Available online at www.sciencedirect.com
ScienceDirect
Forensic Science International: Genetics 1 (2007) e1–e3
FSI GENETICS
www.elsevier.com/locate/fsig

Announcement of Population Data

Frequency assessment of SNPs for forensic identification in different populations

Emiliano Giardina [a], Irene Predazzi [a], Ilenia Pietrangeli [a], Paola Asili [b], Patrizio Marsala [b], Luciano Gabriele [b], Claudio Pipolo [b], Omero Ricci [b], Claudia Martone [a], Aldo Spinella [b], Giuseppe Novelli [a,c,d,*]

[a] *Centre of Excellence for Genomic Risk Assessment in Multifactorial and Complex Diseases, School of Medicine, Tor Vergata University of Rome, Italy*
[b] *Direzione Centrale Anticrimine, Servizio di Polizia Scientifica, Rome, Italy*
[c] *Division of Cardiovascular Medicine, Department of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR 72205, United States*
[d] *Fondazione Livio Patrizi, Roma, Italy*

Received 17 February 2007; accepted 12 May 2007

Abstract

Allele frequencies for 16 previously described autosomal SNPs were tested in 1020 unrelated individuals originating from three different continents (Africa, Asia and Europe). The populations analyzed included Africans from Benin Gulf (180), Asians from Mongolia (160) and Europeans from Italy (680).

Keywords: Single nucleotide polymorphisms; Human identification; Forensic genetics

Population: One thousand and twenty unrelated individuals originating from three different continents (Africa, Asia and Europe). Selected individuals included Africans from Benin Gulf (180), Asians from Mongolia (160) and Europeans from Italy (680).

Extraction: The QIAamp DNA Blood Mini Kit (QIAGEN Inc., Valencia, CA) has been used to extract genomic DNA from whole blood.

SNP selection: Sixteen sequence polymorphisms have been selected from SNPs submitted to the Forensic SNP information site (http://www.cstl.nist.gov/biotech/strbase/SNP.htm) and published in a previous work [1].

The SNPs selected from the Forensic SNP information website were rs420426, rs1482650, rs2311046 rs1453461, rs675236, rs1542931, rs174473, rs1003473. The SNPs selected from previous publication [1] were rs734664, rs734295, rs1075665, rs585070, rs911621, rs999842, rs873289, rs1000322.

SNP typing: SNPs genotyping was performed by *Taq*man allele discrimination assay (Applied Biosystems, Foster City, CA). The list of the SNPs together wtih their relative chromosome localizations, PCR primer sequences and relative Taqman probes are reported in Table 1.

Genotyping confirmation and statistical analysis: Allele frequencies were calculated by direct counting. Genotype assessment for all SNPs tested has been confirmed by direct sequencing of random samples. No departures form Hardy–Weinberg equilibrium for the SNPs considered has been revealed following Arlequin 2.0 calculations (http://lgb.unige.ch/arlequin/). The power of discrimination was calculated as described by Jones [2] using PowerStats (Promega). The identification of the copy number variation regions has been performed after search in the Database of Genomic Variants (http://projects.tcag.ca/variation/).

Results: The allele frequencies observed for the SNPs in the three populations considered are summarized in Table 2. The power of discrimination was 1×10^{-6} in the European population, 5×10^{-6} in the African population and 1×10^{-6} in the Asian population.

* Corresponding author at: Centre of Excellence for Genomic Risk Assessment in Multifactorial and Complex Diseases, School of Medicine, Tor Vergata University of Rome, Italy. Tel.: +39 06 72596072/20900665; fax: +39 06 20900669.
E-mail address: novelli@med.uniroma2.it (G. Novelli).

1872-4973/$ – see front matter
doi:10.1016/j.fsigen.2007.05.002

Table 1
Name, chromosomal localization and 5′-3′ sequences of RT-PCR primers and probes of SNPs selected

SNP name	Chromosomal localization (bp)	PRIMER forward	PRIMER reverse	Probe (1)	Probe (2)
rs420426	chr3:9,127,124	TGCTTCCCTAAAAGAGGGCA	GGTCCGTCCGAGATGGTGTA	AACTGCTGGGTCTGC ACCC	AACTGCTGGGTCTGTACC
rs1482650	chr3:43,459,673	TGAGGATCCACACAGGGAATGA	ATCTTGGAAAGTCACCCATGCA	CAGGGAACCACTAATAT	CAGGGAACCACTTATAT
rs2311046	chr3:79,510,160	CTGGAGGCACCACAGCAT	CCATTCAGAACTAACTAGTCTGGGAAA	CCAGCCTCCTGCATG	CCAGCCTCCAGCATG
rs1453461	chr4:72,682,312	GGGCTTTCCCTTTCCTACCA	GTGCCCCCAAATAGAGCATAGG	CATTTGACCCTCCACAGAC	TTTGACCCACCACAGAC
rs675236	chr7:114,174,892	GCAGATGGGACTCCTGAAAGG	CATCCGGGCCCCAGAAG	CAGCAATTCTTTTTCCCA	CAGCAATTCTTATTCCCA
rs1542931	chr8:91,892,744	GAGCTACAGCAAATGAAGCATCAAT	TCATTTGGCTTAGTGATGCCTCTT	TTGTGAACGCATTGAA	ATTGTGAACCCATTGAA
rs174473	chr11:61,428,811	GCCATGCCTCACCTCCT	GTGTACCGGGAGGACGTG	CTGGACTTCACGATGC	TGGACTTCACAATGC
rs1003473	chr21:35,368,467	CTTTCTCATGTCATGCCACTCTGA	CAAGCATCCTTAAAAGTGGACTCGA	CACTGATGCTTCTGTCTC	CTGATGCTTGTGTCTC
rs734664	chr1:14,869,241	CAAACACAAATGCCTTCAGG	TGCCATTGTACAGGGGATGC	AAAGCAGGACGGAGAAGG	AAAGCAGGATGGAGAAGG
rs734295	chr2:53,681,914	GTGCCCAGCTCCCTAATTTCTT	TTCCACAATCTCTCTTGTGACTTTCATT	CTATCTTCAGATTTGCCCATC	CTATCTTCAGATTTACCCATC
rs1075665	chr6:88,423,321	TCTTCTTCCTCTTTGGTTTCTGCTTT	GTTTGTTCCCTTGACTTGGATTTACC	CATTGATGTCTGCTGGTTT	TCATTGATGTCTACTGGTTT
rs585070	chr10:113,617,876	GGTTCTAGACCTAAATAGTGGCCCTAA	ACATCTCTACTGAAGACAAACTTAGAGGAA	CTTTCCCTACGTATAAAA	CTTTCCCTATGTATAAAA
rs911621	chr14:54,195,466	ACCTGGCCTCTCTGAGATTCA	CCTCAAGGGCCCTGTGA	CACCTGCAAATCAA	ACACCTGTAAATCAA
rs999842	chr15:20,551,713	GCCAGAAACAAGAACTTATAACA CAAAGAAAG	ATTATTAAAAATGTGGTATGTGGATGAAA ATCTCG	ACCATCTGTGGACCTT	CACCATCTATGGACCTT
rs873289	chr19:1,126,396	AGCATCGATTAGAAGTCACTGAATGAATT	GACCCAAGCAGCCACAAG	CCCTGAAACTGGAGTTG	CCCTGAAACTAGAGTTG
rs1000322	chr20:59,491,626	GTAGCTCAAGCACTCTCTCTTTTCA	GGAGTCCAGGTAGATAGGAACACTA	TCACTCATCTCGTCTTAC	TCACTCATCTTGTCTTAC

Table 2
Allele frequencies of selected SNPs in Italian, Benin African and Mongolian Asian populations

SNPs	Alleles	Italians		Total	Benin Africans		Total	Mongolian Asians		Total
		Allele 1	Allele 2		Allele 1	Allele 2		Allele 1	Allele 2	
rs420426	C/T	0.53	0.47	680	0.82	0.18	180	0.60	0.40	160
rs1482650	A/T	0.67	0.33	680	0.91	0.09	180	0.85	0.15	160
rs2311046	A/T	0.62	0.38	680	0.16	0.84	180	0.53	0.47	160
rs1453461	A/T	0.89	0.11	680	0.46	0.54	180	0.87	0.13	160
rs675236	A/T	0.75	0.25	680	0.90	0.10	180	0.78	0.22	160
rs1542931	C/G	0.70	0.30	680	0.72	0.28	180	0.35	0.65	160
rs174473	C/T	0.22	0.78	680	0.42	0.58	180	0.31	0.69	160
rs1003473	C/G	0.73	0.27	680	0.77	0.23	180	0.59	0.41	160
rs734664	C/T	0.74	0.26	680	0.44	0.56	180	0.75	0.25	160
rs734295	C/T	0.60	0.40	680	0.84	0.16	180	0.66	0.34	160
rs1075665	C/T	0.27	0.73	680	0.41	0.59	180	0.38	0.62	160
rs585070	C/T	0.50	0.50	680	0.81	0.19	180	0.61	0.39	160
rs911621	C/T	0.51	0.49	680	0.57	0.43	180	0.70	0.30	160
rs999842	C/T	0.49	0.51	680	0.35	0.65	180	0.60	0.40	160
rs873289	T/C	0.52	0.48	680	0.59	0.41	180	0.63	0.37	160
rs1000322	C/T	0.44	0.56	680	0.30	0.70	180	0.46	0.54	160

Other remarks: It should be outlined that previously published allele frequencies have been evaluated in different populations. Allele frequencies submitted to the Forensic SNP Information website were analyzed in North European samples (86), Indo-Pakistan samples (33) and in Afro-Caribbean (29) samples. Allele frequencies published by Vallone *et al.* [1] were calculated through the analysis of 189 individuals from three different U.S. sample groups: Caucasian (74), African-American (71) and Hispanic (44). Several factors have to be considered for the selection of the best SNPs for forensic identification purposes. The first factor to take into account is the frequency of the SNPs across the populations. The STR markers have many alleles with low frequency in the most of the populations, therefore the probability of matching are not largely dissimilar between the populations. SNPs may show very dissimilar frequencies among different populations, causing a very large dependence of the random match probability from the population frequencies used for the calculation [3]. In particular marker rs1482650 showed low heterozygosity in the African and Asian populations tested; marker rs1453461 showed low heterozygosity in the Italian and Asian populations and marker rs675236 showed low heterozygosity in the African population. A second factor to consider is that forensic SNPs should be located exclusively in the human genome sequence and represented in a single copy. Recently, several research studies have discovered an abundance of submicroscopic copy number variation of DNA segments ranging from kilobases (kb) to megabases (Mb) in size. Deletions, insertions, duplications and complex multi-site variants, collectively termed copy number variations (CNVs) are found in all humans and other mammals [4]. A total of 1447 copy number variable regions (CNVRs), which can encompass overlapping or adjacent gains or losses, covering the 12% of the genome were so far identified. The finding of SNP rs999842 location within a CNVR suggests that it should not be considered a good marker for forensic purposes.

Access to the data: Data will be available on http://www.nacbo.net.

Acknowledgements

This work was supported by financing from EU FP6 projects NACBO (contract no. NMP4-CT-2004-500804). We also thank Olga Rickards and Cristina Martinez-Labarga for their help in recruiting samples.

References

[1] P.M. Vallone, A.E. Decker, J.M. Butler, Allele frequencies for 70 autosomal SNP loci with U.S. Caucasian, African-American, and Hispanic samples, Forensic Sci. Int. 149 (2005) 279–286.

[2] D.A. Jones, Blood samples: probability of discrimination, J. Forensic Sci. Soc. 12 (1972) 355–359.

[3] Q.Y. Huang, F.H. Xu, H. Shen, H.Y. Deng, Y.J. Liu, Y.Z. Liu, J.L. Li, R.R. Recker, H.W. Deng, Mutation patterns at dinucleotide microsatellite loci in humans, Am. J. Hum. Genet. 70 (2002) 625–634.

[4] R. Redon, S. Ishikawa, K.R. Fitch, L. Feuk, G.H. Perry, T.D. Andrews, H. Fiegler, M.H. Shapero, A.R. Carson, W. Chen, E.K. Cho, S. Dallaire, J.L. Freeman, J.R. Gonzalez, M. Gratacos, J. Huang, D. Kalaitzopoulos, D. Komura, J.R. MacDonald, C.R. Marshall, R. Mei, L. Montgomery, K. Nishimura, K. Okamura, F. Shen, M.J. Somerville, J. Tchinda, A. Valsesia, C. Woodwark, F. Yang, J. Zhang, T. Zerjal, J. Zhang, L. Armengol, D.F. Conrad, X. Estivill, C. Tyler-Smith, N.P. Carter, H. Aburatani, C. Lee, K.W. Jones, S.W. Scherer, M.E. Hurles, Global variation in copy number in the human genome, Nature 444 (2006) 444–454.