

Population History of the Gond: The Largest Tribal Population of South Asia

Gyaneshwer Chaubey, *Cytogenetics Laboratory, Department of Zoology, Banaras Hindu University, Varanasi, India and Department of Evolutionary Biology, University of Tartu, Tartu, Estonia*

Rakesh K Upadhyay, *Bharat Adhyayan Kendra, Faculty of Arts, Banaras Hindu University, Varanasi, India*

George van Driem, *Institut für Sprachwissenschaft, Universität Bern, Bern, Switzerland*

Although the tribal populations represent a fraction of the South Asian population, there are only very few tribal populations ranging in the millions. The Central Indian tribal population Gond is one of them, with a census size of approximately 12 million people. Various disciplines of the humanities have drawn conflicting conclusions with regard to their origin. Therefore, in our previous study, we analysed hundreds of thousands of autosomal markers and found out that Gonds share their closest genetic similarity with the Austroasiatic (Munda) populations. While our findings support our previous contention, the current analysis has revealed that the Gonds occupy a transitional position between Dravidian and Munda groups. Sex-specific markers also differentiate the Gond substantially from the Indian Austroasiatic (Munda) and Dravidian (Telugu) speakers. Taken together, we suggest a unique and distinct genetic ancestry of the Gond population of South Asia.

Introduction

Ethnographers, anthropologists and linguists have long gleaned that the human diversity of the Indian subcontinent preserves numerous traces of older strata of population (see also: [Peopling of India: Insights from Genetics](#)). They have also pointed out that there is more ethnolinguistic complexity to the Indian

eLS subject area: Evolution & Diversity of Life

How to cite:

Chaubey, Gyaneshwer; Upadhyay, Rakesh K; and van Driem, George (June 2018) Population History of the Gond: The Largest Tribal Population of South Asia. In: eLS. John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0027474

Advanced article

Article Contents

- Introduction
- Gond: The Largest Tribe of India
- Let Us Recap
- Revisiting the Autosomal Analysis
- Sex-specific Markers Differentiate the Gond Significantly from Munda Speakers
- Acknowledgements

Online posting date: 14th June 2018

subcontinent than the four large linguistic phyla represented by Indo-European (Aryan), Dravidian, Trans-Himalayan (Tibeto-Burman) and Austroasiatic. Linguistic evidence of other older layers of peopling are found in the form of ethnolinguistic relict groups, such as the Burushaski (Berger, 1959; van Driem, 2001) Vedda (Marambe, 1896; de Silva, 1972; van Driem, 2001), Kusunda (Reinhard, 1969; van Driem, 2001; Watters, 2006), Andamanese (Man, 1885; van Driem, 2001; Abbi, 2006, 2012) and Nihali (Koppers, 1947; Kuiper, 1962; van Driem, 2001; Nagaraja, 2014). Additional evidence for other layers of peopling has been gleaned both in the form of phenotypical observations and measurements, which used to be referred to as somatology by the early physical anthropologists, as well as in the complex stratification of the caste system and the many tribal groups who have still survived effectively outside of the caste system.

von Heine-Geldern (1928, 1932) consequently spoke of the primordial demographics of the Indian subcontinent as comprising multiple ancient but distinct populations in Palaeolithic times. He saw the Munda language communities as the result of the linguistic incursion of Austroasiatic peoples from Southeast Asia into the subcontinent (see also: [Origins of the Austro-Asiatic Populations](#)). On the basis of linguistic rather than anthropological arguments, Pinnow (1963) likewise conjectured that a westward migration into the subcontinent had given rise to the ancient Munda. Their old ethnolinguistic hypothesis was borne out by Chaubey *et al.* (2011), with the additional new insight that this linguistic spread was a singularly male-biased intrusion.

A second hypothesis first developed by von Heine-Geldern (1917, 1945) entailed that both a subset of Dravidian populations represented by the various Gond linguistic communities and the local ancestral component of the Munda populations represented an older layer of peopling of the subcontinent. This theory was adopted by Grigson (1938), who proposed that the Gonds were an originally 'pre-Dravidian' or what he called 'proto-Australoid' population that had 'been modified by a considerable Dravidian element', evinced by the Dravidian languages spoken by the Gond populations. von Fürer-Haimendorf (1943, 1945, 1948, 1953) conducted studies of the Gond tribal groups and the closely

allied Dravidian linguistic communities, which led him to view these peoples as remnants of an earlier primordial population that had been linguistically assimilated.

Gond: The Largest Tribe of India

While surveying the Indian scriptures and the census data, we observed both, that the Gond population is mentioned in ancient scriptures and that they exhibit an exceptionally high population size, despite covering a limited geographical region (700 km² territory) (Ramadas, 1925; Chaubey *et al.*, 2015, 2017). Their languages, namely, Gōndi, Konda, Kūi, Kūvi, Pengo and Manda, are non-Indo-Aryan and, taken together, constitute a group of Dravidian languages coordinate with Telugu, with which they comprise the South-Central branch of the Dravidian language family. However, a recent molecular study characterised the Gonds as an ethnic group showing population genetic affinity with the language communities speaking languages of both language families, that is Dravidian and Austroasiatic (Basu *et al.*, 2016). The Gonds encompass a large number of endogamous groups that can be identified by different names, their spoken dialects, occupation, geographical location and varied states of socioeconomic development (Singh, 1997). A large amount of data exists on the analysis of classical genetic markers among various Gond tribes (Bhatia and Rao, 1986). Studies on Y chromosomal and mitochondrial deoxyribonucleic acid (mtDNA) markers have suggested that the Gonds share haplotypes both with the Austroasiatic and Dravidian-speaking populations (Baig *et al.*, 2004; Chaubey *et al.*, 2008a,b; Mittal *et al.*, 2008; Sharma *et al.*, 2012). However, these studies did not aim at providing high-resolution population structure, which is crucial to tracing migrations, admixture and the ethnic history of the Gond populations. Furthermore, it is known that the HbS (sickle haemoglobin) gene occurs with alarming frequencies amongst the Gonds and is highly variable amongst the subgroups that have coinhabited the same geographical regions for several hundred years (Joshi *et al.*, 1990; Rao and Gorakshakar, 1990; Balgir, 2012).

Let Us Recap

Technological advances have led to enormous improvements in the quality of the evidence that we are able to adduce. Our first study analysed hundreds of thousands of markers amongst the Gond and two other tribal populations, that is Kol and Bhil (Chaubey *et al.*, 2015). A combination of various statistical analyses revealed a massive amount of genome sharing amongst Bhil, Kol, Gond and with other ethnic groups of South Asian descent. As only one Gond group was used in this first study, we were unable to evaluate the question of pan-Gondi ancestry of all the Gond groups present in India. Therefore, we further compiled data of four Gond groups from diverse geographical locations (Chaubey *et al.*, 2017). We initially found that all the Gond groups were likely to share a common ancestry with a certain degree of isolation and differentiation. Strikingly, the analysis also pointed out the sharing of substantial genetic ancestry with the Indian Austroasiatic (i.e. Munda) groups rather than with

the other Dravidian groups with whom they share very close linguistic similarity (Chaubey *et al.*, 2017). This study explicitly ruled out that the Gond populations could represent a case of language shift from Austroasiatic to Dravidian on a large scale, keeping in mind their census size and genome sharing with Munda populations, but the study implied instead that the Gond represent distinct indigenous populations that evidently adopted Dravidian language in a separate gradual process of linguistic assimilation.

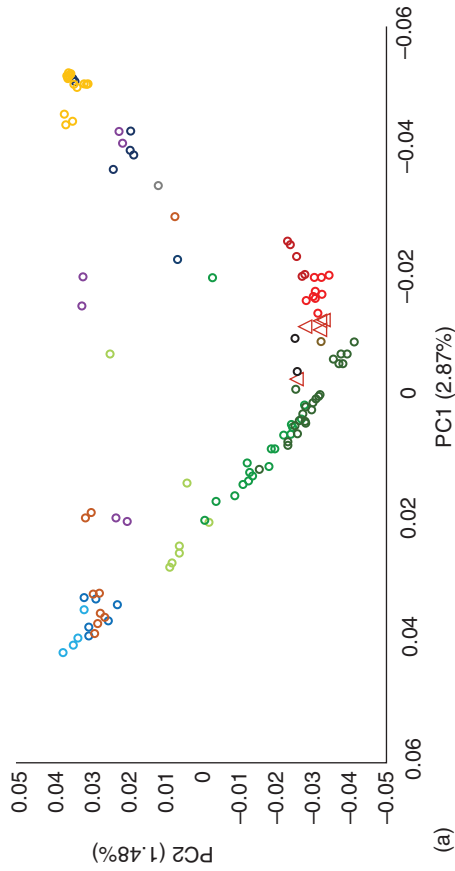
Revisiting the Autosomal Analysis

We took the advantage of available recent genomic data and statistical tools on the Gond and other Indian populations (Basu *et al.*, 2016) and revisited the genomic history of Gonds in the hope of shedding more light on this question. Using the newly compiled data, first we ran the principal component analysis (PCA) and plotted the population-wise mean values of PC1 versus PC2 (**Figure 1**). The plot showed that together with Austroasiatic (Munda) and Tibeto-Burman speakers, Indian transitional and Gond groups do not follow the Indo-European-Dravidian cline. The most likely reason for this pattern is their association with the East/Southeast Asian ancestry (Metspalu *et al.*, 2004; Chaubey *et al.*, 2011). The distance of these ethnic groups from the Indo-European-Dravidian cline is directly related with the proportion of East/Southeast Asian ancestry (**Figure 1**). Importantly, all the Gond groups (except Gond1) form their own distinct clusters which lie in between the Dravidian and Austroasiatic clusters. The self-clustering of the majority of Gond groups supports the common ancestry of the Gond (Chaubey *et al.*, 2017). The outlier nature of Gond1 is likely due to long-term isolation from a parental group, or a process of cultural assimilation, whereby the integration of other tribal groups into the Gond1 may have reshaped their genome, or both.

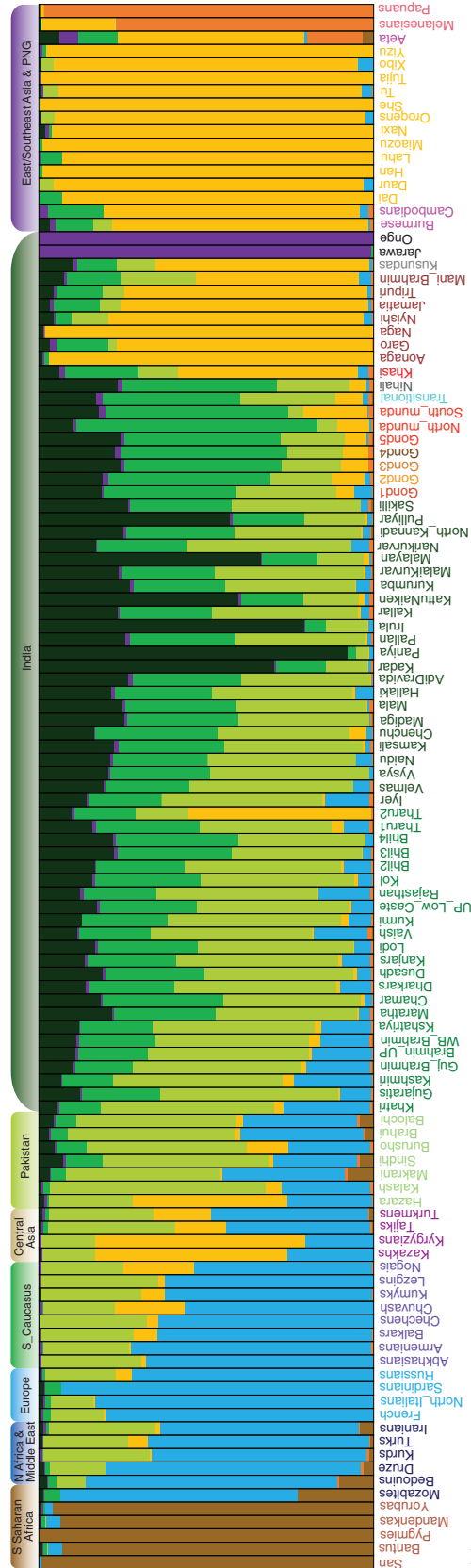
The ADMIXTURE analysis also supports the previous observation by showing the outlier nature of Gond1. Gond1 harbours the least East/Southeast Asian-specific ancestry in comparison with other Gond groups (**Figure 1**). The inclusion of new Indian samples showed an additional ancestry component, segregating from conventional ASI (Ancestral South Indian) component. These components were undifferentiated in previous studies and are collectively known as ASI component (Reich *et al.*, 2009; Chaubey *et al.*, 2011, 2017; Metspalu *et al.*, 2011). However, owing to overwhelming Indian samples as well as a smaller number of SNPs (Single Nucleotide Polymorphisms), this analysis was unable to differentiate the Middle Eastern and European ancestry components (see also: **The Archaeogenetics of European Ancestry**) (**Figure 1**). Therefore, it is worthwhile to consider a sampling balance for ADMIXTURE analysis.

The identity by state (IBS) analysis revealed the allele sharing variation of various Gond groups over the West–East geographical landscape (**Figure 2**). Gond1 shared more IBS with the West Eurasian populations, supporting their admixture with the Indo-European populations. This is also evident in their higher ANI (Ancestral North Indian) proportion (Chaubey *et al.*, 2017).

- Middle East
- Europe
- Caucasus
- Central_Asia
- Pakistan
- Indian Indo-European
- Dravidian
- Gond
- North_Munda
- South_Munda
- Transitional
- Nihali
- Indian Tibeto-Burman
- Khasi
- Kusunda
- Southeast Asia
- East Asia



(a)



(b)

Figure 1 (a) The principal component (PC) and (b) ADMIXTURE analyses of the combined autosomal data set. Both the analyses provided individual-wise information; however, for a simplified understanding, we have plotted the population-wise mean values, which were useful in understanding the clustering pattern of Gonds. The population colour codes in ADMIXTURE plot are the same as given in PC and IBS (Figure 2) plots.

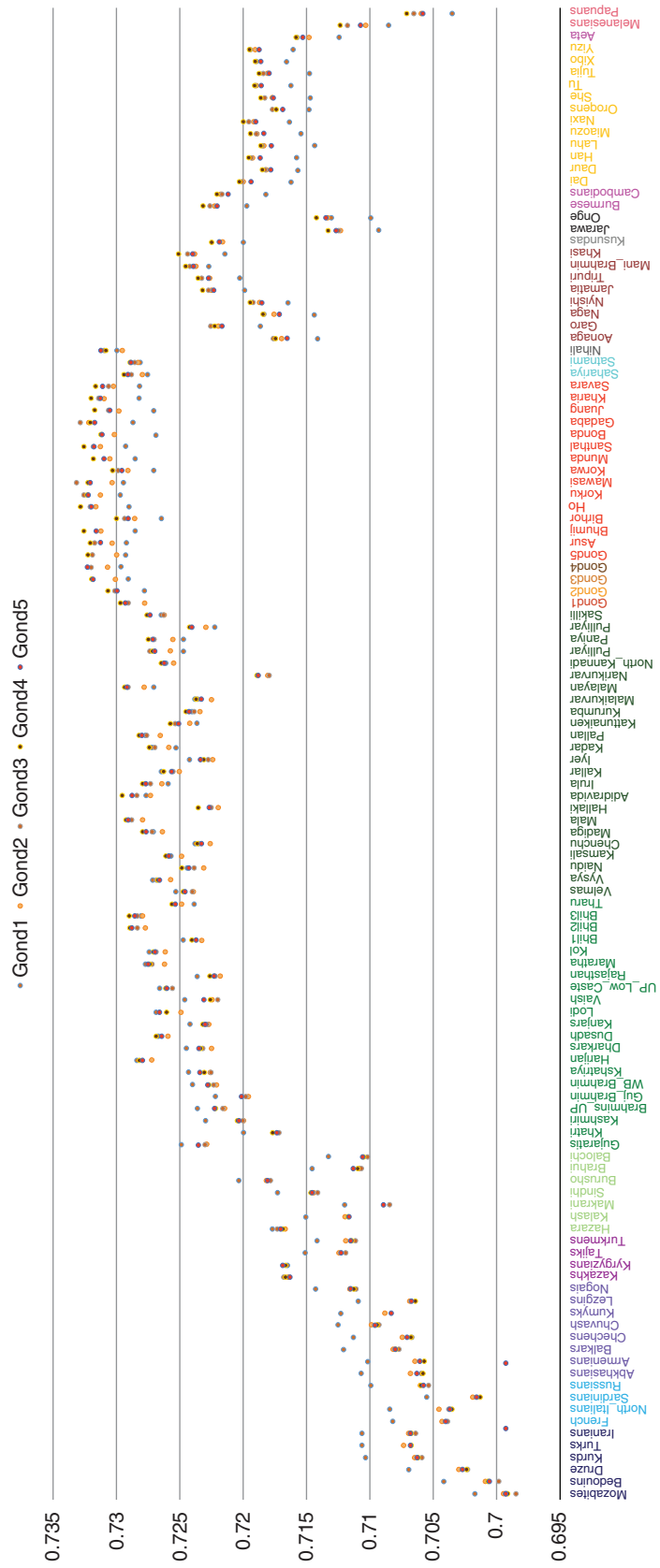
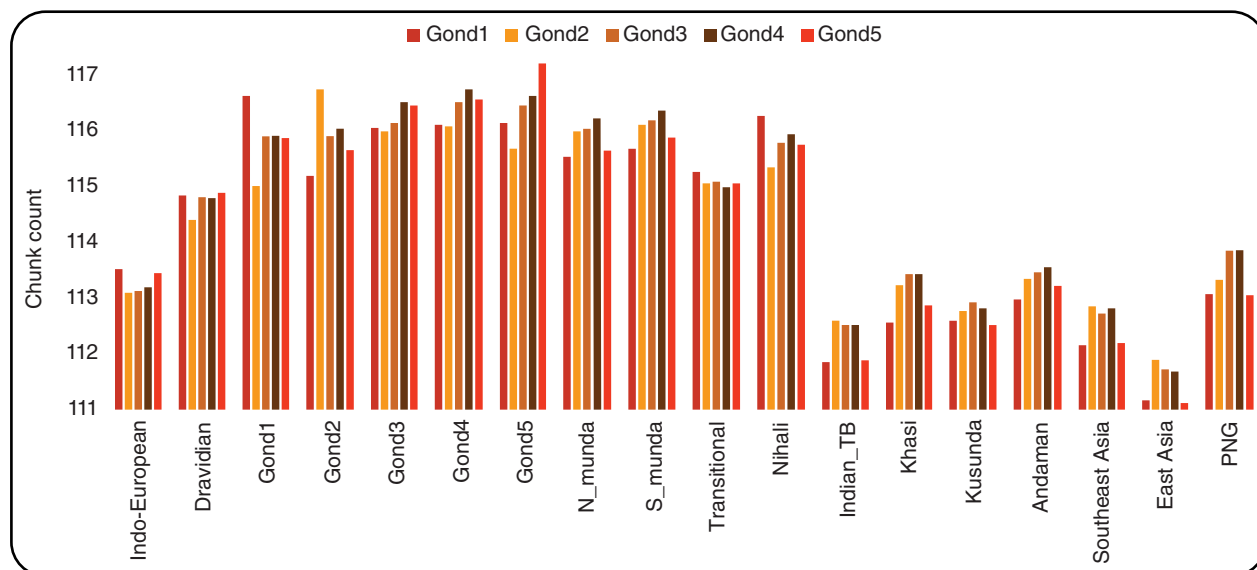


Figure 2 Pairwise IBS (identity by state) analysis showing the IBS distances of various Gond groups with the Eurasian populations.

Table 1 The *D* statistics showing the gene flow among Gond and surrounding populations

W	X	Y	Z	<i>D</i> value	Z score
Gond1	Yoruba	Kol	S_munda	-0.0044	-2.705
Gond2	Yoruba	Kol	S_munda	-0.0168	-10.145
Gond3	Yoruba	Kol	S_munda	-0.0166	-9.647
Gond4	Yoruba	Kol	S_munda	-0.017	-10.441
Gond5	Yoruba	Kol	S_munda	-0.016	-10.824
Gond1	Yoruba	Bhil	S_munda	-0.0053	-3.898
Gond2	Yoruba	Bhil	S_munda	-0.0168	-11.612
Gond3	Yoruba	Bhil	S_munda	-0.0167	-11.158
Gond4	Yoruba	Bhil	S_munda	-0.017	-12.286
Gond5	Yoruba	Bhil	S_munda	-0.0161	-12.774
Gond1	Yoruba	Nihali	S_munda	0.0004	0.151
Gond2	Yoruba	Nihali	S_munda	-0.0094	-3.748
Gond3	Yoruba	Nihali	S_munda	-0.0077	-3.054
Gond4	Yoruba	Nihali	S_munda	-0.0087	-3.441
Gond5	Yoruba	Nihali	S_munda	-0.0067	-3.002
Gond1	Yoruba	N_munda	S_munda	0.0018	1.984
Gond2	Yoruba	N_munda	S_munda	-0.0002	-0.211
Gond3	Yoruba	N_munda	S_munda	-0.0003	-0.303
Gond4	Yoruba	N_munda	S_munda	0	0.009
Gond5	Yoruba	N_munda	S_munda	-0.0007	-0.766

$D = (W, X; Y, Z)$.

**Figure 3** Plot of chunks (DNA segments) donated by the South, East, Southeast Asian and PNG populations to the Gond groups.

Gond3–5 were least differentiated from each other in the plot. To differentiate Gond from neighbouring tribal populations (Kol, Bhil, Nihali and Munda), we evaluated the level of gene flow (Table 1). We observed that most of the Gonds exhibit significant gene flow with Munda speakers when compared with other surrounding major populations. When compared with the Munda speakers, the Gonds remain largely equidistant from North and South Munda groups (Table 1).

In order to have a robust fine-grained population structure and admixture, we used haplotype and LD (linkage disequilibrium)-based method fineSTRUCTURE (Lawson *et al.*, 2012). In this analysis, we specifically looked at and compared the sharing of chunk (segment of deoxyribonucleic acid (DNA)) counts of the Gonds (Figure 3). We grouped populations based on their language and ethnicity, and estimated the number of chunks donated by them to the various Gond

groups. Of great importance is that, in a wider context, we have found that all the Gond groups shared a significantly higher (two-tailed p -value > 0.0001) number of chunks with Papuans and Melanesians than with East or Southeast Asians (**Figure 3**). It is likely that such sharing is due to the ancient genomic history of this region. Consistent with previous work (Chaubey *et al.*, 2017), North and South Munda groups were the topmost chunk donors to all the Gond groups (except Gond1), after excluding self-chunk contribution. Notably, for Gond1, a significantly (two-tailed p -value > 0.001) higher number of chunks were donated by the language isolate Nihali population (**Figure 3**). As both populations live in close vicinity, the gene flow beyond the boundaries of language might be the reason of such genomic sharing.

Sex-specific Markers Differentiate the Gond Significantly from Munda Speakers

The combination of autosomal analyses has suggested that, in spite of the fact that Gonds shared a large amount of genome with the Munda speakers, they still maintain their unique genomic identity. However, it is not clear to what extent this putative gene flow has shaped their maternal and paternal ancestries. Moreover, exploring their maternal (mtDNA) and paternal (Y chromosome) ancestries may help to evaluate the sex-specific admixture, which is remarkable and distinct among Munda speakers (Kumar *et al.*, 2007; Chaubey *et al.*, 2007, 2008a, 2011). The male-biased admixture associated with the overwhelming occurrence of Southeast Asian-specific Y chromosome haplogroup O2a-M95 is quite evident amongst Munda speakers. This was one of the robust pieces of evidence suggesting the Southeast Asian origin of Austroasiatic speakers (Chaubey *et al.*, 2011).

We have compiled and compared the mtDNA and Y chromosome data for the Gond, Dravidian (Telugu) and Austroasiatic (Munda) populations (**Table 2**). For maternal

ancestry, all three groups largely showed a similar composition of continental-specific haplogroups; however, for the Y chromosome, the Munda stand out due to a staggering amount of Southeast/East Asian-specific lineages. An extremely low frequency of Southeast/East Asian-specific lineages amongst the Gond groups supports our analysis based on autosomal data, which rejects the simpler model of language shift from Austroasiatic to Dravidian for the Gonds. Rather, although the Gond language communities are Dravidian, genetically the Gonds appear to represent an old distinct indigenous population. In fact, for their paternal ancestry, the Gonds have been shown to bear the highest frequency of South-Asian-specific Y chromosomal lineages (**Table 2**). Therefore, studying the mtDNA and Y chromosome among Gonds and their neighbours demonstrated that demographic events restricted to only one of the sexes can be of considerable significance in differentiating populations which experienced excessive gene flow in the recent times.

In conclusion, based on our extensive analysis of autosomal and haploid data, we hypothesise that before the arrival of Austroasiatic populations from Southeast Asia, the Gonds were already a distinct indigenous population. Just as the Munda biologically represent an indigenous local population which received the paternal lineage O2a-M95 together with their Austroasiatic father tongue, the Gonds speak Dravidian languages closely affiliated with Telugu, but biologically the Gonds likewise represent a locally indigenous population distinct from the linguistically related Telugu (Chaubey *et al.*, 2017). During the arrival and expansion of Austroasiatic speakers in India, the Gonds received gene flow from them. However, subsequent isolation helped them to maintain their unique identity from surrounding populations. Our results show that the ancestral pre-Dravidian Gond and the original pre-Austroasiatic Munda biologically reflect either two closely affiliated indigenous ancestral populations or perhaps one and the same older population native to this portion of east central India. As all the Gond groups carry the Southeast Asian signature, it is likely that the population differentiation and expansion of Gond with the neighbouring Munda populations continued after the expansion and admixture of Austroasiatic speakers in the sub-continent.

Table 2 mtDNA and Y chromosome haplogroup profiles among Gond and contemporary populations

	<i>n</i>	mtDNA				<i>n</i>	Y chromosome			
		South Asian	E/SE Asian	West Eurasian	Unresolved		South Asian	E/SE Asian	West Eurasian	Unresolved
Gond	88	71.6	0	3.4	25	102	70.6	8.8	4.9	15.7
Dravidian (Telugu)	374	70.9	0	8.8	20.3	508	54	5.3	13	27.7
Munda/Austroasiatic	732	75.4	0	3.6	21	1574	27.3	60.6	4.6	7.5

S-Asian haplogroups: mtDNA: M2-6, M33-65, R5-8 and R31-32; Y chromosome: C5, F, H, L and R2.

SE-Asian haplogroups: mtDNA: A-G, M7-12, R22 and N9; Y chromosome: C2, C3, D and M-O.

Unresolved haplogroups: M*, R*, N* including other lineages, for example M31 and West Eurasian specific; Y chromosome: C*, G, I-K*, P*, Q and R1.

mtDNA references: Roychoudhury *et al.* (2000); Kivisild *et al.* (2003); Cordaux *et al.* (2003); Metspalu *et al.* (2004) and references therein; Thanseem *et al.* (2006); Chaubey *et al.* (2008b).

Y chromosome references: Kivisild *et al.* (2003); Cordaux *et al.* (2004); Sengupta *et al.* (2006); Thanseem *et al.* (2006); Kumar *et al.* (2007); Chaubey *et al.* (2008b); Trivedi *et al.* (2008).

Acknowledgements

GC is supported by the National Geographic Society's grant for Research and Exploration (Grant No. HJ3-182R-18) and Estonian research grant IUT-24-1.

Glossary

Caste system Stratification of the society governed either by hereditary transmission or occupation.

Ethnographers Scientists who study the specific human culture, using methods such as interviews and close observations.

Haplogroup A group of likewise haplotypes derived from a common ancestor.

Haplotype Bunch of linked alleles present on specific DNA segment that are likely to be inherited together.

Language shift Cultural process where a population changes their language in a short period of time, without any major genetic alteration.

References

- Abbi A (2006) *Endangered Languages of the Andaman Islands*. Munich: Lincom Europa.
- Abbi A (2012) *Dictionary of the Great Andamanese Language*. Delhi: Ratna Sagar.
- Baig MM, Khan AA and Kulkarni KM (2004) Mitochondrial DNA diversity in tribal and caste groups of Maharashtra (India) and its implication on their genetic origins. *Annals of Human Genetics* **68**: 453–460.
- Balgir RS (2012) Community expansion and gene geography of sickle cell trait and G6PD deficiency, and natural selection against malaria: experience from tribal land of India. *Cardiovascular & Hematological Agents in Medicinal Chemistry* **10**: 3–13.
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **113**: 1594–1599.
- Berger H (1959) Die Buruškasi-Lehnwörter in der Zigeunersprache. *Indo-Iranian Journal III* **3**: 17–43.
- Bhatia H and Rao V (1986) *Genetic Atlas of the Indian Tribes*. New Delhi: Institute of Immunohaematology, Indian Council of Medical Research.
- Chaubey G, Metspalu M, Kivisild T and Villems R (2007) Peopling of South Asia: investigating the caste-tribe continuum in India. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* **29**: 91–100.
- Chaubey G, Karmin M, Metspalu E, et al. (2008a) Phylogeography of mtDNA haplogroup R7 in the Indian peninsula. *BMC Evolutionary Biology* **8**: 227.
- Chaubey G, Metspalu M, Karmin M, et al. (2008b) Language shift by indigenous population: a model genetic study in South Asia. *International Journal of Human Genetics* **8**: 41.
- Chaubey G, Metspalu M, Choi Y, et al. (2011) Population genetic structure in Indian Austroasiatic speakers: the role of landscape barriers and sex-specific admixture. *Molecular Biology and Evolution* **28**: 1013–1024.
- Chaubey G, Kadian A, Bala S and Rao VR (2015) Genetic affinity of the Bhil, Kol and Gond mentioned in epic Ramayana. *PLoS One* **10**: e0127655.
- Chaubey G, Tamang R, Pennarun E, et al. (2017) Reconstructing the population history of the largest tribe of India: the Dravidian speaking Gond. *European Journal of Human Genetics* **25**: 493–498.
- Cordaux R, Saha N, Bentley G, et al. (2003) Mitochondrial DNA analysis reveals diverse histories of tribal populations from India. *European Journal of Human Genetics* **3**: 253–264.
- Cordaux R, Aunger R, Bentley G, et al. (2004) Independent origins of Indian caste and tribal paternal lineages. *Curr Biol* **14**: 231–5.
- van Driem G (2001) *Languages of the Himalayas*. Leiden: Brill.
- von Fürer-Haimendorf C (1943) *The Chenchus: Jungle Folk of the Deccan*. London: Macmillan.
- von Fürer-Haimendorf C (1945) *The Reddis of the Bison hills: A Study in Acculturation-Aboriginal Tribes of Hyderabad*. London: Macmillan.
- von Fürer-Haimendorf C (1948) *The Raj Gonds of Adilabad: A Peasant Culture of the Deccan (The Aboriginal Tribes of Hyderabad)*. London: Macmillan.
- von Fürer-Haimendorf C (1953) New aspects of the Dravidian problem. *Tamil Culture* **2**: 127–135.
- von Heine-Geldern R (1917) Kopfjagd und Menschenopfer in Assam und Birma und ihre Ausstrahlungen nach Vorderindien. *Mitteilungen der Anthropologischen Gesellschaft in Wien* **XXXVII**: 1–65.
- von Heine-Geldern R (1928) Ein Beitrag zur Chronologie des Neolithikums in Südostasien. In: W. Koppers (ed.), *Festschrift Pater Wilhelm Schmidt: 76 sprachwissenschaftliche, ethnologische, religionswissenschaftliche, prähistorische und andere Studien*, pp. 809–843. Wien: Mechitaristen-Congregations-Buchdruckerei.
- von Heine-Geldern R (1932) Urheimat und früheste Wanderungen der Austronesier. *Vienne, Anthropol.* **XXVII**: 543–619.
- von Heine-Geldern R (1945) Prehistoric research in the Netherlands Indies. In: Honig P and Verdoorn F (eds) *Science and Scientists in the Netherlands Indies*, pp. 129–167. New York: Board for the Netherlands Indies, Surinam and Curaçao of New York City.
- Grigson WV (1938) *The Maria Gonds of Bastar*. London: Oxford University Press.
- Joshi S, Gorakshakar A, Mukherjee M, et al. (1990) Prevalence of HBsAg carriers among some tribes of Madhya Pradesh. *The Indian Journal of Medical Research* **91**: 340–343.
- Kivisild T, Rootsi S, Metspalu M, et al. (2003) The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *American Journal of Human Genetics* **72**: 313–332.
- Koppers W (1947) *Die Geheimnisse des Dschungels*. Luzern: Verlag Josef Stocker.
- Kuiper FBJ (1962) *Nahali: A Comparative Study*. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.
- Kumar V, Reddy AN, Babu JP, et al. (2007) Y-chromosome evidence suggests a common paternal heritage of Austro-Asiatic populations. *BMC Evolutionary Biology* **7**: 47.
- Lawson DJ, Hellenthal G, Myers S and Falush D (2012) Inference of population structure using dense haplotype data. *PLoS Genetics* **8**: e1002453.
- Man E (1885) On the Andaman islands and their inhabitants. *The Journal of the Anthropological Institute of Great Britain and Ireland* **14**: 235–272.
- Marambe AJW (1896) *The Veda Language*. Kandy: Clifton Press.

- Metspalu M, Kivisild T, Metspalu E, *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 Genetics* **5**: 26.
- Metspalu M, Romero IG, Yunusbayev B, *et al.* (2011) Shared and unique components of human population structure and genome-wide signals of positive selection in South Asia. *American Journal of Human Genetics* **89**: 731–744.
- Mittal B, Tripathy V, Aruna M, *et al.* (2008) Mitochondrial DNA variation and substructure among the tribal populations of Andhra Pradesh, India. *American Journal of Human Biology: the Official Journal of the Human Biology Council* **20**: 683–692.
- Nagaraja KS (2014) *The Nihali Language (Grammar, Texts and Vocabulary)*. Mysore: Central Institute of Indian Languages.
- Pinnow H-J (1963) The position of the Munda languages within the Austroasiatic language family. In: *Linguistic Comparison in South East Asia and the Pacific*, pp. 140–152. London: School of Oriental and African Studies.
- Ramadas G (1925) The aboriginal tribes in the Ramayana. *Man in India* **5**: 1–2.
- Rao V and Gorakshakar A (1990) Sick cell hemoglobin, beta-thalassemia and G6PD deficiency in tribes of Maharashtra, India. *Gene Geography: a Computerized Bulletin on Human Gene Frequencies* **4**: 131–134.
- Reich D, Thangaraj K, Patterson N, *et al.* (2009) Reconstructing Indian population history. *Nature* **461**: 489–494.
- Reinhard JG (1969) *Aperçu sur les Kusundâ, peuple chasseur du Népal*. *Objets et Mondes, Revue du Musée de l'Homme*, pp 89–106.
- Roychoudhury S, Roy S, Dey B, *et al.* (2000) Fundamental genomic unity of ethnic India is revealed by analysis of mitochondrial DNA. *Current Science* **79**: 1182–1192.
- Sengupta S, Zhivotovskiy LA, King R, *et al.* (2006) Polarity and temporality of high-resolution y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. *American Journal of Human Genetics* **78**: 202–221.
- Sharma G, Tamang R, Chaudhary R, *et al.* (2012) Genetic affinities of the central Indian tribal populations. *PLoS One* **7**: e32546.
- de Silva MWS (1972) *The Vedda Language: Texts and Lexicon*. Munich: Kitzinger.
- Singh KS (1997) *People of India*. Oxford, UK: Oxford University Press.
- Thanseem I, Thangaraj K, Chaubey G, *et al.* (2006) Genetic affinities among the lower castes and tribal groups of India: inference from Y chromosome and mitochondrial DNA. *BMC Genetics* **7**: 42.
- Trivedi R, Sahoo S, Singh A, *et al.* (2008) Genetic imprints of pleistocene origin of Indian populations: a comprehensive Phylogeographic sketch of Indian Y-chromosomes. *International Journal of Human Genetics* **8**: 97–118.
- Watters D (2006) *Notes on Kusunda Grammar: A language Isolate of Nepal*. Himalayan Languages Archive.

Further Reading

- Chaubey G, Metspalu M, Kivisild T and Villems R (2007) Peopling of South Asia: investigating the caste-tribe continuum in India. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* **29**: 91–100.
- Chaubey G (2010) *The Demographic History of India: A Perspective Based on Genetic Evidence*. <http://hdl.handle.net/10062/15240>
- Chaubey G, Govindraj P, Rai N, *et al.* (2017) The Genome-Wide analysis of the Bhils: the second largest tribal population of India. *Man in India* **95**: 279–290.
- Gita Press (1968). *Srimad Valmiki-Ramayana*. Gorakhpur: Gita Press.
- Gita Press (1961) *Purana*, translated in English by HH Wilson, Calcutta. Gita Press, Gorakhpur.
- Kivisild T, Rootsi S, Metspalu M, *et al.* (2003) The genetics of language and farming spread in India. In: Bellwood P and Renfrew C (eds) *Examining the Farming/Language Dispersal Hypothesis*, pp. 215–222. Cambridge, UK: The McDonald Institute for Archaeological Research.
- Kumar V and Reddy M (2003) Status of Austro-Asiatic groups in the peopling of India: an exploratory study based on the available prehistoric, linguistic and biological evidences. *Journal of Biosciences* **28**: 507–522.
- Moorjani P, Thangaraj K, Patterson N, *et al.* (2013) Genetic evidence for recent population mixture in India. *American Journal of Human Genetics* **93**: 422–438.