Forensic Science International: Genetics 30 (2017) e1–e5

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Short communication

mtDNA sequence diversity of Hazara ethnic group from Pakistan

Allah Rakha[a,b,*], Fatima[b], Min-Sheng Peng[c], Atif Adan[b], Rui Bi[a], Memona Yasmin[b], Yong-Gang Yao[a]

[a] Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan 650223, China
[b] Department of Forensic Sciences, University of Health Sciences, Lahore, Pakistan
[c] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China

ARTICLE INFO

Article history:
Received 12 February 2017
Received in revised form 2 July 2017
Accepted 8 July 2017
Available online 10 July 2017

Keywords:
Pakistan
Hazara
mtDNA
Control region
Haplogroup
Forensic database

ABSTRACT

The present study was undertaken to investigate mitochondrial DNA (mtDNA) control region sequences of Hazaras from Pakistan, so as to generate mtDNA reference database for forensic casework in Pakistan and to analyze phylogenetic relationship of this particular ethnic group with geographically proximal populations. Complete mtDNA control region (nt 16024-576) sequences were generated through Sanger Sequencing for 319 Hazara individuals from Quetta, Baluchistan. The population sample set showed a total of 189 distinct haplotypes, belonging mainly to West Eurasian (51.72%), East & Southeast Asian (29.78%) and South Asian (18.50%) haplogroups. Compared with other populations from Pakistan, the Hazara population had a relatively high haplotype diversity (0.9945) and a lower random match probability (0.0085). The dataset has been incorporated into EMPOP database under accession number EMP00680. The data herein comprises the largest, and likely most thoroughly examined, control region mtDNA dataset from Hazaras of Pakistan.

1. Introduction

Hazaras are among few ethnic groups whose origin is obscure. Ancestral home of Hazaras is known as Hazarajat (also referred as Hazaristan), which lies in the central highlands of Afghanistan, among the Koh-i-Baba mountains and the western extremities of the Hindu Kush [1]. Hazaras speak “Hazaragi” (or Aazragi) language [2] and are overwhelmingly Shia Muslims constituting the third largest ethnic group in Afghanistan [3].

According to historical records, several Turk and Mongol tribes constituted the Hazaras ethnically, hence, these are considered of Turko-Mongol origin. The Hazaras identity is believed to be originated after 1221 CE siege of Bamiyan by the Mongol Empire, led by Genghis Khan in pursuit of Jalal ad-Din Mingburnu, the last ruler of the Khwarezmian Empire, and his forces in Afghanistan. It is commonly believed that the soldiers of army left behind by Genghis Khan to guard important places at Hindu Kush, married to the local Tajik women and gave rise to the ancestors of the Hazaras [2]. Few Mongol words in vocabulary of the language of Hazaras supports their Mongol ancestry. Moreover, there are also lines of evidence that some of the remote tribes of Hazaras spoke Mongol language till last century. Their central Asian facial features including sparse beards, high cheekbones and epicanthic eye folds further supports their Mongol origin [2].

Hazaras have faced history of horrendous persecution in Afghanistan. It dates back to 16th century, with Babur from Kabulistan. It is reported that during reign of Emir Abdur Rahman (1880–1901), thousands of Hazaras were killed, expelled and enslaved [4]. Syed Askar Mousavi, a contemporary Hazara writer, claims that half the population of Hazaras was displaced, shifted to neighboring Baluchistan province of Pakistan [5] and Khorasan Province of Iran [4].

Population data on Hazaras from both countries, Afghanistan and Pakistan, have only been restricted to HVS-I [6,7]. Forensic casework mtDNA analysis requires relevant and reliable databases for estimating the probability of a given haplotype in a population [8]. This population study constitutes high quality entire mitochondrial DNA (mtDNA) control region sequence data of 319 randomly selected individuals from population of Hazara, resident of Baluchistan province of Pakistan.

---

* Corresponding author at: Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan 650223, China.
E-mail addresses: dnaexpert@me.com, dnatypist@gmail.com (A. Rakha).

http://dx.doi.org/10.1016/j.fsigen.2017.07.004

2. Materials and methods

Blood samples were collected from 319 Hazara individuals (males n = 185); (females n = 134) residing in Hazara town, an area on the western outskirts of Quetta, city of Baluchistan Province of Pakistan (Fig. 1). All participants were unrelated individuals within at least three generations.

Written informed consent was obtained from all volunteer donors. Samples were fully anonymized. Ethical approval was obtained from Institutional Review Board of University of Health Sciences, Lahore, Pakistan.

2.1. DNA extraction, amplification and sequencing

Genomic DNA was extracted with Axygen® AxyPrep™ Blood Genomic DNA Miniprep Kit according to the manufacturer's protocol (Axygen Biosciences; CA, USA). Same set of primers and PCR conditions were used to amplify the entire control region (nt 16024-576) as in our previous study [9]. Purification of PCR products was done with Exonuclease I (TaKaRa) and Shrimp Alkaline Phosphatase (TaKaRa) in 10× Exonuclease I Buffer (TaKaRa) following reaction compositions recommended by the manufacturer. The reaction mixture was incubated at 37 °C for 45 min, followed by enzyme deactivation at 95 °C for 15 min. Sequencing primers from [9] were used to perform capillary electrophoresis using the BigDye™ Terminator v3.0 Ready Reaction Cycle Sequencing Kit on Applied Biosystems 3730×l DNA Analyzer (Thermo Fisher Scientific) according to manufacturer's instructions.

2.2. Haplogroup assignment

Forward and reverse sequences were aligned and compared using by SeqManNGen® version 12.0 (DNASTAR. Madison, WI). Sequence variations in each sequence were scored relative to the revised Cambridge Reference Sequence (rCRS) [10]. Quality of sequences was examined manually, and two analysts independently annotated deviations from the reference sequence. The recommended nomenclature for mtDNA typing was used for alignment of variants [11]. The haplogroup assignments were carried out using Mitotool (www.mitotool.org) [12,13], Haplogrep (www.haplogrep.uibk.ac.at) [14], and EMMA (www.empop.online) [15] based on PhyloTree build 17 [16]. Haplogroup assignments were re-evaluated by manual inference and conservative most recent common ancestor (MRCA) status was assigned to each sequence to improve the prediction. All 319 sequences generated here were incorporated to EMPOP database under accession number EMP00680 [17]. The sequences are also available on GenBank via accession numbers KX448813-KX449131.

2.3. Data analysis

Indels at nucleotide positions 16193, 309, 315 and 573 were not taken into account for analysis unless otherwise mentioned. The number of haplotypes were calculated by direct counting method. Discrimination capacity was estimated by dividing the number of different haplotypes by total number of individuals [18]. Haplotype diversity was estimated by DnaSP 5.10.1 [19] and random match probability was calculated as sum of squares of observed haplotype frequencies [18].

Based on the distribution patterns [20] and the latest version of the PhyloTree [16], haplogroups were assigned into three ancestry groups: West Eurasian (H, HV, I, J, K, N, R, R0, T, U1a, U5a, U7, U8a, W, X), South Asian (M*, U2, U4), and East Asian (A, B, C, D, F, G) [7,16,21,22]. Molecular diversity indices, pairwise differences between and within populations, haplotype frequencies, analysis of molecular variance (AMOVA) and pairwise F_{ST} values were calculated using DnaSP 5.10.1 [19] and Arlequin 3.5 (Computational and Molecular Population Genetics Lab, Institute of Ecology, and Evolution, University of Berne, Bern, Switzerland) [23]. F_{ST} calculations were made under the Kimura 2-parameter model [24]. In addition, the data was compared with previously published data of other ethnic groups from Pakistan (Kashmiri [9], Pathan [25], Makrani [26]) and worldwide populations (Uzbek, Kazakh, Turkmen, Kyrgyz, Afghan, Russian, Tajik [27], Chinese [28], Thai [29], Vietnamese [30], Laos [31], Kuwaiti [32], Iraqi [33], Egyptian [34], Dubai [35], Ashkenazi Jews [36], Roma [37], Sze ’kely [37], Greeks [38] Moroccan [39] and Iranian [40]) for which entire

Fig. 1. The red dotted area refers to geographic location of sampling from Quetta of Pakistan. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

control region (nt 16024-576) data could be obtained from the literature. Additional comparative analysis was performed with populations for those available sequences consisted of at least nt 16024–16569 range i.e, Afghanistan (Hazara, Pashtun, Tajik, Turkmen, and Uzbek [7]), China (Kazakh [41] and Mongol [42]), Iran (Iranian [40]), Uzbekistan (Uzbek, Kazakh, Turkmen, Kyrgyz, Afghan, Russian, Tajik [27]), Pakistan (Kashmiri [9], Pathan [25], Makrani [26], Baluch (EU565791-EU565815) and Brahui (EU565766-EU565790)). In order to have uniformity and non-arbitrary comparative analysis for the published data, consistent treatment of selecting base on the basis of phylogenetic weightage was applied for the sequence heteroplasmic positions. Pairwise F_{ST} matrix was imported into R statistical package (http://www.R-project.org/) for plotting heatmap (ComplexHeatmap) and principal components analysis (FactoMineR). This study followed the requirements of this journal [43,44] and the ISFG recommendations for forensic population data publication [43].

3. Results and discussion

To establish reference database for mtDNA applications in forensic, here we present a high quality mtDNA control region sequences dataset of 319 Hazara individuals sampled from Baluchistan Province of Pakistan. mtDNA haplotypes of 319 individuals are given in Table S1. Data showed higher haplotype diversity (0.9945) for Hazaras as compared to Makrani (0.9905) [26] but lower than Kashmiri (0.9977) and Pathan populations [25]. Further, data showed lower random match probability (0.0085) among Hazaras as compared to Makrani population, while a higher value than Kashmiri and Pathan populations (Table 1). Point heteroplasmies were identified in two individuals at two different positions (16230R and 16256Y).

3.1. Haplogroup information

A total of 189 distinct haplotypes were observed in all 319 individuals (Table S2). The most common haplotype among all was shared by 9 individuals (2.82%). Out of 189 haplotypes, 124 were unique while 65 were shared by 2–9 individuals. All these haplotypes were assigned to 156 different haplogroups/sub-haplogroups based on PhyloTree Build 17 [16]. The haplogroups were mainly assigned into three continental groups, namely West Eurasian (51.72%), East & Southeast Asian (29.78%) and South Asian (18.50%) [7,16,21,22]. The most frequent haplogroups were H (14.11%), K (14.73%), D (9.72%) and HV (9.40%). Hazaras had higher percentage of East and South East Asian haplogroups among all other Pakistani populations (Fig. S1). Phylogeographical cross-comparison of mtDNA and Y-chromosomal data from Hazara had already showed that male descendants of Mongols were accompanied by women of East Asian ancestry while settling through Central Asia [6]. Our results indicate that the Hazara are characterized by relatively high frequencies of West Eurasian mtDNAs which are virtually fewer or otherwise absent in bordering West Asian populations.

3.2. Pairwise F_{ST} comparisons and population structure

Comparison between different Pakistani populations (Kashmiri [9], Pathan [25] and Makrani [26]) on the basis of pairwise F_{ST} values using complete mtDNA control region sequences showed the highest distance between Hazara and Makrani followed by Hazara and Pathan populations (Table 2). Analysis of Molecular Variance (AMOVA) showed that 97.95% of the variance was within populations and 2.05% of the variance represented among Pakistani populations (Table S3).

Hazara control region (nt 16024-576) sequence data was compared with regional populations (Uzbek, Kazakh, Turkmen, Kyrgyz, Afghan, Russian, Tajik [27], Chinese[28], Thai [29], Vietnamese [30], Laos [31], Kuwaiti [32], Iraqi [33], Egyptian [34], Dubai [35], Ashkenazi Jews [36], Roma [37] Sze 'kely [37], Greeks [38] Moroccan [39] and Iranian [40]). AMOVA was used to test for significant variation in the mtDNA distribution among regional and global populations. 95.83% of the variance observed within these populations was attributable to differences within populations, and 4.17% of the variance represent differences among populations (Table S4). Heatmap matrix (Fig. S2) and corresponding pairwise F_{ST} values (Table S5) showed that Hazaras, along with other populations of northern areas of Pakistan (i.e. Kashmiri and Pathan), were clustered more closely to Central Asian Turkmen, Uzbek, Kazakh and Tajik populations. Similar results were obtained through Principal Component Analysis (PCA) (Fig. S3). These results supported the fact about Hazaras having their roots in Afghanistan that shares a long common history with Central Asian [45]. More than 2000 km of joint borders unite Afghanistan with Turkmenistan, Tajikistan and Uzbekistan and their multifaceted links might have influenced the mtDNA diversity of Hazaras [45]. It can also be noticed that Afghan and Russian from Uzbekistan [27] do not cluster with rest of the surrounding Central Asian populations including Hazara (Fig. S3). One potential reason could be relatively smaller sample size studied for these two populations.

In order to include available sequences (nt 16024–16569) from Afghanistan (Hazara, Pashtun, Tajik, Turkmen, and Uzbek [7]), an additional comparative analysis was performed by trimming the sequence data from China (Kazakh [41] and Mongol [42]), Iran (Iranian [40]), Uzbekistan (Uzbek, Kazakh, Turkmen, Kyrgyz,

Table 2
Pairwise F_{ST} distances of mtDNA control region sequences among different ethnic groups of Pakistan.

	Kashmiri [9]	Hazara	Makrani [26]	Pathan [25]
Kashmiri	*	0.012	0.031	0.014
Hazara	**0.000**	*	0.035	0.024
Makrani	**0.000**	**0.000**	*	0.032
Pathan	**0.000**	**0.000**	**0.000**	*

Pairwise F_{ST} values (above diagonal) and associated P-values (below diagonal).

Table 1
Forensic parameters and diversity indices of mtDNA control region sequences from Pakistani ethnic groups.

	Hazara	Kashmiri [9]	Pathan [25]	Makrani [26]
No. of Samples	319	317	230	99
No. of Haplotypes	189	251	192	71
No. of Unique haplotypes	124	201	128	54
No. of Polymorphic sites including indels	217	230	215	142
Discrimination Capacity	0.5925	0.7918	0.8348	0.7172
Random Match Probability	0.0085	0.0054	0.0066	0.0195
Haplotype Diversity (Hd)	0.9945	0.9977	0.9978	0.9905
Average Number of Pairwise differences	10.431	10.447	11.144	12.110

Afghan, Russian, Tajik [27]) and Pakistan (Pathan [25], Kashmiri [9], Makrani [26], Baluch (EU565791-EU565815) and Brahui (EU565766-EU565790)). PCA illustrated that Hazaras of Pakistan are more closely related to the Hazaras of Afghanistan while showed comparatively distal clustering with Mongols but in same quadrate (Fig. S4).

Comparative analysis of shared haplotypes of different populations was performed. Although number of shared haplotypes strongly depends on sample size, however, Hazaras shared more haplotypes with Tajiks of Afghanistan as compare to Tajiks from Uzbekistan (Table S6). These findings did not correspond to the historical claim of partial Tajik origin of Hazaras [2].

4. Conclusion

This is the first report of a reasonably sized forensic database from Hazaras including entire mtDNA control region sequences. The Hazaras are genetically found to be heterogeneous and show a higher similarity to the Central Asian populations as compared to Western Asian populations.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

We would like to thank all volunteers for providing their samples used in this study. We acknowledge Muhammad Raza Wakil for his support in sampling, Muhammad Jawad, Saadia Noreen, Li-Li Kong, Lu-Xiu Yang and Ammar Sabir Cheema for technical assistance, and Hansi Weissensteiner (HAPLOGREP, Medical University Innsbruck, Austria) for haplogroup assignments. This project was supported by the President's International Fellowship Initiative (PIFI) Program of the Chinese Academy of Sciences and the National Natural Science Foundation of China (30925021). M.-S.P. thanks the Youth Innovation Promotion Association, Chinese Academy of Sciences, for its support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2017.07.004.

References

[1] Wikipedia, Hazarajat. <https://en.wikipedia.org/wiki/Hazarajat>, 2007 (accessed September 2016.2016).

[2] B. Campbell, Disappearing people? Indigenous groups and ethnic minorities in South and Central Asia, in: Brower Brower, Barbara Rose Johnston (Eds.), International Mountain Society, California, 2007.

[3] L. Dupree, AFGHANISTAN iv. Ethnography, Encyclopaedia Iranica, 2011, pp. 495–501.

[4] A. Monsutti, HAZĀRA ii. HISTORY. <http://www.iranicaonline.org/articles/hazara-2>, 2003 (accessed 7 December 2016.2016).

[5] I. Yusuf, Who Are the Hazara, the Express Tribune, 2011.

[6] L. Quintana-Murci, R. Chaix, R.S. Wells, D.M. Behar, H. Sayar, R. Scozzari, C. Rengo, N. Al-Zahery, O. Semino, A.S. Santachiara-Benerecetti, A. Coppa, Q. Ayub, A. Mohyuddin, C. Tyler-Smith, S. Qasim Mehdi, A. Torroni, K. McElreavey, Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor, Am. J. Hum. Genet. 74 (5) (2004) 827–845.

[7] J. Di Cristofaro, E. Pennarun, S. Mazieres, N.M. Myres, A.A. Lin, S.A. Temori, M. Metspalu, E. Metspalu, M. Witzel, R.J. King, P.A. Underhill, R. Villems, J. Chiaroni, Afghan Hindu Kush: where Eurasian sub-continent gene flows converge, PLoS One 8 (10) (2013) e76748.

[8] W. Parson, A. Brandstatter, A. Alonso, N. Brandt, B. Brinkmann, A. Carracedo, D. Corach, O. Froment, I. Furac, T. Grzybowski, K. Hedberg, C. Keyser-Tracqui, T. Kupiec, S. Lutz-Bonengel, B. Mevag, R. Ploski, H. Schmitter, P. Schneider, D. Syndercombe-Court, E. Sorensen, H. Thew, G. Tully, R. Scheithauer, The EDNAP mitochondrial DNA population database (EMPOP) collaborative exercises: organisation, results and perspectives, Forensic Sci. Int. 139 (2–3) (2004) 215–226.

[9] A. Rakha, M.S. Peng, R. Bi, J.J. Song, Z. Salahudin, A. Adan, M. Israr, Y.G. Yao, EMPOP-quality mtDNA control region sequences from kashmiri of azad jammu & kashmir Pakistan, forensic science international, Genetics 25 (2016) 125–131.

[10] R.M. Andrews, I. Kubacka, P.F. Chinnery, R.N. Lightowlers, D.M. Turnbull, N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA, Nat. Genet. 23 (2) (1999) 147.

[11] W. Parson, L. Gusmao, D.R. Hares, J.A. Irwin, W.R. Mayr, N. Morling, E. Pokorak, M. Prinz, A. Salas, P.M. Schneider, T.J. Parsons, DNA Commission of the International Society for Forensic Genetics: revised and extended guidelines for mitochondrial DNA typing, Forensic science international, Genetics 13 (2014) 134–142.

[12] L. Fan, Y.G. Yao, MitoTool: a web server for the analysis and retrieval of human mitochondrial DNA sequence variations, Mitochondrion 11 (2) (2011) 351–356.

[13] L. Fan, Y.G. Yao, An update to MitoTool: using a new scoring system for faster mtDNA haplogroup determination, Mitochondrion 13 (4) (2013) 360–363.

[14] H. Weissensteiner, D. Pacher, A. Kloss-Brandstatter, L. Forer, G. Specht, H.J. Bandelt, F. Kronenberg, A. Salas, S. Schonherr, HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing, Nucleic Acids Res. 44 (W1) (2016) W58–63.

[15] A.W. Rock, A. Dur, M. van Oven, W. Parson, Concept for estimating mitochondrial DNA haplogroups using a maximum likelihood approach (EMMA), Forensic science international, Genetics 7 (6) (2013) 601–609.

[16] M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation, Hum. Mutat. 30 (2) (2009) E386–94.

[17] W. Parson, A. Dur, EMPOP–a forensic mtDNA database, Forensic science international, Genetics 1 (2) (2007) 88–92.

[18] F. Chen, C.Y. Yin, X.Q. Qian, H.T. Fan, Y.J. Deng, Y.D. Zhang, H.T. Meng, C.M. Shen, C.H. Yang, R. Jin, B.F. Zhu, P. Xu, Single nucleotide polymorphisms of mitochondrial DNA HVS-I and HVS-II in Chinese Bai ethnic group, Electrophoresis 36 (6) (2015) 930–936.

[19] J. Rozas, DNA sequence polymorphism analysis using DnaSP, in: D. Posada (Ed.), Bioinformatics for DNA Sequence Analysis, Humana Press, Totowa, NJ, 2009, pp. 337–350.

[20] P.A. Underhill, T. Kivisild, Use of y chromosome and mitochondrial DNA population structure in tracing human migrations, Annu. Rev. Genet. 41 (2007) 539–564.

[21] M. Metspalu, T. Kivisild, E. Metspalu, J. Parik, G. Hudjashov, K. Kaldma, P. Serk, M. Karmin, D.M. Behar, M.T. Gilbert, P. Endicott, S. Mastana, S.S. Papiha, K. Skorecki, A. Torroni, R. Villems, 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 Genet. 5 (2004) 26.

[22] Y.G. Yao, Q.P. Kong, H.J. Bandelt, T. Kivisild, Y.P. Zhang, Phylogeographic differentiation of mitochondrial DNA in han chinese, Am. J. Hum. Genet. 70 (3) (2002) 635–651.

[23] L. Excoffier, H.E. Lischer, Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, Mol. Ecol. Resour. 10 (3) (2010) 564–567.

[24] M. Kimura, A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences, J. Mol. Evol. 16 (2) (1980) 111–120.

[25] A. Rakha, K.J. Shin, J.A. Yoon, N.Y. Kim, M.H. Siddique, I.S. Yang, W.I. Yang, H.Y. Lee, Forensic and genetic characterization of mtDNA from Pathans of Pakistan, Int. J. Legal Med. 125 (6) (2011) 841–848.

[26] M.H. Siddiqi, T. Akhtar, A. Rakha, G. Abbas, A. Ali, N. Haider, A. Ali, S. Hayat, S. Masooma, J. Ahmad, M.A. Tariq, M. van Oven, F.M. Khan, Genetic characterization of the Makrani people of Pakistan from mitochondrial DNA control-region data, Legal Med. (Tokyo Japan) 17 (2) (2015) 134–139.

[27] J.A. Irwin, A. Ikramov, J. Saunier, M. Bodner, S. Amory, A. Rock, J. O'Callaghan, A. Nuritdinov, S. Atakhodjaev, R. Mukhamedov, W. Parson, T.J. Parsons, The mtDNA composition of Uzbekistan: a microcosm of Central Asian patterns, Int. J. Legal Med. 124 (3) (2010) 195–204.

[28] J.A. Irwin, J.L. Saunier, P. Beh, K.M. Strouss, C.D. Paintner, T.J. Parsons, Mitochondrial DNA control region variation in a population sample from Hong Kong, China, Forensic science international, Genetics 3 (4) (2009) e119–125.

[29] B. Zimmermann, M. Bodner, S. Amory, L. Fendt, A. Rock, D. Horst, B. Horst, T. Sanguansermsri, W. Parson, A. Brandstatter, Forensic and phylogeographic characterization of mtDNA lineages from northern Thailand (Chiang Mai), Int. J. Legal Med. 123 (6) (2009) 495–501.

[30] J.A. Irwin, J.L. Saunier, K.M. Strouss, T.M. Diegoli, K.A. Sturk, J.E. O'Callaghan, C. D. Paintner, C. Hohoff, B. Brinkmann, T.J. Parsons, Mitochondrial control region sequences from a Vietnamese population sample, Int. J. Legal Med. 122 (3) (2008) 257–259.

[31] M. Bodner, B. Zimmermann, A. Rock, A. Kloss-Brandstatter, D. Horst, B. Horst, S. Sengchanh, T. Sanguansermsri, J. Horst, T. Kramer, P.M. Schneider, W. Parson, Southeast Asian diversity: first insights into the complex mtDNA structure of Laos, BMC Evol. Biol. 11 (2011) 49.

[32] M. Scheible, M. Alenizi, K. Sturk-Andreaggi, M.D. Coble, S. Ismael, J.A. Irwin, Mitochondrial DNA control region variation in a Kuwaiti population sample: forensic science international, Genetics 5 (4) (2011) e112–3.

[33] N. Al-Zahery, J. Saunier, K. Ellingson, W. Parson, T.J. Parsons, J.A. Irwin, Characterization of mitochondrial DNA control region lineages in Iraq, Int. J. Legal Med. 127 (2) (2013) 373–375.

[34] J.L. Saunier, J.A. Irwin, K.M. Strouss, H. Ragab, K.A. Sturk, T.J. Parsons, Mitochondrial control region sequences from an Egyptian population sample: forensic science international, Genetics 3 (3) (2009) e97–103.

[35] F. Alshamali, A. Brandstatter, B. Zimmermann, W. Parson, Mitochondrial DNA control region variation in dubai, United Arab Emirates, F orensic science international, Genetics 2 (1) (2008) e9–10.

[36] A. Brandstatter, B. Egyed, B. Zimmermann, A. Tordai, Z. Padar, W. Parson, Mitochondrial DNA control region variation in Ashkenazi Jews from Hungary: forensic science international, Genetics 2 (1) (2008) e4–6.

[37] B. Egyed, A. Brandstatter, J.A. Irwin, Z. Padar, T.J. Parsons, W. Parson, Mitochondrial control region sequence variations in the Hungarian population: analysis of population samples from Hungary and from Transylvania (Romania), Forensic science international, Genetics 1 (2) (2007) 158–162.

[38] J. Irwin, J. Saunier, K. Strouss, C. Paintner, T. Diegoli, K. Sturk, L. Kovatsi, A. Brandstatter, M.A. Cariolou, W. Parson, T.J. Parsons, Mitochondrial control region sequences from northern Greece and Greek Cypriots, Int. J. Legal Med. 122 (1) (2008) 87–89.

[39] R. Aboukhalid, K. Sturk-Andreaggi, M. Bouabdellah, D. Squalli, J.A. Irwin, S. Amzazi, Mitochondrial DNA control region variation from samples of the Moroccan population, Int. J. Legal Med. 127 (4) (2013) 757–759.

[40] M. Derenko, B. Malyarchuk, A. Bahmanimehr, G. Denisova, M. Perkova, S. Farjadian, L. Yepiskoposyan, Complete mitochondrial DNA diversity in iranians, PLoS One 8 (11) (2013) e80673.

[41] W. Shan, Z. Ren, W. Wu, H. Hao, A. Abulimiti, K. Chen, F. Zhang, Z. Ma, X. Zheng, Maternal and paternal diversity in Xinjiang Kazakh population from China, Genetika 50 (11) (2014) 1374–1385.

[42] B. Cheng, W. Tang, L. He, Y. Dong, J. Lu, Y. Lei, H. Yu, J. Zhang, C. Xiao, Genetic imprint of the Mongol: signal from phylogeographic analysis of mitochondrial DNA, J. Hum. Genet. 53 (10) (2008) 905–913.

[43] A. Carracedo, J.M. Butler, L. Gusmao, A. Linacre, W. Parson, L. Roewer, P.M. Schneider, New guidelines for the publication of genetic population data: forensic science international, Genetics 7 (2) (2013) 217–220.

[44] A. Carracedo, J.M. Butler, L. Gusmao, A. Linacre, W. Parson, L. Roewer, P.M. Schneider, Update of the guidelines for the publication of genetic population data: forensic science international, Genetics 10 (2014) A1–2.

[45] A. Vera, L. Marlène, P. Sébastien, The Afghanistan Central. Asia relationship: What role for the EU?, 2013. <http://fride.org/publication/1099/the-afghanistan-central-asia-relationship:-what-role-for-the-eu>(accessed April 2017).

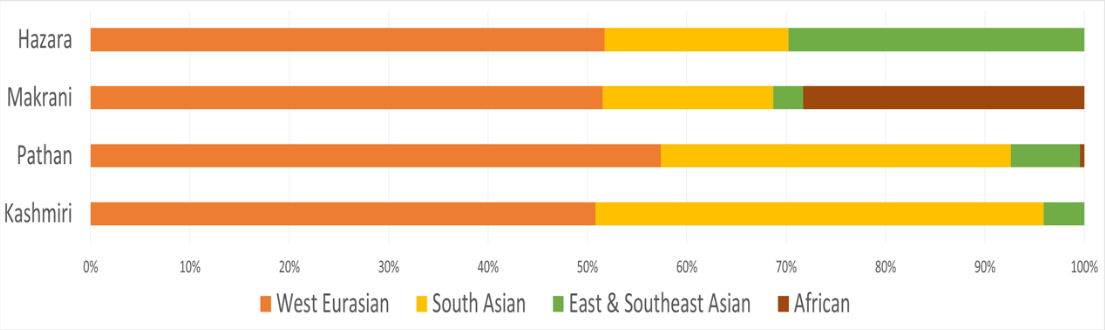

Figure S1. The matrilineal composition of four Pakistani populations based on complete control region sequences (Hazara (319), Kashmiri (317) [9], Pathan (230) [25] and Makrani (99) [26])

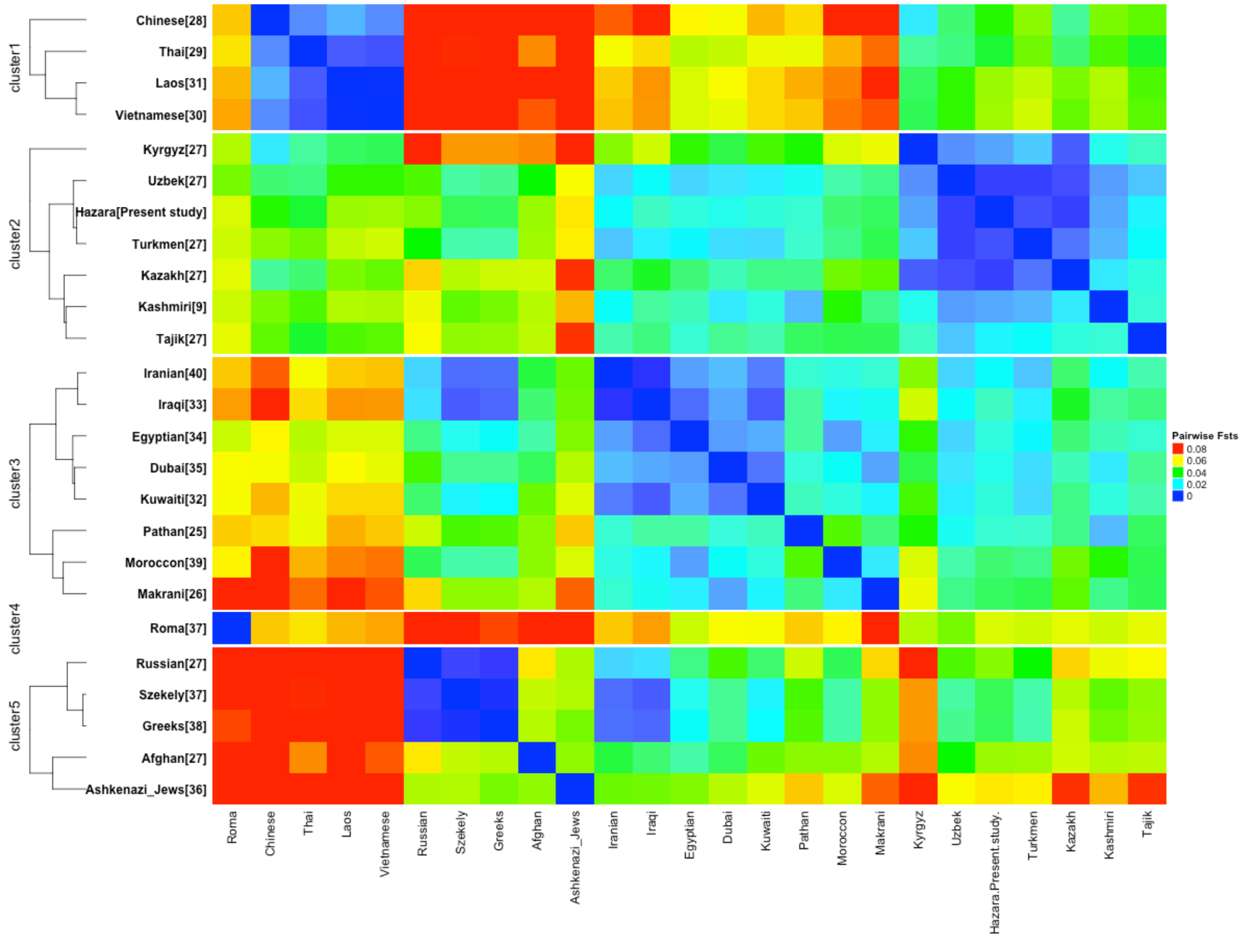

Figure S2. Heatmap matrix of pairwise F_{ST} values for 25 populations including Hazara

Data Sources (mtDNA Complete Control Region):

Populations	References	No. of samples	Populations	References	No. of samples
Hazara	Present study	319	Chinese	[28]	377
Kashmiri	[9]	317	Thai	[29]	190
Pathan	[25]	230	Vietnamese	[30]	187
Makrani	[26]	99	Laos	[31]	214
Uzbek	[27]	328	Kuwaiti	[32]	381
Turkmen	[27]	249	Iraqi	[33]	182
Tajik	[27]	244	Egyptian	[34]	277
Russian	[27]	151	Dubai	[35]	249
Kyrgyz	[27]	249	Ashkenazi Jews	[36]	173
Kazakh	[27]	256	Roma	[37]	205
Afghan	[27]	98	Greeks	[38]	319
Szekely	[37]	182	Iranian	[40]	352
Moroccan	[39]	509			

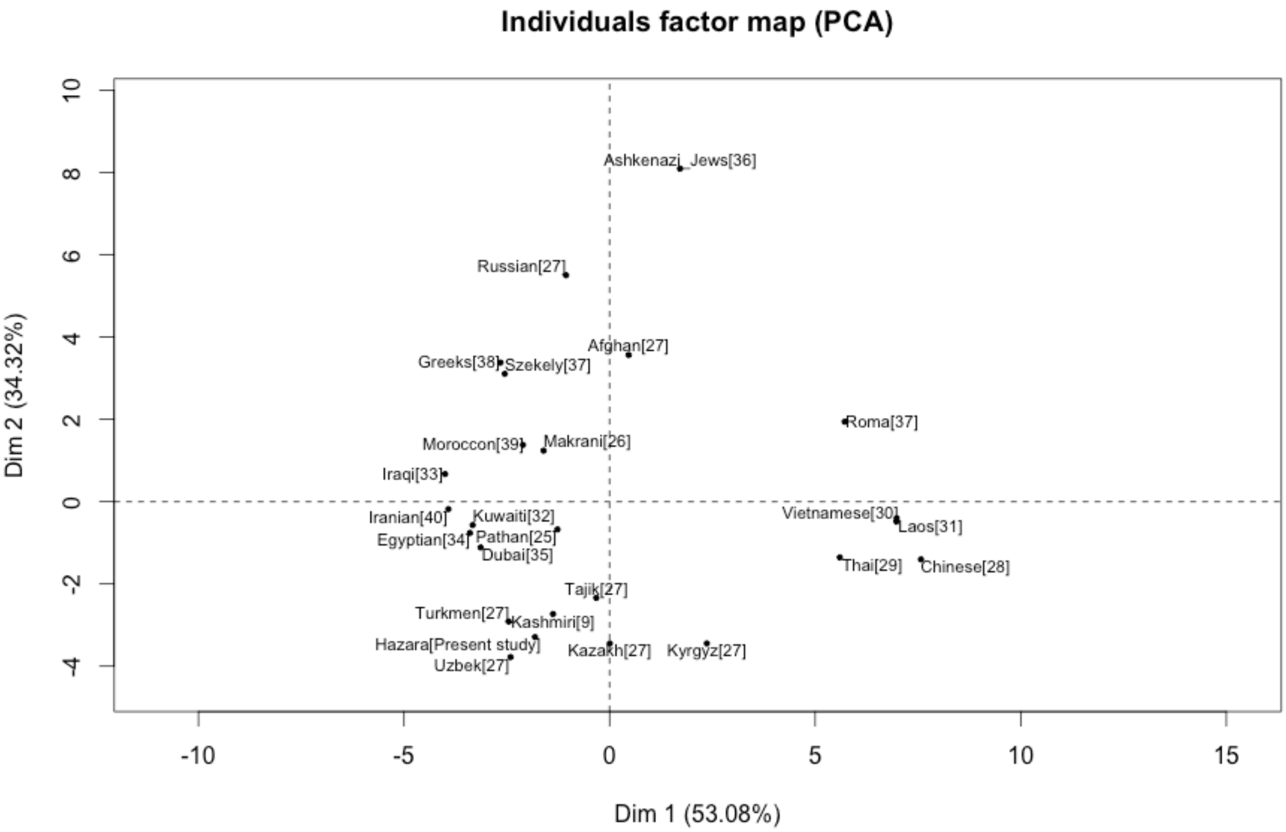

Figure S3: Principal component analysis (PCA) based on pairwise F_{ST} values defined by complete control region sequences

Data Sources:

Populations	**References**	**No. of samples**	**Populations**	**References**	**No. of samples**
Hazara	Present study	319	Chinese	[28]	377
Kashmiri	[9]	317	Thai	[29]	190
Pathan	[25]	230	Vietnamese	[30]	187
Makrani	[26]	99	Laos	[31]	214
Uzbek	[27]	328	Kuwaiti	[32]	381
Turkmen	[27]	249	Iraqi	[33]	182
Tajik	[27]	244	Egyptian	[34]	277
Russian	[27]	151	Dubai	[35]	249
Kyrgyz	[27]	249	Ashkenazi Jews	[36]	173
Kazakh	[27]	256	Roma	[37]	205
Afghan	[27]	98	Greeks	[38]	319
Szekely	[37]	182	Iranian	[40]	352
Moroccan	[39]	509			

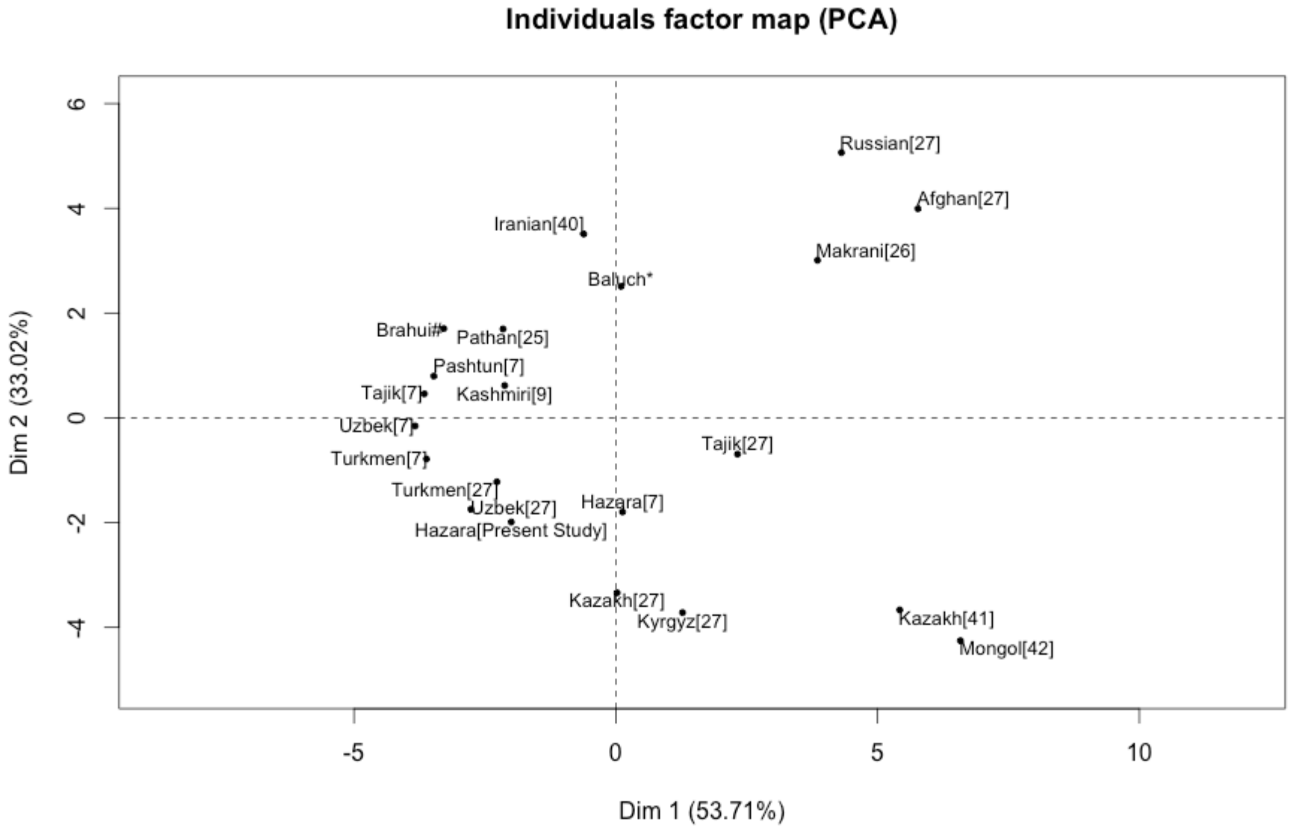

Figure S4: Principal component analysis (PCA) based on pairwise F_{ST} values defined by HVI sequences

Data Sources:

Populations	**References**	**No. of samples**	**Populations**	**References**	**No. of samples**
Hazara	Present study	319	Mongol	[44]	201
Kashmiri	[9]	317	Kazakh	[43]	160
Pathan	[25]	230	Tajik	[7]	146
Makrani	[26]	99	Hazara	[7]	78
Uzbek	[27]	328	Turkmen	[7]	75
Turkmen	[27]	249	Uzbek	[7]	127
Tajik	[27]	244	Baluch*	EU565791-EU565815	25
Russian	[27]	151	Brahui#	EU565766-EU565790	25
Kyrgyz	[27]	249	Iranian	[40]	352
Kazakh	[27]	256	Afghan	[7]	90
Afghan	[27]	98			

Table S1. Mitochondrial DNA control region (16024-576) haplotypes and estimated haplogroups for Hazara ethnic group from Pakistan

SampleI	MRCA	Polymorphisms																					
HZR001	H14b2a	16093C	197G	263G	309.1C	309.2C	315.1C	524.1A	524.2C
HZR002	M43a1	16171G	16223T	16311C	16362C	73G	263G	315.1C	489C
HZR003	B4c1c1	16183C	16189C	16193.1C	16217C	16311C	73G	146C	150T	195C	214G	263G	315.1C
HZR004	A+152+16362	16223T	16290T	16319A	16362C	16519C	73G	152C	235G	263G	315.1C	523d	524d
HZR005	G2a+152	16223T	16227G	16278T	16362C	73G	152C	263G	309.1C	315.1C	489C	523d	524d
HZR006	F1d	16183C	16189C	16193.1C	16304C	16519C	73G	146C	249d	263G	309.1C	315.1C	523d	524d
HZR007	H2a2a	16519C	263G	309.1C	309.2C	315.1C
HZR009	H2a2a	263G	315.1C
HZR010	H1b	16242T	16356C	263G	309.1C	315.1C
HZR012	HV2	16217C	16519C	73G	152C	263G	315.1C	455.1T
HZR014	H10+(16093)	16093C	189G	207A	263G	309.1C	309.2C	315.1C
HZR015	U5a1+@16192	16256T	16270T	16399G	73G	263G	315.1C
HZR016	A23	16183C	16189C	16223T	16290T	16319A	16362C	73G	97T	105d	106d	107d	108d	109d	110d	150T	152C	200G	235G	263G	315.1C	523d	524d
HZR017	B4b1a3a	16086C	16136C	16182C	16183C	16189C	16217C	16519C	73G	146C	263G	315.1C	408A	499A
HZR018	D4p	16223T	16362C	73G	195C	198T	263G	315.1C	489C
HZR019	D4c2a	16176T	16223T	16245T	16362C	16368C	73G	195C	263G	315.1C	489C
HZR020	H14a	16256T	16352C	16362C	263G
HZR021	D4e5a	16223T	16274A	16362C	73G	152C	185A	263G	309.1C	315.1C	489C	573.1C	573.2C
HZR022	U7	16294T	16309G	16318T	16519C	73G	152C	153G	263G	309.1C	315.1C	523d	524d
HZR023	X2	16111T	16183C	16189C	16193.1C	16223T	16278T	16519C	73G	153G	195C	263G	315.1C
HZR025	C4a1a4a	16129A	16150T	16223T	16298C	16327T	16519C	73G	195C	249d	263G	315.1C	333C	489C
HZR028	R2	16051G	16071T	16234T	16519C	73G	150T	152C	182T	263G	315.1C	524.1A	524.2C
HZR029	A+152+16362	16223T	16290T	16319A	16362C	16519C	73G	152C	235G	263G	315.1C	523d	524d
HZR030	M18a	16114G	16223T	16318T	16519C	73G	93G	194T	246C	263G	296T	309.1C	315.1C	489C
HZR031	J1d	16069T	16126C	16193T	16519C	73G	152C	263G	295T	315.1C	316A	462T	489C
HZR032	A+152+16362	16223T	16290T	16319A	16362C	16519C	73G	152C	235G	263G	315.1C	523d	524d
HZR033	X2	16189C	16223T	16278T	16519C	73G	153G	195C	263G	315.1C
HZR034	M30b	16189C	16223T	16278T	16519C	73G	152C	195A	263G	315.1C	489C	523d	524d
HZR035	M12a1b	16172C	16183C	16189C	16209C	16223T	16234T	16290T	16519C	73G	125C	127C	128T	146C	195C	263G	309.1C	315.1C	489C
HZR037	U7a	16274A	16309G	16318T	16519C	73G	151T	152C	263G	315.1C	523d	524d
HZR038	H2a2a	16519C	263G	309.1C	315.1C	513A	524.1A	524.2C	573.1C
HZR039	A+152+16362	16223T	16290T	16319A	16362C	16519C	73G	152C	235G	263G	315.1C	523d	524d
HZR040	U5a2a	16114A	16192T	16256T	16270T	16294T	16526A	73G	189G	263G	309.1C	315.1C
HZR041	D4b2b	16223T	16362C	16519C	73G	194T	263G	315.1C	489C	523d	524d
HZR042	M9a1a1a	16223T	16234T	16291T	16316G	16362C	16519C	73G	263G	309.1C	315.1C	489C
HZR043	D4m2a	16042A	16093C	16214T	16223T	16256T	16362C	12C	73G	263G	315.1C	489C
HZR044	A16	16223T	16290T	16292A	16319A	16362C	73G	152C	199C	235G	263G	298A	309.1C	315.1C	523d	524d
HZR045	A16	16223T	16290T	16292A	16319A	16362C	73G	152C	199C	235G	263G	298A	309.1C	315.1C	523d	524d
HZR046	T2a1b	16126C	16294T	16296T	16324C	16519C	73G	263G	315.1C
HZR047	F1b1b	16172C	16183C	16189C	16193.1C	16232A	16249C	16304C	16311C	73G	152C	204C	249d	263G	309.1C	315.1C	523d	524d
HZR048	M30	16223T	16265G	16355T	16519C	73G	195A	263G	315.1C	489C	523d	524d
HZR049	H2a2a	16519C	263G	309.1C	315.1C
HZR050	G2a+152	16223T	16227G	16278T	16362C	73G	151T	152C	263G	315.1C	489C	523d	524d
HZR052	HV6	16172C	16311C	263G	309.1C	315.1C
HZR053	H2a1	16189C	16192T	16354T	263G	309.1C	309.2C	315.1C
HZR054	M10a1a1b	16093C	16129A	16193T	16223T	16311C	16357C	16497G	73G	146C	263G	309.1C	315.1C	489C	523d	524d	573.1C	573.2C	573.3C
HZR055	F1b1+@152	16129A	16183C	16189C	16193.1C	16215G	16232A	16249C	16304C	16311C	16519C	73G	249d	263G	315.1C	523d	524d
HZR057	U2b	16051G	16114A	73G	146C	263G	309.1C	315.1C
HZR058	C4a2	16223T	16297C	16298C	16327T	16357C	16519C	73G	152C	249d	263G	315.1C	315.2C	489C
HZR059	H13a1b	16243C	16519C	153G	200G	263G	309.1C	315.1C
HZR060	HV1b+152	16067T	16311C	152C	263G	315.1C
HZR061	R0a+60.1T	16126C	16362C	58C	60.1T	64T	152C	263G	309.1C	309.2C	315.1C
HZR062	C4b8a	16093C	16223T	16298C	16327T	16519C	73G	249d	263G	315.1C	489C	523d	524d
HZR063	M8a2+152	16184T	16223T	16293G	16298C	16319A	73G	152C	189G	263G	309.1C	315.1C	489C
HZR064	C4a1a+195	16093C	16129A	16223T	16298C	16327T	16519C	73G	195C	249d	263G	315.1C	489C	523d	524d
HZR065	W3a1+199	16093C	16292T	16519C	73G	189G	194T	195C	199C	204C	207A	263G	315.1C
HZR066	M9a1a1a	16223T	16234T	16316G	16362C	73G	263G	315.1C	489C
HZR067	U2c'd	16051G	16126C	16169T	16234T	16278T	16519C	73G	152C	263G	315.1C
HZR068	C4a1a+195	16129A	16223T	16298C	16327T	16519C	73G	195C	249d	263G	309.1C	315.1C	489C
HZR069	HV14	16311C	263G	309.1C	315.1C	480C
HZR070	H2a2a	16519C	263G	309.1C	309.2C	315.1C
HZR071	J1b8	16069T	16093C	16145A	16189C	16193.1C	16261T	16362C	73G	151T	152C	263G	295T	309.1C	315.1C	462T	489C
HZR072	K1b1a1+199	16224C	16311C	16319A	16463G	73G	152C	199C	263G	309.1C	315.1C	524.1A	524.2C
HZR073	M54	16223T	16304C	16519C	73G	263G	309.1C	315.1C	489C
HZR074	D4+195	16172C	16223T	16362C	73G	195C	263G	309.1C	315.1C	489C
HZR075	B4c1b+16335	16093C	16140C	16182C	16183C	16189C	16217C	16274A	16335G	16519C	73G	150T	263G	309.1C	315.1C
HZR076	R7	16183C	16189C	16193.1C	16319A	16362C	73G	151T	152C	240G	263G	309.1C	315.1C
HZR077	M9a1a1c	16223T	16234T	16291T	16316G	16362C	73G	153G	263G	309.1C	315.1C	489C
HZR078	G2a+152	16223T	16227G	16278T	16362C	16519C	73G	152C	263G	309.1C	315.1C	325T	489C	523d	524d
HZR079	M65a+@16311	16137G	16223T	16289G	16360T	16519C	73G	263G	315.1C	489C	511T
HZR080	K2a5	16224C	16311C	16519C	73G	146C	152C	195C	263G	309.1C	315.1C	324T	430C
HZR081	A+152+16362	16223T	16290T	16319A	16355T	16362C	16519C	73G	152C	235G	263G	315.1C	523d	524d
HZR082	T	16126C	16294T	16304C	16519C	73G	263G	315.1C
HZR083	HV2a	16217C	72C	73G	152C	195C	263G	309.1C	315.1C
HZR084	HV2	16217C	16519C	73G	152C	263G	309.1C	315.1C	455.1T
HZR085	H14a	16249C	16256T	16311C	16352C	16456A	263G	309.1C	309.2C	315.1C	417A
HZR086	H1e1a4	16243C	16311C	16519C	204C	263G	309.1C	309.2C	315.1C
HZR087	W6	16192T	16223T	16292T	16325C	16519C	73G	185A	189G	194T	195C	204C	207A	263G	315.1C
HZR088	HV1b+152	16067T	16311C	152C	263G	315.1C
HZR089	HV2	16217C	16519C	73G	152C	263G	309.1C	315.1C	455.1T	524.1A	524.2C
HZR092	M9a1a1a	16223T	16234T	16316G	16362C	73G	121C	263G	309.1C	315.1C	489C
HZR093	B4b1a3a	16086C	16136C	16182C	16183C	16189C	16193.1C	16193.2C	16217C	16519C	73G	146C	263G	315.1C	408A	499A
HZR094	C4a2	16223T	16297C	16298C	16327T	16357C	16519C	73G	152C	249d	263G	315.1C	315.2C	489C
HZR095	A+152+16362	16223T	16290T	16319A	16355T	16362C	16519C	73G	152C	235G	263G	315.1C	523d	524d
HZR096	J2b1a	16069T	16126C	16193T	16278T	16519C	73G	150T	152C	235G	263G	295T	309.1C	315.1C	489C
HZR097	K1a4c1	16224C	16246T	16311C	16519C	73G	263G	309.1C	315.1C	497T
HZR098	A15	16223T	16290T	16319A	16362C	73G	152C	207A	235G	309.1C	315.1C	523d	524d
HZR099	J1d5	16069T	16126C	16193T	16274A	16519C	73G	146C	152C	263G	295T	309.1C	315.1C	462T	489C	524.1A	524.2C	524.3A	524.4C
HZR100	H2a1	16354T	152C	263G	315.1C
HZR101	U2e2	16051G	16092C	16129C	16153A	16182C	16183C	16189C	16261T	16362C	16519C	73G	151T	152C	200G	217C	263G	315.1C	508G
HZR102	A+152+16362	16223T	16290T	16319A	16362C	16519C	73G	152C	235G	263G	315.1C	523d	524d
HZR103	W+194	16223T	16292T	16519C	73G	189G	194T	195C	204C	207A	263G	315.1C
HZR104	H1e1a4	16243C	16311C	16519C	204C	263G	309.1C	315.1C
HZR105	D4c2b	16189C	16223T	16245T	16362C	73G	263G	315.1C	489C	534T
HZR106	D4c2b	16093C	16172C	16223T	16245T	16362C	73G	263G	315.1C	489C	534T
HZR108	R8a1a3	16292T	16519C	73G	195C	214G	263G	309.1C	309.2C	315.1C
HZR109	T2c1a	16126C	16287T	16292T	16294T	16519C	73G	263G	309.1C	315.1C	573.1C	573.2C	573.3C	573.4C	573.5C
HZR110	D4i	16223T	16294T	16362C	73G	146C	195C	263G	315.1C	489C
HZR111	J1c	16069T	16126C	16259.1A	16299G	73G	185A	228A	263G	295T	315.1C	462T	489C
HZR114	J1d5	16069T	16126C	16193T	16274A	16519C	73G	146C	152C	263G	295T	309.1C	315.1C	462T	489C	524.1A	524.2C	524.3A	524.4C
HZR115	HV2	16217C	16519C	73G	152C	263G	315.1C	455.1T
HZR117	H2a2a	16519C	263G	315.1C
HZR118	HV4b	16069T	263G	309.1C	315.1C	523d	524d
HZR119	H1e1a4	16158T	16311C	16519C	263G	315.1C
HZR120	M5a2a1a	16129A	16223T	16265C	16344T	16519C	73G	263G	315.1C	374G	489C
HZR121	HV1b+152	16067T	16311C	152C	263G	315.1C
HZR122	K2a5	16224C	16311C	16519C	73G	146C	152C	195C	263G	315.1C	324T	430C
HZR123	D4i	16223T	16294T	16362C	73G	146C	195C	263G	315.1C	489C
HZR124	H2a1	16189C	16192T	16354T	263G	309.1C	315.1C
HZR125	U4'9	16129A	16311C	16362C	16519C	73G	195C	263G	315.1C	499A
HZR126	M30c1	16166d	16207G	16223T	16390A	16519C	73G	146C	195A	263G	309.1C	315.1C	489C	523d	524d
HZR127	D4+195	16172C	16223T	16362C	73G	195C	263G	309.1C	315.1C	489C
HZR128	M54	16223T	16304C	16519C	73G	263G	309.1C	309.2C	315.1C	489C	567G
HZR129	U9a1	16051G	16193T	16278T	16357C	73G	263G	315.1C	499A
HZR130	M65a+@16311	16137G	16223T	16289G	16360T	16519C	73G	263G	315.1C	489C	511T
HZR131	M5a2a1a	16129A	16223T	16265C	16519C	73G	263G	315.1C	489C
HZR132	U5a2a	16114A	16192T	16256T	16270T	16294T	16526A	73G	189G	263G	309.1C	315.1C
HZR133	HV2a	16217C	16256T	72C	73G	152C	195C	263G	266C	315.1C
HZR134	H13a1c	16244A	16519C	152C	263G	315.1C
HZR135	A+152+16362+16189	16182C	16183C	16189C	16223T	16290T	16319A	16362C	73G	150T	152C	204C	235G	263G	309.1C	315.1C	523d	524d
HZR136	J1b8	16069T	16093C	16145A	16189C	16193.1C	16261T	16362C	73G	151T	152C	263G	295T	309.1C	315.1C	462T	489C
HZR138	C5c	16093C	16223T	16288C	16298C	16327T	16518T	16519C	16527T	73G	152C	249d	263G	309.1C	315.1C	489C
HZR140	R30b2	16292T	16390A	16497G	16519C	73G	152C	263G	309.1C	315.1C	373G
HZR141	R30b2	16189C	16193.1C	16292T	16497G	16519C	73G	150T	263G	309.1C	315.1C	373G
HZR142	J1c2i	16069T	16126C	16147T	16242T	73G	185A	188G	228A	263G	295T	315.1C	462T	489C
HZR143	R0a+60.1T	16126C	16362C	58C	60.1T	64T	152C	263G	309.1C	309.2C	315.1C
HZR144	U9a1	16051G	16193T	16278T	16357C	73G	263G	315.1C	499A
HZR145	C4a1a+195	16093C	16129A	16223T	16298C	16311C	16327T	16362C	16519C	73G	152C	195C	249d	263G	315.1C	489C
HZR146	U5a2a	16114A	16192T	16256T	16270T	16294T	16526A	73G	234G	263G	309.1C	315.1C
HZR147	HV19	16213A	16260T	16399G	16519C	146C	309.1C	309.2C	315.1C
HZR148	M9a1a1a	16223T	16234T	16316G	16362C	73G	263G	315.1C	489C
HZR149	HV1b+152	16067T	16311C	152C	263G	315.1C
HZR150	U5a2a	16114A	16192T	16256T	16270T	16294T	16526A	73G	189G	263G	309.1C	315.1C
HZR151	J1b1b	16069T	16126C	16145A	16261T	16519C	73G	263G	271T	295T	315.1C	462T	489C	512G
HZRm001	D4c2b	16093C	16223T	16245T	16262C	73G	263G	309.1C	315.1C	489C	534T
HZRm002	Y1a+16189	16126C	16189C	16231C	16266T	16519C	73G	146C	263G	309.1C	315.1C
HZRm003	D4c2a	16176T	16223T	16245T	16362C	16368C	73G	195C	263G	315.1C	489C
HZRm004	U5a1+@16192	16256T	16270T	16399G	73G	263G	315.1C
HZRm005	M43a1	16171G	16223T	16311C	16362C	73G	263G	315.1C	489C
HZRm007	B5a	16140C	16182C	16183C	16189C	16193.1C	16266A	16519C	73G	210G	263G	309.1C	309.2C	315.1C	391C	523d	524d
HZRm008	HV6	16172C	16311C	195C	263G	309.1C	315.1C
HZRm009	HV2a	16217C	72C	73G	152C	195C	263G	315.1C
HZRm011	H2a2a	16519C	263G	309.1C	315.1C	513A	524.1A	524.2C	573.1C
HZRm012	H3b6	16129A	16146G	16318T	16519C	73G	152C	263G	309.1C	315.1C	523d	524d
HZRm013	D2b	16129A	16148T	16223T	16271C	16362C	73G	195C	263G	315.1C	489C
HZRm014	H1o	16355T	150T	263G	309.1C	309.2C	315.1C
HZRm015	D4c2a	16176T	16223T	16245T	16362C	16368C	73G	195C	263G	315.1C	489C
HZRm016	M43a1	16223T	16311C	16362C	16519C	73G	263G	309.1C	315.1C	489C
HZRm017	H2a2a	16519C	263G	309.1C	309.2C	315.1C
HZRm018	M9	16223T	16362C	73G	263G	309.1C	315.1C	489C
HZRm019	H2a5b1	16176T	16519C	249d	263G	315.1C
HZRm020	H2a2a	16519C	263G	309.1C	309.2C	315.1C
HZRm021	W6	16147T	16185T	16192T	16223T	16274A	16292T	16325C	16519C	73G	189G	194T	195C	204C	207A	263G	309.1C	309.2C	315.1C
HZRm022	HV1b+152	16067T	16311C	152C	263G	315.1C
HZRm023	C5c	16093C	16223T	16288C	16298C	16327T	16518T	16519C	16527T	73G	152C	249d	263G	309.1C	315.1C	489C
HZRm024	G2a+152	16223T	16227G	16274A	16278T	16362C	73G	152C	195C	263G	315.1C	489C	523d	524d
HZRm025	F1a1a	16108T	16129A	16162G	16172C	16304C	16519C	73G	150T	249d	263G	309.1C	315.1C	523d	524d
HZRm027	HV1b+152	16067T	16311C	152C	263G	315.1C
HZRm028	K1a4c1	16224C	16246T	16311C	16519C	73G	263G	309.1C	315.1C	497T
HZRm029	D4j+(16286)	16223T	16286T	16362C	73G	263G	309.1C	315.1C	489C
HZRm030	D4h1	16093C	16174T	16209C	16223T	16362C	73G	263G	315.1C	489C	573.1C	573.2C	573.3C	573.4C	573.5C
HZRm031	J1c	16069T	16126C	16259.1A	16299G	73G	185A	228A	263G	295T	315.1C	462T	489C
HZRm032	M65a+@16311	16137G	16223T	16289G	16360T	16519C	73G	263G	315.1C	489C	511T
HZRm033	A15	16223T	16290T	16319A	16362C	73G	152C	207A	235G	309.1C	315.1C	523d	524d
HZRm034	T	16126C	16294T	16304C	16519C	73G	263G	309.1C	315.1C
HZRm035	R0a1a	16126C	16266T	16355T	16362C	44.1C	58C	64T	146C	263G	309.1C	309.2C	315.1C
HZRm036	A+152+16362	16223T	16290T	16319A	16362C	16519C	73G	152C	235G	263G	315.1C	523d	524d
HZRm037	G2a2a	16223T	16227G	16278T	16362C	73G	152C	207A	263G	309.1C	315.1C	489C	523d	524d
HZRm038	C4a1a4a	16129A	16150T	16223T	16298C	16327T	16519C	73G	195C	249d	263G	315.1C	333C	489C
HZRm039	H11a1	16278T	16293G	16311C	16319A	16390A	16519C	195C	263G	309.1C	309.2C	315.1C	573.1C
HZRm041	U7a	16318T	16519C	73G	151T	152C	263G	315.1C	523d	524d
HZRm042	T2b34	16126C	16294T	16296T	16304C	16519C	41T	73G	263G	315.1C	319C
HZRm043	C4a1a+195	16129A	16223T	16298C	16327T	16519C	73G	195C	249d	263G	309.1C	315.1C	489C
HZRm044	T1a1'3	16126C	16147T	16163G	16186T	16189C	16234T	16278T	16294T	16519C	73G	152C	195C	263G	309.1C	315.1C
HZRm045	M9a1a1a	16223T	16234T	16316G	16362C	73G	263G	315.1C	489C
HZRm046	U1a1a1	16182C	16183C	16189C	16249C	73G	263G	285T	309.1C	315.1C	385G	523d	524d	573.1C	573.2C	573.3C
HZRm047	R2	16071T	16519C	73G	152C	263G	315.1C
HZRm048	M5a2a1a	16129A	16223T	16265C	16519C	73G	185A	263G	315.1C	489C
HZRm049	A+152+16362+16189	16189C	16223T	16290T	16319A	16362C	16519C	73G	152C	235G	263G	315.1C	523d	524d
HZRm050	D4c2b	16189C	16223T	16245T	16362C	73G	263G	315.1C	489C	534T
HZRm051	D4c2b	16189C	16223T	16245T	16362C	73G	263G	315.1C	489C	534T
HZRm052	D4g2	16223T	16271C	16362C	16519C	73G	263G	298T	309.1C	315.1C	489C
HZRm053	T2c1a	16126C	16292T	16294T	16519C	73G	263G	309.1C	309.2C	315.1C	573.1C	573.2C	573.3C	573.4C
HZRm054	J1d	16069T	16093C	16126C	16193T	16239T	16519C	73G	152C	263G	295T	315.1C	462T	489C
HZRm055	H1e1a4	16243C	16311C	16519C	204C	263G	309.1C	309.2C	315.1C
HZRm056	N9a1	16111T	16129A	16223T	16257A	16261T	73G	150T	263G	309.1C	315.1C
HZRm057	M7b1a1	16129A	16223T	16297C	73G	150T	159C	199C	263G	315.1C	489C
HZRm058	U7a	16274A	16309G	16318T	16519C	73G	151T	152C	263G	315.1C	523d	524d
HZRm059	HV6	16172C	16311C	263G	309.1C	315.1C
HZRm060	H1+16189	16183C	16189C	16519C	263G	309.1C	315.1C
HZRm061	JT	16126C	16181G	16209C	16362C	16519C	73G	263G	309.1C	315.1C
HZRm062	D4m2a	16042A	16093C	16214T	16223T	16256T	16362C	73G	263G	315.1C	489C
HZRm063	H1e1a4	16243C	16311C	16519C	204C	263G	315.1C
HZRm065	U4'9	16129A	16311C	16362C	16519C	73G	195C	263G	315.1C	499A
HZRm066	D4b1	16223T	16319A	16362C	73G	263G	309.1C	315.1C	489C	523d	524d
HZRm067	U5a2a	16114A	16192T	16256T	16270T	16294T	16526A	73G	234G	263G	309.1C	309.2C	315.1C
HZRm069	C	16223T	16298C	16327T	16519C	73G	249d	263G	309.1C	315.1C	489C
HZRm070	U5a2a	16114A	16192T	16256T	16270T	16294T	16526A	73G	189G	263G	309.1C	315.1C
HZRm071	H2a2a	263G	315.1C
HZRm072	M3a1+204	16126C	16223T	16519C	73G	204C	263G	309.1C	315.1C	482C	489C
HZRm073	U2b2	16051G	16086C	16291T	16305T	16353T	73G	146C	234G	263G	315.1C
HZRm074	B4c1b+16335	16093C	16140C	16182C	16183C	16189C	16193.1C	16217C	16274A	16335G	16519C	73G	150T	263G	309.1C	309.2C	315.1C
HZRm075	R0a1a	16126C	16266T	16355T	16362C	44.1C	58C	64T	146C	263G	315.1C
HZRm076	U1a	16182C	16183C	16189C	16193.1C	16249C	16519C	73G	263G	285T	309.1C	309.2C	315.1C
HZRm077	C	16223T	16287T	16298C	16327T	16519C	73G	189G	249d	263G	309.1C	315.1C	489C
HZRm078	U2e2	16051G	16092C	16129C	16153A	16182C	16183C	16189C	16261T	16362C	16519C	73G	151T	152C	200G	217C	263G	309.1C	315.1C	499A	508G
HZRm079	U1a1a1	16182C	16183C	16189C	16193.1C	16249C	73G	263G	285T	309.1C	309.2C	315.1C	385G	523d	524d	573.1C	573.2C	573.3C	573.4C	573.5C
HZRm080	U2c'd	16051G	16126C	16169T	16234T	16278T	16519C	73G	152C	263G	315.1C
HZRm081	B5b2c	16111T	16140C	16183C	16189C	16234T	16243C	16463G	16519C	73G	103A	131C	263G	309.1C	309.2C	315.1C	481T	523d	524d
HZRm082	H14a	16249C	16256T	16311C	16352C	16456A	263G	309.1C	309.2C	315.1C	417A
HZRm083	D4l2	16223T	16362C	16368C	125C	127C	263G	315.1C	489C
HZRm084	T1a1'3	16126C	16147T	16163G	16186T	16189C	16234T	16278T	16294T	16519C	73G	152C	195C	263G	309.1C	315.1C
HZRm085	H2a2a	263G	315.1C
HZRm086	R0a1a	16126C	16266T	16355T	16362C	44.1C	58C	64T	146C	263G	309.1C	309.2C	315.1C
HZRm087	M8a2a	16184T	16189C	16223T	16293G	16298C	16319A	73G	152C	263G	315.1C	489C
HZRm088	M9a1a1a	16223T	16234T	16291T	16316G	16362C	16519C	73G	263G	309.1C	315.1C	489C
HZRm089	M8a2+152	16184T	16223T	16293G	16298C	16319A	73G	152C	189G	263G	315.1C	489C
HZRm090	N1a1a1a1a	16147A	16172C	16189C	16223T	16248T	16320T	16355T	16519C	73G	152C	199C	204C	263G	309.1C	315.1C	573.1C	573.2C	573.3C	573.4C	573.5C
HZRm091	C4a1a+195	16093C	16129A	16179T	16223T	16298C	16327T	16519C	73G	195C	249d	263G	309.1C	315.1C	489C
HZRm092	C	16223T	16298C	16327T	16519C	73G	249d	263G	309.1C	315.1C	489C
HZRm093	R0a+60.1T	16126C	16311C	16362C	58C	60.1T	64T	152C	263G	309.1C	309.2C	315.1C
HZRm094	A16	16223T	16290T	16292A	16319A	16362C	73G	152C	199C	235G	263G	298A	309.1C	315.1C	523d	524d
HZRm095	U5a2a	16114A	16192T	16256T	16270T	16294T	16526A	73G	234G	263G	309.1C	315.1C
HZRm096	X2	16189C	16223T	16278T	16519C	73G	153G	195C	263G	315.1C
HZRm097	R8a1a1d	16172C	16184A	16390A	73G	152C	195C	263G	309.1C	315.1C
HZRm098	H1m1	16399G	16519C	146C	263G	309.1C	315.1C
HZRm099	U5a1a1+152	16256T	16270T	16399G	73G	152C	263G	309.1C	315.1C
HZRm100	Z+152	16185T	16209C	16260T	16298C	73G	152C	249d	263G	309.1C	315.1C	489C
HZRm101	H1e1a4	16243C	16311C	16519C	204C	263G	309.1C	315.1C
HZRm102	H2a2a	16519C	263G	309.1C	315.1C
HZRm103	U4b3	16356C	16519C	73G	195C	215G	263G	309.1C	309.2C	315.1C	499A	524.1A	524.2C
HZRm104	M9a1a1a	16223T	16234T	16316G	16362C	73G	263G	315.1C	489C
HZRm105	H14a	16249C	16256T	16311C	16352C	16456A	263G	309.1C	309.2C	315.1C	417A
HZRm106	C5c	16093C	16223T	16288C	16298C	16327T	16518T	16519C	16527T	73G	249d	263G	315.1C	489C
HZRm107	K1a4c1	16224C	16246T	16311C	16519C	73G	263G	309.1C	315.1C	497T
HZRm108	M54	16223T	16304C	16519C	73G	263G	309.1C	315.1C	489C
HZRm109	M65a+@16311	16137G	16223T	16289G	16360T	16519C	73G	263G	315.1C	489C	511T
HZRm110	HV2a	16217C	72C	73G	152C	195C	263G	315.1C
HZRm111	H2a1	16189C	16192T	16354T	263G	309.1C	315.1C
HZRm112	U1a1a	16183C	16185T	16189C	16193d	16249C	73G	185A	263G	285T	309.1C	309.2C	315.1C	385G	523d	524d
HZRm113	U2b1	16051G	16168T	16189C	16201T	16390A	16T	73G	146C	215G	263G	279.1T	315.1C
HZRm114	F1b1+@152	16183C	16189C	16215G	16232A	16249C	16304C	16311C	16519C	73G	249d	263G	315.1C	523d	524d
HZRm115	J1b8	16069T	16093C	16145A	16189C	16261T	16362C	73G	151T	152C	263G	295T	309.1C	315.1C	462T	489C
HZRm116	H10+(16093)	16093C	189G	207A	263G	309.1C	315.1C
HZRm117	U1a	16182C	16183C	16189C	16193.1C	16249C	16519C	73G	263G	285T	309.1C	315.1C
HZRm118	U5a1	16192T	16256T	16270T	16399G	16519C	73G	195C	198T	263G	315.1C
HZRm119	H1e1a4	16243C	16311C	16519C	204C	263G	309.1C	315.1C
HZRm120	D4m2a	16042A	16214T	16223T	16256T	16362C	73G	263G	315.1C	489C
HZRm121	HV19	16213A	16260T	16399G	16519C	146C	309.1C	309.2C	315.1C
HZRm122	HV6	16172C	16311C	263G	309.1C	315.1C
HZRm123	U1a1a	16183C	16185T	16189C	16193d	16249C	73G	185A	263G	285T	309.1C	315.1C	385G	523d	524d
HZRm124	B5a	16140C	16182C	16183C	16189C	16193.1C	16266A	16519C	73G	210G	263G	309.1C	315.1C	391C	523d	524d
HZRm125	U5a2a	16114A	16192T	16256T	16270T	16294T	16526A	73G	234G	263G	309.1C	315.1C
HZRm126	D4c2b	16189C	16223T	16245T	16362C	73G	263G	315.1C	489C	534T
HZRm127	H29	16519C	263G	309.1C	315.1C	573.1C
HZRm128	M2a2	16223T	16240C	16274A	16311C	16319A	16519C	73G	263G	309.1C	315.1C	447G	489C
HZRm129	U7a	16274A	16309G	16318T	16519C	73G	151T	152C	263G	315.1C	523d	524d
HZRm130	HV14	16311C	16488T	263G	309.1C	315.1C	459d	480C
HZRm131	U5a2a	16114A	16192T	16256T	16270T	16294T	16526A	73G	234G	263G	309.1C	315.1C
HZRm132	D4c2b	16189C	16223T	16245T	16362C	73G	263G	315.1C	489C	534T
HZRm133	B4b1a3a	16086C	16136C	16182C	16183C	16189C	16193.1C	16217C	16519C	73G	146C	263G	309.1C	309.2C	315.1C	408A	499A	524.1A	524.2C
HZRm134	B4b1a+207	16136C	16182C	16183C	16189C	16217C	16519C	73G	207A	263G	309.1C	309.2C	315.1C	499A
HZRm135	HV2a	16217C	72C	73G	152C	195C	263G	315.1C
HZRm136	D4m2a	16042A	16093C	16214T	16223T	16256T	16362C	12C	73G	263G	315.1C	489C
HZRm137	C5c	16093C	16223T	16288C	16298C	16327T	16518T	16519C	16527T	73G	152C	249d	263G	309.1C	315.1C	489C
HZRm138	M9	16223T	16357C	16362C	73G	263G	309.1C	309.2C	315.1C	489C
HZRm139	B4c1b+16335	16093C	16140C	16182C	16183C	16189C	16193.1C	16217C	16274A	16335G	16519C	73G	150T	263G	309.1C	309.2C	315.1C
HZRm140	R2	16071T	16519C	73G	152C	263G	315.1C
HZRm141	M9a1a1a	16223T	16234T	16316G	16362C	73G	263G	315.1C	489C
HZRm149	D4b1a1	16223T	16239T	16242T	16243C	16319A	16362C	73G	152C	195C	263G	315.1C	489C	523d	524d
HZRm150	G2a+152	16126C	16223T	16227G	16278T	16362C	16519C	73G	152C	263G	309.1C	315.1C	325T	489C	523d	524d
HZRm151a	U8b1b	16183C	16189C	16234T	16324C	16519C	73G	195C	263G	309.1C	309.2C	315.1C
HZRm152a	HV2	16217C	16519C	73G	152C	263G	309.1C	315.1C	455.1T
HZRm151b	J1d	16069T	16193T	16519C	73G	151T	152C	263G	295T	315.1C	462T	489C
HZRm152b	M9a1a1a	16223T	16234T	16316G	16362C	73G	263G	309.1C	315.1C	489C
HZRm153	H2a2a	16519C	263G	309.1C	315.1C
HZRm154	U2c'd	16051G	16126C	16169T	16234T	16278T	16519C	73G	152C	263G	315.1C
HZRm155	U1a1a	16183C	16185T	16189C	16193d	16249C	73G	185A	263G	285T	309.1C	315.1C	385G	523d	524d
HZRm156	M9a1a1a	16223T	16234T	16316G	16362C	73G	263G	309.1C	315.1C	489C
HZRm157	F1a1	16129A	16162G	16172C	16304C	16519C	73G	146C	249d	263G	315.1C	523d	524d
HZRm158	M2a1a	16223T	16270T	16319A	16352C	16519C	73G	195C	204C	263G	309.1C	309.2C	315.1C	447G	489C	523d	524d
HZRm159	M43a1	16223T	16311C	16362C	73G	263G	309.1C	315.1C	489C
HZRm160	HV6	16172C	16311C	263G	309.1C	315.1C
HZRm161	M30e	16093C	16223T	16234T	16519C	73G	150T	152C	195A	263G	315.1C	489C	523d	524d
HZRm162	M9a1a1a	16223T	16234T	16316G	16362C	73G	263G	315.1C	489C
HZRm163	D4h1	16093C	16174T	16209C	16223T	16362C	73G	263G	309.1C	315.1C	489C	573.1C	573.2C	573.3C	573.4C
HZRm164	H5	16261T	16304C	263G	315.1C	456T
HZRm165	U5a1+@16192	16256T	16270T	16399G	73G	263G	315.1C
HZRm166	U8b1b	16183C	16189C	16192T	16234T	16324C	16519C	73G	195C	263G	309.1C	315.1C
HZRm167	U1a	16182C	16183C	16189C	16193.1C	16249C	16519C	73G	263G	285T	309.1C	309.2C	315.1C	389A
HZRm168	M8a2a	16184T	16189C	16223T	16293G	16298C	16319A	73G	152C	263G	310C	489C
HZRm169	M12a1b	16172C	16183C	16193.1C	16209C	16223T	16234T	16290T	16519C	73G	125C	127C	128T	146C	195C	263G	309.1C	315.1C	489C
HZRm170	D4e5b	16223T	16274A	16278T	16291T	16362C	16519C	73G	263G	309.1C	315.1C	489C
HZRm171	A+152+16362	16223T	16290T	16319A	16355T	16362C	16519C	73G	152C	235G	263G	315.1C	523d	524d
HZRm172	M9	16223T	16362C	16519C	73G	263G	309.1C	315.1C	489C
HZRm173	HV1b+152	16067T	16311C	152C	263G	315.1C
HZRm174	HV6	16172C	16311C	16362C	263G	309.1C	315.1C
HZRm175	M9a1a1c	16223T	16234T	16291T	16316G	16362C	73G	153G	263G	309.1C	315.1C	489C
HZRm176	K2a5	16224C	16311C	16519C	73G	146C	152C	195C	263G	309.1C	315.1C	324T	430C
HZRm177	D4b1	16223T	16319A	16362C	73G	263G	309.1C	315.1C	489C	523d	524d
HZRm178	U7a	16274A	16309G	16318T	16519C	73G	151T	152C	263G	315.1C	523d	524d
HZRm179	HV6	16172C	16311C	263G	309.1C	315.1C
HZRm180	T2c1a	16126C	16292T	16294T	16309G	16320T	16519C	73G	263G	315.1C	573.1C	573.2C	573.3C
HZRm181	U5a2a	16114A	16192T	16256T	16270T	16294T	16526A	73G	234G	263G	309.1C	315.1C
HZRm182	C4a1a+195	16093C	16129A	16179T	16223T	16298C	16327T	16519C	73G	195C	249d	263G	309.1C	309.2C	315.1C	489C
HZRm183	U2e1h	16051G	16129C	16183C	16193.1C	16362C	73G	217C	228A	263G	310C	340T	508G
HZRm184	H2a1f	16086C	16189C	16193T	16193.1C	16354T	263G	309.1C	315.1C
HZRm185	M9	16223T	16362C	73G	263G	315.1C	489C
HZRm186	T2c1a	16126C	16292T	16294T	16519C	73G	263G	309.1C	315.1C	573.1C	573.2C	573.3C	573.4C	573.5C	573.6C
HZRm187	N1b1	16145A	16176G	16223T	16390A	16519C	73G	152C	263G	315.1C	523d	524d
HZRm188	HV6	16172C	16311C	263G	309.1C	309.2C	315.1C
HZRm189	H1b	16242T	16356C	263G	309.1C	309.2C	315.1C
HZRm190	C4a1a+195	16129A	16223T	16298C	16327T	16519C	73G	150T	195C	249d	263G	315.1C	333C	489C
HZRm191	U1a1a1	16182C	16183C	16189C	16249C	73G	263G	285T	309.1C	309.2C	315.1C	385G	523d	524d	573.1C	573.2C	573.3C	573.4C	573.5C
HZRm192	A+152+16362	16223T	16290T	16319A	16355T	16362C	16519C	73G	152C	235G	263G	315.1C	523d	524d
HZRm193	N1b1	16145A	16176G	16223T	16390A	16519C	73G	152C	263G	309.1C	315.1C
HZRm194	C4a1a+195	16129A	16223T	16298C	16327T	16519C	73G	150T	195C	249d	263G	315.1C	333C	489C
HZRm195	H1e1a4	16243C	16311C	16519C	204C	263G	309.1C	315.1C
HZRm196	F1b1+@152	16183C	16189C	16193.1C	16215G	16232A	16249C	16304C	16311C	16519C	73G	249d	263G	315.1C	523d	524d

Table S2. mtDNA haplogroup frequencies of Hazara in Pakistan

	Number	Percent	Haplogroup	Number	Percent (%)
A	**18**	**5.6**	A+152+16362	10	3.1
			A+152+16362+16189	2	0.6
			A15	2	0.6
			A16	3	0.9
			A23	1	0.3
B	**11**	**3.4**	B4b1a+207	1	0.3
			B4b1a3a	3	0.9
			B4c1b+16335	3	0.9
			B4c1c1	1	0.3
			B5a	2	0.6
			B5b2c	1	0.3
C	**20**	**6.3**	C	3	0.9
			C4a1a+195	8	2.5
			C4a1a4a	2	0.6
			C4a2	2	0.6
			C4b8a	1	0.3
			C5c	4	1.3
D	**31**	**9.7**	D2b	1	0.3
			D4+195	2	0.6
			D4b1	2	0.6
			D4b1a1	1	0.3
			D4b2b	1	0.3
			D4c2a	3	0.9
			D4c2b	7	2.2
			D4e5a	1	0.3
			D4e5b	1	0.3
			D4g2	1	0.3
			D4h1	2	0.6
			D4i	2	0.6
			D4j+16286	1	0.3
			D4l2	1	0.3
			D4m2a	4	1.3
			D4p	1	0.3
F	**7**	**2.2**	F1a1	1	0.3
			F1a1a	1	0.3
			F1b1+@152	3	0.9
			F1b1b	1	0.3
			F1d	1	0.3

G	**6**	**1.9**	G2a+152	5	1.6
			G2a2a	1	0.3
H	**45**	**14.1**	H1+16189	1	0.3
			H10+16093	2	0.6
			H11a1	1	0.3
			H13a1b	1	0.3
			H13a1c	1	0.3
			H14a	4	1.3
			H14b2a	1	0.3
			H1b	2	0.6
			H1e1a4	8	2.5
			H1m1	1	0.3
			H1o	1	0.3
			H29	1	0.3
			H2a1	4	1.3
			H2a1f	1	0.3
			H2a2a	13	4.1
			H2a5b1	1	0.3
			H3b6	1	0.3
			H5	1	0.3
HV	**30**	**9.4**	HV14	2	0.6
			HV19	2	0.6
			HV1b+152	7	2.2
			HV2	5	1.6
			HV2a	5	1.6
			HV4b	1	0.3
			HV6	8	2.5
J	**14**	**4.4**	J1b1b	1	0.3
			J1b8	3	0.9
			J1c	2	0.6
			J1c2i	1	0.3
			J1d	3	0.9
			J1d5	2	0.6
			J2b1a	1	0.3
			JT	1	0.3
	7	**2.2**	K1a4c1	3	0.9
			K1b1a1+199	1	0.3
			K2a5	3	0.9
K	**47**	**14.7**	M10a1a1b	1	0.3
			M12a1b	2	0.6
			M18a	1	0.3
			M2a1a	1	0.3

			M2a2	1	0.3
			M30	1	0.3
			M30b	1	0.3
			M30c1	1	0.3
			M30e	1	0.3
			M3a1+204	1	0.3
			M43a1	4	1.3
			M54	3	0.9
			M5a2a1a	3	0.9
			M65a+@16311	4	1.3
			M7b1a1	1	0.3
			M8a2+152	2	0.6
			M8a2a	2	0.6
			M9	4	1.3
			M9a1a1a	13	4.1
N	**4**	**1.3**	N1a1a1a1a	1	0.3
			N1b1	2	0.6
			N9a1	1	0.3
R0	**6**	**1.9**	R0a+60.1T	3	0.9
			R0a1a	3	0.9
R0	**8**	**2.5**	R2	3	0.9
			R30b2	2	0.6
			R7	1	0.3
			R8a1a1d	1	0.3
			R8a1a3	1	0.3
T	**10**	**3.1**	T	2	0.6
			T1a1'3	2	0.6
			T2a1b	1	0.3
			T2b34	1	0.3
			T2c1a	4	1.3
U	**46**	**14.4**	U1a	3	0.9
			U1a1a	3	0.9
			U1a1a1	3	0.9
			U2b	1	0.3
			U2b1	1	0.3
			U2b2	1	0.3
			U2c'd	3	0.9
			U2e1h	1	0.3
			U2e2	2	0.6
			U4'9	2	0.6
			U4b3	1	0.3
			U5a1	1	0.3

			U5a1+@16192	3	0.9
			U5a1a1+152	1	0.3
			U5a2a	10	3.1
			U7	1	0.3
			U7a	5	1.6
			U8b1b	2	0.6
			U9a1	2	0.6
W	**4**	**1.3**	W+194	1	0.3
			W3a1+199	1	0.3
			W6	2	0.6
X	**3**	**0.9**	X2	3	0.9
Y	**1**	**0.3**	Y1a+16189	1	0.3
Z	**1**	**0.3**	Z+152	1	0.3

@ corresponds to mutations that are reversions to an ancestral state (back mutations) and are indicated as @ in HaploGrep output

Table S3: Analysis of molecular variance for 4 Pakistani populations based on mtDNA complete control region sequences

Source of variation	d.f	Sum of Squares	Variance components	Percentage of variation
Among population	3	91.189	0.10937 Va	2.05
Within population	961	5021.488	5.22527 Vb	97.95
Total	964	5112.677	5.33464	

Data

Sources:

Populations	References	No. of samples
Kashmiri	[9]	317
Pathan	[25]	230
Makrani	[26]	99
Hazara	Present Study	319

Table S4: Analysis of molecular variance for global populations based on mtDNA complete control region sequence

Source of variation	d.f	Sum of Squares	Variance components	Percentage of variation
Among population	26	1579.624	0.22973 Va	4.17
Within population	6531	34490.714	5.28108 Vb	95.83
Total	6557	36070.338	5.51081	

Data Sources:

Populations	References	No. of samples	Populations	References	No. of samples
Hazara	Present study	319	Chinese	[28]	377
Kashmiri	[9]	317	Thai	[29]	190
Pathan	[25]	230	Vietnamese	[30]	187
Makrani	[26]	99	Laos	[31]	214
Uzbek	[27]	328	Kuwaiti	[32]	381
Turkmen	[27]	249	Iraqi	[33]	182
Tajik	[27]	244	Egyptian	[34]	277
Russian	[27]	151	Dubai	[35]	249
Kyrgyz	[27]	249	Ashkenazi Jews	[36]	173
Kazakh	[27]	256	Roma	[37]	205
Afghan	[27]	98	Greeks	[38]	319
Szekely	[37]	182	Iranian	[40]	352
Moroccan	[39]	509			

Table S5. Pairwise differences among twenty regional populations based on mtDNA complete control region sequences as calculated using Arlequin v3.5

Population Pairwise F_{ST}	Afghan [7]	Baluch *	Brahui #	Hazara (Current Study)	Iranian [40]	Kazakh [27]	Kyrgyz [27]	Kashmiri [9]	Kazakh [41]	Makrani [26]	Mongol [44]	Pashtun [27]	Pathan [25]	Russian [27]	Tajik [27]	Turkmen [27]	Tajik [7]	Turkmen [7]	Uzbek [7]	Uzbek [27]
Population Sample Size	90	25	25	319	352	256	249	317	160	99	201	98	230	151	244	249	146	75	127	328
Afghan [7]	0																			
Baluch *	0.05343	0																		
Brahui #	0.03769	0.00431	0																	
Hazara (Current Study)	0.03777	0.02576	0.01563	0																
Iranian [40]	0.03302	0.01768	-0.00245	0.02428	0															
Kazakh [27]	0.05245	0.03807	0.02369	0.00433	0.03878	0														
Kyrgyz [27]	0.06399	0.03985	0.02704	0.00916	0.04492	0.00378	0													
Kashmiri [9]	0.04284	0.01283	-0.00033	0.01311	0.01655	0.0226	0.02362	0												
Kazakh [41]	0.08377	0.04888	0.04668	0.01981	0.06454	0.01091	0.01255	0.03793	0											
Makrani [26]	0.05363	0.03813	0.01469	0.03616	0.02858	0.04398	0.0505	0.02951	0.06937	0										
Mongol [42]	0.07762	0.06808	0.05135	0.02103	0.07311	0.01025	0.00691	0.04315	0.02074	0.06639	0									
Pashtun [7]	0.0298	0.00598	0.0013	0.00847	0.008	0.02034	0.02654	0.00586	0.04003	0.027	0.04544	0								
Pathan [25]	0.03443	0.00758	-0.00584	0.01609	0.01267	0.02757	0.03211	0.00276	0.04218	0.02886	0.05377	0.00436	0							
Russian [27]	0.04289	0.0439	0.02918	0.03437	0.01028	0.05446	0.06912	0.03858	0.08801	0.05984	0.09745	0.01866	0.03027	0						
Tajik [27]	0.04991	0.05372	0.02468	0.0209	0.0319	0.02618	0.02626	0.03233	0.05168	0.04076	0.03372	0.02437	0.03622	0.0475	0					
Turkmen [27]	0.04205	0.02364	0.00978	0.00706	0.01696	0.00965	0.01269	0.01342	0.02881	0.03561	0.02966	0.00753	0.01617	0.03461	0.02	0				
Tajik [7]	0.02859	0.00908	0.00027	0.00588	0.00787	0.01708	0.02291	0.00572	0.03831	0.02784	0.04176	-0.00153	0.00515	0.01514	0.02147	0.00719	0			
Turkmen [7]	0.0279	0.00908	0.00933	0.00384	0.01591	0.01137	0.01558	0.00816	0.02467	0.0333	0.0328	-0.00167	0.0065	0.02756	0.0225	0.00261	-0.00022	0		
Uzbek [7]	0.03278	0.01209	-0.00301	0.00526	0.00846	0.01219	0.01651	0.0035	0.03207	0.02298	0.03505	0.00084	0.00386	0.02636	0.02324	0.00357	0.00025	0.00002	0	
Uzbek [27]	0.03629	0.02482	0.00559	0.00425	0.01863	0.00539	0.00814	0.01158	0.01922	0.03273	0.02139	0.00876	0.01221	0.03301	0.01717	0.00401	0.00638	0.00336	0.00293	0

F_{ST} p-value	Afghan [7]	Baluch *	Brahui #	Hazara (Current Study)	Iranian [40]	Kazakh [27]	Kyrgyz [27]	Kashmiri [9]	Kazakh [41]	Makrani [26]	Mongol [42]	Pashtun [27]	Pathan [25]	Russian [27]	Tajik [27]	Turkmen [27]	Tajik [7]	Turkmen [7]	Uzbek [7]	Uzbek [27]
Afghan [7]	*																			
Baluch *	0.00000+-0.0000	*																		
Brahui #	0.00781+-0.0024	0.34180+-0.0169	*																	
Hazara (Current Study)	0.00000+-0.0000	0.00098+-0.0010	0.01855+-0.0046	*																
Iranian [40]	0.00000+-0.0000	0.01172+-0.0033	0.61230+-0.0154	0.00000+-0.0000	*															
Kazakh [27]	0.00000+-0.0000	0.00098+-0.0010	0.00684+-0.0027	0.00098+-0.0010	0.00000+-0.0000	*														
Kyrgyz [27]	0.00000+-0.0000	0.00000+-0.0000	0.00586+-0.0022	0.00000+-0.0000	0.00000+-0.0000	0.00977+-0.0026	*													
Kashmiri [9]	0.00000+-0.0000	0.03027+-0.0053	0.46484+-0.0140	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	*												
Kazakh [41]	0.00000+-0.0000	0.00000+-0.0000	0.00098+-0.0010	0.00000+-0.0000	0.00000+-0.0000	0.00098+-0.0010	0.00000+-0.0000	0.00000+-0.0000	*											
Makrani [26]	0.00000+-0.0000	0.00391+-0.0019	0.07227+-0.0064	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	*										
Mongol [42]	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00293+-0.0016	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	*									
Pashtun [7]	0.00000+-0.0000	0.15527+-0.0103	0.36230+-0.0128	0.00098+-0.0010	0.00488+-0.0020	0.00000+-0.0000	0.00000+-0.0000	0.01172+-0.0036	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	*								
Pathan [25]	0.00000+-0.0000	0.11816+-0.0091	0.81836+-0.0103	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.01562+-0.0037	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.03027+-0.0051	*							
Russian [27]	0.00000+-0.0000	0.00000+-0.0000	0.00293+-0.0016	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	*						
Tajik [27]	0.00000+-0.0000	0.00000+-0.0000	0.00586+-0.0022	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	*					
Turkmen [27]	0.00000+-0.0000	0.00977+-0.0026	0.09766+-0.0075	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00781+-0.0028	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	*				
Tajik [7]	0.00000+-0.0000	0.08398+-0.0075	0.43652+-0.0128	0.00098+-0.0010	0.00098+-0.0010	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.72363+-0.0138	0.00098+-0.0010	0.00000+-0.0000	0.00000+-0.0000	0.00098+-0.0010	*			
Turkmen [7]	0.00000+-0.0000	0.12402+-0.0086	0.10742+-0.0088	0.06641+-0.0075	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00293+-0.0016	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.68262+-0.0166	0.01660+-0.0037	0.00000+-0.0000	0.00000+-0.0000	0.11621+-0.0107	0.48340+-0.0161	*		
Uzbek [7]	0.00000+-0.0000	0.03906+-0.0058	0.64844+-0.0139	0.00391+-0.0019	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.02246+-0.0040	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.33203+-0.0149	0.02051+-0.0038	0.00000+-0.0000	0.00000+-0.0000	0.05469+-0.0053	0.38672+-0.0148	0.43164+-0.0156	*	
Uzbek [27]	0.00000+-0.0000	0.00195+-0.0014	0.16895+-0.0100	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00000+-0.0000	0.00293+-0.0016	0.00098+-0.0010	0.07617+-0.0100	0.02734+-0.0049	*

* GenBank Accession ‘EU565791-EU565815’

\# GenBank Accession ‘EU565766-EU565790’

Table S6. Comparison (based on sequence n.p 16024-16569) of Haplotypes from Hazaras with Haplotypes from Central Asian Populations

Haplotype	Hazara (Present Study)	Baluch*	Hazara [7]	Kazakh[27]	Kyrgyz[27]	Kashmiri[9]	Kazakh[41]	Mongol[42]	Pashtun[7]	Pathan[25]	Tajik[27]	Turkmen[27]	Tajik[7]	Turkmen[7]	Uzbek[7]	Uzbek[27]
Hap_27	14	0	3	6	9	17	11	1	4	11	2	23	10	6	4	11
Hap_134	10	0	0	2	2	0	0	0	0	0	0	0	3	0	0	0
Hap_109	9	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Hap_122	9	0	0	2	0	0	1	0	0	0	0	0	0	0	0	1
Hap_120	8	0	0	7	1	0	2	3	0	0	4	3	0	0	1	2
Hap_140	7	0	0	1	0	1	0	0	0	0	0	0	3	0	0	2
Hap_202	6	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
Hap_36	5	0	1	7	16	0	9	9	1	5	0	6	0	1	1	8
Hap_31	4	0	0	1	3	0	0	1	1	0	6	0	1	0	0	2
Hap_75	4	0	2	8	6	0	2	0	0	0	0	2	1	3	2	5
Hap_94	4	0	1	1	0	0	0	0	0	0	5	2	0	0	0	2
Hap_157	4	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Hap_182	4	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Hap_15	3	0	0	2	1	0	2	0	0	3	0	1	2	2	0	6
Hap_77	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Hap_91	3	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0
Hap_105	3	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
Hap_123	3	0	0	0	3	0	0	0	0	0	0	2	0	0	0	1
Hap_124	3	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1
Hap_131	3	0	0	0	0	2	0	0	0	1	0	5	0	0	0	0
Hap_136	3	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0
Hap_153	3	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0
Hap_16	2	0	3	1	1	0	0	0	2	1	0	0	0	0	0	2
Hap_46	2	3	0	0	0	0	0	0	2	2	0	2	0	2	1	1
Hap_82	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Hap_111	2	0	1	1	0	0	0	0	0	0	0	0	0	0	0	3
Hap_112	2	0	1	1	1	0	1	2	0	0	0	1	0	0	0	0
Hap_118	2	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
Hap_145	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Hap_147	2	0	0	0	0	1	0	0	0	2	0	0	0	0	1	2
Hap_148	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hap_154	2	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0
Hap_159	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Hap_166	2	0	0	0	0	1	0	0	0	0	4	0	0	0	0	0
Hap_172	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Hap_179	2	0	0	0	0	0	0	0	1	1	0	2	0	0	0	0
Hap_186	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Hap_209	2	0	0	1	0	0	2	1	0	0	0	0	0	0	0	1
Hap_210	2	0	0	4	7	0	4	6	1	0	1	3	0	0	1	2
Hap_222	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Hap_5	1	2	0	0	1	7	0	0	2	3	0	0	1	0	1	1
Hap_14	1	0	0	0	1	0	1	1	0	0	15	0	0	0	1	2
Hap_18	1	0	0	1	0	0	0	3	0	1	0	1	5	0	1	0
Hap_19	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hap_28	1	0	0	1	1	2	2	0	0	0	0	0	0	0	0	0
Hap_40	1	0	0	2	5	3	2	1	0	8	0	3	2	1	3	6
Hap_44	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hap_72	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Hap_73	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
Hap_74	1	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0
Hap_79	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Hap_93	1	0	1	0	0	0	0	1	0	0	0	0	0	0	1	1
Hap_96	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1
Hap_98	1	0	1	0	4	2	1	0	0	1	0	0	0	1	3	1
Hap_103	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Hap_108	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Hap_121	1	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0
Hap_130	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Hap_132	1	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0
Hap_137	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
Hap_138	1	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0
Hap_161	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Hap_162	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
Hap_165	1	0	0	1	0	0	0	0	0	1	0	5	2	0	1	0
Hap_167	1	0	0	1	0	2	1	0	0	2	14	0	4	2	0	1
Hap_170	1	0	0	0	0	0	0	0	0	0	0	1	0	2	1	1
Hap_175	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Hap_187	1	0	0	0	0	0	1	0	1	0	0	0	1	1	0	0
Hap_192	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
Hap_193	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Hap_196	1	0	0	0	1	0	2	0	0	0	0	0	0	0	0	1
Hap_199	1	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0
Hap_200	1	0	0	1	0	1	2	0	0	2	0	1	0	0	4	3
Hap_206	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
Hap_213	1	0	0	0	2	0	1	0	0	0	0	4	0	1	0	0
Hap_215	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
Hap_219	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
Hap_223	1	0	0	0	1	1	0	0	0	0	0	0	3	0	0	0
Hap_227	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
Hap_229	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Hap_233	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Hap_234	1	0	0	0	0	2	0	0	0	0	0	0	0	0	2	0
Hap_235	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Hap_239	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Hap_242	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1
Population Sample Size	319	25	78	256	249	317	151	196	90	230	244	249	146	75	127	328

* GenBank Accession ‘EU565791-EU565815’