

Genetic affinities of the Jewish populations of India

Gyaneshwer Chaubey ^{1,2§}, Manvendra Singh ^{1§}, Niraj Rai ^{1§}, Mini Kariappa ^{1,3§}, Kamayani Singh ¹, Ashish Singh ¹, Deepankar Pratap Singh ¹, Rakesh Tamang ¹, Deepa Selvi Rani ¹, Alla G Reddy ¹, Vijay Kumar Singh ¹, Lalji Singh ^{1,4} and Kumarasamy Thangaraj ^{1*}

¹CSIR-Centre for Cellular and Molecular Biology, Hyderabad 500 007, India

²Department of Evolutionary Biology, Estonian Biocentre, Riia23b, Tartu, Estonia-51010

³ Department of Anatomy, Amala Institute of Medical Sciences, Thrissur-680555, India

⁴ Genome Foundation, Hyderabad 500007, India

[§] These authors contributed equally to this work

Extended data analysis

Autosomal analysis

First, we sought to investigate the extent of population structure and admixture among the Indian Jewish and other Eurasian populations embedded in their autosomal genomes. We used Affymatrix 6.0 genotyping chips on our collection of Indian Jewish (Cochin Jewish) sample and genotyped 10 individuals. These results were combined with the global data set of 1865 individuals generated with Affymatrix 6.0, Illumina 610K, 650K chips (For a full list of populations and sample sizes see Supplementary Table 1). After excluding SNPs unique to either of these platforms and SNPs from mtDNA, X and Y chromosomes the combined data set had data for 98189 SNPs that was used in subsequent analyses. We have excluded single outlier of Indian Jewish 1 and two outlier samples of Indian Jewish 3 for the population based F_{st} , f_3 , f_4 and Runs of Homozygosity (ROH) analysis. We thinned the dataset by removing one SNP of a pair in strong LD $r^2 > 0.4$ in a window of 200 SNPs (sliding the window by 25 SNPs at a time). Finally we were left with a data matrix of 75594 SNPs.

mtDNA and Y chromosome

The collection of our samples were from Indian Jewish 3 (Cochin Jewish). The Indian Jewish 1 and Indian Jewish 4 data on haploid DNA analysis were published elsewhere¹⁻⁴. For the Indian Jewish 2 the haploid DNA data was not available. For the mtDNA genotyping, we have sequenced first Hypervariable segment I (HVS-I) and diagnostic coding region mutations by utilizing the primers published in Reider et al. ⁵. Individuals with ambiguous haplogroup affiliation were subjected to further sequencing of informative coding region stretches. Similarly, for the Y chromosome genotyping we have used defining SNPs and the published primers to segregate samples in to the major subclades ^{6,7}. Sequencing of mtDNA and Y chromosome markers was carried out in ABI 3730 and 3730XL DNA analyzers. To minimize errors both strands were double-sequenced. Sequences were assembled, and edited using Autoassembler2.0. Mutations were scored

relative to the r-CRS and RSRS ^{8,9}. All the mutations were confirmed by manual checking of their electropherograms.

Extended history of Indian Jewish

India is inhabited by several indigenous populations and there too have been several historical migrations ^{10,11}. Jewish are one of the recent migrants and they have been found in various part of India ¹². The migrations of Jewish populations to India is poorly understood. One of the views is that ten Jewish Families released from Persian Jail by the king in BC605 have arrived Kodungallore, Kerala, India now called as Cochin Jewish group. However, their arrival in India has not been well documented and any contemporary scholars argue their arrival at some time during the early Middle Ages. ¹²⁻¹⁴.

India harbors <5% of world Jewish population ^{15,16}. Among the major communities of Indian Jewish populations, the Cochin Jewish community was thought to be the pioneer among the Indian Jewish who entered to India ^{12,17}. The Bene Israel was considered as second group which has entered to the Konkan area (Maharashtra state) through sea from Middle East around 100 AD In India ^{12,13}. The Paradesi Jewish group arrived to Malabar coast (Kerala) during 15-16th Century after their exile from Spain ^{13,14,17}. More recently the Baghdadi Jewish group from modern Iran, Iraq and adjoining region migrated to Kolkata (East of India) during end of 18th Century ^{13,18,19}.

In the contemporary India the Paradesi Jewish community are integrated with Cochin Jewish and also many Jewish groups are converted to Christianity. Jewish were just another group, not subject to particular exclusion or marginalization. Non-Jewish could, and did, move into the Indian Jewish community, while this was taboo in the Islamic or Christian world ^{12,13,15,17,18}. Bene Israel is the largest group among Indian Jewish community, followed by Cochin Jewish and Baghdadi Jewish ¹⁶.

Because of different population histories and affiliation with various diverse Middle Eastern groups, the relation between different Indian Jewish are rather complex. It was shown that the Indian Jewish communities hardly ever formed a social contact and common identity ¹⁹.

References

1. Metspalu, M. *et al.* Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC Genet.* **5**, 26 (2004).
2. Behar, D. M. *et al.* The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event. *Am J Hum Genet.* **78**, 487-497 (2006).
3. Behar, D. M. *et al.* Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. *PLoS ONE.* **3**, e2062 (2008).
4. Behar, D. M. *et al.* The genome-wide structure of the Jewish people. *Nature.* **466**, 238-242 (2010).
5. Rieder, M. J., Taylor, S. L., Tobe, V. O. & Nickerson, D. A. Automating the identification of DNA variations using quality-based fluorescence re-sequencing: analysis of the human mitochondrial genome. *Nucleic Acids Res.* **26**, 967-73. (1998).
6. YCC A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res.* **12**, 339-348 (2002).
7. Karafet, T. M. *et al.* New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res.* **18**, 830-838 (2008).
8. Andrews, R. M. *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet.* **23**, 147 (1999).
9. Behar, D. M. *et al.* A “Copernican” reassessment of the human mitochondrial DNA tree from its root. *The American Journal of Human Genetics.* **90**, 675-684 (2012).
10. Chaubey, G. The demographic history of India: A perspective based on genetic evidence (<http://hdl.handle.net/10062/15240>). PhD. (Universitatis Tartuensis, Estonia, 2010).
11. Romila, T. A History of India, Vol. 1.
12. Slapak, O. *The Jews of India* (UPNE, 1995).
13. Katz, N. *Who Are the Jews of India?* (Univ of California Press, 2000).
14. Israel, R. R. *The Jews of India* , 2002).
15. Fernandes, E. *The Last Jews Of Kerala* (Granta Books, 2011).
16. Ehrlich, M. A. *Encyclopedia of the Jewish Diaspora* (ABC-CLIO, 2009).

17. Roland, J. G. *The Jewish Communities of India* (Transaction Publishers, 1998).
18. Vail, S. *India's Jewish heritage* (Marg Publications, 2002).
19. Egorova, Y. *Jews and India* (Psychology Press, 2006).

Present affiliation:

Manvendra Singh: Max Delbrueck Center for Molecular Medicine (MDC), Berlin, 13125, Germany. E mail- mnujnu@gmail.com

Kamayani Singh: University of Oxford, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN UK. E mail- singh.kamayani@gmail.com

Deepankar Pratap Singh: Friedrich Miescher Institute for Biomedical Research Maulbeerstrasse 66 4058 Basel, Switzerland. E mail- deepankar.singh@fmi.ch

Ashish Singh: Toxicogenomics and Predictive Toxicology lab, CSIR-Indian Institute of Toxicology Research (IITR) Lucknow, India ashish222985@gmail.com

Rakesh Tamang: Department of Zoology, University of Calcutta, Kolkata 700019, India. E mail- rakesh.tamang7@gmail.com

Vijay Kumar Singh: Dept. of Pediatric Hematology/Oncology, Justus-Liebig. University Giessen, Feulgenstraße 12, 35392 Giessen, Germany. E mail- vijayvishen@gmail.com

Legend for Supplementary Figures:

Supplementary Fig. 1. The ADMIXTURE plot for K=2-9.

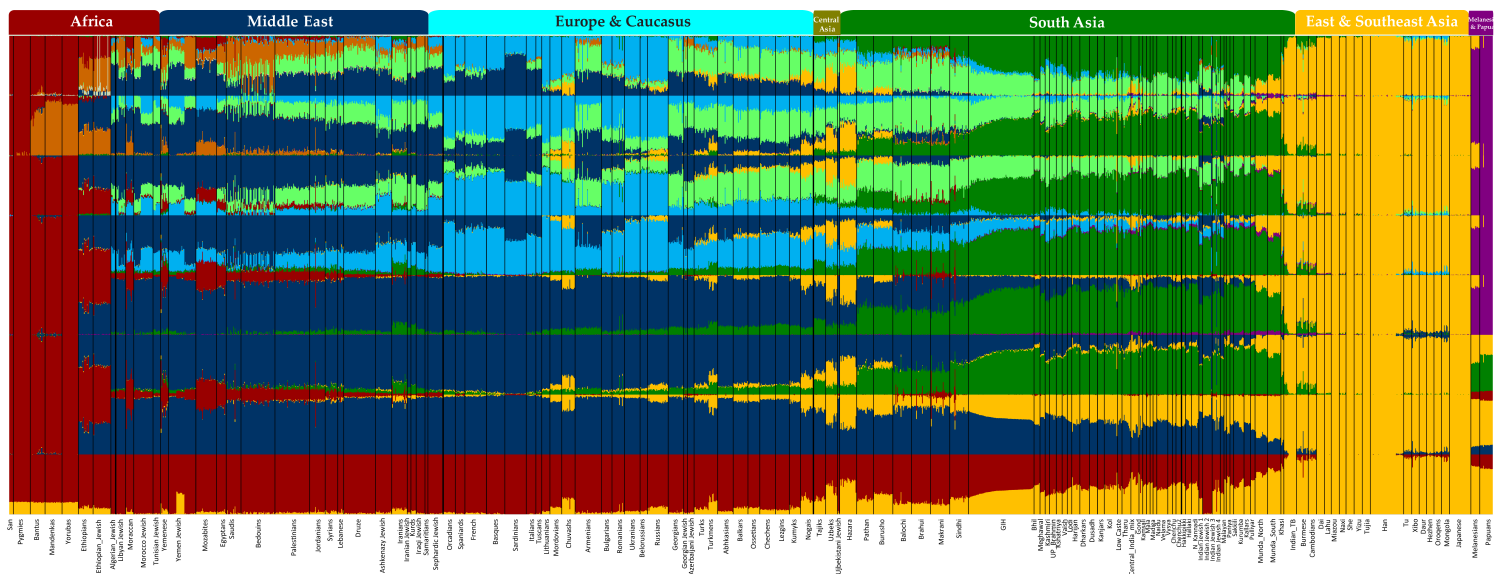
Supplementary Fig. 2. The spatial distribution of Middle Eastern specific ancestry component obtained from ADMIXTURE analysis. The map was generated by using Surfer8 of Golden Software (Golden Software Inc., Golden, Colorado).

Supplementary Fig. 3. The plot of shared drift obtained by the f3 =(Yoruba; Druze, X). The f3 values were plotted on Y axis against the X- targeted populations on X axis.

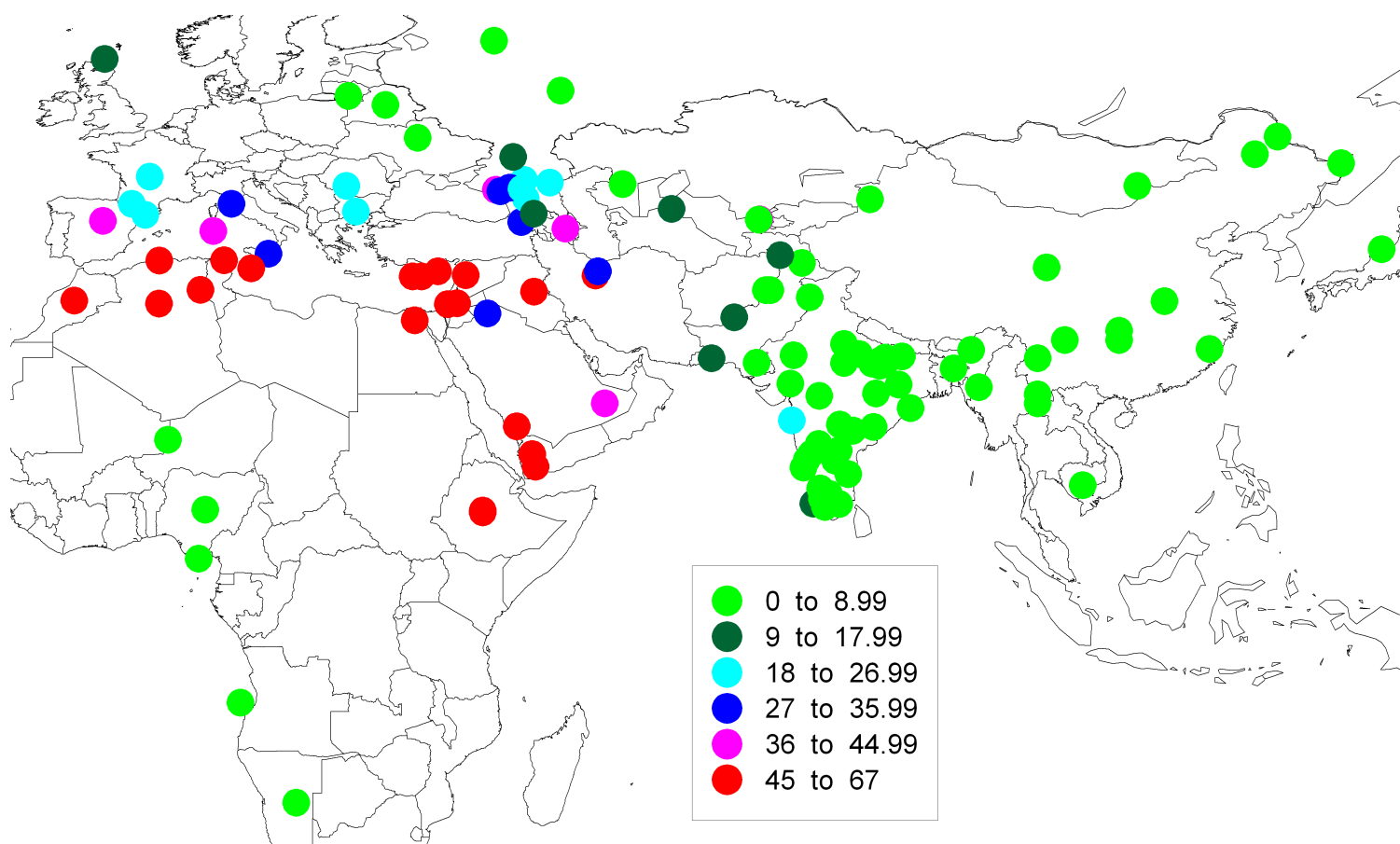
Supplementary Fig. 4. Comparison of Middle East vs Indian ancestry sharing of various Eurasian populations, obtained from f3 shared drift analysis.

Supplementary Fig. 5. The average count of number and length of ROH (Runs of Homozygosity) segments among various populations including Indian Jewish.

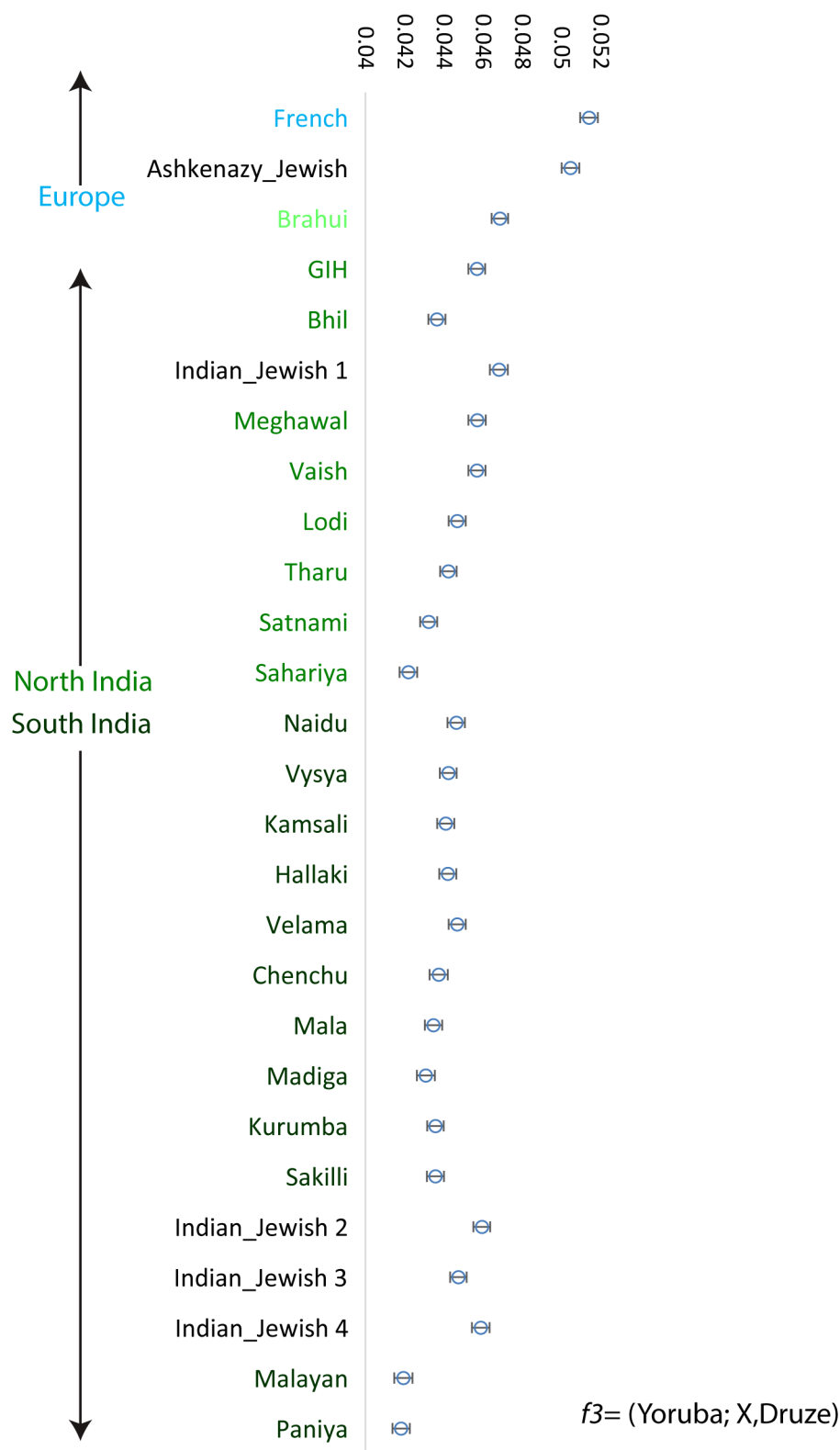
Supplementary Fig. 1



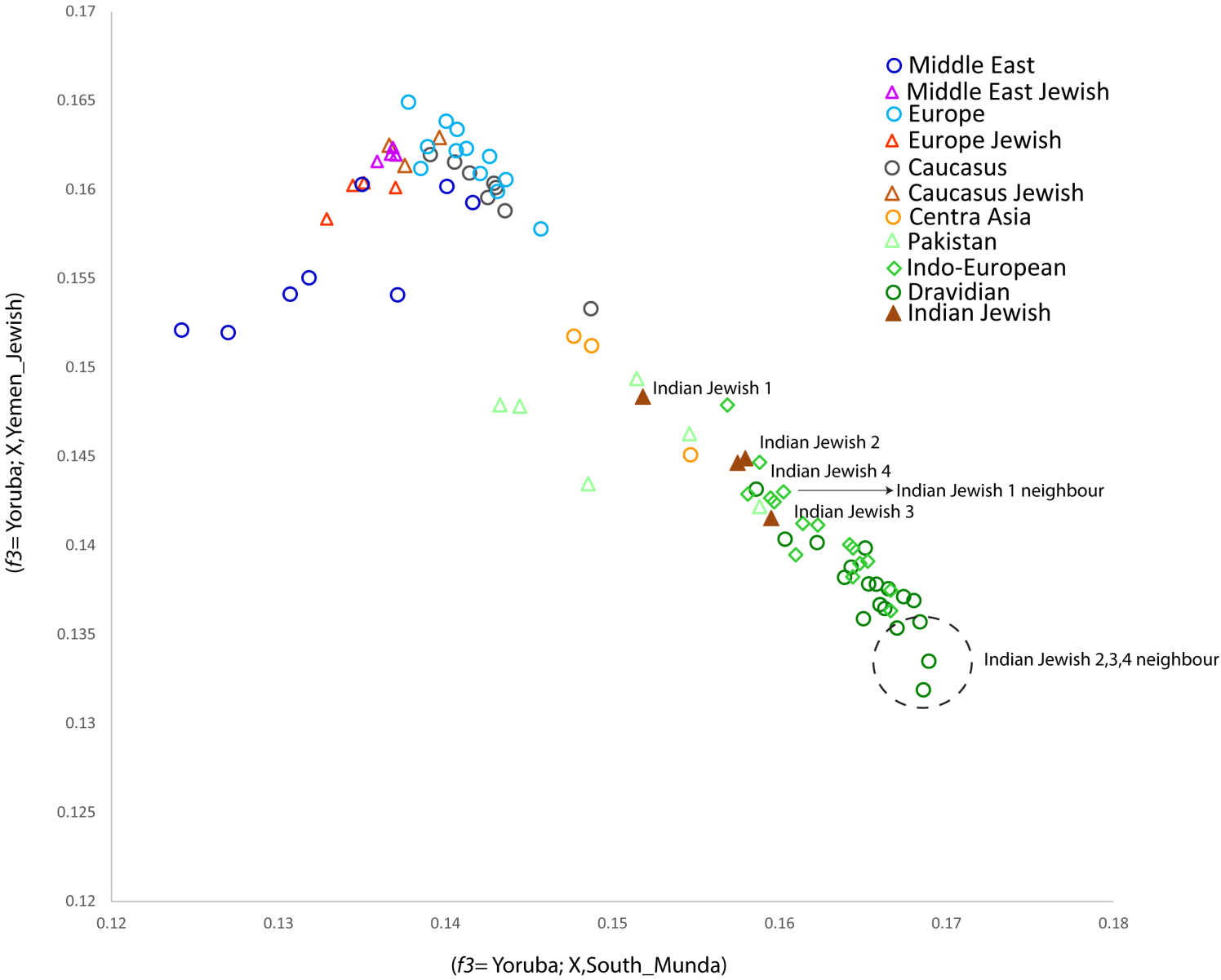
Supplementary Fig. 2



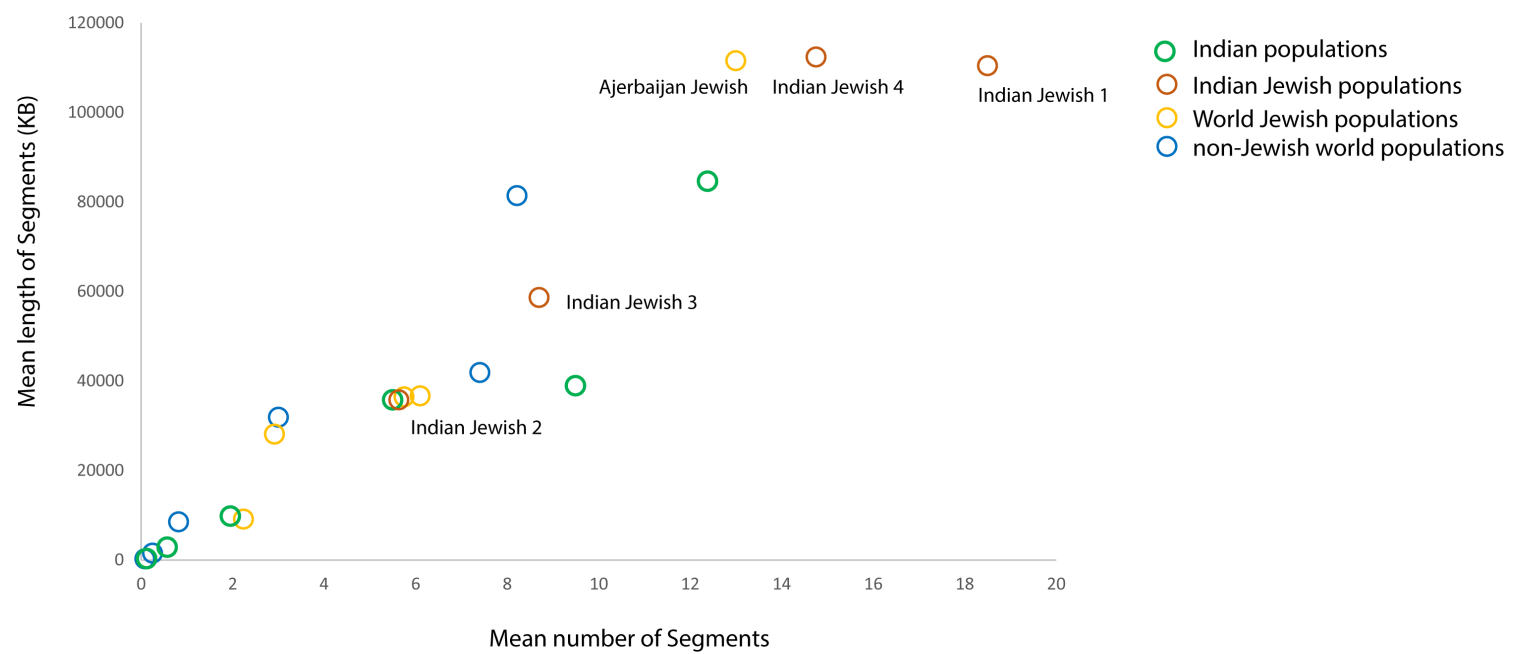
Supplementary Fig. 3



Supplementary Fig. 4



Supplementary Fig. 5



Supplementary Table. 2. The populationwise pairwise F_{st} calculated from autosomal data

[illegible]

Middle East
Jewish
Europe
Caucasus
Central Asia
South Asian
Southeast and East Asian

Supplementary Table. 3. The ANI (Ancestral North Indian) ancestry among studied groups

Population	% ANI (SD)
X= North India	
GIH	47.78 (1.28)
Bhil	26.12 (1.89)
Indian Jewish 1	56.59 (2.27)
X= South India	
Kurumba	24.63 (1.77)
Sakilli	23.50 (2.22)
Malayan	8.45 (3.11)
Paniya	7.45 (2.46)
Indian Jewish 2	48.89 (1.84)
Indian Jewish 3	38.05 (1.65)
Indian Jewish 4	47.25 (2.29)

% ANI = Yoruba,French;X,S_Munda/Yoruba,French;Georgian,S_Munda

Supplementary Table. 4. The extended table showing formal test of Admixture using Alder

Admixed population	Surrogate population 1	Surrogate population 2	Time of Admixture (Generations)	SD	p value	Z score
Jewish 1	GIH	Yemenite Jewish	39.8	5.7	1.9 x 10 ⁻⁴	3.73
Jewish 1	GIH	Druze	37	8.6	6.9x10 ⁻⁵	3.98
Jewish 2	Kurumba	Yemenite Jewish	18.5	4.9	4.6 x 10 ⁻⁴	3.5
Jewish 2	Paniya	Yemenite Jewish	19	7.6	1.2 x 10 ⁻²	2.51
Jewish 2	Kurumba	Druze	16.3	4.7	5 x 10 ⁻⁴	3.48
Jewish 2	Paniya	Druze	17.2	9.8	7.8 x 10 ⁻²	1.76
Jewish 3	Kurumba	Yemenite Jewish	131.7	122.9	0.61	0.5
Jewish 3	Paniya	Yemenite Jewish	82	41.7	4.9 x 10 ⁻²	1.97
Jewish 3	Kurumba	Druze	42.9	76.7	0.48	0.63
Jewish 3	Paniya	Druze	27.4	7.6	8.4 x 10 ⁻⁴	3.34
Jewish 3	Kurumba	-	36.5	11.3	1.6x 10 ⁻⁴	3.23
Jewish 4	Kurumba	Yemenite Jewish	30.8	21.1	0.14	1.46
Jewish 4	Paniya	Yemenite Jewish	10.1	5.2	5.2 x 10 ⁻²	1.94
Jewish 4	Kurumba	Druze	53.1	11.3	3.1x10 ⁻⁵	4.17
Jewish 4	Paniya	Druze	14.8	39.1	0.7	0.38

NKN47	Kattunaiken	Kattunaiken	189-223-519	10398-10400-15043-15301	M	Present Study
NKN48	Kattunaiken	Kattunaiken	189	12705	R	Present Study
NKN48	Kattunaiken	Kattunaiken	189-223-294-519	10398-10400-15043-15301	M3c1b1	Present Study
NKN50	Kattunaiken	Kattunaiken	124-179-183C-189-223-294	10398-10400-15043-15301	M3c1b1	Present Study
NKN51	Kattunaiken	Kattunaiken	223-274-319-320	15043-15301-15670	M2b	Present Study
NKN51	Kattunaiken	Kattunaiken	124-179-183C-189-223-294	10398-10400-15043-15301	M3c1b1	Present Study
NKN51	Kattunaiken	Kattunaiken	179-183C-189-223-274-319-320-519	10398-10400	M2a1a1	Present Study
NKN51	Kattunaiken	Kattunaiken	179-183C-189-223-274-319-320-519	10398-10400	M2a1a1	Present Study
NKN52	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301	M4	Present Study
NKN52	Kattunaiken	Kattunaiken	223-519	10398-10400-15043-15301	M	Present Study
NKN52	Kattunaiken	Kattunaiken	223-519	10398-10400-15043-15301	M	Present Study
NKN53	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301	M4	Present Study
NKN54	Kattunaiken	Kattunaiken	223-239-298-327-357	10398-10400-15043-15301	C4a4a	Present Study
NKN56	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301-15431	M30	Present Study
NKN57	Kattunaiken	Kattunaiken	189-223-294	10398-10400-15043-15301	M3c1b1	Present Study
NKN58	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301-15431	M30	Present Study
NKN61	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301-15431	M30	Present Study
NKN62	Kattunaiken	Kattunaiken	189	12705	R	Present Study
NKN62	Kattunaiken	Kattunaiken	189-230-356-519	15285	U4	Present Study
NKN62	Kattunaiken	Kattunaiken	183C-189-230-356	15385	R5a1	Present Study
NKN63	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301-15431	M30	Present Study
NKN64	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301	M4	Present Study
NKN70	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301-15431	M30	Present Study
NKN71	Kattunaiken	Kattunaiken	093-223-234-262	10398-10400-15043-15301-15431	M30	Present Study
NKN72	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301-15431	M30	Present Study
NKN73	Kattunaiken	Kattunaiken	126-223	482-10398-10400-15043-15301	M3	Present Study
NKN74	Kattunaiken	Kattunaiken	126-223	482-10398-10400-15043-15301	M3	Present Study
NKN75	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301	M4	Present Study
NKN76	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301-15431	M30	Present Study
NKN77	Kattunaiken	Kattunaiken	093-223-234-262	10398-10400-15043-15301-15431	M30	Present Study
NKN78	Kattunaiken	Kattunaiken	183C-189-223-274-319-320	477G-10398-10400	M2a1a1	Present Study
NKN79	Kattunaiken	Kattunaiken	223	10398-10400-15043-15301-15431	M30	Present Study
NKN79	Kattunaiken	Kattunaiken	189-223-274-319-320-519	10398-10400-15043-15301-15670	M2a1	Present Study
NKN80	Kattunaiken	Kattunaiken	124-179-183C-223-294	10398-10400-15043-15301	M3c1b1	Present Study
NKN83	Kattunaiken	Kattunaiken	124-179-183C-189-223-294	15285-15954	M40a	Present Study

Behar, D. M. *et al.* Counting the founders: the matrilineal genetic ancestry of the Jewishish Diaspora. *PLoS ONE*. **3**, e2062 (2008).

