



Article

Characterization of p53 Family Homologs in Evolutionary Remote Branches of Holozoa

Martin Bartas ^{1,†}, Václav Brázda ^{2,†} , Jiří Červený ¹ and Petr Pečinka ^{1,*}

¹ Department of Biology and Ecology, University of Ostrava, 710 00 Ostrava, Czech Republic; bartasmartin@seznam.cz (M.B.); jiri.cerven@osu.cz (J.C.)

² Institute of Biophysics of the Czech Academy of Sciences, 612 65 Brno, Czech Republic; vaclav@ibp.cz

* Correspondence: petr.pecinka@osu.cz; Tel.: +420-553-46-2318

† These authors contributed equally to this work.

Received: 23 October 2019; Accepted: 16 December 2019; Published: 18 December 2019



Abstract: The p53 family of transcription factors plays key roles in development, genome stability, senescence and tumor development, and p53 is the most important tumor suppressor protein in humans. Although intensively investigated for many years, its initial evolutionary history is not yet fully elucidated. Using bioinformatic and structure prediction methods on current databases containing newly-sequenced genomes and transcriptomes, we present a detailed characterization of p53 family homologs in remote members of the Holozoa group, in the unicellular clades Filasterea, Ichthyosporea and Corallochytraea. Moreover, we show that these newly characterized homologous sequences contain domains that can form structures with high similarity to the human p53 family DNA-binding domain, and some also show similarities to the oligomerization and SAM domains. The presence of these remote homologs demonstrates an ancient origin of the p53 protein family.

Keywords: p53; p63; p73; evolution; Holozoa

1. Introduction

p53 is an intensively studied protein due to its role in tumor suppression, where p53 mutations are found in approximately 50% of human tumors [1–3] and dysregulation of p53 activity occurs in most other cancers [4]. It is accepted that p53 family proteins (p53, p63 and p73) arose from a common ancestral gene (most similar to contemporary p63 found in human [5]) and their functions diversified after gene duplications and rearrangements [6]. All three contemporary p53 family members are transcription factors that execute their functions through binding to DNA consensus sequences that are often located within locally structured regions [7–10]. The typical p53 consensus binding site has internal symmetry and two copies of the motif 5'-RRRCWWGYYY-3' [11], although non-canonical half sites [12], targets with long spacers [9] and with various structural features [13,14] are also recognized by p63 and p73 [15]. All p53 family members act as tetramers through their oligomerization domains, and p63 and p73 also contain a C-terminal sterile alpha motif (SAM) [16,17], which has been lost in p53. The DNA binding domain (DBD) of p53 contains four of the five most highly conserved regions of vertebrate p53 and is a hotspot for p53 mutations in cancer, underscoring its importance for function [4,18].

The most detailed evolutionary studies focused on p53 family diversification after gene duplication and rearrangements, which led to the formation of proteins with different functional properties. These diversifications include variations in structurally disordered regions [19], in various secondary structures [20], loss of the SAM domain for p53 [21] and changes of regulatory phosphorylation sites [22]. On the contrary, the core domain containing the DBD retains high homology and preserves similar DNA binding affinities for all known p53 family proteins. The precise DNA-binding properties

of these proteins are regulated not only by changes in this domain, but also in the N-terminal and C-terminal domains (especially for human p53) [23]. Post-translational modifications [24–27] as well as interaction with other proteins [28,29] modulates DNA-binding and subsequent transcriptional activation/suppression of individual target genes.

While p53 family homologs are well described in Metazoa and especially in vertebrates [22], the two earliest branching homologs to date were described outside Metazoa in clade Choanoflagellates [30–33] and some p53 homologs were mentioned in some general studies of transcription factors in further premetazoan clades but did not go deeper into sequence analysis [34–36]. According to [6], the ancestral p53/63/73 protein present in early metazoans (like sea anemones) probably plays a role in germ cell protection. These homologs also have a tetramerization domain, which is present in today's sea anemones (*Nematostella vectensis*) and is similar to the human tetramerization domain, except for missing a glutamine-rich region [20]. The biological role of the remote homologs found in Holozoa is unclear. To shed light onto the evolution of the p53 family, we have characterized the sequences of p53 orthologs in all non-animal eukaryotes, with a focus on the existing genomic and transcriptomic data of unicellular Holozoa. The investigation of sequence databases revealed the presence of p53 homologs in all clades of unicellular Holozoa (Choanoflagellata, Filasterea, Ichthyosporea, Corallochytrata), with two new p53 homologs in *Ichthyophonus hoferi* and *Chromosphaera perkinsii* (both belong to Ichthyosporea). No p53 homolog was found outside Holozoa. We used amino acid sequence homology analyses and 3D modeling predictions to identify structural similarities in evolutionary close relatives and in human proteins.

2. Results

2.1. Homology of the Sequences

Until now, most p53/63/73 homologs were found in the clade Metazoa. Using the Blast algorithm (blastp and tblastn) [37], we examined various databases (non-redundant protein sequences, WGS, EST, STS, GSS, TSA and non-annotated sets of protein sequences from <http://multicellgenome.com/meet-our-organisms>) [38,39]). The results revealed eleven significant hits outside Metazoa (E-value < 0.001), all designated as hypothetical proteins, five from clade Choanoflagellata (*Monosiga brevicollis* XP_001746020.1; *Monosiga brevicollis* XP_001747656.1; *Salpingoeca rosetta* XP_004994590.1; *Salpingoeca rosetta* XP_004991397.1 and *Salpingoeca rosetta* XP_004991396.1), one from clade Filasterea (*Capsaspora owczarzaki*, XP_004365382.2), four from clade Ichthyosporea (*Sphaeroforma arctica*, XP_014156832.1; *Creolimax fragrantissima* CFRG4869T1; *Ichthyophonus hoferi* Ihof_evm3s137; *Chromosphaera perkinsii* Nk52_evm78s1737) and one from clade Corallochytrata (*Corallochytrium limacisporum* Clim_evm153s157) (see Table 1, significant domain homology is indicated by a plus (+) mark). All eleven non-metazoan homologous sequences are given in Supplementary Material 1. To validate the homology of these proteins, we performed reciprocal searches using the phmmer tool [40] against the reference proteome database. These results show significant homology for multiple p53 family proteins, including human p53, p63 and p73 (Supplementary Material 2).

Table 1. List of p53 family homologs in unicellular organisms belonging to Holozoa. “–” absence and “+” presence of particular protein features (E-value < 0.001). Abbreviations: TAD (transactivation domain), DBD (DNA-binding domain), TET (tetramerization motif), SAM (sterile alpha motif), NLS (nuclear localization signal).

Organism Name	ID of p53 Homolog		PFAM Domains				NLS
	Protein Length (aa)	Predicted pI	TAD	DBD	TET	SAM	
<i>Monosiga brevicollis</i> (class Choanoflagellate)	XP_001746020.1						
	523		–	+	–	–	+
	6.68						
	XP_001747656.1						
<i>Salpingoeca rosetta</i> (class Choanoflagellate)	571		–	+	–	–	+
	5.89						
	XP_004991396.1						
	170		–	+	–	–	–
	5.55						
	XP_004991397.1						
<i>Capsaspora owczarzakii</i> (class Filasterea)	351		–	+	–	+	+
	5.74						
	XP_004994590.1						
	613		–	+	+	+	–
<i>Sphaeroforma arctica</i> (class Ichthyosporea)	6.72						
	XP_004365382.2						
	703		–		+	+	+
<i>Creolimax fragrantissima</i> (class Ichthyosporea)	7.05						
	XP_014156832.1						
	525		–	+	+	–	–
<i>Ichthyophonus hoferi</i> (class Ichthyosporea)	9.57						
	CFRG4869T1						
	768		–	+	+	–	–
<i>Chromosphaera perkinsii</i> (class Ichthyosporea)	8.72						
	Ihof_evm3s137						
	498		–	+	+	–	+
<i>Corallochytrium limacisporum</i> (Corallochytrata)	7.75						
	Nk52_evm78s1737						
	756		–	+	–	+	+
<i>Climacisporium</i> (Corallochytrata)	6.29						
	Clim_evm153s157						
<i>Climacisporium</i> (Corallochytrata)	623		–	+	–	–	+
	5.50						

Although these proteins vary in size (170 to 768 aa residues) and in predicted pI (5.5 to 9.57), they all share significant homology with the p53 family DBD. Moreover, five have at least a partial tetramerization motif and four contain a SAM superfamily domain. None have homology with the p53 family transactivation domain (Figure 1).

Due to the evolutionary distance and parallel evolution of these organisms for hundreds of millions of years, the large number of conserved sequences is surprising. We compared the presence of individual amino acids in the DBD of the six most remote p53 family homologs (from clades Filasterea, Ichthyosporea and Corallochytrata) with human p53 (Table 2). The 100% identity of 11 amino acid residues and conservation of positively charged residues at positions 248 and 273 of human p53 points to their crucial importance for p53 family protein structure and function. Fascinatingly, these conserved residues are important for p53-DNA binding in human p53 and/or are hotspot mutation sites in human cancer.

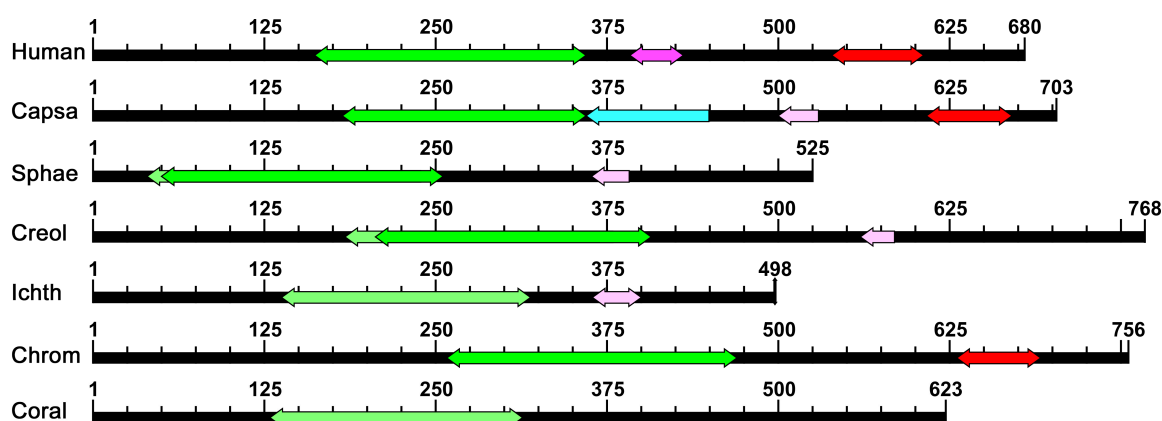


Figure 1. The most evolutionarily distant homologs of human p53 family proteins. Particular domains are highlighted (green and light green for p53 superfamily, purple and light purple for the tetramerization motif, red for SAM domain and light blue for sporulation related domain found only in *Capsaspora owczarzaki*). Lighter shades symbolize lower levels of homology. Abbreviations: Caps (*Capsaspora owczarzaki*), Sphae (*Sphaeroforma arctica*), Creol (*Creolimax fragrantissima*), Ichth (*Ichthyophonus hoferi*), Chrom (*Chromosphaera perkinsii*), Coral (*Corallochytrium limacisporum*).

Table 2. Most conserved amino acid residues in the DBD of the six most remote p53 family homologs related to human p53. For detailed alignment with consensus positions see Supplementary Material 5.

Human p53 Position	Amino Acid	Conservation in All Homologs	Function
142	P	100%	Unknown
173	V	100%	V173L hotspot [41]
175	R	100%	R175H hotspot [42]
176	C	100%	Zn ²⁺ binding [43]
238	C	100%	Zn ²⁺ binding [43]
241	S	100%	S241F hotspot [42]
247	N	100%	N247I hotspot [41]
248	R	+charged aa	R248W hotspot [42]
273	R	+charged aa	R273H hotspot [42]
275	C	100%	Unknown
278	P	100%	Unknown
280	R	100%	R280K hotspot [44]
281	D	100%	D281G hotspot [45]

We inspected the exon-intron structure of p53 homologs in two paralogs of *Monosiga brevicollis*, three paralogs of *Salpingoeca rosetta*, one homolog from *Capsaspora owczarzaki* and one from *Sphaeroforma arctica* (Figure 2). All these homologous genes are very short, maximum length just over 5 kbp in *Salpingoeca rosetta* (XM_004994533.1; protein XP_004994590.1). For comparison, exon-intron structure of three close relative Metazoans are shown and it is evident that the length of introns in p53 homologs increases with the organism's complexity (for example, the human p53 gene is over 250 kbp long). Furthermore, it is interesting that the homologous gene in *Capsaspora owczarzaki* has only two exons, so the whole DBD is located in exon 2. We could not investigate exon-intron structure of homologous genes from more distant clades, because their genomic DNA sequences are unavailable at this time.

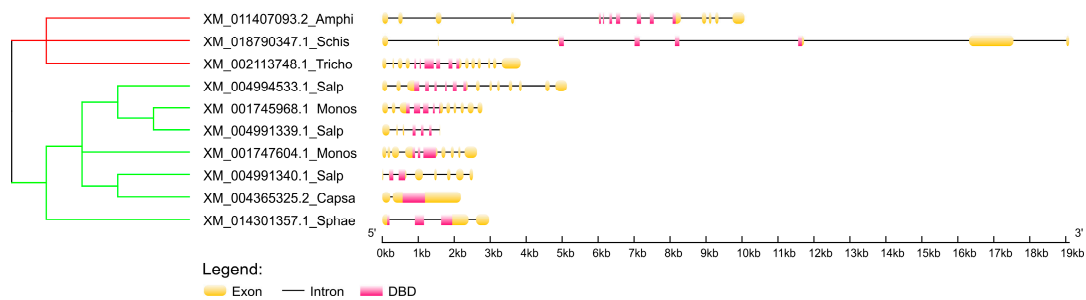


Figure 2. Exon-intron structure of remote human p53 family homologs in the earliest branches of Holozoa. Non-metazoan branches are colored green and metazoan branches red. The tree is based on DNA alignment of cDNAs (T-Coffee aligner) and MrBayes phylogeny. Detailed tree with bootstrapping values and real branch lengths is enclosed in Supplementary Material 6. Abbreviations: Amphi (*Amphimedon queenslandica*), Schis (*Schistosoma mansoni*), Tricho (*Trichoplax adhaerens*), Salp (*Salpingoeca rosetta*), Monos (*Monosiga brevicollis*), Capsa (*Capsaspora owczarzaki*), Sphae (*Sphaeroforma arctica*).

2.2. Homology of Predicted Structures

We used QUARK ab initio protein structure prediction algorithm [46,47] to de novo model the structure of the DBDs. It is clearly visible that all Holozoan p53 family homologs contain functionally important beta sheets in their DBDs (as well as human p63 ab initio modelled structure and most importantly also experimentally verified p53 family DBDs), but the alleged homolog from *Entamoeba histolytica* contains mainly alpha-helices (Figure 3A). Root mean square deviation (RMSD) between experimentally determined structure of human p63 DBD (PDB code: 2rmn) and our ab initio modelled human p63 DBD was only 1.259 Å. Using SWISS-MODEL [48] and UCSF Chimera [49] we visualized 3D models of their DBDs (Figure 3B). Overall RMSD was 0.718 Å (compared to the experimentally determined DBD structure of human p63 as the reference structure). Since the accepted cut off for similarity is <2 Å, this points to a high level of DBD structural conservation in these distant homologs. Simulation of the *E. histolytica* protein was not allowed due its low homology. All predicted models in PDB format are enclosed in Supplementary Material 3 (QUARK) and Supplementary Material 4 (SWISS-MODEL). The fact that the sequences in the DBD might arrange in a similar tridimensional conformation as human p53 suggests these orthologs in unicellular Holozoa could bind to DNA, as observed for animal p53/63/73.

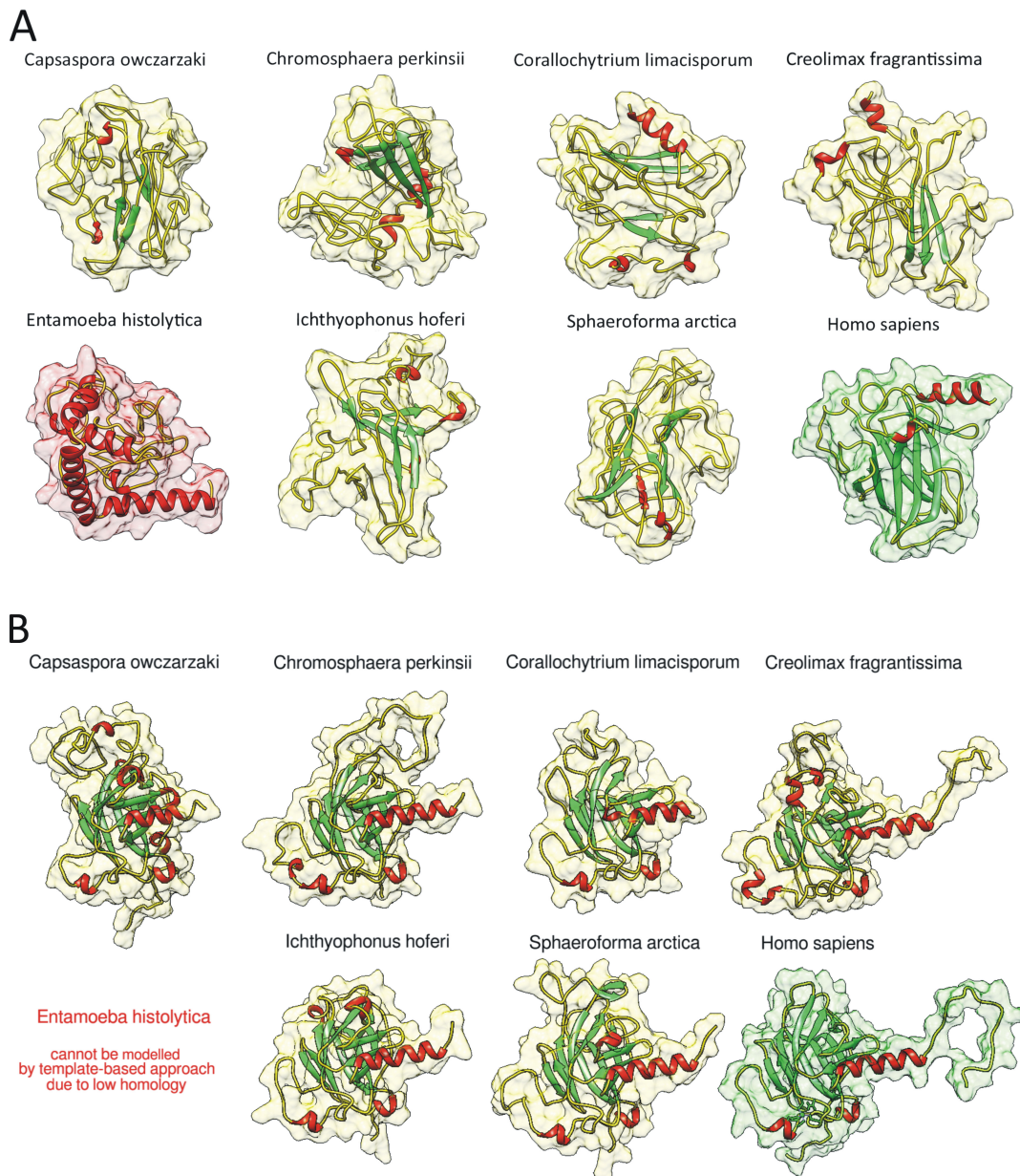


Figure 3. (A) Ab initio modelled DBD structures using the QUARK algorithm [45,46] of the most distant unicellular Holozoa, the alleged homolog from *Entamoeba histolytica* and human p63 DBD (also modelled ab initio from its sequence, to be comparable). (B) The most distant unicellular Holozoan homologs of human p53 family DBDs and experimentally validated p63 DBD. Template based predictions were made using NMR solution model of human p63 DBD (PDB: 2rmn). Beta sheets are colored green, alpha helices red.

The most distant predicted structure of p53 family homolog DBDs was found in *Corallochytrium limacisporum*. This is a single-celled eukaryote living in coral reef lagoons; it was considered to be part of fungi, but contemporary phylogenetic analyses based on DNA sequencing show that this organism is a member of the Holozoa branch [38], as are all known p53 family homologs. Structure-based sequence alignment of primary amino acid sequences of human p63 and the six most distant non-Metazoan organisms (Figure 4) shows high homology (orange rectangles), complete identity is highlighted in the consensus line in red upper case. Structure-based sequence alignment show us combination of sequence alignment and structural features (alpha helices, beta sheets) as determined by the Match Maker tool (3D superimposition of all template-based structures predicted and experimental 2rmn p63).

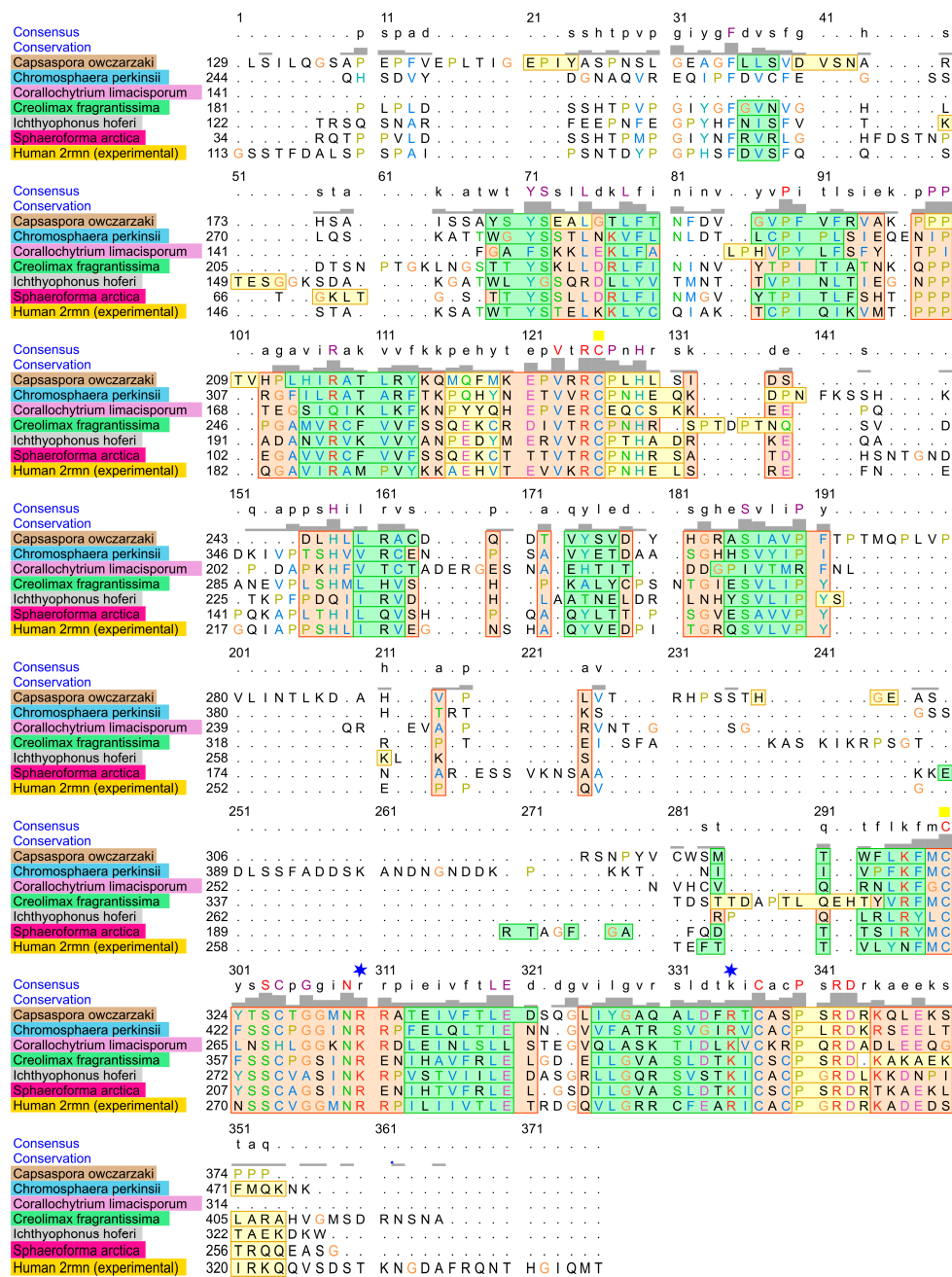


Figure 4. Amino acid alignment of the human p63 DBD and the homologous DBDs from the six most distant non-Metazoan organisms; identical residues are shown above the alignment; pink shaded areas represent helices, green represents beta sheets and yellow represents unstructured parts, orange rectangles show sites of high structural similarity. Text colors represent amino acid similarity. Two blue stars depict positions of the most important DNA binding residues R248 and R273 (related to positions in human p53 sequence). Two yellow squares indicate positions of zinc-coordinating amino acid residues C176 and C238 (related to positions in human p53 sequence).

3. Discussion

The high homology between the core domains of p53 family members suggests that the remote proteins found, containing newly described homologs, have DNA binding features similar to p53 from mammals, amphibians, fish, insects and nematodes [50] (Table 2, alignment of the entire DBD protein sequences of novel homologs with human p53 is shown in Supplementary Material 5). The positive

charge of the key DNA binding amino acid residues 248 and 273 of the human canonical p53 sequence is conserved in all remote homologs; interestingly these residues are always arginines (as in the human sequence) or lysines, or a combination of the two. Furthermore, two of four zinc-coordinated amino acid residues, particularly C176 and C238, are 100% conserved in all remote homologs. Zinc ions play a critical role in stabilizing the architecture of the DBD. We also found significant homology in the tetramerization domains, which indicates the possibility of functional tetramers forming in these homologs. Interestingly, tetramerization domains were not found in *Monosiga brevicollis*, *Chromosphaera perkinsii* and *Corallochytrium limacisporum*, but this may be due to the fact that the tetramerization domain is relatively short and homology lies slightly below selected E-value threshold. Obviously, the binding properties of the remote homologs should be determined by combinations of wet-lab methods.

From our data, we may speculate as to the original function of p53 homologs in unicellular Holozoans and Metazoans. It is supposed that the first evolutionarily role of p53 in primitive Metazoans could be in apoptosis regulatory network via activation by upstream kinase CHK2 [30] and/or in DNA repair via activation of RNR (ribonucleotide reductase) gene expression, which produces deoxyribonucleotide triphosphates (dNTPs) required for DNA replication and repair [30,51]. The absence of an homologous transactivation domain in the most remote p53 family homologs is consistent with previous observations that the transactivation domain first appeared in Placozoa [52]. However, one must be aware that contemporary databases and articles contain misleading information. For example, it is clear that there is no p53 family homolog in plants, but there is a misleading GenBank annotated p53 protein sequence in *Zea mays* (GenBank: AAT42177.1). This is a clear artefact due to human DNA contamination (the *Zea mays* sequence has 100% identity with human p53). The presence of a p53 family homolog has also been reported in *Entamoeba histolytica* (Amoebozoa group) [53]—however our results indicate that this protein is not a p53 family homolog; the protein shows the highest homology to the Alpha kinase superfamily (E-value = 2.5×10^{-8}) and lacks homology with the p53 family DBD domain (Supplementary Material 7, compared to thousands of significant hits for the proteins characterized in this study including human p53, p63 and p63 proteins in Supplementary Material 2).

The finding of an alleged p53 family homolog in *Entamoeba histolytica* provides a good example of the problems inherent in searching for distant protein homologs based solely on overall sequence similarity and has led to misleading information in subsequent papers, including a well-known review on p53 family evolution, where Amoebozoa is included in the evolutionary tree [30]. To avoid such problems, we initially used sequence similarity of the core domain, the most highly conserved region in p53 family evolution and known to be vital for functional activity. We then used structural predictions to model the 3-D organization of the putative homologs and identify those with similar structures. Based on these recent data we have updated the p53 family ancestral tree and show the closest evolutionary branches where p53 family homologs are not present (Figure 5).

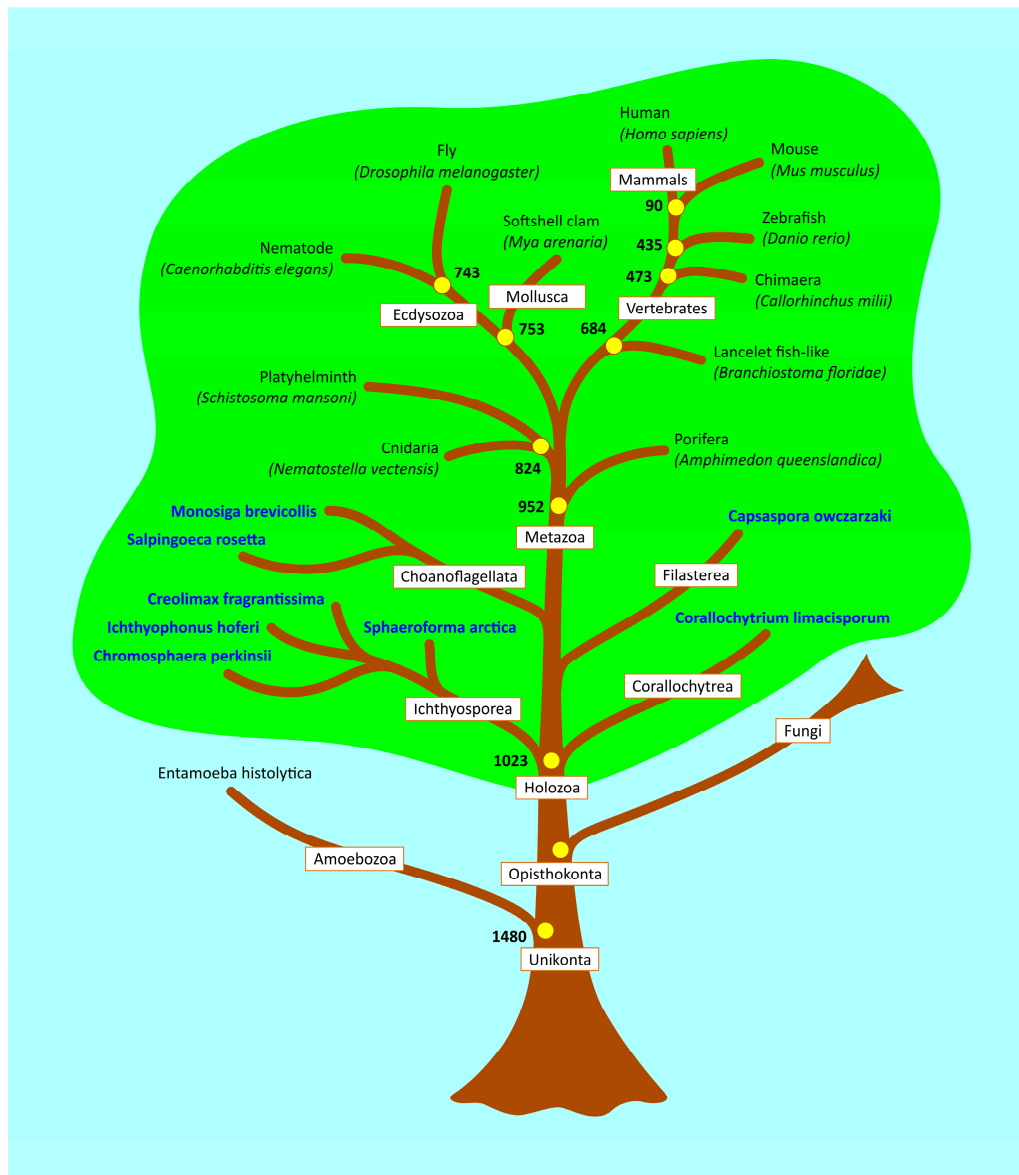


Figure 5. Simplified p53 family tree with our newly characterized p53 family homologs. Yellow circles show the evolutionary age of phylogenetic nodes by TIMETREE database. Non-metazoan organisms with p53 family homologs are in bold. The p53 family world is highlighted in green and the non-p53 family world (including Amoebozoa and Fungi) is in blue.

Recently, two paralogs from *Monosiga brevicollis* and three paralogs from *Salpingoeca rosetta* were mentioned in a paper dealing with the p53 gene family in vertebrates [54]. These results support our data from non-Metazoan analyses. Our results confirm the presence of p53 family homologs in remote members of Holozoa clades Filasterea, Ichthyosporea and Corallochytriales. All newly characterized homologs contain a domain predicted to have high structural similarity to the core DNA-binding domain of p53 family proteins. Moreover, some of these unicellular organisms also contain a tetramerization motif and/or a SAM superfamily domain. We also analyzed exon-intron structures where the genomic sequences were available. It is of interest that the whole DBD of *Capsaspora owczarzaki* is located inside a single exon, suggesting that this may be the ancestral form of gene arrangement. Unfortunately, genomic sequencing data for the most remote branches (Ichthyosporea and Corallochytriales) are not available, so we cannot infer if the *Capsaspora owczarzaki* p53 homolog contains a similar exon-intron arrangement to the ancient form of the gene. As depicted in Table 1,

only the p53 DBD is conserved in all remote homologs, and it is well known that this part of the protein has the critical biological function of binding to DNA. The N-terminal transactivation domain first appeared in Metazoa, as shown by Aberg et al. [52]. The tetramerization motif is too short and variable to be analyzed from the phylogenetic point of view and SAM domains are present in many proteins through all three domains of life (Eukaryota, Archaea, Bacteria) [55], so they are not specific to p53/63/73 evolution. Large bioinformatic analyses suggest that gain of p53 DBD occurred together with the gain of Runt, zf-C2HC and bZIP_Maf and the loss of WRKY transcription factors [35]. The results of this study characterize in detail the presence of p53 family homologs in unicellular evolutionary distant branches of Holozoa and provide important insights for further research on this key family of transcription factors.

4. Materials and Methods

4.1. P53 Family Homolog Search

The human p53 canonical protein sequence (UNIPROT ID: P04637-1) was used for blast search against the contemporary version of the Reference protein database (refseq): <https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins> [37], excluding taxid Metazoa (taxid:33208). To broaden our search we downloaded non-annotated sets of protein sequences (without GenBank IDs) from the database <http://multicellgenome.com/meet-our-organisms> [38,39]. Using UGENE [56] we organized a local blastp database and searched using the newly identified p53 family sequence from *Sphaeroforma arctica*. This approach identified four additional p53 family homologs. All eleven non-metazoan p53 homologous sequences are enclosed in Supplementary Material 1. We also used non-redundant protein sequences, EST, WGS, STS, GSS, TSA and INTERPRO databases to validate and confirm our results. No other homologs have been found in these databases.

4.2. Validation of p53 Homologs (Reciprocal Search)

All putative p53 homologs were reverse searched using the sensitive tool phmmer (<https://www.ebi.ac.uk/Tools/hmmer/search/phmmer>) [40] against “Reference Proteomes” database, E-value cut-off set to 1. Detailed results are enclosed in Supplementary Material 2.

4.3. Isoelectric Point Predictions

Isoelectric point predictions were made using IPC—Isoelectric Point Calculator (<http://isoelectric.org/index.html>) [57].

4.4. Domain Searches and Visualization

Domain annotation was performed using NCBI Web-CD (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) with default parameters except for “search mode” (live search only) and “expect value threshold” (0.001). Results were visualized with the “browse results” option and “compact mode” view [58], *p*-values of all domain hits are enclosed in Supplementary Material 8.

4.5. Nuclear Localization Signal Predictions

Nuclear localization signal predictions were made using cNLS Mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi) [59] with cut-off score = 3 (“localized to both the nucleus and the cytoplasm”).

4.6. Estimation of Evolutionary Age

Information on the evolutionary age of particular phylogenetic nodes was mined from TIMETREE database (<http://www.timetree.org/>) [60,61]. Detailed figures with 95% confidence intervals are enclosed in Supplementary Material 9.

4.7. Validation of *Entamoeba Histolytica* Putative p53 Homolog

Web version of phmmer algorithm was used [40], sequence database was set to “Reference Proteomes”, E-value cut-off set to 1. As a query protein sequence, DBD from putative *Entamoeba histolytica* p53 homolog was used (FASTA file enclosed in Supplementary Material 10). To demonstrate that this putative p53 homolog contains another domain (unrelated to p53 family), SMART tool [62] with PFAM domains search option was used.

4.8. Alignment of DBD from the Most Distant Human p53 Homologs

Alignment was done with MUSCLE [63] under default parameters (flow was used [56]), FASTA sequences of p53 DBD are enclosed in Supplementary Material 11. Conserved human p53 DBD regions II–V were added according to the graph published previously [64].

4.9. p53 Protein Tree

The p53 protein tree is based on eleven unicellular non-Metazoans and twenty-seven other representative closest relative Metazoan species (enclosed in Supplementary Material 12), which were previously aligned by MUSCLE [63], was built using MrBayes [65] with the following parameters: Likelihood model (Number of substitution types: 6 GTR; Substitution model: Dayhoff-protein; Rates variation across sites: gamma + invariable); Markov Chain Monte Carlo parameters (Number of generations: 10,000; Sample a tree every 10 generations); Discard first 250 trees sampled. Resulting p53 protein tree is enclosed in Supplementary Material 13.

4.10. Modelling of 3D Protein Structures

First, we have taken DBD sequences from the most distant holozoan p53 family homologs (clades Filasterea, Ichthyosporea and Corallochytraea) and modelled 3D structures ab initio using QUARK webserver (<https://zhanglab.cmb.med.umich.edu/QUARK2/>) [46,47]. Then we used SWISS-MODEL template-based approach (<https://www.swissmodel.expasy.org/interactive>) [48] to better predict 3D structures using individual FASTA sequences and reference PDB:2rmn as the solution structure of p63 DBD from *Homo sapiens* [66]. All resulting PDB files are enclosed in Supplementary Material 3. Predicted structures of p53 family homologs were visualized in UCSF Chimera 1.12 [49,67]. Structural alignment of human and most distant unicellular Holozoan DBDs was made using MatchMaker [67].

4.11. Comparison of Exon-Intron Composition of the Selected p53 Homologs

Complementary DNAs (cDNAs) and related genomic regions were downloaded from NCBI and curated as two FASTA files (separate FASTA file for all cDNAs and separate FASTA file for all corresponding genomic regions). These FASTA files were uploaded to Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn/>) [68] and DBDs were annotated using tabular file manually filled with PFAM data on all inspected cDNAs obtained previously. The DNA tree was reconstructed using Phylogeny.fr [69] webserver and “A la carte” mode of analysis with T-Coffee aligner [70] selected for alignment of cDNAs and MrBayes [65] phylogeny with subsequent parameters; Likelihood model: Number of substitution types “GTR”, Substitution model “Default”, Rates variation across sites “Invariable + gamma”, Markov Chain Monte Carlo parameters “10 000 generations and sampling tree every 10 generations”, Discard first “250” trees sampled. The detailed resulting DNA tree is enclosed in Supplementary Material 6.

Supplementary Materials: Supplementary materials can be found at <http://www.mdpi.com/1422-0067/21/1/6/s1>.

Author Contributions: Conceptualization, M.B. and V.B.; Data curation, M.B.; Formal analysis, M.B. and V.B.; Funding acquisition, P.P.; Investigation, J.Č.; Methodology, M.B. and V.B.; Project administration, P.P.; Resources, P.P.; Supervision, P.P.; Validation, J.Č.; Visualization, M.B.; Writing—original draft, M.B. and V.B.; Writing—review & editing, V.B., J.Č. and P.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The Czech Science Foundation (18-15548S); by the Ministry of Education, Youth and Sports of the Czech Republic in the “National Feasibility Program I”, project LO1208 TEWEP”; by the EU structural funding Operational Programme Research and Development for innovation, project No. CZ.1.05/2.1.00/19.0388; and by the project SGS/09/PrF/2019 financed by University of Ostrava.

Acknowledgments: We thank Philip Coates for proofreading and editing.

Conflicts of Interest: The authors declare no conflicts of interests.

References

1. Hollstein, M.; Sidransky, D.; Vogelstein, B.; Harris, C. P53 Mutations in Human Cancers. *Science* **1991**, *253*, 49–53. [[CrossRef](#)] [[PubMed](#)]
2. Goh, A.M.; Coffill, C.R.; Lane, D.P. The role of mutant p53 in human cancer. *J. Pathol.* **2011**, *223*, 116–126. [[CrossRef](#)] [[PubMed](#)]
3. Stiewe, T.; Haran, T.E. How mutations shape p53 interactions with the genome to promote tumorigenesis and drug resistance. *Drug Resist. Update* **2018**, *38*, 27–43. [[CrossRef](#)] [[PubMed](#)]
4. Joerger, A.C.; Fersht, A.R. The p53 pathway: Origins, inactivation in cancer, and emerging therapeutic approaches. *Annu. Rev. Biochem.* **2016**, *85*, 375–404. [[CrossRef](#)] [[PubMed](#)]
5. Strano, S.; Rossi, M.; Fontemaggi, G.; Munarriz, E.; Soddu, S.; Sacchi, A.; Blandino, G. From p63 to p53 across p73. *FEBS Lett.* **2001**, *490*, 163–170. [[CrossRef](#)]
6. Belyi, V.A.; Levine, A.J. One billion years of p53/p63/p73 evolution. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 17609–17610. [[CrossRef](#)]
7. Beckerman, R.; Prives, C. Transcriptional Regulation by P53. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, 1–18. [[CrossRef](#)]
8. Brázda, V.; Coufal, J. Recognition of Local DNA Structures by p53 Protein. *Int. J. Mol. Sci.* **2017**, *18*, 375. [[CrossRef](#)]
9. Vyas, P.; Beno, I.; Xi, Z.; Stein, Y.; Golovenko, D.; Kessler, N.; Rotter, V.; Shakked, Z.; Haran, T.E. Diverse p53/DNA binding modes expand the repertoire of p53 response elements. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 10624–10629. [[CrossRef](#)]
10. Brázda, V.; Fojta, M. The Rich World of p53 DNA Binding Targets: The Role of DNA Structure. *Int. J. Mol. Sci.* **2019**, *20*, 5605. [[CrossRef](#)]
11. El-Deiry, W.S.; Kern, S.E.; Pietenpol, J.A.; Kinzler, K.W.; Vogelstein, B. Definition of a consensus binding site for p53. *Nat. Genet.* **1992**, *1*, 45–49. [[CrossRef](#)] [[PubMed](#)]
12. Tebaldi, T.; Zaccara, S.; Alessandrini, F.; Bisio, A.; Ciribilli, Y.; Inga, A. Whole-genome cartography of p53 response elements ranked on transactivation potential. *BMC Genom.* **2015**, *16*, 464. [[CrossRef](#)] [[PubMed](#)]
13. Brázda, V.; Čechová, J.; Battistin, M.; Coufal, J.; Jagelská, E.B.; Raimondi, I.; Inga, A. The structure formed by inverted repeats in p53 response elements determines the transactivation activity of p53 protein. *Biochem. Biophys. Res. Commun.* **2017**, *483*, 516–521. [[CrossRef](#)] [[PubMed](#)]
14. Coufal, J.; Jagelská, E.B.; Liao, J.C.; Brázda, V. Preferential binding of p53 tumor suppressor to p21 promoter sites that contain inverted repeats capable of forming cruciform structure. *Biochem. Biophys. Res. Commun.* **2013**, *441*, 83–88. [[CrossRef](#)]
15. Dötsch, V.; Bernassola, F.; Coutandin, D.; Candi, E.; Melino, G. p63 and p73, the Ancestors of p53. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*. [[CrossRef](#)]
16. Arrowsmith, C.H. Structure and function in the p53 family. *Cell Death Differ.* **1999**, *6*, 1169–1173. [[CrossRef](#)]
17. Joerger, A.C.; Rajagopalan, S.; Natan, E.; Veprintsev, D.B.; Robinson, C.V.; Fersht, A.R. Structural evolution of p53, p63, and p73: Implication for heterotetramer formation. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 17705–17710. [[CrossRef](#)]
18. Pavletich, N.P.; Chambers, K.A.; Pabo, C.O. The DNA-binding domain of p53 contains the four conserved regions and the major mutation hot spots. *Genes Dev.* **1993**, *7*, 2556–2564. [[CrossRef](#)]
19. Xue, B.; Brown, C.J.; Dunker, A.K.; Uversky, V.N. Intrinsically disordered regions of p53 family are highly diversified in evolution. *BBA Proteins Proteom.* **2013**, *1834*, 725–738. [[CrossRef](#)]
20. Joerger, A.C.; Wilcken, R.; Andreeva, A. Tracing the Evolution of the p53 Tetramerization Domain. *Structure* **2014**, *22*, 1301–1310. [[CrossRef](#)]
21. Yang, A.; Kaghad, M.; Caput, D.; McKeon, F. On the shoulders of giants: p63, p73 and the rise of p53. *Trends Genet.* **2002**, *18*, 90–95. [[CrossRef](#)]

22. Dos Santos, H.G.; Nunez-Castilla, J.; Siltberg-Liberles, J. Functional diversification after gene duplication: Paralogue specific regions of structural disorder and phosphorylation in p53, p63, and p73. *PLoS ONE* **2016**, *11*, e0151961. [[CrossRef](#)] [[PubMed](#)]
23. Laptenko, O.; Shiff, I.; Freed-Pastor, W.; Zupnick, A.; Mattia, M.; Freulich, E.; Shamir, I.; Kadouri, N.; Kahan, T.; Manfredi, J. The p53 C terminus controls site-specific DNA binding and promotes structural changes within the central DNA binding domain. *Mol. Cell* **2015**, *57*, 1034–1046. [[CrossRef](#)] [[PubMed](#)]
24. Gu, W.; Roeder, R.G. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell* **1997**, *90*, 595–606. [[CrossRef](#)]
25. Pospíšilová, S.; Brázda, V.; Kucharíková, K.; Luciani, M.G.; Hupp, T.R.; Skládal, P.; Palecek, E.; Vojtesek, B. Activation of the DNA-binding ability of latent p53 protein by protein kinase C is abolished by protein kinase CK2. *Biochem. J.* **2004**, *378*, 939–947. [[CrossRef](#)]
26. Ou, Y.-H.; Chung, P.-H.; Sun, T.-P.; Shieh, S.-Y. p53 C-terminal phosphorylation by CHK1 and CHK2 participates in the regulation of DNA-damage-induced C-terminal acetylation. *Mol. Biol. Cell* **2005**, *16*, 1684–1695. [[CrossRef](#)]
27. Muller, P.; Chan, J.M.; Simoncik, O.; Fojta, M.; Lane, D.P.; Hupp, T.; Vojtesek, B. Evidence for allosteric effects on p53 oligomerization induced by phosphorylation. *Protein Sci.* **2018**, *27*, 523–530. [[CrossRef](#)]
28. Brázda, V.; Paleček, J.; Pospíšilová, S.; Vojtěsek, B.; Paleček, E. Specific modulation of p53 binding to consensus sequence within supercoiled DNA by monoclonal antibodies. *Biochem. Biophys. Res. Commun.* **2000**, *267*, 934–939. [[CrossRef](#)]
29. Weinberg, R.L.; Freund, S.M.V.; Veprintsev, D.B.; Bycroft, M.; Fersht, A.R. Regulation of DNA binding of p53 by its C-terminal domain. *J. Mol. Biol.* **2004**, *342*, 801–811. [[CrossRef](#)]
30. Lu, W.-J.; Amatruda, J.F.; Abrams, J.M. p53 ancestry: Gazing through an evolutionary lens. *Nat. Rev. Cancer* **2009**, *9*, 758–762. [[CrossRef](#)]
31. Belyi, V.A.; Ak, P.; Markert, E.; Wang, H.; Hu, W.; Puzio-Kuter, A.; Levine, A.J. The origins and evolution of the p53 family of genes. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a001198. [[CrossRef](#)] [[PubMed](#)]
32. Nedelcu, A.M.; Tan, C. Early diversification and complex evolutionary history of the p53 tumor suppressor gene family. *Dev. Genes Evol.* **2007**, *217*, 801–806. [[CrossRef](#)] [[PubMed](#)]
33. King, N.; Westbrook, M.J.; Young, S.L.; Kuo, A.; Abedin, M.; Chapman, J.; Fairclough, S.; Hellsten, U.; Isogai, Y.; Letunic, I.; et al. The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature* **2008**, *451*, 783–788. [[CrossRef](#)] [[PubMed](#)]
34. Sebé-Pedrós, A.; de Mendoza, A.; Lang, B.F.; Degnan, B.M.; Ruiz-Trillo, I. Unexpected repertoire of metazoan transcription factors in the unicellular holozoan *Capsaspora owczarzaki*. *Mol. Biol. Evol.* **2011**, *28*, 1241–1254. [[CrossRef](#)] [[PubMed](#)]
35. de Mendoza, A.; Sebé-Pedrós, A.; Šestak, M.S.; Matejčić, M.; Torruella, G.; Domazet-Lošo, T.; Ruiz-Trillo, I. Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E4858–E4866. [[CrossRef](#)] [[PubMed](#)]
36. Sebé-Pedrós, A.; Peña, M.I.; Capella-Gutiérrez, S.; Antó, M.; Gabaldón, T.; Ruiz-Trillo, I.; Sabidó, E. High-throughput proteomics reveals the unicellular roots of animal phosphosignaling and cell differentiation. *Dev. Cell* **2016**, *39*, 186–197. [[CrossRef](#)]
37. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
38. Grau-Bove, X.; Torruella, G.; Donachie, S.; Suga, H.; Leonard, G.; Richards, T.A.; Ruiz-Trillo, I. Dynamics of genomic innovation in the unicellular ancestry of animals. *eLife* **2017**, *6*, e26036. [[CrossRef](#)]
39. de Mendoza, A.; Suga, H.; Permanyer, J.; Irímia, M.; Ruiz-Trillo, I. Complex transcriptional regulation and independent evolution of fungal-like traits in a relative of animals. *eLife* **2015**, *4*, e08904. [[CrossRef](#)]
40. Potter, S.C.; Luciani, A.; Eddy, S.R.; Park, Y.; Lopez, R.; Finn, R.D. HMMER web server: 2018 update. *Nucleic Acids Res.* **2018**, *46*, W200–W204. [[CrossRef](#)]
41. Chang, F.-L.; Lai, M.-D. Various forms of mutant p53 confer sensitivity to cisplatin and doxorubicin in bladder cancer cells. *J. Urol.* **2001**, *166*, 304–310. [[CrossRef](#)]
42. Olivier, M.; Hollstein, M.; Hainaut, P. TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a001008. [[CrossRef](#)] [[PubMed](#)]
43. Eldar, A.; Rozenberg, H.; Diskin-Posner, Y.; Rohs, R.; Shakked, Z. Structural studies of p53 inactivation by DNA-contact mutations and its rescue by suppressor mutations via alternative protein–DNA interactions. *Nucleic Acids Res.* **2013**, *41*, 8748–8759. [[CrossRef](#)] [[PubMed](#)]

44. Gomes, A.S.; Trovão, F.; Andrade Pinheiro, B.; Freire, F.; Gomes, S.; Oliveira, C.; Domingues, L.; Romão, M.J.; Saraiva, L.; Carvalho, A.L. The Crystal Structure of the R280K Mutant of Human p53 Explains the Loss of DNA Binding. *Int. J. Mol. Sci.* **2018**, *19*, 1184. [[CrossRef](#)] [[PubMed](#)]
45. Dittmer, D.; Pati, S.; Zambetti, G.; Chu, S.; Teresky, A.K.; Moore, M.; Finlay, C.; Levine, A.J. Gain of function mutations in p53. *Nat. Genet.* **1993**, *4*, 42–46. [[CrossRef](#)] [[PubMed](#)]
46. Xu, D.; Zhang, Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins Struct. Funct. Bioinform.* **2012**, *80*, 1715–1735. [[CrossRef](#)] [[PubMed](#)]
47. Xu, D.; Zhang, Y. Toward optimal fragment generations for ab initio protein structure assembly. *Proteins Struct. Funct. Bioinform.* **2013**, *81*, 229–239. [[CrossRef](#)] [[PubMed](#)]
48. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [[CrossRef](#)]
49. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)]
50. Lion, M.; Raimondi, I.; Donati, S.; Jousson, O.; Ciribilli, Y.; Inga, A. Evolution of p53 Transactivation Specificity through the Lens of a Yeast-Based Functional Assay. *PLoS ONE* **2015**, *10*, e0116177. [[CrossRef](#)]
51. Jordan, A.; Reichard, P. Ribonucleotide Reductases. *Annu. Rev. Biochem.* **1998**, *67*, 71–98. [[CrossRef](#)] [[PubMed](#)]
52. Åberg, E.; Saccoccia, F.; Grabherr, M.; Ore, W.Y.J.; Jemth, P.; Hultqvist, G. Evolution of the p53-MDM2 pathway. *BMC Evol. Biol.* **2017**, *17*, 177. [[CrossRef](#)] [[PubMed](#)]
53. Mendoza, L.; Orozco, E.; Rodriguez, M.A.; Garcia-Rivera, G.; Sanchez, T.; Garcia, E.; Gariglio, P. Ehp53, an Entamoeba histolytica protein, ancestor of the mammalian tumour suppressor p53. *Microbiology* **2003**, *149*, 885–893. [[CrossRef](#)] [[PubMed](#)]
54. Biscotti, M.A.; Barucca, M.; Carducci, F.; Forconi, M.; Canapa, A. The p53 gene family in vertebrates: Evolutionary considerations. *J. Exp. Zool. Part B* **2019**, *332*, 171–178. [[CrossRef](#)]
55. Aravind, L.; Anantharaman, V.; Iyer, L.M. Evolutionary connections between bacterial and eukaryotic signaling systems: A genomic perspective. *Curr. Opin. Microbiol.* **2003**, *6*, 490–497. [[CrossRef](#)]
56. Okonechnikov, K.; Golosova, O.; Fursov, M.; Team, U. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **2012**, *28*, 1166–1167. [[CrossRef](#)]
57. Kozłowski, L.P. IPC—Isoelectric Point Calculator. *Biol. Direct* **2016**, *11*, 55. [[CrossRef](#)]
58. Marchler-Bauer, A.; Derbyshire, M.K.; Gonzales, N.R.; Lu, S.; Chitsaz, F.; Geer, L.Y.; Geer, R.C.; He, J.; Gwadz, M.; Hurwitz, D.I.; et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **2015**, *43*, D222–D226. [[CrossRef](#)]
59. Kosugi, S.; Hasebe, M.; Tomita, M.; Yanagawa, H. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10171–10176. [[CrossRef](#)]
60. Hedges, S.B.; Marin, J.; Suleski, M.; Paymer, M.; Kumar, S. Tree of life reveals clock-like speciation and diversification. *Mol. Biol. Evol.* **2015**, *32*, 835–845. [[CrossRef](#)]
61. Kumar, S.; Stecher, G.; Suleski, M.; Hedges, S.B. TimeTree: A resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* **2017**, *34*, 1812–1819. [[CrossRef](#)] [[PubMed](#)]
62. Letunic, I.; Bork, P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* **2018**, *46*, D493–D496. [[CrossRef](#)] [[PubMed](#)]
63. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)] [[PubMed](#)]
64. de Anta, J.M.; Jassem, E.; Rosell, R.; Martínez-Roca, M.; Jassem, J.; Martínez-López, E.; Monzó, M.; Sánchez-Hernández, J.J.; Moreno, I.; Sánchez-Céspedes, M. TP53 mutational pattern in Spanish and Polish non-small cell lung cancer patients: Null mutations are associated with poor prognosis. *Oncogene* **1997**, *15*, 2951. [[CrossRef](#)] [[PubMed](#)]
65. Huelsenbeck, J.P.; Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **2001**, *17*, 754–755. [[CrossRef](#)] [[PubMed](#)]
66. Enthart, A.; Klein, C.; Dehner, A.; Coles, M.; Gemmecker, G.; Kessler, H.; Hagn, F. Solution structure and binding specificity of the p63 DNA binding domain. *Sci. Rep.* **2016**, *6*, 26707. [[CrossRef](#)] [[PubMed](#)]

67. Meng, E.C.; Pettersen, E.F.; Couch, G.S.; Huang, C.C.; Ferrin, T.E. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinform.* **2006**, *7*, 339.
68. Hu, B.; Jin, J.; Guo, A.-Y.; Zhang, H.; Luo, J.; Gao, G. GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* **2015**, *31*, 1296–1297. [[CrossRef](#)]
69. Dereeper, A.; Guignon, V.; Blanc, G.; Audic, S.; Buffet, S.; Chevenet, F.; Dufayard, J.-F.; Guindon, S.; Lefort, V.; Lescot, M. Phylogeny. fr: Robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **2008**, *36*, W465–W469. [[CrossRef](#)]
70. Notredame, C.; Higgins, D.G.; Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **2000**, *302*, 205–217. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).