International Congress Series 1288 (2006) 79–81

www.ics-elsevier.com

Investigation of single-nucleotide polymorphisms associated with ethnicity

Runa Daniel *, Simon J. Walsh, Anita Piper

Department of Cell and Molecular Biology, Centre for Forensic Science, University of Technology Sydney, NSW, Australia

Abstract. Autosomal SNPs were investigated as markers of ethnicity. Loci were selected based on strong associations with major ethnic groups in the Australian population. The results from this preliminary research indicate that distinct genotype distributions are evident among the sub-populations under study. © 2005 Published by Elsevier B.V.

Keywords: SNP; Ethnicity; Allele frequency; SNaPshot

1. Introduction

Single nucleotide polymorphisms (SNPs) are the most frequently observed variation in the human genome. SNPs are stably inherited markers and reliable indicators of human biogeographical ancestry as a result of their low mutation rate. SNP based systems for inferring ethnicity enable investigators to narrow a pool of suspects when they are unable to match crime scene samples to database profiles.

As the first component of an investigation into the utility of autosomal SNPs as markers of ethnicity, this study aimed to develop an initial ethnicity multiplex. SNPs were selected to distinguish major ethnic groups in the Australian population.

2. Materials and methods

Population samples: Participants self-declared biogeographical ancestry over three generations. The study group consisted of Asian ($N=29$), Caucasian ($N=26$), Middle Eastern ($N=22$) and Sub-Continental Asian ($N=18$) samples. Samples from participants whose parents were both from the same ethnic group were analyzed.

Selection of SNP loci: Six autosomal SNPs were chosen from an extensive literature survey. The SNP loci are located in the Interleukin-10 (IL-10) [1], Interleukin-6 (IL-6) [2], Duffy [3] and DEFB1 [4] genes.

* Corresponding author. PO Box 123, Broadway NSW 2007, Australia. Tel.: +61 2 95144139; fax: +61 2 95144003.
E-mail address: runa.daniel@student.uts.edu.au (R. Daniel).

0531-5131/ © 2005 Published by Elsevier B.V.
doi:10.1016/j.ics.2005.11.131

Table 1
Allele and genotype frequencies for the six SNP loci and four populations under study

SNP locus		Populations				SNP locus		Populations			
		Asian (N=29)	Caucasian (N=26)	Middle Eastern (N=22)	Sub-Continental Asian (N=18)			Asian (N=29)	Caucasian (N=26)	Middle Eastern (N=22)	Sub-Continental Asian (N=18)
IL-10-1082	A	0.94	0.62	0.72	0.91	Duffy -46	C	–	–	0.04	0.11
	G	0.06	0.38	0.28	0.09		T	1.00	1.00	0.96	0.89
	AA	0.89	0.46	0.56	0.82		CC	–	–	–	–
	AG	0.11	0.31	0.33	0.18		CT	–	–	0.08	0.22
	GG	–	0.23	0.11	–		TT	1.00	1.00	0.92	0.78
IL-10-819	C	0.32	0.71	0.78	0.86	hBD1 322	A	0.62	0.33	0.71	0.50
	T	0.68	0.29	0.22	0.14		T	0.38	0.68	0.29	0.50
	CC	0.07	0.47	0.67	0.86		AA	0.31	0.10	0.41	0.29
	CT	0.50	0.47	0.22	–		AT	0.62	0.45	0.59	0.43
	TT	0.43	0.06	0.11	0.14		TT	0.08	0.45	–	0.29
IL-6-174	C	0.06	0.45	0.13	0.17	hBD1 1754	A	0.33	0.17	0.11	0.35
	G	0.94	0.55	0.88	0.83		G	0.67	0.83	0.89	0.65
	CC	–	0.16	–	–		AA	0.19	0.10	0.07	0.20
	CG	0.13	0.58	0.25	0.33		AG	0.29	0.14	0.07	0.30
	GG	0.88	0.26	0.75	0.67		GG	0.52	0.76	0.86	0.50

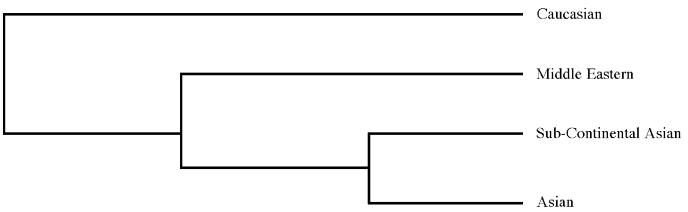

Fig. 1. Neighbour-joining tree for all populations (N_{pops}=4, N=95).

Table 2
Pairwise genetic distance matrix for all populations

	Sub-Continental Asian	Middle Eastern	Caucasian	Asian
Sub-Continental Asian		0.0289	0.0797	0.1178
Middle Eastern	0.0293		0.0697	0.1245
Caucasian	0.0831	0.0722		0.2362
Asian	0.1253	0.1330	0.2695	

SNP genotyping and detection: Primer extension reactions were carried out using the ABI Prism® SNaPshot™ Multiplex Kit (Applied Biosystems). SNaPshot products were capillary electrophoresed on an ABI PRISM™ 310 Genetic Analyzer (Applied Biosystems) and analyzed using GeneScan® Analysis Software Version 3.7 (Applied Biosystems).

3. Results and discussion

Observable differences in allele and genotype frequencies exist (Table 1), notably between Asian and Caucasian populations at the IL-10-1082 and IL-10-819 loci. Caucasian observed frequencies at IL-6-174 are also distinct from other populations.

A neighbour-joining (NJ) phylogenetic tree was constructed [5] from θ estimates for the full set of populations (Fig. 1). Pairwise genetic distance values for the set of four populations are presented in Table 2. The θ estimates for each pair of populations are shown above the diagonal and the genetic distance based on those θ values are below. Pairwise θ estimates in the order of 0.24 (Asian-Caucasian) are high and akin to those observed at non-autosomal markers [6].

The results indicate a considerable degree of diversity, particularly between the Caucasian and Asian sub-populations, which are two populations of principle importance in the Australian context. Further characterisation of additional informative loci will enable us to define ethnic groups specific to the Australian population, providing a valuable intelligence tool for forensic investigators.

References

[1] D. Kube, et al., Cytokine 7 (1995) 1–7.
[2] D. Fishman, et al., Journal of Clinical Investigation 102 (1998) 1376–1639.
[3] C. Tournamille, et al., Blood 92 (1998) 2147–2156.
[4] R. Jurevic, et al., Genetic Testing 6 (2002) 261–269.
[5] P.O. Lewis, D. Zaykin. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). http://lewis.eeb.uconn.edu/lewishome/software.html.
[6] L. Jorde, W. Watkins, M. Bamshad, Human Molecular Genetics 10 (2001) 2199–2207.