Molecular Genetic Study on the Status of Transitional Groups of Central India: Cultural Diffusion or Demic Diffusion?

Vikrant Kumar[1], A.N.S. Reddy[1], Pradeep Babu[1], T. Nageswar Rao[1], K. Thangaraj[2], A.G. Reddy[2], Lalji Singh[2] and B. Mohan Reddy[1*]

[1]*Molecular Anthropology Group, Biological Anthropology Unit, Indian Statistical Institute, Hyderabad, India and [2]Centre for Cellular and Molecular Biology, Hyderabad, India*

ABSTRACT Two different models of diffusion - demic and cultural - have been proposed as an explanation for the spread of languages. Recent studies have shown that in some cases the dispersal of the language was due to the demic diffusion while for others it is purely due to the process of acculturation. There are four major linguistic families in India which have largely their own geographic domain. However, there are a few situations in which the populations affiliated to different linguistic families cohabit. For example, we find the spread of the Indo-European and Dravidian tribal populations in the core/peripheral areas of the Mundari Austro-Asiatic groups. These non-Mundari groups have been termed as transitional populations to indicate that these populations originally were probably Mundaris. However, there has been no attempt to ascertain if these populations are genetically Austro-Asiatics or do they belong to the linguistic groups that they are currently affiliated to. To examine this we have analysed Y-SNPs and STRs data of the 13 Mundari and 7 transitional groups and compared with the other populations of relevant linguistic groups. The results suggest that the Indo-European transitional groups are genetically Mundari and have acquired the present language through the process of cultural diffusion, while in the case of Dravidian transitional groups, the spread of language seems to be due to the process of both, the demic and cultural diffusion.

INTRODUCTION

The spread of culture and language in human populations is explained by two alternative models (Fig. 1): the demic diffusion model requiring mass movement of people and the cultural diffusion model or acculturation, which refers to dispersal of cultural traits between populations entailing limited or no genetic exchange between them (Cavalli-Sforza et al. 1994; Wen et al. 2004). A striking correlation has been found between linguistic and genetic diversity suggesting that linguistic variation could account for most of the genetic diversity (Cavalli-Sforza et al. 1994). This implies that language has been spread by the process of demic diffusion rather than the process of acculturation and this has been unambiguously demonstrated by two recent genetic studies (Cordaux et al. 2004a; Wen et al. 2004). Wen et al. (2004) have shown the spread of culture and language through the demic diffusion of Han populations while Cordaux et al. (2004a) suggested the spread of agriculture in India through the demic diffusion

of South Indian caste populations. However, there are few studies which invoke the process of cultural diffusion to explain the replacement of language in the region of Caucasus (Nasidze et al. 2001, 2003, 2006). The above studies suggest that both demic and cultural diffusion processes have led to the spread of culture and language, and different regions of the world have been perhaps affected by one of the processes.

There are four major linguistic families in India which have, by and large, their own non-overlapping geographic domain. For instance, the populations speaking the language affiliated to Dravidian linguistic family are largely confined to Southern India where as the Indo-European and Tibeto-Burman groups are mostly distributed in North and East India and, Northeast India, respectively. Reddy et al. (2005) based on autosomal Short Tandem Repeat (STR) loci have shown that populations of India have their own linguistic and geographic clusters while Kumar et al. (2007) based on Y-chromosome have suggested a separate genetic identity of the Austro-Asiatic groups, implying that these linguistic groups have their own genetic characteristics. However, there are certain populations which co-inhabit or have overlapping geographic distribution with the populations affiliated to a different linguistic family, other than

Corresponding Author's Address: Dr. B. Mohan Reddy
Molecular Anthropology Group, Biological Anthropology Unit, Indian Statistical Institute, Street No. 8, Habsiguda Hyderabad 500 007, Andhra Pradesh, India
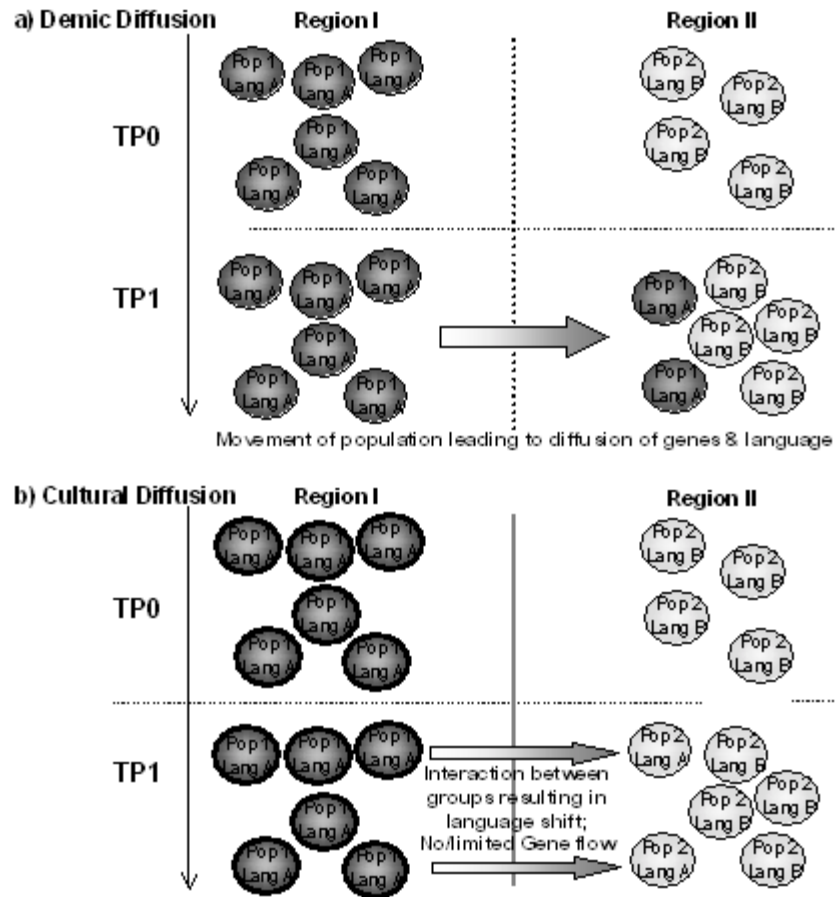*E-mail:* bmrisi@gmail.com

Fig. 1. A schematic representation of the demic and cultural diffusion models. 'Pop', Population; 'Lang', Language. 'TP', Time Period. In b) circle with thick and thin lines are dominant and weaker social groups, respectively. Vertical broken and solid lines indicate movement and no movement of genes, respectively.

their own. For example, tribes speaking Dravidian and Indo-European languages inhabit the Central and Central-Eastern region (Chota-Nagpur area) which is dominated by the tribes speaking Mundari languages (a sub-linguistic branch of Austro-Asiatic family). The spread of the Indo-European and Dravidian populations in the realm of Mundari tribes is enigmatic and needs to be understood in the background of demic and cultural diffusion models.

Kumar et al. (2003, 2007) have described the Indo-European and Dravidian populations in the region populated mainly by the Mundari populations as transitional groups indicating that these populations were probably originally

Mundari whereas Thangaraj et al. (2006) have designated the Dravidian speaking Oraon populations as Mundari. However, there has been no specific attempt to explore if these populations are genetically Austro-Asiatics, or do they belong to the linguistic groups that they are presently affiliated to or are they a conglomeration of both. Although Kumar et al. (2003) based on traditional genetic and anthropometric markers studied some of the transitional populations the results, at the best, can be tentative. Based on molecular genetic data we examine here the genetic affinities of 3 Dravidian and 4 Indo-European speaking transitional populations with the Dravidians and Indo-Europeans of South

and North India, respectively, vis-a-vis their genetic affinity with the geographically proximate Mundari populations, to infer the genetic status of transitional groups. Given the patrilocality of Indian populations, despite strict endogamy a possibility of female gene flow between these groups (Bamshad et al. 1998; Kivisild et al. 2003) and negligible male gene flow across the ethnic groups of India (Bhattacharyya et al. 1999) is expected. Therefore, we used Y-Single Nucleotide Polymorphisms (SNPs) and STRs so that there is minimal genetic exchange between the Mundari and North Indian Indo-European and South Indian Dravidian groups, which may help us in correctly identifying the parental genetic source of these transitional groups.

MATERIAL AND METHODS

Sampling: Intravenous blood samples were collected from a total of 994 unrelated males belonging to 20 tribal populations out of which 13 are Mundari groups and 7 are transitional groups which are considered to have certain

geographical and historical affinities with the Austro-Asiatic groups and were also presumed to be formerly Austro-Asiatic (for population detail see Kumar et al. 2007). Of the 7 transitional groups, 3 are speakers of Dravidian languages (Oraon, Nagesia and Paharia) and 4 speak Indo-European languages (Bathudi, Bhuiyan, Kanwar and Pando). To investigate if the transitional populations were originally Mundari speakers (based on Y-STRs), we compared them with 3 relevant sets of populations; 2 Indo-European language speaking Jat populations (Jat-Sikh and Jat-Haryana) from Northwest India (Nagy et al. 2006), 5 (Akuthota, Kapu, Pokanati, Panta and vanne) and 2 (Chenchu and Koya) Dravidian caste (Kumar et al. 2006) and Tribal (Kivisild et al. 2003) populations, respectively, from Andhra Pradesh (South India). The population details are furnished in Table 1 and the region of sampling is depicted in Figure 2. We also examined the relationship of transitional population based on Y-SNPs and the details of the SNPs are given at Kumar et al. (2007). We chose both tribal and caste populations from

Fig. 2. Map of India showing the areas of sampling. The symbols represent the regions of sampling of the different linguistic groups.

Table 1: Number of populations along with the sample size, number of Y-STRs available along with the geographic distribution of linguistic groups.

Linguistic groups	No. of populations	Total sample size (Range)	No of Y-STRs available
Mundari	13	695 (24-109)	16
Dravidian (Transitional)	3	116 (11-91)	16
Indo-European (Transitional)	4	183 (23-83)	16
Dravidian (caste)	5	106 (16-25)	6
Dravidian (Tribal)	2	43 (21-22)	6
Indo-European	2	164 (84-80)	10

Southern India because they have been shown to be genetically distinct (Cordaux et al. 2004a, b). Further, since there has been no evidence of these comparative populations having any genetic connection or admixture with Mundari, the close genetic affinity, if any, found between them and the transitional populations would be only because of the common genetic heritage or due to extensive gene flow. To distinguish between the transitional populations from the comparative populations we shall refer them as Indo-European transitional populations and Dravidian transitional populations.

Genetic Analysis: We have typed the following 16 Y-Short Tandem Repeats (STRs): DYS19, DYS388, DYS389I, DYS389b, DYS390, DYS391, DYS392, DYS393, DYS426, DYS437, DYS438, DYS439, DYS447, DYS448, DYS460 and H4. Y-STRs were amplified by multiplex PCR (Butler et al., 2002) and were analysed on ABI 3730 sequencer. The fragment sizes of the alleles were converted into repeat units as suggested by Butler et al. (2002). Allele length for DYS389b, were obtained by subtracting the allele length of DYS389I from DYS389II. Since the number of STRs typed for the comparative populations were only a subset of what was typed in the Mundari and the transitional populations, we also made subset of STRs for the purpose of comparison. For example, data for only 10 out of the 16 STRs we have typed for Austro-Asiatic groups were available for the Indo-European populations. Therefore, we used same set of 10 STRs for our populations. We used similar sub sets of Y-STRs for the Dravidian populations.

Statistical Analysis: The genetic distances for STRs were calculated using the R_{ST} statistics as implemented in ARLEQUIN (Excoffier et al. 2005). This statistic is based on the squared differences in allele sizes and utilizes the additional information content from allele-size variance in the STR data set (Slatkin 1995). Based on the genetic distance matrix, Neighbor-Joining

(NJ) tree and Multidimensional (MDS) plot were drawn. Further, the genetic structure was explored by the analysis of molecular variance (AMOVA), as implemented in ARLEQUIN (Excoffier et al. 2005). We also performed admixture analysis using the ADMIX 1.0 software (Bertorelle and Excoffier 1998) to estimate the amount of genetic admixture in the transitional groups using the Mundari and Indo-European populations as the parental groups for the Indo-European transitional groups, and Mundari and Caste/tribal Dravidian groups as the parental groups for Dravidian transitional populations. We computed the estimator mY described in Bertorelle and Excoffier (1998) along with the estimators of the admixture coefficient mC (Chakraborty et al. 1992) and mR (Robert and Hiorns 1965). The main difference between mY and the other two estimators is that mY also takes into account the molecular distance between the different alleles (Bertorelle and Excoffier 1998).

RESULTS

Distribution of Y-chromosome Haplogroups: The group-wise distribution of Y-haplogroup frequency is furnished in Table 2. Of all the haplogroups, the frequency of O-M95 is highest (62%) in Indo-European transitional groups followed by H-M69 (22%), similar to the average frequency found for O-M95 (61%) and H-M69 (26%) as the two most common haplogroups in the Mundari groups. However, the populations under Indo-European group did not show any O-M95 haplogroup and show a very low frequency of H-M69 (7%). Further, the most common haplogroup in the Indo-European group is R-M173 (49%) which is quite low (7%) in the transitional populations of Indo-European category. In the Dravidian populations of transitional category, the most common haplogroup is H-M69 (55%) followed by O-M95 (35%), where as the frequency of O-M95 is low in the caste

Table 2: Frequency distribution of Y-chromosome haplogroups for different linguistic groups.

Linguistic groups	No. of populations	C-RPS4Y	DE-YAP	F-M89	H-M69	J-M172	K-M9	L-M20	O-M95	P-M45	R-M173
Indo-European Transitional[1]	4	0.00	0.00	0.05	0.22	0.07	0.00	0.00	0.62	0.01	0.03
Dravidian Transitional[1]	3	0.00	0.00	0.02	0.55	0.01	0.00	0.00	0.35	0.01	0.06
Mundari[1]	13	0.00	0.00	0.03	0.26	0.02	0.00	0.00	0.61	0.02	0.07
Indo-European[2]	4	0.05	0.00	0.04	0.07	0.17	0.02	0.11	0.00	0.05	0.49
Dravidian Caste[3]	7	0.06	0.00	0.12	0.15	0.10	0.07	0.22	0.01	0.10	0.18
Dravidian Tribe[4]	18	0.06	0.00	0.12	0.29	0.08	0.03	0.04	0.08	0.13	0.13

Comparative Data from [1]Kumar et al. (2007); [2]Kivsild et al. (2003); [3]Ramana et al. (2001), Wells et al. (2001), Cordaux et al. (2004b); [4]Ramana et al. (2001), Kivsild et al. (2003), Cordaux et al. (2004b), Thanseem et al. (2006).

(1%) and moderate in tribal (8%) populations in the non-transitional category. Although, the frequency of H-M69 is comparably high (29%) in the Dravidian tribes, it has haplogroups such as C-RPS4Y (6%) and P-M45 (13%) with reasonable frequency, which are almost absent in the Dravidian transitional group.

Genetic Relationship Based on MDS and N-J Tree: Based on the Pairwise values of R_{ST} distances computed on Y-STR frequencies, both the MDS and NJ tree were drawn for different sets of populations. The MDS plot of Mundari, transitional Dravidian and Dravidian caste populations (Fig. 3) shows a good fit between the two-dimensional MDS plot and the source data (Pairwise value of R_{ST}), as evident by the stress value of 0.055. The MDS and also the NJ tree (fig. not shown) suggest two distinct clusters one formed by the Dravidian caste populations and the other by most of the Mundari and the transitional Dravidian populations except Paharia, suggesting that caste Dravidians have negligible genetic affinity with the transitional Dravidian populations. However, another MDS plot (stress value: 0.059) based on Mundari, transitional Dravidian and Dravidian tribal populations (Fig. 4) does not show separate clusters of Dravidian tribal and transitional groups. Although Koya

Fig. 3. Plot of Mundari, Dravidian transitional and caste populations on the first two Dimensions derived from the multidimensional scaling of the R_{ST} distance based on Y-STRs. Circle diamond and square denotes Dravidian castes, Mundari and Dravidian transitional groups, respectively

population is an outlier**,** the position of Oraon vis-à-vis with Chenchu and the Mundari populations suggest some genetic affinity between them, as also evident from the N-J tree where the status of Oraon is ambiguous (fig. not shown). It is also apparent from both the plots that the Dravidian transitional populations neither form their own cluster nor are distinctly separated from the Mundari populations except the Paharia.

The N-J tree (Fig. 5) and the MDS plot (not shown) based on the Indo-European, Mundari and the transitional Indo-European populations shows unequivocal and clear separation of Indo-European populations from the rest of the populations. Further, like the transitional Dravidian populations, the Indo-European transitional populations neither form their own cluster nor separated from the Mundari populations.

Analysis of Molecular Variance and Admixture Analysis: The Analysis of Molecular Variance (AMOVA) based on Y-STRs (Table 3) suggests that Indo-European and transitional Indo-European populations are well differentiated ($F_{ST} = 0.65$). However, the F_{ST} value turns out to be zero between Mundari and the

Table 3: Analysis of molecular variance based on Y-STRs

Groups	F_{ST}
Mundari Vs TTDR*	0.01
Mundari Vs TTIE**	0.00
IE Vs TTIE	*0.65*
DR (caste) Vs TTDR	*0.27*
DR (tribe) Vs TTDR	0.12

*TTDR, Transitional Dravidian populations;
**TTIE, Transitional Indo-European populations; all the p-values are >0.05 except for the italicized and underlined.

transitional Indo-European populations, suggesting no genetic differences between them. Similarly, the differentiation between the Mundari and the transitional Dravidian population is virtually nil (0.01) as compared to that between the transitional Dravidian and caste Dravidian groups with a high F_{ST} value (0.27). However, among group differentiation between the transitional Dravidian group and the tribal Dravidian group is high but not significant (0.12).

The admixture analysis (Table 4) of transitional Indo-European populations considering Mundari and Indo-European groups as parental populations suggests no genetic contribution
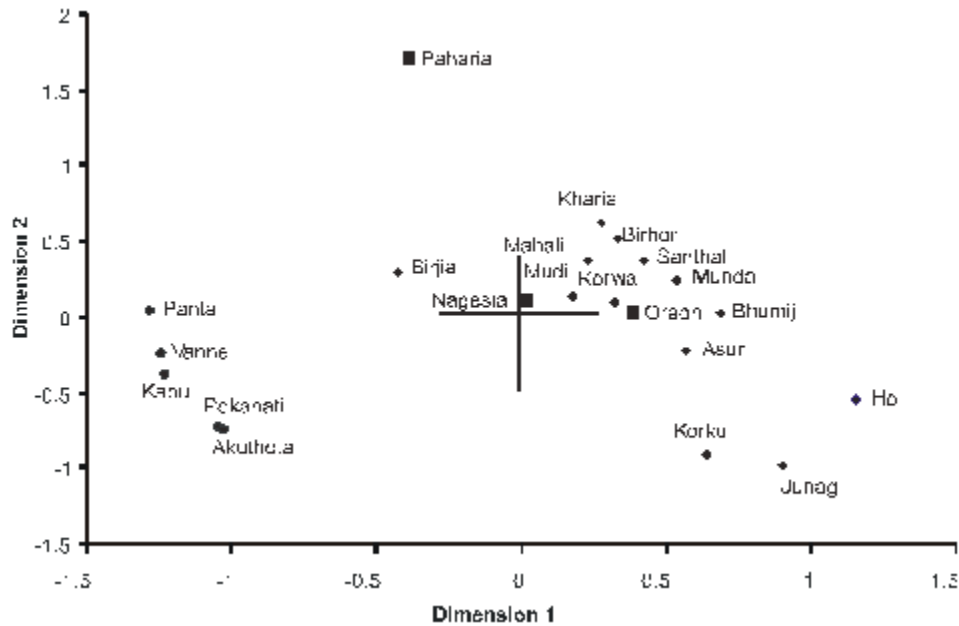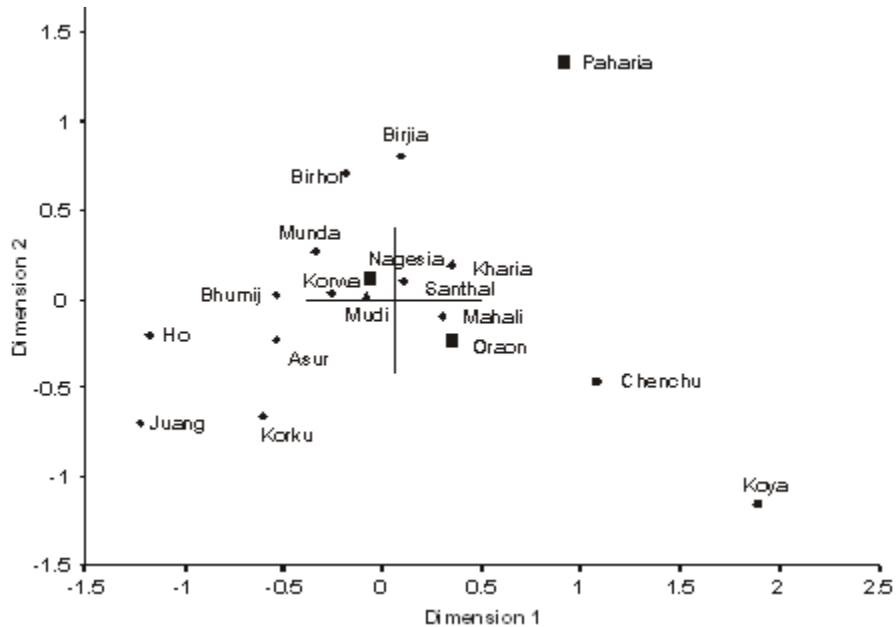

Fig. 4. Plot of Mundari, Dravidian transitional and tribal populations on the first two Dimensions derived from the multidimensional scaling of the R_{ST} distance based on Y-STRs. Circle diamond and square denotes Dravidian tribe, Mundari and Dravidian transitional groups, respectively

Fig. 5. N-J tree of Mundari, Indo-European and transitional Indo-European populations based on the R_{ST} distance of Y-STRs. Circle and square denotes Indo-Europeon caste & transitional groups, respectively

or gene flow from the Indo-European populations. In fact, the estimates of all admixture coefficients were similar suggesting that transitional Indo-European populations had only Mundari as parental source population. The analysis of transitional Dravidian groups with Mundari and Dravidian caste groups as parental populations suggest a low contribution of ~13% (except mR which suggest 25%) from the Dravidian caste groups. However, when the analysis was repeated using Dravidian tribal group as one of the parental population the contribution from this was quite high (~30-40%). Overall, the average genetic constitution of the transitional Dravidian groups still reflects predominantly Mundari in origin (75%).

DISCUSSION

We try to ascertain here if the transitional Indo-European and Dravidian populations are genetically Mundari or do they reflect genetic parentage of the linguistic groups that they are presently identified with. Alternatively, do they represent genetic conglomeration of heterogeneous units, derived from both the sources. For this purpose we analyzed Y-chromosome SNP and STR data of Mundari, transitional tribes and the relevant comparative data. Overall, the analysis based on both Y-SNPs and STRs strongly suggests that Indo-European transitional groups (Bathudi, Bhuiyan, Kanwar and Pando) are derived from the Mundari source populations. The haplogroups found among the Indo-European transitional groups are almost similar to what is found among the Mundari tribes (Table 2) where as they are quite different from the Indo-European Jat populations, which show absence of O-M95 and presence of high average frequency of R-M173. The high average frequency of O-M95 and H-M69 and negligible frequency of R-M173 among both the Mundari and transitional Indo-Europeans suggests that Indo-European transitional populations are genetically very similar to Mundari populations. The results of AMOVA (Table 3), N-J tree (Fig. 5)

Table 4: Admixture analysis of transitional populations based on three different admixture statistics using two parental populations

Admixed group	Contribution from Parental populations (S.D.)[$]					
	Mundari			Group 2		
	m R	m C	m Y	m R	m C	m Y
Transitional Indo-European[1]	1.00 (0.03)	1.00 (0.04)	1.00 (0.01)	0.00	0.00	0.00
Transitional Dravidian[2]	0.77 (0.04)	0.87 (0.06)	0.88 (0.04)	0.23	0.13	0.12
Transitional Dravidian[3]	0.71 (0.04)	0.71 (0.07)	0.60 (0.05)	0.29	0.29	0.40

[1]Parental Group 2 is Indo-Europeans; [2]Parental Group 2 is Caste Dravidians; [3]Parental Group 2 is tribal Dravidians. The Parental group 1 in each of the above three analyses is Mundari tribes. [$]S.D. based on 1000 bootstraps.

and the admixture analysis (Table 4) also conform to the above pattern. These findings may suggest that the Indo-European transitional groups were earlier the speaker of Mundari language and there has been, at a later point of time, shift/ replacement in their language from Mundari to Indo-European. This is consistent with the 'elite dominance' model suggested by Renfrew (1992, 2000) in the mosaic-linguistic zones (Fig. 1). In this process, which is a specific case of accul-turation, the language spoken by the popu-lations practicing agricultural economy is considered to be dominant and is adopted by the erstwhile hunter-gatherer populations. In the present case, the Indo-European transitional groups who were traditionally hunter-gatherers adopted the language of neighboring Indo-European agricultural groups without the involvement of gene flow between them. Taken as a whole, the genetic analysis reflects a shift in the language without any exchange of gene(s) between the groups.

On the other hand, the Dravidian transitional groups do not show a similar picture. Both the Dravidian transitional and the Mundari groups have a high average frequency of H-M69 besides the relatively high frequency of O-M95. However, H-M69 is also found to be high in the Dravidian tribal groups suggesting that the Dravidian transitional groups perhaps have a genetic connection with the Dravidian tribal groups. The MDS plot (Fig. 4) and the N-J tree suggests that out of the 3 Dravidian transitional groups, it is the Oraon population which shows a genetic affinity with the Chenchu, while Nagesia seems to be genetically Mundari. The position of Paharia as an outlier seems to be intriguing and it may be because of small sample size (n = 11) resulting in ascertainment bias. The results of AMOVA (Table 3) although shows a relatively high among group differentiation between them ($F_{ST} = 0.12$), the differentiation is not significant, whereas the differentiation between transitional Dravidian and Mundari groups is negligible ($F_{ST} = 0.01$). However, when the AMOVA was done based on only the Oraon and the Dravidian tribal populations the differentiation was significantly lower (0.04) than that obtained for Dravidian transitional and tribal groups, whereas between the Nagesia and the Dravidian tribal population was quite large (0.21). This strongly indicates that the Oraon tribe probably had duel genetic origin, both from the Mundari and the tribal Dravidian populations. The admixture analysis also suggests a high degree of genetic overlap (~30-40%) between the Dravidian tribes and the transitional Dravidian groups. When the analysis was repeated based on only the Oraon population the estimate of admixture coefficient increases substantially by 20-30% (except for mY which surprisingly decreases). The foregoing analyses suggest that not all the Dravidian transitional populations are genetically purely Mundari, unlike in the case of Indo-European transitional populations. Among the transitional Dravidian tribes, only Nagesia, which does not show any genetic affinity with the Dravidian tribal or caste groups, conform to the elite dominance model of language shift. The Oraon population seems to be either genetically a Dravidian tribe or had an extensive admixture with the Dravidian tribal populations. Given that there are no Dravidian tribal populations in the vicinity form which gene flow could have occurred into the Oraons, there is a strong possibility that this population is originally Dravidian population with extensive gene flow from the neighboring Mundari groups and in the process might have acquired O-M95 haplogroup from them. This argument is also consistent with the absence of O-M95 among the Chenchu and the Koya. In fact, O-M95 is present only in 5 populations (Ramana et al. 2001, Thanseem et al. 2006) out of the 18 Dravidian tribal groups studied so far. Even in these 5 populations, 2 of them are geogra-phically close to the region inhabited by the Mundari populations and probably they may have acquired O-M95 from them.

Bellwood (2001) and Renfrew (1996) had suggested a farming/language dispersal model where by the languages have spread because of the farming dispersals, generally through the expansion of populations of farmers by a process of colonization or demic diffusion. Given that the Dravidian populations are considered to be the farming populations, which have migrated from Central Asia to South India (Cavalli-Sforza et al. 1994), some of the Dravidian groups might have, while migrating to southern India, stayed in the Central-Eastern India (the Chota-Nagpur region) and over a period of time had extensive gene flow from the neighboring Mundari populations. Overall, the results suggest that the Indo-European transitional groups are genetically Mudari and have acquired the present language, while in the case of Dravidian transitional groups,

the spread of language seems to be due to the process of both, the demic and cultural diffusion (Fig. 1).

REFERENCES

Bamshad MJ, Watkins WS, Dixon ME, Bhaskara BR, Naidu JM et al. 1998. Female gene flow stratifies Hindu castes. *Nature,* **395:** 651–652.

Bellwood P 2001. Early agriculturalist population diasporas? farming, languages and genes. *Annual Review of Anthropology,* **30**: 181-207.

Bertorelle G, Excoffier L 1998. Inferring admixture proportions from molecular data. *Mol Biol Evol,* **10**:1298-1311.

Bhattachayya NP, Basu P, Das M, Pramanik S, Banerjee R et al. 1999. Negligible male gene flow across ethnic boundaries in India, revealed by analysis of Y-chromosomal DNA polymorphisms. *Genome Res,* **9:** 711–719.

Butler JM, Schoske R, Vallone PM, Kline MC, Redd AJ et al. 2002. A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. *Forensic Sci Int,* **129**: 10-24.

Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. *The history and geography of human genes.* Princeton, NJ: University Pres).

Chakraborty R, Kamboh MI, Nwankwo M, Ferrell RE 1992. Caucasian genes in American blacks: new data. *Am J Hum Genet*, **1**: 145-155.

Cordaux R, Deepa E, Vishwanathan H, Stoneking M 2004a. Genetic evidence for the demic diffusion of agriculture to India. *Science*, **304**: 1125.

Cordaux R, Aunger R, Bentley G, Nasidze I, Sirajuddin SM et al. 2004b. Independent origins of Indian caste and tribal paternal lineages. *Curr Biol,* **14**: 231-235.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1**: 47-50.

Kivisild T, Rootsi S, Metspalu M, Mastana S, Kaldma K et al. 2003. The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *Am J Hum Genet,* **72**: 313-332.

Kumar V, Reddy BM 2003. Status of Austro-Asiatic groups in the peopling of India: An exploratory study based on the available prehistoric, Linguistic and Biological evidences. *J Biosci,* **28:** 507-522.

Kumar V, Langstieh BT, Madhavi KV, Naidu VM, Singh HP et al. 2006. Global patterns in human mitochondrial DNA and Y-chromosome variation caused by spatial instability of the local cultural processes. *PLoS Genet,* **24**: e53.

Kumar V, Reddy AN, Babu JP, Rao TN, Langstieh BT et al. 2007. Y-chromosome evidence suggests a common paternal heritage of Austro-Asiatic populations. *BMC Evol Biol,* **7**: e47.

Nagy M, Henke L, Henke J, Chatthopadhyay PK, Volgyi A et al. 2006. Searching for the origin of Romanies: Slovakian Romani, Jats of Haryana and Jat Sikhs Y-STR data in comparison with different Romani populations. *Forensic Sci Int,* (In press).

Nasidze I, Stoneking M. 2001. Mitochondrial DNA variation and language replacements in the Caucasus. *Proc Biol Sci,* **7**: 1197-206.

Nasidze I, Sarkisian T, Kerimov A, Stoneking M 2003.Testing hypotheses of language replacement in the Caucasus: evidence from the Y-chromosome. *Hum Genet,* **3**: 255-261.

Nasidze I, Quinque D, Rahmani M, Alemohamad SA, Stoneking M 2006. Concomitant replacement of language and mtDNA in South Caspian populations of Iran. *Curr Biol,* **16**: 668-673.

Ramana GV, Su B, Jin L, Singh L, Wang N et al. 2001. Y-chromosome SNP haplotypes suggest evidence of gene flow among caste, tribe, and the migrant Siddi populations of Andhra Pradesh, South India. *Eur J Hum Genet,* **9**: 695-700.

Reddy BM, Naidu VM, Madhavi VK, Thangaraj LK, Kumar V et al. 2005. Microsatellite diversity in Andhra Pradesh, India: genetic stratification versus social stratification. *Hum Biol*, **6**: 803-823.

Renfrew C 1992. World languages and human dispersals: a minimalist view. In: JA Hall and I.C. Jarvie(eds.): *Transition to Modernity, Essays on Power, Wealth and Belief*. Cambridge: Cambridge University Press pp. 11-68.

Renfrew C 1996. Language families and the spread of farming. In: DR Harris (ed.): *The Origins and Spread of Agriculture and Pastoralism in Eurasia*. London: University College London Press pp. 70-92.

Renfrew C 2000. At the edge of knowability: towards a prehistory of languages. *Cambridge Archaeological Journal,* **10(1):** 7-34.

Robert DF, Hiorns RW 1965. Methods of analysis of the genetic composition of a hybrid population. *Hum Biol*, **37**: 38-43.

Slatkin M 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics,* **139**: 457-462.

Thangaraj K, Chaubey G, Singh VK, Vanniarajan A, Thanseem I et al. 2006. In situ origin of deep rooting lineages of mitochondrial Macrohaplogroup 'M' in India. *BMC Genomics,* **7**: e151.

Thanseem I, Thangaraj K, Chaubey G, Singh VK, Bhaskar LV 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.

Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I et al. 2001. The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci USA,* **98**: 10244-10249.

Wen B, Li H, Lu D, Song X, Zhang F et al. 2004. Genetic evidence supports demic diffusion of Han culture. *Nature,* **431**: 302-305.