Full Length Research Article
Advancements in Life Sciences – International Quarterly Journal of Biological Sciences

ARTICLE INFO

Date Received: 10/04/2017;
Date Revised: 17/05/2017;
Date Published Online: 25/05/2017;

Authors' Affiliation:
1- Department of Genetics, Hazara University, Garden Campus, Mansehra - Pakistan
2- Department of Anthropology, University of Alaska, Fairbanks, Fairbanks, AK - United States
3- Department of Biochemistry, Faculty of Science, King Abdulaziz University Jeddah 21589 - Saudi Arabia
4- Islamia College University, Peshawar, Khyber Pakhtunkhwa – Pakistan

***Corresponding Author:**
Inam Ullah
Email:
inamgenetics@gmail.com

How to Cite:
Ullah I, Ahmad H, Hemphill BE, Nadeem MS, Tariq M, Tabassum S (2017). Mitochondrial genetic characterization of Gujar population living in the Northwest areas of Pakistan. Adv. Life Sci. 4(3): 84-91.

Keywords:
Pakistan, Swat, Gujar, mtDNA control region, Haplotyping

Open Access

Mitochondrial genetic characterization of Gujar population living in the Northwest areas of Pakistan

Inam Ullah*[1], Habib Ahmad[1,4], Brian E. Hemphill[2], Muhammad Shahid Nadeem[3], Muhammad Tariq[4], Sadia Tabassum[1]

Abstract

Background: Diversity of communities with specific cultural, ethnic, lingual and geographical backgrounds makes Pakistani society a suitable study subject to unravel the early human migrations, evolutionary history of population having about 18 ethnic groups. Gujars are mostly Indic-speaking nomadic herders with the claims of multiple origins in the sub-continent. Present study was aimed at the determination of maternal lineage of Gujars by mitochondrial DNA analysis.

Methods: Total DNA from the human buccal cells was isolated using modified phenol chloroform method. Purified DNA was used for the PCR amplification of mitochondrial Hyper Variable Region 1 and 2 (HVR1 & 2). The nucleotide sequences of amplified PCR products were used to explore the maternal lineage of the Gujar population residing in Northern Pakistan.

Results: Haplotypes, allele frequencies and population data of the mitochondrial control region was determined in 73 unrelated individuals belonging to Gujar ethnic group of Northwest areas of Pakistan. Total 46 diverse haplotypes were identified out of which 29 were found unique with (0.9223) genetic diversity and (0.9097) power of discrimination. Haplogroup R was the most frequent (48%) followed by haplogroup M (45%) and N (7%).

Conclusion: We found that the Gujar population has multiple maternal gene pool comprising of South Asian, West Eurasian, East Eurasian, Southeast Asian and fractions of Eastern Asian, Eastern Europe and Northern Asian lineages. This study will contribute for the development of mitochondrial DNA database for Pakistani population.

Scan QR code to view this publication on your mobile device

Introduction

Pakistan is located in the western part of the Indian subcontinent, with Afghanistan and Iran to the west, India to the east, the Arabian Sea to the south and covers an area of approximately 796,095 sq. km (figure 1). About 46,8000 sq. km of this area is in west and north comprises mountains lands and plateau, while the remaining 328,000 km^2 is in the form of plains [1]. Pakistan has a diverse communities distributed into variety of ethnic groups, having variety of cultures, languages and geographical backgrounds, which make this land suitable for unraveling early human migrations, population study and evolutionary history having 18 ethnic groups further divided into casts and sub-casts [2,3].

Figure 1: Map of Khyber Pakhtunkhwa, Pakistan and its neighbor countries. The "shaded area" shows study area.

Gujars are Indic-speaking nomadic herders whose origins are claimed to be in Rajasthan and adjacent regions of Gujarat in India and the Indus Valley of Pakistan [4]. Following irrigation efforts in the Indus Valley by the British administration, Gujars were forced northwards in the late-19th century into the foothills rimming the northern margin of the Indus Valley and beyond into Khyber Pakhtunkhwa, Jammu and Kashmir. Some historians says that Gujars probably first appeared in the area about 400 years ago [5, 6]. Gujars are considered as 'Aryas' and their arrival to this part of the world is traced back to 242 and 300 BCs. Gujars invaded India in third century B.C. and they are actually inhabitants of Gujarustan which is still called as Gujarustan or Gorgia [7]. First time the word Gujar was used by a pioneer Ramchand with his name [8].

Various studies have proved that human DNA is a direction to explore historical movements of populations by studying their genetic make-up. Mitochondrial DNA is a proper tool for the human migration, geographic distribution and population origin due to its high evolutionary importance [9, 10].

To investigate all possible lineages among various ethnic groups, we obtained data for the Hyper Variable Region 1&2 (HVR1&2) of mtDNA from 73 Gujar individuals from the Swat district of Khyber Pakhtunkhwa Pakistan. mtDNA haplogroups affiliations have been diagnosed by using different computer software and servers and finally we compared the mtDNA distribution among the various subpopulations, including regional ethnic groups from Pakistan and neighboring countries.

Methods

Saliva samples were collected in sterile collecting cups from 73 unrelated Gujar volunteers belongs to different areas of district Swat of Northwest Pakistan (figure 1). All participants gave their informed consent verbally or in writing after explaining the aims and procedures of the study to them. The consent form was designed according to the ethical review board of Hazara University. Genomic DNA from the human buccal cells was obtained using DNA isolation method [11]. The isolated genomic DNA was used for the PCR amplification of HVR1 & 2 of mtDNA with two sets of reverse and forward primers (table 1). The PCR reaction mixture included 2.0µL of 10pM/µL F-Primer, 2.0µL of 10pM/µL R-Primer, 0.5µL of Taq DNA Polymerase enzyme (5U/µL) "Fermentas", and 2.0µL of DNA template with a final volume of 25.0µL. Thermal cycling was conducted using an Applied Bio system 2720 (95°C for 4 min; 35 cycles of 94°C for 40 s, 56°C for 1 minute, and 72°C for 1 minutes; and a final extension at 72°C for 5 min). The gel containing PCR products were purified using the procedure adopted from GeneAll Gel Elution Kit (SV) Cat. no. 102-101. Sequencer machine (ABI Prism 3730XL) was used for sequencing the purified products.

Oligo Name	Sequence (5'-3')
HVR-1 (Forward)	CTCCACCATTAGCACCCAAAGCTAAG
HVR-1 (Reverse)	GATATTGATTTCACGGAGGATGGTGGTC
HVR-2 (Forward)	AGGTCTATCACCCTATTAACCACTCACG
HVR-2 (Reverse)	GGTGTCTTTGGGGTTTGGTTGGTTC

Table 1: The sequences of primers used in the present study.

Data analysis

Haplotypes for the corresponding HVR1 and HVR2 sequences were then identified with the help of online software, MitoTool [12], HaploGrep [13] and Mitomaster [14] using PhyloTree Build 16 (http://www.phylotree.org) as classification tree to assess the quality of mtDNA data [10]. The sequences of Gujar mitochondrial DNA were assign to haplogroup according to phylotree [10] and published data [15-18]. The population statistics *i.e.* Genetic Diversity (GD), Power of Discrimination (PD) and Random Match Probability (RMP) were also calculated using computational tools [19, 20].

Results

A total of 73 samples were analyzed for the mitochondrial DNA control region of Gujar population belongs to District Swat of Khyber Pakhtunkhwa (KP) Province of Pakistan. Haplogroup frequencies were calculated for the characterization of mtDNA variation in the individuals of the present study population. Forty six different haplotypes were observed during the present study among which 29 were unique while 17 haplotypes were shared by more than one individual, while the corresponding mtDNA genetic diversity was (0.9223), power of discrimination (0.9097) and random match probability (0.0903) table 2. The observed haplogroup frequencies, their respective variants and geographic position are given in table 3.

Population statistics	
Total number of samples	73
No of haplotypes	46
No of unique haplotypes	29
Random match probability	0.0903
Power of discrimination	0.9097
Genetic diversity	0.9223

Table 2: Statistical analysis of Gujar population residing in Northwest areas of Pakistan

By comparing the genetic parameters of the reported population living in Pakistan with the current studied Gujar population, we found that the Gujars of Swat have a moderate unique haplotypes (29) consistent with the other population of Pakistan (table 4). The moderate frequency of unique haplotypes reflected in high genetic diversity (0.922) in the Gujar ethnic group of the present study as compared to the other reported ethnic groups from Pakistan except Kalash with (0.851) genetic diversity (table 4). However, the highest number of unique haplotypes (128) has also been reported in Pakhtuns of Pakistan due to large number of sample size (n= 230) table 4.

The obtained sequences of mtDNA control region (1-574, 15974-16425) of the present Gujar population were compared with revised Cambridge Reference Sequence (rCRS) [21]. The results of sequences revealed that at nucleotide position 16023np 95% (G/A), at 16061np 91% (C/A), at 16163np 95% (G/A) , at 32np 92% (A/G), at 38np 98.5% (G/A) and at 278np 100% (A/G) had transition mutations while transversion mutations were scored at 16036np 99% (G/C), 16172np 100% (G/T), 16219np 97% (A/G), 33np 95% (C/G), 44np 93% (C/A) respectively.

In the present study we observed South Asian haplogroups (42%), West Eurasian (37%), East Eurasian (11%), Southeast Asian (4%), Eastern Asian (2.7%), Eastern Europe (1.4%) and Northern Asian (1.4%). Among south Asian haplogroups, haplogroup M6 occurred (7%), M30 (4%), M37 (4%), M5c (4%), M3 (2.7%), M3a (2.7%), M5 (2.7%), M52a (2.7%), R5a (2.7%), M30d (1.4%), M3c (1.4%), M53 (1.4%), M54 (1.4%), M7c (1.4%) and R22 (1.4%). West Eurasian haplogroups includes H2a (4%), T2b (4%), H14a (2.7%), H5 (2.7%), K1a (2.7%), U7a (2.7%), H1 (1.4%), H1a (1.4%), H1e (1.4%), H3p (1.4%), N (1.4%), T (1.4%), T1a (1.4%), U2a (1.4%), U4a (1.4%), U5b (1.4%), U7 (1.4%), V9a (1.4%) and W3a (1.4%). East Eurasian haplogroups includes B4a (5%), D4b (1.4%), D4e (1.4%), D4g (1.4%) and D4p (1.4%). Southeast Asian haplogroups includes F1 (1.4%), G2b (1.4%) and S (1.4%). Eastern Asian haplogroups includes A (2.7%); Eastern Europe H7i (1.4%) and Northern Asian include haplogroup J (1.4%) respectively. The frequencies of each haplogroups are given in (figure 2).

The haplotypes of Gujar population were assigned to mega haplogroups which revealed that the most frequent among them was R with the frequency of (48%) followed by haplogroup M (45%) and N (7%) (figure 3).

Sr. No	Frequency	Variants	Hg	HGO
GS1	2	A73G, T152C, A234G, A235G, A263G, C309CCT, T310C, AC523d, C560A, T16105C, C16115A, C16223T, C16290T, T16311C, G16319A, T16362C, T16413A	A	EA
GS2	4	A73G, T195C, A263G, C309CCT, T310C, AC523d, C560A, A16100T, T16189C	B4a	EEA
GS3	1	C16115A, C16223T, G16274A, A16307T, C16332T, C16355T, T16362C, A16367G, G16384T, A16387G	D4b	EEA
GS4	1	T152C, T155A, A165T, A178T, C16083A, T16090A, A16100T, C16223T, G16274A, T16362C	D4e	EEA
GS5	1	C151T, T152C, A263G, A290T, C298T, C308T, C315CC, G323T, C324T, C332T, T334TT, C340T, C349T, C356T	D4g	EEA
GS6	1	A73G, T195C, C198T, A263G, C309CCT, T310C, T482C, T489C, AC523d, T16097C, G16110A, G16414A	D4p	EEA
GS7	1	A16183C, T16189C, A16194C, T16195A, C16197G, C16201A, A16203AA, C16205A, T16209A, C16211A, C16214A, A16230T, C16234A, C16236A, T16243G, C16245G, A16258C, A16265C, A16269C, T16271A, T16276A, C16282A, A16293C, C16301A, T16304C, C16306A, T16308A, T16311A, C16313A, A16322T, C16332A, C16339A, A16340T, T16347C, C16358T, T16359C, A16367T, T16368G, T16372A	F1	SEA
GS8	1	G62GG, A73G, G184A, A200G, A263G, T310C, T310TTC, G380T, G389T, A396T, G410T, A425T, T430C, C445T, C465T, T16094C, T16117A, T16189C, C16192CT, A16194G, T16195G, A16212T, A16220C, C16223T, C16239G, C16245G, A16258C, A16265T, A16269G, T16276A, A16277C, A16285C, A16293T, C16294T, C16296T, A16305T, A16316G, A16326C, T16330G, A16333T, T16334A, G16346A, T16347C, A16351T, T16362C, A16367G, T16368G	G2b	SEA
GS9	1	G53GC, A263G, T310C, T310TTC, T16154C, G16156C, C16159T, A16166C, T16189C, A16402C, T16413C	H1	WE
GS10	2	G71GG, T72G, A263G, C315CC, A16180C, C16256T, T16352C, G16414A	H14a	WE
GS11	1	G92A, A111d, G124T, A126T, T131A, G184A, G185T, G187A, A200G, G203T, C231T, A241T, A248T	H1a	WE
GS12	1	T89TT, C150T, A263G, C264T, A300G, T310A, G316C, C317CC, G329T, C330T, C332T, A339T, A351T, A357T, A360T	H1e	WE
GS13	3	A263G, C315CC, T16075A, C16223T, C16234T, G16274A	H2a	WE
GS14	1	G53GC, A263G, C315CC, T16154C, G16156C, A16166C, C16168A, C16169A, C16174A, C16222T, C16242T, G16273A, T16356C	H3p	WE
GS15	2	A263G, C315CC, G366A, G389A, T408TT, A419C, A426T, A428T, C436A, C438T, A439T, A443T, C445A, A446T, C456T, C462T	H5	WE
GS16	1	G124T, T125A, T133A, A178T, A215d, G228A	H7i	EEU
GS17	1	G184A, A191T, A200G, C222T, A240T, A263G, C295T	J	NA
GS18	2	T63A, A73G, C150T, T199C, A263G, C315CC, G366T, C371T, G380A, T391A, A395T, A415T, A419T, T424TT, C438T, C441T, A443T, A451T, T452A, T453G, C459T, C462T, C467A, C476T, A478T, G16129A, T16224C, C16301T, A16312C, C16321T, C16328T	K1a	WE
GS19	2	A73G, T195C, A263G, T310C, T310TTC, G366A, T414G, T482C, T489C, AC523d, A561C	M3	SA
GS20	3	A73G, T125C, T127C, T195A, A263G, C309CCT, T310C, T489C, AC523d, C560A, T16075A, A16078T, C16223T, C16234T, G16274A, G16414A	M30	SA
GS21	1	A73G, T195A, A263G, C315CC, T489C, AC523d, C16179d, C16223T, A16302G	M30d	SA
GS22	1	T199C, A263G, A278T, A281T, A291T, C311T, G16096C, T16097C, C16223T, T16304C, T16362C	M35b	SA
GS23	3	A73G, C151T, T152C, A263G, C309CCT, T310C, T489C, T16075A, C16085A, C16221T, C16223T	M37	SA
GS24	2	C194T, T195C, T204C, G260T, A263G, C271T, A272T, C273A, C315CC, A331T, C332A, C349T, A16074G, T16126C, C16192T, C16223T, A16312G	M3a	SA
GS25	1	A73G, T195C, A263G, C309CCT, T310C, T482C, T489C, AC523d, T16126C, T16154C, C16223T, T16224C	M3c	SA
GS26	2	G53GT, A73G, T195C, A263G, C309CCCT, T310C, T489C, C560A, G16129A, C16223T	M5	SA
GS27	2	A73G, C78CA, G79C, T195C, A237T, A263T, C268T, C269T, A281T, A287T, C16223T, C16266T, A16275G, C16327A, G16390A	M52a	SA
GS28	1	T16154C, A16164C, A16165C, T16189C, C16192T, C16223T, C16294T, A16316G, T16362C, G16384A, T16386A	M53	SA
GS29	1	A73G, A263G, C315CC, T489C, C560A, A16070C, G16129A, C16223T, T16304C, T16325C, G16414A	M54	SA
GS30	3	A73G, C150T, A263G, C315CC, T489C, C560A, G16110A, C16111A, C16115A, G16118A, T16126C, G16129A, T16209C, C16223T, T16311C	M5c	SA

Sr. No	Frequency	Variants	Hg	HGO
GS31	5	G54GG, A73G, T152C, A214G, A263G, C315CC, C461T, T489C, AC523d, T16140A, T16152C, T16154C, A16155C, A16164C, A16165C, C16174G, C16223T, G16274A, T16323A, A16351T, T16362C, C16376T, G16384T, A16387G, C16404T	M6	SA
GS32	1	T16068C, A16070C, A16074G, A16078T, G16110A, T16117A, T16140d, T16144A, C16147T, A16158d, T16161A, A16171T, T16172C, A16182T	M7c	SA
GS33	1	C16223T	N	WE
GS34	1	A73G, A188G, C194T, T204C, G207A, A263G, G316C, C317CC, G329A, A339T, T344C, C353G, A358T	R22	SA
GS35	2	G85T, G94A, G107T, A111T, G124A, T146C, T152C, A210T, G229T, A263G, G275T, A278T, A281T, C299T, A301T, C307T, C312CT, C312T, T16094C, G16096C, T16097C, C16099T, C16266T, T16304C, T16311C, T16356C, C16393T, C16404T	R5a	SA
GS36	1	T152C, A263G, C315CC, T455TT, A492T, A515T, CAC516d, C558T, A16066T, C16069T, A16070T, A16074G, C16176T, C16185T, C16223T, A16246T, A16309T, G16346T, C16348T, A16402T	S	SEA
GS37	1	A87G, A263G, C315CC, T16126C, T16143G, C16151G, C16188T, T16189C, A16194G, A16207G, A16216T, C16234T, T16263C, A16277T, C16279T, A16284T, A16289T, C16294T, G16303T, C16321T, C16327T, C16337T, T16342A, A16343T, C16353T, A16367G, C16382T, T16386G, A16387G, C16393T, C16395G	T	WE
GS38	1	A73G, T152C, T195C, A263G, C309CCT, T310C, C16174A, C16186T, T16189C, C16294T	T1a	WE
GS39	3	A73G, A263G, C285T, T310C, G316C, C317CCCC, T321C, C324G, C332A, C343T, C362A, G366A, T372C, A379C, G380A, T383C, T391A, C394T, C404T, C411G, G429C, T430C, A432C, A448T, T460C, A464C, T471C, C473A, T474A, T482C, T489A, A492C, A523C, C527G	T2b	WE
GS40	1	A73G, A183G, A188G, C194T, G207A, A263G, C315CC, G545A, G16110A, C16115A, G16129A, A16206C, T16362C	U2a	WE
GS41	1	A73G, T99TT, G124T, T199C, A263G, A270T, C296T, A300T	U4a	WE
GS42	1	A73G, C150T, A263G, C315CC, C560A, T16093C, T16094C, T16097C, G16110A, C16111A, C16115A, T16131G, C16270T, G16412C	U5b	WE
GS43	1	C16114A, C16115A, A16309G, A16318T, A16416T	U7	WE
GS44	2	A73G, G94T, G97T, G103T, G121T, C151T, T152C, A183T, G187A, A189T, T208A, T233A, A243T, A249T, G260T, A263G, G275A, T16121C, T16126C, T16131G, T16263C, A16269G, T16288C, T16304A, A16309G, T16311A, A16318T, T16359C, T16362C, T16372C, T16396C	U7a	WE
GS45	1	T119C, A189G, T195C, T204C, G207A, A263G, C315CC, C516T, C530T, T16093C, T16094C, T16097C, T16105C, G16213A, G16274A, G16319A, T16362C, G16390A	V9a	WE
GS46	1	A73G, G143A, A189G, C194T, T195C, T199C, T204C, G207A, A263G, C315CC, C16223T, C16292T, G16414A	W3a	WE

Hg, haplogroup; Hgo, Haplogroup origin; Eastern Asian, EA; South East Asian, SEA; West Eurasian,WE; Eastern Europe, EEU; Northern Asian, NA; South Asian, SA; East Eurasian, EEA.

Table 3: Haplogroups frequencies and their respective variants of Gujar population of Swat

Parameters	**Gujars present study**	**Mak [26]**	**Sk [23]**	**Pt [25]**	**Bl [24]**	**Br [24]**	**Hz [24]**	**Hb [24]**	**Ks [24]**	**Ps [24]**	**Sd [24]**	**Pk [24]**
No of samples	73	100	85	230	39	38	23	44	44	44	23	100
No of haplotypes	46	70	63	157	26	22	21	32	12	22	21	77
No of unique haplotypes	29	54	58	128	18	15	19	25	5	12	19	63
Genetic diversity	0.922	0.97	0.96	0.99	0.97	0.95	0.99	0.98	0.851	0.95	0.992	0.992

Mak, Makrani; Sk, Saraiki; Pt, Pakhtuns; Bl, Baluch; Br, Brahui; Hz, Hazara; Hb, Hunza burusho; Ks, Kalash; Ps, Pasrsi; Sd, Sindhi; Pk, Pakistan Karachi

Table 4: Diversity comparison of Gujar population with the other reported ethnic groups of Pakistan

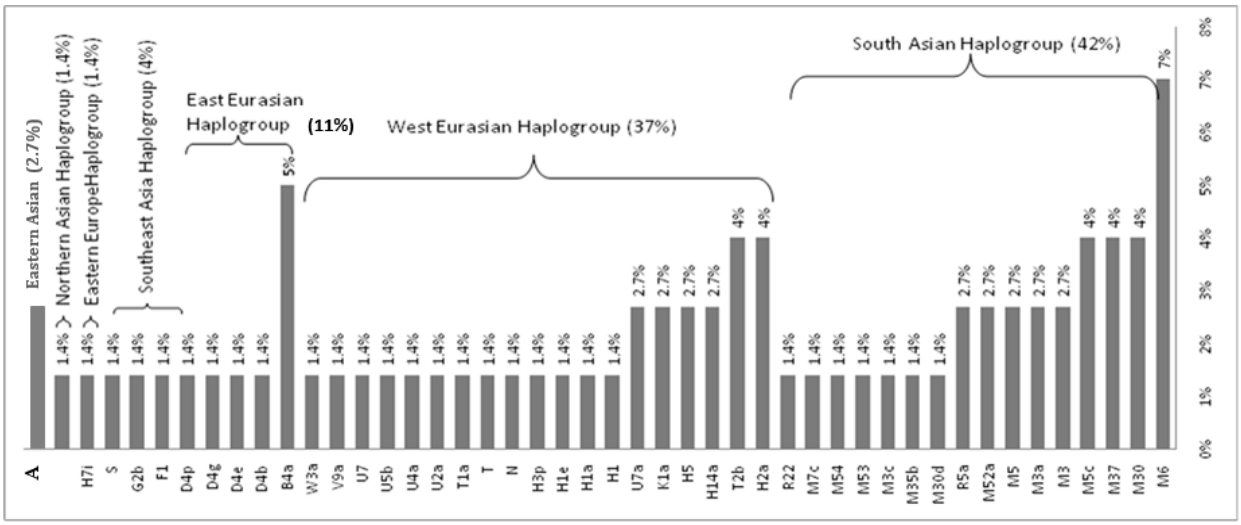

Figure 2: Graphical representation of the mtDNA haplogroup composition of the Gujar residing Northwest areas of Pakistan

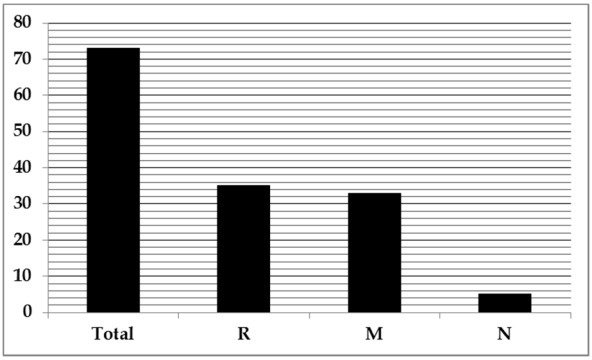

Figure 3: Mega haplogroup frequencies among the individuals of Gujars population. Distribution of HCV Positive Diabetic Patients

Discussion

In the present study 73 unrelated samples from the Gujars were characterized for maternal linage and other genetic structure. The genetic structure of the present studied population was compared with the previously reported data of Pakistani ethnic groups. The haplotypic diversity of the Gujar population (GD=0.9223) observed shows a high genetic diversity in comparison with the other reported population of Pakistan except Kalash [22-25]. Genetic diversity is due the reflection of unique haplotypes distribution. The numbers of unique haplotypes identified in the present studied population were 63%, which were found somehow consistent with Burusho 78%, Hazara 76%, Makrani 76%, Baluchi 69% and Brahui 68% among the other reported population of Pakistan, while moderately lower from Saraiki 92%, Sindhi 90% and Pathan 81% [22-25]. Members of Gujars population revealed high frequency (42%) of South Asian lineage. The proportion of South Asian lineages in the other reported Pakistani populations were 48% in Sindhi, 39.1% in Pathan, 36% Pashtun, 29.4% in Saraiki and 24% in Makrani [22, 24-27]. Low frequency of South Asian lineages among the major ethnic groups of Afghanistan have also been reported with the prevalence of 15% in Hazara, 13.3% in Baluch and 7.1% in Pashtun, while absent in Tajik [28]. The presence of south Asian mtDNA haplogroups in the present study population revealed that the population residing in this region are the true inhabitants and are remolded in the past by local demographic events [17]. The West Eurasian haplogroup was the second most prevalent haplogroup accounting for (37%) in the individuals of the present study population. Its frequency among the Pathans of Pakistan was reported 55%and 26% in Makranis [24, 25]. Furthermore, the frequency of West Eurasian haplogroup in Indian Punjabis population were reported from (40-50%), in Kashmiris and Gujrathis 30%, while the least were observed in Indian Uttar Pradesh and West Bengal [17,29]. Greater proportion of West Eurasian lineages were also reported among the major ethnic groups of Afghanistan with the frequencies of 40% in Hazara, 89% in Tajik, 74% in Baluch and 64% in Pashtun [28] . The presence of these lineages revealed that, the gene flow in the past to this region may occur from the west through Iran or from the North through Central Asia [23], through the invasion by different invaders *i.e.* Alexander, Arabians, Muslims and the British [30]. The mega haplogroup R, M and N identified in the Gujars population are said to be South Asian in origin and has been originated approximately

60000-75000 years ago in South Asia [31], suggesting their maternal gene pool as South Asian in origin.

Acknowledgments

The authors would like thank the Ethnogenetic Project (No. 20-1409) titled "Ethnogenetic elaboration of KP through dental morphology and DNA analysis" at Hazara University, Mansehra, Pakistan for assisting in sample collection. The research was funded by the Indigenous 5000 Ph.D. Fellowship Program of the Higher Education Commission of Pakistan.

References

1. Ahmad K, Hussain M, Ashraf M, Luqman M, Ashraf MY, *et al.* Indigenous vegetation of Soone Valley; At the risk of extinction. Pak J Bot, (2007); 39(3): 679-690.
2. Ayub Q, Tyler-Smith C. Genetic variation in South Asia: assessing the influences of geography, language and ethnicity for understanding history and disease risk. Briefings in functional genomics & proteomics, (2009); 8(5): 395-404.
3. Grimes BF. Ethnologue: languages of the world: 1992. Dallas, Texas: Summer Institute of Linguistics. Inc
4. Grierson GA. 1903–1928. Linguistic Survey of India, Vols I-XI.: 1968. Calcutta[Reprint 1968, Delhi: Motilal Benarsidass]
5. Barth F. Ecologic relationships of ethnic groups in Swat, North Pakistan. American Anthropologist, (1956); 58(6): 1079-1089.
6. Rome S. Forestry in the princely state of Swat and Kalam (North-West Pakistan). 2005: pp. 1-125.
7. Ali I. Mapping and documentation of the cultural assets of Kaghan Valley, Mansehra. United Nations Educational, Scientific and Cultural Organization, Islamabad, (2005); 5-6.
8. Chauhan RAH. A short history of the Gurjars: past and present/by Rana Ali Hasan Chauhan. (2001).
9. Nesheva D. Aspects of ancient mitochondrial DNA analysis in different populations for understanding human evolution. science, (2014); 1(5): 5-14.
10. Van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Human mutation, (2009); 30(2): E386-E394.
11. Akbar N, Ahmad H, Nadeem MS, Ali N, Saadiq M. An Efficient Procedure for DNA Isolation and Profiling of the Hyper Variable MtDNA Sequences. Journal of Life Sciences, (2015); 9530-534.
12. Fan L, Yao Y-G. MitoTool: a web server for the analysis and retrieval of human mitochondrial DNA sequence variations. Mitochondrion, (2011); 11(2): 351-356.
13. Kloss-Brandstätter A, Pacher D, Schönherr S, Weissensteiner H, Binna R, *et al.* HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. Human mutation, (2011); 32(1): 25-32.
14. Brandon MC, Ruiz-Pesini E, Mishmar D, Procaccio V, Lott MT, *et al.* MITOMASTER: a bioinformatics tool for the analysis of mitochondrial DNA sequences. Human mutation, (2009); 30(1): 1-6.
15. Behar DM, Villems R, Soodyall H, Blue-Smith J, Pereira L, *et al.* The dawn of human matrilineal diversity. The American Journal of Human Genetics, (2008); 82(5): 1130-1140.
16. Elmadawy MA, Nagai A, Gomaa GM, Hegazy HM, Shaaban FE, *et al.* Investigation of mtDNA control region sequences in an Egyptian population sample. Legal Medicine, (2013); 15(6): 338-341.
17. Metspalu M, Kivisild T, Metspalu E, Parik J, Hudjashov G, *et al.* Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. BMC genetics, (2004); 5(1): 26.
18. Oven M, Vermeulen M, Kayser M. Multiplex genotyping system for efficient inference of matrilineal genetic ancestry with continental resolution. Investigative Genetics, (2011); 2(6): 1-14.
19. Prieto L, Zimmermann B, Goios A, Rodriguez-Monge A, Paneto G, *et al.* The GHEP–EMPOP collaboration on mtDNA population data—A new resource for forensic casework. Forensic Science International: Genetics, (2011); 5(2): 146-151.
20. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, (1989); 123(3): 585-595.
21. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nature genetics, (1999); 23(2): 147-147.
22. Hayat S, Akhtar T, Siddiqi MH, Rakha A, Haider N, *et al.* Mitochondrial DNA control region sequences study in Saraiki population from Pakistan. Legal Medicine, (2015); 17(2): 140-144.
23. Quintana-Murci L, Chaix R, Wells RS, Behar DM, Sayar H, *et al.* Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. The American Journal of Human Genetics, (2004); 74(5): 827-845.
24. Rakha A, Shin K-J, Yoon JA, Kim NY, Siddique MH, *et al.* Forensic and genetic characterization of mtDNA from Pathans of Pakistan. International journal of legal medicine, (2011); 125(6): 841-848.
25. Siddiqi MH, Akhtar T, Rakha A, Abbas G, Ali A, *et al.* Genetic characterization of the Makrani people of Pakistan from mitochondrial DNA control-region data. Legal Medicine, (2015); 17(2): 134-139.
26. Bhatti S, Aslamkhan M, Abbas S, Attimonelli M, Aydin HH, *et al.* Genetic analysis of mitochondrial DNA control region variations in four tribes of Khyber Pakhtunkhwa, Pakistan. Mitochondrial DNA Part A, (2016); 1-11.
27. Bhatti S, Aslamkhan M, Attimonelli M, Abbas S, Aydin HH. Mitochondrial DNA variation in the Sindh population of Pakistan. Australian Journal of Forensic Sciences, (2017); 49(2): 201-216.
28. Whale J. Mitochondrial DNA analysis of four ethnic groups of Afghanistan. (2012). University of Portsmouth.
29. Ahmed M. Ancient Pakistan-an archaeological history. 2014 Amazon.
30. McElreavey K, Quintana-Murci L. A population genetics perspective of the Indus Valley through uniparentally-inherited markers. Annals of human biology, (2005); 32(2): 154-162.

31. Kivisild T, Rootsi S, Metspalu M, Mastana S, Kaldma K, *et al*. The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. The American Journal of Human Genetics, (2003); 72(2): 313-332.