COMMENTARY

A Commentary on Genetic affinity and admixture of northern Thai People along their migration route in Northern Thailand: evidence from autosomal STR loci

M Esther Esteban Torné

A good story is the first step towards captivating people's interest in disciplines such as scenic arts and literature, but also in scientific articles. However, even the most interesting issue should be complemented with a well-built script to ensure success. The article 'Genetic affinity and admixture of northern Thai People along their migration route in Northern Thailand: evidence from autosomal STR loci' by Kutanan et al.[1] is an example of that. The authors make good use of a set of commercial STRs (AmpFISTR Identifiler Kit, Applied Biosystems, Foster City, CA, USA), which are not exactly the state-of-the-art markers in human genetics, to draw an interesting hypothesis about the peopling of northern Thailand, thanks to a valuable collection of samples and a very detailed statistical approach.

The authors combine historical, ethnical, linguistic and genetic evidences to analyze 10 Khon Mueang populations, which constitute the majority of northern Thai people, together with 10 other neighboring populations from seven different ethnic groups. The main conclusions emphasize the admixed origin of the Khon Mueang from parental Tai and Mon-Khmer groups, the high heterogeneity among Khon Mueang groups, and the lack of correlation between geographic and genetic distances. Different historical and demographical factors are used to explain the human settlement of such a complex geographical area.

The AmpFISTR Identifier kit is widely used for forensic purposes.[2] The variability of the 15 STRs that constitute this kit has been described in numerous world populations,[3] but some continental parts of Oceania, the Americas and South East Asia are still poor covered. In this way, the data contained in the work of Kutanan et al.[1] will contribute to fill this gap. A basic search using the word 'AmpFISTR' in PubMed can retrieve hundreds of publications of this theme, most of them published in legal medicine or forensic science journals as data announcements. However, some of these articles go beyond the mere description of data and use the variability of STRs to address issues relevant in anthropology, such as the peopling and origin of diverse human groups. Regarding Asian populations, the recent work of Lin et al.[4] in Chinese ethnic minorities and the work commented here are good examples of how microsatellite markers could yield interesting results in genetic studies on the human population, apart from its unquestionable utility in forensic disciplines.

The ability of STRs, as any other genetic marker, to detect fine relationships among populations depends to a great extent on the number and quality of samples. Thus, these lines are used to encourage researchers to invest their time and effort on accurate sampling. The exhaustive fieldwork conducted by Kutanan et al.[1] could serve to illustrate the potentiality of a large and thorough sampling. The review of Kalsbeek and Heiss[5] emphasizes the need for rigorous population-sampling methods to face up the increasingly complex statistical analyses performed in epidemiology and human population genetics.

If valuable sampling is one of the pillars of successful research in population genetics, the
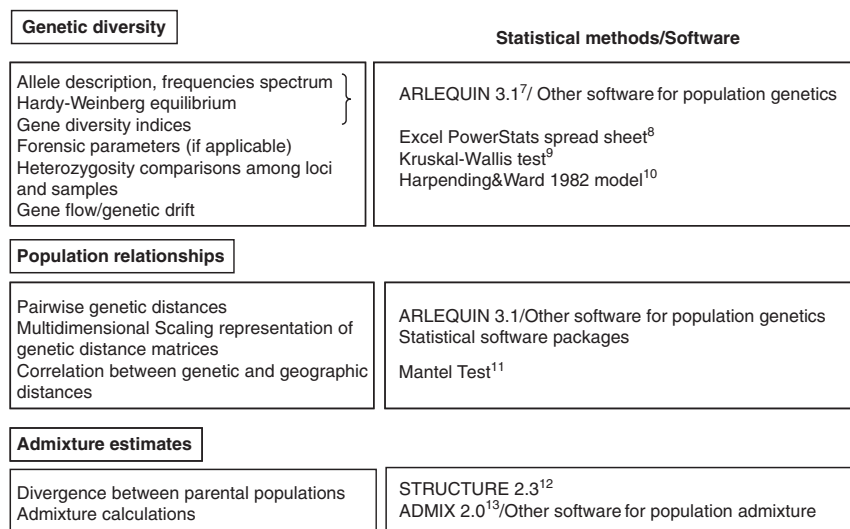
Dr ME Esteban Torné is at the Department of Animal Biology, Anthropology Section, Faculty of Biology, University of Barcelona, Barcelona 08028, Spain.
E-mail: mesteban@ub.edu

Genetic diversity	Statistical methods/Software
Allele description, frequencies spectrum Hardy-Weinberg equilibrium Gene diversity indices Forensic parameters (if applicable) Heterozygosity comparisons among loci and samples Gene flow/genetic drift	ARLEQUIN 3.1[7]/ Other software for population genetics Excel PowerStats spread sheet[8] Kruskal-Wallis test[9] Harpending&Ward 1982 model[10]
Population relationships	
Pairwise genetic distances Multidimensional Scaling representation of genetic distance matrices Correlation between genetic and geographic distances	ARLEQUIN 3.1/Other software for population genetics Statistical software packages Mantel Test[11]
Admixture estimates	
Divergence between parental populations Admixture calculations	STRUCTURE 2.3[12] ADMIX 2.0[13]/Other software for population admixture

Figure 1 Statistical methods to detect human population genetic parameters.

statistical tools used to analyze data are also essential for the final outcome. In this regard, the article commented upon here includes a large set of statistical analyses. The different questions addressed in the article, and the statistical tools used to answer them, are summarized in Figure 1. Even at the risk of seeming naïve, the diagram depicted might be very useful as a beginners' manual in population genetics, and could be completed with the helpful review of Excoffier and Heckel.[6]

Following the example used in the first paragraph, the work of Kutanan et al.[1] could be used to illustrate how the use of good materials (the samples) and methods (the statistical approach), together with a solid script, produces a convincing story: in this case, a complete understanding of the population history of Khon Mueang, northern Thailand. New studies capable of com- bining highly informative genetic markers with adequate statistical tools and interesting anthropological questions are encouraged.

1 Kutanan, W., Kampuansai, J., Colonna, V., Nakbunlung, S., Lertvicha, P., Seielstad, M. et al. Genetic affinity and admixture of northern Thai People along their migration route in Northern Thailand: evidence from autosomal STR loci. J. Hum. Genet. 56, 130–137 (2011).

2 Phillips, C., Fernandez-Formoso, L., Garcia-Magariños, M., Porras, L., Tvedebrink, T., Amigo, J. et al. Analysis of global variability in 15 established and 5 new European Standard Set (ESS) STRs using he CEPH human genome diversity panel. Forensic Sci. Int. Genet. (e-pub ahead of print 5 March 2010; doi:10.1016/j.fsigen.2010.02.003).

3 ALFRED allele frequency database. http://alfred.med. yale.edu/alfred/.

4 Lin, H., Fan, H., Huang, X., Lin, K., Shi, L., Hu, S. et al. Genetic relationships of ethnic minorities in Southwest China revealed by microsatellite markers. PLoS One 5, e9895 (2010).

5 Kalsbeek, W. & Heiss, G. Building bridges between populations and samples in epidemiological studies. Ann. Rev. Public Health 21, 147–169 (2000).

6 Excoffier, L. & Heckel, G. Computer programs for population genetics data analysis: a survival guide. Nat. Rev. Genet. 7, 745–758 (2006).

7 Excoffier, L., Laval, G. & Schneider, S. Arlequin ver. 3.0: an integrated software package for population genetic data analysis. Evol. Bioinform. Online 1, 47–50 (2005).

8 Excel PowerStats spread sheet. http://www.promega. com/geneticidtools/powerstats.

9 Zar, J. H. Biostatistical Analysis 4th edn (Prentice Hall, New Jersey, 1999).

10 Harpending, H. C. & Ward, R. in Biochemical Aspects of Evolutionary Biology (ed. Nitecki, M.) 213–256 (University of Chicago Press, Chicago, 1982).

11 Mantel, N. The detection of disease clustering and a generalized regression approach. Cancer Res. 27, 209–220 (1967).

12 Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. Genetics 155, 945–959 (2000).

13 Dupanloup, I. & Bertorelle, G. Inferring admixture proportions from molecular data: extension to any number of parental populations. Mol. Biol. Evol. 18, 672–675 (2001).