



Computational Analysis of Evolutionary Relationship of a Family of Cold Shock Proteins in Ten Mammalian Species

E. A. Okon^{1*}, E. V. Ikpeme², O. U. Udensi², E. E. Ekerette², H. E. Etta¹,
E. P. Willie² and M. Ozoje²

¹Department of Biological Science, Cross River University of Technology, Calabar, Nigeria.

²Department of Genetics and Biotechnology, University of Calabar, Calabar, Nigeria.

Authors' contributions

This work was carried out in collaboration between all authors. Authors MO, EAO and EVI designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Authors EEE, EPW and OUU managed the analyses of the study. Author HEE managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JABB/2017/36326

Editor(s):

(1) Andrzej Kowalski, Department of Biochemistry and Genetics, Institute of Biology, Jan Kochanowski University, Kielce, Poland.

Reviewers:

(1) J. O. Egahi, Federal University of Agriculture, Nigeria.

(2) G. K. Sivaraman, Central Institute of Fisheries Technology, India.

(3) Marilia Danyelle Nunes Rodrigues, Rural Federal University of the Amazon, Brazil.

Complete Peer review History: <http://www.sciencedomain.org/review-history/22365>

Original Research Article

Received 23rd August 2017
Accepted 21st September 2017
Published 19th December 2017

ABSTRACT

Aims: This study was carried out to evaluate the evolutionary relationship of a family of cold shock proteins (CSP) in ten mammalian species using bioinformatics tools and soft wares such as Genbank, FASTA, BLAST and MEGA 5.

Sample: Twenty protein sequences of both RBM3 and CIRP proteins of some selected mammalian species were downloaded from NCBI database.

Study Design: Computational analysis to evaluate the evolutionary relationship of the CSP was carried out by estimating the phylogenetic relationship of CSP in the different mammalian species studied.

Place and Duration of Study: This study was carried out at the Department of Genetics and Biotechnology, Calabar.

*Corresponding author: E-mail: essyokon@yahoo.com;

Methodology: The molecular evolution and genetic analysis, version 5 (MEGA 5) software was used to determine the evolutionary relationship of both CIRP and RBM3 in the ten mammalian species studied by constructing phylogenetic tree using the amino acid sequences of protein retrieved from NCBI.

Results: The highest identity (100%) was observed between *Ovis aries* and *Bos Taurus*; *Rattus norvegicus* and *Mus-musculus* while the least percentage identity was observed between *Pan troglodytes* and *Bos taurus* (84%). The phylogenetic relationship using UPGMA based on Jones-Taylor-Thornton (JTT) matrix model revealed high relationship.

Conclusion: It was observed that evolutionary relationship of CIRP and RBM3 revealed high relatedness among the mammalian species studied.

Keywords: Cold shock proteins (CSP); cold inducible RNA-binding protein (CIRP); RNA binding motif 3 (RBM3); molecular evolutionary and genetic analysis, version 5 (MEGA 5); National center for biotechnology information (NCBI).

1. INTRODUCTION

All organisms possess the ability to adapt to environmental changes such as the availability of nutrients and also susceptibility to the changes of physical parameters including osmotic pressure and temperature. The most common stress that living organisms constantly confront in nature is likely to result from temperature changes which has resulted in dispersion of organisms to different environments in various parts of the world.

Certain organisms tend to survive in cold regions and some in temperate and hot (desert) regions of the world. Some organisms possess certain proteins called cold shock protein. This protein helps the organism to survive in temperatures lower than the optimum growth temperature. To survive in cold regions, some organisms have developed a capacity for hypothermia in which they lower their body temperature to a few degrees below the optimum level [1–4]. Cold shock proteins (CSPs) comprise of a family of small proteins that are structurally highly conserved and bind to single-stranded nucleic acid through their nucleic acid binding motif (Ribonucleoprotein I and Ribonucleoprotein II). The structural unit characteristic of cold shock protein occurs as a cold shock domain (CSD) and can be found in wide variety of organisms from bacteria to vertebrate.

Cold shock protein in mammals consists of a family of proteins consisting of one amino-terminal consensus sequence RNA-binding domain (CS-RBD) and one carboxyl-terminal glycine-rich domain. This protein family is referred to as a Glycine-rich RNA-binding Protein (GRP) family and some of them have been demonstrated to be induced by cold stress as described by [5–7].

Generally, at reduced temperatures living organisms tend to undergo stress which is accompanied by reduction in metabolic activity. Also protein synthesis is repressed, thereby switching the cellular program from cell growth to cell preservation. This results in the production of a small group of proteins called cold shock proteins (CSPs) which include cold-inducible RNA-binding protein (CIRP) and RNA binding motif protein 3 (RBM3) that increases in hibernating animals [8,9]. Various studies have indicated that CIRP and RBM3 also have important functions in cell protection under general endogenous and environmental stresses at normal temperatures [10].

This study was carried out *in silico* using bioinformatic tools such as the Genbank, FASTA, BLAST and other softwares e. g. MEGA 5 to determine and evaluate by phylogenetic analysis the evolutionary relationship of the family of cold shock proteins (CSPs) in ten mammalian species. The molecular knowledge of the mechanism of expression obtained in this study will help in the selection of suitable species for breeding and other purposes. Studies on phylogenetic of cold shock protein would help to predict the evolutionary changes which could occur with time as well as determine the relationships among the different mammalian species.

2. MATERIALS AND METHODS

2.1 Retrieval of Amino Acids Sequences

The Amino acid sequences of CIRP and RBM3 of *Homo sapiens* (Human), *Sus scrofa* (pig), *pan troglodytes* (chimpanzee), *Ovis aries* (Sheep), *Bos Taurus* (Cattle), *Camelus ferus* (camel), *Canis lupus familiaris* (Dog) *Felis catus*

(Cat) *Rattus norvegicus* (Rat), and *Mus-musculus* were downloaded from Genbank (Tables 1 & 2). The FASTA format of amino acids sequences of these genes were retrieved from the National Center for Biotechnology Information (NCBI) U.S.A database. Also the organism name, Gene ID number, Genbank accession numbers of these proteins sequence as well as sequence lengths were also retrieved.

2.2 Determination of Percentage Identity and Similarity

Percentage identity and similarity (Tables 3, 4, 5 & 6) among the cold inducible RNA binding protein (CIRP) and RNA Binding motif 3 (RBM3) in *Homo sapiens* (Human), *Sus scrofa* (pig), *pan troglodytes* (chimpanzee), *Camelus ferus* (Camel), *Ovis aries* (Sheep), *Bos taurus* (Cattle), *Canis lupus familiaris* (Dog) *Felis catus* (Cat) *Rattus norvegicus* (Rat) and *Mus-musculus* was determined by conducting a multiple sequence comparison of these sequences using BLAST.

2.3 Phylogenetic Analysis and Time of Divergence

The Molecular Evolution and Genetic Analysis, version 5 (MEGA 5) software was used to determine the evolutionary relationship both CIRP and RBM3 in the ten mammalian species studied by constructing phylogenetic tree using the amino acid sequence of protein retrieved from NCBI (Figs. 1 & 2). The evolutionary relationship was inferred using unweighted pair group method with arithmetic mean (UPGMA) analysis based on the Jones-Taylor-Thornton (JTT) matrix-base model. The reliability of the inferred cladogram was evaluated with bootstrap analysis of 1000 replications. The time of divergence of CIRP and RBM3 protein for each species was also calculated.

2.4 Prediction of Protein Domain Architecture

Domain architecture of CIRP and RBM3 proteins of *Homo sapiens* (Human), *Sus scrofa* (pig), *pan troglodytes* (chimpanzee), *Camelus ferus* (Camel), *Ovis aries* (Sheep), *Bos Taurus* (Cattle), *Canis lupus familiaris* (Dog) *Felis catus* (Cat) *Rattus norvegicus* (Rat) and *Mus-musculus* was predicted using SMART software (Tables 7 & 8). The amino acids sequence obtained from NCBI database was used to determine the domain architecture.

3. RESULTS AND DISCUSSION

3.1 Results

3.1.1 Amino acid sequences of CIRPs

The length of amino acid sequences of CIRP is shown in Table 1 ranging from *Homo sapiens* to *Rattus norvegicus* (Rat). The longest sequence length was observed in *Ovis aries* (Sheep) being 313 amino acids while the shortest amino acid length of 83 amino acids was observed in *Camelus ferus* (Camel). The amino acid sequence of *Homo sapiens*, *Mus-musculus* and *Rattus norvegicus* were observed to be the same (172 amino acids).

3.1.2 Amino acid sequences of RBM3 proteins

The length of amino acid sequences of RBM3 is shown in Table 2 for the different mammalian species studied. It varied from 90 amino acids (found in cattle) to 162 amino acids (found in Dog). However, the same length of amino acid sequence of 160 amino acids was observed in *Ovis aries* (Sheep) and *Felis catus* (cat). *Homo sapiens* (human) and *Pan troglodytes* (Chimpanzee) also recorded similar length of 157 amino acids. The shortest length of amino acid sequences of 90 amino acids was observed in *Bos taurus* (Cattle) while *Canis lupus familiaris* (Dog) had the longest sequence length of 162 amino acids.

3.1.3 Percentage identity among the amino acid sequence of CIRPs

The percentage identity of CIRP is tabulated in Table 3 and revealed high identity among the amino acid sequences. The highest identity (100%) was observed between *Ovis aries* and *Bos Taurus*; *Rattus norvegicus* and *Mus-musculus* while the least percentage identity was observed between *Pan troglodytes* and *Bos taurus* (84%). *Sus scrofa* (Pig) with *Bos taurus*, *Canis lupus familiaris* (Dog) and *Ovis aries* (Sheep); *Ovis aries* with *Felis catus* (cat), *Canis lupus familiaris* all shared 99% identity. *Homo sapiens*, *Felis catus*, *Bos taurus*, *Canis lupus familiaris*, *Camelus ferus* (Camel) and *Ovis aries* shared 98% identity.

3.1.4 Percentage similarity among the amino acid sequence of CIRP

The percentage similarity is tabulated in Table 4 and revealed the highest similarity of 100%

between *Pan troglodytes* and *Homo sapiens*; *Sus scrofa* and *Mus-musculus*; *Ovis aries*, *Rattus norvegicus*, *Camelus ferus* and; *Felis catus* and *Canis lupus familiaris*; *Canis lupus familiaris* and *Felis catus*; *mus-musculus*, *Rattus norvegicus*, *Pan troglodytes* and *Homo sapiens*. The least observed similarity of 53% was seen between *Ovis aries* and *Felis catus*, *Mus-musculus*, *Bos taurus*, *Canis lupus familiaris*, *Pan troglodytes*, *Sus scrofa* and *Homo sapiens*.

3.1.5 Percentage identity among the amino acid sequence of RBM3

The percentage identity of RBM3 protein obtained is tabulated in Table 5 and revealed the highest identity between *Pan troglodytes* and *Homo sapiens*. *Ovis aries* and *Bos taurus* which shared 99% identity; *Rattus norvegicus* and *Mus-musculus* shared 98% identity; *Pan troglodytes* and *Rattus norvegicus*; *Rattus norvegicus*, *Pan troglodytes* and *Homo sapiens* all shared 97% identity. *Canis lupus familiaris* and *Felis catus*; *Felis catus*, *Rattus norvegicus* and *Canis lupus familiaris* all shared 96% identity. The least identity was observed between *Camelus ferus* and *Sus scrofa* (87%).

3.1.6 Percentage similarity among the amino acid sequence of RBM3

The percentage similarity of RBM3 obtained is tabulated in Table 6 and revealed great similarities between the mammalian species studied. *Homo sapiens*, *Sus scrofa*, *Ovis aries*, *Pan troglodytes*, *Camelus ferus*, *Canis lupus familiaris*, *mus-musculus*, *Felis catus* and *Rattus norvegicus* was observed to share the highest similarity (100%). However, the least percentage similarity was observed between *Bos taurus* and *Rattus norvegicus* (82%).

3.1.7 Domain architecture of CIRP

The domain architecture of *Homo sapiens*, *Pan troglodytes*, *Sus scrofa*, *Canis lupus familiaris*, *Felis catus*, *Ovis aries*, *Bos taurus*, *Camelus ferus*, *Rattus norvegicus* and *Mus-Musculus* were predicted using SMART software and tabulated in Table 7. CIRP domain contains mainly two types of domain namely: The RNA Recognition Motif and low complexity region. CIRP domain in human was observed at position 7-80 with low complexity region at 89-125. Camel CIRP was at position 8-81. Rat and mouse CIRP

domain was at position 7-80 with low complexity region at 89-168.

3.1.8 Domain architecture of RBM3

RBM3 contain mainly two types of domain namely: RNA Recognition Motif and low complexity region. According to Table 8; RBM3 domain in Human was observed at position 7 -80 with low complexity region at 82-135. Chimpanzee RBM3 domain was observed at position 7-80 with low complexity region at 82-138. Cattle RBM3 domain was observed at position 7-82.

3.1.9 Phylogenic relationship of CIRP proteins

The evolutionary relationship among the mammalian species CIRP and RMB3 is shown in Figs. 1 and 2. It was determined by Unweighted Pair Group Method with Arithmetic mean (UPGMA) analysis based on the Jones Taylor-Thornton (JTT) matrix-based model. In Fig. 1 the data revealed that *Mus musculus* was more related to *Rattus norvegicus* sharing the same clade and having 67G consensus. *Bos taurus* and *Homo sapiens* clustered in the same clade having 14G consensus. *Canis lupus familiaris* and *Ovis aries* clustered in the same clade having 17G consensus. The same Fig. 1 also showed that *Homo sapiens* and *Bos taurus* shared the same lineage with *Ovis aries* and *Canis lupus familiaris* with 28G consensus but with longer time of divergence.

3.1.10 Phylogenic relationship of RBM3 proteins

The evolutionary relationship among the mammalian species of RMB3 is shown in Fig. 2. Phylogenetic analysis of RBM3 protein revealed that *Homo sapiens* RBM3 protein was more related to *Pan troglodytes* which clustered in the same clade. *Felis catus* and *Ovis aries* also clustered in the same clade.

3.2 Discussion

Changes in the DNA and proteins caused by mutations are essential to evolution. The building blocks of these biological macromolecules, nucleotide bases, and amino acids form linear sequences determine the primary structure of the molecules [11].

Table 1. Retrieved proteins sequence of CIRBP gene of the mammalian species with their accession numbers and sequence length obtained from Genbank, National Center for Biotechnology Information (NCBI)

Species names	Gene name	Gene ID / accession number	protein name	GenBank accession number for protein	Amino acid sequence length
<i>Homo sapiens</i> (Human)	CIRBP mRNA	1153 NC_000019.10	CIRP	AAC04895.1	172
<i>Sus scrofa</i> (Pig)	CIRBP mRNA	100522851 NC_010444.3	CIRP	NP_001233126.1	144
<i>Ovis aries</i> (Sheep)	CIRBP mRNA	101117758 NC_019462.2	CIRP	XP_012033837.1	313
<i>Pan troglodytes</i> (Chimpanzee)	CIRBP mRNA	742386 NC_006486.4	CIRP	NP_001267442.1	225
<i>Camelus ferus</i> (Camel)	CIRBP mRNA	102522728 NW_006218594.1	CIRP	XP_006195343.2	83
<i>Canis lupus familiaris</i> (Dog)	CIRBP mRNA	476755 NC_006602.3	CIRP	XP_013977603.1	185
<i>Bos Taurus</i> (Cattle)	CIRBP mRNA	507120 AC_000164.1	CIRP	NP_001029450.1	213
<i>Mus-musculus</i> (Mouse)	CIRBP mRNA	12696 NC_000076.6	CIRP	BAA11213.1	172
<i>Felis catus</i> (Cat)	CIRBP mRNA	101085782	CIRP	XP_006928229.2	209
<i>Rattus norvegicus</i> (Rat)	CIRBP mRNA	81825 NC_005106.4	CIRP	BAA19092.1	172

Table 2. Retrieved proteins sequence of RBM3 gene of the mammalian species with their accession numbers and sequence length obtained from Genbank, National Center for Biotechnology Information (NCBI)

Species names	Gene name	Gene ID/ accession number	protein name	GenBank accession number for protein	Amino acid sequence length
<i>Homo sapiens</i> (Human)	RBM3 mRNA	5935 NC_000023.1	RBM3	NP_006734.1	157
<i>Sus scrofa</i> (Pig)	RBM3 mRNA	100627807 NC_010461.4	RBM3	NP_001230348.1	156
<i>Ovis aries</i> (Sheep)	RBM3 mRNA	101110819 NC_019484.2	RBM3	XP_011962149.1	160
<i>Pan troglodytes</i> (Chimpanzee)	RBM3 mRNA	465617 NC_006491.4	RBM3	JAA11238.1	157
<i>Camelus ferus</i> (Camel)	RBM3 mRNA	102519265 NW_006211533.1	RBM3	XP_006190470.1	159
<i>Canis lupus familiaris</i> (Dog)	RBM3 mRNA	609457 NC_006621.3	RBM3	XP_013967121.1	162
<i>Bos Taurus</i> (Cattle)	RBM3 mRNA	509771 AC_000187.1	RBM3	AAI03474.1	90
<i>Mus-musculus</i> (Mouse)	RBM3 mRNA	19652 NC_000086.7	RBM3	NP_001280587.1	154
<i>Felis catus</i> (Cat)	RBM3 mRNA	101095080 NC_018741.2	RBM3	XP_011290011.1	160
<i>Rattus norvegicus</i> (Rat)	RBM3 mRNA	114488 NC_005120.4	RBM3	NP-446148.1	156

Table 3. Percentage identity (%) of CIRP protein among the mammalian species studied using BLAST

Species names	<i>Homo sapiens</i> (Human)	<i>Sus scrofa</i> (Pig)	<i>Ovis aries</i> (Sheep)	<i>Pan troglodytes</i> (Chimpanzee)	<i>Camelus ferus</i> (Camel)	<i>Canis lupus familiaris</i> (Dog)	<i>Bos taurus</i> (Cattle)	<i>Mus-musculus</i> (Mouse)	<i>Felis catus</i> (Cat)	<i>Rattus norvegicus</i> (Rat)
<i>Homo sapiens</i> (Human)	100	97	98	97	98	98	98	95	98	95
<i>Sus scrofa</i> (Pig)	98	100	99	97	98	99	99	97	98	96
<i>Ovis aries</i> (Sheep)	98	99	100	98	98	99	100	96	99	96
<i>Pan troglodytes</i> (Chimpanzee)	97	96	84	100	98	94	85	93	93	94
<i>Camelus ferus</i> (Camel)	98	98	98	98	100	98	98	93	97	93
<i>Canis lupus familiaris</i> (Dog)	98	97	99	98	98	100	97	96	97	96
<i>Bos Taurus</i> (Cattle)	98	98	100	98	98	97	100	96	93	96
<i>Mus-musculus</i> (Mouse)	95	97	96	93	93	96	96	100	95	100
<i>Felis catus</i> (Cat)	98	98	99	98	93	97	93	95	100	95
<i>Rattus norvegicus</i> (Rat)	95	96	96	93	93	96	96	100	95	100

Table 4. Percentage similarity of CIRP protein among the mammalian species studied using BLAST

Species names	<i>Homo sapiens</i> (Human)	<i>Sus scrofa</i> (Pig)	<i>Ovis aries</i> (Sheep)	<i>Pan troglodytes</i> (Chimpanzee)	<i>Camelus ferus</i> (Camel)	<i>Canis lupus familiaris</i> (Dog)	<i>Bos Taurus</i> (Cattle)	<i>Mus-musculus</i> (Mouse)	<i>Felis catus</i> (Cat)	<i>Rattus norvegicus</i> (Rat)
<i>Homo sapiens</i> (Human)	100	97	97	100	99	97	97	100	98	100
<i>Sus scrofa</i> (Pig)	99	100	99	98	97	99	99	100	99	97
<i>Ovis aries</i> (Sheep)	53	53	100	53	100	53	53	53	53	100
<i>Pan troglodytes</i> (Chimpanzee)	100	74	84	100	97	82	89	78	82	78
<i>Camelus ferus</i> (Camel)	97	97	97	97	100	97	97	99	84	99
<i>Canis lupus familiaris</i> (Dog)	90	94	90	90	97	100	95	97	100	97
<i>Bos Taurus</i> (Cattle)	77	77	77	77	97	82	100	77	87	77
<i>Mus-musculus</i> (Mouse)	100	100	97	100	99	97	97	100	95	100
<i>Felis catus</i> (Cat)	79	88	79	80	84	88	88	95	100	95
<i>Rattus norvegicus</i> (Rat)	100	97	97	100	99	97	97	100	95	100

Table 5. Percentage identity (%) of RBM3 protein among the mammalian species studied using BLAST

Species names	<i>Homo sapiens</i> (Human)	<i>Sus scrofa</i> (Pig)	<i>Ovis aries</i> (Sheep)	<i>Pan troglodytes</i> (Chimpanzee)	<i>Camelus ferus</i> (Camel)	<i>Canis lupus familiaris</i> (Dog)	<i>Bos taurus</i> (Cattle)	<i>Mus-musculus</i> (Mouse)	<i>Felis catus</i> (Cat)	<i>Rattus norvegicus</i> (Rat)
<i>Homo sapiens</i> (Human)	100	89	89	100	91	91	88	94	93	97
<i>Sus scrofa</i> (Pig)	89	100	91	89	93	91	91	94	93	94
<i>Ovis aries</i> (Sheep)	89	90	100	93	93	90	99	90	94	91
<i>Pan troglodytes</i> (Chimpanzee)	100	89	89	100	91	91	88	94	93	97
<i>Camelus ferus</i> (Camel)	91	87	92	91	100	91	91	90	91	93
<i>Canis lupus familiaris</i> (Dog)	92	90	91	92	91	100	90	93	96	94
<i>Bos Taurus</i> (Cattle)	89	90	91	89	91	90	100	89	90	93
<i>Mus-musculus</i> (Mouse)	92	94	91	92	91	93	91	100	94	98
<i>Felis catus</i> (Cat)	93	94	94	93	94	96	94	94	100	96
<i>Rattus norvegicus</i> (Rat)	97	94	91	97	92	94	93	98	96	100

Table 6. Percentage similarity of RBM3 protein among the mammalian species studied using BLAST

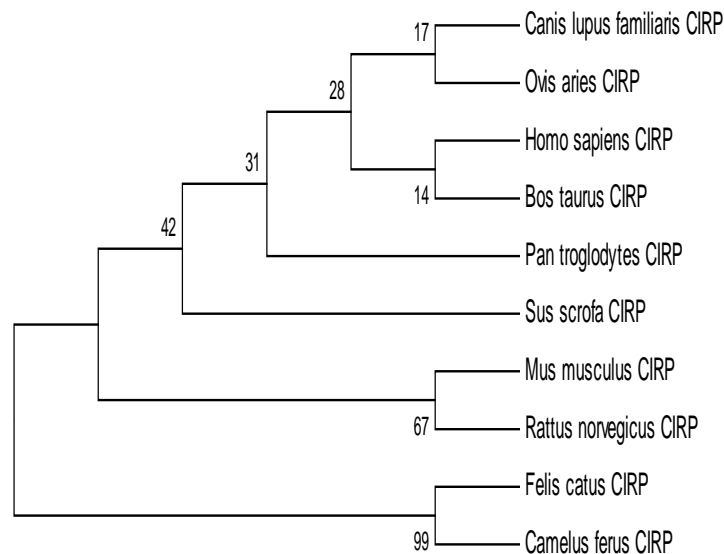
Species names	<i>Homo sapiens</i> (Human)	<i>Sus scrofa</i> (Pig)	<i>Ovis aries</i> (Sheep)	<i>Pan troglodytes</i> (Chimpanzee)	<i>Camelus ferus</i> (Camel)	<i>Canis lupus familiaris</i> (Dog)	<i>Bos Taurus</i> (Cattle)	<i>Felis catus</i> (Cat)	<i>Mus-musculus</i> (Mouse)	<i>Rattus norvegicus</i> (Rat)
<i>Homo sapiens</i> (Human)	100	100	100	100	100	100	90	100	100	100
<i>Pan troglodytes</i> (Chimpanzee)	100	100	100	100	100	100	90	100	100	100
<i>Ovis aries</i> (Sheep)	100	100	100	100	100	100	90	100	100	100
<i>Sus scrofa</i> (Pig)	100	100	100	100	100	100	90	100	100	100
<i>Camelus ferus</i> (Camel)	100	100	100	100	100	100	100	100	100	100
<i>Canis lupus familiaris</i> (Dog)	100	100	100	100	100	100	100	100	100	100
<i>Bos taurus</i> (Cattle)	90	90	90	90	100	100	100	88	90	82
<i>Mus-musculus</i> (Mouse)	100	100	100	100	100	100	90	100	100	100
<i>Felis catus</i> (Cat)	100	100	100	100	100	100	88	100	100	100
<i>Rattus norvegicus</i> (Rat)	100	100	100	100	100	100	82	100	100	100

Table 7. Domain architecture of CIRP using SMART

Domain name	RNA recognition motif		E-value	Low complexity	
	Start	End		Start	End
<i>Homo sapiens</i> (Human)	7	80	1.32e-32	89	125
<i>Pan troglodytes</i> (Chimpazee)	7	80	2.11e-31	89	128
<i>Sus scrofa</i> (Pig)	7	80	5.19e-31	89	144
<i>Bos taurus</i> (cattle)	7	81	1.32e-32	89	167
<i>Ovis aries</i> (Sheep)	7	80	1.32e-32	89	170
<i>Canis lupus familiaris</i> (Dog)	7	80	1.32e-32	89	172
<i>Felis catus</i> (Cat)	31	104	2e-32	113	196
<i>Cemelus ferus</i> (Camel)	8	81	2.63e-32		
<i>Rattus norvegicus</i> (Rat)	7	80	1.62e-32	89	168
<i>Mus musculus</i> (Mouse)	7	80	1.62e-32	89	168

Table 8. Domain architecture of RBM3 using SMART

Domain name	RNA Recognition Motif		E-value	Low complexity	
	Start	End		Start	End
<i>Homo sapiens</i> (human)	7	80	6.84e-31	82	135
<i>Pan troglodytes</i> (chimpazee)	7	80	6.84e-31	82	138
<i>Sus scrofa</i> (Pig)	7	80	1.83e-31	86	142
<i>Bos taurus</i> (Cattle)	7	82	3.64e-26		
<i>Ovis aries</i> (Sheep)	7	80	4.84e-31	82	117
<i>Canis lupus familiaris</i> (Dog)	7	80	8.42e-31	86	154
<i>Felis catus</i> (Cat)	7	80	2.78e-31	82	153
<i>Cemelus ferus</i> (Camel)	7	80	6.48e-32	87	151
<i>Rattus norvegicus</i> (Rat)	7	80	2.78e-31	86	142
<i>Mus musculus</i> (Mouse)	7	80	2.78e-31	82	140

**Fig. 1. Cladogram (Phylogenetic tree) (Topology) of CIRP showing the evolutionary relationship among the mammalian species using Molecular Evolution and Genetic Analysis, version 5 (MEGA 5)**

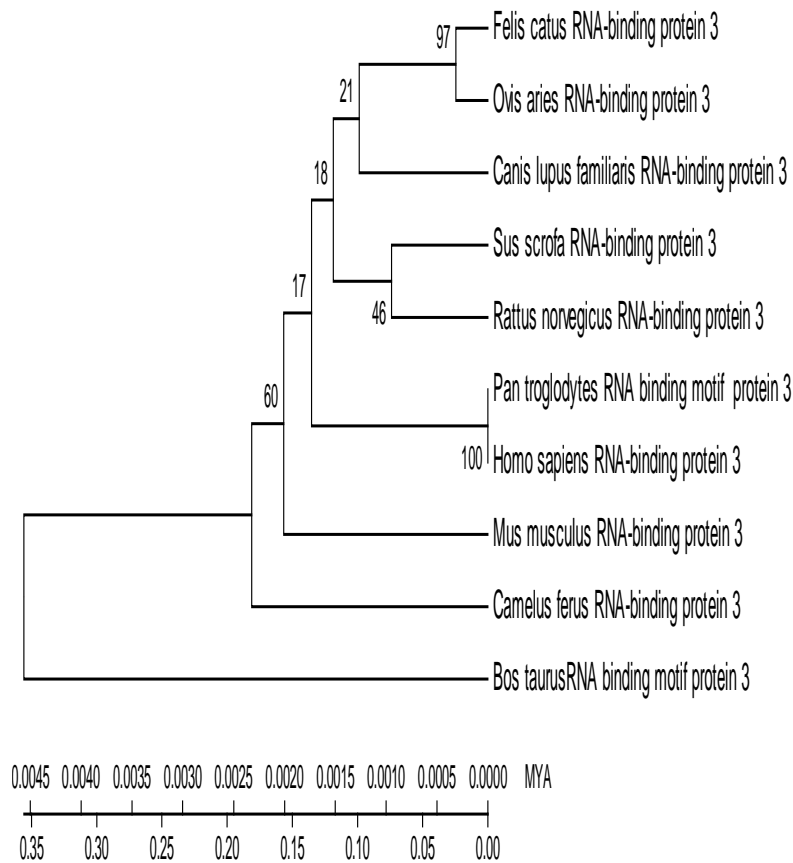


Fig. 2. Cladogram (Phylogenetic tree) of RBM3 showing the evolutionary relationship among the mammalian species using Molecular Evolution and Genetic Analysis, version 5 (MEGA 5)

The sequence length variation observed in CIRP and RBM3 protein of the ten mammalian species compared with each other might have resulted from evolution, mutation and differentiation. According to [12] sequence length variation is caused by insertion and deletion.

The presence of similar length of amino acid sequences between species is an indication of evolution. Following the aforementioned, the protein sequence length of CIRP of *Homo sapiens*, *Rattus norvegicus* and *Mus musculus* was found to be the same; 172 amino acids (a.a.). The protein sequence length of RBM3 was also observed to be similar between *Homo sapiens* and *Pan troglodytes*; 157 (a.a.); *Sus scrofa* and *Rattus norvegicus*; 156 (a.a.) and *Ovis aries* and *Felis catus*; 160 (a.a.) (Tables 1&2). However, the percentage identity of CIRP revealed the highest value of 100 between *Bos taurus* and *Ovis aries*; *Mus musculus* and *Rattus norvegicus*. A 99% similarity was observed between *Sus scrofa*, *Bos taurus* and *Ovis aries*;

Ovis aries and *Felis catus*, *Canis lupus familiaris* and *Sus scrofa*. Also, the percentage identity of RBM3 revealed that *Homo sapiens* and *Pan troglodytes* were closely related (Tables 3, 4, 5 & 6) Moreover, the percentage similarity of RBM3 revealed a high percentage among most of the organisms while a low percentage was observed between *Bos taurus* and some of the other organisms. For CIRP 100% similarity was revealed between *Homo sapiens* and *Mus musculus*; *Rattus norvegicus* and *Pan troglodytes*; *Sus scrofa* and *Mus-musculus*; *Ovis aries* and *Camelus ferus*; *Canis lupus familiaris* and *Felis catus*. The interpretation is that all these organisms were related and had conserved protein sequence.

Based on the results obtained, the phylogenic relationship of CIRP among these organisms showed a close relationship among the species investigated with the exception of *Camelus ferus* and *Felis catus* which could be considered as an out-group and thus distantly related with other

species studied. *Bos taurus* and *Felis catus* were closely related and shared some ancestry with other mammalian species. However, the phylogenetic tree of RBM3 protein revealed that *Homo sapiens* and *Pan troglodytes* were more closely related and had equal time of divergence (Figs. 1&2). Also *Ovis aries* and *Felis catus* were closely related while *Bos taurus* could be regarded as an out group distantly related with other mammalian species studied.

Thus, the same time of divergence of RBM3 was revealed between *Homo sapiens* and *Pan troglodytes* of less than 5 TYA (thousands of years ago) while *Felis catus* and *Ovis aries* had 97G consensus with divergence time of about 3 TYA. *Sus scrofa* and *Rattus norvegicus* had 46G consensus and divergence time of about 9.5 TYA. *Canis lepus familiaris*, *Felis catus* and *Ovis aries* shared 21G consensus and divergence time of about 12.5 TYA.

Domains are evolutionarily more conserved than other regions of a protein and tend to evolve as units, which are gained, lost, or shuffled as one module. The identification of motifs and domains in proteins is an important aspect of the classification and functional annotation of protein sequences [13].

The domain architecture revealed that these proteins contained mainly RNA Recognition motif and a region of low complexity. The domain start position generally was 7 and terminated at 80 amino acids with exception of *Camelus ferus* of 8–81. *Felis catus* recorded 31–104 (Tables 7 & 8). The proteins also contained region of little diversity in the sequence of amino acid composition called low complexity region. Low complexity region evolved rapidly through mitotic replication slippage and meiotic recombination event. The region of low complexity of CIRP started generally from position 89 with the exception of *Felis catus* which commenced at position 113 (Tables 7 & 8).

4. CONCLUSION

Based on the results obtained, evolutionary relationship of CIRP and RBM3 revealed high relatedness among the mammalian species studied. Similarly, a high similarity and identity was observed among these mammalian species. Furthermore, it was found that the proteins were highly conserved and contained mainly RNA Recognition Motif domain which generally started

at 7 and ended at position 80 with low complexity regions.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Al-Fageeh MB, Smales CM. Control and regulation of the cellular response to cold shock: The responses in yeast and mammalian systems. *Biochemistry Journal*. 2006;397:247-259.
2. Carey HV, Andrews MT, Martin SL. Mammalian hibernation: Cellular and molecular responses to depressed metabolism and low temperature. *Physiology Research*. 2003;83(4):1153-1181.
3. Storey KB. Out cold: Biochemical regulation of mammalian hibernation—a mini-review. *Gerontology*. 2010;56(2):220-230.
4. Milsom WK, Jackson DC. Hibernation and gas exchange. *Comprehensive Physiology*. 2011;1(1):397-420.
5. Bergeron D, Beauseigle D, Bellemare G. Sequence and expression of a gene encoding a protein with RNA-binding and glyceric domains in *Brassica napus*. *Biochemistry and Biophysics Acta*. 1993; 1216:123-125.
6. Carpenter CD, Kreps JA, Simon AE. Genes encoding glycine-rich *Arabidopsis thaliana* proteins with RNA-binding motifs are influenced by cold treatment and an endogenous circadian rhythm. *Plant Physiology*. 1994;104:1015-1025.
7. Heintzen C, Melzer S, Fischer R, Kappeler S, Apel K, et al. A light- and temperature-entrained circadian clock controls expression of transcripts encoding nuclear proteins with homology to RNA binding proteins in meristematic tissue. *Plant Journal*. 1994;5:799-813.
8. Williams DR, Epperson LE, Li W, Hughes MA, Taylor R, et al. Seasonally hibernating phenotype assessed through transcript screening. *Physiology Genomics*. 2005; 24(1):13-22.
9. Sano Y, Shiina T, Naitou K, Nakamori H, Shimizu Y. Hibernation-specific alternative splicing of the mRNA encoding cold-

- inducible RNA-binding protein in the hearts of hamsters. *Biochemistry and Biophysics Research Communication*. 2015;462(4): 322-325.
10. Leonart ME. A new generation of proto-oncogenes: Cold-inducible RNA binding proteins. *Biochemica and Biophysica Acta*. 2010;1805(1):43-52.
 11. Jin X. *Essential bioinformatics*. Cambridge University Press. 2006;1:3-27.
 12. Kedersha NL, Gupta M, Li W, Miller I, Anderson P. RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 α to the assembly of mammalian stress granules. *Journal of Cellular Biology*. 1999; 147:1431–1442.
 13. Jiang W, Hou Y, Inouye M. CspA, the major cold-shock protein of *Escherichia coli*, is an RNA chaperone. *Journal of Biology and Chemistry*. 1997;272:196-202.

© 2017 Okon et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here:
<http://sciencedomain.org/review-history/22365>