Journal of Computational Methods in Molecular Design, 2013, 3 (2):32-63

Scholars Research Library
(http://scholarsresearchlibrary.com/archive.html)

ISSN : 2231- 3176
CODEN (USA): JCMMDA

In silico Proteomics of NF-κB protein from different mammals

Ranganatha S[1,3], Prerna Singh[2], Shruthi SD[1,2], Govindappa[3] and Ramachandra YL[1]*

[1]*Department of P.G. Studies and Research in Biotechnology, Kuvempu University, Shankaraghatta , Karnataka, India.*
[2]*P.G. Department of Biotechnology, The Oxford College of Science, Bangalore, Karnataka, India*
[3]*Sri Shridevi Research Centre, Sira Road, Tumkur, Karnataka*

ABSTRACT

NF-kappa-B is a pleiotropic transcription factor present in almost all cell types and is the endpoint of a series of signal transduction events that are initiated by a vast array of stimuli related to many biological processes such as inflammation, immunity, differentiation, cell growth, tumorigenesis and apoptosis. In silico proteomics study of NF-κB was conducted to find out the similarities and differences between sequences from different mammals. And also to find out the functional sites, physico chemical parameters, domains, secondary and tertiary structures of the protein. Hence, the study conducted will give insight regarding the function and components of NF-κB sequences from different mammals taken for the research activity.

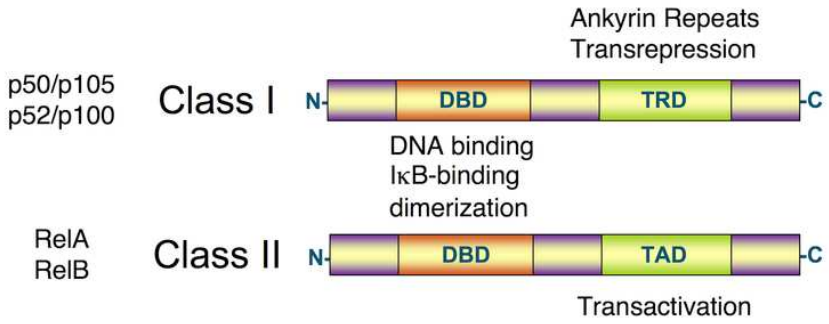
Keywords: function, inflammation, phylogenetic analysis, secondary and tertiary structures

---

INTRODUCTION

NF-κB (nuclear factor kappa-light-chain-enhancer of activated B cells) is a protein complex that controls the transcription of DNA. NF-κB is found in almost all animal cell types and is involved in cellular responses to stimuli such as stress, cytokines, free radicals, ultraviolet irradiation, oxidized LDL, and bacterial or viral antigens. NF-κB plays a key role in regulating the immune response to infection (kappa light chains are critical components of immunoglobulins) [1,2]**.** Incorrect regulation of NF-κB has been linked to cancer, inflammatory and autoimmune diseases, septic shock, viral infection, and improper immune development. NF-κB has also been implicated in processes of synaptic plasticity and memory. All proteins of the NF-κB family share a Rel homology domain in their N-terminus [3]. A subfamily of NF-κB proteins, including RelA, RelB, and c-Rel, has a transactivation domain in their C-termini. In contrast, the NF-κB1 and NF-κB2 proteins are synthesized as large precursors, p105, and p100, which undergo processing to generate the mature NF-κB subunits, p50 and p52, respectively. The processing of p105 and p100 is mediated by the ubiquitin/proteasome pathway and involves selective degradation of their C-terminal region containing ankyrin repeats.

Schematic diagram of NF-κB protein structure [4]



Both classes of proteins contain N-terminal DNA-binding domain (DBD), which serves as a dimerization interface to other NF-κB transcription factors and, in addition, binds to the inhibitory IκBα protein. The C-terminus of class I proteins contains a number of ankyrin repeats and has transrepression activity. In contrast, the C-terminus of class II proteins has a transactivation function,

In silico is an expression used to mean "performed on computer or via computer simulation". It is based on the analysis of some papers that present scientific applications which rely on *in silico* experiments [5,6]. There are basically two ways of viewing them. The first view is that, it is a computer program that realizes some specific operations which constitutes some particular experimental conditions, allowing investigating biological phenomena, and complementing those present in *in vivo* and *in vitro* experiments [7]. According to the second view, *in silico* corresponds more closely to the meaning of 'simulation' where its identity is linked to that of the model used to construct such simulation. *In silico* organisms grows in nanoseconds so experiments that would normally take months can be performed in minutes. *In silico* model may be used to identify new drug targets, particularly those capable of killing persistent bacilli. Several useful predictions have been obtained from such models. The large volume of genome-scale data that is being produced and made available in databases on the World Wide Web is demanding the development in *in silico* computer representation [8,9].

In the growing field of proteomics, tools for the *in silico* analysis of proteins and even of whole proteomes are of crucial importance to make best use of the accumulating amount of data. To utilise this data for healthcare and drug development, understanding the characteristics of proteomes of entire species is necessary along with their differentiation between individuals can be surveyed [10].

The present study is conducted to predict the secondary structures, scalar parameters, hydrophobicity, cleavage sites, subcellular localization, transmembrane helix, functional annotation and tertiary structure of the NF-κB protein from different mammals namely**,** *Homo sapiens* (human), *Mus musculus* (mouse), *Bos Taurus* (bovine), *Ailuropoda melanoleuca* (giant panda), *Sus scrofa* (pig), *ovis aries* (sheep), *Pongo abelii* (Sumatran orangutan), *Macaca mulatta* (Rhesus macaque), *Rattus norvegicus* (Rat) and *Canis familiaris* (Dog) and compare the uniformity between the characters.

MATERIALS AND METHODS

NCBI- National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/)
NCBI advances science and health by providing access to biomedical and genomic information. The fasta sequences of all NF-κB proteins were retrieved in NCBI and further used for other analysis.

UniProt KB/Swiss-Prot (http://web.expasy.org/)
UniProtKB/Swiss-Prot is the manually annotated and reviewed section of the UniProt Knowledgebase (UniProtKB). It has high quality annotated and non-redundant protein sequences, which brings together experimental results, computed features and scientific conclusions. The tool was used for the determination of biological process, domain, cellular component etc. And also for the prediction of functional and post translational modifications of NF-κB in each organism.

PROTPARAM (http://web.expasy.org/protparam/)
Is a tool which allows the computation of various physical and chemical parameters of a given protein. The determination of molecular weight, theoretical pI, atomic composition, half life etc. of NF-κB were predicted using this tool.

PSORT (http://www.psort.org/)
It is a bioinformatics tool used for the prediction of protein localization sites in cells. The tool was used to determine the sub-cellular localization of NF-κB in all the different organisms taken for the study.

PubMed (www.ncbi.nlm.nih.gov/pubmed)
PubMed is a free database accessing primarily the MEDLINE database of references and abstracts on life sciences and biomedical topics. The database was used for doing literature survey on the amount of work already done and also for the tools and softwares used in *in silico* proteomic studies.

TOPPRED (http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html)
Is a tool which is used for the examining the number of membrane - spanning segments present in a particular sequence. For the determination of the number of transmembrane site and the level of hydrophobicity, this tool was used.

InterPro (http://www.ebi.ac.uk/interpro/)
InterPro provides functional analysis of proteins by classifying them into families and predicting domains and important sites. We combined protein signatures from a number of member databases into a single searchable resource, capitalising on their individual strengths to produce a powerful integrated database and diagnostic tool. For the determination of the number of domains, repeats and detailed signature matches this tool was used.

GOR-Garnier-Osguthorpe-Robson (http://gor.bb.iastate.edu/cdm/)
It is an information theory-based method for the prediction of secondary structures in proteins. For the secondary structure determination of the NF-κB in selected organisms, the GOR was used.

PDB (http://www.rcsb.org/pdb/)
It is a repository for the 3-D structural data of large biological molecules, such as proteins and nucleic acids. The data, typically obtained by X-ray crystallography or NMR spectroscopy and submitted by biologists and biochemists are maintained by its member organizations such as PDBe, PDBj and RCSB. The three dimensional structures of NF-κB were retrieved from PDB.

Clustal Omega (http://www.ebi.ac.uk/Tools/services/web_clustalo)
Clustal Omega is a new multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments. It produces biologically meaningful multiple sequence alignments of divergent sequences.

ClustalW2 Phylogeny (http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/)
It is a commonly used phylogenetic tree generation method provided by the ClustalW2 program. It accepts the multiple sequence alignment in any supported format and provides the tree in Clustal, Distance Matrix and NEXUS format. The above tool is used to construct rooted and unrooted phylogenetic tree of NF-κB sequences.

RESULTS AND DISCUSSION

NF-κB (nuclear factor-κB) is a rapidly-acting primary transcription factor, which is present in cells in an inactive state and drastically converted to an active form in the absence of any protein synthesis. This characteristic allows NF-κB to be a first responder to various cellular stimuli, and positions the NF-κB pathway as a key component in the regulation of a variety of cellular processes. NF-kB has been found to play an active role in inflammatory responses, cellular growth, and apoptosis as well as being present in diseases such as cancer, arthritis, asthma, and others. NF-kB functions as a fast messenger. Under normal circumstances, it is sequestered in the cytoplasm by an inhibitor protein. Additionally, the p50 subunit is in the inactive elongated p105 state. Once a viral infection is recognized, the inhibitor is phosphorylated, releasing the complex, and the repeated chain is cleaved. Because the

proteins originate in the cytoplasm already assembled and proceed to the nucleus where they act as the transcription factor, the cellular message is very fast [11,12.13].

In the study conducted the taxonomic differences between different mammals and of NF-κB sequences were understood using NCBI which is shown in **Table 1**. All the sequences taken for study are shown in **Figure 2**. The alternative names of NF-κB, biological processes in which it is involved, cellular component, domains present, molecular function and post translational modifications were studied using UniProtKB/Swiss-Prot (**Table 2**). The physico-chemical parameters of all the NF kappa B sequences like their molecular weight, therotical pI, negatively and positively charged residues, composition of atoms such as carbon, hydrogen, nitrogen, oxygen and sulfur, extension co-efficient, half-life, instabality and aliphatic index were tabulated in **Table 3** as obtained from ProtParam. The sub-cellular localization of NF-κB were predicted using PSORT. It is shown to be localised in cytoplasm, where as in sheep and orangutan it is found to be unknown (**Table 4**). The transmembrane sites present in NF-κB were predicted using TOPPRED, which were found very much similar with all the sequences tested (**Figure 3**). The functional sites and domains present in NF-κB were predicted using INTERPRO (**Figure 4**). The secondary structures such as alpha, beta, coiled coil etc of NF-κB sequences were predicted using GOR and represented in **Figure 5**. The three dimensional structures of NF-κB sequences were retrieved from PDB (**Figure 6**). The structural details such as PDB ID, experimental method, PubMed entry, resolution [Å] and polymers like RELB, P105, DNA, p100, RNA present in it along with the length of polymers are as tabulated in **Table 5**. Multiple sequence alignment of NF-κB sequences were performed using Clustal Omega to find out the similarity between the sequences (**Figure 7**). **Figure 8** and **Figure 9** represents the unrooted and rooted tree of the sequences built using the tool ClustalW2 Phylogeny. The tree illustrates that the NF-κB of human beings are closely related to that of dog and pig. Whereas distantly related to cattle and monkey.

Meffert and colleagues have investigated a possible role for NF-kB transcription factors in mammalian behavior using mice lacking the p65 subunit of NF-kB [1]. NF-kB-like molecules or homologs have been reported in several model organisms, including Drosophila, Aplysia and crabs. Some of the data suggests a role for NF-kB transcription factors in behavior emerged from a habituation test that serves as a model for long-term memory in the crab *Chasmagnathus*. In mammals, the studies using super-repressor-expressing mice have yielded seemingly inconsistent results with regards to the role of NF-kB transcription factors in spatial learning. Nevertheless, experimental differences could account for the discrepancy and might also help to elucidate complexities in the function of NF-kB transcription factors in the mammalian metabolism.

CONCLUSION

In addition to their roles in immune and inflammatory responses, NF-kB family members are well known as crucial regulators of cell proliferation, differentiation, apoptosis and oncogenesis. However, whilst much is known about the signalling pathways that result in NF-κB activation in transformed cells and in mice. Cutting edge technologies, such as small inhibiting (si) RNA, will doubtless also give great insights into the functional roles of these proteins in the future. The exquisitely specific NF-κB response induced by different stimuli in different cells gives hope that treatments can be developed to specifically target NF-κB activation. Many exciting avenues remain to be explored in the investigation of NF-kB function.

Table 1: Showing the 10 different mammals and of NF-κB sequences taken for study from NCBI.

Sl.No.	Organisms	Common name	Accession ID of NCBI	Definition of entry	Number of residues	Taxonomy
1	*Homo sapiens*	Human	P19838	HUMAN Nuclear factor NF-kappa-B p105 subunit	968 aa	Eukaryota; Metazoa; Chordata;Craniata; Vertebrata; Euteleostomi;Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;Catarrhini; Hominidae; Homo
2	*Mus musculus*	Mouse	P25799	NFKB1_MOUSE Nuclear factor NF-kappa-B p105 subunit	971 aa	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus; Mus
3	*Bos taurus*	Domestic Cattle	156523218	NF-kappa-B-repressing factor	690 aa	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
4	*Ailuropoda melanoleuca*	Giant Panda		PREDICTED: mitochondrial ubiquitin ligase activator of NFKB 1-like	352 aa	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Caniformia; Ursidae; Ailuropoda.
5	*Sus scrofa*	Pig	Q0PHA8	PIG Nuclear factor kappa-B 1	959 aa	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;Mammalia;Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;Sus.
6	*Ovis aries*	Sheep	Q9GLG5	SHEEP NF-kappa-B (Fragment)	58 aa	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Caprinae; Ovis.
7	*Pongo abelii*	Orangutan	395731028	PREDICTED: mitochondrial ubiquitin ligase activator of NFKB 1	316 aa	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Pongo.
8	*Macaca mulatta*	Rhesus Monkey	386781067	NF-kappa-B-repressing factor	690 aa	Eukaryota; Metazoa; Chordata;Craniata; Vertebrata; Euteleostomi;Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;Catarrhini;Cercopithecidae; Cercopithecinae; Macaca.
9	*Rattus norvegicus*	Rat	Q63369	RAT Nuclear factor NF-kappa-B p105 subunit	522 aa	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
10	*Canis familiaris*	Dog	Q6F3J0	CANFA Nuclear factor NF-kappa-B p105 subunit	972 aa	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Caniformia; Canidae; Canis.

Figure 2: Showing 10 sequences taken for study belonging to different sources of NF-κB.

>sp|P19838|NFKB1_HUMAN Nuclear factor NF-kappa-B p105 subunit OS=Homo sapiens GN=NFKB1 PE=1 SV=2

MAEDDPYLGRPEQMFHLDPSLTHTIFNPEVFQPQMALPTDGPYLQILEQPKQRGFRFRYVCEGPSHGGLPG
ASSEKNKKSYPQVKICNYVGPAKVIVQLVTNGKNIHLHAHSLVGKHCEDGICTVTAGPKDMVVGFANLGI
LHVTKKKVFETLEARMTEACIRGYNPGLLVHPDLAYLQAEGGGDRQLGDREKELIRQAALQQTKEMDLSV
VRLMFTAFLPDSTGSFTRRLEPVVSDAIYDSKAPNASNLKIVRMDRTAGCVTGGEEIYLLCDKVQKDDIQIR
FYEEEENGGVWEGFGDFSPTDVHRQFAIVFKTPKYKDINITKPASVFVQLRRKSDLETSEPKPFLYYPEIKDK
EEVQRKRQKLMPNFSDSFGGGSGAGAGGGGMFGSGGGGGGTGSTGPGYSFPHYGFPTYGGITFHPGTTKS
NAGMKHGTMDTESKKDPEGCDKSDDKNTVNLFGKVIETTEQDQEPSEATVGNGEVTLTYATGTKEESAG
VQDNLFLEKAMQLAKRHANALFDYAVTGDVKMLLAVQRHLTAVQDENGDSVLHLAIIHLHSQLVRDLLE

VTSGLISDDIINMRNDLYQTPLHLAVITKQEDVVEDLLRAGADLSLLDRLGNSVLHLAAKEGHDKVLSILLK
HKKAALLLDHPNGDGLNAIHLAMMSNSLPCLLLLVAAGADVNAQEQKSGRTALHLAVEHDNISLAGCLL
LEGDAHVDSTTYDGTTPLHIAAGRGSTRLAALLKAAGADPLVENFEPLYDLDDSWENAGEDEGVVPGTTP
LDMATSWQVFDILNGKPYEPEFTSDDLLAQGDMKQLAEDVKLQLYKLLEIPDPDKNWATLAQKLGLGILN
NAFRLSPAPSKTLMDNYEVSGGTVRELVEALRQMGYTEAIEVIQAASSPVKTTSQAHSLPLSPASTRQQIDE
LRDSDSVCDSGVETSFRKLSFTESLTSGASLLTLNKMPHDYGQEGPLEGKI

>sp|P25799|NFKB1_MOUSE Nuclear factor NF-kappa-B p105 subunit OS=Mus musculus GN=Nfkb1 PE=1 SV=2
MADDDPYGTGQMFHLNTALTHSIFNAELYSPEIPLSTDGPYLQILEQPKQRGFRFRYVCEGPSHGGLPGASS
EKNKKSYPQVKICNYVGPAKVIVQLVTNGKNIHLHAHSLVGKHCEDGVCTVTAGPKDMVVGFANLGILH
VTKKKVFETLEARMTEACIRGYNPGLLVHSDLAYLQAEGGGDRQLTDREKEIIRQAAVQQTKEMDLSVVR
LMFTAFLPDSTGSFTRRLEPVVSDAIYDSKAPNASNLKIVRMDRTAGCVTGGEEIYLLCDKVQKDDIQIRFY
EEEENGGVWEGFGDFSPTDVHRQFAIVFKTPKYKDVNITKPASVFVQLRRKSDLETSEPKPFLYYPEIKDKE
EVQRKRQKLMPNFSDSFGGGSGAGAGGGGMFGSGGGGGSTGSPGPGYGYSNYGFPPYGGITFHPGVTKSN
AGVTHGTINTKFKNGPKDCAKSDDEESLTLPEKETEGEGPSLPMACTKTEPIALASTMEDKEQDMGFQDNL
FLEKALQLARRHANALFDYAVTGDVKMLLAVQRHLTAVQDENGDSVLHLAIIHLHAQLVRDLLEVTSGLI
SDDIINMRNDLYQTPLHLAVITKQEDVVEDLLRVGADLSLLDRWGNSVLHLAAKEGHDRILSILLKSRKAA
PLIDHPNGEGLNAIHIAVMSNSLPCLLLLVAAGAEVNAQEQKSGRTALHLAVEYDNISLAGCLLLEGDAHV
DSTTYDGTTPLHIAAGRGSTRLAALLKAAGADPLVENFEPLYDLDDSWEKAGEDEGVVPGTTPLDMAAN
WQVFDILNGKPYEPVFTSDDILPQGDMKQLTEDTRLQLCKLLEIPDPDKNWATLAQKLGLGILNNAFRLSP
APSKTLMDNYEVSGGTIKELMEALQQMGYTEAIEVIQAAFRTPATTASSPVTTAQVHCLPLSSSSTRQHIDE
LRDSDSVCDSGVETSFRKLSFTESLTGDSPLLSLNKMPHGYGQEGPIEGKI

>gi|156523218|ref|NP_001096023.1| NF-kappa-B-repressing factor [Bos taurus]
MEKILQMAEGIDIGEMPSYDLMLSKASKGQKRHLSTCDGQNPPKKQAGSKFHVRPRFEPVHFVASSSKDER
QEDPYGPQAKERNEQTHFANMPRDIYQDYTQDSFSIQDGNSQYCDSSGFIFTKDKPVTANMYFDSGNPAPS
STSQQADSQSPPEPSPSQTFPESVVAEKQYFIEKLTATIWKNLSNPEMTSGSDKINYTYMLTRCIQACKTNPE
YIYAPLKEIPPADIPKNKKLLTDGYACEVRCQNIYLTTGYAGSKNGSRDRATELAVKLLQKRIEVRVIRRKF
KHTFGEDLVVCQIGMPSYDFPPALKPPEELVVLAKDASGQPIFNASAKHWTNFILTENANDAIGILNNSASY
NKMSVEYKYEMMPNRTWRCRVFLQDHCLAEGYGTKKTSKHAAADEALKILQKTQPTYPSVKSSQCQAGS
SPRGSGKKKDIKDLVVYENSSNPVCTLNDTAQFNRMTVEYVYERMTGLRWKCKVILESEVIAEAVGVKKT
VKYEAAGEAVKTLKKTQPTVINNLKKGAIEDVISRNEIQGRSAEEAYKQQIKEDNIGNQLLRKMGWTGGG
LGKSGEGIREPISVKEQHKREGLGLDVERVNKIAKRDIEQIIRNYARSESHTDLTFSTELTNDERKQIHQIAQK
YGLKSKSHGVGHDRYLVVGRKRRKEDLLDQLKQEGQVGHYELVMPQAN

>gi|301759711|ref|XP_002915681.1| PREDICTED: mitochondrial ubiquitin ligase activator of NFKB 1-like [Ailuropoda melanoleuca]
MESGGRPSLGQFILLGTSSVVTAVLYSVYRQKAQVAQELKGAKRIHLGEDLKNILSEAPGKCVPYAVIEGA
VRSVKETLNSQFVENCKGVIQRLTLQEHKMVWNRTTHLWNDYSKIIHQRTNTVPFDLVPHEDGVAVAVR
VLKPLDSQDLGLETVYEKFHPSIQSFTDVIGHYISGERPKGIQETEEMLKVGATLTGVGELVLDNNSVRLQP
PKQGMQYYLSSQDFESLLQRQDSSVRLWKVLTLVFGFATCAALFFLLRRHYLQRQERRRLQQMENEFRHE
AQLLSRAKPEDRESLKSACVVCLSSFKSCVFLECGHVCSCAECYRALPEPKRCPICRQAITRVIPLYNS

>tr|Q0PHA8|Q0PHA8_PIG Nuclear factor kappa-B 1 OS=Sus scrofa GN=NFKB1 PE=2 SV=1
MFHLDPLNHTIFNPELFQPEMPLPTADGPYLQILEQPKQRGFRFRYVCEGPSHGGLPGASSEKNKKSYPQVK
ICNYVGPAKVIVQLVTNGKNIHLHAHSLVGKHGEDGICTVTAGPKDMVVGFANLGILHVTKKKVFETLEA
RMTEACIRGYNPGLLVHPDLAYLQAEGGGDRQLTDREKEIIRQAALQQTKEMDLSVVRLMFTAFLPDSTGS
FTRRLEPVVSDAIYDSKAPNASNLKIVRMDRTAGCVTGGEEIYLLCDKVQKDDIQIRFYEEEENGGIWEGFG
DFSPTDVHRQFAIVFKTPKYKDVNITKPASVFVQLRRKSDLETSEPKPFLYYPEIKDKEEVQRKRQKLMPNF
SDSFGGGSGAGAGGGGMFGSGGGGGGAGSTGPGYGFPHYGFPTYGGITFHAGTTKSNAGMKHGTVDTPS
KNDPEDCDKSDDREAVNLSGKVTETTEQDKESSNGEDEVHLTYSVGVKEENYRFQDNLFLEKAMQLAKQ
HANALFDYAVTGDVKMLLAVQRHLTAVQDENGDSVLHLAIIHLHAQLVRDLLEVTSGLISDEIINMRNDL
YQTPLHLAVITKQEAVVEDLLRAGADLSLLDRLGNSVLHLAAKEGHDKILSILLKHKKAALLINHPNGEGL
NAIHVAMMSNSLPCLLLLMAAGADVNAQERKSGRTALHLAVELDNISLAGCLLLEGDAHVDSTTYDGTTP

LHIAAGRGSTRLAALLKAAGADPLVENFEPLYDLDDSWDEDGEDEGVVPGTTPLDMATNWQVFDILNGKP
YEPEFTSDDLLAQGDMKQLTEDTKLQLYKLLEIPDPDKNWATLAQKLGLGILNNAFRLSAAPSKLMDNYE
VSGGTIKELVEALRQMGYTEAIDVIQAAFCTSGTAATSPVKTTSQAHSLPFSPASTRQQIDELRDDSICDSGV
ETSFRKLSFTESLTSSSSLLTLNKVPHDFGQEGPLEGKI

>tr|Q9GLG5|Q9GLG5_SHEEP NF-kappa-B (Fragment) OS=Ovis aries GN=NfkB PE=2 SV=1
KVQKDDIEVRFYEDDENGWQAFGDFSPTDVHKQYAIVFRTPPYHKMKIERPVTVFLQL

>gi|395731028|ref|XP_003775827.1| PREDICTED: mitochondrial ubiquitin ligase activator of NFKB 1 [Pongo abelii]
MRQAGAKKVHLGEDLKSILSEAPGKCVPYAVIEGAVRSVKETLNSQFVENCKGVIQRLTLQEHKMVWNRT
THLWNDCSKIIHQRTNTVPFDLVPHEDGMDVAVRVLKPLDSVDLGLETVYEKFHPSIQSFTDVIGHYISGER
PKGIQETEEMLKVGATLTGVGELVLDNNSVRLQPPKQGMQYYLSSQDFDSLLQRQESSVRLWKVLALVFG
FATCATLFFILRKQYLQRQERLRLKQMQEEFQEHEAQLLSRAKPEDRESLKSACVVCLSSFKSCVFLECGHV
CSCTECYRALPEPKKCPICRQAITRVIPLYNS

>gi|386781067|ref|NP_001247822.1| NF-kappa-B-repressing factor [Macaca mulatta]
MEKILQMAEGIDIGEMPSYDLVLSKPSKGQKRHLSTCDGQNPPKKQAGSKFHARPRFEPVHFVASSSKDER
QEDPYGPQTKEVNEQTHFASLPRDIYQDYTQDSFSIQDGNSQYCDSSGFILTKDQPVAANMYFDSGNPAPSS
TSQQANSQSTPEPSPSQTFPESVVAEKQYFIEKLTATIWKNLSNPEMTSGSDKINYTYMLTRCIQACKTNPEY
IYAPLKEIPPADIPKNKKLLTDGYACEVRCQNIYLTTGYAGSKNGSRDRATELAVKLLQKRIEVRVVRRKFK
HTFGEDLVVCQIGMSSYEFPPALKPPEDLVVLGKDASGQPVFNASAKHWTNFVITENANDAIGILNNSASFN
KMSIEYKYEMMPNRTWRCRVFLQDHCLAEGYGTKKTSKHAAADEALKILQKTQPTYPSVKSSQCHTGSSP
RGSGKKKDIKDLVVYENSSNPVCTLNDTAQFNRMTVEYVYERMTGLRWKCKVILESEVIAEAVGVKKTV
KYEAAGEAVKTLKKTQPTVINNLKKGAVEDVISRNEIQGRSAEEAYKQQIREDNIGNQLLRKMGWTGGGL
GKSGEGIREPISVKEQHKREGLGLDVERVNKIAKRDIEQIIRNYARSESHTDLTFSRELTNDERKQIHQIAQK
YGLKSKSHGVGHDRYLVVGRKRRKEDLLDQLKQEGQVGHYELVMPQAN

>sp|Q63369|NFKB1_RAT Nuclear factor NF-kappa-B p105 subunit (Fragment) OS=Rattus norvegicus GN=Nfkb1 PE=2 SV=1
REILNPPEKETQGEGPSLFMASTKTEAIAPASTMEDKEEDVGFQDNLFLEKALQLAKRHANALFDYAVTGD
VKMLLAVQRHLTAVQDENGDSVLHLAIIHLHAQLVRDLLEVTSGSISDDIINMRNDLYQTPLHLAVITKQE
DVVEDLLRVGADLSLLDRWGNSVLHLAAKEGHDKILGVLLKNSKAALLINHPNGEGLNAIHIAVMSNSLS
CLQLLVAAGAEVNAQEQKSGRTALHLAVEYDNISLAGCLLLEGDALVDSTTYDGTTPLHIAAGRGSTRLA
ALLKAAGADPLVENFEPLYDLDDSWEKAGEDEGVVPGTTPLDMAANWQVFDILNGKPYEPVFTSDDILPQ
GDIKQLTEDTRLQLCKLLEIPDPDKNWATLAQKLGLGILNNAFRLSPAPSKTLMDNYEVSGGTIKELVEALR
QMGYTEAIEVIQAAFRTPETTASSPVTTAQAHLLPLSSSSTRQHIDELRDNDSVCDSGVETSFRKLSFSESLT
GDGPLLSLNKMPHNYGQDGPIEGKI

>sp|Q6F3J0|NFKB1_CANFA Nuclear factor NF-kappa-B p105 subunit OS=Canis familiaris GN=NFKB1 PE=2 SV=2
MAEDDTYLGAHEQMFHLDPLTHTIFNPELFQPEMPLPTADGPYLQILEQPKQRGFRFRYVCEGPSHGGLPG
ASSEKNKKSYPQVKICNYVGPAKVIVQLVTNGKNIHLHAHSLVGKHCEDGICTVTAGPKDMVVGFANLGI
LHVTKKKVFETLEARMTEACTKGYNPGLLVHPDLAYLQAEGGGDRQLTDREKEIIRQAALQQTKEMDLSV
VRLMFTAFLPDSTGSFTRRLEPVVSDAIYDSKAPNASNLKIVRMDRTAGCVTGGEEIYLLCDKVQKDDIQIR
FYEEEENGGIWEGFGDFSPTDVHRQFAIVFKTPKYKDVNITKPASVFVQLRRKSDLETSEPKPFLYYPEIKDK
EEVQRKRQKLMPNFSDSFGGGSGAGAGGGGMFGSGGGGGGAGSTGPGYGFPHYGFPTYGGITFHPGTTKS
NAGMKHGTIDTPSKNDSEGCGKNVDREAVNLSGKVTEPTEQDKESSMGVDEVTLTYTVGIKEENSRFQDN
LFLEKAMQLAKRHANALFDYAVTGDVKMLLAVQRHLTAVQDENGDSVLHLAIIHLHAQLVRDLLEVTSG
LISDDIINMRNDLYQTPLHLAVITKQEAVVDDLLRAGADLSLLDRLGNSVLHLAAKEGQDKILSILLKHKKA
ALLMDHPNGEGLNAIHIAVMSNSMPCLLLLVAAGADVNAQERKSGRTALHLAVEHDNISLAGCLLLEGDA
HVDSTTYDGTTPLHIAAGRGSTRLAALLKAAGADPLVENFEPLYDLDDSWEKDGEDEGVVPGTTPLDMAT
NWQVFDILNGKPYEPEFTSDDLLAQGDMKQLTEDAKLQLYKLLEIPDPDKNWATLAQKLGLGILNNAFRL
SPAPSKTLMDNYEVSGGTIKELVEALRQMGYTEAIEVIQAAFCAPETAAPSPGKGAPQTLSLPLSSASTRSPV
DEVRDDSICDSGVETSFRKLSFTESLTSGSSLLTLNKAPHEYGQEGPIEGKI

Table 2: Showing necessary annotation of NF kappa B as obtained from UniProtKB/Swiss-Prot.

Sl.No.	Organisms	Alternative names	Biological process	Cellular component	Domain	Molecular function	PTM
1	*Homo sapiens*	DNA-binding factor KBF1 EBP-1, Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	Apoptosis Transcription Transcription regulation	Cytoplasm Nucleus	ANK repeat Repeat	Activator	Acetylation Hydroxylation Lipoprotein Phosphoprotein S-nitrosylation Ubl conjugation
2	*Mus musculus*	DNA-binding factor KBF1 EBP-1 NF-kappa-B1 p84/NF-kappa-B1 p98 Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	Apoptosis Transcription Transcription regulation	Cytoplasm Nucleus	ANK repeat Repeat	Activator	Acetylation Hydroxylation Lipoprotein Phosphoprotein S-nitrosylation Ubl conjugation
3	*Bos taurus*	IkB kinase-associated protein 1	Transcription Transcription regulation	Cytoplasm Nucleus	Coiled coil Zinc-finger	metal ion binding	Disulphide bond Isopeptide bond Phosphoprotein
4	*Ailuropoda melanoleuca*	–	–	–	–	–	–
5	*Sus scrofa*	–	signal transduction .	Nucleus	ANK repeat Repeat	sequence-specific DNA binding transcription factor activity	–
6	*Ovis aries*	Toll-like receptor 2	Immunity Inflammatory response Innate immunity	Membrane	Leucine-rich repeat Repeat Signal Transmembrane Transmembrane helix	Receptor	Disulphide bond Glycoprotien
7	*Pongo abelii*	–	–	–	–	–	–
8	*Macaca mulatta*	NF-kappa-B inhibitor-like protein 1	Cellular response to lipopolysaccharide.	Nucleus	–	–	–
9	*Rattus norvegicus*	DNA-binding factor KBF1 EBP-1 Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	Transcription Trancription regulation	Cytoplasm Nucleus	ANK repeat Repeat	Activator	Hydroxylation Phosphoprotien
10	*Canis familiaris*	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	Transcription Transcripton regulation	Cytoplasm Nucleus	ANK repeat Repeat	Activator	Acetylation Hydroxylation Lipoprotein Phosphoprotein S-nitrosylation

Table 3: Showing necessary physico-chemical parameters of NF kappa B as obtained from ProtParam.

Organism	Molecular weight	Therotical pI	Negatively charged residue	Positively charged Residue	Atomic Composition	Extension Coefficient	Half-life	Instability index	Aliphatic index
Homo sapiens	105356.0	5.20	133	93	Carbon C 4643 Hydrogen H 7343 Nitrogen N 1271 Oxygen O 1458 Sulfur S 33	61365	30 hours	38.15 (stable)	-0.339
Mus Musculus	105615.4	5.20	130	92	Carbon C 4658 Hydrogen H 7360 Nitrogen N 1272 Oxygen O 1457 Sulfur S 35	68605	30 hours	41.25 (unstable)	-0.304
Bos Taurus	77752.0	8.93	86	100	Carbon C 3407 Hydrogen H 5420 Nitrogen N 968 Oxygen O 1057 Sulfur S 28	71460	30 hours	52.81 (unstable)	-0.746
Ailuropoda melanoleuca	39800.8	8.76	37	44	Carbon C 1764 Hydrogen H 2821 Nitrogen N 501 Oxygen O 513 Sulfur S 17	33640	30 hours	59.06 (unstable)	-0.245
Sus scrofa	104450.9	5.22	132	93	Carbon C 4610 Hydrogen H 7279 Nitrogen N 1261 Oxygen O 1443 Sulfur S 31	59875	30 hours	34.92 (stable)	-0.330
Ovis aries	6973.8	5.52	10	8	Carbon C 321 Hydrogen H 478 Nitrogen N 82 Oxygen O 91 Sulfur S 1	9970	1.3 hours	70.94 (unstable)	-0.707
Pongo abelii	35947.5	8.18	37	40	Carbon C 1590 Hydrogen H	29170	30 hours	53.05 (unstable)	-0.267

					2546 Nitrogen N 444 Oxygen O 466 Sulfur S 19				
Macaca mulatta	77624.7	8.95	85	99	Carbon C 3399 Hydrogen H 5411 Nitrogen N 971 Oxygen O 1057 Sulfur S 26	69970	30 hours	52.10 (unstable)	-0.734
Rattus norvegicus	56553.8	4.67	77	41	Carbon C 2488 Hydrogen H 3986 Nitrogen N 676 Oxygen O 798 Sulfur S 13	35660	1 hours	41.66 (unstable)	-0.177
Canis familiaris	105631.4	5.23	133	95	Carbon C 4661 Hydrogen H 7372 Nitrogen N 1272 Oxygen O 1457 Sulfur S 34	61490	30 hours	37.03 (stable)	-0.322

Table 4: Showing sub-cellular localization of NF-κB as predicted using PSORT.

Organism	Cytoplasmic	Cytoplasmic Membrane	Cell wall	Extracellcular	Final Prediction
Human	7.50	1.00	0.87	0.63	Cytoplasm 7.50
Mouse	7.50	1.00	0.87	0.63	Cytoplasm 7.50
Cattle	7.50	1.00	0.87	0.63	Cytoplasm 7.50
Panda	7.50	1.00	0.87	0.63	Cytoplasm 7.50
Pig	7.50	1.00	0.87	0.63	Cytoplasm 7.50
Sheep	2.50	2.50	2.50	2.50	Unknown
Orangutan	2.50	2.50	2.50	2.50	Unknown
Monkey	7.50	1.00	0.87	0.63	Cytoplasm 7.50
Rat	7.50	1.00	0.87	0.63	Cytoplasm 7.50
Dog	7.50	1.00	0.87	0.63	Cytoplasm 7.50

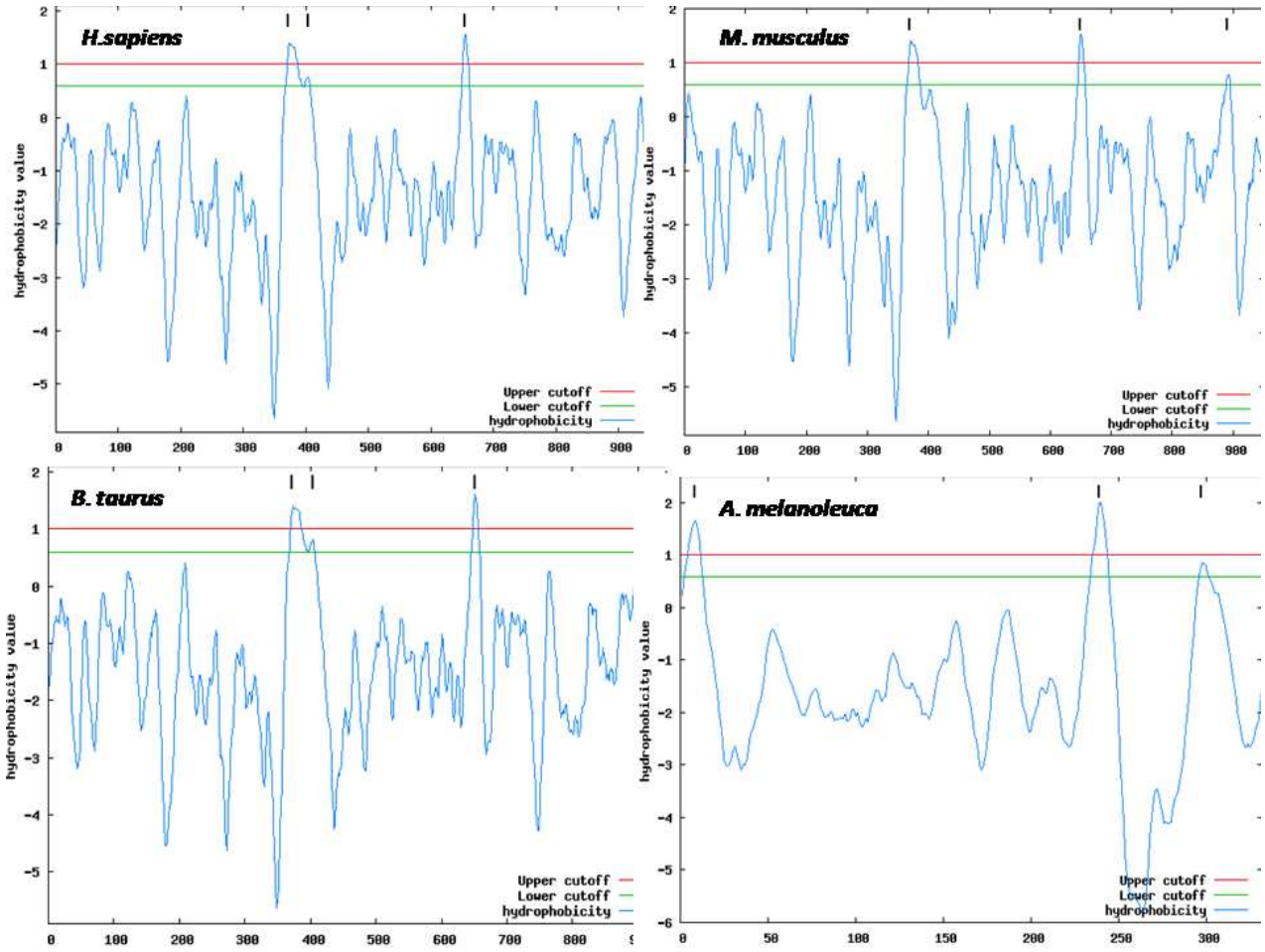


Figure 3A: Showing transmembrane sites of NF-κB as predicted using TOPPRED.

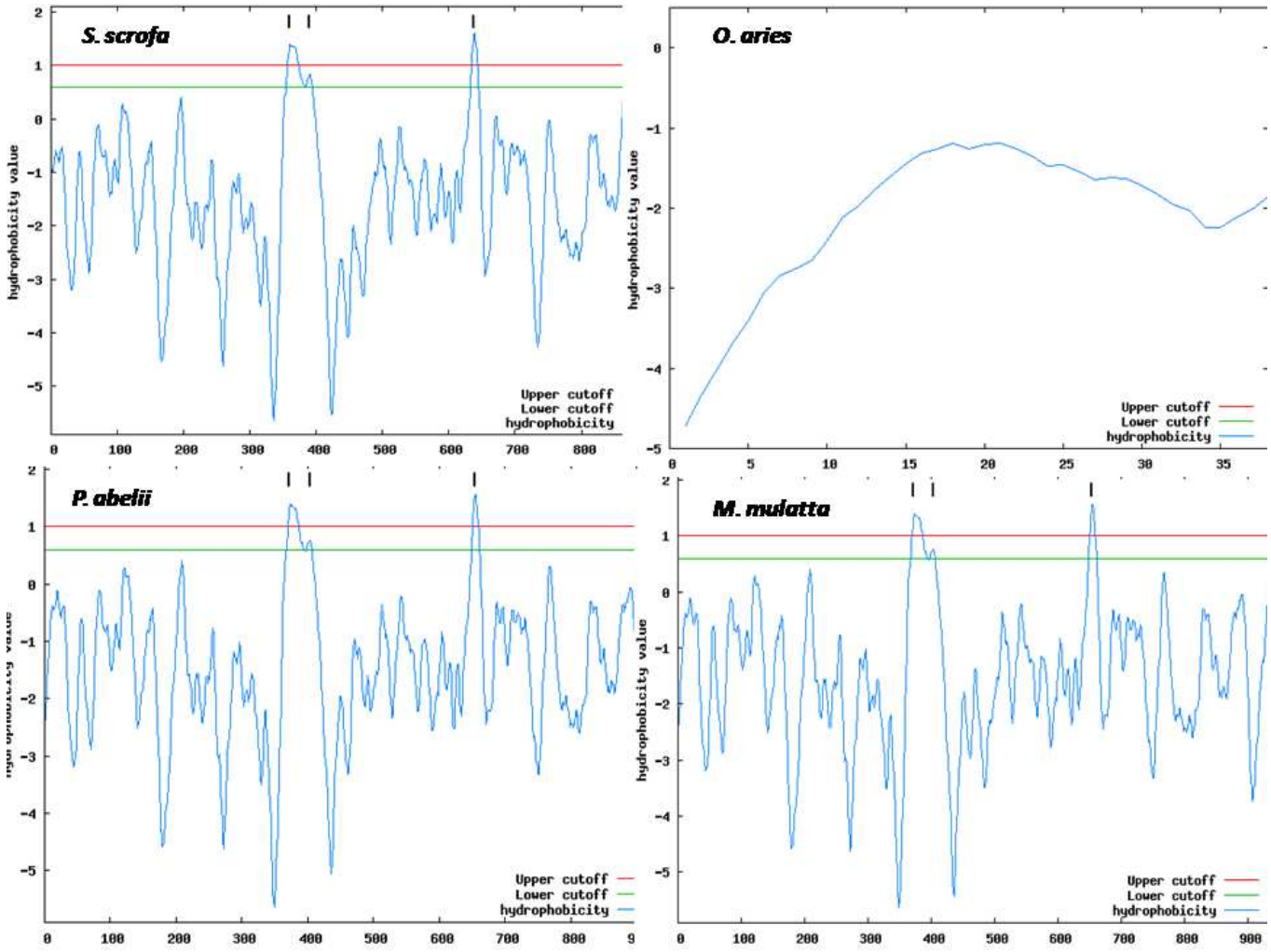


Figure 3B: Showing transmembrane sites of NF-κB as predicted using TOPPRED.

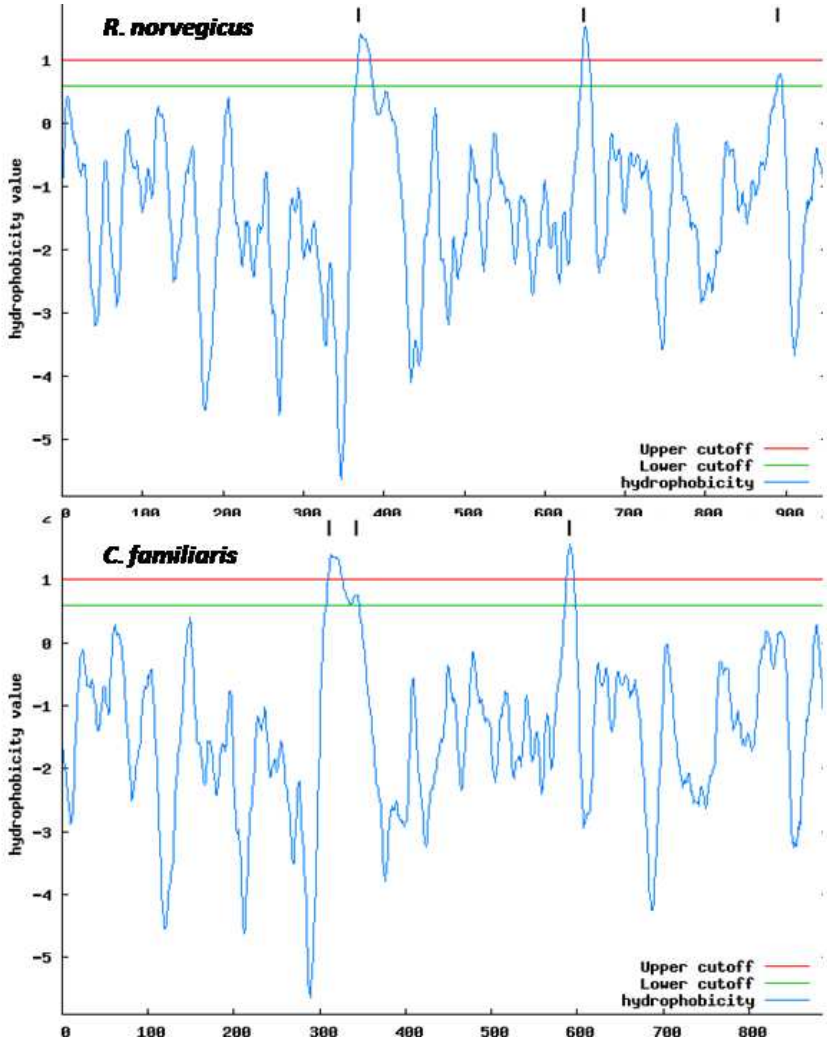


Figure 3C: Showing transmembrane sites of NF-κB as predicted using TOPPRED.

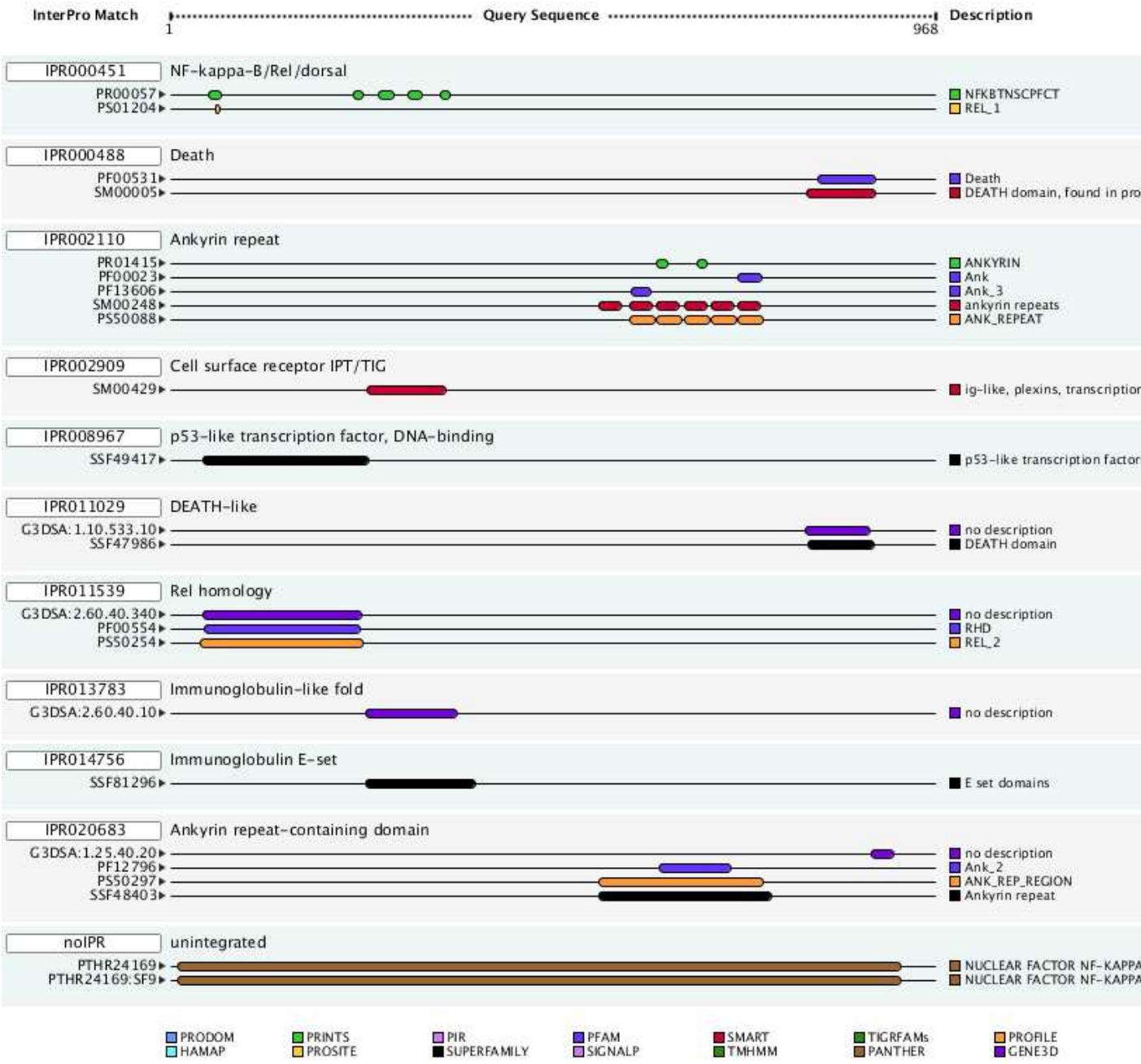


Figure 4A: Showing functional sites and domains of NF-κB from *H. sapiens* as predicted using INTERPRO.

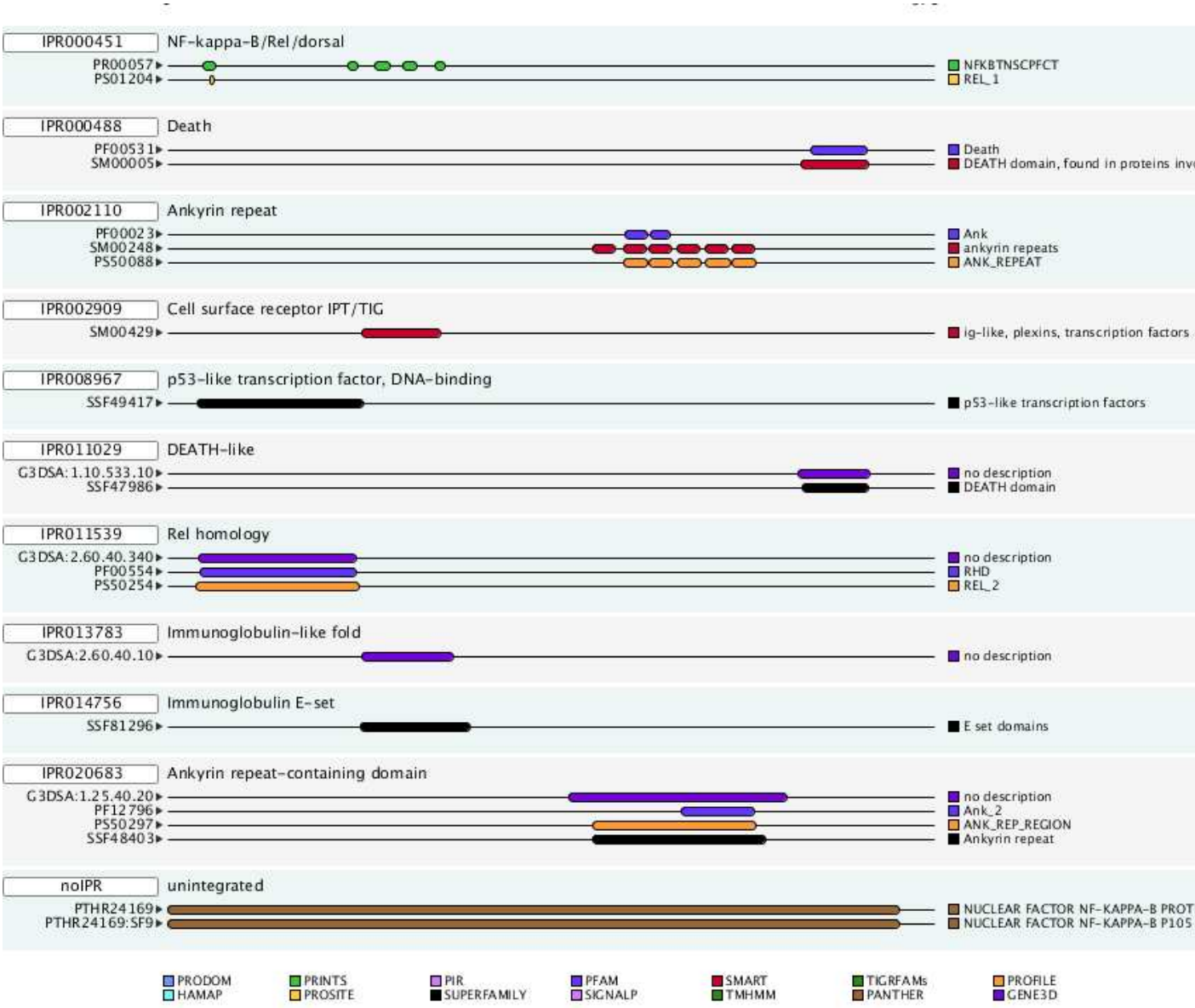


Figure 4B: Showing functional sites and domains of NF-κB from *M. musculus* as predicted using INTERPRO.

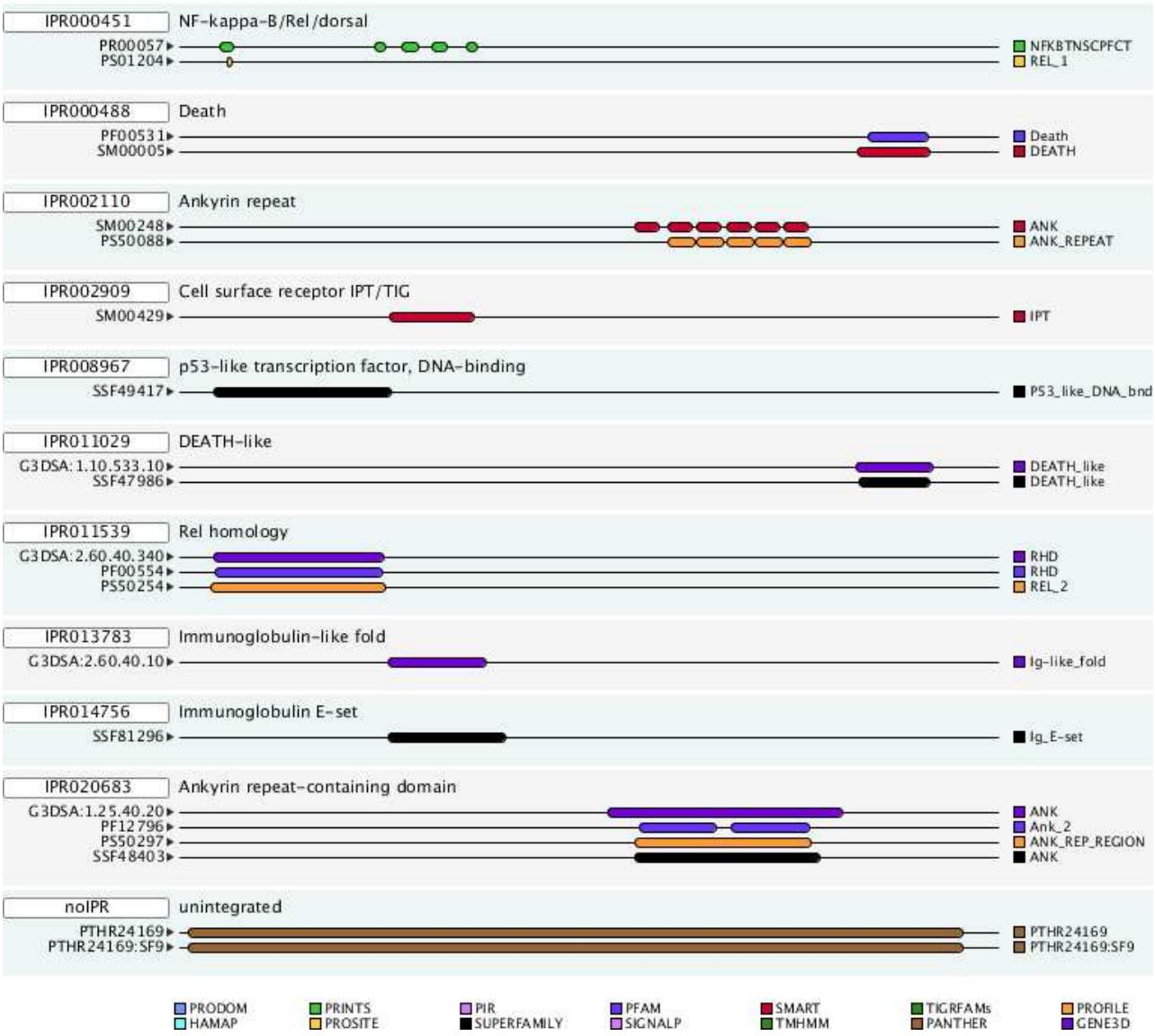


Figure 4C: Showing functional sites and domains of NF-κB from *B. taurus* as predicted using INTERPRO.

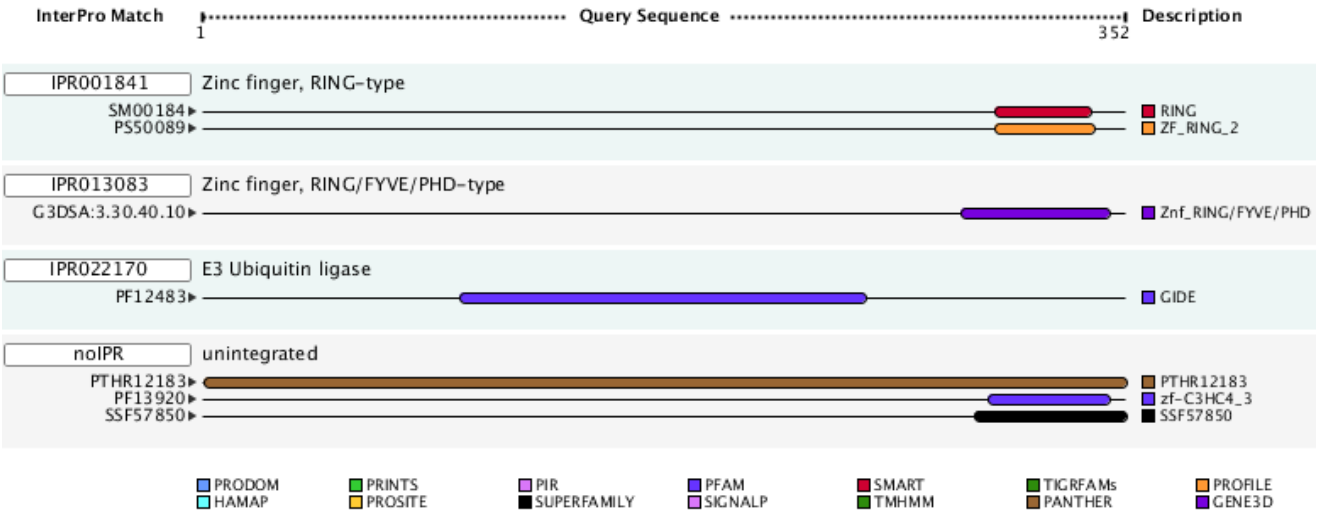


Figure 4D: Showing functional sites and domains of NF-κB from *A. melanoleuca* as predicted using INTERPRO.

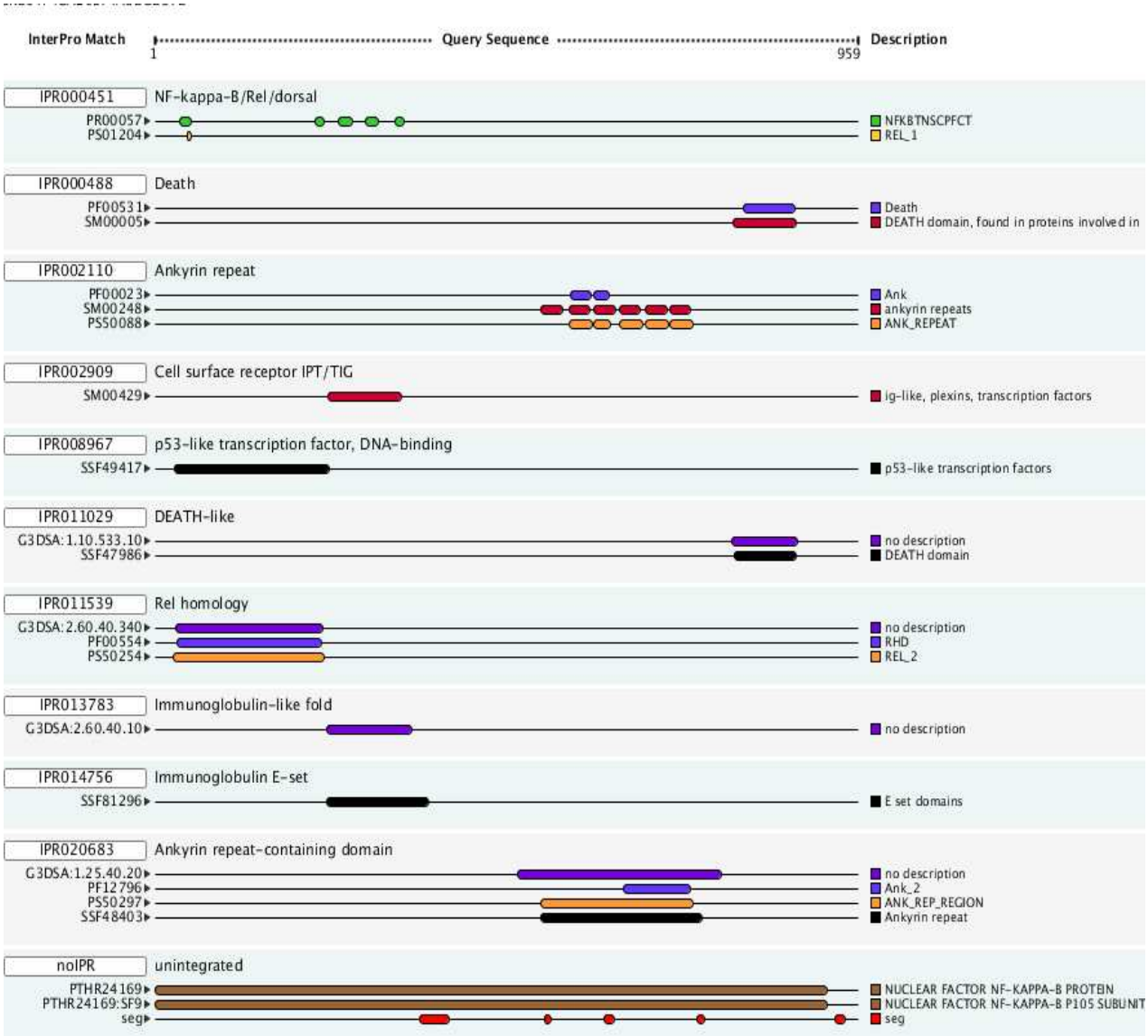


Figure 4E: Showing functional sites and domains of NF-κB from *S. scrofa* as predicted using INTERPRO.

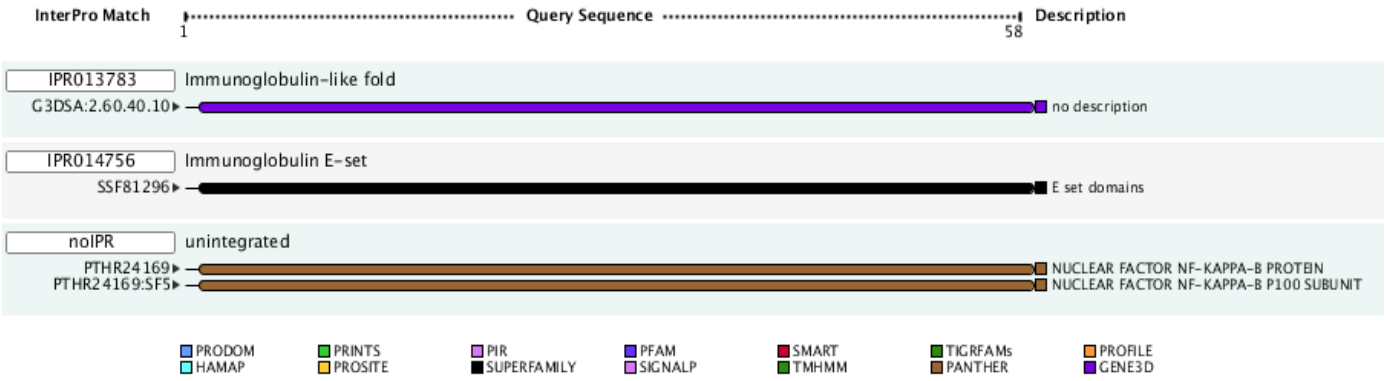


Figure 4F: Showing functional sites and domains of NF-κB from *O. aries* as predicted using INTERPRO.

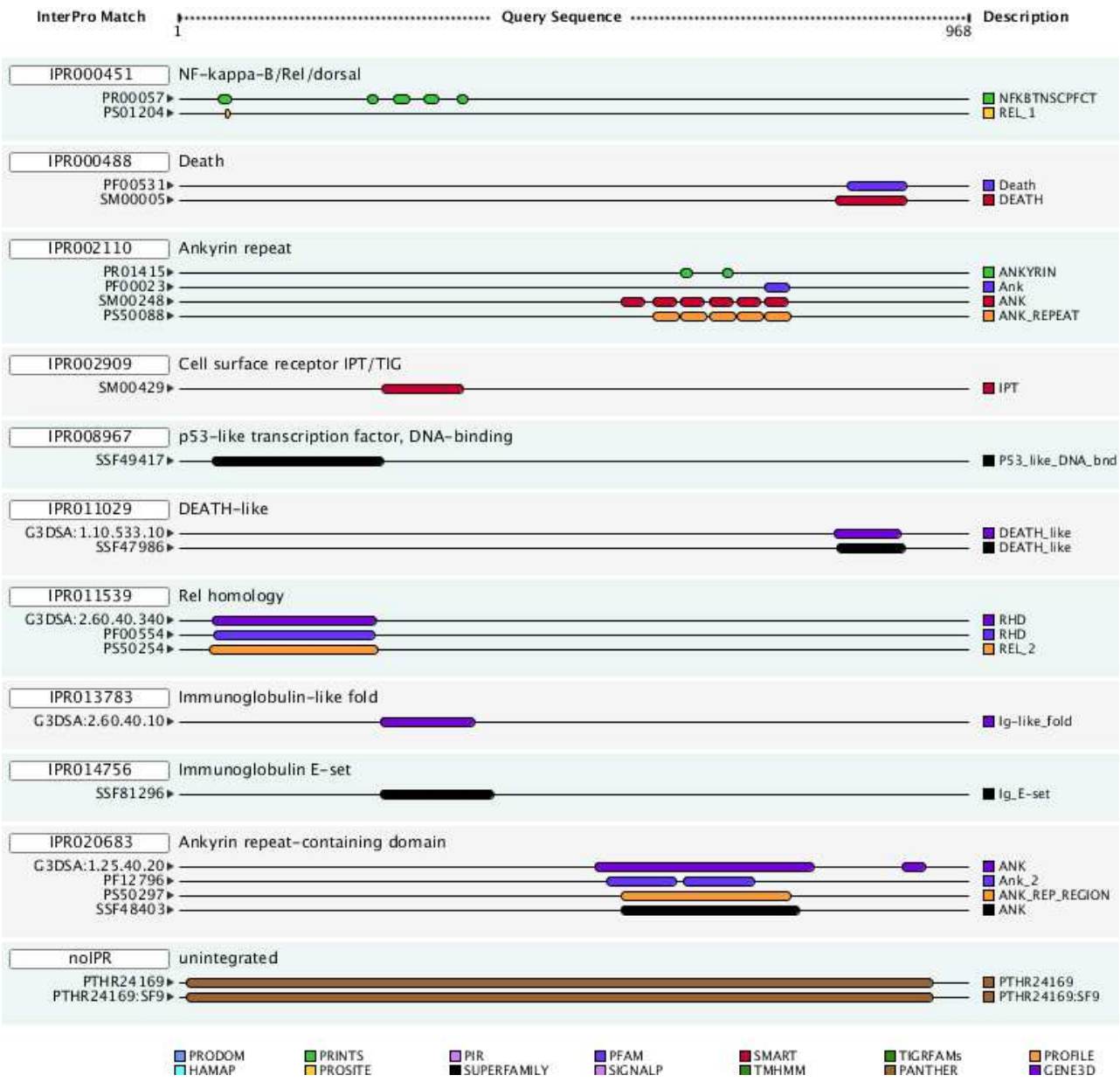


Figure 4G: Showing functional sites and domains of NF-κB from *P. abelii* as predicted using INTERPRO.

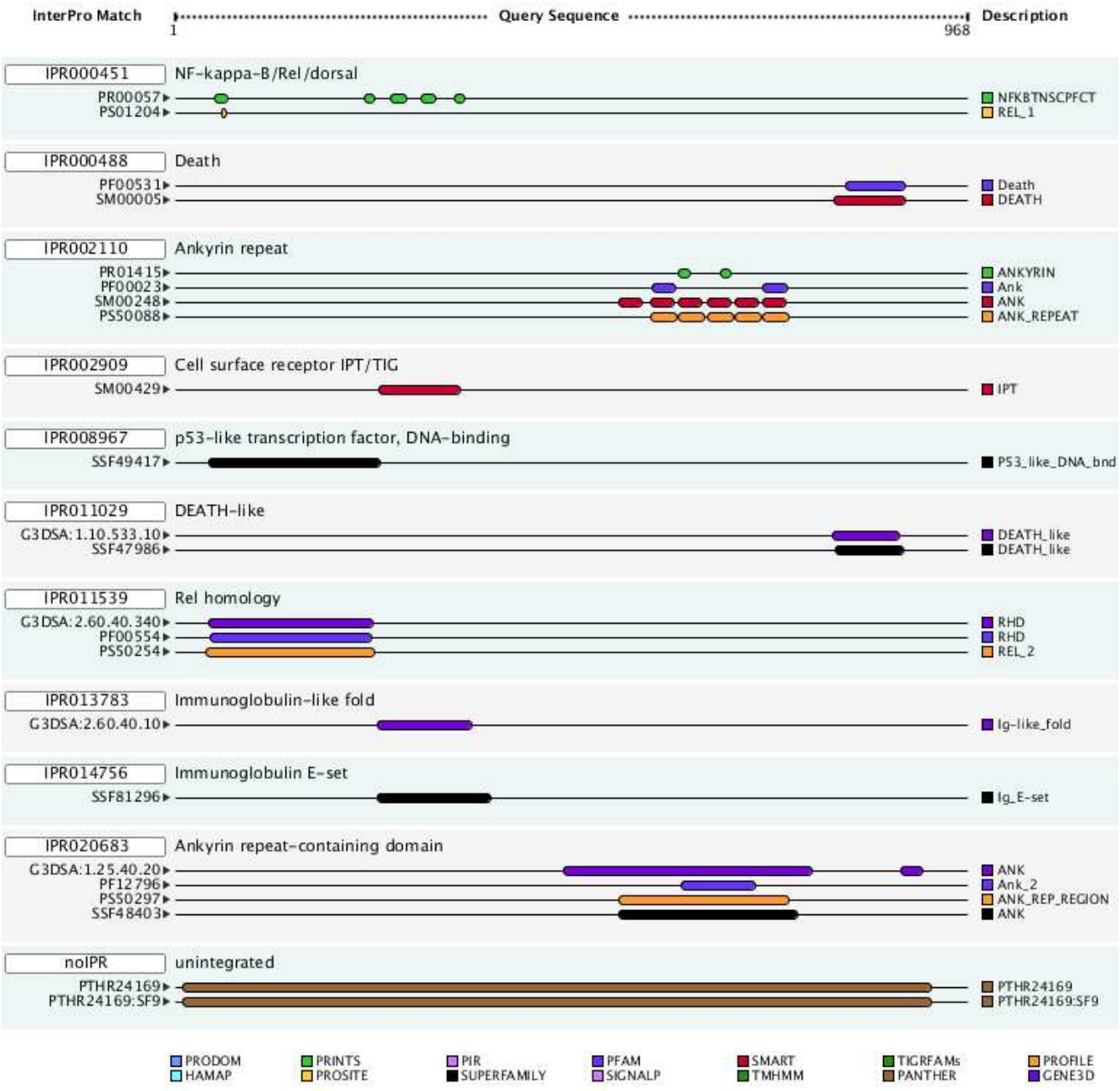


Figure 4H: Showing functional sites and domains of NF-κB from *M. mulatta* as predicted using INTERPRO.

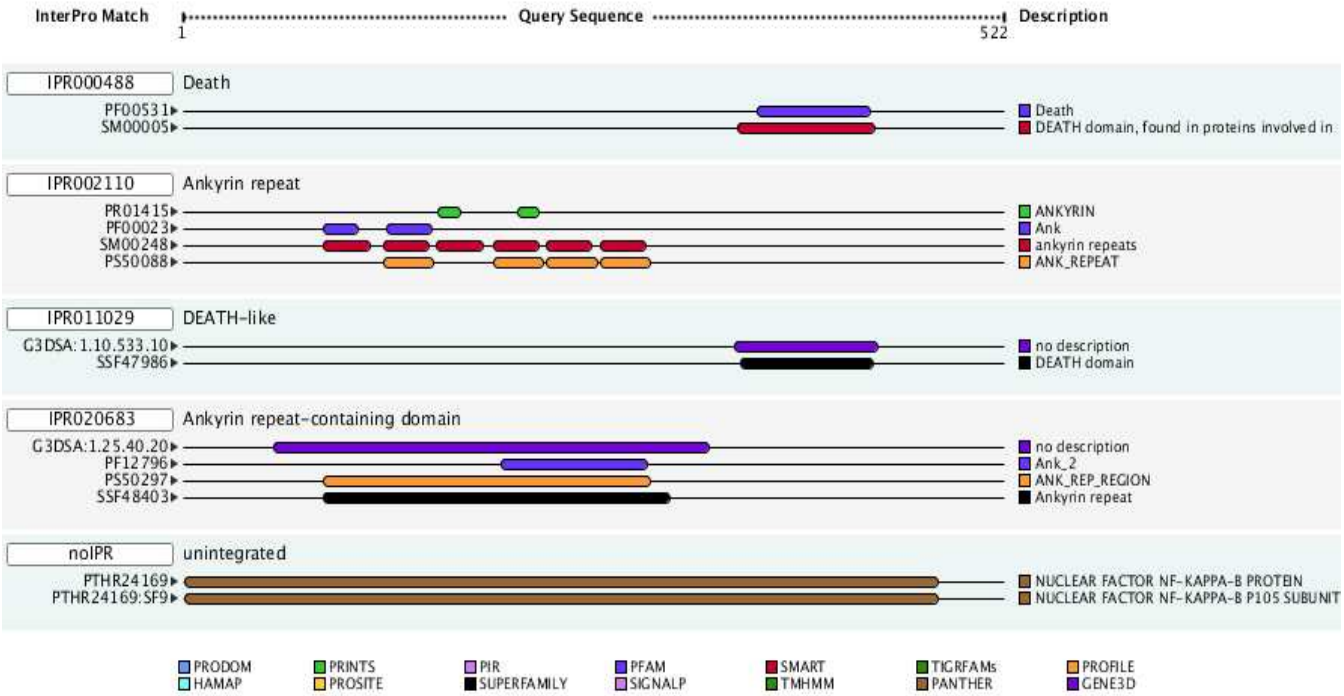


Figure 4I: Showing functional sites and domains of NF-κB from *R. norvegicus* as predicted using INTERPRO.

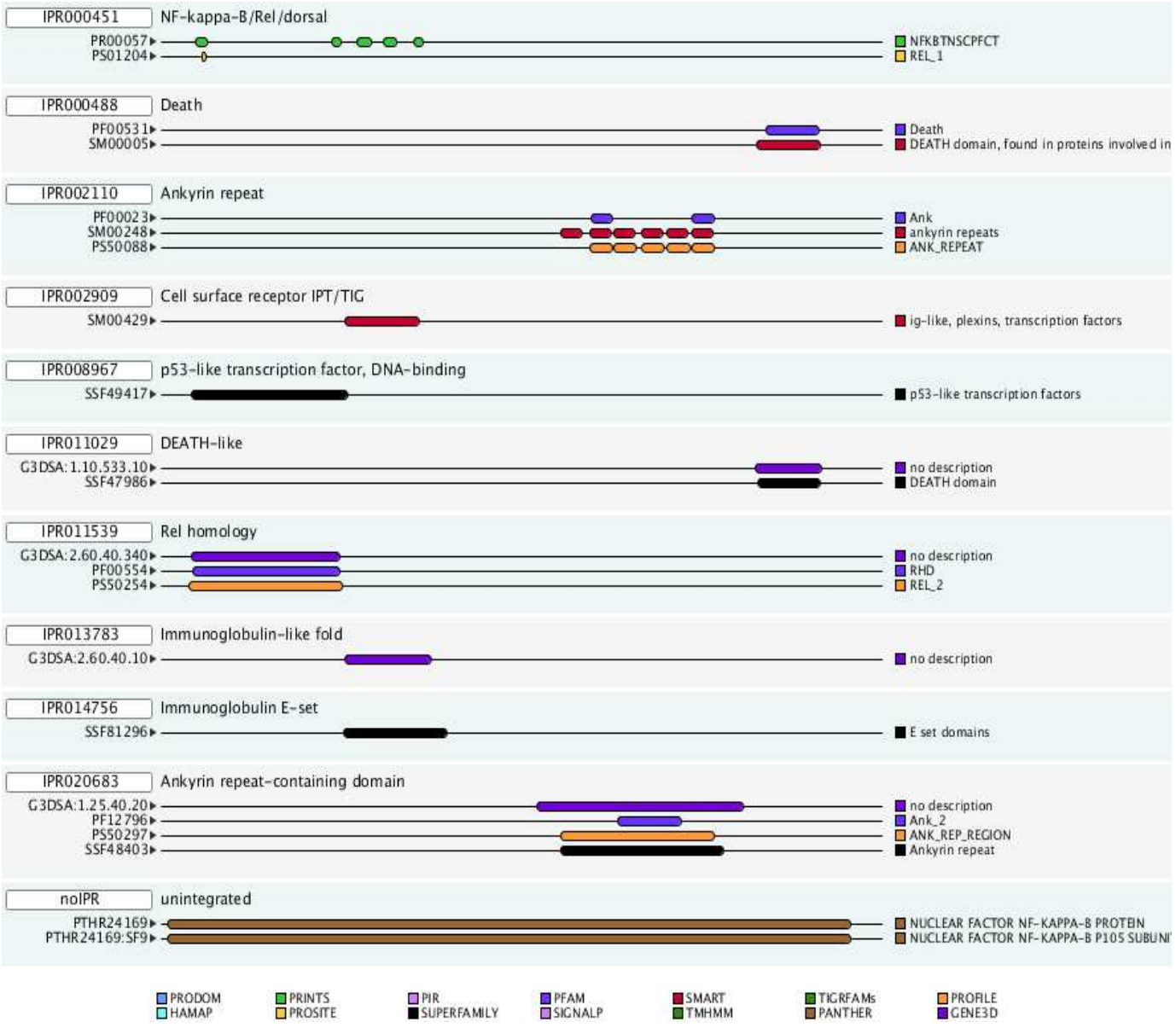


Figure 4J: Showing functional sites and domains of NF-κB from *C. familiaris* as predicted using INTERPRO.

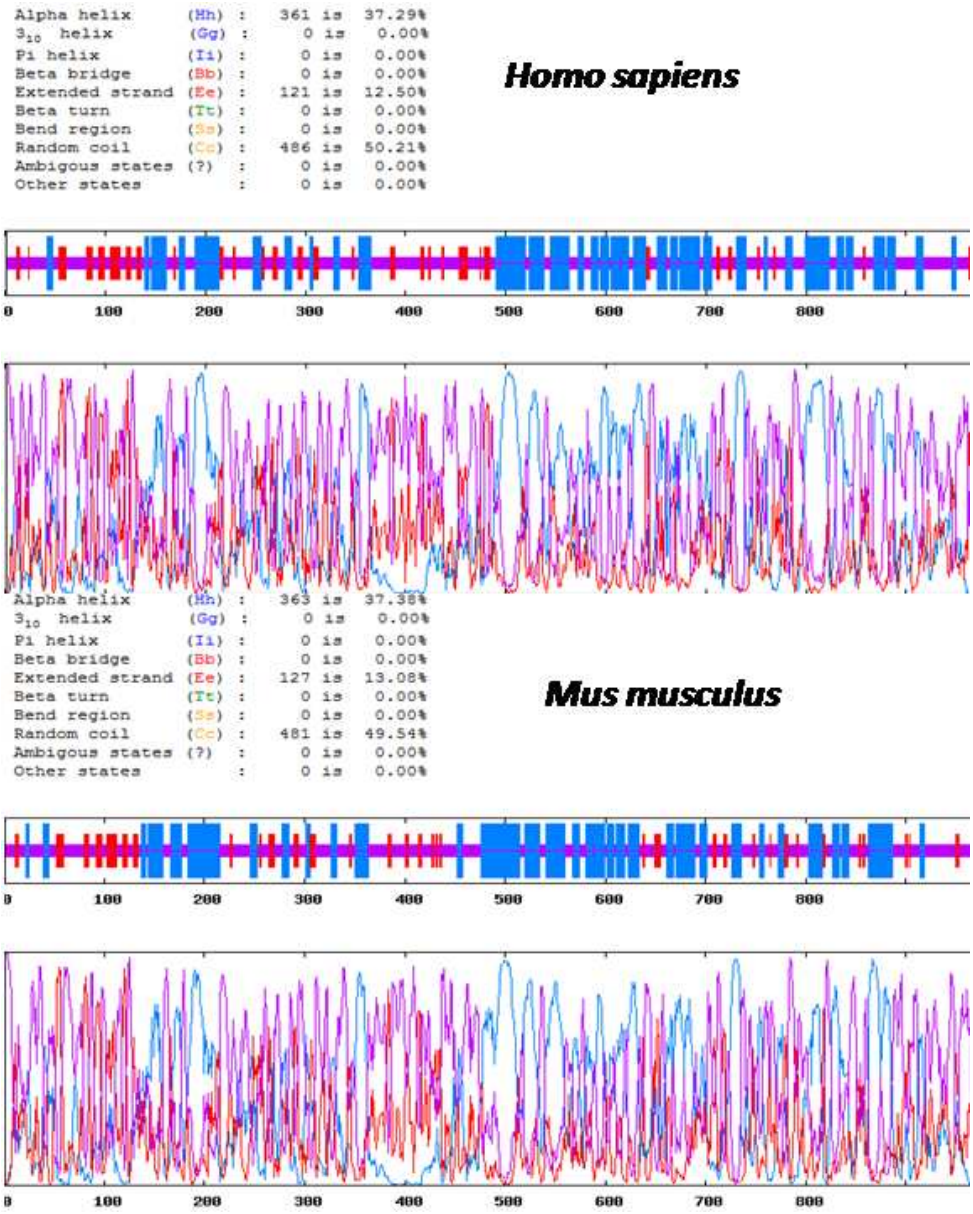


Figure 5A: Secondary structures of NF-κB sequences predicted using GOR

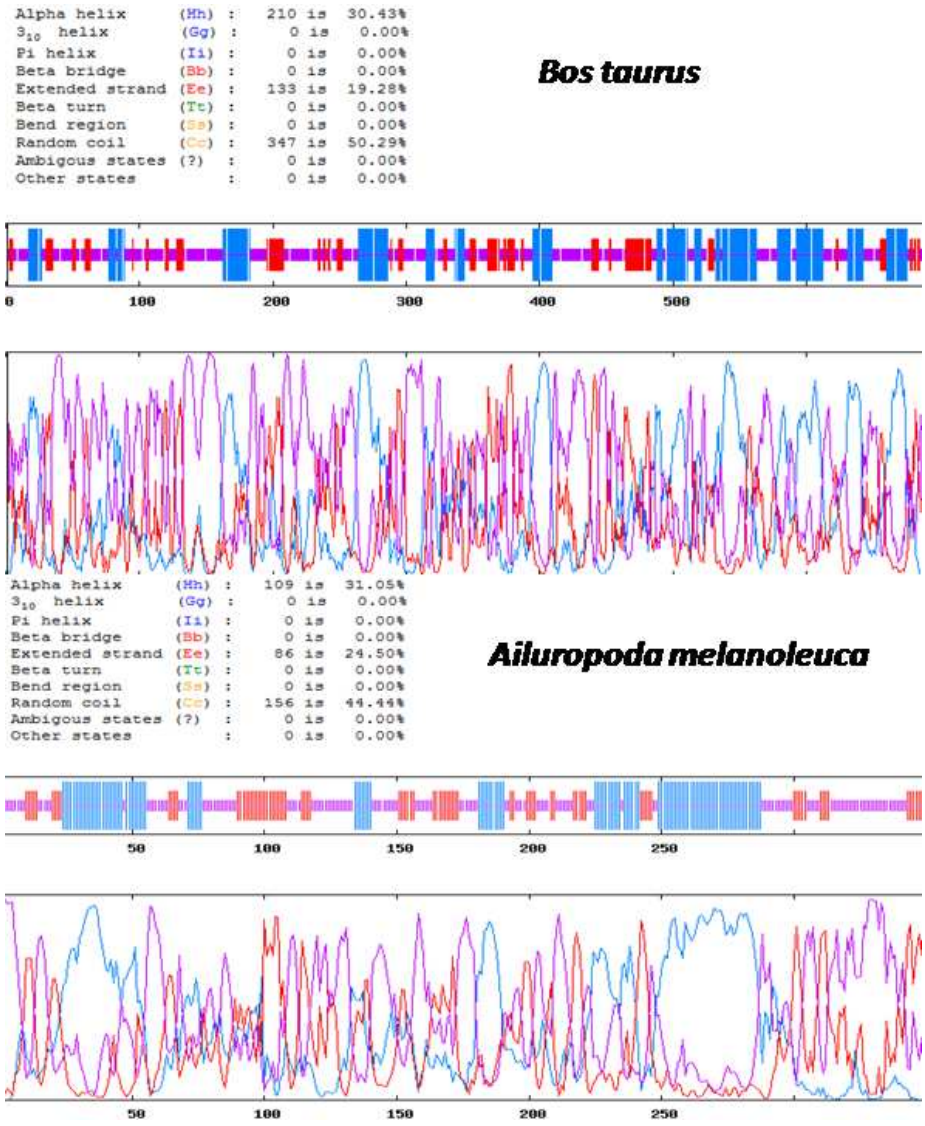


Figure 5B: Secondary structures of NF-κB sequences predicted using GOR

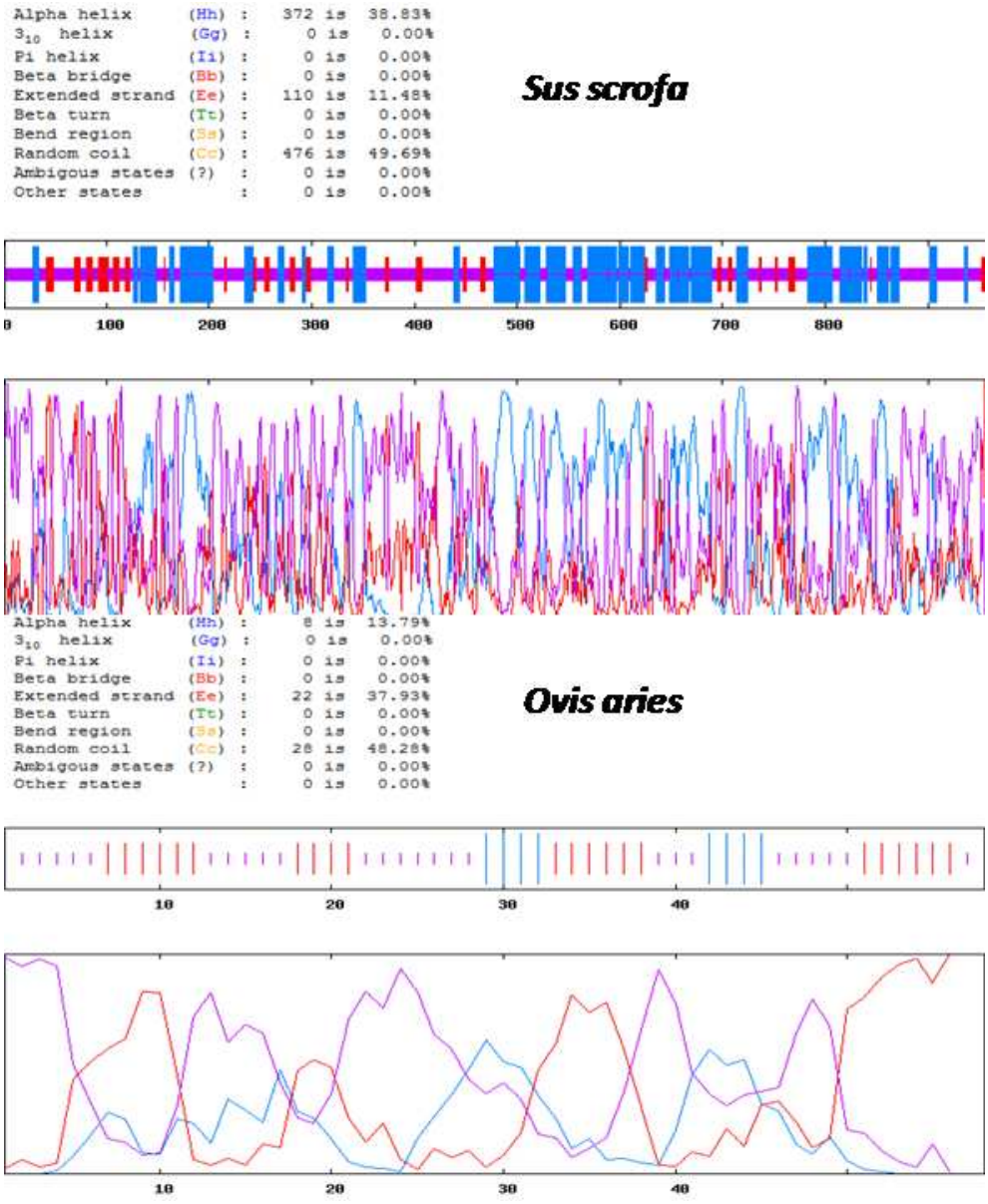


Figure 5C: Secondary structures of NF-κB sequences predicted using GOR

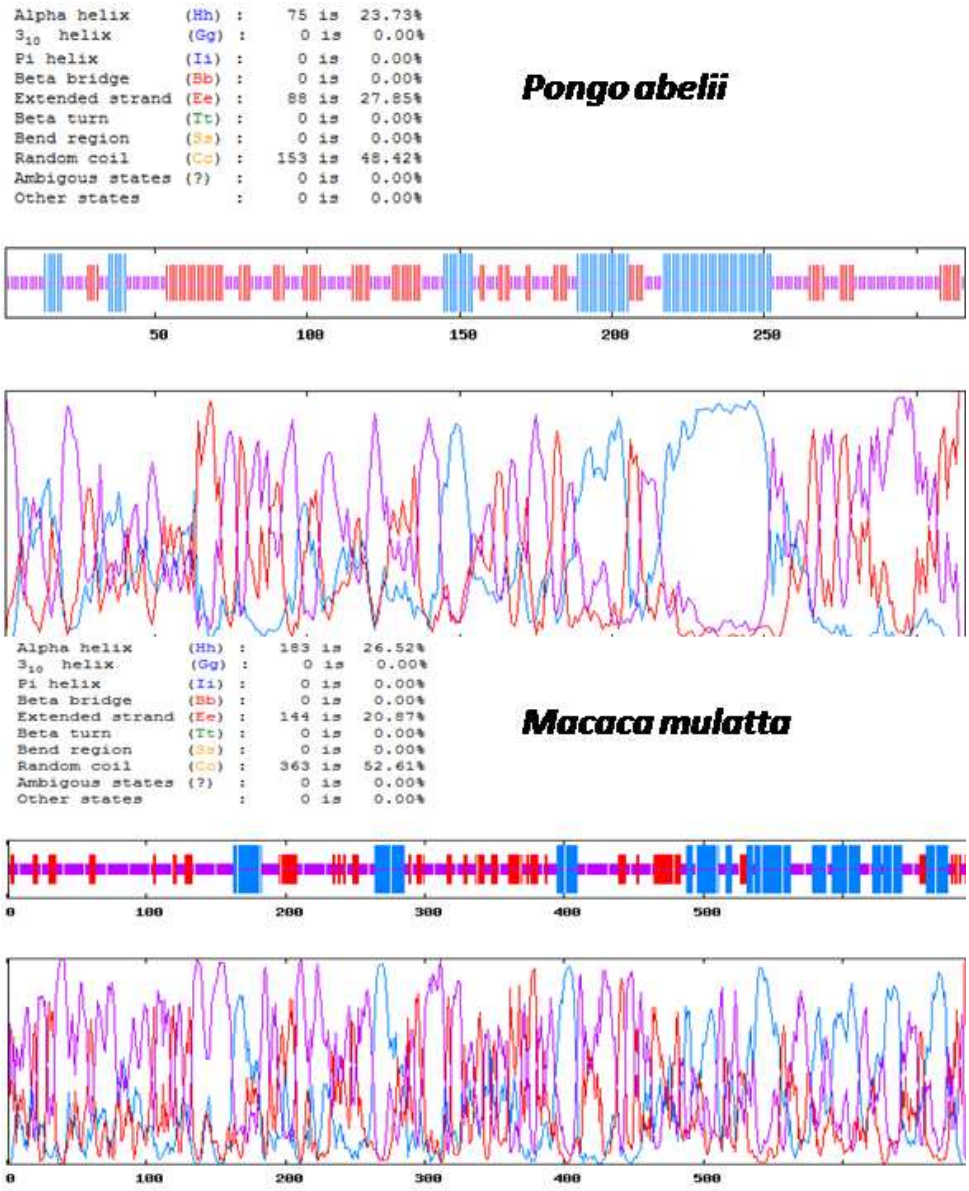


Figure 5D: Secondary structures of NF-κB sequences predicted using GOR

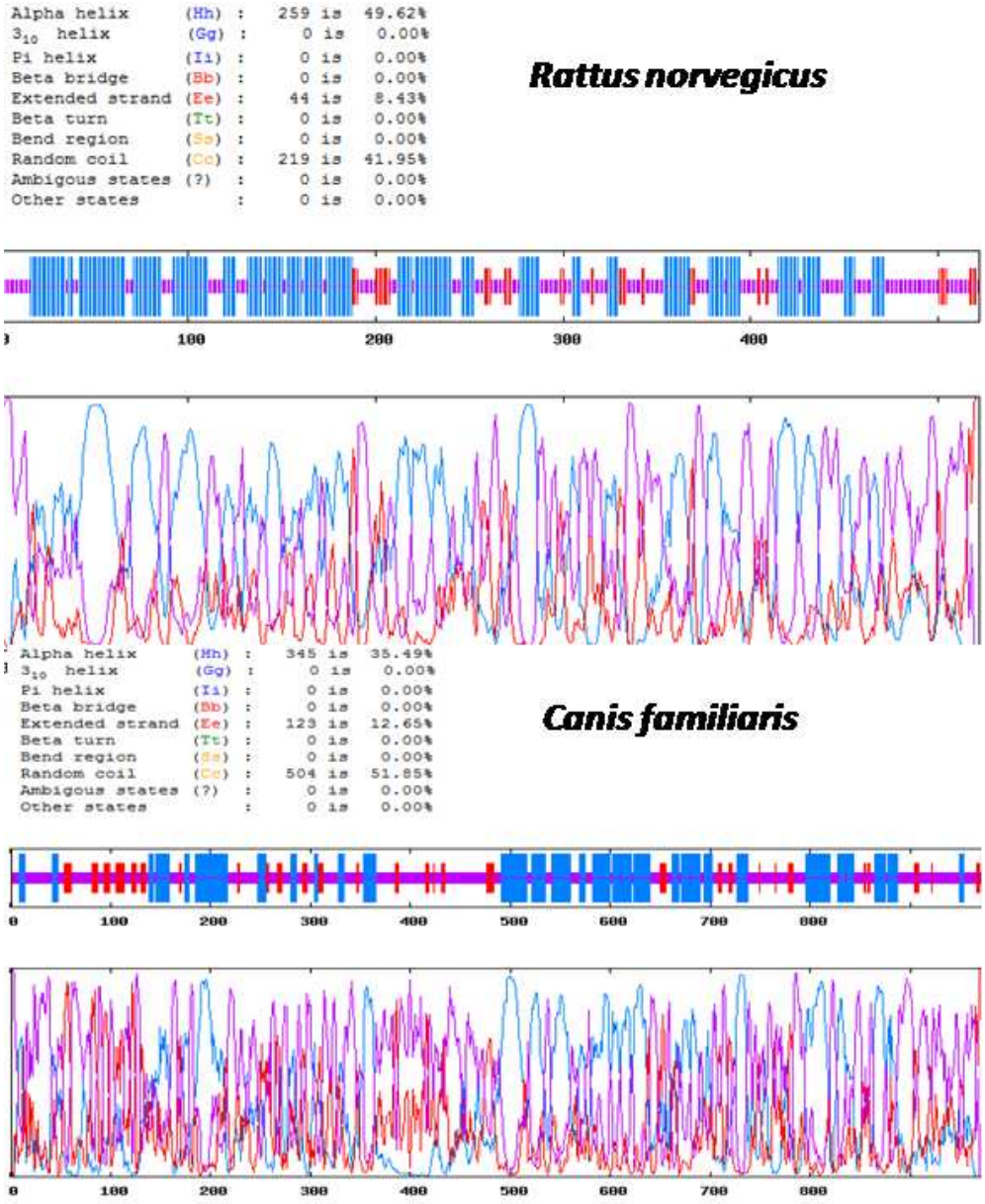


Figure 5E: Secondary structures of NF-κB sequences predicted using GOR

Table 5: Showing the structural details of NF-κB sequences as obtained from PDB

Sl.No.	Organisms	PDB ID	Experimental method	PubMed entry	Resolution[Å]	Polymer-length
1	*Homo sapiens*	2V2T	X-RAY DIFFRACTION	17869269	3.05	RELB-288 P105-326 DNA-11
2	*Mus musculus*	1SVC	X-RAY DIFFRACTION	7830764	2.60	DNA-19 Protein- 365
3	*Bos taurus*	1MSZ	SOLUTION NMR	12547203	--	Protein- 86
4	*Ailuropoda melanoleuca*	3T6P	X-RAY DIFFRACTION	22021857	1.90	Protein- 345
5	*Sus scrofa*	1NFK	X-RAY DIFFRACTION	7530332	2.30	DNA- 11 Protein-325
6	*Ovis aries*	3JV5	X-RAY DIFFRACTION	--	2.65	p100- 104
7	*Pongo abelii*	3EB5	X-RAY DIFFRACTION	18784070	2.00	Protein-74
8	*Macaca mulatta*	3JWE	X-RAY DIFFRACTION	19962385	2.70	Protein- 320
9	*Rattus norvegicus*	2DBF	SOLUTION NMR	--	--	p105- 100
10	*Canis familiaris*	1OOA	X-RAY DIFFRACTION	12886018	2.45	RNA- 29 p105- 326

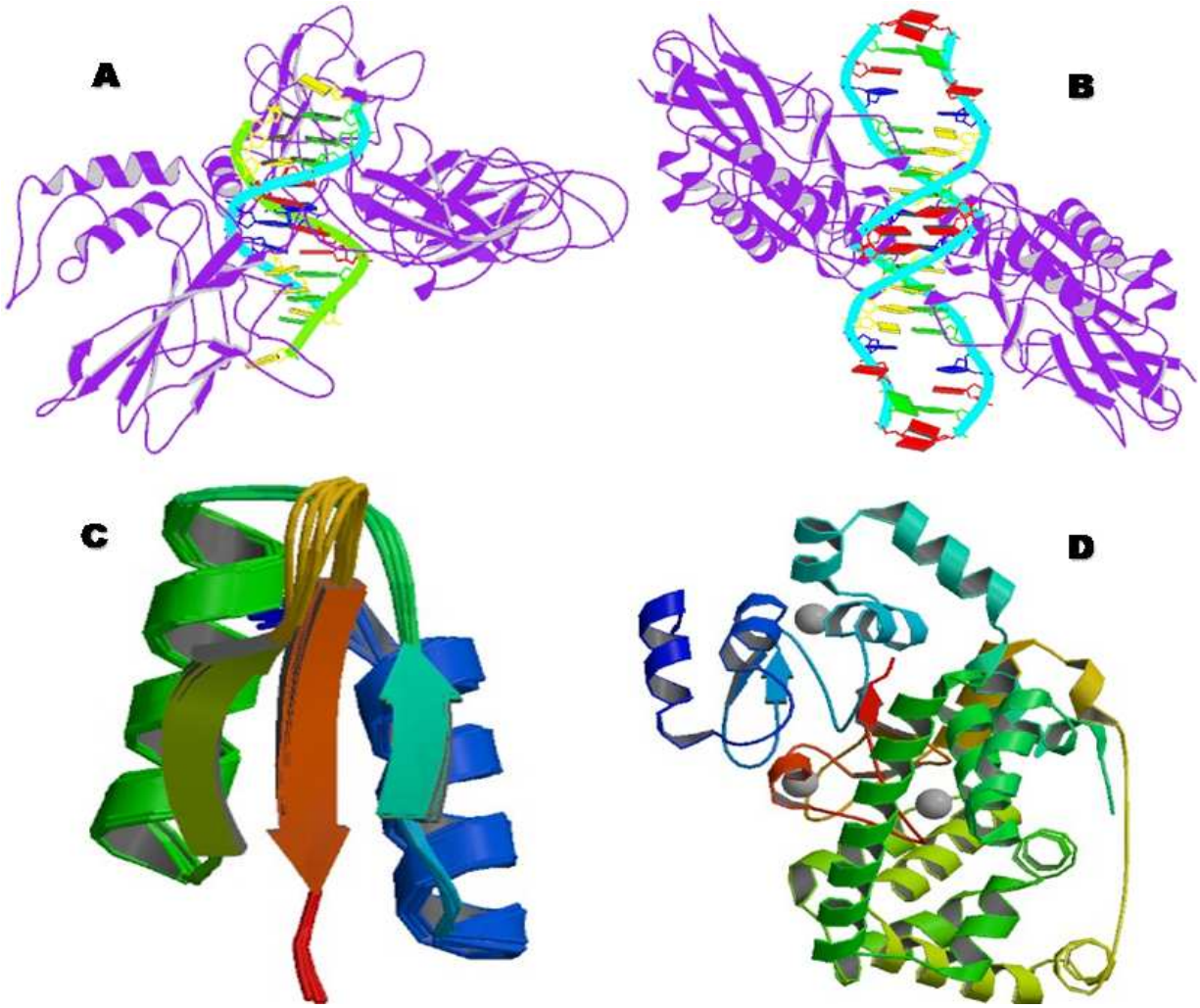


Figure 6A: Three dimensional structures of NF-κB sequences (A)- *H. sapiens;* (B)- *M. musculus;* (C)- *B. Taurus* and (D)- *A. Melanoleuca*

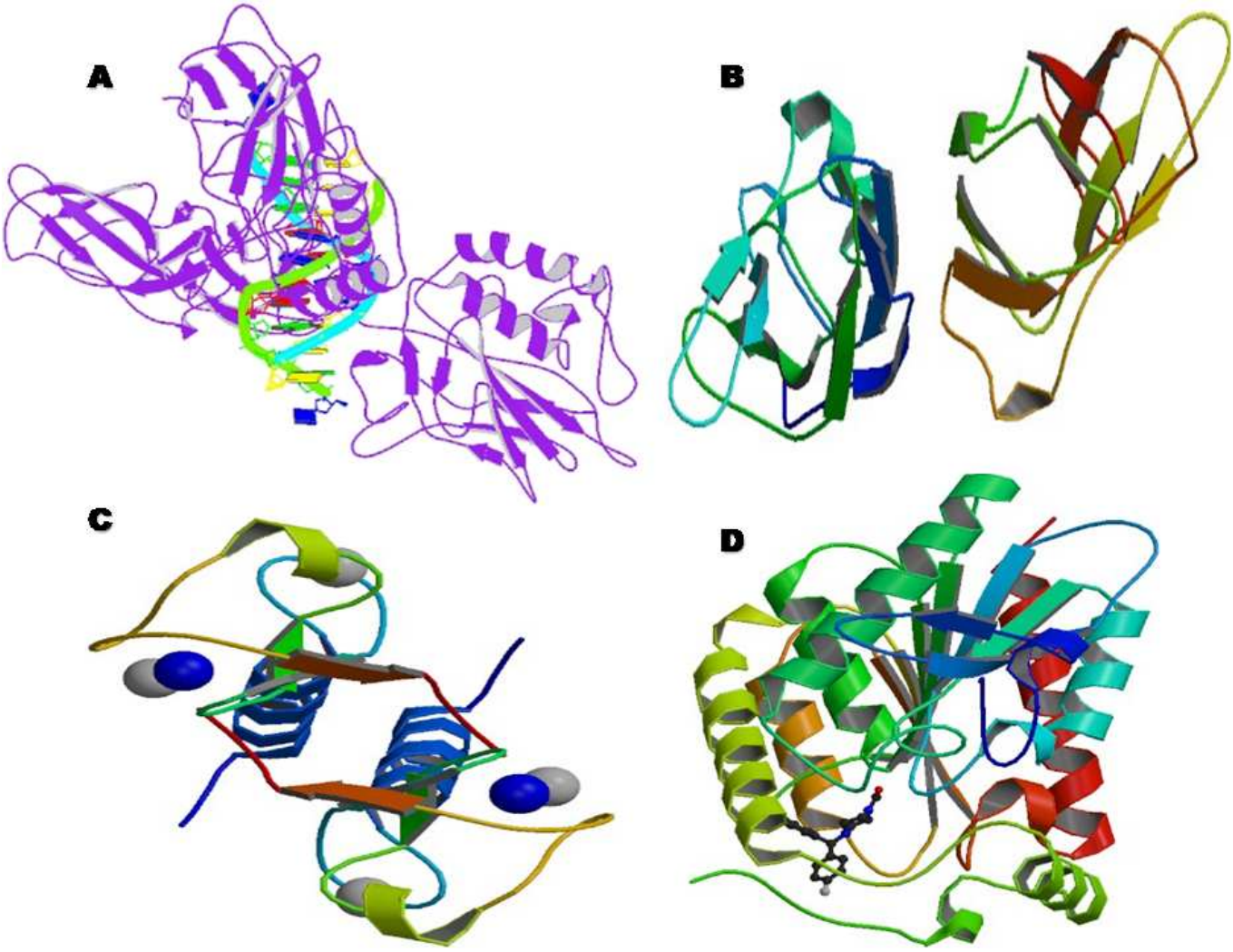


Figure 6B: Three dimensional structures of NF-κB sequences (A)- *S. scrofa;* (B)- *O. aries;* (C)- *P. abelii* and (D)- *M. mulatta.*

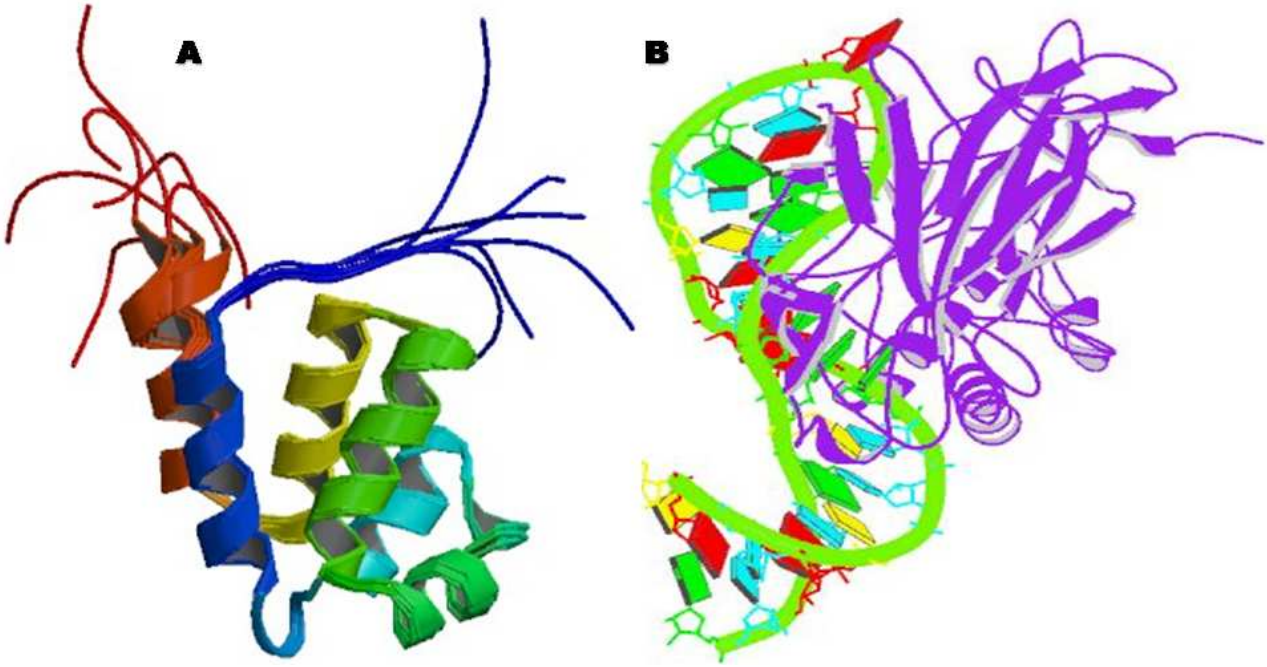


Figure 6C: Three dimensional structures of NF-κB sequences (A)- *R. norvegicus* and (B)- *C. Familiaris*

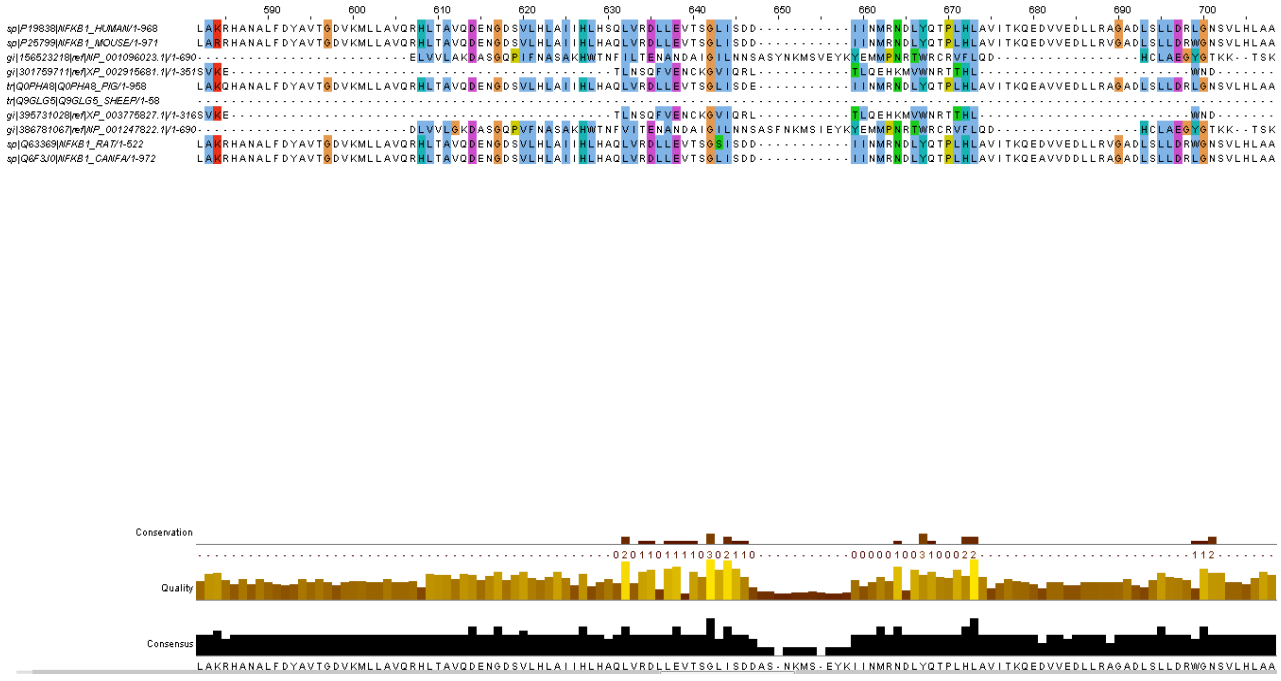


Figure 7: Sequence alignment of NF-κB sequences as obtained in Clustal Omega.

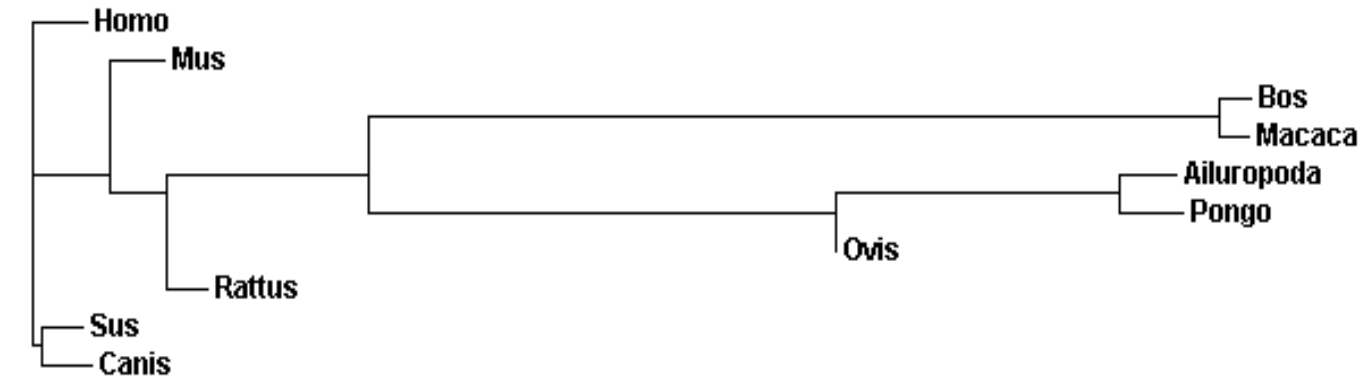


Figure 8: Phylogenetic tree of NF-κB sequences constructed using ClustalW2 Phylogeny (Unrooted).

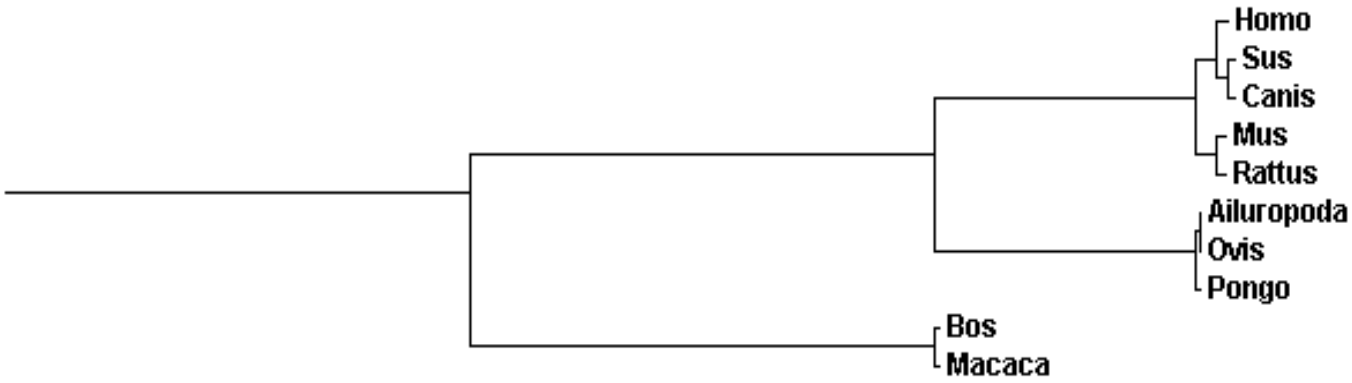


Figure 9: Phylogenetic tree of NF-κB sequences constructed using ClustalW2 Phylogeny (Rooted).

REFERENCES

[1] MK Meffert; JM Chang; BJ Wiltgen; MS Fanselow; D Baltimore. *Nat Neurosci*, **2003**, 6(10), 1072–8.
[2] TD Gilmore. *Oncogene*, **2006**, 25(51), 6680–4.
[3] G Dantas; C Corrent; S Reichow; J Havranek; Z Eletr; N Isern; B Kuhlman; G Varani. *Journal of Molecular Biology*, **2007**, 366(4), 1209.
[4] J John Haddad; E Nisreen Abdel-Karim. *Cellular immunology*, **2011**, 271(1), 5–14.
[5] HB Sieburg. *Studies in the Sciences of Complexity*, **1990**, 12, 321–342.
[6] J Benoit Van den Eynde. *Journal of medicinal chemistry*, **2010**.
[7] W Markus Covert. *Trends in Biochemical science*, **2001**; 26(3), 179-186.
[8] D Nathan Price. *Trends of biotechnology*, **2003**, 21(4), 162 -169.
[9] Nicola Jane Mulder; Manuela Pruess; Rolf Apweiler. *The Handbook of proteomics*, **2005**, 619-627.
[10] Soumen Basak. *Cell*, **2007**, 128(2), 369-381.
[11] T Huxford. *Cell*, **1998**, 98, 759.
[12] FE Chen. *Nature*, **1998**, 391, 410.
[13] S Ghosh. *Annu Rev Immunol*, **1998**, 16, 225-260.