



Fertility of hybrids of dromedary and Bactrian camels: A possible role of conserved architecture of zinc finger domain of recombination regulator *PRDM9*

SONIKA AHLAWAT¹, REKHA SHARMA¹, REENA ARORA¹, HIMANI SHARMA¹, RENUKA SEHRAWAT¹,
ANNU SHARMA¹, KARAN VEER SINGH¹ and RAMESH KUMAR VIJH¹

ICAR-National Bureau of Animal Genetic Resources, Karnal, Haryana 132 001 India

Received: 23 July 2021; Accepted: 3 October 2022

ABSTRACT

Recombination regulator, *PRDM9*, has been regarded as the most rapidly evolving gene in the genomes of many metazoans, in addition to being acknowledged as the sole speciation gene in vertebrates. It has become the focus of many scientific investigations because of exceptional numerical and sequence variability in its zinc finger (ZF) domain within and across species that contributes to reproductive isolation between species. This study is the maiden attempt to explore the architecture of *PRDM9* ZF domain in two Camelid species (*Camelus dromedarius* and *Camelus bactrianus*). Sequence analysis revealed highly conserved domain architecture with presence of 3 and 4 ZFs in dromedary and Bactrian camels, respectively. Typical evolutionary features of *PRDM9* ZF domain i.e. concerted evolution and positive selection were invariably absent in both the one-humped dromedary and the two-humped Bactrian camels. Fertility of hybrids of dromedary and Bactrian camels, despite being taxonomically distinct species can be attributed to the lack of sequence variability in *PRDM9* in these species. Phylogenetic analysis underpinned clear demarcation of camels from other livestock species. The results of the present study defy what has been learnt so far about *PRDM9* and add to the enigma surrounding the most intriguing gene in the genome.

Keywords: Bactrian, Camel, Dromedary, *PRDM9*, Recombination, Zinc finger

Meiotic recombination regulator, *PRDM9*, is considered the most enigmatic gene in the genome across diverse taxa. By virtue of being the determinant of the location of recombination hotspots during crossing over between maternal and paternal chromosomes, it has assumed paramount significance in evolution of eukaryotic genomes (Baudat *et al.* 2013). Of greatest interest over the last decade has been the DNA binding domain of *PRDM9*, which is a coding minisatellite harboring a tandem array of C2H2 zinc fingers (ZF) that establishes the positioning of double strand break sites during meiosis. High variability in the copy number of ZF repeats and remarkable sequence polymorphism in diverse species have contributed to its recognition as the fastest evolving gene in the genome (Ponting 2011). Interestingly, it is thus far, the sole gene identified in vertebrates that contributes to the process of speciation or reproductive isolation between species (Mihola *et al.* 2009). Therefore, the attributes that have made *PRDM9* the focus of many scientific investigations over the last decade are well comprehensible.

Exploration of diversity of *PRDM9* in mice, humans, primates, equines, small and large ruminants has unraveled many facets of this multihued gene (Buard *et al.* 2014,

Myers *et al.* 2010, Groeneveld *et al.* 2012, Steiner and Ryder 2013, Ahlawat *et al.* 2016a, 2017). These include immense numerical and sequence variability in the ZFs and signals of positive selection and concerted evolution in the ZF domain across species. The ZF of *PRDM9* is typically 28 amino acids long and each sequential ZF of the minisatellite array has affinity for sequential trinucleotide on the target DNA molecule. This binding leads to recruitment of double strand break machinery during meiotic crossing over (Baudat *et al.* 2013). Rapid turnover of hotspot locations has been considered to be the driving factor for swift evolution of *PRDM9* zinc fingers, a phenomenon referred to as 'recombination hotspot paradox' (Baudat *et al.* 2010). Moreover, hybrid sterility observed in certain mouse subspecies is considered to be the manifestation of allelic differences in *PRDM9* (Mihola *et al.* 2009). Surprisingly, the reason for absence or disruption of *PRDM9* from the genomes of birds, fruit fly and canids is still inscrutable (Oliver *et al.* 2009).

Despite huge interest in deciphering the evolutionary features of *PRDM9* across different metazoans, there is still no information available from camels. Camels are regarded as sustainable livestock genetic resources despite their reduced relevance as draft power in the current scenario. Resilience to climate change and high production efficiency of camels has ensured their survival in diverse ecosystems against the harsh reality of destruction of their

Present address: ¹ICAR-National Bureau of Animal Genetic Resources, Karnal, Haryana. ✉Corresponding author email: Sonika.Ahlawat@icar.gov.in

habitats (Satyanarayana *et al.* 2022). They constitute an important means of transport in the desert landscapes, in addition to being acknowledged for their unique ability to tolerate hostile conditions, unusual immune system and beneficial attributes of their milk (Ali *et al.* 2019, Sharma *et al.* 2020). India has 9 registered breeds of the dromedary or one-humped camels that inhabit the hot deserts of arid and semi arid zones of India (www.nbagr.res.in). Additionally, a few Bactrian or two-humped camels (<200 in number) that endure the cold deserts of the northern temperate region are also present in the country (Sharma *et al.* 2018). This study is the first report on characterization of ZF domain of *PRDM9* in the two camel species, *Camelus dromedarius* and *Camelus bactrianus*.

MATERIALS AND METHODS

Samples: Blood samples were collected from 20 animals each of all the 9 defined breeds of dromedary camels (Bikaneri, Jaisalmeri, Jalori, Kharai, Kutchi, Malvi, Marwari, Mewari and Sindhi) and 8 Bactrian camels from their respective breeding tracts in EDTA containing vacutainers with the help of trained veterinarians. Sampling of unrelated animals for each breed/population was ensured to cover maximum genetic diversity. Genomic DNA was extracted from the blood samples according to the protocol by Sambrook *et al.* (1989). The quality and quantity of the isolated DNA were determined by running on agarose gels as well as by Nanodrop spectrophotometer (ND-1000, Thermo Scientific).

PCR amplification of *PRDM9* ZF array: *PRDM9* ZF array was amplified by designing primers (Forward: CACCTGAACAACAGCCATTAAGCTCC and Reverse: GGTGGCCTTATCGCTGAAGCTTCG) on the basis of *Camelus dromedarius* gene sequence (GenBank accession JWIN03000010.1). PCR cocktail consisted of 12.5 µL of DreamTaq Green PCR Mastermix (2×), 100 ng of DNA, and 1 mM of each primer and nuclease free water to make the final volume to 25 µL. PCR cycling conditions included initial denaturation at 95°C for 5 min; 32 cycles of denaturation at 95°C for 30 sec, annealing at 63°C for 30 sec, extension at 72°C for 30 sec, and final extension at 72°C for 7 min. PCR products were run along with 100 plus 500 bp DNA marker on 1% agarose gel, followed by staining with ethidium bromide. Specific amplicon (approximately 1100 bp) was obtained using the designed primers in all 188 samples. To decipher the sequence variability in the ZF domain, the amplified fragments from 62 animals (six animals each of 9 Indian breeds and 8 Bactrian camels) were enzymatically purified using Exonuclease I and Antarctic Phosphatase and subsequently sequenced in ABI 3100 Automated DNA Sequencer (Applied Biosystems, USA). The sequences of this study were submitted to NCBI and accession numbers were obtained for camel *PRDM9* ZF domain (MT926082- MT926116).

Sequence analysis: All 62 camel sequences were trimmed so as to include only the ZF portion of *PRDM9* for further analysis using BIOEDIT 7.0 sequence alignment

editor (Hall 1999). The sequences were then subjected to multiple sequence alignment using CLUSTAL OMEGA to analyze sequence variations. Sequence Manipulation Suite Version 2 was used to translate the DNA sequences (Stothard 2000). SMART (<http://smart.emblheidelberg.de>) was used for functional domain prediction of the camel *PRDM9* protein. *PRDM9* sequences for other species that emerged as the best hits in BLAST analysis were retrieved from NCBI. For construction of phylogenetic tree, all the sequences of this study and those taken from NCBI were subjected to multiple sequence alignment using MUSCLE tool found in MEGA6 (<http://www.megasoftware.net>), and subsequently, Maximum Likelihood method was used for the construction of the evolutionary tree (Tamura and Nei 1993).

RESULTS AND DISCUSSION

In this study, amplification of ZF domain of *PRDM9* in 188 camels (180, dromedary and 8, Bactrian) yielded amplicons of the same size possibly suggesting lack of numerical variability in the ZF repeats. Sequence analysis of the amplified products revealed an intriguing observation that not only was the size of *PRDM9* ZF domain same in all animals, the DNA as well as the amino acid sequence was also found to be highly conserved in dromedary camels (Fig. 1).

However, there were 11 nucleotide variations between dromedary and Bactrian camels, out of which 5 were non-synonymous mutations. This observation of lack of sequence variability in ZF domain is contrary to reports from many investigated vertebrate species, including livestock wherein tremendous variability in the ZF repeats has been a hallmark of evolution of *PRDM9*. The translated sequences were subjected to SMART to predict the functional domains in the encoded protein. Because of non-synonymous SNPs in Bactrian camels, the ZF domain was observed to possess 3 C2H2 ZFs in all the dromedary camels and 4 in Bactrian camels.

These ZFs were composed of 23 amino acids in both species. The location of ZFs has been highlighted in Fig. 1. This again was a fascinating finding since, in all the other species investigated so far, ZF array contains tandem repeats of 84 nucleotides or 28 amino acids. These repeats have been documented to exhibit perfect homology at both DNA and protein levels in many species (Baudat *et al.* 2010). Amino acids at positions -1, 3, and 6 in the ZF domain which are involved in binding to bases 3, 2, and 1, respectively, in the primary DNA strand during recombination are known to be under positive selection (Oliver *et al.* 2009). Since, camel ZF domain was highly conserved, it is important to highlight that concerted evolution and positive selection consistently typifying the evolutionary features of *PRDM9* in different species were not evident in our study in camels.

Subsequently, we carried out phylogenetic analysis of the camel *PRDM9* sequences with other species that were observed as the best hits in BLAST analysis. To evaluate



Fig. 1. Multiple sequence alignment of *PRDM9* ZF domain in camels. ZFs have been highlighted in boxes. Since the sequence was highly conserved, only two animals of each breed of dromedary camel and 2 Bactrian camels have been included in the figure.

evolutionary relationship with other livestock species, the sequences from large ruminants (cattle, buffalo, yak and mithun), small ruminants (sheep and goat) and equines (horse, asinus and zebra) were also included in the analysis. In the evolutionary tree, camels appeared to be the most distinct species forming a separate cluster. Astoundingly, camels showed proximity to Siberian tiger and high sequence dissimilarity with various livestock species (Fig. 2). Although all ruminant species clustered together, their high genetic distance with camels was conspicuous in the analysis. Percent identity was maximum with Siberian tiger (82.5%) and least with large ruminants such as cattle, yak and mithun (64-65%). As expected, the two camel species shared 98.92% sequence identity.

Comprehensive analysis of *PRDM9* across diverse animal species has unveiled three classes to which these species belong (Ponting 2011). The first class includes species such as humans, chimpanzees, and mice, which encode a ZF domain that displays fast evolution of ZFs in response to short lived recombination hotspots. Therefore, it's no wonder that remarkable diversity in the number of ZFs ranging from 7-17 in bonobos and chimpanzees (Groeneveld *et al.* 2012), 5-14 in equines (Steiner and Ryder 2013), 7-17 in mice (Buard *et al.* 2014), 8-12 in small

ruminants (Ahlawat *et al.* 2016a), 6-9 in large ruminants (Ahlawat *et al.* 2017), 8-16 in humans (Baudat *et al.* 2010), 7-15 in primates and 7-12 in rodents (Oliver *et al.* 2009), with positive selection at DNA binding amino acids and concerted evolution have been consistently observed. In contrast, there are species such as ray-finned fishes and tunicates that have persistent recombination hotspots and consequently, have gradually evolving ZF array and therefore, comprise the second class. Intriguingly, the third class includes chicken, fruit fly and dog, in which *PRDM9* gene is either absent or is disrupted by multiple stop codons. In case of livestock species, remarkable diversity has been reported in *PRDM9* which substantiates the reports from other vertebrates, but multiple disruptive mutations have been reported in its paralog, *PRDM7* (Ahlawat *et al.* 2016b). However, the results of our study in camels do not support their conformity to any of the known categories because neither could we observe any variability in the number/sequence of ZF array nor its coding sequence was disrupted. Does this imply that on the basis of domain architecture of *PRDM9* ZF array, camels represent an altogether separate class? Species such as humans and mice which have rapidly evolving *PRDM9* ZF domain have recombination hotspots localized in discrete clusters,

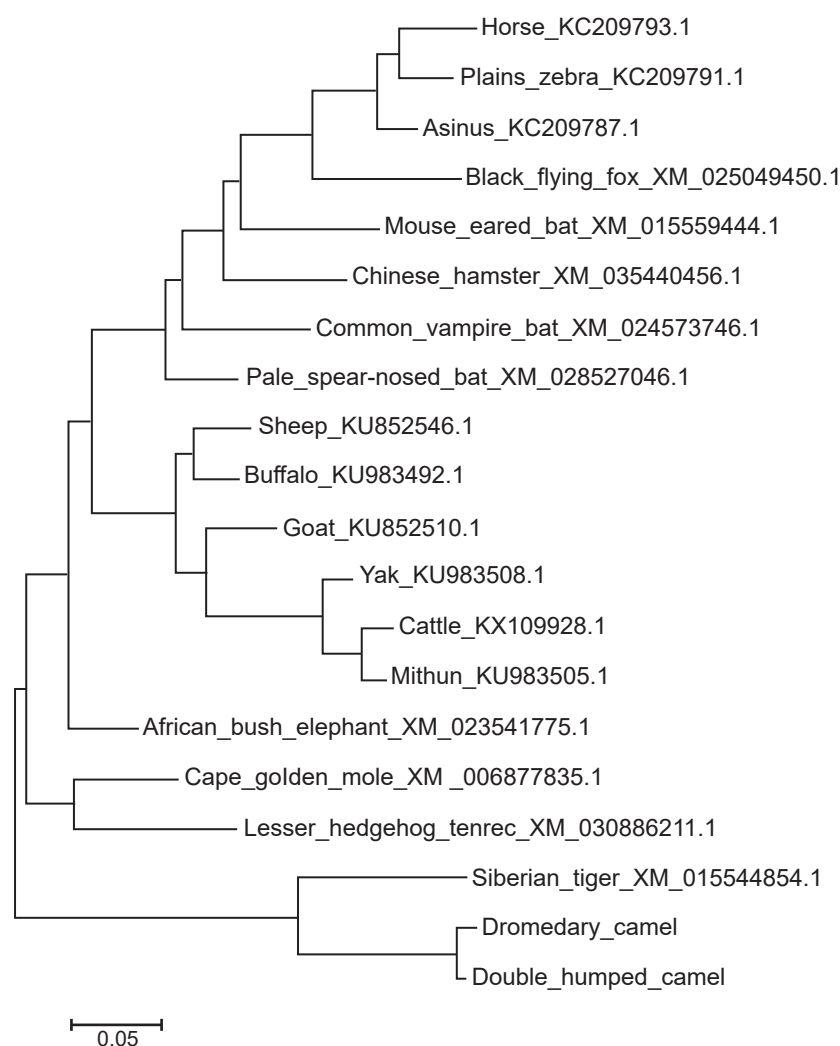


Fig. 2. Inferred phylogeny of camel *PRDM9* and its orthologs.

whereas mammals like canids that harbour functionally inert versions of *PRDM9* have more uniform distribution of crossover sites in the genome (Baudat *et al.* 2013, Axelsson *et al.* 2012). Since camels have a conserved *PRDM9* ZF domain, it would be interesting to explore its implications in determining the recombination landscape in future.

Till date, *PRDM9* is the only gene that has been implicated in hybrid sterility in some mouse subspecies. Variability in the *PRDM9* ZF domain in *Mus m. domesticus* and *Mus m. musculus* has been identified to be a key determinant of the reproductive isolation (Mihola *et al.* 2009). There are many other instances where inter species breeding is practiced to exploit hybrid vigour, which is manifested in the form of higher productivity, better draftability and greater adaptability of hybrids. Common examples from livestock species are the crosses of horses and donkeys to produce mules as well as interbreeding of cattle (*Bos taurus*) and yaks (*Bos grunniens*) (Steiner and Ryder 2013, Ahlawat *et al.* 2017). In both these cases, the male hybrids are sterile. However, crossbreeding of the dromedary and Bactrian camel has been attempted since ages and is still routinely carried out in countries like Turkey and

Kazakhstan. Despite being taxonomically distinct species, they produce fertile F1 hybrids, which in turn are used to produce F2 and F3 hybrids (Dioli *et al.* 2020). Thus, there is a possibility that conserved architecture of *PRDM9* has a role to play in ensuring the reproductive capability of dromedary and Bactrian camel hybrids.

Various physiological and genomic studies have deciphered many inimitable attributes that enable camels to survive in the inhospitable desert conditions. These include storage of energy in the form of fat in their humps, ability to endure severe dehydration and two folds higher blood glucose levels than other ruminants, etc to underscore a few. Moreover, camels are the only mammals which have unusually elliptical shaped enucleated erythrocytes and produce non-conventional heavy-chain antibodies which are devoid of light chains (Ali *et al.* 2019, Tillib *et al.* 2014). It is thus a surprise that camels harbour such unique features despite their small genome size (2.45 Gb) as compared to other mammals. Other prominent differences are lower repetitive DNA (34%) and far greater number of 'rapidly evolving genes' in their genome that are involved in metabolism (carbohydrate and lipid) and

signaling pathways (insulin and adipocytokine) (Jirimutu *et al.* 2012). Surprisingly, our study on the fastest evolving gene in the genome (*PRDM9*) presents lack of variability in camels. Since genetic recombination has great implications in evolution and adaptation, it is plausible to speculate the role of *PRDM9* in the latter in camels, if not, in the former.

Put together, the present study for the first time reports absence of sequence variability in the *PRDM9* ZF domain in two camelid species (Dromedary and Bactrian), despite numerical variability between them. Our results are not in concordance with the evolutionary dynamics of *PRDM9* that has been reported in other metazoans. We put forth an important question as to what can be the possible role of *PRDM9* in the evolution of the camel genome. It would be exciting for researchers in future to explore what evolutionary advantages the conserved architecture of *PRDM9* imparts to camels.

ACKNOWLEDGEMENTS

This study was made possible by the financial support provided by Indian Council of Agricultural Research. The authors gratefully acknowledge the facilities provided by Director, ICAR-National Bureau of Animal Genetic Resources for carrying out the research work.

REFERENCES

- Ahlawat S, De S, Sharma P, Sharma R, Arora R, Kataria R S, Datta T K and Singh R K. 2017. Evolutionary dynamics of meiotic recombination hotspots regulator *PRDM9* in bovids. *Molecular Genetics and Genomics* **292**: 117–31.
- Ahlawat S, Sharma P, Sharma R, Arora R and De S. 2016a. Zinc finger domain of the *PRDM9* gene on chromosome 1 exhibits high diversity in ruminants but its paralog *PRDM7* contains multiple disruptive mutations. *PLoS ONE* **11**(5): e0156159.
- Ahlawat S, Sharma P, Sharma R, Arora R, Verma N K, Brahma B, Mishra P and De S. 2016b. Evidence of positive selection and concerted evolution in the rapidly evolving *PRDM9* zinc finger domain in goats and sheep. *Animal Genetics* **47**: 740–51.
- Ali A, Baby B and Vijayan R. 2019. From Desert to Medicine: A review of camel genomics and therapeutic products. *Frontiers in Genetics* **10**: 17.
- Axelsson E, Webster M T, Ratnakumar A, Ponting C P and Lindblad-Toh K. 2012. Death of *PRDM9* coincides with stabilization of the recombination landscape in the dog genome. *Genome Research* **22**: 51–63.
- Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G and de Massy B. 2010. *PRDM9* is a major determinant of meiotic recombination hotspots in humans and mice. *Science* **327**: 836–40.
- Baudat F, Imai Y and de Massy B. 2013. Meiotic recombination in mammals: localization and regulation. *Nature Reviews Genetics* **14**: 794–806.
- Buard J, Rivals E, Dunoyer de Segonzac D, Garres C, Caminade P, de Massy B and Boursot P. 2014. Diversity of *PRDM9* zinc finger array in wild mice unravels new facets of the evolutionary turnover of this coding minisatellite. *PLoS One* **9**: e85021.
- Burger P A, Ciani E and Faye B. 2019. Old World camels in a modern world – a balancing act between conservation and genetic improvement. *Animal Genetics* doi: 10.1111/age.12858.
- Dioli M. 2020. Dromedary (*Camelus dromedarius*) and Bactrian camel (*Camelus bactrianus*) crossbreeding husbandry practices in Turkey and Kazakhstan: An in-depth review. *Pastoralism: Research, Policy and Practice* **10**: 6.
- Groeneveld L F, Atencia R, Garriga R M and Vigilant L. 2012. High diversity at *PRDM9* in chimpanzees and bonobos. *PLoS One* **7**: e39064.
- Hall T A. 1999. BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95–98.
- Jirimutu, Wang Z, Ding G, Chen G, Sun Y, Sun Z *et al.* 2012. Genome sequences of wild and domestic bactrian camels. *Nature Communications* **3**: 1202.
- Mihola O, Trachtulec Z, Vlcek C, Schimenti J C and Forejt J. 2009. A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science* **323**: 373–75.
- Myers S, Bowden R, Tumian A, Bontrop R E, Freeman C, MacFie T S, McVean G and Donnelly P. 2010. Drive against hotspot motifs in primates implicates the *PRDM9* gene in meiotic recombination. *Science* **327**: 876–79.
- Oliver P L, Goodstadt L, Bayes J J, Birtle Z, Roach K C, Phadnis N, Beatson S A, Lunter G, Malik H S and Ponting C P. 2009. Accelerated evolution of the *PRDM9* speciation gene across diverse metazoan taxa. *PLoS Genetics* **5**: e1000753.
- Ponting C P. 2011. What are the genomic drivers of the rapid evolution of *PRDM9*? *Trends in Genetics* **27**(5): 165–71.
- Sambrook J, Fritsch E F and Maniatis T. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbour Laboratory, New York.
- Satyanarayana D S, Ahlawat S, Sharma R, Arora R, Sharma A, Tania M S and Viji R K. 2022. Mitochondrial DNA diversity divulges high levels of haplotype diversity and lack of genetic structure in the Indian camels. *Gene* **820**: 146279.
- Sharma R, Ahlawat S, Sharma H, Prakash V, Shilpa, Khatak S, Sawal R K and Tania M S. 2020. Identification of a new Indian camel germplasm by microsatellite markers based genetic diversity and population structure of three camel populations. *Saudi Journal of Biological Sciences* **27**: 1699–1709.
- Sharma R, Sharma H, Ahlawat S, Aggarwal R A K, Vij P K and Tania M S. 2018. First attempt on somatic cell cryopreservation of critically endangered *Camelus bactrianus* of India. *Gene Reports* **10**: 109–115.
- Steiner C C and Ryder O A. 2013. Characterization of *PRDM9* in equids and sterility in mules. *PLoS One* **8**: e61746.
- Stothard P. 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *BioTechniques* **8**: 1102–04.
- Tamura K and Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10**: 512–26.
- Tillib S V, Vyatchanin A S and Muyldermans S. 2014. Molecular analysis of heavy chain-only antibodies of *Camelus bactrianus*. *Biochemistry* **79**: 1382–90.