

Pattern of Y-chromosome and mtDNA Diversity in Northern Chinese Populations

Song Xiufeng

(Center for Anthropological Studies at Fudan University, Shanghai 200433)

Abstract

19 SNPs within the non-recombination region of Y chromosome were genotyped in a total of 884 male individuals including 11 northern populations, 6 southern populations and 5 ethnic groups from northwestern China. HVS I of mtDNA were sequenced and haplogroups A—G as well as 9bp deletion in CO II/tRNA^{ly} were genotyped in all samples. For dating SNP lineages, the following microsatellite: DYS19, DYS388, DYS389 and DYS392 were analyzed. Principal component maps were drawn with northern- southern populations and northern-Central Asian populations respectively. The first principal component distinguishes the east and west, and the second principal component distinguishes the north and south in both PC maps, which show clear correlation between genetic data and geographical distances. In these two PC maps, the first two principal components account for more than half of the variation in haplotype frequency, and thus represent the most important influences on patterns of Y chromosome diversity in the populations studied here. To further analyze the genetic pattern of 4 Altaic ethnic populations and 1 Tibetan population, another three-dimensional PC map was drawn to demonstrate the correlation between genetic data and linguistic classification. Results show that there are clear divergence between Altaic and Sino-Tibetan populations. And among Altaic populations, Turkic and Tunguzic show distinctive clusters, while the mongolian cluster is much closer with Sino-Tibetan, with the exception of westmongolian which is closer with Turkic. The age of M120 is estimated by microsatellite variance, and which proved that this lineage is very young and probably was derived from M45 after entering East Asia. A network was constructed with microsatellite haplotypes under the mutation of M89, which shows that the Central Asian populations are more diverse than northern Chinese populations, with Altaic ethnic populations closer with the former. Result of HVS I of mtDNA shows that the average nucleotide diversity of northern Chinese populations is similar with Central Asian populations, while those of Altaic populations and southern Han populations are higher. The HSV I motif of different haplogroups are summarized and the frequencies of some nucleotide within control region are compared with Central Asian populations. Admixture results show that the male contribution of Central Asia to northern population is smaller than that of female by comparing the data set of Y chromosome and mtDNA. All of the above evidences indicate that the most part of the gene pool of northern Chinese populations is from southern populations, and Central Asia has made contribution to it during later times, especially Altaic populations, through large-scale migrations associated with trade along the Silk Road, which eventually had an influence on diversity in East Asia. Whether the influence is much earlier than the establishment of Silk Road needs further studies. Estimating genetic influence of Central Asia on these populations, and further investigating the genetic divergence between north and south may help us with the understanding of population history of East Asia.