Forensic Science International: Genetics 22 (2016) 81–88

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Research paper

The Global AIMs Nano set: A 31-plex SNaPshot assay of ancestry-informative SNPs

M. de la Puente[a], C. Santos[a], M. Fondevila[a], L. Manzo[a], The EUROFORGEN-NoE Consortium, Á. Carracedo[a,b], M.V. Lareu[a], C. Phillips[a,*]

[a] Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain
[b] Center of Excellence in Genomic Medicine Research, King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 7 October 2015
Received in revised form 20 January 2016
Accepted 21 January 2016
Available online 25 January 2016

Keywords:
SNPs
AIMs
Biogeographical ancestry
SNaPshot
Population-specific Divergence

ABSTRACT

A 31-plex SNaPshot assay, named 'Global AIMs Nano', has been developed by reassembling the most differentiated markers of the EUROFORGEN Global AIM-SNP set. The SNPs include three tri-allelic loci and were selected with the goal of maintaining a balanced differentiation of: Africans, Europeans, East Asians, Oceanians and Native Americans. The Global AIMs Nano SNP set provides higher divergence between each of the five continental population groups than previous small-scale AIM sets developed for forensic ancestry analysis with SNaPshot. Both of these characteristics minimise potential bias when estimating co-ancestry proportions in individuals with admixed ancestry; more likely to be observed when using markers disproportionately informative for only certain population group comparisons. The optimised multiplex is designed to be easily implemented using standard capillary electrophoresis regimes and has been used to successfully genotype challenging forensic samples from highly degraded material with low level DNA. The ancestry predictive performance of the Global AIMs Nano set has been evaluated by the analysis of samples previously characterised with larger AIM sets.

© 2016 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Although STR profiling has been successfully applied to the majority of forensic DNA analyses for many years, there are still situations when STR typing is unable to inform criminal investigations, for example, with no matching profile found in DNA database searches or when no suspect is apprehended. For this reason, there is interest in developing DNA tests that can provide investigative leads, focused on panels of single nucleotide polymorphisms (SNPs) to predict external visible characteristics (ECVs), including common variation in pigmentation [1], or to infer an individual's biogeographical ancestry [2].

With the recent availability of bench-top systems for massively parallel sequencing (MPS) that are applicable to forensic DNA analysis, it is now possible to assemble multiplexes of 400–500 markers [3]. Such enlarged forensic multiplexes can include a portion of carefully chosen ancestry informative markers (AIMs), e.g., the Illumina ForenSeq panel [4], or can be exclusively composed of AIMs [5–7]. Both approaches raise the level of geographic resolution that can be obtained from tests that keep the necessary forensic sensitivity. However, the forensic community will take time to adopt, optimise and validate MPS technology as a routine analysis system. Therefore, it is important to continue to develop small-scale AIM sets suited to short-amplicon marker genotyping with validated, universally applicable capillary electrophoresis (CE) analysis regimes [8–10]. However, one drawback with use of small-scale AIM sets is the potential for over-estimation of co-ancestry proportions in individuals with admixed ancestry, stemming from the analysis of genotypes strongly differentiated for some populations but not others. The phenomenon of biased estimation of co-ancestry components was detected in a study of Bolivian populations [11] using 46 ancestry-informative Indels [8] compared with a much larger panel of 446 AIM-SNPs [12]. The Indel set consistently over-estimated European co-ancestry and under-estimated Native American co-ancestry using STRUCTURE-based analyses, indicating that the higher European differentiation of the Indel genotypes inflated the estimates of European co-ancestry proportions. Bearing in mind this effect, construction of a dedicated AIM-SNP set for MPS by the EUROFORGEN Consortium [5] sought to carefully balance the cumulative population-specific Divergence values for the five continental population groups of Africa, Europe, East Asia, Native America and Oceania.

* Corresponding author.
E-mail address: c.phillips@mac.com (C. Phillips).

http://dx.doi.org/10.1016/j.fsigen.2016.01.015
1872-4973/© 2016 Elsevier Ireland Ltd. All rights reserved.

The emphasis on keeping balanced population group Divergence values provided the main focus for the new ancestry informative SNP panel reported here. We took the most differentiated AIM-SNPs from the EUROFORGEN Global AIMs panel [5] and assembled a compact 31-plex assay genotyped with SNaPshot® single base extension technology. The SNP set, named 'Global AIMs Nano' (herein Nano) was designed to be applicable to forensic analyses where several different admixture combinations may be commonly encountered, e.g. in Australia; where comparisons of European, Oceanian and East Asian co-ancestry components will be routinely necessary. As well as preserving a comparable level of differentiation amongst the five population groups, the Nano assay aimed to provide a single CE-based test that is sufficiently informative for all five groups.

2. Materials and methods

2.1. Reference population SNP genotype data and DNA samples

SNP variation data from representative populations without high levels of admixture was obtained from 1000 Genomes Phase III [13] and from the Stanford University HGDP-CEPH SNP analysis [14] using the SPSmart frequency browser [15]. SNP genotype data was compiled from 108 YRI Africans (AFR: Yoruba in Ibadan, Nigeria); 99 CEU Europeans (EUR: Utah Residents with North and Western European ancestry); 103 CHB East Asians (EAS: Han Chinese in Beijing, China); 28HGDP-CEPH Oceanians (OCE: 17 Papuan from New Guinea and 11 Melanesian from Bougainville); and 64 HGDP-CEPH Native Americans (AMR: 14 Karitiana, 8 Surui from Brazil; 21 Maya, 14 Pima from Mexico; and 7 Piapoco from Colombia). Phase III 1000 Genomes populations were also analysed, comprising: as a test set, 99 AFR LWK (Luhya in Webuye, Kenya); 113 AFR GWD (Gambian in Western Divisions in the Gambia); 85 AFR MSL (Mende in Sierra Leone); 99 AFR ESN (Esan in Nigeria); 107 EUR TSI (Toscani in Italia); 99 EUR FIN (Finnish in Finland); 91 EUR GBR (British in England and Scotland); 107 EUR IBS (Iberian Population in Spain); 104 EAS JPT (Japanese in Tokyo, Japan); 105 EAS CHS (Southern Han Chinese); 99 EAS KHV (Kinh in Ho Chi Minh City, Vietnam); 93 EAS CDX (Chinese Dai in Xishuangbanna, China); plus admixed populations 61 ASW (Americans of African Ancestry in SW USA); 96 ACB (African Caribbeans in Barbados); 104 PUR (Puerto Ricans from Puerto Rico), 94 CLM (Colombians from Medellin, Colombia); 64 MXL (individuals with Mexican Ancestry from Los Angeles USA); 85 PEL (Peruvians from Lima, Peru).

To evaluate the forensic sensitivity of the Nano assay, challenging casework samples plus control DNAs were analysed, comprising: (i) five DNA samples each from separate population groups, previously used in an ancestry analysis collaborative exercise [10]; (ii) highly degraded skeletal DNA extracts; (iii) a doubling dilution series of 1 ng/μL; 0.5 ng/μL; 0.25 ng/μL; 0.125 ng/μL; 0.064 ng/μL; 0.032 ng/μL; and 0.016 ng/μL of the 9947A forensic kit DNA standard.

2.2. AIM-SNP selection and SNaPshot assay design

Ancestry-informative SNPs were selected directly from the EUROFORGEN Global AIM-SNP set according to the following criteria: (i) differentiation of five population groups to comparable levels to produce population-specific Divergence (PSD) values as balanced as possible (use of capitalised Divergence distinguishes the metric from the phenomenon of population divergence); (ii) inclusion of certain informative tri-allelic SNPs to allow a level of mixed DNA detection; (iii) genomic separation of component SNPs by a minimum inter-marker distance of 1 Mb to minimise the effects of linkage on likelihood calculations that assume independence for the loci analysed.

From the selected SNPs, a 31-plex SNaPshot® single base extension assay was designed and optimised following established guidelines [16]. Locus details and summary allele frequencies for component SNPs are summarised in Table 1. PCR and single base extension (SBE) primers are detailed in Supplementary Table S1.

Table 1
Description, reference allele frequencies and population-specific/pairwise Divergence values (I_n) of the 31 Nano SNPs. Chr: chromosome; RA: reference allele. All positions from genome build 37.1 (GRCh37). SNPs are ranked according to their Pop vs. Other Pop I_n (highlighted in grey) inside each population informative (Informat.) group.

SNP details					Reference allele frequency					Population-specific Divergence					Pairwise Divergence									
Informat.	SNP ID	Chr	Position	RA	AFR	EUR	EAS	OCE	AMR	AFR	EUR	EAS	OCE	AMR	AFR - EUR	AFR - EAS	AFR - OCE	AFR - AMR	EUR - EAS	EUR - OCE	EUR - AMR	EAS - OCE	EAS - AMR	OCE - AMR
AFR	rs2814778	1	159174683	A	0.005	1.000	1.000	1.000	0.992	0.672	0.131	0.134	0.083	0.108	0.663	0.663	0.634	0.656	0.000	0.672	0.000	0.000	0.000	0.001
	rs1871534	8	145639681	C	0.981	0.000	0.000	0.000	0.000	0.641	0.128	0.131	0.081	0.107	0.631	0.632	0.603	0.624	0.000	0.641	0.000	0.000	0.000	0.000
	rs2789823	9	136769688	G	0.935	0.000	0.000	0.000	0.000	0.565	0.121	0.123	0.076	0.100	0.556	0.557	0.527	0.548	0.000	0.565	0.000	0.000	0.000	0.000
EUR	rs1426654	15	48426484	A	0.014	1.000	0.029	0.000	0.039	0.117	0.622	0.093	0.081	0.069	0.641	0.001	0.000	0.003	0.611	0.117	0.595	0.001	0.000	0.002
	rs16891982	5	33951693	C	1.000	0.020	0.985	1.000	0.984	0.126	0.620	0.104	0.073	0.087	0.629	0.001	0.000	0.002	0.606	0.126	0.603	0.000	0.000	0.000
	rs12142199	1	1249187	G	0.977	0.177	0.971	1.000	1.000	0.076	0.402	0.068	0.061	0.083	0.393	0.000	0.000	0.002	0.383	0.076	0.423	0.001	0.003	0.000
	rs8072587	17	19211073	C	0.986	0.182	1.000	1.000	0.817	0.099	0.358	0.113	0.069	0.003	0.405	0.001	0.000	0.047	0.425	0.099	0.218	0.000	0.058	0.043
	rs9522149	13	111827167	T	0.972	0.237	0.995	1.000	0.977	0.063	0.353	0.092	0.055	0.056	0.334	0.004	0.001	0.000	0.377	0.063	0.341	0.002	0.003	0.000
	rs4749305	10	28391596	A	0.389	0.909	0.078	0.036	0.008	0.001	0.309	0.095	0.100	0.155	0.162	0.072	0.105	0.141	0.404	0.001	0.514	0.004	0.017	0.005
EAS	rs17822931	16	48258198	C	1.000	0.869	0.029	0.875	0.650	0.190	0.053	0.432	0.039	0.000	0.039	0.613	0.037	0.129	0.428	0.190	0.034	0.434	0.251	0.036
	rs1229984	4	100239319	A	0.000	0.015	0.709	0.071	0.000	0.089	0.070	0.320	0.018	0.071	0.002	0.336	0.018	0.000	0.313	0.089	0.001	0.238	0.328	0.015
	rs3827760	2	109513601	T	1.000	1.000	0.063	0.946	0.109	0.219	0.209	0.319	0.098	0.203	0.000	0.559	0.012	0.500	0.558	0.219	0.499	0.471	0.003	0.415
	rs6437783	3	108172817	C	0.259	0.146	0.995	0.589	0.891	0.078	0.153	0.268	0.001	0.104	0.010	0.359	0.057	0.223	0.459	0.078	0.311	0.157	0.031	0.082
	rs12594144	15	64161351	C	1.000	0.889	0.121	0.607	0.177	0.237	0.097	0.222	0.000	0.130	0.032	0.487	0.149	0.430	0.334	0.237	0.283	0.136	0.003	0.101
	rs4657449	1	165465281	G	0.912	0.909	0.102	0.000	0.117	0.176	0.164	0.177	0.210	0.130	0.000	0.380	0.497	0.364	0.376	0.176	0.360	0.017	0.000	0.022
OCE	rs9908046	17	53563782	C	0.958	0.929	0.883	0.018	0.992	0.020	0.008	0.000	0.528	0.042	0.002	0.010	0.562	0.006	0.003	0.020	0.015	0.463	0.030	0.626
	rs3751050	11	9091244	A	0.972	0.924	0.966	0.089	0.961	0.020	0.002	0.016	0.451	0.011	0.006	0.000	0.477	0.000	0.004	0.020	0.003	0.467	0.000	0.459
	rs2139931	1	84590527	A	0.898	0.753	0.879	0.018	0.898	0.017	0.002	0.011	0.433	0.014	0.019	0.000	0.480	0.000	0.013	0.017	0.019	0.458	0.000	0.480
	rs715605	22	30640308	T	0.866	0.914	0.985	0.089	1.000	0.000	0.003	0.039	0.422	0.041	0.003	0.030	0.345	0.036	0.015	0.000	0.020	0.502	0.001	0.516
	rs6054465	20	6673018	T	0.972	0.859	0.743	0.036	0.859	0.062	0.005	0.005	0.408	0.004	0.022	0.061	0.553	0.022	0.011	0.062	0.000	0.306	0.011	0.408
	rs9809818	3	71480566	C	0.019	0.116	0.869	0.982	0.820	0.245	0.120	0.172	0.227	0.102	0.021	0.446	0.603	0.399	0.319	0.245	0.276	0.026	0.002	0.042
AMR	rs12498138	3	121459589	G	1.000	0.949	0.922	0.911	0.094	0.085	0.034	0.020	0.011	0.443	0.011	0.020	0.024	0.519	0.002	0.085	0.437	0.000	0.401	0.387
	rs10483251	14	21671277	G	0.921	0.798	0.898	0.712	0.024	0.055	0.006	0.040	0.000	0.429	0.016	0.001	0.038	0.497	0.010	0.055	0.370	0.028	0.469	0.301
	rs2080161	7	13331150	T	0.981	0.758	0.689	0.820	0.000	0.144	0.007	0.001	0.036	0.424	0.064	0.091	0.044	0.624	0.003	0.144	0.366	0.033	0.314	0.462
	rs8137373	22	41729216	G	0.833	0.707	0.927	0.982	0.023	0.020	0.000	0.068	0.093	0.406	0.011	0.011	0.037	0.402	0.043	0.020	0.298	0.009	0.506	0.593
	rs1557553	22	44760984	C	0.949	0.904	0.714	0.786	0.094	0.076	0.042	0.000	0.002	0.325	0.004	0.053	0.031	0.436	0.030	0.076	0.380	0.003	0.219	0.270
	rs12402499	1	101528954	G	1.000	0.919	1.000	1.000	0.258	0.061	0.007	0.059	0.032	0.324	0.021	0.000	0.000	0.361	0.021	0.061	0.252	0.000	0.361	0.335
	rs4792928	17	42105174	T	1.000	0.960	0.345	0.804	0.195	0.171	0.111	0.108	0.011	0.178	0.008	0.297	0.064	0.413	0.239	0.171	0.350	0.113	0.014	0.199
Triallelic	rs2069945	20	33761837	CG	0.153 / 0.796	0.480 / 0.420	0.680 / 0.277	0.036 / 0.536	0.766 / 0.234	0.114	0.002	0.045	0.201	0.087	0.078	0.156	0.118	0.210	0.022	0.114	0.065	0.289	0.017	0.384
	rs4540055	4	38803255	AC	0.069 / 0.514	0.793 /0.010	0.301 / 0.068	0.536 / 0.250	0.605 / 0.000	0.217	0.148	0.055	0.024	0.084	0.355	0.147	0.142	0.300	0.130	0.217	0.027	0.098	0.063	0.101
	rs5030240	11	32424389	CA	0.278 / 0.389	0.712 /0.055	0.228 / 0.039	0.093 / 0.278	0.258 / 0.023	0.083	0.117	0.062	0.062	0.054	0.125	0.123	0.052	0.135	0.135	0.083	0.124	0.067	0.001	0.085
								CUMULATIVE VALUES		4.739	4.404	3.392	3.986	4.374	5.263	6.111	6.210	8.029	6.274	4.739	7.184	4.323	3.106	6.350

Download English Version:

https://daneshyari.com/en/article/98708

Download Persian Version:

https://daneshyari.com/article/98708

Daneshyari.com