

Original Research Article

<https://doi.org/10.20546/ijcmas.2021.1001.335>

Phylogenetic and Protein Domain Architecture Analysis of Toll-like Receptor 4-10 Genes in Camel (*Camelus dromedarius*)

Jyoti Choudhary^{1*}, Mukul Purva¹, Mahendra Milind¹, Kritika Gahlot¹,
Taruna Bhati¹, Sunil Maherchandani¹ and Amit Pandey²

¹Department of Veterinary Microbiology and Biotechnology, ²Department of Biochemistry,
College of Veterinary and Animal Science, RAJUVAS, Bikaner, Rajasthan, India

*Corresponding author

ABSTRACT

Keywords

Camel, domain analysis, Leucine rich repeats (LRRs), phylogenetic analysis, Toll like receptors (TLRs)

Article Info

Accepted:
18 December 2020
Available Online:
10 January 2021

Toll-like receptors (TLRs) are evolutionary conserved protein and most extensively studied pathogen recognition receptors (PRRs) which provoke innate immune responses after recognition of invading pathogens. The present research is the first report of determination of domain architecture of camel (*camelus dromedarius*) TLRs4-10 proteins by Simple Modular Domain Architecture (SMART) tool. The study revealed that all the TLR proteins showed the typical structure containing Leucine Rich Repeats (ectodomain), transmembrane domain (TM), Toll/interleukin-1 receptor (TIR) domains except for TLR9 which was devoid of TM domain. The TLR 7,8,9 possessed more Leucine Rich Repeats (LRRs) motifs in their ectodomains than the other TLRs (4, 5, 6, 10). Further, the TM domain in TLRs (4-10) was exactly of similar length. However, the TIR domain was constituted of almost equal number of amino acids (140-144aa). Based on a phylogenetic analysis of the camel TLRs (1-10), the intracellularly expressed genes TLR7, TLR8 and TLR9 were clustered together whereas extracellularly expressed TLR1, 2, 6, and 10 branched out close to each other.

Introduction

Toll-like receptors (TLRs) are important pattern recognition receptors that play a crucial role in innate immunity in vertebrates (Kawai and Akira, 2011). They belong to the type I transmembrane glycoprotein receptor family present on the cell surface or membrane compartments of immune and non-immune cells (e.g., epithelial cells) (Barton and Kagan, 2009). TLRs specifically recognize and bind a variety of highly

conserved pathogen-associated molecular patterns (PAMPs) that are essential for pathogen survival, subsequently triggering protective immune responses against pathogen infections (Wallet *et al.*, 2018). They recognize a wide variety of PAMPs, present in bacteria, virus and fungi, which include glycolipids such as bacterial lipopolysaccharides (LPS), proteins such as bacterial flagellin and viral nucleic acids (Kannaki *et al.*, 2011). Numerous studies on human immunity have indicated that TLR

genetic variation correlates closely with disease (Mattila *et al.*, 2016; Netea *et al.*, 2012). Mammalian TLR genes were previously classified into six subfamilies according to their phylogenetic relationship and recognition of microbial component: the TLR 2 subfamily (TLR1, TLR2, TLR6 and TLR10); the TLR 7 subfamily (TLR7, TLR8 and TLR9); the TLR 11 subfamily (TLR11, TLR12 and TLR13); the TLR3 subfamily; the TLR4 subfamily and the TLR5 family (Hoshino *et al.*, 1999). The TLR is composed of three domains which includes ectodomain, transmembrane (TM) domain and intracellular cytoplasmic TIR domain. The ectodomains are highly variable in different TLRs and directly involved in the recognition of a variety of PAMPs. It is made up of tandem repeats of leucine-rich regions termed as leucine-rich repeats (LRRs). The arrangement of LRRs confers a unique combinatorial code to each TLR allowing it to bind a specific ligand. They are expressed by a variety of cell types such as dendritic cell, macrophage, Plasmacytoid dendritic cells and lymphocytes (Dowling and Dellacasagrande, 2016). Further, TLR members can form association outside TLR family, e.g., MD-2 co-receptor and CD14 form complex with mammalian TLR4 to form a functional TLR1 (Akira and Takeda, 2004) and this further increase diversity in PAMPs recognizing ability of the TLRs. The TIR domain is conserved across all TLRs and critical for recruiting adaptor proteins, play a role in linking TLRs to nuclear transcription factors that further initiates TLR signalling which leads to inflammation, release of cytokines and cell activation (Bowie and Haga, 2005).

Analysis of camel TLR genes have provided the opportunity to examine their role in disease resistance as well as susceptibility to infections and the adaptive evolution of TLRs. Although, in dromedary camel there is paucity of information about TLRs (4-10)

structure, therefore it is the first report where we have generated domain analysis of camel TLRs4-10 and phylogenetic relationship between all the ten TLRs.

Materials and Methods

The coding region of the camel TLRs (1-10) were used in this study and these nucleotide sequences were retrieved from NCBI database with following accession number- TLR1 (MG655186), TLR2(JQ979305), TLR3 (MN756614), TLR4(MN807289), TLR5 (MN807290), TLR6(MN807291), TLR7 (MN807292), TLR8(MN807293), TLR9 (MN967525), TLR10(MN807294). These coding regions were translated into a protein sequence by using ExPasy software. Multiple sequence alignment of these sequences was done using ClustalW multiple alignment program (<http://www.ebi.ac.uk/Tools/clustalw2>). Then, the aligned protein sequences of TLR genes were used to build the phylogenetic trees in MEGA 6.0(<http://www.megasoftware.net>) software and the neighbour-joining (NJ) algorithm was used to build the tree. Deduced TLRs (4-10) protein domains were deduced by results from the Simple Modular Architecture Research Tool (SMART) databases. The normal mode of SMART was used for all protein domain searches and predictions.

Results and Discussion

The mRNA sequence of *Camelus dromedarius* TLRs(4-10) was composed of 841, 870, 831, 1050, 1030, 1324, and 811 amino acids respectively and showed a typical structure. Schematic representation of TLR genes domains of camel was predicted by SMART (Fig. 1). According to SMART predictions, the pattern of the large extracellular domain, a transmembrane domain and the TIR domains among camel TLRs revealed no significant differences

except TLR9 which devoid of the TM domain but additionally consisted of two actin depolymerization factors (ADF) which was declared as an unknown region (Dhar *et al.*, 2016). This region can be further analysed in

future studies. As evident in Fig.1 the number of domains was different among different TLR genes in camel. TLR 7 subfamily (7,8,9) contain more LRR motifs in their ectodomains than the other TLRs.

Table.1 Transmembrane structures of *Camelus dromedarius* TLR4 predicted domain, repeats and motifs done by SMART tool

Name	Start	End	E-value
LRRNT	28	58	2.08
LRR	57	76	97.7
LRR	77	100	1.86
LRR	101	124	4.2
LRR	149	169	495
LRR	174	192	20.3
LRR	201	225	20.3
LRR	372	393	318
LRR	398	423	26.2
LRR	470	494	101
LRR	495	518	0.249
LRR	519	535	509
LRR	543	566	225
LRRCT	579	629	2.52e-7
Transmembrane region	635	657	N/A
TIR	674	819	5.32e-41

Table.2 Transmembrane structures of *Camelus dromedarius* TLR5 predicted domain, repeats and motifs done by SMART tool

Name	Start	End	E-value
Low complexity	13	26	N/A
Low complexity	81	89	N/A
LRR TYP	106	129	0.0000529
LRR	156	174	93.5
LRR	181	207	587
LRR TYP	323	346	0.0205
LRR	347	373	124
LRR	371	394	57.3
LRR	395	416	201
LRR TYP	513	536	0.00016
LRR	537	560	14.9
LRR-CT	591	642	3.93e-9
Transmembrane region	656	678	N/A
TIR	704	851	1.66e-28

Table.3 Transmembrane structures of *Camelus dromedarius* TLR6 predicted domain, repeats and motifs done by SMART tool

Name	Start	End	E-value
LRR	110	133	0.286
LRR	411	430	221
LRR	485	506	57.3
LRR	507	531	16.2
LRR-CT	564	618	8.11e-11
Transmembrane region	622	644	N/A
TLR	676	819	5.53e-39

Table.4 Transmembrane structures of *Camelus dromedarius* TLR7 predicted domain, repeats and motifs done by SMART tool

Name	Start	End	E-value
LRRNT	33	69	2.08
LRR	126	149	11.2
LRR	171	193	395
LRR	203	226	14.9
LRR	224	245	93.5
LRR_TYP	289	312	0.00894
LRR	313	334	85.9
LRR	339	364	455
LRR	396	416	53.1
LRR	420	440	353
LRR	497	516	240
LRR	517	541	75.9
LRR	542	568	9.81
LRR	596	619	276
LRR_TYP	650	673	0.000973
LRR_TYP	699	722	0.0000608
LRR	724	746	38.7
LRR	747	768	306
LRRCT	784	835	0.00221
Transmembrane region	844	866	N/A
TIR	891	1037	2.41e-7

Table.5 Transmembrane structures of *Camelus dromedarius* TLR8 predicted domain, repeats and motifs done by SMART tool

Name	Start	End	E-value
LRR	59	80	169
LRR_TYP	120	143	0.0822
LRR	165	194	261
LRR	195	234	19.8
LRR	281	304	4.57
LRR	305	329	103
LRR	331	356	128
LRR	361	384	137
LRR	388	411	11.9
LRR	412	433	128
LRR	519	542	336
LRR	573	597	16.2
LRR_TYP	627	650	0.00447
LRR	676	699	10.6
LRR	700	723	6.77
LRR	724	744	201
LRRCT	761	812	7.75e-8
Transmembrane region	815	837	N/A
TIR	868	1014	0.0000616

Table.6 Transmembrane structures of *Camelus dromedarius* TLR-9 predicted domain, repeats and motifs done by SMART tool

Name	Start	End	E-value
ADF	11	139	1.69e-22
ADF	184	312	1.04e-8
LRR	415	434	0.108
LRR	491	510	138
LRR	512	532	73.3
LRR	576	599	137
LRR	600	628	78.5
LRR	656	679	89.8
LRR	683	706	19.1
LRR	769	787	73.8
LRR	788	810	153
LRR	813	836	0.465
LRR	837	860	301
LRR	867	895	105
LRR	920	943	35.5
LRR	969	988	147
LRR	993	1022	4.3
LRR	1018	1040	27.6
LRR	1042	1069	60.4
TIR	1160	1307	0.00121

Table.7 Transmembrane structures of *Camelus dromedarius* TLR10 predicted domain, repeats and motifs done by SMART tool

Name	Start	End	E-value
LRR	71	94	6.96
LRR	116	139	14.9
LRR	371	395	7.16
LRR	441	464	261
LRR	465	487	1.73
LRRCT	522	575	9.54e-8
Transmembrane region	577	599	N/A
TIR	633	778	1.94e-34

Fig.1 Domain organization of camel TLR4-10, defined by SMART analysis (Green box-LRRs, blue box- TM domain, pink box- low complexity region, ADF- actin depolymerization factor)

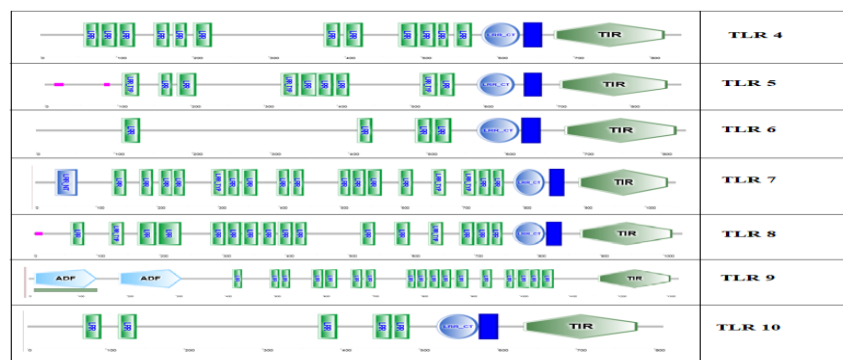
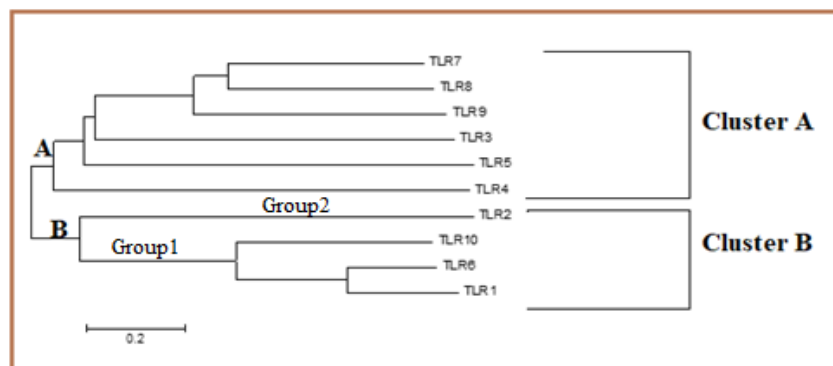


Fig.2 Phylogenetic analysis of camel TLR genes 1–10 at amino acid level by Neighbour-Joining method



The TLR4 protein was consisted of N-terminal LRR domain of around 31 amino acids (aa) (LRR-NT, residues 28-58) followed by twelve LRR domains (residue 57-566) and

one C – terminal LRR domain of around 51 aa (LRR-CT, residues 579-629) in the extracellular region. The transmembrane domain (TM) was composed of 23 aa (residue

635-657) and TIR domain of around 146 aa long (residues 674-819) in the cytoplasmic region as shown in Fig.1 and Table1. The TIR domain of camel was found to be similar to that of pig, horse, human and orangutan (Shinkai *et al.*, 2006).

Camel TLR5 had two Low Complexity Regions (residues 13-26, 81-89). Some regions highly enriched in one or a few amino acids, known as low complexity regions (LCRs), are strikingly abundant in protein sequences and their composition is an important determinant of the evolutionary pattern (O'Neill *et al.*, 2013). In the ectodomain region, nine LRRs domains were located (between 106-560 aa) followed by one C – terminal LRR domain of around 52 aa (LRR-CT, residues 591-642). The TM region was composed of 23 aa (from position 656-678) and a TIR domain of 148 aa (residues 704-851) in the cytoplasmic region as shown in Fig1 and Table2. TLR6 was made up of four LRR domains (residue 110-531aa) tracked by one C – terminal LRR domain of around 55 aa (LRR-CT, residues 564-618) in the ectodomain region. The TM region was composed of 23 aa (residue 622-644) and a TIR domain of 144 aa (residues 676-819) was in the cytoplasmic region (Fig.1, Table3). A total of seventeen LRRs were found in camel TLR7. The ectodomain region consisted of an N-terminal LRR domain of 37 aa (LRRNT, residues 33-69) followed by LRR domains (between 126th-768thaa) and C-terminal LRR domain of around 52 aa (LRR-CT, residues 784-835). The TM region was composed of 23 aa (from position 844-866). The cytoplasmic region which comprised the TIR domain was about 147 aa long (residues 891-1037) (Fig1. and Table4).

The camel TLR8 consisted of C – terminal LRR domain of around 52aa (LRR-CT, residues 761-812) followed by sixteen LRR domains (between 59-744) whereas, the TM region was composed of 23 aa (from position

815-837). The cytoplasmic region contained a TIR domain of 147 aa (residues 868-1014) (Fig1. and Table5). A total seventeen number of LRR domains were located in camel TLR9. It contained two actin depolymerization factors (ADF) (residue 11-139, 184-312) at N-terminal followed by LRRs domains (between 415-1069 aa) in the ectodomain region. ADF/cofilins are small actin-binding proteins composed of a single ADF-H domain. They bind both actin-monomers and filaments, promote rapid filament turnover in cells by depolymerizing/fragmenting actin filaments. The role of the actin cytoskeleton is in controlling receptor compartmentalization, dynamics, and clustering as a means to regulate receptor signalling through controlling the interactions with protein partners (Mattila *et al.*, 2016). Polymerization of the actin cytoskeleton has been found to be essential for B-cell activation (Hao and August, 2005). The transmembrane region was not presented in TLR9. In the cytoplasmic region, TIR domain was composed of 148 aa (residues 1160-1307) (Fig1. and Table6). A total of five LRRs were found in camel TLR 10. The extracellular region consisted of LRRs domains (between position 71-487) followed by C – terminal LRR domain around 54 amino acids (LRR-CT, residues 522-575). TM region was made up of 23 aa (from position 577-599) however, the TIR domain was 146 aa long which was situated (residues 633-778) in the cytoplasmic region (Fig1. and Table7).

These analyses presented that; number of domains were different among different TLR genes in camel ranging from 4 to 17 (Fig. 1). This supported the role of LRRs in specific ligand recognition. The results also showed that the viral TLR genes (TLR7, 8, 9) tended to have more numbers of LRRs (17-18) than the nonviral TLR genes. After the SMART analysis, we concluded that TM domain in TLRs (4-10) was exactly similar in length (23aa). However, the TIR domain was

constructed by an almost equal number of amino acids (140-144aa). It means the variation in length was mainly due to ligand binding LRR ectodomain region. In a previous study, a point mutation in the TIR domain of C3H/HeJ mouse strain created dominant negative effect in TLR-mediated immune response suggesting its importance in signalling (Hoshino *et al.*, 2014).

Phylogenetic analysis

To define TLR subfamilies and demonstrate the evolutionary history of camel TLR members, a phylogenetic tree was constructed. The Neighbour-joining method algorithm which is a distance-based method was chosen as the tool to evaluate phylogenetic relationship among all the TLRs (1-10) genes amino acid sequences of coding region.

As shown in Fig. 2, two clades, cluster-A and cluster-B, were obtained by the evolutionary analysis. It revealed TLR3, 4, 5, 7, 8 and 9 gene in cluster-A and the other TLR1, 2, 6 and 10 gene in cluster-B. Genes of cluster-B belong to TLR2 family whereas, TLR7, 8 and 9 belong to TLR7 family (Banerjee *et al.*, 2012; Roach *et al.*, 2005). The TLR7 and 8 genes were more closely related to each other which in turn was in close association with TLR9. The TLR3 was a close member of the clade TLR7 (TLR7/8/9) family. In clade of TLR2 family, TLR2 gene is related to TLR1, 6 and 10 via a single branch, whereas TLR10 is more distantly related to both TLR1 and 6 genes. Studies to date have shown that all known TLRs from group2 are able to form functional heterodimers with the known TLRs from group1 on the cell surface, and together these TLRs recognize hydrophobic molecules such as lipids and lipoproteins (Guan *et al.*, 2010; Keestra *et al.*, 2007). The phylogenetic tree showed that genes were present in single family close to one another indicating that

these genes might have arisen from the gene duplication because they were present on the same chromosome (chr) *i.e.* TLR4, 6 located on the 2nd chromosome and TR7, 8 on the chromosome X. This same grouping as clades may have also been due to the more conserved nature of the TIR domains that play a role in cell signalling. In previous study, documented that TLRs 7, 8 and 9 are endosomal and TLRs 1, 6 and 10 acts as heterodimers for TLR ligand binding, so they belonged to the same group respectively.

In conclusion to our knowledge, this is the first study of the determination of domain architecture and phylogenetic analysis of camel TLRs (4-10). This information may be useful for studies on evolutionary lineages, and immune functions associated with bacterial, viral and fungal infections of camels.

References

- Akira S, Takeda K. Toll-like receptor signalling. *Nature reviews immunology*. 2004; 4(7): 499-511.
- Areal H, Abrantes J, Esteves PJ. Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC evolutionary biology*. 2011; 11(1): 368.
- Banerjee P, Gahlawat SK, Joshi J, Sharma U, Tantia MS, Vijh RK. Sequencing, Characterization and Phylogenetic analysis of TLR genes of *Bubalus bubalis*. *DHR-IJBLS*. 2012; 3: 137-58.
- Barton GM, Kagan JC. A cell biological view of Toll-like receptor function: regulation through compartmentalization. *Nature Reviews Immunology*. 2009; 9(8): 535-42.
- Bowie AG, Haga IR. The role of Toll-like receptors in the host response to viruses. *Molecular immunology*. 2005; 42(8): 859-67.
- Dhar D, Dey D, Basu S. Insights into the evolution of extracellular leucine-rich repeats in metazoans with special reference to Toll-like receptor 4. *Journal of biosciences*. 2019;44(1):18.

- Dowling JK, Dellacasagrande J. Toll-like receptors: ligands, cell-based models, and readouts for receptor action. In *Toll-Like Receptors*. 2016; 3-27. Humana Press, New York, NY.
- Guan Y, Ranao DR, Jiang S, Mutha SK, Li X, Baudry J, Tapping RI. Human TLRs 10 and 1 share common mechanisms of innate immune sensing but not signaling. *The Journal of Immunology*. 2010; 184(9): 5094-103.
- Hao S, August A. Actin depolymerization transduces the strength of B-cell receptor stimulation. *Molecular biology of the cell*. 2005; 16(5): 2275-2284.
- Hoshino K, Takeuchi O, Kawai T, Sanjo H, Ogawa T, Takeda Y, Takeda K, Akira S. Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product. *The Journal of Immunology*. 1999; 162(7): 3749-52.
- Kannaki TR, Shanmugam M, Verma PC. Toll-like receptors and their role in animal reproduction. *Animal reproduction science*. 2011; 125(1-4): 1-2.
- Kawai T, Akira S. Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. *Immunity*. 2011; 34(5): 637-50.
- Keestra AM, de Zoete MR, Bouwman LI, Vaezirad MM, van Putten JP. Unique features of chicken Toll-like receptors. *Developmental & Comparative Immunology*. 2013; 41(3): 316-23.
- Keestra AM, de Zoete MR, van Aubel RA, van Putten JP. The central leucine-rich repeat region of chicken TLR16 dictates unique ligand specificity and species-specific interaction with TLR2. *The Journal of Immunology*. 2007; 178(11): 7110-9.
- Mattila PK, Batista FD, Treanor B. Dynamics of the actin cytoskeleton mediates receptor cross talk: An emerging concept in tuning receptor signaling. *Journal of Cell Biology*, 2016; 212(3): 267-280.
- Netea MG, Wijmenga C, O'Neill LA. Genetic variation in Toll-like receptors and disease susceptibility. *Nature immunology*. 2012; 13(6):535-42.
- O'Neill LA, Golenbock D, Bowie AG. The history of Toll-like receptors—redefining innate immunity. *Nature Reviews Immunology*. 2013; 13(6): 453-60.
- Radó-Trilla N, Albà M. Dissecting the role of low-complexity regions in the evolution of vertebrate proteins. *BMC evolutionary biology*. 2012 Dec 1;12(1):155.
- Roach JC, Glusman G, Rowen L, Kaur A, Purcell MK, Smith KD, Hood LE, Aderem A. The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences*. 2005; 102(27): 9577-82.
- Shinkai H, Tanaka M, Morozumi T, Eguchi-Ogawa T, Okumura N, Muneta Y, Awata T, Uenishi H. Biased distribution of single nucleotide polymorphisms (SNPs) in porcine Toll-like receptor 1 (TLR1), TLR2, TLR4, TLR5, and TLR6 genes. *Immunogenetics*. 2006; 58(4): 324-30.
- Wallet SM, Puri V, Gibson FC. Linkage of infection to adverse systemic complications: periodontal disease, toll-like receptors, and other pattern recognition systems. *Vaccines*. 2018; 6(2): 21.
- Werling D, Jann OC, Offord V, Glass EJ, Coffey TJ. Variation matters: TLR structure and species-specific pathogen recognition. *Trends in immunology*. 2009; 30(3): 124-30.

How to cite this article:

Jyoti Choudhary, Mukul Purva, Mahendra Milind, Kritika Gahlot, Taruna Bhati, Sunil Maherchandani and Amit Pandey. 2021. Phylogenetic and Protein Domain Architecture Analysis of Toll-like Receptor 4-10 Genes in Camel (*Camelus dromedarius*). *Int.J.Curr.Microbiol.App.Sci*. 10(01): 2887-2895.
doi: <https://doi.org/10.20546/ijcmas.2021.1001.335>