CHAPTER 5

CONCLUSION

The aims of this research were to choose the SNPs that were specific in term of genetic representation for each population, to test the SNPs population classification efficiency, and to examine the genetic relationship among populations.

From 58,960 SNP genotyping data generating from 13 ethnic groups residing in northern Thailand, the top ranked 100 SNPs, with highest mutual information values, were selected for the decision tree classification. The training set of 60 SNPs which had the highest classification accuracy of 89.84% on 13 populations differentiation, was selected.

The selected 60 SNPs were tested for genetic relationship analysis among 13 ethnic groups, using genetic distance analysis: Nei's Standard, and Cavalli-Sforza distance, phylogenetic tree reconstruction and multidimensional scaling plot. According to the population relationship when look at on multidimensional scaling, the outcome of the present method exhibits the similar population relationship to the HUGO Pan SNP Consortium study (2009) for the Tai, but in the Mon-Khmer, the Paluang and Mlabri are separated from the main cluster. The outcome indicates that the Tai and Mon-Khmer are genetically different.