ B-like proteins emerged earlier than previously known, prior to the split between animals and fungi. We further highlight conserved, animal-like architecture in these proteins from diverse opisthokonts. Together, our results suggest that localization and regulatory mechanisms found in animal Rel/NF- κ B proteins were likewise present in the last common ancestor of animals and fungi.

Materials and Methods

Raw Hidden Markov Models (HMMs) of Rel homology DNA-binding domain (RHD_DNA_bind v.21, PF00554), Rel homology Dimerization domain (RHD_dimer v.4, PF16179), Death domain (Death v.21, PF00531), and the IKK component domains Inhibitor of Kappa B Kinase Beta NEMO-binding

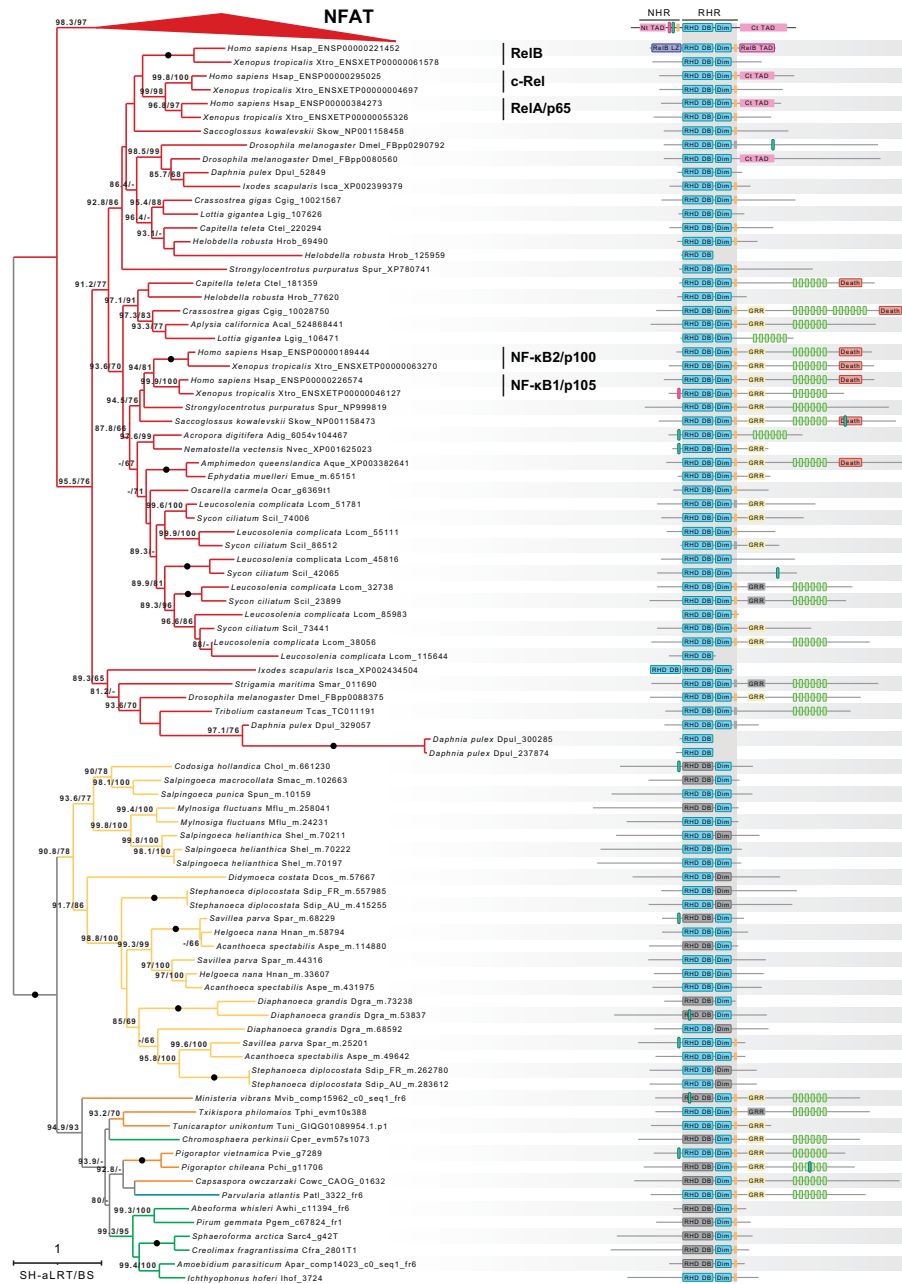


Fig. 2.—Rel homology domain phylogeny and domain architecture of Rel homology domain-containing proteins in NF- κ B and NFAT families. Only SH-aLRT and nonparametric bootstrap support values above 80 and 60, respectively, are shown. Fully supported bipartitions are indicated with filled circles. Sequences are color-coded according to taxonomic group: red, Metazoa; yellow, Choanoflagellata; orange, Filasterea; green, Ichthyosporia; and blue, Holomycota (Fungi and their closest relatives). The two side lengths of the collapsed NFAT clade are proportional to the distances between the node and its closest and furthest leaves. Schematic representation of Pfam domains and conserved regions related to Rel/NF- κ B and NFAT proteins in the RHR are depicted to the right of each gene identifier; colors and abbreviations as in figure 1A. Domains or regions depicted in gray match domains or regions that were identified in previous Pfam versions or that share high sequence similarity in the protein alignment (at least four Lysine or Arginine residues in the case of NLS).

domain (IKKbetaNEMObind v.7, PF12179) and NF-Kappa B Essential Modulator (NEMO v.7, PF11577) were retrieved from Pfam v.34.0 (Mistry et al. 2021), and used as queries in hmmscan (hmmer 3.1b2-2; Eddy 1998; Söding 2005) searches against a panukaryotic predicted proteome database enriched in holozoan representatives (supplementary

table S1, Supplementary Material online). BLAST searches for IKK complex components were carried out using *Homo sapiens* (GenBank accession numbers O15111.2 [Inhibitor of nuclear factor kappa-B kinase subunit alpha], O14920.1 [Inhibitor of nuclear factor kappa-B kinase subunit beta], Q9Y6K9.2 [NF-kappa-B essential modulator], Q14164.1

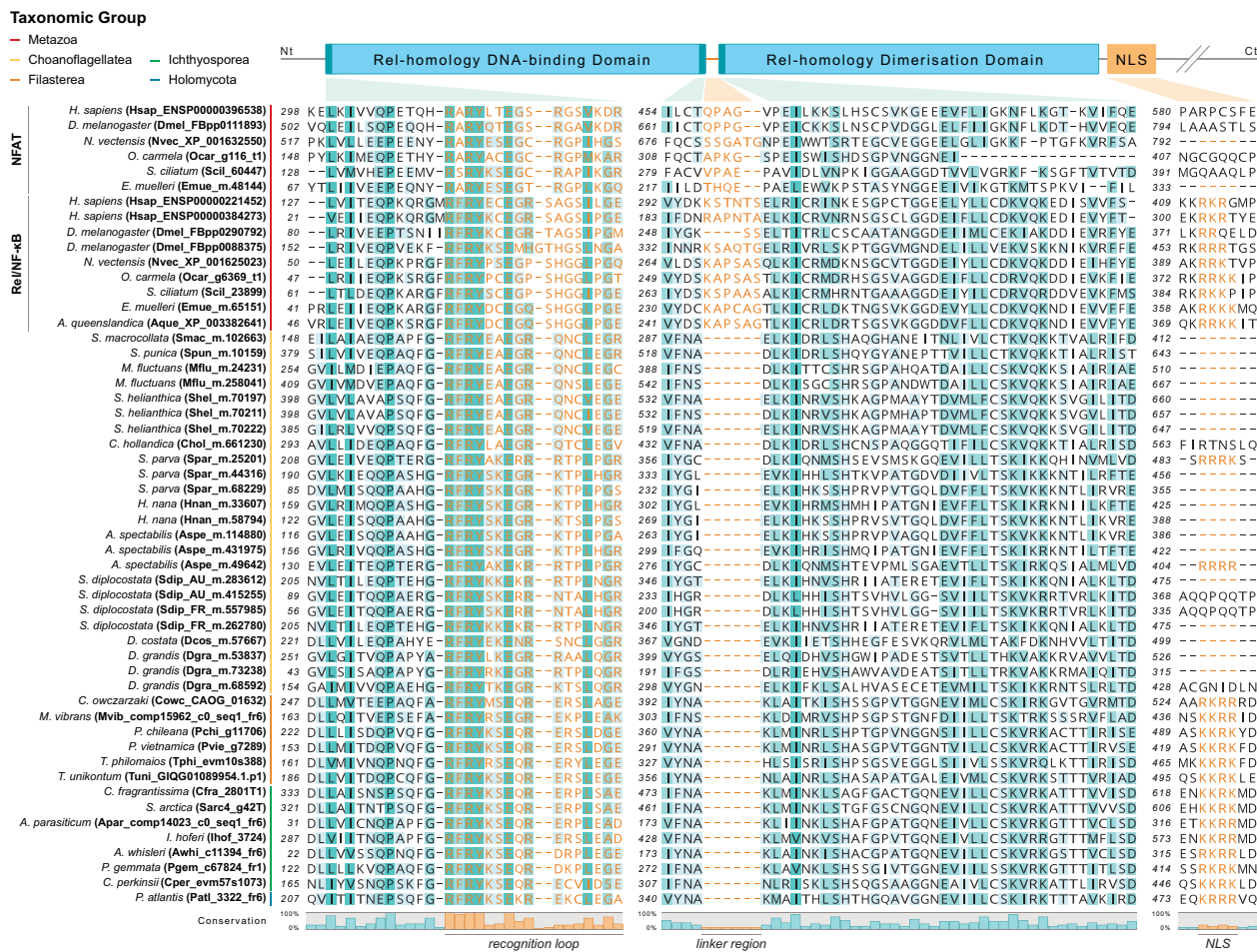


Fig. 3.—Partial multiple protein sequence alignment of key Rel/NF-κB domains in Opisthokonta. Partial protein sequence alignment (generated by MAFFT v7.299b E-INS-i multiple sequence alignment with the gap extension parameter set to 0) depicting the N- and C-terminal regions of the Rel homology DNA-binding domain, the N-terminal region of the Rel homology dimerization domain and the complete NLS. Background shading of individual amino acids reflects the degree of conservation at a given position; this degree of conservation is also depicted in the histogram shown below the alignment. Dashes indicate gaps in the alignment. Key DNA-binding amino acid residues in the recognition loop, residues comprising the complete linker region between the Rel homology DNA-binding domain and dimerization domain, and key residues in the NLS are highlighted in orange. The amino acid positions in the original sequences are shown to the left of each alignment series.

[Inhibitor of nuclear factor kappa-B kinase subunit epsilon] and *Nematostella vectensis* (ADQ57374.1 [single IKK-like protein]) IKK complex components as queries.

Using custom Perl scripts, the resulting output files were parsed and analyzed using PfamScan v.1.5 (Gish and States 1993), and all sequences containing a Rel homology DNA-binding domain were retrieved and examined using reciprocal best BlastP searches against the nonredundant protein database (nr) of the National Center for Biotechnology Information (NCBI). The domain architecture of all retrieved sequences was inferred with PfamScan using the gathering threshold as cutoff value. The number of ankyrin repeats was verified using InterProScan 5.26-65.0 (Jones et al. 2014).

Sequences were aligned using MAFFT v7.299b E-INS-i (Katoh et al. 2002, 2005; Katoh and Standley 2013) with the gap extension parameter set to 0, trimmed using BMGE

v.1.0 (Crisuolo and Gribaldo 2010) using the BLOSUM45 matrix, and alignments and trimming were verified by eye. Partial sequences with fewer than 50% of positions represented in the final trimmed alignment were excluded. Preliminary phylogenies were constructed using FastTree v. 2.1.9 (Price et al. 2009, 2010); the number of metazoan representatives was reduced. Final alignments were constructed using MAFFT E-INS-i with the gap extension parameter set to 0, and trimmed using trimAl v1.4.rev22 build[2015-05-21] (Capella-Gutiérrez et al. 2009), and final phylogenies were constructed using IQ-TREE multicore version 2.0-rc1 (Nguyen et al. 2015; Minh et al. 2020) with 1,000 ultrafast bootstrap resamplings (Minh et al. 2013; Hoang et al. 2018), using LG + F + R7 for Rel homology domain proteins (chosen by ModelFinder [Kalyaanamoorthy et al. 2017]) as the best fitting model according to the Bayesian

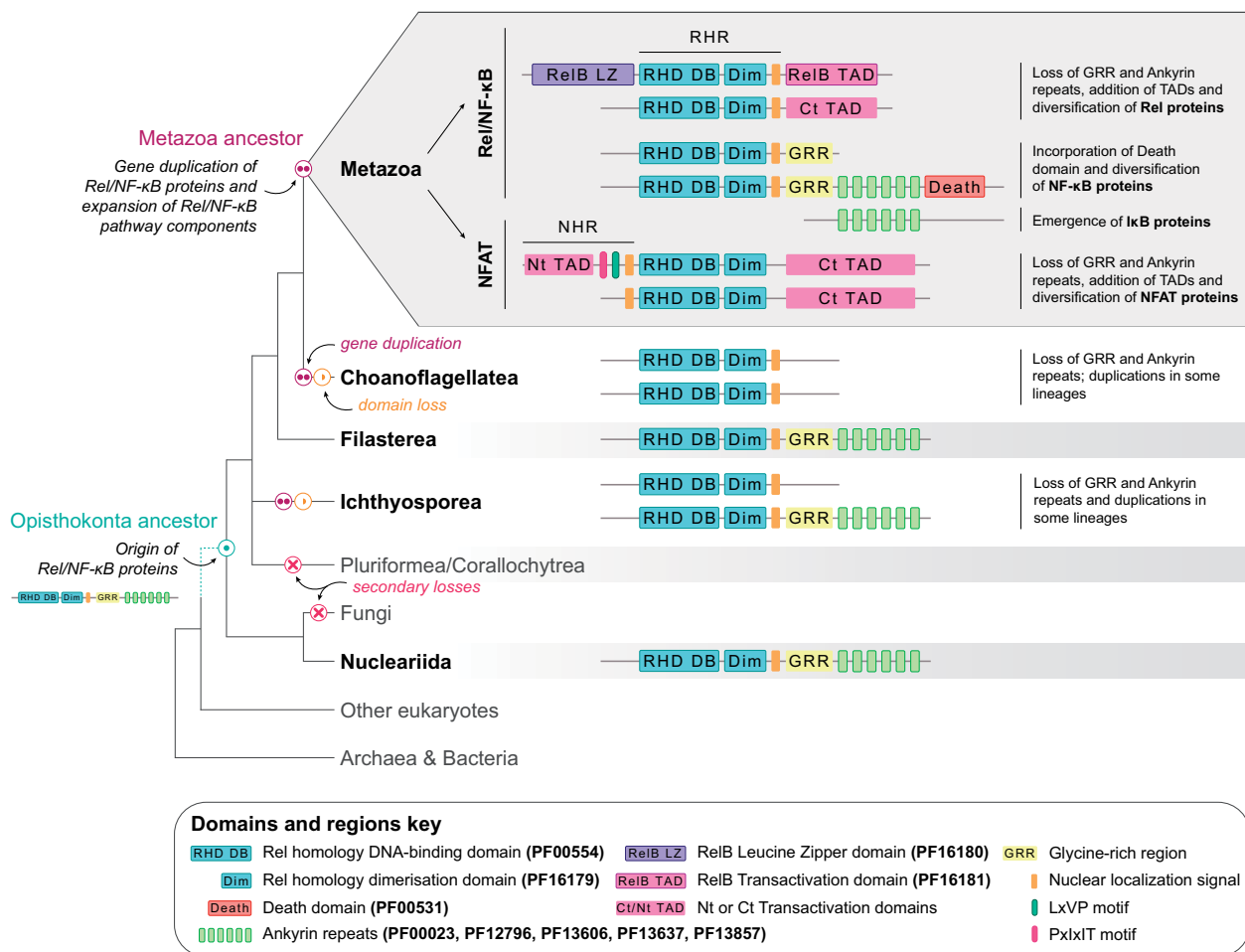


FIG. 4.—Evolution and diversification of Rel/NF- κ B proteins in Opisthokonta. The Rel/NF- κ B proteins originated along the Opisthokonta stem and diversified through gene duplication events and secondary losses and gains of key domains. Based on low sequence conservation of some regions and the lack of experimental information, we cannot rule out the possibility of TADs being present in other nonmetazoan species.

Information Criterion). Subsequently, 100 nonparametric bootstrap replicates were also performed under the same model.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

This work was supported by a European Research Council Consolidator Grant (ERC-2012-CO-616960) grant and a grant (BFU2017-90114-P) from Ministerio de Economía y Competitividad (MINECO), Agencia Estatal de Investigación (AEI), and Fondo Europeo de Desarrollo Regional (FEDER) to I.R.-T. M.M.L. was supported by a Marie Skłodowska-Curie Individual Fellowship under the EU Framework Programme for Research and Innovation Horizon 2020 (Project ID

747789) and an Ayuda Juan de la Cierva-Incorporación post-doctoral fellowship (IJC2018-036657-I) from the Spanish Ministry of Science and Innovation. N.R.-R. was supported by a “Formación del Profesorado Universitario (FPU13/01840)” PhD scholarship from the Spanish Ministerio de Educación, Cultura y Deporte (MECD). We thank the anonymous reviewers for their helpful and insightful feedback.

Data Availability

The data underlying this article are available in the supplementary materials (supplementary table S2 and dataset S1, Supplementary Material online); the data sources are listed in supplementary table S1, Supplementary Material online.

Literature Cited

Alegado RA, et al. 2012. A bacterial sulfonolipid triggers multicellular development in the closest living relatives of animals. *Elife* 1:e00013.

- Blair WS, Bogerd HP, Madore SJ, Cullen BR. 1994. Mutational analysis of the transcription activation domain of RelA: identification of a highly synergistic minimal acidic activation module. *Mol Cell Biol.* 14(11):7226–7234.
- Bull P, Morley KL, Hoekstra MF, Hunter T, Verma IM. 1990. The mouse c-rel protein has an N-terminal regulatory domain and a C-terminal transcriptional transactivation domain. *Mol Cell Biol.* 10(10):5473–5485.
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25(15):1972–1973.
- Crisuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol.* 10(1):210.
- de Mendoza A, et al. 2013. Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *Proc Natl Acad Sci U S A.* 110(50):E4858–E4866.
- de Mendoza A, Sebé-Pedrós A. 2019. Origin and evolution of eukaryotic transcription factors. *Curr Opin Genet Dev.* 58–59:25–32.
- Dobrzanski P, Ryseck RP, Bravo R. 1993. Both N- and C-terminal domains of RelB are required for full transactivation: role of the N-terminal leucine zipper-like motif. *Mol Cell Biol.* 13(3):1572–1582.
- Eddy SR. 1998. Profile hidden Markov models. *Bioinformatics* 14(9):755–763.
- Ghosh G, Duyn GV, Ghosh S. 1995. Structure of NF- κ B p50 homodimer bound to a KB site. *Nature* 373(6512):303–310.
- Ghosh S, Hayden M. 2012. Celebrating 25 years of NF- κ B Research. *Immunol Rev.* 246(1):5–13.
- Gilmore TD, Temin HM. 1986. Different localization of the product of the v-rel oncogene in chicken fibroblasts and spleen cells correlates with transformation by REV-T. *Cell* 44(5):791–800.
- Gilmore TD, Wolenski FS. 2012. NF- κ B: where did it come from and why? *Immunol Rev.* 246(1):14–35.
- Gish W, States DJ. 1993. Identification of protein coding regions by database similarity search. *Nat Genet.* 3(3):266–272.
- Grau-Bové X, et al. 2017. Dynamics of genomic innovation in the unicellular ancestry of animals. *eLife* 6:e26036.
- Gross I, Georgel P, Oertel-Buchheit P, Schnarr M, Reichhart J-M. 1999. Dorsal-B, a splice variant of the *Drosophila* factor Dorsal, is a novel Rel/NF- κ B transcriptional activator. *Gene* 228(1–2):233–242.
- Hayden MS, Ghosh S. 2008. Shared principles in NF- κ B signaling. *Cell* 132(3):344–362.
- Hayden MS, Ghosh S. 2011. NF- κ B in immunobiology. *Cell Res.* 21(2):223–244.
- Hehenberger E, et al. 2017. Novel predators reshape Holozoan phylogeny and reveal the presence of a two-component signaling system in the ancestor of animals. *Curr Biol.* 27(13):2043–2050.e6.
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol.* 35(2):518–522.
- Jones P, et al. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9):1236–1240.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 14(6):587–589.
- Kanarek N, London N, Schueler-Furman O, Ben-Neriah Y. 2010. Ubiquitination and degradation of the inhibitors of NF- κ B. *Cold Spring Harb Perspect Biol.* 2(2):a000166.
- Karin M. 1999. How NF- κ B is activated: the role of the I κ B kinase (IKK) complex. *Oncogene* 18(49):6867–6874.
- Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33(2):511–518.
- Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30(14):3059–3066.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Kucharczak J, Simmons MJ, Fan Y, Gélinas C. 2003. To be, or not to be: NF- κ B is the answer—role of Rel/NF- κ B in the regulation of apoptosis. *Oncogene* 22(56):8961–8982.
- Leadbeater BSC. 2015. The choanoflagellates: evolution, biology, and ecology. Cambridge: Cambridge University Press.
- Ledoux AC, Perkins ND. 2014. NF- κ B and the cell cycle. *Biochem Soc Trans.* 42(1):76–81.
- Lee N, Kim D, Kim W-U. 2019. Role of NFAT5 in the immune system and pathogenesis of autoimmune diseases. *Front Immunol.* 10:270.
- Lin L, Ghosh S. 1996. A glycine-rich region in NF- κ B p105 functions as a processing signal for the generation of the p50 subunit. *Mol Cell Biol.* 16(5):2248–2254.
- Lingappan K. 2018. NF- κ B in oxidative stress. *Curr Opin Toxicol.* 7:81–86.
- López-Escardó D, López-García P, Moreira D, Ruiz-Trillo I, Torruella G. 2018. *Parvularia atlantis* gen. et sp. nov., a nucleariid filose amoeba (Holomycota, Opisthokonta). *J Eukaryot Microbiol.* 65(2):170–179.
- Mendoza L, Taylor JW, Ajello L. 2002. The class mesomycetozoa: a heterogeneous group of microorganisms at the animal-fungal boundary. *Annu Rev Microbiol.* 56:315–344.
- Mikhailov KV, et al. 2009. The origin of Metazoa: a transition from temporal to spatial cell differentiation. *Bioessays* 31(7):758–768.
- Minh BQ, et al. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol.* 37(5):1530–1534.
- Minh BQ, Nguyen MAT, von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol.* 30(5):1188–1195.
- Mistry J, et al. 2021. Pfam: the protein families database in 2021. *Nucleic Acids Res.* 49(D1):D412–D419.
- Moorthy AK, et al. 2006. The 20S proteasome processes NF- κ B p105 into p50 in a translation-independent manner. *EMBO J.* 25(9):1945–1956.
- Müller CW. 1995. Structure of the NF- κ B p50 homodimer bound to DNA. *Nature* 373:311–317.
- Müller MR, Rao A. 2010. NFAT, immunity and cancer: a transcription factor comes of age. *Nat Rev Immunol.* 10(9):645–656.
- Mylnikov AP, Tikhonenkov DV, Karpov SA, Wylezich C. 2019. Microscopical studies on *Ministeria vibrans* Tong, 1997 (Filasterea) highlight the cytoskeletal structure of the common ancestor of Filasterea, Metazoa and Choanoflagellata. *Protist* 170(4):385–396.
- Napetschnig J, Wu H. 2013. Molecular basis of NF- κ B signaling. *Annu Rev Biophys.* 42(1):443–468.
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274.
- Park S, Uesugi M, Verdine GL. 2000. A second calcineurin binding site on the NFAT regulatory domain. *Proc Natl Acad Sci U S A.* 97(13):7130–7135.
- Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol.* 26(7):1641–1650.
- Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5(3):e9490.
- Richter DJ, Fozouni P, Eisen M, King N. 2018. Gene family innovation, conservation and loss on the animal stem lineage. *eLife* 7:e34226.
- Salminen A, Hyttinen JMT, Kauppinen A, Kaarniranta K. 2012. Context-dependent regulation of autophagy by IKK-NF- κ B signaling: impact on the aging process. *Int J Cell Biol.* 2012:849541.

- Sebé-Pedrós A, de Mendoza A, Lang BF, Degnan BM, Ruiz-Trillo I. 2011. Unexpected repertoire of metazoan transcription factors in the unicellular holozoan *Capsaspora owczarzaki*. *Mol Biol Evol.* 28(3):1241–1254.
- Sebé-Pedrós A, et al. 2013. Regulated aggregative multicellularity in a close unicellular relative of metazoa. *Elife* 2:e01287.
- Sen R, Baltimore D. 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell* 46(5):705–716.
- Serfling E, et al. 2004. NFAT and NF- κ B factors—the distant relatives. *Int J Biochem Cell Biol.* 36(7):1166–1170.
- Shaw JP, et al. 1988. Identification of a putative regulator of early T cell activation genes. *Science* 241(4862):202–205.
- Söding J. 2005. Protein homology detection by HMM–HMM comparison. *Bioinformatics* 21(7):951–960.
- Stibbs HH, Owczarzak A, Bayne CJ, DeWan P. 1979. Schistosome sporocyst-killing amoebae isolated from *Biomphalaria glabrata*. *J Invertebr Pathol.* 33(2):159–170.
- Tikhonenkov DV, et al. 2020. New lineage of microbial predators adds complexity to reconstructing the evolutionary origin of animals. *Curr Biol.* 30(22):4500–4509.e5.
- Tong SM. 1997. Heterotrophic flagellates and other protists from Southampton Water, U.K. *Ophelia* 47(2):71–131.
- Torruella G, et al. 2015. Phylogenomics reveals convergent evolution of lifestyles in close relatives of animals and fungi. *Curr Biol.* 25(18):2404–2410.
- Urrutia A, et al. 2021. *Txikispora philomaos* n. sp., n. g., a micro-eukaryotic pathogen of amphipods, reveals parasitism and hidden diversity in Class Filasterea. *bioRxiv*, preprint. doi:10.1101/2021.01.19.427289.
- Wigington CP, et al. 2020. Systematic discovery of short linear motifs decodes calcineurin phosphatase signaling. *Mol Cell* 79(2):342–358.e12.
- Williams LM, Gilmore TD. 2020. Looking down on NF- κ B. *Mol Cell Biol.* 40(15):e00104-20.

Associate editor: Sandra Baldauf