

MAPPING THE HAPLOGROUP OF INDIAN AND ZOROASTRIAN - PARSI MITOCHONDRIAL GENOME AND PHYLOGENETIC ANALYSIS TO DETECT THE INCIDENCES OF ADMIXTURE

Jennifer John¹, Deepika Iyyengar², L, Kiran Kumar D³, Manohar GM⁴

Post Graduate, Biotechnology Department, Nrupathunga University, Bangalore, India¹

Post Graduate, Biotechnology Department, Nrupathunga University, Bangalore, India²

Associate Professor, BT Department, Nrupathunga University, Bangalore, India³

Secretary cum Mentor, HalliLabs, Ragihalli, Bangalore, India⁴

Abstract: Zoroastrian-Parsi's people belong to the Persian community, who emigrated to India between 785 and 936 A.D. after the Arab invasion. They are the proponents of Zoroastrianism faith and have endogamous culture. Since they have settled in India for a long time we hypothesized that though they were endogamous there might be incidences of admixture with the Indian population. In order to test this, we built a phylogenetic tree using mitochondrial sequences of both the Indian and Parsi population. We used 97 Indian mitochondrial sequences representing different communities from different regions of India and 64 Zoroastrian-Parsi mitochondrial genome sequences from different parts of India and a single African mitochondrial genome sequence as an outgroup. We built a maximum likelihood phylogenetic tree and mapped sequence specific haplogroups and their variants on the phylogenetic tree. The phylogenetic analysis showed admixture to have taken place at many instances. A total of 6 major haplogroups with a maximum of 12 sub-haplogroups were identified as mapped on the tree. The co-clustering of common as well as unique variants of different haplogroups in some clades suggest multiple admixture events to have happened between the Zoroastrian-Parsi and Indian populations, which raise remarks on their endogamous culture.

Keywords: Mitochondrial genome, Haplogroups, Phylogenetic analysis, Admixture.

I. INTRODUCTION

The Parsi community in India are Indo-European followers of the Zoroastrian faith, one of the ancient theism that emerged in the pre-Islamic Persian period [1]. They are the followers of the prophet Zarathustra (Zoroaster in Greek) of Iran [2], who's deemed to have preached the message about the one true God - Ahura Mazda which is detailed in the Gathas, the consecrated Zoroastrian text known as the core of Avesta [3],[4],[5]. Zoroastrians who migrated to Gujarat, India were known as the "Parsis" meaning people from Persia [6],[7]. By the 10th Century this community became a religious minority group and migrated to other regions [8]. Zoroastrian migration to India is known as the New Persian- Qessa-ye Sanjān "Story of Sanjan"- a legendary account of the journey by sea from Iran to India [9],[10]. The Zoroastrians practiced endogamous culture when migrating to India, which preserved their genetic identity to remain restricted among their community. This is attributed as a major reason for the population rate decline. Also their religious belief was against inter-caste marriages, if any individual marries within another community he/she is banished from their sacred beliefs and culture [11]. Bombay became the hub of residence for the Parsi community but the recent population is <52000 in India as per the 2011 Census, Govt of India [12].

TABLE I PARSIS POPULATION IN BOMBAY AND INDIA,1881-2001-KULKS, E (1974)

SL NO.	CENSUS YEAR	PARSIS IN MUMBAI	PARSIS IN INDIA (INCLUDING MUMBAI)	TOTAL POPULATION OF INDIA	PARSIS AS A% OF INDIA'S POPULATION
1.	1881	48,597	85,397	253,891,821	0.034
2.	1891	47,458	88,887	287,314,671	0.031
3.	1901	46,231	93,952	294,361,056	0.032
4.	1911	50,931	100,096	315,156,396	0.032
5.	1921	52,234	101,778	318,942,480	0.032
6.	1931	57,765	111,853	352,837,778	0.032
7.	1941	59,813	114,890	388,997,955	0.030
8.	1951	68,660	111,790	361,088,090	0.031
9.	1961	70,065	100,772	439,234,771	0.023
10.	1971	64,667	91,266	548,159,652	0.017
11.	1981	50,053	71,630	683,329,097	0.010
12.	1991	53,794	76,382	846,302,688	0.009
13.	2001	46,657	69,601	1,028,737,436	0.007

Admixed populations provide intense opportunities for mapping ancestral diseases as well as in studying signatures of selection such as inbreeding in a population [13].The Indian population harbours substantial proportions of genetic diversity and also local admixture incidents[14].Studies reveal that a recent admixture has taken place between the Indian and African population [15]. Asian populations ,the Uygurs in China also report mixed populations[16].Admixture of the Parsis with Indian and Iranian populations are known to have occurred ~1,200 years ago[17].

A. Enumerating incidences of Admixture between Indian and Zoroastrian-Parsi Population through mitochondrial genome analysis.

Human Mitochondrial Genomes - Enroute ancestry findings.

Mitochondrial genome based analysis have become paramount in the recent decades to decipher human evolution , population genetic studies and ancestral disease linkage studies.

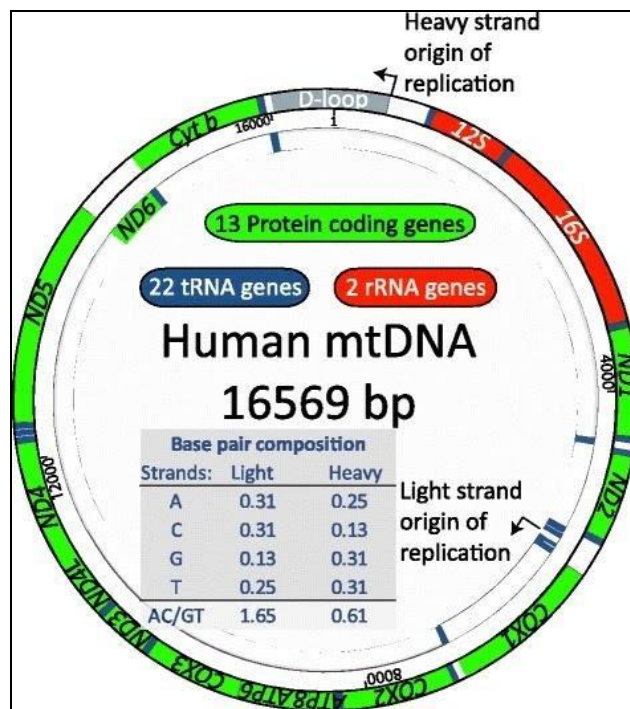


Fig 1: Functional map of Mt DNA -[17]

Human mitochondrial DNA (mt DNA) is a double stranded, circular (16,569 kb) genome of bacterial origin ,which encodes 13 structural polypeptides responsible for oxidative phosphorylation . In addition, 22 tRNAs and 2 rRNAs are also encoded by the mtDNA[18].

Mitochondrial DNA sequences have also played a prime role in enunciating the ancestral origins of European, Asian, African, Iranian, South East Asian, Indian populations and also in reconstructing the phylogenetic trees that have been previously ascribed [19].The characteristic features such as maternal inheritance, high rate of mutation ,lack of recombination, high copy number per cell, offer the opportunity to explore ancestral relationships among individuals and to study matrilineal differences among the human populations at continental and regional level [20],[21],[22],[23]. Lack of genetic recombination in mtDNA results in maternally inherited singlenucleotide polymorphisms(SNPs). Accumulation of these SNPs or more commonly known as variants would result in stable mutations to form haplotypes that become markers for tracing phylogenetic relationships.[24], [25].

B.Haplogroups

A group of haplotypes that share common ancestral polymorphic variations are referred to as haplogroups. In the 1990s the concept of mtDNA haplogroup nomenclature was introduced, with alphabetical labels assigned to different geographical areas [26].The mitochondrial haplogroups are subdivided in to three Top Level Haplogroups of lineages L, M, and N;

Lineage L ("African"): L0, L1, L2, L3, L4, L5, L6.

Lineage M ("Asian"): C, D, E, G, M, Q, Z.

Lineage N ("Eurasian"): A, B, F, H, HV, I, J, K, N, O, P, R, S, T, U, V, W, X, Y.

wherein A to G haplogroups belong to Asian and American lineages , H to K belong to Europe and L confined to the African lineage.

The Iranian mitochondrial DNA belongs to haplogroups HV, TJ, U, N1, N2 and X, commonly found in West Eurasia.The Indian mitochondrial DNA majorly belongs to M, N,H and R along with the subclass M2,M3,M4 and M5 [27].In this study ,we have proposed to analyse the Indian- Zoroastrian-Parsi mitochondrial genomes to pinpoint the proportion of admixture.

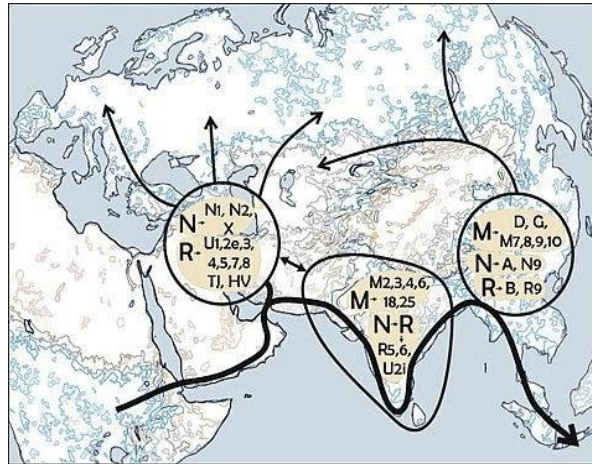


Fig. 2 Haplogroup distribution in Eurasia [27]

II. MATERIALS AND METHODS

A. Collection of Sequences for Analysis

Hundred Indian mitochondrial genome sequences (AY713976 to AY714050 and AY922253 to AY922277) were retrieved from Genbank. These sequences were representative of the following communities: Reddy and Thogataveera from Andhrapradesh, South India; Bhargava, Chaturvedi and other Brahmin from Uttarpradesh, North India; Rajbhansi from West Bengal and Khasi population from Meghalaya both located in Northeast India [28], [29]. 105 Zoroastrian-Parsi mitochondrial genome sequences (MT506242 to MT506346) were retrieved from GenBank. The sequences had been collected from samples of the resident Parsi population of Hyderabad, Secunderabad, and Bangalore, India [12] and [28]. An African mitochondrial genome sequence belonging to the zulu community was chosen as the outgroup and retrieved from Genbank (MH981686) [30].



Fig. 3 The geographical distribution and sampling locations of modern and ancient Parsi samples. The population data are obtained from the Parzor Foundation, New Delhi, India (<http://unescomparzor.com/>)

B. Sequence Specific Haplogroup Identification:

Sequence specific haplogroups and variants of Parsi , Indian and the African mitochondrial genomes were retrieved using the MITOMAP Tool [31].

Mitomap tools can be accessed through the following webpage : <https://www.mitomap.org/MITOMAP>

C. Indo-Parsi Phylogeny

Redundant Sequences containing the same haplogroups and same variants were eliminated manually. As a result, 97 of 100 Indian mitochondrial genome sequences and 64 of 105 Parsi mitochondrial genome sequences were selected for further analysis. These sequences were aligned using MUSCLE [31]. alignment with the African mitochondrial genome sequence as an outgroup in the MEGA-X [33] Software. The obtained fasta format of the aligned mitochondrial sequences were employed to construct the maximum likelihood phylogenetic tree using RAxML [34] with GTR - GAMMA +I model with bootstrap value set to 1000 , and other parameters were kept default. The MUSCLE program, source code and PREFAB test data are freely available at <http://www.drive5.com/muscle>. MEGA-X Software : www.megasoftware.net free of charge.

RAxML Program :The code to download this program is available under GNU GPL at <https://github.com/stamatak/standard-RAxML>.

D. Evolutionary Model Analysis:

Partitionfinder 2.0 [35] was used to find the best evolutionary model for the aligned Indian and Zoroastrian- Parsi sequences , parameters such as input data was given as phylip format, Greedy algorithm [36] was used with model selection scheme as "AICC" and PhyML [37] .Among the several models seventeen evolutionary models were listed as a result of the analysis.

Here we present the data of evolutionary models provided for the whole mitochondrial genome sequences. The models are subject to a specific gene or a charset of genes of the genome.

TABLE II BEST EVOLUTIONARY MODELS LIST FOR INDIAN-ZOROASTRIAN -PARSI MITOCHONDRIAL GENOME

SL NO	BEST EVOLUTIONARY MODELS	GENES
1.	TIM+I+G	D'LOOP;
2.	K81UF+I+G	t RNA-Phe ,r RNA (12SrRNA), t RNA-Val,rRNA(16S rRNA), tRNA-Leu; tRNA-Ser,tRNA-Asp
3.	HKY+I	NADH dehydrogenase subunit 2;NADH dehydrogenase subunit 4L (ND4L) NADH dehydrogenase subunit 1(ND1) NADH dehydrogenase subunit 3 (ND3) NADH dehydrogenase subunit 4(ND4)
4.	TRN+G	NADH dehydrogenase subunit 1(ND1) cytochrome c oxidase subunit I (COX I) NADH dehydrogenase subunit 4L (ND4L) NADH dehydrogenase subunit 2 (ND2) cytochrome c oxidase subunit III (COX III)
5.	TRN	cytochrome c oxidase subunit III (COX III) NADH dehydrogenase subunit 1(ND1) cytochrome c oxidase subunit II (COX II)
6.	TRN	tRNA-Trp, tRNA-Ala, tRNA-Asn, tRNA-Cys, tRNA-Tyr ; tRNA-Ile, tRNA-Gln,tRNA-Met
7.	TRN+I	NADH dehydrogenase subunit 2 (ND2) tRNA-Glu NADH dehydrogenase subunit 5(ND5)

		NADH dehydrogenase subunit 6(ND6) NADH dehydrogenase subunit 4(ND4)
8.	TVM	NADH dehydrogenase subunit 5(ND5) cytochrome c oxidase subunit III (COX III) cytochrome c oxidase subunit II (COX II) cytochrome c oxidase subunit I (COX I) cytochrome b (CYTB)
9.	K80	cytochrome c oxidase subunit I (COX I)
10.	TRN	cytochrome c oxidase subunit II (COX II) NADH dehydrogenase subunit 6(ND6)
12.	TIM+G	ATP synthase F0 subunit 8 (ATP8)
13.	TRN+I	tRNA-Arg;tRNA-Gly
14.	TRN+G	NADH dehydrogenase subunit 5(ND5) cytochrome b (CYTB) NADH dehydrogenase subunit 4(ND4) NADH dehydrogenase subunit 3 (ND3)
15.	JC	NADH dehydrogenase subunit 4L (ND4L)
16.	TRN	NADH dehydrogenase subunit 6(ND6)
17.	TRN+G	tRNA-Thr,tRNA-Pro

E. Partition finder 2.0 Software

PartitionFinder is a free open source software that allows the users to select best-fitting partitioning schemes and models of molecular evolution for phylogenetic analyses.

PartitionFinder 2 can be installed by downloading it from the webpage www.robertlanfear.com/partitionfinder. The source code is available at <https://github.com/brettc/partitionfinder> or installing it via GitHub. It requires the installation of Python and few dependencies.

III. RESULT

Our admixture analysis among the Parsi and Indian population with mitochondrial genome sequences has provided some insight regarding the haplogroups specific to these populations and also provide information about the ancestral haplogroup markers.

Venn diagram illustrates that there is an overlap of 23 variants between the Indian and Parsi haplogroups which depicts a probable admixture incident to have taken place and that the African haplogroup doesn't share remarkable variants with both these populations.

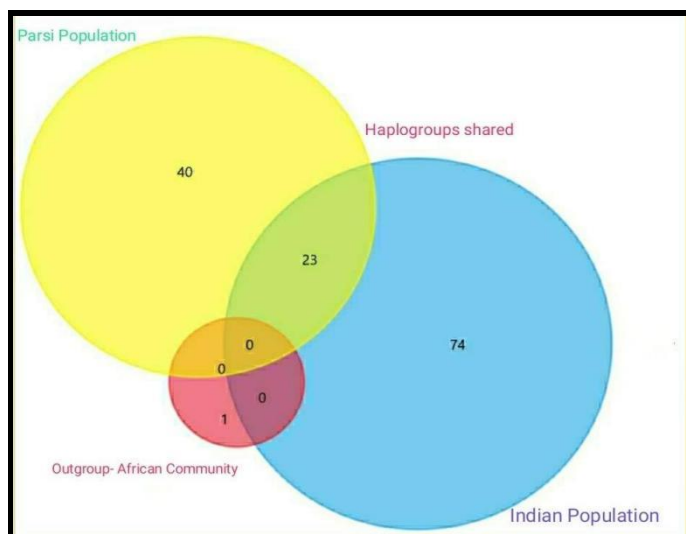


Fig. 4 Venn Diagram representing the number of Haplogroups found in both Parsi and Indian populations built using the Funrich tool [38].

A. Variants Distribution

Haplogroup M2b (M2b1) of the Parsi population was found to have a higher variant number of 55 and haplogroup R30b1 of the Indian population was found to have a higher variant number of 47 in the analysis.

B. Ancestral Haplogroup Markers

Most SNPs are widespread across all lineages. These variants play a role in haplotyping and are referred to as ancestral haplogroup markers.

In our analysis we came across 8 haplogroup markers (Tab 3). Ancestral haplogroups of both Indian and Zoroastrian-Parsi populations provide information about their lineage.

TABLE III REPRESENTATION OF COMMON VARIANTS OF DIFFERENT HAPLOGROUPS PRESENT IN BOTH INDIAN-PARSI POPULATION

Types of variants observed	Percentage of variants in the populations studied
A73G	89%
A263G	99%
A750G	99%
A1438G	99%
A2706G	91%
A4769G	99%
C7028T	95%
G11719A	86%

C. Haplogroup Distribution

A total of 5 major haplogroups were identified from the phylogenetic tree analysis, with 69 sub-haplogroups of the Indo- Zoroastrian-Parsi population. The number of variants of the haplogroups ranged between 10 - 55. Analysis of the sub haplogroups depicted M haplogroup to be most prevalent with 29 distinct sub haplogroups. (Tab 4)

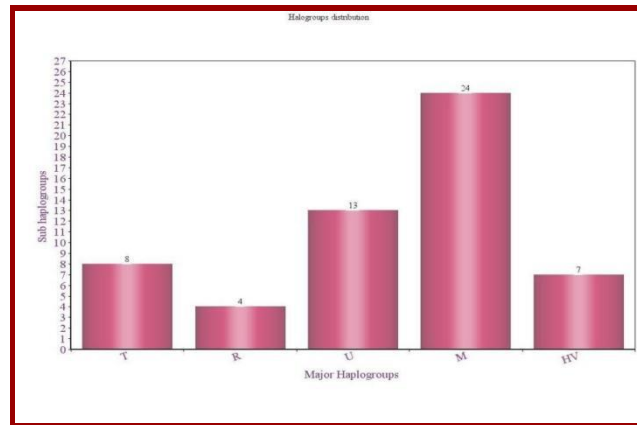


Fig 5: Bar graph showing the distribution of major haplogroups depicting admixture.

D. Phylogenetic Tree

The phylogenetic tree (fig6) obtained showed incidents of admixture of 23 haplogroups among the Parsi and Indian populations representing five distinct major haplogroups, revealed by the presence of the ancestral haplotype of each similar haplogroups.(Tab 3)

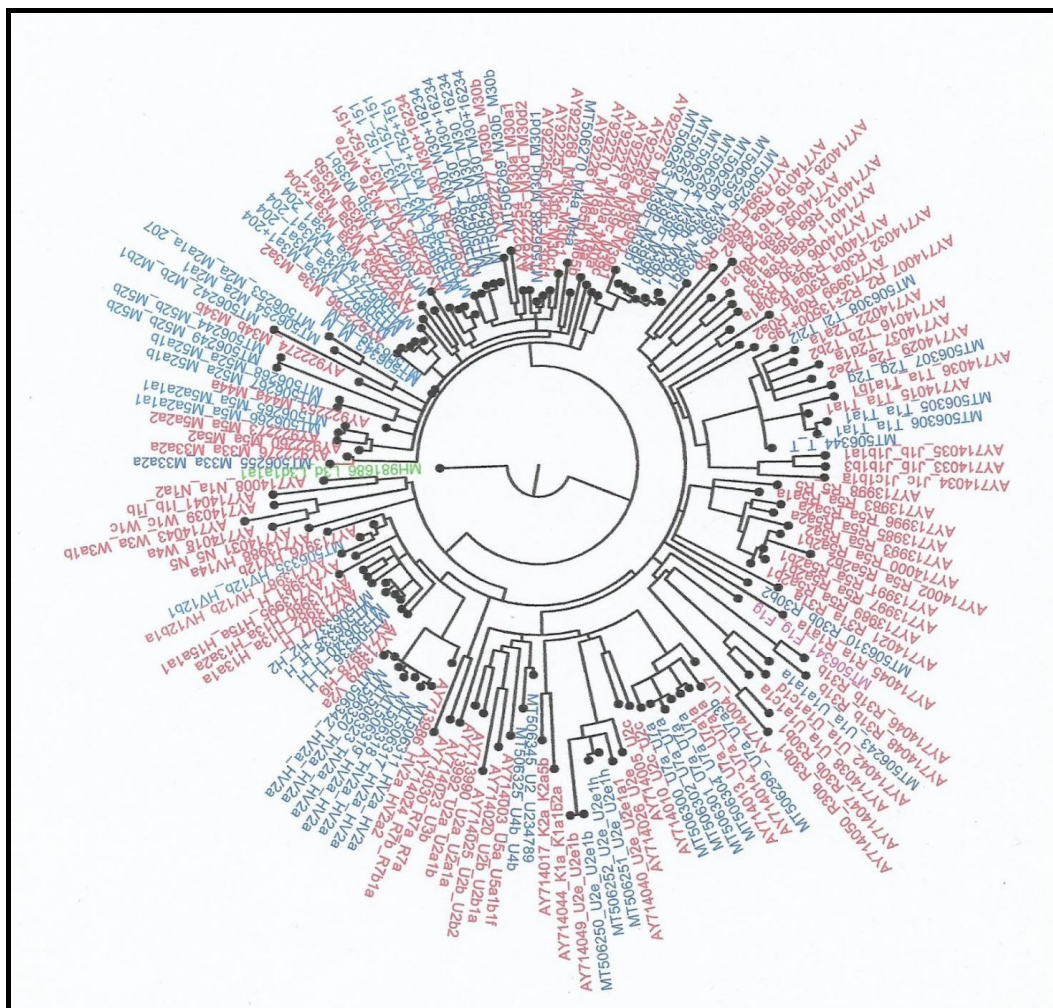


Fig 6 The phylogenetic tree constructed based on the haplogroups of Indian mitochondrial genome (Red) and Zoroastrian-parsi mitochondrial genome sequences (Blue). The clades marked with different colors indicate admixture event.

E. Inferences derived from Indian-Zoroastrian-Parsi phylogeny

TABLE IV PARSI AND INDIAN HAPLOGROUPS THAT CLUSTERED TOGETHER IN THE SAME CLADE AS INFERRED FROM THE PHYLOGENETIC TREE. F1g EURASIAN HAPLOGROUPS WAS FOUND ONLY IN THE ZOROASTRIAN PARSI POPULATION CONTROVERSIAL

Indian and Parsi Haplogroups	
Major haplogroups	Sub haplogroups
T	T1a1 ; T1a1b1 T2a1a ;T2b2;T2d1a; T2e2; T2g; T2i
R	R31a ;R30b2 R30b1; R31b;
U	U1a1a1a ;U1a1c1a;U1a1c1d U2b1a;U2b2;U2a1a;U2a1b;U5a1b1f; U24789; U4b ; U3b U2e; U2e1b U2e1a1; U2e1h U7; U7a ;U7a1a; U7a3b
M	M2a;M2a1a+207;M2b1; M34b;M M3a3a2; M3a3a1+204 M4a ; M65aM5a2a1a; M5a2a2; M5a2 M30+16234 M30a1;M30b; M30c1; M30d1;M30d2; M33a2a M35b, M35a,M35b1 M37+152+151; M37e M39b; M39b1 ; M39c M44a; M52a1b;M52b
HV	HV2a,HV2a2 H;HV12b;HV12b1; HV12b1a ;HV14a H2; H ;H9

M Haplogroup

The M - Haplogroup depicted the higher proportion of admixture among the Indian- Zoroastrian-Parsi population. A maximum of 30 sub-haplogroups were found co-clustered in the phylogenetic tree, which proposes that this haplogroup shares a major role in admixture.

U Haplogroup:

The U Haplogroup comprises 19 sub-haplogroups co-clustered to indicate incidences of admixture with haplogroup U7a being prevalent in both the study populations.

T Haplogroup:

The T Haplogroup comprises 8 sub-haplogroups that provided incidents of admixture.

H Haplogroup:

H haplogroup consisted of 10 sub- Haplogroups that showed incidents of admixture. HV12b and HV12b1 are found to be present in both populations.

R Haplogroup:

R Haplogroup had the least no. of sub-haplogroups i.e. only 4, that depicted admixture to have happened between the Zoroastrian-Parsi and Indian population.

IV. CONCLUSION

Our studies to elucidate the questionnaire of endogamous stability of present day Zoroastrian Parsis among its counter Indian community has been noteworthy. This study has unfolded the reduction of homogeneity of the extant Zoroastrian population. Known for their sacred zoroastrian belief, the Parsis have lost their endogamous practice on the course of civilisation, migration, through intermarriages among Parsis and Non-Parsis. Genetic diversity observed through mitochondrial haplogroup analysis reveals the extent of admixture.

This study has also highlighted the census of the Zoroastrian-Parsi population which has steeply increased thereby characterising admixed population rate increase. Further research is required to predict the period when admixture of these populations could have taken place.

REFERENCES

- [1]. Patgambar, Perochak. "The Parsis of India: Preservation of Identity in Bombay City." (2005): 197-200.
- [2]. Mirza, Hormazdyar Dastur Kayoji. *Outlines of Parsi history*. Mirza, 1974.
- [3]. Boyce, Mary. *Zoroastrianism: its antiquity and constant vigour*. Vol. 7. Mazda Publ., 1992.
- [4]. Karaka, Dosabhai Framji. History of the Parsis: including their manners, customs, religion, and present position. Vol. 2. Macmillan and Company, 1884.
- [5]. Falah, Shubber M. "The Shrine that Consumed Its Town: The Role of Religion and Politics in Reshaping the Iraqi City of Najaf." PhD diss., University of Cincinnati, 2018.
- [6]. Stewart, Sarah, Firoza Punthakey Mistree, Ursula Sims-Williams, Almut Hintze, and Pheroza Godrej. *The Everlasting Flame, Zoroastrianism in History and Imagination*. IB Tauris, 2013.
- [7]. William W Malandra "Zoroastrianism I. Historical Review Up To The Arab Conquest" Encyclopaedia Iranica, online edition 2005
- [8]. Jenny Rose " Zoroastrianism from the early modern period" Discovering Sacred Texts, Theme:Sacred texts, Zoroastrianism; published 23 Sep 2019
- [9]. E Eduljee (2005-17) Qissa I Sanjan - Heritage Institute <http://www.heritageinstitute.com/zoroastrianism/history/qissa1.html>
- [10]. Williams, Alan. The Zoroastrian Myth of Migration from Iran and Settlement in the Indian Diaspora: Text, Translation and Analysis of the 16th Century Qesse-ye Sanjān 'The Story of Sanjan'. Brill, 2009.
- [11]. Shroff, Zubin C., and Marcia C. Castro. "The potential impact of intermarriage on the population decline of the Parsis of Mumbai, India." *Demographic Research* 25 (2011): 545-564.
- [12]. Viloo Morawala-Patell, Naseer Pasha, et al. The first complete Zoroastrian-Parsi mitochondrial reference genome and genetic signatures of an endogamous non-smoking population, *Meta Gene*, Volume 28, 2021, 100882, ISSN 2214-5400.
- [13]. Zhu, Xiaofeng, Amy Luke, Richard S. Cooper, Tom Quertermous, Craig Hanis, Tom Mosley, C. Charles Gu et al. "Admixture mapping for hypertension loci with genome-scan markers." *Nature genetics* 37, no. 2 (2005): 177-181.
- [14]. Basu, Anabha, Namita Mukherjee, Sangita Roy, Sanghamitra Sengupta, Sanat Banerjee, Madan Chakraborty, Badal Dey et al. "Ethnic India: a genomic view, with special reference to peopling and structure." *Genome research* 13, no. 10 (2003): 2277-2290.
- [15]. Sun, Chang, Qing-Peng Kong, Malliya gounder Palanichamy, Suraksha Agrawal, Hans-Jürgen Bandelt, Yong-Gang Yao, Faisal Khan, Chun-Ling Zhu, Tapas Kumar Chaudhuri, and Ya-Ping Zhang. "The dazzling array of basal branches in the mtDNA macrohaplogroup M from India as inferred from complete genomes." *Molecular Biology and Evolution* 23, no. 3 (2006): 683-690.
- [16]. Xu, Shuhua, Wei Huang, Ji Qian, and Li Jin. "Analysis of genomic admixture in Uyghur and its implication in mapping strategy." *The American Journal of Human Genetics* 82, no. 4 (2008): 883-894.
- [17]. Chaubey, Gyaneshwer, Mait Metspalu, Toomas Kivisild, and Richard Villems. "Peopling of South Asia: investigating the caste-tribe continuum in India." *Bioessays* 29, no. 1 (2007): 91-100.
- [18]. Anderson, Sharon, Alan T. Bankier, Bart G. Barrell, Maarten HL de Bruijn, Alan R. Coulson, Jacques Drouin, Ian C. Eperon et al. "Sequence and organization of the human mitochondrial genome." *Nature* 290, no. 5806 (1981): 457-465.
- [19]. Sun, Chang, Qing-Peng Kong, Malliya gounder Palanichamy, Suraksha Agrawal, Hans-Jürgen Bandelt, Yong-Gang Yao, Faisal Khan, Chun-Ling Zhu, Tapas Kumar Chaudhuri, and Ya-Ping Zhang. "The dazzling array of basal branches in the mtDNA macrohaplogroup M from India as inferred from complete genomes." *Molecular Biology and Evolution* 23, no. 3 (2006): 683-690.
- [20]. Hutchison, Clyde A., John E. Newbold, S. Steven Potter, and Marshall Hall Edgell. "Maternal inheritance of mammalian mitochondrial DNA." *Nature* 251, no. 5475 (1974): 536-538.

- [21]. Brown, Wesley M., Matthew George Jr, and Allan C. Wilson. "Rapid evolution of animal mitochondrial DNA." *Proceedings of the National Academy of Sciences* 76, no. 4 (1979): 1967-1971.
- [22]. Michaels, George S., William W. Hauswirth, and Philip J. Laipis. "Mitochondrial DNA copy number in bovine oocytes and somatic cells." *Developmental biology* 94, no. 1 (1982): 246-251.
- [23]. Hagström, Erik, Christoph Freyer, Brendan J. Battersby, James B. Stewart, and Nils-Göran Larsson. "No recombination of mtDNA after heteroplasmy for 50 generations in the mouse maternal germline." *Nucleic acids research* 42, no. 2 (2013): 1111-1116.
- [24]. Lee, William T., Xin Sun, Te-Sha Tsai, Jacqueline L. Johnson, Jodee A. Gould, Daniel J. Garama, Daniel J. Gough, Matthew McKenzie, Ian A. Trounce, and Justin C. St John. "Mitochondrial DNA haplotypes induce differential patterns of DNA methylation that result in differential chromosomal gene expression patterns." *Cell death discovery* 3, no. 1 (2017): 1-11.
- [25]. Wallace, Douglas C., Michael D. Brown, and Marie T. Lott. "Mitochondrial DNA variation in human evolution and disease." *Gene* 238, no. 1 (1999): 211-230.
- [26]. Rishishwar, Lavanya, and I. King Jordan. "Implications of human evolution and admixture for mitochondrial replacement therapy." *BMC genomics* 18, no. 1 (2017): 1-11.
- [27]. Metspalu, M., T. Kivisild, E. Metspalu, J. Parik, G. Hudjashov, K. Kaldma, and P. Serk. "karmin M, Behar DM, Gilbert MTP, Endicott P, Mastana S, Papiha SS, Skorecki K, Torroni A, Villems R: Most of the extant mt DNA boundaries in South and Southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans." *BMC Genetics* 5 (2004): 26.
- [28]. Gounder Palanichamy, Malliya, Chang Sun, Suraksha Agrawal, Hans-Jürgen Bandelt, Qing-Peng Kong, Faisal Khan, Cheng-Ye Wang, Tapas Kumar Chaudhuri, Venkatramana Palla, and Ya-Ping Zhang. "Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia." *The American Journal of Human Genetics* 75, no. 6 (2004): 966-978.
- [29]. Sun, Chang, Qing-Peng Kong, Malliya gounder Palanichamy, Suraksha Agrawal, Hans-Jürgen Bandelt, Yong-Gang Yao, Faisal Khan, Chun-Ling Zhu, Tapas Kumar Chaudhuri, and Ya-Ping Zhang. "The dazzling array of basal branches in the mtDNA macrohaplogroup M from India as inferred from complete genomes." *Molecular Biology and Evolution* 23, no. 3 (2006): 683-690.
- [30]. McCrow, John P., Desiree C. Petersen, Melanie Louw, Eva KF Chan, Katherine Harmeyer, Stefano Vecchiarelli, Ruth J. Lyons, MS Riana Bornman, and Vanessa M. Hayes. "Spectrum of mitochondrial genomic variation and associated clinical presentation of prostate cancer in South African men." *The Prostate* 76, no. 4 (2016): 349-358.
- [31]. Lott, Marie T., Jeremy N. Leipzig, Olga Derbeneva, H. Michael Xie, Dimitra Chalkia, Mahdi Sarmady, Vincent Procaccio, and Douglas C. Wallace. "mtDNA variation and analysis using mitomap and mitomaster." *Current protocols in bioinformatics* 44, no. 1 (2013): 1-23.
- [32]. Edgar, Robert C. "MUSCLE: multiple sequence alignment with high accuracy and high throughput." *Nucleic acids research* 32, no. 5 (2004): 1792-1797.
- [33]. Kumar, Sudhir, Glen Stecher, Michael Li, Christina Knyaz, and Koichiro Tamura. "MEGA X: molecular evolutionary genetics analysis across computing platforms." *Molecular biology and evolution* 35, no. 6 (2018): 1547.
- [34]. Stamatakis, Alexandros. "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies." *Bioinformatics* 30, no. 9 (2014): 1312-1313.
- [35]. Lanfear, Robert, Paul B. Frandsen, April M. Wright, Tereza Senfeld, and Brett Calcott. "PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses." *Molecular biology and evolution* 34, no. 3 (2017): 772-773.
- [36]. Lanfear, Robert, Brett Calcott, Simon YW Ho, and Stéphane Guindon. "PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses." *Molecular biology and evolution* 29, no. 6 (2012): 1695-1701.
- [37]. Guindon, Stéphane, Jean-François Dufayard, Vincent Lefort, Maria Anisimova, Wim Hordijk, and Olivier Gascuel. "New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0." *Systematic biology* 59, no. 3 (2010): 307-321.
- Pathan, Mohashin, Shivakumar Keerthikumar, Ching-Seng Ang, Lahiru Gangoda, Camelia YJ Quek, Nicholas A. Williamson, Dmitri Mouradov et al. "FunRich: An open access standalone functional enrichment and interaction network analysis tool." *Proteomics* 15, no. 15 (2015): 2597-2601.