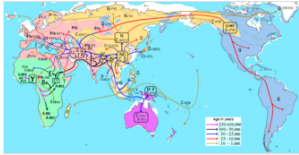


# Y-DNA haplogroups in populations of South Asia

**Y-DNA haplogroups in populations of South Asia** are haplogroups of the male Y-chromosome found in South Asian populations.

## Major Y-chromosome DNA haplogroups in South Asia



World map of early migrations of modern human beings based on the Y-chromosome DNA.

South Asia, located on the crossroads of Western Eurasia and Eastern Eurasia, accounts for about 39.49% of Asia's population,<sup>[1]</sup> and over 24% of the world's population.<sup>[2]</sup> It is home to a vast array of people who belong to diverse ethnic groups, who migrated to the region during different periods of time.<sup>[3]</sup>

The presence of Himalayas in northern and eastern borders of South Asia have limited migrations from Eastern Eurasia into Indian subcontinent in the past. Hence most of the male-mediated migrations into South Asia occurred from Western Eurasia into the region, as seen in the Y-chromosome DNA Haplogroup variations of populations in the region.<sup>[4]</sup>

The major paternal lineages of South Asian populations, represented by Y chromosomes, are haplogroups **R1a1**, **R2**, **H**, **L**, and **J2**,<sup>[5]</sup> as well as **O-M175** in some parts (northeastern region) of the Indian subcontinent.<sup>[6]</sup>

Haplogroup R is the most observed Y-chromosome DNA haplogroup among the populations of South Asia,<sup>[4]</sup> followed by H, L, and J, in the listed order.<sup>[4]</sup> These four haplogroups together constitute nearly 80% of all male Y-chromosome DNA haplogroups found in various populations of the region.<sup>[4]</sup>

The Y-chromosome DNA Haplogroups R1a1, R2, L, and J2, which are found in higher frequencies among various populations of the Indian subcontinent, are also observed among various populations of Europe, Central Asia, and Middle East.<sup>[7]</sup> Y-chromosome DNA haplogroups found in South Asia show strong affinity towards West Eurasian uniparental male line of ancestry while the mitochondrial DNA (mtDNA) haplogroups found in most of the populations of the region show an admixture of West Eurasian and indigenous ancient South Asian uniparental female lines of ancestries in contrast.<sup>[4]</sup>

Some researchers have argued that Y-DNA Haplogroup R1a1 (M17) is of autochthonous South Asian origin.<sup>[8]</sup> However, proposals for a Eurasian Steppe origin for R1a1 are also quite common and supported by several more recent studies.<sup>[9]</sup> The spread of R1a1 in Indian subcontinent is associated with Indo-Aryan migrations into the region from South Central Asia that occurred around 3,500-4,000 years before present.<sup>[4]</sup> The R1a-Z93 paternal genetic in Romani people was also discovered.<sup>[10]</sup> Indian-Brahmin origin of paternal haplogroup R1a1\*.<sup>[11]</sup>

The Haplogroup R2 is mostly restricted to various populations of the Indian subcontinent, in addition to some populations of South Central Asia and parts of Middle East where it is observed in low frequencies.<sup>[4]</sup> R2 has high frequency among the speakers of the Dravidian languages of South India and it spread throughout the Indian subcontinent during the spread of agriculture associated with the Neolithic Revolution, before the spread of R1a1 in the region.<sup>[4]</sup>

The Haplogroup H (also known as the "Indian marker"<sup>[4]</sup>), which is a direct descendant of the Upper Paleolithic Eurasian Haplogroup HIJK, is mostly restricted to South Asian populations of the Indian subcontinent,<sup>[4]</sup> in addition to some populations of South Central Asia and eastern Iranian plateau, where it is found in low frequencies.<sup>[5]</sup> It originated somewhere in the Middle East or South Central Asia and travelled to Indian subcontinent and adjoining areas of the eastern Iranian plateau around 40,000-50,000 years before present.<sup>[4]</sup>

The Haplogroup L, which is thought to have originated near Pamir mountains of present-day Tajikistan in South Central Asia,<sup>[4]</sup> travelled throughout Indian subcontinent during the Neolithic period, and it is associated with the spread of the Bronze Age Indus Valley Civilisation (IVC) in South Asia, which existed around 3,300-5,300 years before present.<sup>[4]</sup> It is also observed among many populations of the Iranian plateau. The spread of the Haplogroup J2 from Iranian plateau into Indian subcontinent also occurred during the Neolithic period, alongside L.<sup>[4]</sup>

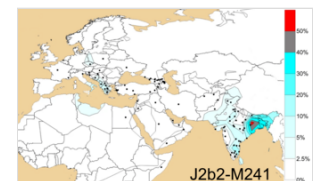
The Haplogroup O-M175, which is a major haplogroup observed among the populations of East and Southeast Asia, is found among many Austroasiatic and Tibeto-Burman speakers of northeastern regions of the Indian subcontinent.<sup>[6]</sup>

## Frequencies in South Asian ethnic groups

Listed below are some notable groups and populations from South Asia by human Y-chromosome DNA haplogroups based on various relevant studies.

The samples are taken from individuals identified with specific linguistic designations (**IE**=Indo-European, **Dr**=Dravidian, **AA**=Austro-Asiatic, **ST**=Sino-Tibetan) and individual linguistic groups, the third column (**n**) gives the sample size studied, and the other columns give the percentage of the respective haplogroups.

Majority of the Indo-European (**IE**) speakers of South Asia speak Indo-Aryan languages, followed by Iranian languages, both of which belong to Indo-Iranian branch of the Indo-European language family.<sup>[13][14]</sup> They form around 75% of the South Asian populations.<sup>[13]</sup>



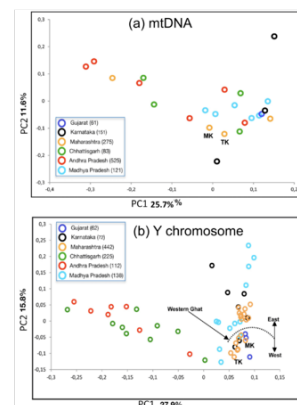
frequency distribution of Y-DNA haplogroup J2b2-M241

The Dravidian (Dr) speakers of South Asia are mostly clustered in South India and Balochistan.<sup>[15]</sup> They form around 20% of the South Asian populations.<sup>[15]</sup>

The Sino-Tibetan (ST) speakers in northeastern parts of the Indian subcontinent speak various languages belonging to Tibeto-Burman branch of the Sino-Tibetan language family.<sup>[16]</sup>

The Austroasiatic (AA) speakers of South Asia are scattered in parts of East India and Bangladesh.

*Note:* The converted frequencies from some old studies conducted in 2000s may lead to unsubstantial frequencies below. Table below has been sorted in alphabetical order based on the name of the population.



Principal component (PC) scatter plot of mtDNA (a) and Y chromosomal (b) among the population of Maharashtra and other contemporary populations. <sup>[12]</sup>

Frequencies of Y chromosome DNA haplogroups in some selected ethi

Population	Language (if specified)	n	<u>C</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>H</u>	<u>I</u>	<u>J</u>	<u>K</u>	<u>L</u>	<u>N</u>	<u>O</u>
<b>Andhra Tribals (Andhra Pradesh, India)</b>	Dr (Telugu)	29	0	0	10.3%	3.4%	34.5%	0	3.4%	0	6.9%	0	6.9%
<b>Balochi (Balochistan, Pakistan)</b>	IE (Balochi)	25	0	8%	0	0	4%	0	16%	0	24%	0	0
<b>Bengalis (Bangladesh)</b>	IE (Bengali)	42	7.14%	0	0	4.72%	35.71%	0	11.9%	0	4.76%	0	9.52%
<b>Bengalis (Bangladesh)</b>	IE (Bengali)	649	0	2.62%	0	1.08%	17.72%	4.16%	15.72%	0	17.41%	0.46%	0
<b>Bengalis (West Bengal, India)</b>	IE (Bengali)	31	3.2%	0	3.2%	3.2%	9.7%	0	9.7%	0	0	0	3.2%
<b>Bhargavas (India)</b>	IE (Indo-Aryan)	96	6.3%	0	4.2%	0	16.7%	0	6.3%	4.2%	0	0	2.1%
<b>Bihar Paswan (Bihar, India)</b>	IE (Bhojpuri)	27	0	0	3.7%	11.11%	14.81%	0	3.7%	0	7.41%	0	3.7%
<b>Bihar Brahmins (Bihar, India)</b>	IE (Maithili)	38	2.63%	0	0	0	0	0	2.63%	5.26%	13.16%	0	0
<b>Brahmins (India)</b>	IE (Indo-Aryan), Dr (Dravidian)	118	10.2%	0	2.5%	1.7%	11.9%	0	11.9%	1.7%	2.5%	0	4.2%
<b>Brahui (Balochistan, Pakistan)</b>	Dr (Brahui)	25	4%	0	0	16%	4%	0	28%	0	8%	0	4%
<b>Burusho (Gilgit-Baltistan, Pakistan)</b>	Isolate (Burushaski)	97	8.2%	0	1%	1%	4.1%	0	8.2%	0	16.5%	0	3.1%
<b>Chakma (Bangladesh)</b>	IE (Chakma)	116	0	1.72%	0	3.45%	4.31%	6.03%	10.34%	0	52.59%	1.72%	0
<b>Chamar (Hindi Belt, India)</b>	IE (Indo-Aryan)	18	5.6%	0	0	0	44.4%	0	0	0	0	0	0
<b>Chaturvedis (Hindi Belt, India)</b>	IE (Hindi)	88	9.1%	0	3.4%	0	6.8%	0	12.5%	4.5%	2.3%	0	2.3%
<b>Chenchu (South India)</b>	Dr (Telugu)	41	4.9%	0	0	0	36.6%	0	7.3%	0	14.6%	0	0
<b>Dawoodi Bohra (Gujarat, India)</b>	IE (Gujarati)	50	2%	0	4%	0	26%	0	8%	0	6%	0	0
<b>Dawoodi Bohra (Tamil Nadu, India)</b>	IE (Urdu)	26	0	0	0	0	38.5%	0	0	0	7.7%	0	0
<b>Gujarat Bhils (Gujarat, India)</b>	IE (Bhil)	22	9.09%	0	0	0	27.27%	0	18.18%	0	18.18%	0	0
<b>Gujarat Brahmins (Gujarat, India)</b>	IE (Gujarati)	64	3.33%	3.33%	0	10.94%	1.56%	0	15.63%	3.13%	7.81%	3.13%	0
<b>Gujarati Indians (USA)</b>	IE (Gujarati)	58	20.69%	0	0	3.45%	20.69%	0	8.62%	0	8.62%	0	0

Population	Language (if specified)	n	<u>C</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>H</u>	<u>I</u>	<u>J</u>	<u>K</u>	<u>L</u>	<u>N</u>	<u>O</u>
<b>Gujaratis (Gujarat, India)</b>	IE (Gujarati)	29	17.2%	0	0	0	13.8%	0	20.7%	0	10.4%	0	0
<b>Hazara (Balochistan, Pakistan)</b>	IE (Hazaragi)	25	40%	0	0	0	0	4%	4%	0	0	0	8%
<b>Himachal Brahmin (Himachal Pradesh, India)</b>	IE (Pahari)	19	5.26%	0	15.79%	0	10.53%	0	5.26%	0	5.26%	0	0
<b>India</b>	IE (Indo-Aryan), Dr (Dravidian), AA (Austroasiatic), ST (Tibeto-Burman)	728	1.8%	0	5.2%	1.2%	26.4%	0	9.3%	0	18.7%	0	23.9%
<b>India</b>	IE (Indo-Aryan), Dr (Dravidian), AA (Austroasiatic), ST (Tibeto-Burman)	1152	1.4%	0	3%	0.1%	23%	0	9.1%	0	17.5%	0	18%
<b>India</b>	IE (Indo-Aryan), Dr (Dravidian), AA (Austroasiatic), ST (Tibeto-Burman)	1,615	0	0	0	0	0	0	0	0	0	0	0
<b>India (Central)</b>	IE (Indo-Aryan), Dr (Dravidian)	50	0	0	4%	2%	20%	0	4%	0	4%	0	8%
<b>India (East)</b>	IE (Indo-Aryan), AA (Austroasiatic), Dr (Dravidian)	367	0.8%	0	2.7%	0	19.3%	0	4.1%	0	1.9%	0	20.7%
<b>India (West)</b>	IE (Indo-Aryan)	204	5.4%	0	0.5%	0	33.3%	0	11.3%	0	11.8%	0	0
<b>India (North)</b>	IE (Indo-Aryan), ST (Tibeto-Burman)	180	0	0	1.1%	0.6%	24.5%	0	7.8%	0	1.7%	0	2.3%
<b>India (Northeast)</b>	ST (Tibeto-Burman)	108	0	0	0	0	0.9%	0	0	0	0	0	79.7%
<b>India (S. Gujarat tribals)</b>	IE (Indo-Aryan)	284	8.5%	0	4.2%	0	40.1%	0	10.2%	2.8%	3.2%	0	0
<b>India (South)</b>	Dr (Dravidian)	372	1.9%	0	4%	0	27.5%	0	19.7%	0	10.8%	0	0
<b>India Shia (India)</b>	IE (Urdu)	161	3.7%	1.9%	5.6%	5%	8.6%	0	28.7%	0	0	0	0.6%
<b>Indian Dravidians</b>	Dr (Dravidian)	353	1.7%	0	9.3%	2.3%	32.9%	0	19.7%	0	11.6%	0	13.6%
<b>Indian Indo-Europeans</b>	IE (Indo-Aryan)	205	2.4%	0	2.4%	0.5%	28.8%	0	11.3%	0	3.9%	0	4.9%
<b>Indian Munda</b>	AA (Munda)	892	0	0	4%	0	23.1%	0	3.9%	0	0	0	57.2%
<b>Indian Sunni</b>	IE	129	0	0	3.1%	2.3%	14%	0	20.1%	2.3%	3.1%	0	0

Population	Language (if specified)	n	C	E	F	G	H	I	J	K	L	N	O
(India)	(Urdu)												
<b>Indian tribes</b>	IE (Indo-Aryan), Dr (Dravidian), AA (Austroasiatic), ST (Tibeto-Burman)	505	2.2%	0	2%	0.2%	21.2%	0	2.6%	0	3.2%	0	40.6%
<b>Indian Sino-Tibetans</b>	ST (Tibeto-Burman)	87	1.1%	0	0	0	2.3%	0	0	0	0	0	86.2%
<b>Indian Telugus (UK)</b>	Dr (Telugu)	60	1.67%	0	0	0	26.67%	0	6.67%	0	1.67%	1.67%	0
<b>India's Lower Castes</b>	IE (Indo-Aryan), Dr (Dravidian)	261	0.8%	0	4.6%	0	27.6%	0	3.1%	0	5.4%	0	0.4%
<b>India's Middle Castes</b>	IE (Indo-Aryan), Dr (Dravidian)	175	0.6%	0	5.1%	0	21.1%	0	9.7%	0	5.7%	0	0
<b>India's Upper Castes</b>	IE (Indo-Aryan), Dr (Dravidian)	211	0.9%	0	1.9%	0	23.3%	0	10%	0	11.4%	0	0
<b>Indo-Aryan Castes (India)</b>	IE (Indo-Aryan)	29	5.3%	0	3.6%	0.6%	6.5%	0	16.6%	1.2%	11.2%	0	1.2%
<b>Iranian Shia (India)</b>	IE (Indo-Aryan)	25	0	0	16%	8%	4%	0	24%	28%	0	0	0
<b>Jats (Punjab)</b>	IE (Punjabi)	302	0	1%	0	3.6%	3.6%	1%	9.6%	0	36.8%	0	0
<b>J&amp;K Kashmir Gujars (Jammu and Kashmir, India)</b>	IE (Gujari)	49	2.04%	0	4.08%	0	10.2%	0	6.12%	8.16%	16.33%	0	0
<b>J&amp;K Kashmiri Pandits (Jammu and Kashmir, India)</b>	IE (Kashmiri)	51	1.96%	0	3.92%	1.96%	9.8%	0	9.8%	9.8%	5.88%	0	0
<b>Kalash (Khyber Pakhtunkhwa, Pakistan)</b>	IE (Kalasha)	44	0	0	0	18.2%	20.5%	0	9.1%	0	25%	0	0
<b>Kathmandu (Nepal)</b>	IE (Indo-Aryan), ST (Tibeto-Burman)	77	7.8%	0	0	0	11.7%	0	10.4%	0	0	0	20.8%
<b>Khasi (Meghalaya, India)</b>	AA (Khasi)	92	0	0	10.9%	0	6.5%	0	0	0	0	0	72.8%
<b>Kodava (Karnataka, India)</b>	Dr (Kodava)	50	2%	0	2%	0	8%	0	16%	0	18%	0	0
<b>Khasi (Bangladesh)</b>	AA (Khasi)	90	0	12.75%	0	0	3.92%	0	9.80%	0	45.10%	0	0
<b>Kokanastha Brahmin (Maharashtra, India)</b>	IE (Marathi)	25	0	0	0	0	12%	0	28%	0	4%	0	0
<b>Konkanastha Brahmins (Maharashtra, India)</b>	IE (Marathi)	43	2.3%	0	2.3%	0	9.3%	0	14%	2.3%	18.6%	0	0

Population	Language (if specified)	n	<u>C</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>H</u>	<u>I</u>	<u>J</u>	<u>K</u>	<u>L</u>	<u>N</u>	<u>O</u>
<b>Koraga Tribals (Karnataka, India)</b>	Dr (Koraga)	33	0	0	0	0	87.9%	0	0	0	0	0	0
<b>Koya (Andhra Pradesh and Telangana, India)</b>	Dr (Koya)	41	0	0	26.8%	0	70.7%	0	0	0	0	0	0
<b>Lambadi (Rajasthan, India)</b>	IE (Lambadi)	35	17.1%	0	2.9%	0	8.6%	0	5.7%	0	17.1%	0	0
<b>Madhya Pradesh Brahmins (Madhya Pradesh, India)</b>	IE (Hindi)	42	0	0	2.38%	0	7.14%	0	23.81%	0	7.14%	2.38%	0
<b>Madhya Pradesh Gonds (Madhya Pradesh, India)</b>	Dr (Gondi)	64	0	0	0	0	62.5%	0	0	6.25%	0	0	6.25%
<b>Madhya Pradesh Saharia (Madhya Pradesh, India)</b>	IE (Hindi)	57	0	0	5.27%	0	33.33%	0	0	3.51%	0	0	0
<b>Maharashtra Brahmins (Maharashtra, India)</b>	IE (Marathi)	30	3.33%	0	0	3.33%	10%	0	16.67%	3.33%	10%	3.33%	3.33%
<b>Manipuri (Bangladesh)</b>	ST (Meitei)	102	0	0	0	5.56%	2.22%	10.00%	11.11%	0	25.56%	1.11%	0
<b>Mappila Muslims (Kerala, India)</b>	Dr (Malayalam)	40	0	0	0	0	27.5%	0	20%	5%	10%	0	0
<b>Mundari (East India)</b>	AA (Munda)	789	0	0	3.3%	0	25.4%	0	4.4%	0	0	0	55%
<b>New Delhi Hindus (New Delhi, India)</b>	IE (Hindi)	49	0	0	0	2%	18.3%	0	8.1%	0	4.1%	0	2%
<b>Pakistan</b>	IE (Indo-Aryan, Iranian)	176	7.4%	0	0	6.2%	6.2%	0	15.3%	0	13.1%	0	2.3%
<b>Pakistan</b>	IE (Indo-Aryan, Iranian), Dr (Brahui)	638	3%	0	0.8%	2.7%	2.5%	0	20.2%	0	11.6%	0	0.5%
<b>Parsis (Western India – Lay)</b>	IE (Persian)	122	0	5.7%	0	0	0	0	54.9%	1.6%	2.5%	0	0
<b>Parsis (Western India – Priest)</b>	IE (Persian)	71	0	1.4%	0	0	0	0	4.2%	0	54.9%	0	0
<b>Pashtun (Khyber Pakhtunkhwa, Pakistan)</b>	IE (Pashto)	96	0	0	2.1%	11.5%	4.2%	0	6.2%	0	12.5%	0	5.2%
<b>Pathan (Khyber Pakhtunkhwa, Pakistan)</b>	IE (Pashto)	21	4.8%	0	0	9.5%	14.3%	0	0	0	9.5%	0	0

Population	Language (if specified)	n	<u>C</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>H</u>	<u>I</u>	<u>J</u>	<u>K</u>	<u>L</u>	<u>N</u>	<u>O</u>
<u>Punjab Brahmin (Punjab, India)</u>	IE (Punjabi)	28	3.58%	0	3.57%	3.57%	0	0	21.43%	0	7.14%	0	0
<u>Punjabi (Punjab, India)</u>	IE (Punjabi)	66	3%	0	0	0	4.6%	0	21.2%	0	12.1%	0	0
<u>Punjabis (Punjab, Pakistan)</u>	IE (Punjabi)	48	2.08%	0	0	8.33%	6.25%	0	27.08%	0	4.17%	0	0
<u>Rajput (North India)</u>	IE (Indo-Aryan)	29	3.4%	0	3.4%	0	20.7%	0	17.2%	0	6.9%	0	3.4%
<u>Rakhine (Bangladesh)</u>	ST (Arakanese)	107	0	0	0	1.87%	6.54%	1.87%	13.08%	0	29.91%	3.74%	0
<u>Shia (India)</u>	IE (Urdu)	154	9.1%	11%	3.3%	9.7%	7.8%	0	19.5%	3.3%	3.9%	0	2%
<u>Sindhi (Sindh, Pakistan)</u>	IE (Sindhi)	21	0	0	0	0	0	0	33.3%	0	4.8%	0	0
<u>Sinhalese (Sri Lanka)</u>	IE (Sinhala)	39	0	0	10.3%	0	10.3%	0	10.3%	0	18%	0	0
<u>Sri Lanka</u>	IE (Sinhala)	74	0	0	9%	0	15%	0	14%	1%	19%	0	0
<u>Sourashtrians (Brahmin) (Tamil Nadu, India)</u>	IE (Gujarati), Dr (Tamil)	46	6.5%	0	4.4%	0	15.2%	0	2.2%	0	26.1%	0	2.2%
<u>South Indian Tribals (South India)</u>	Dr (Dravidian), AA (Austroasiatic)	315	8.6%	0	18.1%	0	31.1%	0	2.9%	0	7%	0	6.7%
<u>South castes (South India)</u>	Dr (Dravidian)	447	5.1%	0	11.9%	0	14.1%	0	9.8%	5.6%	21.9%	0	1.6%
<u>Sri Lanka</u>	IE (Sinhala), Dr (Tamil)	91	3.3%	0	9.9%	5.5%	25.3%	0	19.8%	0	15.4%	0	1.1%
<u>Sri Lankan Tamils (UK)</u>	Dr (Tamil)	55	1.8%	0	0	0	29.1%	0	14.5%	0	18.2%	0	0
<u>Sunni (India)</u>	IE (Urdu)	104	7.7%	0	2%	5.8%	10.6%	0	15.4%	2%	3.8%	0	2.9%
<u>Tamil Nadu (India)</u>	Dr (Tamil)	1680	4.4%	0.3%	16.3%	3.1%	23.6%	0	2.1%	14.0%	2.0%	0	0.4%
<u>Terai Hindus (Terai, Nepal)</u>	IE (Nepali)	26	11.5%	0	0	0	3.8%	0	3.8%	0	0	0	3.8%
<u>Tharu (Nepal)</u>	IE (Tharu)	171	0.6%	0	0	0	25.7%	0	14%	0	2.3%	0	36.8%
<u>Tripuri (Bangladesh)</u>	ST (Kokborok)	107	0	9.41%	0	1.18%	3.53%	3.53%	10.59%	0	65.88%	0	0
<u>Uttar Pradesh (South) Kols (Uttar Pradesh, India)</u>	AA (Munda)	54	0	0	0	0	11.11%	0	33.34%	0	0	0	40.74%
<u>Uttar Pradesh (South) Gonds (Uttar Pradesh, India)</u>	Dr (Gondi)	37	0	0	0	0	59.46%	0	18.92%	10.81%	0	2.7%	8.11%
<u>Uttar Pradesh Brahmin (Uttar Pradesh, India)</u>	IE (Hindi)	31	0	0	0	0	16.13%	0	3.23%	0	3.21%	0	0