





Historic migration to South Asia in the last two millennia: A case of Jewish and Parsi populations

AJAI K PATHAK^{1,2} , ANSHIKA SRIVASTAVA³, PRAJVAL PRATAP SINGH³, DEBASHRUTI DAS³,
AUDDITIYA BANDOPADHYAY³, POOJA SINGH³ and GYANESHWER CHAUBEY^{3*} 

¹*Estonian Biocentre, Institute of Genomics, University of Tartu, Riia 23b, Tartu 51010, Estonia*

²*Department of Evolutionary Biology, Institute of Cell and Molecular Biology,
University of Tartu, Tartu 51010, Estonia*

³*Cytogenetics Laboratory, Department of Zoology, Banaras Hindu University, Varanasi 221005, India*

*Corresponding author (Email, gyaneshwer.chaubey@bhu.ac.in)

The South Asian populations have a mosaic of ancestries likely due to the interactions of long-term populations of the landmass and those of East and West Eurasia. Apart from prehistoric dispersals, there are some known population movements to India. In this study, we focussed on the migration of Jewish and Parsi populations on temporal and spatial scales. The existence of Jewish and Parsi communities in India are recorded since ancient times. However, due to the lack of high-resolution genetic data, their origin and affiliation with other Indian and non-Indian populations remains shrouded in legends. Earlier genetic studies on populations of Indian Jews have found evidence for a minor shared ancestry of Indian Jews with Middle Eastern (Jews) populations, whereas for Parsis, the Iranian link was proposed. Recently, in our high-resolution study, we were able to quantify the admixture dynamics of these groups, which has suggested a male-biased admixture. Here, we added the newly available ancient samples and revisited the interplay of genes and cultures. Thus, in this study we reconstructed a broad genetic profile of Indian Jews and Parsis to paint a fine-grained picture of these ethnic groups.

Keywords. Autosomes; cultural assimilation; Kerala Jews; Mumbai Jews; Parsi; Zoroastrian

Recent genetic studies on South Asia have highlighted several intriguing aspects, one among which was the large number of endogamous populations who follow various levels of marriage restrictions (Chaubey *et al.* 2007; Basu *et al.* 2016; Nakatsuka *et al.* 2017; Narasimhan *et al.* 2018). Since the practice is several millennium older, we can assume that it has strongly affected the admixture and assimilation dynamics of the populations who have migrated from outside (Behar *et al.* 2010; Shah *et al.* 2011; Chaubey *et al.* 2016, 2017; Lopez *et al.* 2017). Here, we undertook a study to reconstruct the admixture and assimilation of populations (Jewish and Parsi) who have a known record of migration to the Indian subcontinent (Chaubey *et al.* 2016, 2017).

1. Cultural and genetic assimilation of Jews in India

The Jewish diaspora is one of the major unique ethnic group of the world (Ehrlich 2009). Several anthropological and genetic studies have been conducted on them (Katz 2000; Behar *et al.* 2006, 2010; Moorjani *et al.* 2011); however, Indian Jews are among the least studied ethnic groups

among all the Jewish diaspora (Slapak 1995; Katz 2000; Israel 2002). In India, three Jewish groups are well known: in southern India, the Jews of Cochin in Kerala; in the western part of India, the Bene Israel in Mumbai, and in eastern India, Kolkata, the Baghdadi Jews (Katz 2000). There are several legendary stories about their migrations to India; nonetheless, owing to the lack of written records and inscriptions, the origin and migrations of Indian Jews remain shrouded in legends (Atzmon *et al.* 2010; Fernandes 2011).

Unlike other Jewish communities worldwide, an extensive study on Indian Jews was not done until recently (Chaubey *et al.* 2016; Waldman *et al.* 2016). In earlier studies they have been mainly considered as a Jewish outlier (Behar *et al.* 2006, 2010). Their maternal and paternal ancestries were contrasting: The maternal ancestry was much more related with the local South Asian gene pool, whereas paternal ancestry was mixed with the visible inputs from Levant (Metspalu *et al.* 2004; Behar *et al.* 2008, 2010). Indian Jews have hitherto been one of the most ancient offshoots of the Jewish diaspora. Their complex history, the presence of multiple subgroups in various regions of India and the persistence of an ancestral social system embellish them with several unique characteristics (Roland 1998).

Recent studies have testified their unique genetic profile where Indian Jewish groups were much more closely related with the local South Asian groups (Chaubey *et al.* 2016; Waldman *et al.* 2016). We used principal component analysis (PCA) to capture the genetic variation of Indian Jews along the two axes covering the Eurasian landscape (figure 1). Indian Jews were clustered over the South Asian Indo-European-Dravidian cline. On a closer look, we could see that their placement along the South Asian cline was attracted towards West Eurasia. This attraction was not observed among their local neighbouring populations (figure 1). The admixture plot also showed an additional substantial amount of Middle-Eastern-related ancestry, which is otherwise negligible among their local neighbouring populations (figure 1). This ancestry shows a gradient among Jewish populations, where Mumbai Jews carried higher levels of Middle-Eastern-related ancestry compared to the Kerala Jews. This can be explained either through the isolation-by-distance model or by the relatively later migration of Mumbai Jews (Chaubey *et al.* 2016). The Indian ancestry was overwhelmingly dominant among Indian Jews and showed a range of Middle-Eastern-related component from 3% to 20% (Chaubey *et al.* 2016).

To get a clearer perspective of a fine-scaled population structure, we used ChromoPainter and performed

FineStructure analysis (Lawson *et al.* 2012). All the Indian Jews form their own sub-cluster, nested into the South Asian large cluster, whereas the other Jews form a cluster with the Middle Eastern populations (figure 2). It is also evident that the Indian Jews share higher number of chunks with the Middle Eastern and Jewish populations compared to their neighbouring Indian populations. Consistent with the previous analysis, the Mumbai Jews carry more Middle-Eastern-related chunks compared to the Kerala Jews (figure 2). The time of admixture of Indian and Middle Eastern ancestries for various Indian Jews fall in the time frame between 1100 and 1600 years (Chaubey *et al.* 2016).

For the maternal ancestry, the Indian Jews share 4.6% East Eurasian and 21.1% West Eurasian lineages, whereas the paternal ancestries of Indian Jews were also united with frequent Middle-East-specific haplogroups (E, G, J(\times J2) and I) (Chaubey *et al.* 2016).

2. Parsi: ‘Like sugar in milk’

Similar to the Jews, Parsis were also migrants from West Eurasia and travelled to South Asia in the overlapping time frame. Parsis are one of the smallest ethnoreligious

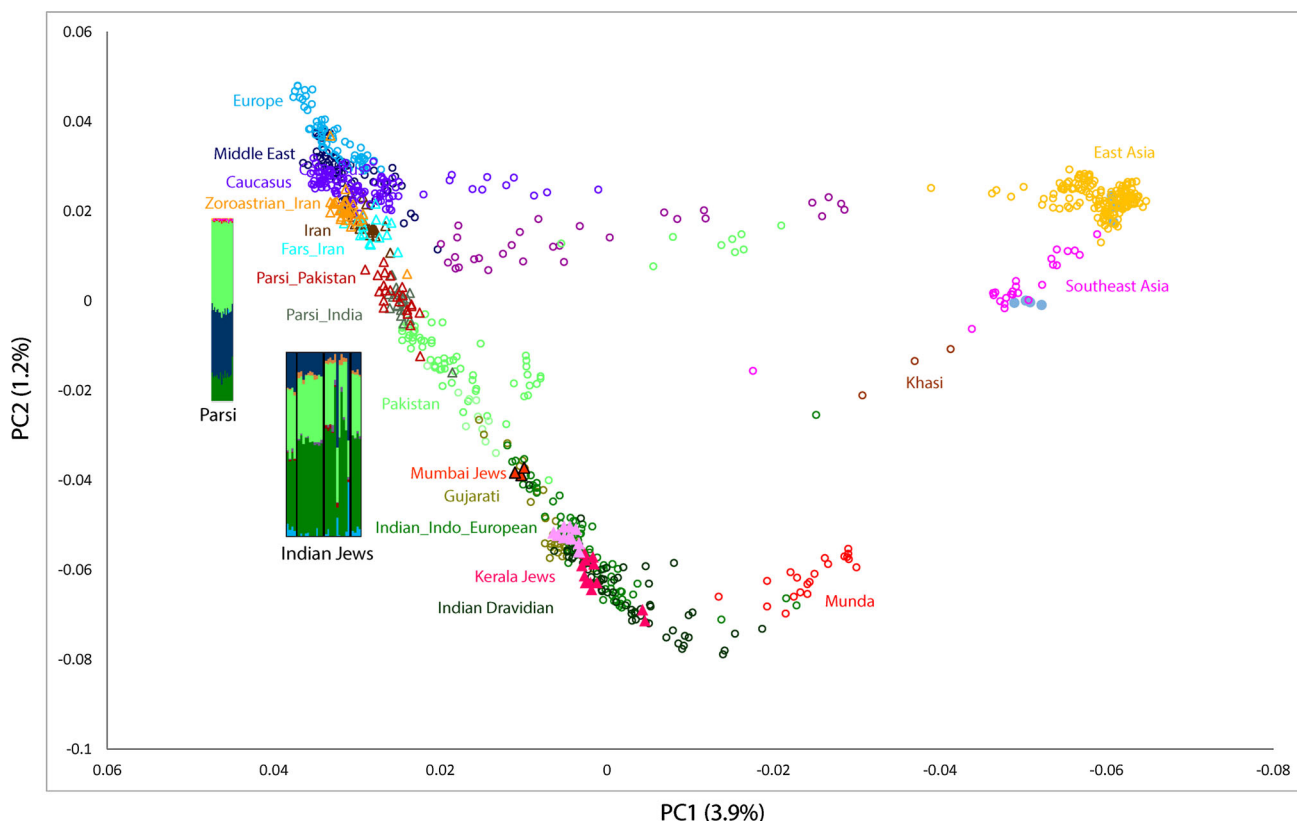


Figure 1. PCA of Eurasian populations showing the placement of Indian Jewish and Parsi groups along with the South Asian cline. It is noteworthy that they were attracted towards West Eurasian populations, suggesting a higher level of ancestry related to West Eurasians complying with their neighbouring local South Asian populations. The ADMIXTURE proportions of both of the studied populations were also shown. The colour codes are consistent with our previous studies (Chaubey *et al.* 2016, 2017).

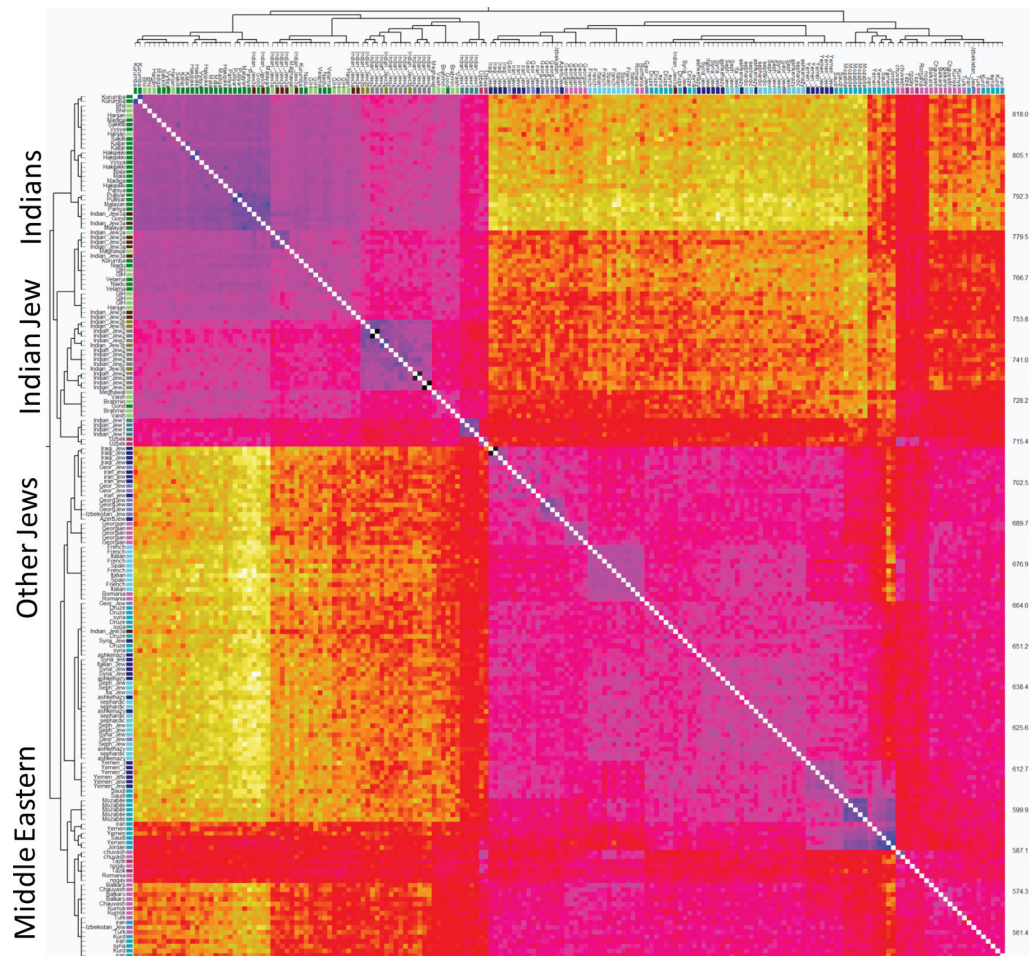


Figure 2. FineStructure analysis showing clustering pattern as well as the chunk sharing of Indian Jews with the other Eurasian population.

communities of the world, with around 57,264 members as per the 2011 Indian census (down from 69,601 in 2001 and 114,000 in 1940). With the declining population of only around 57,000 individuals in India (mainly Mumbai city), the Parsi has been a remarkable exception in the general South Asian demographic scenario (Bachi Karkaria 2016). They trace their origin to Persia (present day Iran), from where they fled during the 8–10th century due to Islamic conquest (Darukhanawala 1963). A group of Zoroastrians first landed in Sanjan (Gujarat, India). They were granted asylum by Jadi Rana, a local king, with a promise to blend culturally with native Hindus ‘like sugar in milk’. Their fertility and mortality rates have steadily declined over the past century, thus making them vulnerable. To control this decline, the Indian Government has formulated ‘Jiyo Parsi’ (literally meaning long live Parsis) campaign in the year 2013. They make up less than 0.005% of India’s population but as described by a famous Indian writer, Amitav Ghosh, ‘they have essentially created modern India’ (Amitav Ghosh 2011).

Earlier studies of Parsi populations have relied mainly on small sample sizes and low-resolution genetic markers,

restraining the power of the analysis. Moreover, the major Parsi group (living in India) has been underrepresented, and likewise, no high-resolution autosomal evidence has been considered in this debate. Therefore, by using a combination of high-resolution uniparental and biparental genetic markers on modern Parsi samples as well as mtDNA markers on ancient Parsi samples, we specifically looked at the question of whether the current Parsi people living in India and Pakistan maintained any common genetic basis among them and also with their homeland (Iranian) population, or whether their genetic compositions had been affected by the neighbouring Indian and Pakistani populations, as we have seen for the Jewish populations. This study has also examined runs of homozygosity (RoH) to study consanguinity (Chaubey *et al.* 2017). Moreover, with a particular attempt to address the extent to which the current Parsi populations of India assimilated local South Asian females during their long formation history, this study has generated first successful ancient DNA data from ancient remains excavated in Sanjan (Gujarat, India) (Chaubey *et al.* 2017).

We have observed a distinct population structure of Parsis in comparison with their present neighbouring South Asian populations. Both Indian and Pakistani Parsi populations were genetically closest to each other, followed by the Iranians. The admixture with the local South Asian populations was minimal in comparison with other historical migrants to India (e.g. Jews (Chaubey *et al.* 2016) and Siddis (Shah *et al.* 2011)). The modern Iranian population exhibited marked difference from Parsis mainly in carrying additional European and significantly lower South-Asian-specific ancestries (figure 1). The haplotype-based analysis of all the Zoroastrian communities showed a complex history of the Parsi population in which the admixture with the local South Asians was limited in terms of amalgamation of cultural traits (figure 3). Further analyses have supported that this genomic difference was likely due to Islamic conquest of Persia. Besides the alleles sharing compassion of Parsis with modern populations, it is important to note that the Neolithic Iranian samples shared higher number of alleles with Parsis than modern day Iranians (Chaubey *et al.* 2017).

It was shown previously that long RoH segments characterize the case of consanguinity and population isolation, and consequently provide a distinctive record of the demographic history for a particular population. As expected, both Parsi populations made a ‘right shift’

compared to their putative parental and present neighbours, likely due to small population size and high levels of inbreeding. To have a better understanding of Parsi admixture and its temporal impact over the historical interpretations, this study for the first time quantitatively estimated a date for the admixture event. It was shown that a single wave of admixture coincides with the historical date of the arrival of Parsis to South Asia (Chaubey *et al.* 2017).

In addition, this was the first successful ancient DNA study from India which analysed the composition of mtDNA haplogroup of ancient remains excavated in Sanjan (Gujarat, India). Interestingly, 48% South Asian indigenous lineages among the ancient Parsi samples were observed, which is likely due to the assimilation of local females during the initial settlement. For maternal ancestry, the Parsi remains closer to the Indian and Pakistani clusters, whereas for paternal ancestry, they align in between Iranian and Pakistani populations. Such contrasting patterns of maternal and paternal ancestries supported a female-biased admixture from the South Asian populations to the Parsis.

In search of disease-related genotype, the *in silico* analysis showed very high frequency of an ancestral allele in Parsis, which is associated with decreased risk of active brucellosis among Iranians (Chaubey *et al.* 2017).

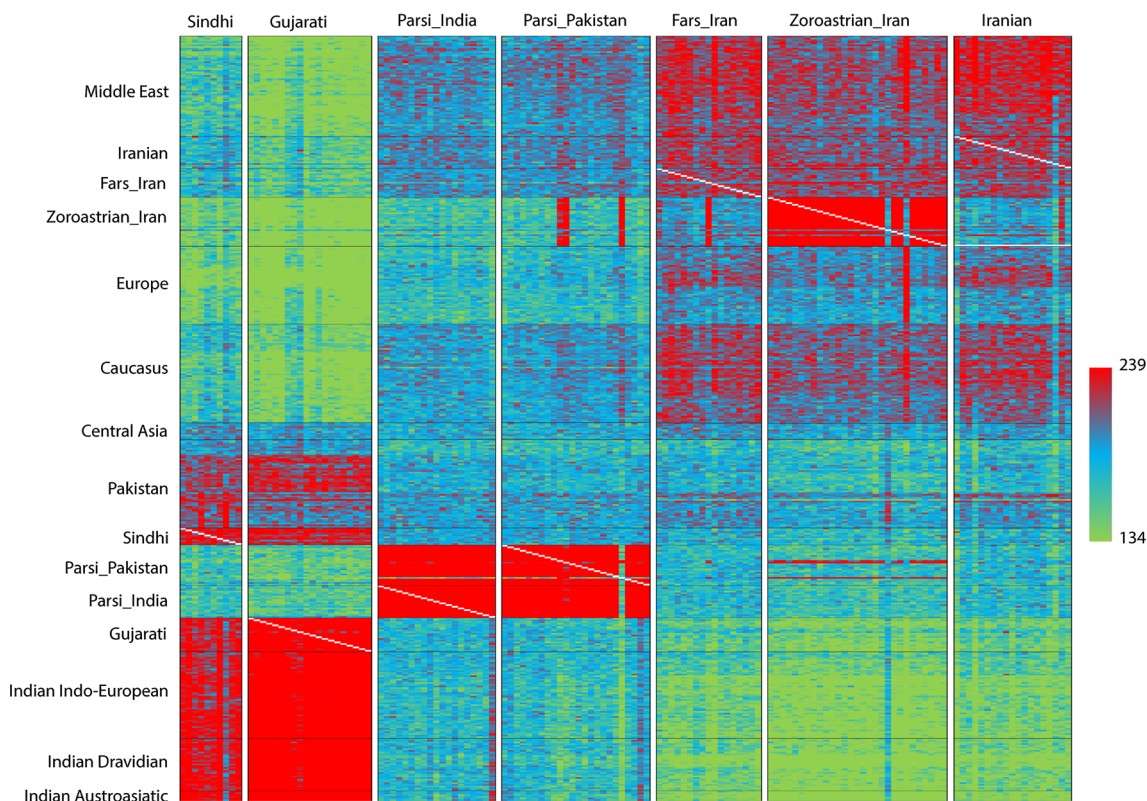


Figure 3. FineStructure analysis of the Parsis and Iranian Zoroastrians comparing the chunk shared with their neighbouring and putative parental populations.

3. Indian Jews and Parsis in the context of ancient individuals from West Eurasia and South Asia

With technological advancement, we now have data published from several ancient samples mainly from West Eurasia (Mathieson *et al.* 2015; Mittnik *et al.* 2018; Narasimhan *et al.* 2018). Although the climatic conditions in India is a big hurdle for obtaining good quality DNA (Narasimhan *et al.* 2018; Pathak *et al.* 2018), we hope that in the next five years further improvement in technology will

enable us to produce adequate amount of ancient samples from India to test various assumptions. Since both the populations migrated from West Eurasia, we could compare them with the published ancient samples (Mathieson *et al.* 2015; Mittnik *et al.* 2018; Narasimhan *et al.* 2018). Our analysis on the outgroup f_3 test suggested a higher level of allele sharing of Parsis with the ancient West Eurasian samples, which is consistent with modern data (figure 4). Among the Indian Jews groups, Steppe_EMBA, Armenian_EBA, and EHG were the top three populations who

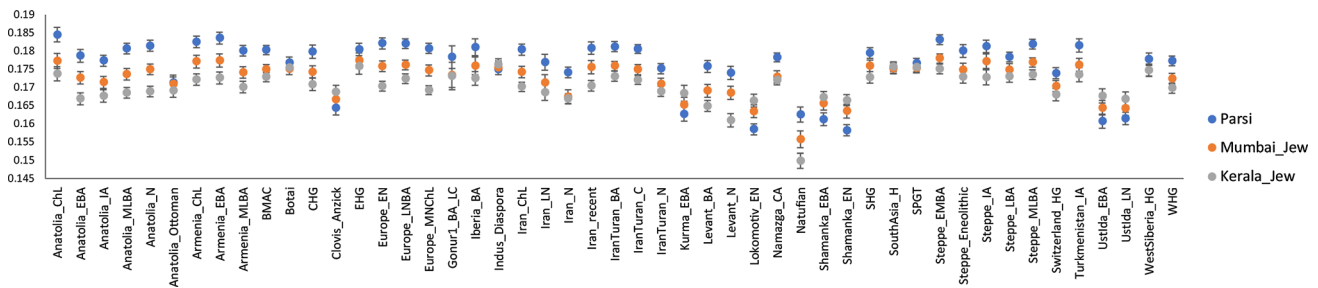


Figure 4. Outgroup f_3 statistics of Indian Jewish and Parsi populations showing the shared drift with the ancient samples.

Table 1. Top 10 related ancient samples to Parsis, Mumbai_Jews and Kerala_Jews populations with respect to Iran_Neolithic and Kurumba, respectively

Population A	Population B	Population C	Population D	f_4	Std. Err.	Z score	SNPs
Mumbai_Jews	Yorubas	Steppe_EMBA	Kurumba	0.001216	0.000302	4.033	264648
Mumbai_Jews	Yorubas	EHG	Kurumba	0.001052	0.000479	2.198	235189
Mumbai_Jews	Yorubas	Armenia_EBA	Kurumba	0.001047	0.000369	2.838	234382
Mumbai_Jews	Yorubas	Steppe_IA	Kurumba	0.001017	0.000445	2.285	206723
Mumbai_Jews	Yorubas	Armenia_ChL	Kurumba	0.000972	0.000344	2.825	239601
Mumbai_Jews	Yorubas	Steppe_MLBA	Kurumba	0.000865	0.000273	3.167	264837
Mumbai_Jews	Yorubas	Anatolia_ChL	Kurumba	0.000697	0.00038	1.833	137434
Mumbai_Jews	Yorubas	Europe_LNBA	Kurumba	0.000655	0.000295	2.22	264808
Mumbai_Jews	Yorubas	Turkmenistan_IA	Kurumba	0.000613	0.00033	1.857	250200
Mumbai_Jews	Yorubas	IranTuran_BA	Kurumba	0.000575	0.000337	1.704	249472
Kerala_Jews	Yorubas	Indus_Diaspora	Kurumba	-0.000244	0.000316	-0.772	199486
Kerala_Jews	Yorubas	EHG	Kurumba	-0.00027	0.000474	-0.571	235189
Kerala_Jews	Yorubas	SouthAsia_H	Kurumba	-0.000457	0.000247	-1.852	258336
Kerala_Jews	Yorubas	Gonur1_BA_LC	Kurumba	-0.000489	0.000785	-0.623	27787
Kerala_Jews	Yorubas	SPGT	Kurumba	-0.000511	0.000207	-2.468	260807
Kerala_Jews	Yorubas	Botai	Kurumba	-0.000585	0.000389	-1.504	264914
Kerala_Jews	Yorubas	Steppe_EMBA	Kurumba	-0.00067	0.000277	-2.418	264648
Kerala_Jews	Yorubas	WestSiberia_HG	Kurumba	-0.000704	0.000398	-1.768	201635
Kerala_Jews	Yorubas	Steppe_Eneolithic	Kurumba	-0.00102	0.000379	-2.692	175828
Kerala_Jews	Yorubas	Steppe_IA	Kurumba	-0.001157	0.000396	-2.919	206723
Parsis	Yorubas	Anatolia_ChL	Iran_N	0.002921	0.000565	5.171	123969
Parsis	Yorubas	Armenia_EBA	Iran_N	0.002677	0.000409	6.547	192146
Parsis	Yorubas	Steppe_EMBA	Iran_N	0.002585	0.000338	7.648	201909
Parsis	Yorubas	Armenia_ChL	Iran_N	0.002497	0.000388	6.428	194646
Parsis	Yorubas	Europe_EN	Iran_N	0.002394	0.000364	6.577	201912
Parsis	Yorubas	Europe_LNBA	Iran_N	0.002246	0.000345	6.509	201912
Parsis	Yorubas	Anatolia_N	Iran_N	0.002235	0.000362	6.179	201746
Parsis	Yorubas	Steppe_MLBA	Iran_N	0.002177	0.000339	6.419	201913
Parsis	Yorubas	Anatolia_MLBA	Iran_N	0.002121	0.000384	5.519	200514
Parsis	Yorubas	Iran_ChL	Iran_N	0.002087	0.000372	5.618	191819

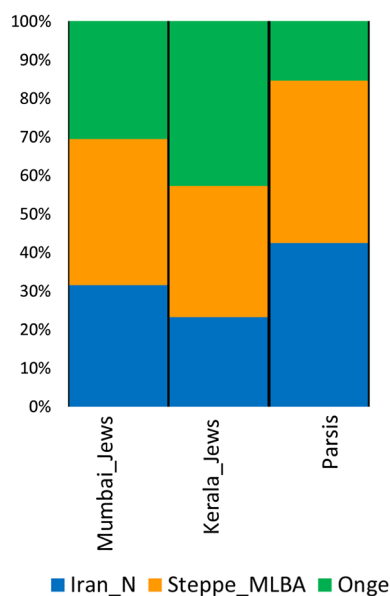


Figure 5. Plot of ancestry proportions modelled in the studied populations. The p -values and other parameters are shown in table 2.

shared the highest drift with the Mumbai Jews, whereas Indus_periphery, EHG (East European Hunter Gatherer), and South Asia_H were the top three groups who shared the highest drift with the Kerala Jews group (figure 4). For Parsis, the top three groups with the highest shared drift were Anatolian_ChI, Armenia_EBA, and Steppe_EMBA.

To obtain information about other ancient ancestries, we conducted four population tests fixing their most genetically closer populations (table 1). For the Mumbai Jews, Steppe_EMBA/Steppe_MLBA were significantly predominant (Z -score >3), whereas Armenia_EBA and EHG-related ancestry also were prevalent, and Z -scores were non-significant. In case of the Kerala Jews, Indus_Periphery and EHG-related ancestries were major; however, none of them were significant at the Z -score. For the Parsis, genetic affinity with Steppe_EMBA, Anatolia_ChI, and Armenia_EBA-related ancestries were highly common, supported with significant Z -score (table 1). We also modelled the Jewish and Parsi populations with the same ancient source groups characterized previously as a good fit to model South Asian populations in a distal (distant in time and space) model (Narasimhan *et al.* 2018; Pathak *et al.* 2018). Both the

Jews and Parsis were successfully modelled (p -value > 0.05). The Mumbai Jews were observed to have $\sim 70\%$ West-Eurasian-related (31.6% Iran_Neolithic and 37.8% Steppe_MLBA) and 30% Onge-related ancestries (figure 5, table 2). Kerala Jews harbour relatively lower $\sim 57\%$ West-Eurasian-related and $\sim 43\%$ Onge-related ancestries. Parsis showed $\sim 15\%$ of Onge-related ancestry (figure 5, table 2). Since the Onge represent only a part of South-Asian-related ancestry (Chaubey and Endicott 2013), the actual South Asian ancestry (as a proxy of Onge) present in these populations is likely to be higher than the estimated value.

We also plotted the shared drift values of the studied populations along the two axes representing Ust'-Ishim, and Iran_Neolithic ancestries, keeping in mind to place the studied populations along the East- and West-Eurasian-related ancestries (figure 6). South Asians formed a cline from Munda to Parsis with a decreasing Ust'-Ishim-related ancestries, and vice versa for the Iran_Neolithic ancestries. Compared to modern Iranians, Parsis showed elevated ancestries for both the sources, whereas both groups of Jews also were attracted towards higher Iran_Neolithic-related ancestries compared to their neighbours.

Overall, our pooled analysis of genetic variation among various groups of Indian Jewish and Parsi populations, involving autosomal and sex-linked markers, helped us to understand the demographic history of these migrant groups, which is biased towards male-mediated migration. Parsis were more restricted to mingle with the local populations, whereas the Indian Jews overwhelmingly carried the South Asian ancestry. The analyses of autosomal data revealed a high level of heterogeneity among the Indian Jews. Therefore, for the South Asian Jewish communities we can summarize that the migration of the Indian Jews from Middle East was followed by extensive admixture and assimilation with the local populations. For the Parsis, estimates of admixture based on the analyses of autosomal markers genotyped across the genome, markers from Y chromosome and mtDNA pointed to largely female-mediated episodic gene flow from the Indian populations ~ 1200 years ago followed by extensive isolation and inbreeding within the Parsi population. With detailed oral and written records, smaller population size and longer RoH, it can also be deduced that the number of founders for this population might have been limited. Moreover, high levels of endogamy have enriched the high occurrence of recessive diseases in these communities.

Table 2. Modelled ancestries of Jewish and Parsi populations

Target	p -Value	Mix coeff1	Mix coeff2	Mix coeff3	Std err1	Std err2	Std err3
		Iran_N	Steppe_MLBA	Onge	Iran_N	Steppe_MLBA	Onge
Mumbai_Jews	0.968977	0.316	0.378	0.306	0.062	0.055	0.034
Kerala_Jews	0.687954	0.233	0.340	0.427	0.060	0.054	0.031
Parsis	0.832763	0.424	0.422	0.154	0.057	0.049	0.029

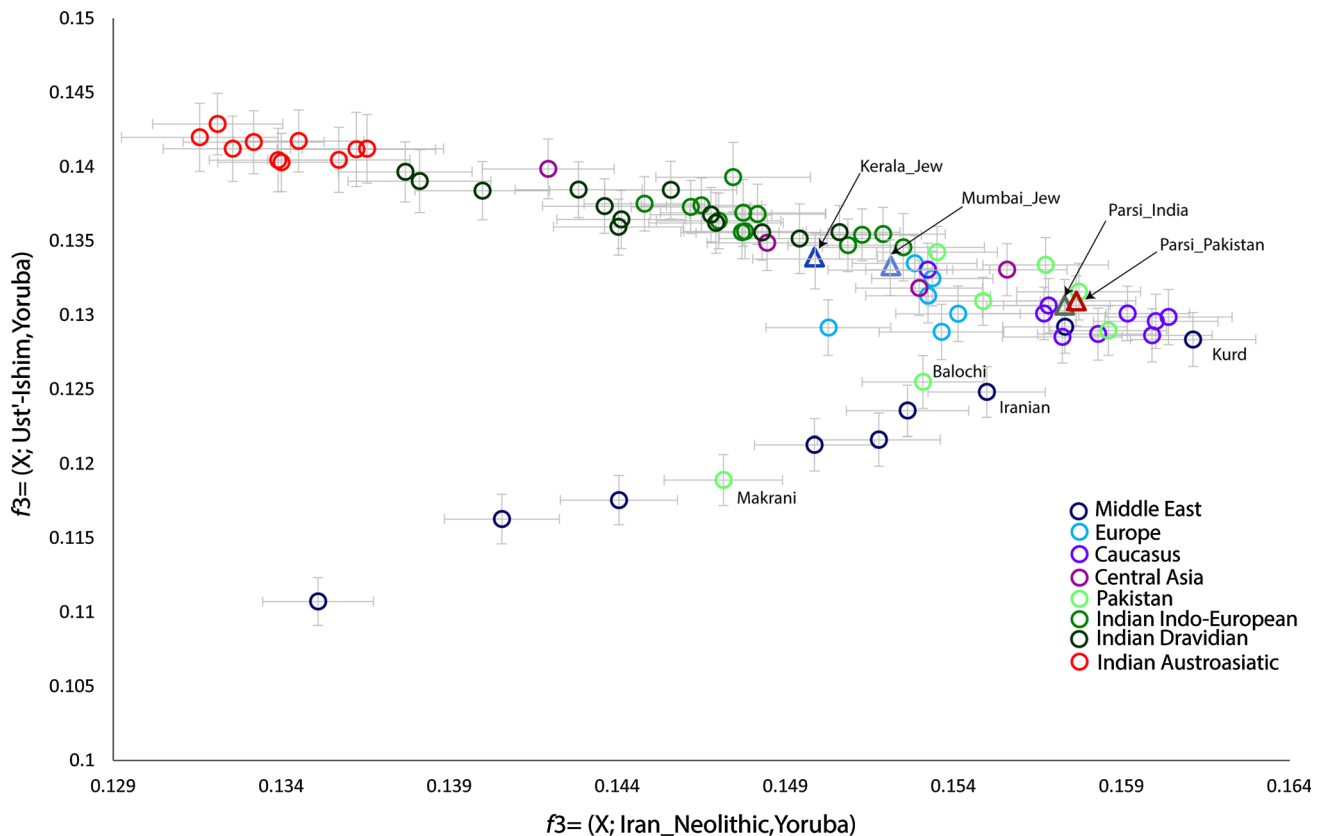


Figure 6. Plot of South Asian and other relevant neighbours along the axes of Ust'-Ishim vs Iranian_Neolithic ancestries.

Acknowledgements

This work is supported by the National Geographic explore grant HJ3-182R-18. AKP is supported by Estonian Institutional Research grants IUT24-1 and by the European Social Fund's Doctoral Studies and Internationalization Programme DoRa. AS and DD are supported by the UGC-RET doctoral fellowship and PPS is supported by the CSIR-JRF doctoral fellowship. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

- Amitav Ghosh 2011 Parsi community has created modern India. <https://parsi-khabar.net/india/parsi-community-has-created-modern-india-amitav-ghosh/3224/>
- Atzmon G, Hao L, Pe'er I, Velez C, Pearlman A, Palamara PF, Morrow B, Friedman E, Oddoux C, Burns E, *et al.* 2010 Abraham's children in the genome era: major Jewish diaspora populations comprise distinct genetic clusters with shared Middle Eastern ancestry. *Am. J. Hum. Genet.* **86** 850–859
- Bachi Karkaria 2016 Why is India's wealthy Parsi community vanishing? <http://www.bbc.com/news/world-asia-india-35219331>
- Basu A, Sarkar-Roy N and Majumder PP 2016 Genomic reconstruction of the history of extant populations of India reveals five distinct ancestral components and a complex structure. *Proc. Natl. Acad. Sci. USA* **113** 1594–1599
- Behar DM, Metspalu E, Kivisild T, Achilli A, Hadid Y, Tzur S, Pereira L, Amorim A, Quintana-Murci L, Majamaa K, *et al.* 2006 The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event. *Am. J. Hum. Genet.* **78** 487–497
- Behar DM, Metspalu E, Kivisild T, Rosset S, Tzur S, Hadid Y, Yudkovsky G, Rosengarten D, Pereira L, Amorim A, *et al.* 2008. Counting the founders: the matrilineal genetic ancestry of the Jewish diaspora. *PLoS ONE* **3** e2062
- Behar DM, Yunusbayev B, Metspalu M, Metspalu E, Rosset S, Parik J, Rootsi S, Chaubey G, Kutuev I, Yudkovsky G, *et al.* 2010 The genome-wide structure of the Jewish people. *Nature* **466** 238–242
- Chaubey G and Endicott P 2013 The Andaman islanders in a regional genetic context: reexamining the evidence for an early peopling of the archipelago from south Asia. *Hum. Biol.* **85** 153–172
- Chaubey G, Metspalu M, Kivisild T and Villems R 2007 Peopling of south Asia: investigating the caste-tribe continuum in India. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **29** 91–100
- Chaubey G, Singh M, Rai N, Kariappa M, Singh K, Singh A, Pratap Singh D, Tamang R, Selvi Rani D, Reddy AG, *et al.* 2016 Genetic affinities of the Jewish populations of India. *Sci. Rep.* **6** 19166
- Chaubey G, Ayub Q, Rai N, Prakash S, Mushrif-Tripathy V, Mezzavilla M, Pathak AK, Tamang R, Firasat S, Reidla M, *et al.* 2017 'Like sugar in milk': reconstructing the genetic history of the Parsi population. *Genome Biol.* **18** 110

- Darukhanawala HD 1963 *Parsi lustre on Indian soil* (G. Claridge, Bombay)
- Ehrlich MA 2009 Encyclopedia of the Jewish diaspora. ABC-CLIO. http://books.google.com/books?id=NoPZu79hqaEC&pg=PA1223&dq=Jews+of+India&ie=ISO-8859-1&cd=7&source=gbs_gdata
- Fernandes E 2011 *The last Jews of Kerala* (Granta Books) http://books.google.com/books?id=yJI32yLYykcC&pg=PT104&dq=Jews+of+India&ie=ISO-8859-1&cd=29&source=gbs_gdata
- Israel RR 2002 The Jews of India. http://books.google.com/books?id=3KVtAAAAAAAJ&q=Jews+of+India&dq=Jews+of+India&ie=ISO-8859-1&cd=5&source=gbs_gdata
- Katz N 2000 *Who are the Jews of India?* (University of California Press) http://books.google.com/books?id=ZWX6pF2PTJwC&printsec=frontcover&dq=Jews+of+India&ie=ISO-8859-1&cd=2&source=gbs_gdata
- Lawson DJ, Hellenthal G, Myers S and Falush D 2012 Inference of population structure using dense haplotype data. *PLoS Genet.* **8** e1002453
- Lopez S, Thomas MG, van Dorp L, Ansari-Pour N, Stewart S, Jones AL, Jelinek E, Chikhi L, Parfitt T, Bradman N, et al. 2017 The genetic legacy of Zoroastrianism in Iran and India: insights into population structure, gene flow and selection. *bioRxiv* [Internet]. <http://biorxiv.org/content/early/2017/05/05/128272.abstract>
- Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, et al. 2015 Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528** 499–503
- Metspalu M, Kivisild T, Metspalu E, Parik J, Hudjashov G, Kaldma K, Serk P, Karmin M, Behar DM, Gilbert MTP, et al. 2004 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** 26
- Mittnik A, Wang C-C, Pfrengle S, Daubaras M, Zariņa G, Hallgren F, Allmāe R, Khartanovich V, Moiseyev V, Törv M, et al. 2018 The genetic prehistory of the Baltic sea region. *Nat. Commun.* **9** 442
- Moorjani P, Patterson N, Hirschhorn JN, Keinan A, Hao L, Atzmon G, Burns E, Ostrer H, Price AL and Reich D 2011 The history of African gene flow into southern Europeans, Levantines, and Jews. *PLoS Genet.* **7** e1001373
- Nakatsuka N, Moorjani P, Rai N, Sarkar B, Tandon A, Patterson N, Bhavani GS, Girisha KM, Mustak MS and Srinivasan S 2017 The promise of discovering population-specific disease-associated genes in south Asia. *Nat. Genet.* **49** 1403
- Narasimhan VM, Patterson NJ, Moorjani P, Lazaridis I, Mark L, Mallick S, Rohland N, Bernardos R, Kim AM, Nakatsuka N, et al. 2018 The genomic formation of south and central Asia. <https://www.biorxiv.org/content/early/2018/03/31/292581>
- Pathak AK, Kadian A, Kushniarevich A, Montinaro F, Mondal M, Ongaro L, Singh M, Kumar P, Rai N, Parik J, et al. 2018 The genetic ancestry of modern Indus valley populations from northwest India. *Am. J. Hum. Genet.* **103** 918–929
- Roland JG 1998 *The Jewish communities of India* (Transaction Publishers) http://books.google.com/books?id=kHJccZ92IecC&printsec=frontcover&dq=The+Jewish+Communities+of+India:+identity+in+a+colonial+era&ie=ISO-8859-1&cd=1&source=gbs_gdata
- Shah AM, Tamang R, Moorjani P, Rani DS, Govindaraj P, Kulkarni G, Bhattacharya T, Mustak MS, Bhaskar LVKS, Reddy AG, et al. 2011 Indian Siddis: African descendants with Indian admixture. *Am. J. Hum. Genet.* **89** 154–161
- Slapak O 1995 The Jews of India. UPNE. http://books.google.com/books?id=qhKGPprbQaYC&printsec=frontcover&dq=Jews+of+India&ie=ISO-8859-1&cd=1&source=gbs_gdata
- Waldman YY, Biddanda A, Davidson NR, Billing-Ross P, Dubrovsky M, Campbell CL, Oddoux C, Friedman E, Atzmon G, Halperin E, et al. 2016 The genetics of Bene Israel from India reveals both substantial Jewish and Indian ancestry. *PLoS ONE* **11** e0152056