Trends in Bioinformatics

ISSN 1994-7941

science alert

ANSInet
an open access publisher
http://ansinet.com

OPEN ACCESS Trends in Bioinformatics

ISSN 1994-7941
DOI: 10.3923/tb.2017.1.10

Research Article
In silico Analysis of BRCA1 Gene and its Phylogenetic Relationship in some Selected Domestic Animal Species

[1]E.V. Ikpeme, [1]O.U. Udensi, [1]M.E. Kooffreh, [2]H.E. Etta, [1]B.B. Ushie, [1]E. Echea and [1,3]M. Ozoje

[1]Department of Genetics and Biotechnology, University of Calabar, PMB Calabar, Nigeria
[2]Department of Biological Science, Cross River State University of Science and Technology, Calabar, Nigeria
[3]Department of Animal Science, Federal University of Agriculture, Abeokuta, Nigeria

Abstract

Background: Breast cancer type 1 (BRCA1) gene also known as breast cancer type 1 susceptibility protein homologue plays important role in DNA double-strand break repair, homologous recombination, chromatin remodeling, cell cycle regulation and transcription regulation. **Objective:** Considering the importance of this protein, the present study was undertaken to analyze BRCA1 gene of different mammalian species by assessing the identity and similarity, phylogenetic relationship, physicochemical properties, predict the motifs, secondary and the tertiary structures. **Materials and Methods:** Seventeen nucleotide and protein sequences of BRCA1 gene of different mammalian species were retrieved from National Centre for Biotechnology Information (NCBI). Multiple sequence alignment was done using CLUSTAL W software, while identity and similarity was determined by constructing a pair wise comparison. **Results:** Results obtained at the end of the experiment showed that there was no percentage identity or similarity that was less than 70%. The phylogenetic relationship of BRCA1 gene of the mammalian species clustered into aquatic, herbivores, carnivores and omnivores, respectively. The highest time of divergence (95MYA) of the BRCA1 gene was observed between the BRCA1 gene of killer whale and human, while the least time of divergence (4.6MYA) was observed between the BRCA1 gene of cattle and American buffalo. Physicochemical properties of BRCA1 proteins in the five mammalian species (cattle, sheep, pig, American buffalo and human) were found to be unstable, hydrophilic and intracellular in nature. The following motifs were present at various positions in the BRCA1 gene of the five mammalian species; zinc finger RING, BRCT domain, N-myristoylation site, N-glycosylation, cAMP and cGMP dependent protein kinase phosphorylation site, tyrosine kinase phosphorylation, casein kinase 2 phosphorylation site, protein kinase C phosphorylation site and cell attachment sequence but leucine zipper pattern and microbodies c-terminal motifs were found in only human BRCA1 gene. The BRCA1 secondary structure contained the alpha helix, extended strand and random coil. **Conclusion:** Based on the results obtained, it can be deduced that BRCA1 gene has identical homologue, functional similarity and highly conserved in these mammalian species.

Key words: Mammals, BRCA1 gene, evolution, physicochemical analysis, divergence time

Received: September 11, 2016 Accepted: November 14, 2016 Published: December 15, 2016

Citation: E.V. Ikpeme, O.U. Udensi, M.E. Kooffreh, H.E. Etta, B.B. Ushie, E. Echea and M. Ozoje, 2017. *In silico* analysis of BRCA1 gene and its phylogenetic relationship in some selected domestic animal species. Trends Bioinform., 10: 1-10.

Corresponding Authors: E.V. Ikpeme and O.U. Udensi, Department of Genetics and Biotechnology, University of Calabar, PMB Calabar, Nigeria
Tel: +2348037245930

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Breast cancer 1 early onset (BRCA1) also known as breast cancer type 1 susceptibility protein homologue is a breast and ovarian tumour suppressor[1]. BRCA1 genes are involved in the process of DNA damage repair, regulation of cell cycle and transcription as well as other important pathway in the bid to inhibit tumour and make sure genome stability is maintained[2,3]. However, protein-truncating mutations across BRCA1 have been implicated with an increased cumulative risk of breast (60-80%) and ovarian (20-40%) cancers[4-7]. According to Shen and Vadgama[8], nearly half of the reported changes in BRCA1 are frame-shift mutations[9] and thus expected to be disease-linked. Yuan *et al*.[10] reported bovine BRCA1 gene as a new candidate gene for bovine mastitis, although BRCA1 gene polymorphism associated with bovine mastitis in various cattle breeds have not been investigated. Nash *et al*.[11] and Ruegg[12] independently observed that mastitis is a complex and common disease of dairy cattle, which causes major economic losses to the dairy industry. The genetic association of bovine mastitis notwithstanding[13], highlighted other factors that might cause bovine mastitis such as pathogens, poor management practices and health of the dairy cattle. Mastitis can manifest as clinical and subclinical forms. The clinical mastitis shows noticeable symptoms such as red and swollen mammary glands, while subclinical mastitis do not reveal any obvious signs but is characterized by high Somatic Cell Count (SCC), a normal or elevated body temperature and milk samples that should test positive on culture[14].

The bovine BRCA1 gene has been mapped to chromosome 19 (BTA19)[6] and this location was within a region of similar gene order as the BRCA locus in human (chromosome 17) and mouse (chromosome 11)[6,15,16]. According to Madsen *et al*.[17], BRCA1 gene within mammalian species shows high degrees of divergence, which probably might affect its functionality and structural architecture in different organisms. With this in mind, in the present study, we x-rayed BRCA1 gene of different mammalian species with particular emphasis on the domestic farm species (cattle, sheep, pig and American buffalo) in comparison with the BRCA1 gene of human in terms of its physicochemical properties, phylogenetic relationship with time of divergence, percentage identity and similarity of amino acid sequences as well as to predict the protein motifs. This will essentially expand the knowledge about the functions, structure and evolutionary relationships of the gene.

MATERIALS AND METHODS

Retrieval of nucleotides and amino acids sequences: Nucleotides and amino acid sequences of *Bos taurus* (cattle), *Ovis aries* (sheep), *Sus scrofa* (pig*), Bison bison* (American buffalo), *Homo sapien* (Human), *Tursiops truncatus* (Bottlenose dolphin), *orcinus orca* (Killer whale), *Physeter catodon* (sperm whale), *Odobenus rosmarus divergens* (Walrus), *felis catus* (Cat), *Ailuropoda melanoleuca* (Giant panda), *Pongo abelii* (Sumatran orangutan), *Canis lupus* (Dog), *Colobus guereza* (Colobus monkey), *Panthera tigris* (Tiger), *Pan troglodytes* (Chimpanzee) and *Camelus ferus* (Camel) were downloaded from NCBI. Basic Local Alignment Search Tool (BLAST) was used to obtain similar sequences in other organisms. Multiple sequence alignment was carried out on all obtained amino acid sequences using CLUSTALW software[18].

Determination of percent identity and similarity: The percent identity and similarity among the amino acid sequences of BRCA1 gene in cattle, sheep, pig, American buffalo and human and other mammalian species was determined by conducting a pairwise comparison of the sequences using BLAST.

Phylogenetic analysis and time of divergence: Phylogenetic relationship using the amino acid sequences of the gene retrieved from NCBI was done according the method of Tamura *et al*.[19]. The evolutionary relationship was inferred using unweighted pair group method with arithmetic mean (UPGMA) based on the Jones-Taylor-Thornton (JTT) matrix-based model. The reliability of the inferred phylogenetic tree was evaluated with bootstrap analysis of 1000 replications. The divergence time of BRCA1 gene for each species was calculated based on the nucleotide percent substitution per site.

Determination of physicochemical properties of BRCA1 protein: Physicochemical properties of BRCA1 protein of the five mammalian species were determined using the Expert Protein Analysis System (Expasy), which is the proteomic server of Swiss Institute of Bioinformatics (SIB) (web.expasy.org). We chose only cattle, sheep, pig and American buffalo as domestic farm animals while human was used as a reference point.

Prediction of protein motifs and protein structures: Motifs in the amino acid sequence of BRCA1 gene of the five selected domestic farm animals and human were predicted using

PROSITE software[20] while the protein secondary structures were predicted using GORIV[21]. Additionally, the protein tertiary structure (protein 3D structure) of BRCA1 gene was predicted based on the canonical amino acid sequence obtained from NCBI database using phyre2 according to Kelley and Sternberg[22].

RESULTS

Retrieved nucleotide and amino acids sequences of BRCA1 gene: The lengths of the retrieved nucleotide and amino acid sequences showed variations. The length of the nucleotide sequence of BRCA1 gene varied from 5550-5777 bps while the length of the amino acid sequence varied from 1849-1863 amino acid residues. Among the five selected domestic farm mammals, the length of the nucleotide sequence for pig BRCA1 was the longest (5777 bps), followed by human (7224 bps) while that of the cattle was the shortest (5550 bps). The amino acid residues for pig and human were the same (1863) while for cattle and A. buffalo had 1849 and 1850 amino acid residues, respectively. For emphasis however, the amino acid sequence length for Sumatran Orangutan and chimpanzee were the same with that of pig and human (Table 1).

Percent identity and similarity among the amino acid sequence of BRCA1 gene: Percentage identity among amino acid sequences of human BRCA1 gene with cattle, sheep, pig, A. buffalo revealed over 70% identity while percent similarity showed more than 80% (Table 2). Within the domestic farm mammals, sheep and A. buffalo had 99% identity with the least recorded for human (72%). This trend was also maintained for amino acid sequence similarity of cattle with sheep (95%), A. buffalo (99%), pig (84%) and human (81%). Generally, however, BRCA1 gene amino acid sequence of human has higher percent similarity with that of cattle, sheep, pig as well as A. buffalo when compared with its percent identity.

Phylogenetic relationship of BRCA1 gene among the five domestic farm mammals: The phylogenetic relationship among the BRCA1 genes of the different mammalian species showed that the gene came from common ancestry root but diverged into two major clades A and B in the course of evolution. Though our report laid more emphasis on four domesticated mammalian species (cattle, sheep, pig and A. buffalo) as well as human, however, the clustering was based on feeding habits-omnivores (clade A), carnivores (Sub-clade B_1) and herbivores (Sub-clade B_2). It is also reported that BRCA1 gene of cattle, sheep, pig and A. buffalo are more related as compared to that of human BRCA1 gene (Fig. 1).

Time of divergence: We looked at the time of divergence of BRCA1 gene of human and the BRCA1 genes of cattle, sheep,

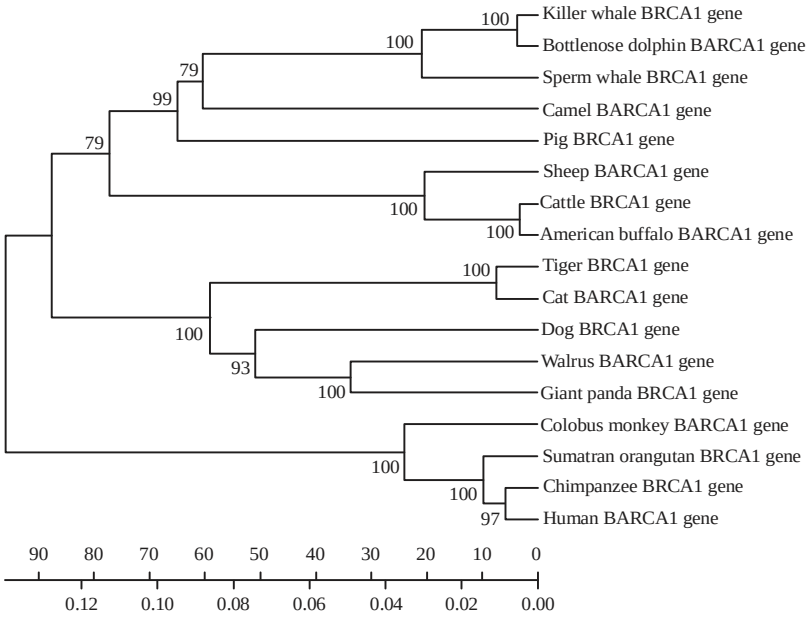

Fig. 1: Phylogenetic tree showing the evolutionary relationship among the mammalian species BRCA1

Table 1: Retrieved nucleotide and amino acid sequences of BRCA1 gene of the mammalian species with their accession numbers and sequence length

Species	Gene name	GenBank accession No.	Protein name	GenBank accession No. for protein	Nucleotide sequence length	Amino acids sequence length
**Cattle (*Bos taurus*)	BRCA1, mRNA	NM178573.1	BRCA1	NP848668.1	5550	1849
**Pig (*Sus scrofa*)	BRCA1, mRNA	AB271921.1	BRCA1	BAF622961.1	5777	1863
**A. buffalo (*Bison bison*)	BRCA1, mRNA	XM010842468.1	BRCA1	XP010840770.1	6829	1850
** *Ovis aries* (sheep)	BRCA1, mRNA	XM015098805.1	BRCA1	XP014954291.1	6685	1862
** Human (*Homo sapiens*)	BRCA1, mRNA	NM007294.3	BRCA1	NP009225.1	7224	1863
Monkey (*Colobus guereza*)	BRCA1, mRNA	KM017623.1	BRCA1	AIG88487.1	5595	1864
Walrus (*Odobenus rosmarus divergens*)	BRCA1, mRNA	XM012565923.1	BRCA1	XP012421377.1	5731	1881
Cat (*Felis catus*)	BRCA1, mRNA	XM011289122.1	BRCA1	XP011287424.1	6531	1871
Giant panda (*A. melanoleuca*)	BRAC1, mRNA	XM011216747.1	BRCA1	XP011215049.1	6702	1869
Sumatran orangutan (*Pongo abelii*)	BRCA1, mRNA	XM009251702.1	BRCA1	XP009249977.1	5794	1863
Sperm whale (*Physeter catodon*)	BRCA1, mRNA	XM007102091.1	BRCA1	XP007102153.1	7022	1860
Killer whale (*Orcinus orca*)	BRCA1, mRNA	XM012537731.1	BRCA1	XP012393185.1	5854	1862
Bottlenose dolphin (*T. truncatus*)	BRCA1, mRNA	XM004312200.1	BRCA1	XP004312248.1	5794	1815
Tiger (*Panthera tigris*)	BRCA1, mRNA	XM007087694.1	BRCA1	XP007037756.1	6460	1874
Chimpanzee (*Pan troglodytes*)	BRCA1, mRNA	XM009432082.7	BRCA1	XP009430357.1	7261	1863
Camel (*Camelus ferus*)	BRCA1, mRNA	XM006175151.2	BRCA1	XP006175213.1	5736	1859
Dog (*Canis lupus familaris*)	BRCA1, mRNA	NM001013416.1	BRCA1	NP001013434.1	5637	1878

**Mammalian species of interest

Table 2: Percent identity and similarity among the amino acid sequence of cattle, sheep, pig, A. buffalo and human BRCA1 gene

Species	Identity to cattle	Identity to sheep	Identity to pig	Identity to A. buffalo	Identity to human	Similarity to cattle	Similarity to sheep	Similarity to pig	Similarity to A. buffalo	Similarity to human	E-value
**Cattle	100	94	77	99	72	100	95	85	99	81	0.00
**Sheep	99	100	76	94	72	95	100	85	96	82	0.00
**Pig	76	76	100	76	74	84	85	100	84	83	0.00
**A. buffalo	99	94	76	100	72	99	96	85	100	82	0.00
**Human	72	72	74	72	100	81	82	83	82	100	0.00
Chimpanzee	72	72	74	72	98	82	82	84	82	98	0.00
Camel	77	77	81	77	76	86	85	87	86	85	0.00
Tiger	74	74	77	74	76	82	82	84	82	84	0.00
Sperm whale	80	80	83	80	77	87	87	88	87	84	0.00
Killer whale	80	80	82	80	76	86	86	88	87	84	0.00
Bottlenose dolphin	79	79	81	79	76	86	86	87	86	84	0.00
Walrus	75	75	77	75	77	83	83	85	83	84	0.00
Cat	74	74	76	74	75	82	82	84	82	83	0.00
Giant panda	74	73	76	74	75	83	83	84	83	83	0.00
Sumatran orangutan	72	72	74	72	97	81	81	83	82	97	0.00
Dog	72	72	74	72	74	81	80	82	81	82	0.00
Colobus monkey	71	71	73	71	93	81	81	83	81	93	0.00

**Mammalian species of interest

Table 3: Physicochemical properties of BRCA1 protein in five selected mammalian species

Physicochemical properties	Cattle value	Sheep value	Pig value	A. buffalo value	Human value
Theoretical pl	5.40	5.49	5.60	5.43	5.21
Molecular weight	206278.6	207372.1	206747.4	206552.9	207720.8
No. of amino acids	1849	1862	1863	1850	1862
Aliphatic index	68.84	68.46	66.71	68.43	69.01
Grand average hydropathicity (GRAVY)	-0.778	-0.767	-0.815	-0.794	-0.785
Total No. of negatively charged residue (aspartate+glutamine)	282	277	272	283	283
Total No. of positively charged residue (arginine+lysine)	217	217	221	220	213
Atomic composition	C (8865), H (14191), N (2541), O (2979), S (70)	C (8908), H (14251), N (2563), O (2978), S (77)	C (8851), H (14178), N (2584), O (2988), S (67)	C (8874), H (14213), N (2549), O (2983), S (69)	C (8908), H (14246), N (2554), O (3014), S (74)
Total No. of atoms	28646	28777	28668	28688	28796
Extinction coefficient	0.459	0.477	0.458	0.452	0.500
Instability index	55.51	54.13	54.63	56.34	54.68

pig as well as A. buffalo. It is revealed that human BRCA1 gene diverged from the BRCA1 genes of the domestic mammals selected about 88MYA. Important to mention for emphasis is the fact that the clade where the canine BRCA1 genes diverged from the BRCA1 genes of the omnivores about 75MYA. Figure 1 showed that there were varying time of divergence among the four selected domestic farm animals, the least being approximately 4MYA (Fig. 1).

Physicochemical properties of BRCA1 protein: The analysis of BRCA1 protein of cattle, sheep, pig, A. buffalo and human showed varying physicochemical properties when compared. The result revealed that the least theoretical pl was for human BRCA1 protein (5.21) while the highest was for pig BRCA1 protein (5.60). The molecular weights of BRCA1 proteins from the 5 mammalian species revealed that human BRCA1 protein weigh 207720.8 kDa while sheep BRCA1 protein weigh 207372.1 kDa. Cattle BRCA1 protein had a weight of 206278.6 kDa, which was the least. The number amino acids in BRCA 1 protein of the species showed the following: Cattle (1849), sheep (1862), pig (1863). The A. buffalo (1850) and human (1862), implying that the number of amino acids in the BRCA1 protein of human and sheep are the same (Table 3).

Predicted motifs in BRCA1 proteins: Nine motifs were reported as predicted in the BRCA1 gene of the different mammalian species selected, which include zinc finger ring, BRCT domain, N-myristoylation site, N-glycosylation site, cAMP and cGMP dependent protein kinase phosphorylation site, tyrosine kinase phosphorylation site, cell attachment sequence, leucine zipper pattern as well as micro-bodies c-terminal targeting signal motifs. However, cell attachment sequence motif was absent in the BRCA1 proteins of sheep and human while human BRCA1 protein possess leucine zipper pattern and micro-bodies c-terminal targeting signal motifs, which were lacking in those of cattle, sheep, pig as well as A. buffalo (Table 4).

Protein secondary structures BRCA1 gene: The secondary structures of BRCA1 protein was predicted using GORIV software for the five selected mammalian species. It is showed that their BRCA1 protein contain basically alpha helix, extended strand as well as random coil. However, the alpha helix for human BRCA1 protein was 32.15% while for cattle it was 30.18%. The least was for sheep whose alpha helix occupies 28.68% of the structure. Conversely, the extended strand of the BRCA1 protein for pig was the longest (14.65%) with that of the human being the shortest (12.56%). Ther were

Table 4: Motifs in cattle, sheep, pig, A. buffalo and human BRCA1 protein

	Positions				
Motif names	Cattle	Sheep	Pig	A. buffalo	Human
Zinc finger RING	24-65	24-65	22-63	24-65	24-65
BRCT domain	1642-1729, 1749-1848	1648-1735, 1755-1854	1645-1732, 1752-1852	1643-1730, 1750-1849	1642-1736, 1756-1855
N-myristoylation site	98-103, 160-165, 183-188, 263-268, 401-406, 402-407, 534-539, 544-549, 675-680, 1042-1047, 1049-1054, 1079-1084, 1133-1138, 1252-1257, 1260-1265, 1266-1271, 1287-1292, 1377-1382, 1417-1422, 1487-1492, 1699-1704, 1756 -1761	98-103, 160-165, 183-188, 268-273, 406-411, 407-412, 549-554, 640-685, 1048-1053, 1055-1060, 1139-1144, 1266-1271, 1272-1277, 1293-1298, 1423-1428, 1493-1498, 1705-1710, 1762-1767	96-101,182-187,198-203, 202-207, 257-262, 402-407, 530-535, 538-543, 540-545, 671-676, 769-774, 808-813, 943-948, 1041-1046, 1048-1053, 1109-1114, 1132-1137, 1287-1292, 1418-1423, 1488-1493, 1702-1707, 1759-1764	98-103,160-165,183-188, 263-268, 401-406, 402-407, 534-539, 544-549, 675-680, 1043-1048, 1050-1055, 1134-1139, 1261-1266, 1267-1272, 1288-1293, 1378-1383, 1418-1423, 1488-1493, 1700-1705, 1757-1762, 1778-1783	98-103,160-165,183-188, 263-268, 312-317, 323-328, 394-399, 484-489, 535-540, 543-548, 677-682, 774-779, 778-783, 813-818, 949-954, 960-965, 1048-1053, 1055-1060, 1138-1143, 1332-1337, 1350-1355, 1422-1427, 1492-1497, 1515-1520, 1591-1596, 706-1711, 1763-1768
N-glycosylation site	106-109, 175-178, 287-290, 306 -309, 376-379, 471-474, 546-549, 554-557, 633-636, 679-682, 708-711, 734-737, 760-763, 816-819, 1614-1617	106-109, 175-178, 292-295, 311-314, 381-384, 551-554, 55-562, 580-583, 592-595, 638-641, 685-688, 714-717, 740-743, 766-769, 822-825, 1040-1043, 1620-1623	104-107, 174-177, 281-284, 300-303, 370-373, 467-470, 542-545, 550-553, 695-698, 705-708, 757-760, 815-818, 829-832, 989-992, 1008-1011, 1097-1100, 1444-1447, 1515-1518	106-109, 175-178, 287-290, 306-309, 378-379, 471-474, 520-523, 546-549, 554-557, 633-636, 680-683, 709-712, 735-738, 761-764, 817-820, 1615-1618	247-250, 287-290, 306-309, 376-379, 473-476, 537-340, 547-550, 555-558, 635-638, 714-717, 824-827, 838-841, 913-916, 924-927, 961-964, 1215-1218
cAMP and cGMP-dependent protein kinase phosphorylation site	464-467, 465-468, 504-507, 610-613, 611-614, 650-653, 858-861, 1641-1644	469-472, 470-473, 509-512, 615-618, 616-619, 655-658, 864-867, 1647-1650	461-464, 500-503, 606-609, 607-610, 646-649, 856-859, 928-931, 1644-1647	464-467, 465-468, 610-613, 611-614, 650-653, 859-862, 1642-1645	467-470, 506-509, 612-615, 613-616, 865-868, 893-896, 1648-1651
Tyrosine kinase phosphorylation site	1548-1556, 1755-1762	341-344, 353-356	1551-1559	1549-1557, 1756-1763	823-831, 1555-1563, 1762-1769
Cell attachment sequence	1730-1732		663-665, 1733-1735	1731-1733	
Leucin zipper pattern					1209-1230, 1732-1739
Microbodies c-terminal targeting signal					1861-1863

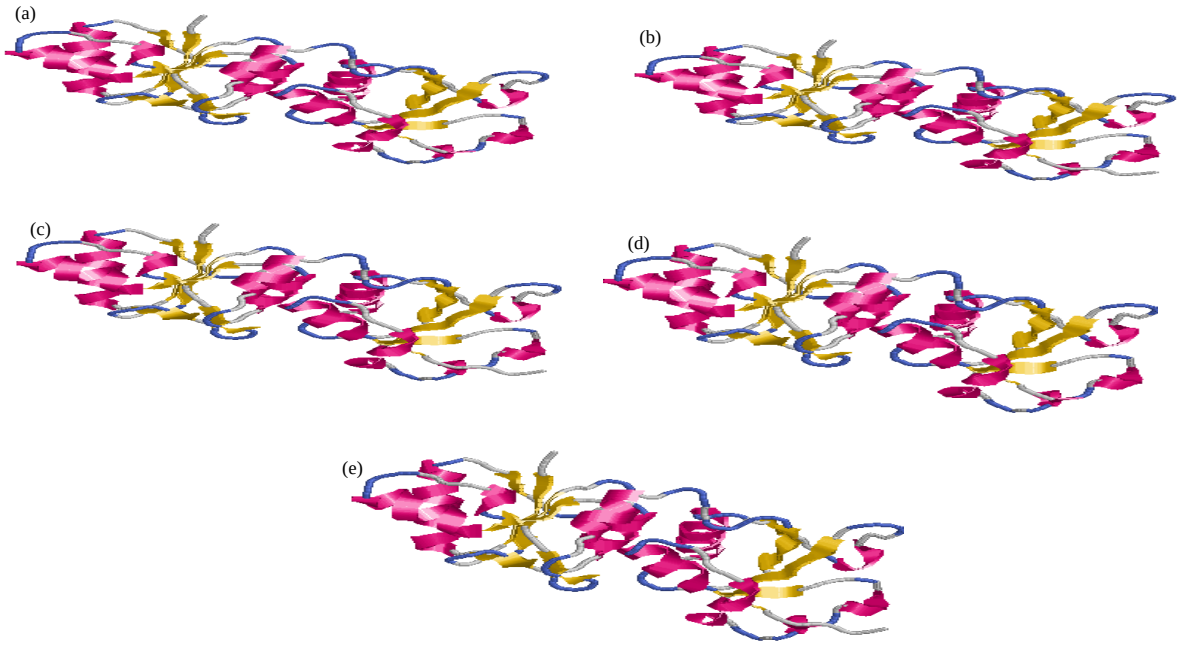

Fig. 2(a-e): Tertiary structures of (a) Cattle, (b) Sheep, (c) Pig, (d) A. buffalo and (e) Human BRCA1 proteins in cartoon model as displayed by Rasmol
The alpha helix (pink spiral sheet), the random coil (blue strand), extended strands (white strands), beta sheet (yellow)

Table 5: Protein secondary structure of BRCA1 protein as calculated using GORIV

Secondary structure element (%)	Cattle	Sheep	Pig	A. buffalo	Human
Alpha helix	30.18	28.68	28.77	30.11	32.15
310 helix	0.00	0.00	0.00	0.00	0.00
Pi helix	0.00	0.00	0.00	0.00	0.00
Beta bridge	0.00	0.00	0.00	0.00	0.00
Extended strand	13.30	14.12	14.65	12.97	12.56
Beta turn	0.00	0.00	0.00	0.00	0.00
Bend region	0.00	0.00	0.00	0.00	0.00
Random coil	56.52	57.20	56.58	56.92	55.29
Ambiguous states	0.00	0.00	0.00	0.00	0.00
Other states	0.00	0.00	0.00	0.00	0.00

variations on the random coil of the secondary structure of BRCA1 protein of the five species with the random coil of sheep BRCA1 protein occupying 57.20% of the structure (Table 5, Fig. 2).

DISCUSSION

The BRCA1 genes are reportedly involved in processes such as DNA repair, regulation of cell cycle, transcription processes as well as other important pathways leading to the inhibition of tumour and to ensure that the organism's genome maintain stability[2,3]. Regrettably, protein-truncating mutations across BRCA1 gene have predisposed organisms, especially human to increased cumulative risk of breast (60-80%) and ovarian (20-40%) cancers[4-7].

Yuan *et al.*[10] reported that bovine BRCA1 gene has been implicated as a new candidate gene for bovine mastitis, however, its polymorphism in various cattle breeds have not been investigated. The thrust of the present report was to evaluate the phylogenetic relationship of BRCA1 genes in different mammalian species but with more emphasis on domesticated species. The essence is that the disease that may be linked to the BRCA1 genes of these domestic mammals might spell doom on food security and nutrition as it causes major economic losses to the dairy industry. Rajasekaran *et al.*[9] reported that 50% of the reported changes in BRCA1 gene are frameshift mutations. Here, we report nucleotide amino acids sequence lengths variations comparing the cattle, sheep, pig, A. buffalo and the human BRCA1 gene. It is also affirmed that

sequence length variations are caused by indels mutations, which might have accumulated during evolution.

Organisms with high percentage similarity in their genes have a similar pattern of evolution as well as differentiation[23]. From our result, percentage identity among cattle, sheep, pig and A. buffalo when compared with that of human BRCA1 amino acids ranged 72-74% showing that they may had similar pattern of evolution. This was also the case for percentage similarity (81-83%). Importantly, sequences with percent similarity approximately 70% imply that identical homology, function similarity and very high conservation in BRCA1 gene. Additionally, Joshi and Xu[24] observed that if two sequences have sequence identity greater than 70%, it is suggested that they have about 90% probability or more to share the same biological processes and functions. Proteins in the same family share at least more than 30% amino acid sequence similarity. This might lead to sharing of some structural characteristics[25]. The implication of zero E-value is that the percent identity and similarity obtained were significant[26] suggesting that the result reported here is correct.

The phylogenetic relationship of the mammalian BRCA1 genes is such that species of similar eating habits are clustered together. These result suggested that the closer the percentage identity and similarity among species, the more possibility of being clustered together. Yuan *et al.*[10] reported that BRCA1 gene is the candidate gene for bovine mastitis, this may not imply that it should be implicated for other tumour-related diseases in other mammalian species found in the same sub-clade such as sheep, pig and A. buffalo. To buttress the above, Ma *et al.*[27] reported that BRCA1 gene knockout in mouse model provoked embryonic lethality while conditional knockout in breast tissue resulted to tumour development after a long latency. It might be pertinent here to stress that though the different domestic farm mammals chosen have BRCA1 gene at different chromosome locations, their breast biology[28] as well as cancer genetics[29] vary widely. It thus becomes difficult to use one model to infer the function of BRCA1 gene in another species even when mutated.

Laud *et al.*[30] opined that BRCA1 gene has regulatory effects on steroid and growth hormones, especially in the induction of mammary epithelial cell proliferation. Taking together with other functions of BRCA1 genes earlier mentioned, the implication is that BRCA1 gene might play different role in different mammalian species even when mutated. This might be dependent on the motifs where the action takes place. To underscore the adequacy of the phylogenetic tree, the time of divergence of the mammalian species corroborates the previous reports of Stone *et al.*[23] Babb *et al.*[31] and Psouni *et al.*[32]. According to Kyte and Doolittle[33], GRAVY value greater than zero indicates a relatively hydrophobic protein. However, the GRAVY values obtained in this study for cattle, sheep, pig, A. buffalo and human BRCA1 protein are less than zero, which implies that they are rather hydrophilic in nature. The fact that BRCA1 protein in all the five mammalian species discussed in this report have higher negatively charged residues than positively charged residues makes BRCA1 protein in these species intracellular[34]. According to Guruprasad *et al.*[35], instability index, which is the measure of the stability of a protein *in vitro* when it is greater than 40, it means that the protein may be stable. This as reported by Sharma *et al.*[36] is accounted for by the abundance of cysteine as a result of the formation of disulphide bond in the protein molecule. Regrettably, we report instability index that are greater than 50, which might suggest that this protein is probably unstable *in vitro.*

Sequence motifs are short recurring patterns in the DNA that are presumed to have a biological function. Most often they indicate sequence-specific binding sites for proteins such as nucleases and transcription factors. However, others may be involved in important processes at the RNA level, including ribosome binding, mRNA processing as well as transcription termination. Out of the nine motifs predicted in the mammalian species investigated, leucine-zipper pattern and microbodies c-terminal targeting signal were only present in the BRCA1 protein of human suggesting that they may be linked to DNA binding and protein dimerization[37] and give assistance in the conversion of lipids to carbohydrate. For emphasis, other motifs that are found in the mammalian species are very important in the correct functioning of the system. Chung *et al.*[38], Whisler *et al.*[39], Maurer-Stroh *et al.*[40] and Jin *et al.*[41] reported lay credence to the above claim.

Our result showed that on the secondary structure of BRCA1 protein of the five mammalian species, secondary elements were only alpha helix, extended strands as well as random coil. These three structures have been implicated for the folding stability and function of the protein.

CONCLUSION

The implication of results taking together is that BRCA1 gene is found in diverse mammals but with varying gene and protein characteristics. Though BRCA1 gene has been indicted to cause breast/ovarian cancer in human and as candidate gene for bovine mastitis, it might be difficult to generalize that the gene may also be responsible for any mammary tumour in these mammals evaluated.

SIGNIFICANCE STATEMENTS

- BRCA1 gene has been implicated in breast and ovarian cancer in humans
- Different domestic mammals are also susceptible to this disease
- There is every possibility that the gene implicated in humans may have evolutionary relationship, which may be correlated with the disease
- It therefore becomes imperative to use bioinformatics tools to understand this relationship, which may help veterinary experts and provide baseline information as regards prevention and treatment

ACKNOWLEDGMENTS

The author would like to thank the Bioinformatics and Genome Science Unit of the Department of Genetics and Biotechnology, University of Calabar, Nigeria for technical assistance as well as Dr. Durosaro of FUNNAB, Abeokuta, Nigeria.

REFERENCES

1. Donninger, H., K. Hobbing, M.L. Schmidt, E. Walters, L. Rund, L. Schook and G.J. Clark, 2015. A porcine model system of BRCA1 driven breast cancer. Frontiers Genet., Vol. 6. 10.3389/fgene.2015.00269.
2. Scully, R. and D.M. Livingston, 2000. In search of the tumour-suppressor functions of BRCA1 and BRCA2. Nature, 408: 429-432.
3. Xu, J., B. Wang, Y. Zhang, R. Li, Y. Wang and S. Zhang, 2012. Clinical implications for BRCA gene mutation in breast cancer. Mol. Biol. Rep., 39: 3097-3102.
4. Antoniou, A.C., S.A. Gayther, J.F. Stratton, B.A.J. Ponder and D.F. Easton, 2000. Risk models for familial ovarian and breast cancer. Genet. Epidemiol., 18: 173-190.
5. Kennedy, R.D., J.E. Quinn, P.G. Johnston and D.P. Harkin, 2002. BRCA1: Mechanisms of inactivation and implications for management of patients. Lancet, 360: 1007-1014.
6. Krum, S.A., J.E. Womack and T.F. Lane, 2003. Bovine BRCA1 shows classic responses to genotoxic stress but low *in vitro* transcriptional activation activity. Oncogene, 22: 6032-6044.
7. Whitehouse, C., J. Chambers, A. Catteau and E. Solomon, 2004. BRCA1 expression is regulated by a bidirectional promoter that is shared by the Nbr1 gene in mouse. Gene, 326: 87-96.
8. Shen, D. and J.V. Vadgama, 1999. BRCA1 and BRCA2 gene mutation analysis: Visit to the breast cancer information core (BIC). Oncol. Res. Featuring Preclinical Clin. Cancer Ther., 11: 63-69.
9. Rajasekaran, R., C. Sudandiradoss, C.G.P. Doss and R. Sethumadhavan, 2007. Identification and in silico analysis of functional SNPs of the BRCA1 gene. Genomics, 90: 447-452.
10. Yuan, Z., J. Li, J. Li, L. Zhang, X. Gao, H.J. Gao and S. Xu, 2012. Investigation on BRCA1 SNPs and its effects on mastitis in Chinese commercial cattle. Gene, 505: 190-194.
11. Nash, D.L., G.W. Rogers, J.B. Cooper, G.L. Hargrove and J.F. Keown, 2003. Heritability of intramammary infections at first parturition and relationships with sire transmitting abilities for somatic cell score, udder type traits, productive life and protein yield. J. Dairy Sci., 86: 2684-2695.
12. Ruegg, P.L., 2003. Investigation of mastitis problems on farms. Vet. Clin. N. Am.: Food Anim. Pract., 19: 47-73.
13. Zhang, L.P., Q.F. Gan, T.H. Ma, H.D. Li and X.P. Wang *et al*., 2009. Toll-like receptor 2 gene polymorphism and its relationship with SCS in dairy cattle. Anim. Biotechnol., 20: 87-95.
14. Awale, M.M., G.B. Dudhatra, A. Kumar, B.N. Chauhan and D.R. Kamani *et al*., 2012. Bovine mastitis: A threat to economy. Open Access Scient. Rep., Vol. 1.
15. Miki, Y., J. Swensen, D. Shattuck-Eidens, P.A. Futreal and K. Harshman *et al*., 1994. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science, 266: 66-71.
16. Lane, T.F., C. Deng, A. Elson, M.S. Lyu, C.A. Kozak and P. Leder, 1995. Expression of BRCA1 is associated with terminal differentiation of ectodermally and mesodermally derived tissues in mice. Genes Dev., 9: 2712-2722.
17. Madsen, O., M. Scally, C.J. Douady, D.J. Kao and R.W. DeBry *et al*., 2001. Parallel adaptive radiations in two major clades of placental mammals. Nature, 409: 610-614.
18. Thompson, J.D., D.G. Higgins and T.J. Gibson, 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res., 22: 4673-4680.
19. Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar, 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. Mol. Biol. Evol., 28: 2731-2739.
20. De Castro, E., C.J.A. Sigrist, A. Gattiker, V. Bulliard and P. S. Langendijk-Genevaux *et al*., 2006. ScanProsite: Detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucl. Acids Res., 34: W362-W365.
21. Garnier, J., J.F. Gibrat and B. Robson, 1996. GOR method for predicting protein secondary structure from amino acid sequence. Methods Enzymol., 266: 540-553.
22. Kelley, L.A. and M.J.E. Sternberg, 2009. Protein structure prediction on the web: A case study using the Phyre server. Nat. Protocols, 4: 363-371.

23. Stone, A.C., F.U. Battistuzzi, L.S. Kubatko, G.H. Perry, E. Trudeau, H. Lin and S. Kumar, 2010. More reliable estimates of divergence times in Pan using complete mtDNA sequences and accounting for population structure. Philos. Trans. R. Soc. London B: Biol. Sci., 365: 3277-3288.
24. Joshi, T. and D. Xu, 2007. Quantitative assessment of relationship between sequence similarity and function similarity. BMC Genomics, Vol. 8. 10.1186/1471-2164-8-222
25. Lehninger, A.L., D.L. Nelson and M.M. Cox, 1993. Principles of Biochemistry. 2nd Edn., Worth Publisher, New York.
26. Lesh, A.M., 2002. Introduction to Bioinformatics. 2nd Edn., Oxford University Press, Oxford, Pages: 255.
27. Ma, Y., S. Fan, C. Hu, Q. Meng and S.A. Fuqua *et al*., 2010. BRCA1 regulates acetylation and ubiquitination of estrogen receptor-α. Mol. Endocrinol., 24: 76-90.
28. Dine, J. and C.X. Deng, 2013. Mouse models of BRCA1 and their application to breast cancer research. Cancer Metastasis Rev., 32: 25-37.
29. Kendall, S.D., C.M. Linardic, S.J. Adam and C.M. Counter, 2005. A network of genetic events sufficient to convert normal human cells to a tumorigenic state. Cancer Res., 65: 9824-9828.
30. Laud, K., I. Gourdou, L. Belair, D.H. Keisler and J. Djiane, 1999. Detection and regulation of leptin receptor mRNA in ovine mammary epithelial cells during pregnancy and lactation. FEBS Lett., 463: 194-198.
31. Babb, P.L., E. Fernandez-Duque and T.G. Schurr, 2010. AVPR1A sequence variation in monogamous owl monkeys (*Aotus azarai*) and its implications for the evolution of platyrrhine social behavior. J. Mol. Evolution., 71: 279-297.
32. Psouni, E., A. Janke and M. Garwicz, 2012. Impact of carnivory on human development and evolution revealed by a new unifying model of weaning in mammals. PloS One, Vol. 7. 10.1371/journal.pone.0032452.
33. Kyte, J. and R.E. Doolittle, 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol., 157: 105-132.
34. Andrade, M.A., S.I. O'Donoghue and B. Rost, 1998. Adaptation of protein surfaces to sub cellular location. J. Mol. Biol., 276: 517-525.
35. Guruprasad, K., B.V.B. Reddy and M.W. Pandit, 1990. Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting *in vivo* stability of a protein from its primary sequence. Protein Eng. Des. Sel., 4: 155-161.
36. Sharma, N., R. Kushwaha, J.S. Sodhi and T.C. Bhalla, 2009. *In silico* analysis of amino acid sequences in relation to specificity and physicochemical properties of some microbial nitrilases. J. Proteom. Bioinform., 2: 185-192.
37. Sessa, G., C. Steindler, G. Morelli and I. Ruberti, 1998. The Arabidopsis athb-8,-9 and genes are members of a small gene family coding for highly related HD-ZIP proteins. Plant Mol. Biol., 38: 609-622.
38. Chung, M., J. Lee, T. Terasawa, J. Lau and T.A. Trikalinos, 2011. Vitamin D with or without calcium supplementation for prevention of cancer and fractures: An updated meta-analysis for the U.S. preventive services task force. Ann. Internal Med., 155: 827-838.
39. Whisler, R.L., M.A. Goyette, I.S. Grants and Y.G. Newhouse, 1995. Sub lethal levels of oxidant stress stimulate multiple serine/threonine kinases and suppress protein phosphatases in Jurkat T cells. Arch. Biochem. Biophys., 319: 23-35.
40. Maurer-Stroh, S., B. Eisenhaber and F. Eisenhaber, 2002. N-terminal N-myristoylation of proteins: Prediction of substrate proteins from amino acid sequence. J. Mol. Biol., 317: 541-557.
41. Jin, Y., Y. Wang, D.L. Walker, H. Dong, C. Conley, J. Johansen and K.M. Johansen, 1999. JIL-1: A novel chromosomal tandem Kinase implicated in transcriptional regulation in *Drosophila*. Mol. Cell, 4: 129-135.