



B26 Y-STR Diversity in Pakistani Populations and Their Phylogenetic Relationships

Sibte Hadi, PhD, Louisiana State University Health Sciences Center, New Orleans, LA; William Goodwin, PhD, Department of Forensic Medicine & Science, University of Glasgow, Glasgow, Scotland; AMT Linacre, PhD, Forensic Science Unit, University of Strathclyde, Department of Pure & Applied Chemistry, Glasgow, Scotland; and Peter Vanezis, PhD, Department of Forensic Medicine & Science, University of Glasgow, Glasgow, Scotland*

Y chromosome specific STR database for the Pakistani population was required for forensic analyses. Such a database would also allow the study of the phylogenetic relationships of these highly endogamous and old ethnic groups. A male sample of 562 individuals from six ethnic groups was profiled for seven Y-STRs and genotyped using in-house allelic ladder. The haplotype analyses using various software revealed interesting patterns of diversity. The data was used to elicit the phylogenetic relationship between these ethnic groups.

In all human societies the majority of violent crimes are committed by the males, thus the characteristics that could identify and/or exclude a male have been of acute forensic interest. Most of the Y chromosome does not undergo recombination during meiosis thus the paternally inherited chromosome bears the genetic prints along the whole paternal lineage. Y specific short tandem repeats are now popular forensic markers that can be used as an adjunct to other markers and have also proved to be a useful forensic tool on their own. Various STRs have been described and many of these markers have been applied for building up population databases across the world in order to apply them to forensic casework. Due to the mode of inheritance, Y chromosome markers have also been utilized as a tool for studying phylogenetic relationships. The population of Pakistan is peculiar in that it is comprised of distinct groups, which are largely endogamous due to cultural and linguistic differences. These populations inhabit the area of one of the world's most ancient civilizations, i.e., the Indus Valley, thus it would be of interest to compare them with one another, and populations of other regions.

In this study 562 males belonging to six Pakistani populations, Punjabi, Sindhi, Pustoon, Baluchi, Makrani, Brosho, and Kalash were profiled for seven Y specific STR markers (DYS 391, 392, 393, 19, 389I, 389II and 390) to generate a haplotype (Yh1) in order to assess the forensic utility of this new tool. These STR markers were amplified in two multiplex reactions and an in-house allelic ladder was developed to genotype the amplified products. Genotyping was performed on ABI 310 DNA sequencer.

The haplotype diversity for the Pakistani population was found to be higher as compared to other populations, thus Yh1 would perform better in Pakistani than the European Caucasian population. For the individual populations the frequency of most common haplotype (MCH) in this database was not more than 6% in the mainland populations while it was only 3% in Sindhi and Punjabis which are the two major populations. The Y STR analysis revealed higher Yh1 haplotype diversity for Pakistani population in comparison to Caucasian populations. The discrimination capacity of this haplotype was only slightly lower than that of the extended haplotype which includes two more loci *DYS385* and *YCAII* to the seven locus Yh1, in European populations. The Y STR analysis in Pakistani populations would be thus very valuable even when the Yh1 is used for forensic casework.

The AMOVA analyses of the haplotype data performed using Arlequin software. AMOVA showed that the variation between the individuals of populations was much greater than that between the populations or to the arbitrary groups formed for the analysis. The variation between the major Pakistani ethnic groups was very low when the northern and southern ethnic groups were compared. Locus by locus AMOVA was also performed which showed that for all the groups major contribution of diversity was by the locus *DYS393*. The significant finding of pair-wise *F_{st}* showed that for the paternal lineages Punjabis were more closely related to Sindhis (as were Baluchis to Makranis) than to any other population. Mean pair-wise differences were calculated for all population pairs, highest values were obtained for the Baluchi population, while the lowest were for the Kalash. Molecular relationships of the component ethnic groups were studied using phylogenetic techniques. Distance matrices were used for generating phylogenetic trees. 'UPGMA' and 'Neighbourhood Joining' trees generated by PHYLIP were viewed with TREEVIEW program. Data from African populations was used to root the trees. The basic structure of the trees was similar. The Punjabi and Sindhi population cluster together showing common origin of the Y chromosome of these two populations; a similar cluster was seen for the Baluchi and Makrani.

In conclusion the analyses have revealed that a high degree of substructure exists in the Pakistani populations for Y STRs and the work has generated a large database of Y STRs for component ethnic groups of the Pakistani population. The AMOVA and the phylogenetic analyses have shown closer genetic relationship of Punjabi/Sindhi and Baluchi/Makrani populations which is consistent with the recent history of these populations.

Pakistani, Forensic, Y-STR