



### **King's Research Portal**

DOI: 10.1016/j.fsigen.2015.06.004

Document Version Peer reviewed version

Link to publication record in King's Research Portal

Citation for published version (APA):

Santos, C., Fondevila, M., Ballard, D., Banemann, R., Bento, A. M., Børsting, C., Branicki, W., Brisighelli, F., Burrington, M., Capal, T., Chaitanya, L., Daniel, R., Decroyer, V., England, R., Gettings, K. B., Gross, T. E., Haas, C., Harteveld, J., Hoff-Olsen, P., ... Phillips, C. (2015). Forensic ancestry analysis with two capillary electrophoresis ancestry informative marker (AIM) panels: Results of a collaborative EDNAP exercise. *Forensic Science International-Genetics*, *19*, 56-67. https://doi.org/10.1016/j.fsigen.2015.06.004

#### Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

#### General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
You may not further distribute the material or use it for any profit-making activity or commercial gain
You may freely distribute the URL identifying the publication in the Research Portal

#### Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

## Forensic ancestry analysis with two capillary electrophoresis ancestry informative marker (AIM) panels: Results of a collaborative EDNAP exercise

C. Santos<sup>a,1</sup>, M. Fondevila<sup>a,1</sup>, D. Ballard<sup>b,1</sup>, R. Banemann<sup>c</sup>, A.M. Bento<sup>d</sup>, C. Børsting<sup>e,1</sup>, W. Branicki<sup>f,2</sup>, F. Brisighelli<sup>g</sup>, M. Burrington<sup>h</sup>, T. Capal<sup>i</sup>, L. Chaitanya<sup>j</sup>, R. Daniel<sup>k</sup>, V. Decroyer<sup>l</sup>, R. England<sup>m</sup>, K.B. Gettings<sup>n</sup>, T.E. Gross<sup>o,1</sup>, C. Haas<sup>p</sup>, J. Harteveld<sup>q</sup>, P. Hoff-Olsen<sup>r</sup>, A. Hoffmann<sup>c</sup>, M. Kayser<sup>j</sup>, P. Kohler<sup>r,2</sup>, A. Linacre<sup>s</sup>, M. Mayr-Eduardoff<sup>t,1</sup>, C. McGovern<sup>m</sup>, N. Morling<sup>e,1 l</sup>, G. O'Donnell<sup>h</sup>, W. Parson<sup>t,u,1</sup>, V.L. Pascali<sup>g</sup>, M.J. Porto<sup>d</sup>, A. Roseth<sup>r</sup>, P.M. Schneider<sup>o,1</sup>, T. Sijen<sup>q</sup>, V. Stenzl<sup>i</sup>, D. Syndercombe Court<sup>b,1</sup>, J.E. Templeton<sup>s</sup>, M. Turanska<sup>V</sup>, P.M. Vallone<sup>n</sup>, R.A.H. van Oorschot<sup>k</sup>, L. Zatkalikova<sup>V</sup>, The EUROFORGEN- NoE Consortium; Á. Carracedo<sup>a,1</sup>, C. Phillips<sup>a,1</sup>\*

<sup>a</sup> Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Santiago de Compostela, Spain

<sup>b</sup> Department of Forensic and Analytical Science, Faculty of Life Science, King's College London, UK

<sup>C</sup> Federal Criminal Police Office, Wiesbaden, Germany.

<sup>d</sup> Forensic Genetic and Biology Service, Centre Branch, National Institute of Legal Medicine and Forensic Sciences , Coimbra, Portugal

<sup>e</sup> Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Frederik V's Vej 11, Copenhagen, Denmark

f Section of Forensic Genetics, Institute of Forensic Research, Kraków, Poland

<sup>g</sup> Forensic Genetics Laboratory, Institute of Legal Medicine, Università Cattolica del Sacro Cuore, Rome, Italy

h Forensic Science Laboratory, Dublin, Ireland

<sup>1</sup> Department of Forensic Genetics, Institute of Criminalistics, Prague, Czech Republic

<sup>J</sup> Department of Forensic Molecular Biology, Erasmus MC University Medical Centre Rotterdam, Rotterdam, The Netherlands

<sup>k</sup> Office of the Chief Forensic Scientist, Forensic Services Department, Victoria Police, Australia

l National Institute of Criminalistics and Criminology, Chaussée de Vilvoorde 100, Brussels, Belgium

<sup>m</sup> ESR, Private Bag 92021, Auckland, New Zealand

<sup>n</sup> Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA

<sup>0</sup> Institute of Legal Medicine, Faculty of Medicine, University of Cologne, Cologne, Germany

<sup>p</sup> Zurich Institute of Forensic Medicine, University of Zurich, Zurich, Switzerland

<sup>q</sup> Department of Human Biological Traces, Netherlands Forensic Institute, The Hague, The Netherlands

r Department of Forensic Biology, Norwegian Institute of Public Health, Oslo, Norway

<sup>s</sup> School of Biological Sciences, Flinders University, Adelaide, South Australia 5042, Australia

t Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria

<sup>U</sup> Forensic Science Program, The Pennsylvania State University, University Park, PA, USA

<sup>V</sup> Institute of Forensic Science, Ministry of the Interior, Department of Biology and DNA Analysis, Slovenská Lupca, Slovakia

<sup>1</sup> Also participated in the EUROFORGEN-NoE inter-laboratory ancestry analysis pilot exercise

<sup>2</sup> Participated in the EUROFORGEN-NoE pilot exercise only

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

• Nineteen laboratories completed a collaborative EDNAP exercise to evaluate two forensic ancestry informative marker (AIM) assays and accompanying statistical tools to infer ancestry from the genotype data.

• Laboratories were sent primers, reference data and five test DNAs of undisclosed origin plus an unmarked DNA mixture (but reported to be one of the samples).

• Fourteen laboratories successfully genotyped the DNAs with a 34-plex SNP assay using SNaPshot, achieving 96.1% profile completeness and 93.5% genotype concordance.

• All laboratories successfully genotyped the DNAs with a 46-plex Indel assay using dye-labelled PCR primers, achieving 99.8% profile completeness and genotype concordance.

• All laboratories identified the mixed DNA sample, indicated by disrupted peak height ratios in the Indel profile and three-allele patterns in SNP rs5030240. 18/19 laboratories assigned the correct ancestry to each of the test DNAs of unknown origin, obtaining likelihood ratios from 80 markers in the range: 1.25E+07 to 1.78E+41.

#### 1 2

2 3

## Forensic ancestry analysis with two capillary electrophoresis ancestry informative marker (AIM) panels: Results of a collaborative EDNAP exercise

#### 4 Abstract

5 There is increasing interest in forensic ancestry tests, which are part of a growing number of DNA analyses 6 that can enhance routine profiling by obtaining additional genetic information about unidentified DNA 7 donors. Nearly all ancestry tests use single nucleotide polymorphisms (SNPs), but these currently rely on 8 SNaPshot single base extension chemistry that can fail to detect mixed DNA. Insertion-deletion 9 polymorphism (Indel) tests have been developed using dye-labeled primers that allow direct capillary 10 electrophoresis detection of PCR products (PCR-to-CE). PCR-to-CE maintains the direct relationship 11 between input DNA and signal strength as each marker is detected with a single dye, so mixed DNA is 12 more reliably detected. We report the results of a collaborative inter-laboratory exercise of 19 13 participants (15 from the EDNAP European DNA Profiling group) that assessed a 34-plex SNP test using 14 SNaPshot and a 46-plex Indel test using PCR-to-CE. Laboratories were asked to type five samples with 15 different ancestries and detect an additional mixed DNA sample. Statistical inference of ancestry was 16 made by participants using the Snipper online Bayes analysis portal plus an optional PCA module that 17 analyzes the genotype data alongside calculation of Bayes likelihood ratios. Exercise results indicated 18 consistent genotyping performance from both tests, reaching a particularly high level of reliability for the 19 Indel test. SNP genotyping gave 93.5% concordance (compared to the organizing laboratory's data) that 20 rose to 97.3% excluding one laboratory with a large number of miscalled genotypes. Indel genotyping 21 gave a higher concordance rate of 99.8% and a reduced no-call rate compared to SNP analysis. All 22 participants detected the mixture from their Indel peak height data and successfully assigned the correct 23 ancestry to the other samples using Snipper, with the exception of one laboratory with SNP miscalls that 24 incorrectly assigned ancestry of two samples and did not obtain informative likelihood ratios for a third. 25 Therefore, successful ancestry assignments were achieved by participants in 92 of 95 *Snipper* analyses. 26 This exercise demonstrates that ancestry inference tests based on binary marker sets can be readily 27 adopted by laboratories that already have well-established CE regimes in place. The Indel test proved to 28 be easy to use and allowed all exercise participants to detect the DNA mixture as well as achieving 29 complete and concordant profiles in nearly all cases. Lastly, two participants successfully ran parallel next-30 generation sequencing analyses (each using different systems) and achieved high levels of genotyping 31 concordance using the exercise PCR primer mixes unmodified.

32

33 Keywords: Ancestry; SNPs; Indels; AIMs; Bayes analysis; Principal Component Analysis (PCA)

#### 34 1. Introduction

35

36 DNA-based forensic ancestry tests have the capacity to provide key information about unidentified DNA 37 donors, which can be particularly useful when police investigators do not have reliable eyewitness 38 descriptions or if the STR profiling data fails to give a DNA database match [1]. Therefore, tests for the 39 inference of ancestry can be grouped alongside forensic DNA phenotyping (FDP) tests such as HIrisPlex [2] 40 in a growing array of new technologies that have the potential to take forensic DNA analysis well beyond 41 simple identification [3,4]. For such tests to be effective in routine forensic use they must be sensitive; 42 easy to run using validated DNA detection instruments; and, being mainly composed of binary loci, they 43 should have a reasonably robust way to detect mixed DNA so that apparent heterozygotes are not 44 mistyped. In addition, the genetic data obtained must be easy to interpret. Ideally, it should be 45 straightforward to use the genotypes to calculate a set of Bayes likelihoods for particular ancestries (or 46 phenotypes) in comparison to reference populations whose patterns of genetic variation are already well 47 defined. Although STRs can provide a degree of ancestry information [5,6] and Y-chromosome/mtDNA 48 variation is highly differentiated geographically, there are widely discussed reasons why stand-alone 49 autosomal SNP tests provide more reliable indications of a person's ancestry [7-9].

50 For the last ten years, forensic SNP genotyping has relied on the SNaPshot single base extension system to 51 create relatively large-scale PCR and extension multiplexes followed by capillary electrophoresis (CE) of 52 the dye-labeled products using standard run conditions. In this way, FDP and ancestry analysis tests [2,8-53 12] have been developed using single-tube amplification reactions that are highly sensitive and use 54 validated CE regimes [10,13]. One drawback of SNP genotyping with SNaPshot is the inability to 55 distinguish the highly skewed heterozygote peaks often seen in normal DNA with this technique, from the 56 imbalanced peaks common to mixtures. This is mainly due to the SNaPshot terminator chemistry using 57 dyes with much stronger blue/green fluorescence (G/A) compared to yellow/red (C/T) [14]. Therefore, 58 despite their widespread use and evident sensitivity, forensic SNaPshot tests can be inefficient in 59 detecting mixtures. Indel tests have been developed in recent years for identification [15-17] and ancestry 60 analysis [18-20] detecting dye-labeled PCR products sent directly to CE from the amplification stage (PCR-61 to-CE). The benefits of short amplicon lengths and high levels of multiplexing that SNPs provide, are kept 62 with Indel genotyping in this way. However, peak height ratios in heterozygotes are more balanced within 63 any one locus than those of SNaPshot so mixed DNA is more easily detected from the resulting 64 imbalanced patterns [17]. Two CE-based forensic ancestry tests have been established that offer 65 complimentary characteristics: a SNaPshot assay of 34 ancestry informative marker (AIM) SNPs containing 66 some of the most population-differentiated loci (herein 34-plex, [11]) plus a PCR-to-CE assay of 46 AIM-67 Indels [19] that offers comparable population differentiation to AIM-SNPs, but much greater sensitivity to 68 mixed DNA. This report describes the use of these two assays in an inter-laboratory exercise of 15 69 participants from the European DNA Profiling (EDNAP) group, and 4 overseas participants, organized by 70 the University of Santiago de Compostela (USC). As a preamble to the EDNAP exercise, the EUROFORGEN-71 NoE Consortium ran a similar small-scale inter-laboratory exercise to establish the test framework and 72 gauge the transportability of the assay primer sets. As part of the Consortium's networking remit, the 73 primer mixes used for the EDNAP exercise were purchased, optimized and packaged by USC along with 74 test DNAs with known ancestries (undisclosed to participants). These test components are freely available 75 in trial quantities for the forensic community to assess for themselves (available from USC upon request).

The exercise had three main goals: i) for laboratories to assess the relative ease-of-use and reliability of
 the two assays by genotyping test DNAs, whenever possible, using each participant's own CE regimes; ii)

78 for laboratories to use the statistical ancestry inference tools developed at USC and part of the Snipper 79 data analysis portal [11]; iii) to assess the ability of each assay to detect mixtures by including an 80 unmarked mixed-donor sample amongst the test DNAs. This third goal was analyzed further by assessing 81 the Indel heterozygote peak height balance in normal DNA across the range of participant's laboratory 82 setups, in comparison to peaks in the mixed sample. As well as the 15 European laboratories including 83 USC, two participants were from Australia, one from New Zealand and one from the USA. All but three 84 laboratories had participated in the preceding EDNAP IrisPlex exercise that applied SNaPshot analysis to 85 the genotyping of six FDP SNPs [21]. Five EDNAP laboratories, were part of the EUROFORGEN-NOE pilot 86 ancestry exercise.

#### 87 2. Materials and methods

#### 88 2.1. Primer sets, test DNA samples and assay protocols

89 Six quantified DNA samples (10  $\mu$ l volumes at 0.5 ng/ $\mu$ l) plus primer mixes sufficient for 20 reactions were 90 sent to participants who used their own PCR and SNaPshot reaction components. For the Indel assay, 91 PCRs only required the combination of 2x Qiagen Multiplex PCR Master Mix (Qiagen, Hilden, Germany) 92 with the primer mix and DNA. The SNaPshot PCR and extension primer sets plus the Indel PCR primer mix 93 were prepared as previously described [11,19] and were dispatched with the DNA samples at ambient 94 temperature. Some package transit times outside Europe exceeded one week, but the stability of both 95 SNP and Indel primer sets had been previously assessed for the EUROFORGEN-NOE pilot exercise by 96 carefully testing the profile quality obtained from batches of primers originally sent to the US participant 97 and one in Australia, who were also part of the subsequent exercise.

98 The test DNAs were given anonymized codes and comprised five volunteer donors, each with a different 99 continental origin of: East Asian, European, Oceanian, Native American or African ancestries. With the 100 geographic distribution of these samples, examples of all alleles in 80 markers were observed when 101 genotyped by USC, except SNP: rs1573020 (all A homozygotes) and Indels: rs35451359 and rs33974167 102 (all short-allele 'A' homozygotes) plus rs2307998 (all long-allele 'C' homozygotes). In this way, more than 103 97.6% of component marker alleles could be identifiable in the profiles of the test DNAs. A rare third allele 104 in Indel: rs25584 was found in one test DNA. The sixth test sample was an artificial mixture combining a 105 1:3 ratio of additional European and East Asian volunteer donors (herein M1 and M3 respectively). Note 106 that 34-plex has two tri-allelic SNPs and one: rs5030240 showed three allele patterns in the mixed DNA 107 sample (other examples in [11]).

Participants were told that one sample was mixed and were asked to identify it, then assign ancestries to the others using Bayes analysis and Principal Component Analysis (PCA) in *Snipper*, as detailed in section 2.2. The above primer volumes were sufficient to allow participants to begin their analyses with the 9947A positive control DNA used in many STR kits.

Protocols for PCR, SNaPshot extension reactions and CE were sent in the form of an Excel laboratory calculator (Supplementary File S1) plus fragment mobility panels-and-bins files (Supplementary Files S2) that formed templates for participants to adapt to their own CE regimes when necessary. The 9947A DNA acted as a universal point of reference for the peak patterns typical of both assays and example electropherograms were provided to participants, as shown in Fig. 1. Although Indel amplified fragments separate well using all POP polymer types, participants were recommended to use POP-4 for 34-plex genotyping as peak positions are less well separated at the low size range using POP-7. Supplementary 119 Table S1 lists the CE regimes chosen by participants, indicating that most applied a 3130 or 3500 detector

- 120 with POP-4 (13 and 3 respectively), although two used a 3130 with POP-7 and one successfully typed SNPs
- 121 with a 3100 and POP-6. Lastly, participants were advised that Indel PCR products could require dilution
- 122 prior to CE to obtain optimum peak patterns free from excessive signal pull-up.

#### 123 2.2. Preliminary ancestry checks of test DNAs and use of the Snipper data analysis portal

Although this section reports ancestry analysis results, these analyses were made by USC to evaluate the ancestry of the exercise test samples prior to their dispatch. This process also checked the reference population data supplied and ensured test samples were suitably representative of each of the ancestries the participants were asked to identify.

- 128 The Snipper portal provides a Bayes classifier accessing population reference data in place in the website, 129 including fixed training sets for three, four or five main continental HGDP-CEPH population groups, for 34 130 SNPs and/or 46 Indels (these training set genotypes are provided in Supplementary File S3.2). The fixed 131 data options assess one uploaded profile at a time, which is compared to a training set selected by the 132 user. Partial profiles can be uploaded with NN genotypes (or partial genotypes, e.g. 'CN'). Indel data has 133 an identical framework but with 'AC coded' genotypes, where A=short alleles, C=long and is reserved for 134 novel third alleles. Participants were asked to use the fixed training set option in *Snipper* to make ancestry 135 inferences. However, no guidance was given on choice of training set, which influences calculation of the 136 likelihood ratios (herein LRs). For example, selecting a five-group training set for 34-plex SNP data will lead 137 to lower LRs for East Asian assignments as this marker set lacks AIMs sufficiently differentiated to 138 distinguish Oceanians and Native Americans from East Asians. As a rule-of-thumb, 34-plex profiles are 139 optimally analyzed with three-group data (Africa-Europe-East Asia), Indel profiles provide high ancestry 140 assignment LRs for these groups plus Americans, as this differentiation was targeted in their original 141 selection [19]. When combined 80-marker data is used, the differentiation of the fifth Oceanian 142 population group can be accomplished, although Indel data alone can also distinguish Oceanians with 143 minimal error [19].
- 144 To check the Snipper fixed training sets and test samples used, three ancestry analyses were applied to 145 the genotype data prior to the exercise and results are summarized in Fig. 2. First, the 80-marker 146 reference data was cross-validated with Snipper (each training set profile removed and classified by 147 remaining data). The Fig. 2 upper plot shows the distribution of probabilities in ranked order of log<sub>10</sub> LR 148 values, i.e. the lowest LR from five population comparisons (data in Supplementary File S3.3). The grey 149 line of LR=1 represents balanced odds, so points below this line show misclassifications. East Asian 150 training set profiles gave five misclassifications, all assigned as American (5/226=2.2% error). However, 151 none of their LRs exceeded 750, so applying a threshold value of 1000 led to error-free East Asian 152 assignments, but a non-classification rate of 3.54% (8/226). Fig. 2 indicates the LRs for test samples, 153 mixture donors and 9947A tend to fall in the middle to upper range of training set LRs in nearly all cases.

154 In addition to obtaining LRs, it can be helpful to compare patterns of variation in reference population 155 data to samples of unknown geographic origin by applying *STRUCTURE* and PCA. Both provide an intuitive 156 way to make such comparisons [3,22,23] and can be useful to alert the analyst that a forensic sample of 157 unknown origin may be from an admixed individual with co-ancestry. Following review of the 158 EUROFORGEN-NOE ancestry exercise results, a two-dimensional PCA module (plotting the first two 159 principal components or PCs) was developed for *Snipper* that allows analysis of multiple profiles plotted 160 directly onto reference data. The *Snipper* output lists the Bayes analysis data for each profile and their positions are labeled on a PC1-PC2 PCA plot (no PC3 estimates are currently made). Participants were provided with the input file of training set genotypes and a link to the *Snipper* PCA module to enable graphical analysis of test DNAs and 9947A.

164 The middle graphics of Fig. 2 show STRUCTURE cluster plots (an optimum K=5 genetic clusters inferred 165 from data) matched to the order of training set LRs charted above. The enlarged cluster plots for samples 166 A-E on the right indicate an absence of co-ancestry, i.e. their cluster plots have almost no membership to 167 multiple genetic clusters. Likewise, cluster plots on the left for mixture components M1 and M3 show no 168 multiple cluster membership, whereas sample F has approximately equal joint membership to the 169 relevant clusters. The lower graphics show three PCAs made with Snipper, with reference cluster colours 170 matched to the STRUCTURE data. PCA plot A is a 3-group analysis of sample F and M1-M3 components. 171 The position of F highlights the fact that population admixture and mixed DNA can give indistinguishable 172 PCA patterns, emphasizing the need to efficiently detect mixed DNA in forensic ancestry analysis. PCA plot 173 B is a 5-group analysis showing samples A-E plus 9947A are distributed into their expected clusters, 174 although in these 2-PC plots the Oceanian and American clusters show some overlap with East Asians. To 175 better differentiate these three groups, a PCA can be made of just three possible groups to obtain a more 176 distinct separation, as shown in PCA plot C analyzing the three test samples from less differentiated 177 population groups.

178

- 179 **3. Results**
- 180

182

#### 181 3.1. Genotyping performance of the SNP assay

183 Supplementary Table S1 summarizes the CE regimes used by participants and indicates five did not pursue 184 SNaPshot genotyping of SNPs but elected to just analyze and report Indel genotypes. Given the 185 complexities of reading electropherograms consisting of 32 peak pairs plus two triple-peak positions, this 186 was considered to be a reasonable decision and Indel data alone was collected from these laboratories. 187 The number of SNaPshot no-calls and miscalls recorded for the five test samples A-E, from 14 participants 188 reporting SNP data, are summarized in Fig. 3A. SNPs are listed in order of decreasing genotyping 189 performance for participants, by ranking loci in increasing miscall rate followed by increasing no-calls. 190 Therefore, rs2065160, rs3785181 and rs8986788 are the most robustly genotyped SNPs in 34-plex, with 191 all 14 laboratories identifying peaks in five samples, although laboratory #17 had one genotype miscall in 192 each SNP. At the other extreme, rs239031 was both the most difficult SNP to genotype and the least 193 reliably genotyped, with laboratories #8 and #21 not assigning genotypes to all or most samples, bringing 194 the overall call rate down to 80%, well below those of the other 33 SNPs. Genotyping concordance for 195 rs239031 was also the lowest, with 81.4% of genotypes correctly called. High no-call rates for certain 196 other SNPs tended to cluster with participants: rs1573020 was not genotyped in laboratory #21; rs881929 197 in #20; rs1886510 and rs2304925 in #19, despite other laboratories genotyping these SNPs without 198 problems. Only 1 of 5 genotypes was called by laboratory #13 for rs182549. Average SNP call and 199 genotype concordance rates shown at the bottom right of Fig. 3A reached 96.3% and 93.5% respectively. 200 The genotype completeness of ~96% equates to approximately one missing SNP call per 34-plex profile. 201 Laboratory #17 had evident problems recognizing and accurately calling their SNaPshot 202 electropherograms with less than half the successful genotype calls made by the other participants 203 reporting SNPs. Therefore, when considering concordance amongst 13/14 participants, the value rose to 204 97.3%. Although one other laboratory #6 had slightly below-average genotyping concordance, no obvious 205 connection could be made between the CE regimes used by participants and miscalls seen in certain SNPs. 206 Nevertheless, there are known issues previously recognized at USC in some 34-plex SNPs and several of 207 these were observed in the electropherograms from participants. Certain mobility or non-specific peak 208 patterns can explain a proportion of the genotype miscalls and these are outlined next.

209

210 Examples of three different challenges for SNP genotyping with 34-plex are shown in Supplementary Fig. 211 S1. First, SNPs rs10843344-rs239031 run to positions very close together, with the C peak of rs239031 212 often having a mobility shift that places it very close to the much higher C peak of rs10843344 213 (Supplementary Fig. S1.1). The same signal imbalance can be seen in the T peaks but the electrophoretic 214 separation of these peaks remains more distinct. Examination of participant's SNaPshot profiles indicated 215 some laboratories had missed the lower, shifted rs239031-C peak. Second, rs182549, rs881929 and 216 rs3827760 have particularly low signal strengths (Supplementary Fig. S1.2) and the three SNPs show 217 higher than average no-call rates. In the case of rs3827760, there is a very marked disparity in peak 218 heights between the higher East Asian-informative G allele and the A allele (> 10:1 peak height ratio in the 219 example shown), so this SNP requires particular care. Third, rs2304925 shows an artifactual G signal in the 220 negative control very close to the G peak of rs5030240 (Supplementary Fig. S1.3). This peak is much 221 higher than the T peak of rs2304925 when it is a true allelic extension product but much lower when 222 artifactual. All participants ran a negative control and most recognized the extra G signal running close to 223 the G peak of rs5030240, although as this is a tri-allelic SNP, when a homozygous A or C allele is present the genotypes can be mistyped as an AG or CG in the absence of the stronger G peak with which to compare the artifact signal.

226

#### 227 3.2. Genotyping performance of the Indel assay

228 All 19 participants successfully completed the genotyping of the samples with Indels. Supplementary 229 Table S1 shows that almost half of the laboratories chose to dilute the PCR products 1:5-1:20 prior to CE 230 detection to control signal pull up. Supplementary Figs. S1.4-5 show two examples of minor challenges 231 with genotyping of Indels, consisting of the occurrence of dye blobs (broad non-specific peaks around 232 allele peak positions), identifiable in the negative control, plus signal pull-ups that can occur when the 233 Indel PCR products are not sufficiently diluted. However, there was no evidence that these two profile 234 phenomena interfered with the genotyping performance of the Indel tests in any of the 19 laboratories. In 235 fact, the genotyping completeness and concordance were very high when considering that most 236 participants were running the test for the first time and required reading 46 different peak sets in each 237 electropherogram.

Fig. 3B summarizes the Indel genotyping performance and shows participants achieved a very high overall genotyping completeness and concordance rate of 99.8%. Fourteen participants did not have miscalls or no-calls in any test sample profiles. A slight degree of clustering of genotyping miscalls and no-calls is discernible in Fig 3B; for example, laboratory #20 chose not to call 3/5 rs2307922 genotypes, and laboratories #1 and #7 mistyped more than one Indel. It is notable that all 19 participants successfully identified the rare third allele of rs25584 present in test sample C.

#### 244 *3.3. Inference of ancestry*

245 All participants identified F as the mixed DNA sample and made Bayes analysis to infer the ancestry of 246 samples A-E using Snipper. The majority, but not all, also made comparisons of the genotypes from A-E 247 with the Snipper PCA module using the supplied reference population data. This section summarizes 248 results for all laboratories using both statistical approaches to illustrate that the SNP and Indel data has a 249 degree of ancestry-informativeness redundancy, i.e. the Bayes LRs or PCA positions of samples A-E are 250 very similar despite some genotype miscalls or missing data. Therefore, the ancestry inferences made by 251 participants were correct in all cases apart from those of laboratory #17 that made incorrect ancestry 252 inferences for two samples and had PCA positions markedly displaced from the others in most cases.

253 Fig. 4A summarizes SNP profile quality (bar-charts, left-hand scale); Bayes LRs (points superimposed on 254 bars, right-hand scale); and PCA positions for the SNaPshot assay data of 14 participants, analyzing 255 samples A-E. Bayes LRs and PCAs from SNP data alone compare African, European and East Asian 256 ancestries; consequently C and D give lower LRs and edge-of-cluster PCA positions that suggest East Asian 257 ancestry despite these being Oceanian and American in origin. For 13/14 laboratories, samples A, B and E 258 give mid-cluster PCA positions and high LRs that varied by four orders of magnitude between 1E+14 to 259 1E+18 correctly assigning A as East Asian and B as European, and 1E+22 to 1E+26 correctly assigning E as 260 African. The LR values obtained by coordinating laboratory USC for SNP and Indel data are outlined in 261 Table 1 (the 80-marker LRs for all samples are given separately in Fig. 2). Table 1 indicates sample C gave a 262 high LR for Oceanian ancestry with just Indel data used in a 5-group comparison.

Fig. 4B summarizes Indel profile quality, Bayes LRs and PCA positions for a four group comparison using the Indel data of all participants. A sixth PCA plot, bottom right, shows the combined 80-marker analysis 265 for Oceanian sample C. Apart from African sample E, Indel data gives lower LRs than SNPs and the LRs for 266 samples A and E are from different population likelihoods (bold values in Table 1). The improved 267 genotyping consistency of Indels amongst participants is reflected in more uniform sets of Bayes LRs and 268 PCA positions that mainly overlay each other (i.e. seen as single points on plots). For the two laboratories 269 with three Indel miscalls, an effect is seen in the Bayes LRs for American sample D and African sample E, 270 with some PCA displacement, indicating that even with just two markers miscalled it can sometimes affect 271 the statistical inference made from other correctly called genotypes (~97% of the data). The Oceanian 272 sample C was correctly identified by 18 participants, with many using both Indel and combined data to 273 make the inference.

Therefore, 18 of 19 laboratories were able to successfully assign ancestry to five samples of undisclosed
 geographic origin, obtaining unequivocal Bayes LRs and, in most, cases participants constructed PCA plots
 providing supplementary analyses with good matches to the Bayes results.

#### 277 3.4. Mixture detection and analysis of participant's Indel peak height data

278 Although the exercise was not a fully blinded test (i.e. where the presence of a mixed sample is not 279 disclosed), all participants were able to identify sample F as the mixture from the observation of 280 imbalanced signals in the heterozygote peak pairs of the Indel profile. Therefore, despite a lack of 281 familiarity with Indel peak patterns in most laboratories, there was sufficient contrast between the mixed 282 sample F and the unmixed A-E DNAs for the mixture to be discernible by all participants. In addition, 7 of 283 14 laboratories reported an ACG triple-peak pattern in the tri-allelic SNP rs5030240, one reported an AC 284 with possible G, one a GG result and the other five gave no-calls. A typical sample F peak pattern for 285 rs5030240 is shown in Supplementary Fig. S1.6.

286 The detection of peak height imbalances that can indicate mixed DNA has been stated to be an advantage 287 of direct PCR-to-CE Indel genotyping compared to SNaPshot tests [15,17,19], however such patterns have 288 not been properly assessed across a range of CE detectors. For this reason, we decided to ask participants 289 to provide their heterozygote peak height data and then compiled the variation in peak height ratios 290 (PHRs, highest/lowest peaks) recorded in the five unmixed and single mixed DNAs from the range of CE 291 regimes used. Furthermore, when analyzing binary markers the number of heterozygotes observed in 292 mixtures is invariably higher than normal unmixed samples. Although PHR values were distinct between 293 A-E and F, three factors complicated the straightforward statistical comparison of patterns of 294 heterozygosity observed amongst the test samples. First, there was variation in the number of 295 heterozygotes recorded in sample F. Specifically, laboratory #1 identified 18 heterozygotes; #15: 17; #18: 296 21; and #20: 17, compared to an average number of heterozygotes identified by the other fifteen 297 laboratories of 27. Second, the lower number of identified heterozygotes for F in some participant's data 298 affects the minimum-maximum and average PHR values, particularly when the PHR is extreme and a very 299 low peak is discounted when reading the profile. Four example peak pairs that were recorded as single 300 allele genotypes by one participant but as heterozygotes by the others, are shown in Fig. 5A. Third, due to 301 the contrasting frequencies of most of the 46 Indels between population groups, sample A showed lower 302 numbers of heterozygotes and sample B higher numbers than those seen in C-E.

The numbers of heterozygotes and PHR values are plotted in Fig. 5B. This chart shows data from 15/19 laboratories (excluding #1, #15, #18 and #20). The same chart with all 19 participant's data is shown in Supplementary File S4.A. The dark grey bars mark the data from 3500 detectors and indicates that no difference in peak height ratios are discernible in comparison with 31xx CE data. 307 Statistical assessment of the number of heterozygotes in A-E vs. F was made with a unilateral 2-sample 308 test for equality-of-proportions (with continuity correction). The resulting grid of p-values for pairwise 309 comparisons across all 19 laboratories is shown in Supplementary File S4.B, along with the Fig 5B chart re-310 plotted for full data from all laboratories (Supplementary File S4.A). It can be seen from the 311 Supplementary File S4.A chart that the numbers of sample F heterozygotes recorded by laboratories #1, 312 #15 and #20 is lower than the average number in unmixed sample B. Inclusion of this data has a direct 313 effect on the distribution of significant *p*-values obtained from pairwise comparisons. Laboratories #1, #15 314 and #20 sample F heterozygote numbers are significantly different to those of most of the other 315 laboratories, but not different to heterozygote numbers in unmixed samples B-E, while #18 data for 316 sample F is not significantly different to samples B and C. The high number of heterozygotes in sample B is 317 reflected in significant differences only found for comparisons to those of laboratories #8, #13, #14, and 318 #5, who recorded 29 or more heterozygotes in their sample F profiles. Therefore, we opted to remove #1, 319 #15, #18 and #20 data from the statistical assessment of PHR differences between A-E and F.

320 The average PHRs shown in Fig. 5B indicate a quite distinct contrast between samples A-E and F, with 321 values of 1.15 compared to 3.14 respectively, which suggests a ratio of 1:2.73 that approximates the 322 actual 1:3 contributor ratio well. Although the PHR values give a clearly discernible difference between 323 mixed and unmixed samples, we completed a formal statistical test of this difference. An ANOVA test is a 324 standard approach for assessing continuous values such as PHR measurements, but a Shapiro-Wilks test 325 indicated some of the data was not normally distributed (data not shown). Therefore, a Kruskal-Wallis 326 rank sum test was applied and the grid of pairwise p-values comparing the average PHRs of A-E with 327 individual PHRs of F is shown in Supplementary File S4.C. The results are completely consistent: the 328 pairwise comparisons of mixed vs. unmixed PHRs give significant p-values in every case and none were 329 detected for comparisons within each sample set.

In summary, despite the need to adjust statistical comparisons by removing 4 of 19 participant's data due to under-reported heterozygote peak pairs, the other laboratories provided a ratio of average peak heights close to 1:3. This ratio is consistent with the mixture that was constructed for the exercise and is statistically significant for all signal strength comparisons made.

334 3.5. Additional Next Generation Sequencing experiments applied to test DNAs by two laboratories

Two laboratories decided to use their remaining PCR primers to genotype one or both marker sets with different Next Generation Sequencing (NGS) systems, as outlined in Supplementary File S5. One assessed 34-plex SNP typing using an unmodified PCR followed by library preparation and massively parallel sequencing with the Illumina MiSeq system. The other assessed 34-plex SNPs and Indel genotyping in the same way (unmodified PCR in each case) with the Thermo Fisher Scientific-Life Technologies (TFS-LT) Ion PGM<sup>™</sup> system.

341 The 34-plex SNP sequence analyses were successful to a very large degree, as all genotypes were 342 identified and almost fully concordant with each laboratory's SNaPshot data. Sample F was observed to 343 be distinct in a major proportion of its allele-pair sequence ratios (defined as the second allele exceeding 344 10% of sequence reads), compared to A-E. Supplementary File S5 indicates there were only 5/14 345 sequence ratios of 1.5 or less (i.e. in the range: 0.4:0.6-0.5:0.5) in the Ion PGM™ data and 3/17 in the 346 MiSeq data. This equates to 64% and 82% of sequence ratios exceeded those of most normal DNA 347 heterozygotes seen in Ion PGM<sup>™</sup> and MiSeq respectively, giving unequivocal signals of a mixture in F. 348 Both systems also detected displaced sequence ratios in each of the two tri-allelic SNPs.

349 The Indel analysis with NGS gave three discordant genotypes in samples B and C, plus an average 8.7% no-350 calls (coverage too low) and 2.9% missing data (undetected sequence), although not all samples gave the 351 same non-detection rates. Overall, 84% of the NGS genotypes matched the CE calls. However, the 352 alignment of sequences that contain short insertions and deletions is particularly challenging in NGS 353 sequence analysis and it was not possible to be sure how many miscalls or no-calls were due to 354 misalignment issues. Supplementary File S5 shows assessments of Indel sequence ratios for sample F 355 compared to A-E. Given that sequence coverage was low in some loci and this is the first NGS experiment 356 with this type of forensic marker, results need cautious interpretation. However, patterns suggest a 357 degree of displacement in F away from the perfect sequence balance midline (0.5:0.5) compared to many

358 of the heterozygote sequence ratios detected in A-E.

#### 359 4. Discussion

360 As forensic NGS analysis gains greater traction, it is the right moment for the forensic community to use 361 inter-laboratory exercises to assess the binary marker sets that will start to add complementary genetic 362 data to conventional STR polymorphisms. Ancestry inference is seen as a key part of the enhanced 363 characterization of forensic DNA that NGS will allow. Therefore, it is important to evaluate the robustness 364 of existing CE-based ancestry-informative SNP and Indel multiplexes in terms of how easily they can be 365 adopted in laboratories not previously experienced with binary marker genotyping. The statistical analysis 366 of the genotype data obtained from AIMs also needs to be easy to use and interpret by forensic 367 laboratories. The most straightforward approach for inferring ancestry uses Bayesian LR comparisons 368 between the two geographic origin hypotheses with the highest likelihoods. Lastly, binary variation has a 369 reduced capacity to detect mixtures since homozygotes in combination can look like heterozygotes and 370 only a few non-binary SNPs or Indels currently offer the chance to observe more than two alleles. 371 Therefore, the exercise findings for genotyping reliability, ease-of-use of the recommended ancestry 372 inference tools and ability to detect mixed DNA are all relevant to the progress towards adoption of AIMs 373 in forensic analysis.

The principal finding of this exercise was that each of the participants readily established the AIM-Indel 46-plex test in their laboratory. All participants achieved good quality profiles that reached the high level of genotyping concordance of 99.8% and then efficiently detected mixed sample F. In contrast, SNaPshot typing was both more challenging and for many participants less reliable, despite most laboratories having successfully genotyped six SNPs for the preceding *IrisPlex* EDNAP exercise [21]. Miscalled genotypes with SNaPshot produced an overall genotyping concordance rate of 97.3% when a single participant's results were excluded (13/14 laboratories).

381 We have no explanation for the very high number of SNP miscalls from this one laboratory but it resulted 382 in their statistical analyses producing the only incorrect ancestry inferences for two test DNAs and one 383 uninformative LR of 1.2. All other participants produced correct ancestry predictions from the Bayes LRs 384 calculated in *Snipper* and, for those that created PCA plots, obtained cluster patterns and profile positions 385 that corresponded to these LRs. Therefore, from the review of exercise reports returned from 19 386 laboratories, we can recommend the use of both of these statistical approaches to ancestry inference, as 387 these proved easy to use and allowed correct ancestry assignments of samples with undisclosed 388 geographic origin in 92 of 95 cases.

389 Mixture detection achieved from Indel peak patterns was particularly successful, with sample F giving a 390 clear signal of mixed DNA for all participants. Our analysis of peak height ratios made after the exercise 391 finished, gave a good approximation of the actual mixture component ratio, averaging 1.15 and 3.14 for 392 PHRs in unmixed samples and the mixture respectively. The much higher number of heterozygotes in F 393 could mainly be due to the different ancestries of the mixture contributors. Nevertheless, recording a 394 higher number of heterozygotes than in normal DNA samples and observing PHRs markedly above ~1.2 395 gives a simple and easily adopted system to detect mixtures with Indels. SNaPshot does not offer the 396 same direct relationship between peak heights and input DNA so there is a risk that simple two-person 397 mixtures mimic the patterns seen in individuals with co-ancestry due to population admixture, as revealed 398 by the PCA plot of sample F in Fig. 2 (plot A). Obviously, single sample experiments are not fully indicative 399 of how well Indels will perform with a range of forensic samples, mixture ratios or component ancestry 400 combinations, but the fact that most participants were running Indels for the first time and all detected 401 the mixture indicates sensitivity to mixed DNA with this assay.

402 Although the NGS findings from two participants are a set of parallel genotyping experiments using 403 exercise materials that were not part of the study plan, results are included in this report to highlight the 404 enhanced sensitivity to mixtures obtained for SNP analysis with NGS. It is also interesting to note that 405 existing optimized forensic multiplexes work very well in NGS without the need for any modification, 406 confirming the results of a recent study that found the 34-plex PCR primers, amongst four other forensic 407 SNP multiplexes, provide good quality output with the Ion PGM<sup>™</sup> system [24]. In addition, the relative 408 success of the initial Indel genotyping experiments with NGS indicate dye-labeled PCR primers do not 409 interfere with library preparation and subsequent sequencing chemistry of the Ion PGM<sup>™</sup>. This suggests 410 existing forensic CE multiplexes for a range of markers, including STR kits, could be used to prepare target 411 DNA for experimental NGS sequencing runs.

Until NGS systems that incorporate AIMs are widely adopted for forensic use, the results from this EDNAP inter-laboratory exercise indicate the PCR-to-CE Indel test is by far the best current option for forensic ancestry analysis. The Indel multiplex provides a simple, reliable and informative test from a comparatively large marker set that is analyzed using validated CE regimes. Detection of simple twocomponent mixed DNA from scrutiny of Indel peak patterns was a task accomplished by all exercise participants and gives Indel genotyping a key additional advantage over SNP-based ancestry tests.

#### 418 Acknowledgements

419

420 This work was partially supported through a grant to CS (SFRH/BD/75627/2010) awarded by the 421 Portuguese Foundation for Science and Technology (FCT) and co-financed by the European Social Fund 422 (Human Potential Thematic Operational Programme). The organizing laboratory of the Forensic Genetics 423 Unit, University of Santiago de Compostela, would like to thank Fabio Oldoni for his assistance in piloting 424 the exercise, in partial fulfillment of an MSc in Forensic Science, Department of Forensic Science and Drug 425 Monitoring, King's College, London. The two anonymous donors for the mixture are also thanked for their 426 willingness to provide samples. Zurich Institute of Forensic Medicine thanks Corinne Moser for technical 427 assistance. The pilot studies, Indel primers and test DNAs used in the exercise were funded by the 428 European Forensic Genetics Network of Excellence (EUROFORGEN-NOE, Grant Agreement No. 285487).

- 429 References
- 430

[1] C. Phillips, L. Prieto, M. Fondevila, A. Salas, A. Gomez-Tato, J.A. Alvarez-Dios, A. Alonso, A. BlancoVerea, M. Brión, M. Montesino, Á. Carracedo, M.V. Lareu, Ancestry analysis in the 11-M Madrid bomb
attack investigation. PLoS One 4 (2009) e6583.

434 [2] S. Walsh, F. Liu, A. Wollstein, L. Kovatsi, A. Ralf, A. Kosiniak-Kamysz, W. Branicki, M. Kayser, The
435 HIrisPlex system for simultaneous prediction of hair and eye colour from DNA, Forensic Sci. Int. Genet. 7
436 (2013) 98–115.

437 [3] M. Kayser, P. de Knijff, Improving human forensics through advances in genetics, genomics and438 molecular biology, Nat. Rev. Genet. 12 (2011) 179–192.

[4] M. Kayser, P.M. Schneider, DNA-based prediction of human externally visible characteristics in
forensics: motivations, scientific challenges, and ethical considerations, Forensic Sci. Int. Genet. 3 (2011)
154–161.

442 [5] C. Phillips, L. Fernandez-Formoso, M. Garciañas, L. Porras, T. Tvedebrink, J. Amigo, M. Fondevila, A.
443 Gomez-Tato, J. Alvarez-Dios, A. Freire-Aradas, A. Gomez-Carballa A. Mosquera-Miguel, Á. Carracedo, M.V.
444 Lareu, Analysis of global variability in 15 established and 5 new European Standard Set (ESS) STRs using
445 the CEPH human genome diversity panel, Forensic Sci. Int. Genet. 5 (2011) 155–169.

[6] C. Phillips, M. Gelabert-Besada, L. Fernandez-Formoso, M. García-Magariños, C. Santos, M.Fondevila,
D. Ballard, D. Syndercombe Court, Á. Carracedo, M.V. Lareu, "New turns from old STaRs": Enhancing the
capabilities of forensic short tandem repeat analysis, Electrophoresis 35 (2014) 3173-3187.

[7] T.E. King, E.J. Parkin, G. Swinfield, F. Cruciani, R. Scozzari, A. Rosa, S.K. Lim, Y. Xue, C. Tyler-Smith, M.A.
Jobling, Africans in Yorkshire? The deepest-rooting clade of the Y phylogeny within an English genealogy,
Eur. J. Hum. Genet. 15 (2007) 288-293.

[8] O. Lao, P.M. Vallone, M.D. Coble, T.M. Diegoli, M. van Oven, K.J. van der Gaag, J. Pijpe, P. de Knijff, M.
Kayser, Evaluating self-declared ancestry of U.S. Americans with autosomal, Y-chromosomal and
mitochondrial DNA, Hum. Mutat. 31 (2010) E1875–E1893.

[9] D. Corach, O. Lao, C. Bobillo, K. van Der Gaag, S. Zuniga, M. Vermeulen, K. van Duijn, M. Goedbloed,
P.M. Vallone, W. Parson, P. de Knijff, M. Kayser, Inferring continental ancestry of Argentineans from
autosomal, Y-chromosomal and mitochondrial DNA, Ann. Hum. Genet. 74 (2010) 65–76.

[10] C. Bouakaze, C. Keyser, E. Crubézy, D. Montagnon, B. Ludes, Pigment phenotype and biogeographical
ancestry from ancient skeletal remains: inferences from multiplexed autosomal SNP analysis, Int. J. Leg.
Med. 123 (2009) 315-325.

[11] M. Fondevila, C. Phillips, C. Santos, A. Freire Aradas, P.M. Vallone, J.M. Butler, M.V. Lareu, Á.
Carracedo, Revision of the SNP*for*ID 34-plex forensic ancestry test: assay enhancements, standard
reference sample genotypes and extended population studies, Forensic Sci. Int. Genet. 7 (2013) 63-74.

464 [12] C. Phillips, A. Freire Aradas, A.K. Kriegel, M. Fondevila, O. Bulbul, C. Santos, F. Serrulla Rech, M.D.

- 465 Perez Carceles, Á. Carracedo, P.M. Schneider, M.V. Lareu, Eurasiaplex: A forensic SNP assay for 466 differentiating European and South Asian ancestries, Forensic Sci. Int. Genet. 7 (2013) 359–366.
- 467 [13] M. Fondevila, C. Phillips, N. Naveran, L. Fernandez, M. Cerezo, A. Salas, Á. Carracedo, M.V. Lareu,
  468 Case report: identification of skeletal remains using short-amplicon marker analysis of severely degraded
  469 DNA extracted from a decomposed and charred femur, Forensic Sci. Int. Genet. 2 (2008) 212–218.
- [14] C. Børsting, E Rockenbauer, N. Morling, Validation of a single nucleotide polymorphism (SNP) typing
  assay with 49 SNPs for forensic genetic testing in a laboratory accredited according to the ISO 17025
  standard, Forensic Sci. Int. Genet. 4 (2009) 342–42.
- 473 [15] R. Pereira, C. Phillips, C. Alves, A. Amorim, Á. Carracedo, L. Gusmão, A new multiplex for human
  474 identification using insertion/deletion polymorphisms, Electrophoresis 30 (2009) 3682–3690.
- [16] S.L. Friis, C. Børsting, E. Rockenbauer, L. Poulsen, S.F. Fredslund, C. Tomas, N. Morling, Typing of 30
  insertion/deletions in Danes using the first commercial indel kit—Mentype<sup>®</sup> DIPplex, Forensic Sci. Int.
  Genet. 6 (2012) e72-e74.
- [17] M. Fondevila, C. Phillips, C. Santos, R. Pereira, L. Gusmão, Á. Carracedo, J.M. Butler, M.V. Lareu, P.M.
  Vallone, Forensic performance of two insertion-deletion marker assays, Int. J. Legal Med. 126 (2012) 725737.
- [18] N.P. Santos, E.M. Ribeiro-Rodrigues, A.K. Ribeiro-dos-Santos, R. Pereira, L. Gusmão, A. Amorim, J.F.
  Guerreiro, M.A. Zago, C. Matte, M.H. Hutz, S.E. Santos, Assessing individual interethnic admixture and
  population substructure using a 48-insertion-deletion (INDEL) ancestry- informative marker (AIM) panel.
  Hum. Mutat. 31 (2010) 184–190.
- [19] R. Pereira, C. Phillips, N. Pinto, C. Santos, C.E.B. Santos, A. Amorim, A. Carracedo, L. Gusmão,
  Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion
  deletion multiplexing, PLoS One 7 (2012) e29684.
- 488 [20] D. Zaumsegel, M.A. Rothschild, P.M. Schneider, A 21 marker insertion deletion polymorphism panel
  489 to study bio-geographical ancestry, Forensic Sci. Int. Genet. 7 (2013) 305-312.
- 490 [21] L. Chaitanya, S. Walsh, J.D. Andersen, R. Ansell, K. Ballantyne, D. Ballard, R. Banemann, C.M. Bauer,
  491 A.M. Bento, F. Brisighelli, et al., Collaborative EDNAP exercise on the IrisPlex system for DNA-based
  492 prediction of human eye colour, Forensic Sci. Int. Genet. 11 (2014) 241-251.
- 493 [22] J. Novembre, M. Stephens, Interpreting principal component analyses of spatial population genetic494 variation, Nat. Genet. 40 (2008) 646-649.
- 495 [23] D. Reich, A. Price, N. Patterson, Principal component analysis of genetic data, Nat. Genet. 40 (2008)496 491–492.
- 497 [24] R. Daniel, C. Santos, C. Phillips, M. Fondevila, R.A. van Oorschot, Á. Carracedo, M.V. Lareu, D.
  498 McNevin, A SNaPshot of next generation sequencing for forensic SNP analysis, Forensic Sci. Int. Genet. 14
  499 (2015) 50–60.

**Table 1.** Lowest LR values produced from *Snipper* Bayes analysis of the full SNP and Indel profiles of samples A-E and 9947A with their ancestry inferences. Participant LR values for the same samples are plotted in Figs. 4A/B. Bold values for A and E highlight different population ratios giving the lowest LRs when SNP, 3-group or Indel, 4-group comparisons are made. With Indel, 4-group comparisons the second lowest LRs for samples A and E are based on the same population ratios as the lowest LRs for SNP, 3group comparisons. Sample C is correctly inferred to be Oceanian with Indel data alone but most participants reported the LR from 80 marker data.

507

Inference:	34-plex SNPs, 3-group	
European	9947A is 2,118,840,589,047,061,020,672 times more likely EUROPEAN than E ASIAN	
East Asian	A is 361,148,635,069,545,024 times more likely E ASIAN than EUROPEAN	
European	B is 64,191,487,284,485,608 times more likely EUROPEAN than E ASIAN	
East Asian	C is 13,115,706 times more likely E ASIAN than AFRICAN	
East Asian	D is 248,539,593,557 times more likely E ASIAN than EUROPEAN	
African	E is 556,454,701,312,037,054,117,314,560 times more likely AFRICAN than E ASIAN	
	46-plex Indels, 4-group	46-plex, 4-group (second lowest LR)
European	9947A is 1,937,432,967,198 times more likely EUROPEAN than E ASIAN	
East Asian	A is 6,993,957 times more likely E ASIAN than AMERICAN	A is 37,290,377,821,078,192,128 times more likely E ASIAN than EUROPEAN
European	B is 143,659,679,122 times more likely EUROPEAN than E ASIAN	
LR too low	C is 131 times more likely E ASIAN than EUROPEAN	
American	D is 944,698,134 times more likely AMERICAN than E ASIAN	
African	E is 3,229,841,442,838,053,650,432 times more likely AFRICAN than EUROPEAN	E is 5,715,694,248,335,998,122,459,136 times more likely AFRICAN than E ASIAN
	46-plex, 5-group	80 Markers, 5-group
Oceanian	C is 24,880,402 times more likely OCEANIAN than E ASIAN	C is 153,747,536,542,653 times more likely OCEANIAN than E ASIAN

508

- 509 Figure legends
- 510

Fig. 1. Electropherograms from the Indel test (upper panel) and the 34-plex SNP test for the 9947A control
 DNA. Peak positions are labeled with the internal codes used for each marker (internal code-rs-number
 lists are provided in Fig. 3A; Supplementary Files S3; *Snipper* and in [11]).

514

515 Fig. 2. Ancestry analysis of exercise test samples. 80-marker genotypes were analyzed and HGDP-CEPH 516 training set data was as supplied to participants (Supplementary File S3.2). Top plot shows ranked 517 Snipper Bayes analysis LRs from training set cross validation or test profile analysis (black points). Grey 518 points in East Asians/Oceanians indicate LRs below a threshold value of 1000 (the grey shaded log LR 519 range around balanced odds line of LR=1). Red points indicate East Asian training set LRs that misclassified 520 as Americans. Middle plots show STRUCTURE analysis aligned directly to the LR distributions above 521 with separate plots for mixture components, left and test samples, right. Lower plots show 2D PCA 522 analyses of test samples in 3-group or 5-group comparisons. Plot A shows a 3-group comparison of sample 523 F, positioned mid-cluster between contributors M1 and M3. Plot B shows the full 5-group PCA of samples 524 A-E plus 9947A. Plot C shows a restricted comparison of just East Asian, Oceanian and American data to 525 obtain better differentiation of reference population clusters and A, C, D; all more closely distributed in 526 plot B.

527

Fig. 3. (A) Genotyping performance of the 34-plex test arranged by SNP (rows) and by 14 participants (columns). Cells record miscalls on the left, and no-calls right. The bar plots on the right summarize total genotype completeness and concordance for each SNP and at the bottom, for each participant. SNPs are ordered by diminishing performance (i.e. decreasing concordance then completeness). Overall genotype concordance is given for 14 and 13 laboratories separately, excluding participant #17 with a very high number of SNP miscalls. (B) Genotyping performance grid for Indel test data from all 19 laboratories. 534 Miscalls are shown as dark grey cells, no-calls light grey.

535

536 Fig. 4. (A) Participant's SNP-based Bayes LRs and PCA positions for three-group comparisons (AFR-EUR-E 537 ASN) analyzing samples A-E. Genotype completeness and concordance rates are shown as bar charts (left-538 hand scales) and ancestry assignment LRs (i.e. lowest values) as overlaid points (right-hand scales). 539 Laboratories with some displacement of a sample position from the main PCA cluster are individually 540 labeled and incorrect positions/assignments from miscalled genotypes are shown in red. (B) Participant's 541 Indel-based Bayes LRs and PCAs for 4-group comparisons (including Americans) analyzing A-E. The sixth 542 plot, lower right, shows a 5-group PCA of sample C (adding Oceanian reference data) using 80-marker 543 genotypes. Laboratories only reporting Indel data have Bayes LRs shown in green and one uninformative 544 LR shown in blue.

545

Fig. 5. (A) Example Indel peak pairs for sample F discounted as heterozygotes by one participant. (B)
 Numbers of Indel heterozygotes (bars) and their peak height ratios (PHR: points) recorded by 15
 participants. Unmixed samples A-E are average values from all data and sample F values are shown
 individually as different numbers of peak pairs were recognized as heterozygotes amongst laboratories.

- 550 Supplementary Files

**Supplementary Figs. S1.** Examples of genotyping challenges in 34-plex or Indel profiles.

- **Supplementary Table S1.** Capillary electrophoresis (CE) details for participating laboratories. Grey bars for 34-plex denote five participants not completing SNP genotyping with SNaPshot.

**Supplementary File S1.** Laboratory protocol guide in the form of an Excel calculator for reaction setups 558 provided to exercise participants.

**Supplementary File S2.** CE fragment mobility panels-and-bins files provided to exercise participants.

**Supplementary Files S3.** SNP and Indel genotypes used in the exercise as reference population data plus test DNA data established by USC. Worksheets are:

- **File S3.1** PCA input
- **File S3.2** The 5-group training set data for Bayes analysis
- **File S3.3** Cross-validation data plotted in Fig. 2.
- 567 Note that to use the files for *Snipper* analysis each must be moved to 'worksheet position 1'.
- 568 PCA: <u>http://mathgene.usc.es/snipper/analysismultipleprofiles.html</u>,
- 569 Bayes custom or fixed training set data: <u>http://mathgene.usc.es/snipper/analysispopfile\_new.html</u> 570 http://mathgene.usc.es/snipper/popchoosing5groups.html).
- **Supplementary File S4.** Statistical analysis of participant's Indel heterozygote peak height ratio data.
- **Supplementary File S5.** Details and results of NGS analyses of 34-plex and Indel markers made by two
- 575 laboratories.





## Figure 3A Click here to download high resolution image

PPP       AUBLENT       R10       R10 <thr10< th="">       R10       R10       <t< th=""><th></th><th></th><th>1</th><th>-2</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>15</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th><th>21</th><th>Geno</th><th>Gar</th><th></th><th></th></t<></thr10<>			1	-2	4	5	6	7	8	9	15	12	13	14	15	16	17	18	19	20	21	Geno	Gar		
no.         no. <td>PHE</td> <td>112065160</td> <td>010</td> <td>618</td> <td>818</td> <td>-</td> <td>818</td> <td>BLB</td> <td>1818</td> <td>818</td> <td>818</td> <td>_</td> <td>818</td> <td>-</td> <td></td> <td></td> <td>118</td> <td>818</td> <td>818</td> <td>818</td> <td>818</td> <td>100</td> <td>98.57</td> <td>-</td> <td>-</td>	PHE	112065160	010	618	818	-	818	BLB	1818	818	818	_	818	-			118	818	818	818	818	100	98.57	-	-
Phi with with birling         Obi         Obi<	P10	193780181	818	018	110		010	418	· ALCH	818	010		81.0				118	018	010	818	#10	100	96.57	and an	1
number         number<	124	19856786	010	010	010		010	U)=	010	ole	010		010					010	010	010	010	100	96.57	the state of the s	1
Sin         widesset         Bit         Bi	PTR	191573038	110	818	818			eis.	010	018	610		818					818	nin		min	92.66	98.57	to the second	1
No.       N	224	101420654	818	818	018		Dia		818	810	010		019				210	210	HID		010	100	97.14	100	1
mm       matrix       0 </td <td>16a</td> <td>142572367</td> <td>010</td> <td>414</td> <td>010</td> <td></td> <td>819</td> <td>410</td> <td>010</td> <td>010</td> <td>pig</td> <td></td> <td>eie.</td> <td></td> <td></td> <td></td> <td>219</td> <td>014</td> <td>82.0</td> <td>ate</td> <td>010</td> <td>100</td> <td>97.54</td> <td>-</td> <td>-</td>	16a	142572367	010	414	010		819	410	010	010	pig		eie.				219	014	82.0	ate	010	100	97.54	-	-
PND       rt200070       D12       D12 <thd12< th="">       D12       D12       &lt;</thd12<>	104	152614778	010	816	618		1018	111	818	- 010	1010		010				218	810	81.0		910	100	97.14	1000	<u> </u>
Ath       H1000000       Ath	P24	rs730870	010	sin	0.0		with.	niz	110	218	414						218	min	win.		210	100	97.54	_	_
max       m	47.8	141886810	818	818	010		014		ate		e.iu		110				218	#18	815	nin	210	88.57	97.14		-
All       NOME	128	183827700	618	414-1	018		818	018	014		819		414				818	#18	010	811	810	87.14	97.14	and see	-
MA       MA <t< td=""><td>ART</td><td>102040411</td><td>818</td><td>010</td><td>WUR.</td><td></td><td>210</td><td>010</td><td>10.0</td><td>818</td><td>010</td><td></td><td></td><td></td><td></td><td></td><td></td><td>*1*</td><td>010</td><td></td><td>110</td><td>100</td><td>95.71</td><td>1111</td><td>_</td></t<>	ART	102040411	818	010	WUR.		210	010	10.0	818	010							*1*	010		110	100	95.71	1111	_
$ \begin{array}{c} \mathbf{r} \mathbf{r} \mathbf{r} \mathbf{r} \mathbf{r} \mathbf{r} \mathbf{r} r$	44.4	112065682	818	818	RUB.		818		818	B18	818		818				810		810	818	818	100	95.71	-	1
No. 1975/260       0 0	POR	157067580	818	1010	018		818	018	0.000	- 10			818				218		010	212	210	100	95.71	(maintenant)	<i></i>
Normalization       Normalinstation       Normalization       Normalizatio	094	191978600	010		010				10.04	010	110		010				*18		110		810	10.57	05.71		_
miseries       niseries       nistit       nistit       nis	heal	1273026727	610	416	01.0		818	414		010	616		010				- 218				110	97.14	95.71	(and an	-
min       m	-	110091002	818	818	010		810		HIR.	010	010		010				418		NICO		010	100	94.29	-	-
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	100	1912013632	110	818	818			B10	818	210	810						210		HIR	1011	#10	98.57	94.25		_
n172       n172       n17       n18       n18 <th< td=""><td>Pete</td><td>(81301333</td><td>010</td><td>010</td><td>010</td><td></td><td>010</td><td>010</td><td>010</td><td>010</td><td>816</td><td></td><td>010</td><td></td><td></td><td></td><td>410</td><td>010</td><td>010</td><td></td><td>010</td><td>85.57</td><td>94.29</td><td>less line</td><td><u> </u></td></th<>	Pete	(81301333	010	010	010		010	010	010	010	816		010				410	010	010		010	85.57	94.29	less line	<u> </u>
PTT       nabotitie       010	em	19772658	818	818	018		min	010	win.		818		010				418	BIR	013			88.57	94.25	-	
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	PET	1122003794	818	0.14	814			41.0		818	0.18						410	810	1111	ate.	818	98.57	92.66	-	
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	A82	m1335873	010	aia	010		010	010	818	010	810		eis.				410	810	010	nie	110	97.14	92.86	(continue)	1
number of the second of the	201	111406444	610	010	410		with 1	818	110	818	818		110				210	110	010	-	210	WT.14	92.86	Sec. 1	
number of the state of the state	PDA	194540058	110		11.0			1110	11.0		010		410						1111	at a	818	101.71	92.86	ter	
View         View <th< td=""><td>100</td><td>190907008</td><td>010</td><td>111</td><td>41.0</td><td></td><td>010</td><td>110</td><td></td><td></td><td>810</td><td></td><td></td><td></td><td></td><td></td><td>210</td><td>010</td><td>010</td><td>811</td><td>110</td><td>95.71</td><td>02.86</td><td>and the second</td><td>-</td></th<>	100	190907008	010	111	41.0		010	110			810						210	010	010	811	110	95.71	02.86	and the second	-
And       institution       010	P20	10001025	610	011	ain		old	010	114	010	010		012				410	010	019	#1#	010	85.71	92.86	_	-
varbolezoate       vice       vice <td>A215</td> <td>101024110</td> <td>010</td> <td>0.10</td> <td>010</td> <td></td> <td>210</td> <td>Nim</td> <td>41.0</td> <td>010</td> <td>-</td> <td></td> <td>010</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Him</td> <td>010</td> <td>910</td> <td>98.57</td> <td>91.43</td> <td>Contraction of</td> <td>-</td>	A215	101024110	010	0.10	010		210	Nim	41.0	010	-		010						Him	010	910	98.57	91.43	Contraction of	-
HIZ         HIZ <td>100.0</td> <td></td> <td>110</td> <td>818</td> <td>818</td> <td></td> <td>110</td> <td>ain</td> <td>110</td> <td>- 10</td> <td>nia</td> <td></td> <td>010</td> <td></td> <td></td> <td></td> <td>410</td> <td>818</td> <td></td> <td>mia.</td> <td>- 118</td> <td>92.06</td> <td>90.00</td> <td>-</td> <td><i>.</i></td>	100.0		110	818	818		110	ain	110	- 10	nia		010				410	818		mia.	- 118	92.06	90.00	-	<i>.</i>
PP1       int0141763       Q10	112	rx182548	818	0.10	-		410	81.0		811	818		814				816		81.0		810	02.00	90.00	-	-
All       est220000       11       810       610       810	P+1	110141763	810	1010	010		310	410	+10	aia	pig		ais				410	arist.	82.0	aie	010	100	66.57	-	1
NUT         NUT <td>A71</td> <td>CA.720088</td> <td>411</td> <td>110</td> <td>018</td> <td></td> <td>10.0</td> <td></td> <td>110</td> <td>216</td> <td>1010</td> <td></td> <td>814</td> <td></td> <td></td> <td></td> <td>010</td> <td>810</td> <td></td> <td></td> <td>210</td> <td>87.14</td> <td>46.57</td> <td>position.</td> <td><i>.</i></td>	A71	CA.720088	411	110	018		10.0		110	216	1010		814				010	810			210	87.14	46.57	position.	<i>.</i>
P01         sezoeses         010         110         110         01	842	146030340	818	818	Him		1.10	MIN.	BLP	018	411		110				818	1011	NIN.			92.66	47.14	-	-
All       Her       All       Bib       B	PER	112304825	810	110	+10		0.14		21.0		814		410				210		814		4.14	90.00	87.54		1
Ph2         Ph2 <td>A27</td> <td>10017110</td> <td>114</td> <td></td> <td>010</td> <td></td> <td></td> <td>018</td> <td>110</td> <td>810</td> <td>818</td> <td></td> <td>414</td> <td></td> <td></td> <td></td> <td></td> <td>818</td> <td>010</td> <td></td> <td>810</td> <td>98.57</td> <td>40.75</td> <td>-</td> <td>-</td>	A27	10017110	114		010			018	110	810	818		414					818	010		810	98.57	40.75	-	-
completeness 98.82 98.24 100 100 96.82 90.59 18.24 96.47 98.24 99.41 91.16 66.47 82.94 96.3 e concordance 16.29 97.06 99.41 90.69 98.82 94.12 19.41 98.24 96.41 98.24 96.41 98.24 95.29 93.5	inc.	102300031	213	110	018		210		1014	211	010		nin				210	wie:				80.00	81.42		_
completeness         96.82         96.34         100         100         96.82         96.34         96.47         98.24         99.41         91.36         96.3           i concordance         95.29         97.06         99.41         90.42         96.41         96.34         96.34         96.35								20.000										1000		0.000					
concordance 85.29 07.06 99.41 90.99 98.82 94.12 99.41 98.24 44.71 98.82 99.41 88.24 95.29 93.5	tomp	keteness	98.82	98.24	108		100	90.82	90.59		98.24		96.47				98.24	99.41	91.58	86.47	\$2.84	96.3			
	000	cordance	95.23	87.06	99.47		90.99	96.82	94.12		99.41		98,34				44.71	98.82	99.41	16.24	95.29		93.5		(74

#### Figure 3B Click here to download high resolution image

16 Indel	s miscalls	no-ca	alis								19 lab	s										ampletenes
	1	124		1940		1	Ŧ	100		- 11	49	- 12	1.44	- 15	10	17	18	- 10	-96		Centoh	pe control pe con
	-	-										-				11					Gor	Ge
And Control	10000000	010						910	918							610					100	100
maight start	1010003	414	ALC: N				010	414	010				ALC: NO.			010		010			100	100
Marg. and		CONTRACT OF	010		010	010	810	010	010	010	010	010	010	010	010	610	410	010	010	010	100	
		ALC: NO	1010		010		010	919	414		100	1010	1110		010	010	-				100	100
A411.4.24	100407000	110	100				919	918	414		414	100	010	414			-				800	100
		010	NI.			210	010	410		41.0	0.0	nie			inin.	210	min	410			500	100
10-2011	101100000	010	010				010	010	010		818	818	010			010		410			100	100
0.0010		010	ate.		-		810	010	414	410	410		61.0	-		nia	010				100	100
441.000						814	810	414	4.4						214	nia	010	414			100	100
AUT. 794		81.0	818		818		818	418	818		-		618			818				818	100.004	100
1.100	0.2067280	010	-				818	ate		414						212	- 10				100	100
101-10111	0.7508067	818	110					310	1110	818	818	110	81.0		810	010	8116			111	HE 95	100
MED-17	(1412)	110	NI.			214	8.18	0:0	818	10.0	His	Him	Hin.	1010		010	010	810		1010	100	100
10.2114	103054057	010	010	010	810	pia	810	010	610	aie	610	010	010	aia	014	610	010	eie	aia.	810	100	100
	102307845	win .	win		-		0.10	ain				0.0				210	-			-	100	100
10.3444	1000013434	121140	10.00			210	ania	ain	810	1.1	818				210	a in	010	810			100	88.95
its stres	193013003	0.10	NUM			010		410	0.0	414	0.0	010			nin	wite	min	410	414	414	100	100
-	1116364	010	010				810	010		410	010	010	010			010		*14	818		100	100
0.000	1034011475	aim	nin	- 10	-	- 10	210	010	0.10	410	410	110	10.0		010	010	010	610	ala.		100	100
441. 111		014				810	010	010	414		0.0	0.00			810	616	010	616			-	100
m. stork	103048215	414					818	410			010	814	414								300	100
40.907	1423621	010	110			810		ala		81.0	010	010				010	816	414			100	100
et. tattal	01007033	110	110				0.10	410		414	0.10	110	1110			atim					100	100
MET	1416343	110	110				010	410		81.0	110	110	1110			010					100	100
15.2411	0.00111015	010	010		810	810	810	0.0	010	aie	010	nie	010	010	010	810	410	010	aia	210	10.05	100
	(a14120927	010	min		-		<b>BEFORE</b>	aia	818	ale :	010	nie				010		818	als.		100	87.60
-	(1133082	121110	010			10.0	1.1	ain	818		with the	<b>HEYPER</b>	1.010			ale.					100	117.60
A11. 198	110400	0.10	NID		0.14	810	pia	010	010	with a	010	0.10	Him	ain.	aia	210	-	610	414	mill	100	100
MD-15	194181	818	818				min			110	nin.	0.0	010	aia	nin		010	aia.	aia		100	100
0.2141	153030539		010	818	816	818	818	010	818	818		818		818	818	DIG.	010	818	110		100	100
400-018	11102708			818	010	816	818	010	0.0		1110		41.0	818	816	ein	010	016		10116	100	100
um sus	191511225	81.0	818	818		818	SCITCS.	aia	818	818	818	818	818	818		618	819	818	818	818	100	96.96
HTD-THD	1115438	010	218			810	010	010	815		010	010			819	010	010	016		218	100	100
10-2006	rs200enen	818	111	818	818	310	010	010	816	818	818	818	110	818	810	010	pip	916	NIN.	818	100	100
MID-000	1915667	818	U.E.			Dit.	818	010	818	ULE	818	818	UIE	1818		010	010		win:	211	100	100
10-1402	112307988	010	010	010	810	pin		010	610	nin	<b>nin</b>	010	810	816	010	010	010	010	nin	610	100	100
	192307803	wis.	818			018	010	gip	810	818	0.00	010		218	318	010	010	sie			100	100
10-1234	142307900	018		818	ein	210	010	0.0	810	NIE	018	010			218	010	0.0	810	NIE.	818	100	100
ALL-405	/1020600	010	vie.	eie	eie	810	010	016	910	11.0	018	010	UIP.		010	010	010	010	010	#18	100	100
0-1386	193307582	818	818	818	818	818	818	di0	818	ain	818	010	818	010		610	010	ein		818	100	100
0-1729	142307922	818	110	818	818	018	010	0.0	818	818	018	010	818	818	- 018	618	610	810	*12	010	96.54	100
0-0429	1911207925	010		818		810	918	010	818		010	010	818		810	eia	010	010		918	100	100
anti-see	1121584	010	818	818		010	4.4	910	910	818	0.10	MIR				019	0.0	110	414	818	100	100
(D-1022)	rs2307799	010	10.0	815	#10		==10	010	618	318	110	81.0		818	87.0	010	010	616	111	218	100	100
40-2719	1934541393	0(8		:0)0		0(0	0(0	0 0	010		010	010	0(0	810	D)0	0(0	0.0	010		010	100	100
		100.00	1440	0.522	1994	11222	1000		2021	444	1440	19222	112225	1222	- 200	1.222	200	1000	-	1444	00.0	
and the second se	A REAL PROPERTY AND A REAL	498.3	100	1940	100	100	5999	104	1004	100	100	Profession in concerning of the	1000	100	100	100	100	100 T	444.3	100		

#### Figure 4A Click here to download high resolution image











в

AIM-Indel Heterozygosity and peak height ratios in single donor (average, 15 participants) and mixed samples



Average PHR A Maximum PHR

Average

Supplementary Figs S1

Click here to download e-component: Supplementary Figs S1.pdf Supplementary Fig. S1 Examples of genotyping challenges in 34-plex or Indel profiles

S1.1 SNPs P06a-P07 (rs10843344-rs239031) peak pairs run very close together



# **S1.2** SNPs P12/P20/P28 (rs182549/rs1881929/rs3827760) show very low peaks for one or both alleles, particularly allele A in SNP P28.







**S1.3** P01 (rs2304925) shows an artifactual G signal in the negative control very close to the G peak of P27 (rs5030240-rs2304925). This peak is much higher than the T peak when a true allelic extension product but much lower when artifactual.



**S1.4** Indel dye-blobs present in non-allelic positions in a typical negative template control (NTC) profile. One example of a dye-blob is shown in the inset top-left. This is very close to an allele of the short-amplicon Indel MID1470 (rs2307666), influencing its estimated mid-peak position slightly.



S1.5 Example of strong signal pull-up in an Indel profile due to an overloaded sample



**S1.6** SNP P27 (rs5030240) is a tri-allelic marker that showed three alleles in mixed sample F but has low relative peak height for the C allele compared to those of A and G alleles. Peak patterns shown left give another example of the close peak positions shown in S1.3.





**Supplementary Table S1.** Capillary electrophoresis (CE) details for participating laboratories. Grey bars for 34-plex denote five participants not completing SNP genotyping with SNaPshot.

Lah	CE Detector	Bolymor	Dilution factor							
Lap.	CE Delector	Polymer	34-plex*	AIM-Indels						
1	3130xl	POP-4	None	None						
2	3130xl	POP-4	None	1:10 (A-E) 1:5 (F & NTC)						
4	3100	POP-6	None	None						
5	3130xl	POP-4		None						
6	3130xl	POP-7	None	1:10						
7	3130	POP-4	None	None						
8	3130xl	POP-4	None	1:20						
9	3130xl	POP-7	None	None						
11	3130xl	POP-4	None	1:10 (E & F)						
12	3500	POP-4		None						
13	3130	POP-4	1:10	1:10						
14	3500xl	POP-4		1:5						
15	3130xl	POP-4		(Not reported)						
16	3130xl	POP-4		None						
17	3130xl	POP-4	None	1:20						
18	3130	POP-4	None	None						
19	3500xl	POP-4	None	1:20						
20	3130xl	POP-4	None	1:10						
21	3130	POP-4	None	None						

\* 5/19 laboratories marked did not run the 34-plex SNP assay, but this does not show a relationship to the choice of polymer or CE detector used.
S       < add prefered % pipetting top-u	5       < add prefered % pipetting top-up here         16       < add sample multiple here         11.6 μl       11.59       11.60         11.6 μl       1.5       1.50         11.6 μl       27.38       27.40         27.4 μl       27.38       27.40         7.2 μl       7.22       7.20         16.8 μl       16.80       16.80         16.8 μl       16.80       16.80         17. μl       0.86       0.90       14.45         14.4 μl       5.40       5.40         0.5       < add DNA concentration of the state of the st			1	< µl decimal	s (1 or 2)		
x1 sample       16       < add sample multiple here	16       < add sample multiple here         11.6 μl       11.59       11.60         11.6 μl       1.5       1.50         11.6 μl       27.38       27.40         27.4 μl       27.38       27.40         7.2 μl       7.22       7.20         16.8 μl       16.80       16.80         16.8 μl       16.80       16.80         17.4 μl       0.86       0.90       14.45         14.4 μl       5.40       5.40         •       0.5       < add DNA concentra         •       Optimum DNA input is 0.75 ng       1.5 μl       □       6.9 μl	PCR mix:		5	< add prefer	ed % pipettin	ig top-up he	ere
Buffer 10x $0.69 \ \mu l$ $11.6 \ \mu l$ $11.59$ $11.60$ BSA (1.6 \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	11.6 µl       11.59       11.60         11.6 µl       1.5       1.50         27.4 µl       27.38       27.40         7.2 µl       7.22       7.20         7.2 µl       7.22       7.20         16.8 µl       16.80       16.80         16.8 µl       16.80       14.45         1.7 µl       0.86       0.90       14.45         14.4 µl       5.40       5.40         Optimum DNA input is 0.75 ng       1.5 µl       G.9 µl         Total Volume       1.5 µl       Total Volume		x1 sample	16	< add sampl	e multiple he	re	
BSA (1.6 µg/µl) $0.69 µl$ $11.6 µl$ MgCl2 (25 mM) $1.63 µl$ $27.4 µl$ $27.38$ $27.40$ dNTPs (10 mM) $0.43 µl$ $7.2 µl$ $7.22$ $7.20$ PCR primer mix $1 µl$ $16.8 µl$ $16.8 µl$ $1.68$ $1.50$ AmpliTaq Gold $0.1 µl$ $1.7 µl$ $1.68$ $1.70$	11.6 µl       27.38       27.40         27.4 µl       27.38       27.40         7.2 µl       7.22       7.20         1.50       1.50         16.8 µl       16.80         16.8 µl       1.50         1.7 µl       0.86       0.90         14.4 µl       5.40       5.40         •       0.5       < add DNA concentration of the state of the	Buffer 10x	0.69 µl	11.6 µl	11.59 1.5	11.60 1.50		
MgCl2 (25 mM) $1.63 \mu$ l $27.4 \mu$ l $27.38$ $27.40$ dNTPs (10 mM) $0.43 \mu$ l $7.2 \mu$ l $7.22$ $7.20$ PCR primer mix $1 \mu$ l $16.8 \mu$ l $16.80$ $1.50$ AmpliTaq Gold $0.1 \mu$ l $1.7 \mu$ l $1.68$ $1.70$	27.4 μl       27.38       27.40         7.2 μl       7.22       7.20         1.50       1.50       1.50         16.8 μl       16.80       16.80         1.7 μl       1.68       1.70         1.7 μl       0.86       0.90       14.45         14.4 μl       5.40       5.40         •       0.5       < add DNA concentration of the state of the st	BSA (1.6 μg/μl)	0.69 µl	11.6 µl	1			
dNTPs (10 mM)       0.43 μl       7.2 μl       7.20         PCR primer mix       1 μl       16.8 μl       16.80         AmpliTaq Gold       0.1 μl       1.7 μl       1.68       1.70	7.2 μl       7.22       7.20         1.50       1.50         16.8 μl       16.80         1.7 μl       1.68         1.7 μl       0.86         0.86       0.90         14.4 μl       5.40         5.40       5.40         0.5       < add DNA concentra	MgCl2 (25 mM)	1.63 µl	27.4 µl	27.38	27.40		
PCR primer mix         1 μl         16.8 μl         16.80         16.80           AmpliTaq Gold         0.1 μl         1.7 μl         1.68         1.70	16.8 µl       16.80       16.80         16.8 µl       1.50         1.7 µl       1.68       1.70         0.86       0.90       14.45         14.4 µl       5.40       5.40         0.5       < add DNA concentra	dNTPs (10 mM)	0.43 µl	7.2 µl	7.22	7.20 1.50		
AmpliTaq Gold         0.1 μl         1.7 μl         1.68         1.70           0.86         0.90         1	1.7 μl       1.68       1.70         0.86       0.90       14.45         14.4 μl       5.40       5.40         •       0.5       < add DNA concentra	PCR primer mix	1 µl	16.8 µl	16.80	16.80 1.50		
	14.4 μl     0.86     0.90     14.45       14.4 μl     5.40     5.40       0.5     < add DNA concentra       +     Optimum DNA input is 0.75 ng     1.5 μl	AmpliTag Gold	0.1.ul	1 7 ul	1.68	1.70		
	14.4 μl 5.40 5.40 0.5 < add DNA concentra 0.5 < add DNA concentra 4 Optimum DNA input is 0.75 ng 1.5 μl U Total Volume	Ampiring dold	0.1 µ1	pi	0.86		14.45	
H <sub>2</sub> O 0.9 μl 14.4 μl 5.40 5.40	0.5 < add DNA concentra + Optimum DNA input is 0.75 ng 1.5 μl □ Total Volume	H <sub>2</sub> O	0.9 µl	14.4 µl	5.40	5.40		
□ 0.5 < add D	+ DNA input 1.5 µl □ Total is 0.75 ng Volume				J	0.5	< add DNA	concentra
5.4 µl Mix Optimum Total + DNA input 1.5 µl □ Volume is 0.75 ng			5.4 µl Mix Total Volume	+	Optimum DNA input is 0.75 ng	1.5 µl		6.9 <i>µ</i> l Total Volume
		xo-SAP purific	ation:		low-cost			
Exo-SAP purification: <i>low-cost</i>	low-cost		x1 sample		x1 sample (r	ion evidentia	I DNA)	
Exo-SAP purification: <i>Iow-cost</i> x1 sample x1 sample (non evidential DNA)	<i>low-cost</i> x1 sample (non evidential DNA)	ExoSAPit	1.3 µl		0.65 µl			
Exo-SAP purification:     Iow-cost       x1 sample     x1 sample (non evidential DNA)       ExoSAPit     1.3 μl	<b>low-cost</b> x1 sample (non evidential DNA) 0.65 μl	PCR product	2.5 µl		1.25 µl			

x1 sample 16 x

Version	CM v 3	Λ	
Chomic	tru Kit	AIM ind	lolplov
DinCot	li y Kit	AIM ind	lolplov
Diliset I Danal M	amo	AIM ind	lolplov
Markor	Mama	AIM-IIIC	leipiex
1	Name 60.17	14/0	0.76
1	00.17	0.5	0.70
Z Maulaan	05.1 Nama	0.5	0.5
Marker	Name	///	0 5
1	69.33 72.22	0.5	0.5
	/2.33	0.5	0.5
Marker	Name	196	0 5
1	/8.3/	0.5	0.5
2	81.37	0.5	0.5
Marker	Name	881	~ -
1	86.37	0.5	0.5
2	90.37	0.5	0.5
Marker	Name	3122	
1	96.18	0.5	0.5
2	100.18	0.5	0.5
Marker	Name	548	
1	105.8	0.5	0.5
2	107.8	0.5	0.5
Marker	Name	659	
1	114.3	0.5	0.5
2	116.3	0.5	0.5
Marker	Name	2011	
1	125.29	0.5	0.5
2	130.29	0.5	0.5
Marker	Name	2929	
1	147.91	0.5	0.5
2	149.91	0.5	0.5
– Marker	Name	593	010
1	154 56	05	05
2	156.56	0.5	0.5
2 Markor	Name	798	0.5
1	166 28	05	05
1 2	172.28	0.5	0.5
4 Markor	172.20 Namo	1102	0.5
1	170 <i>1</i> 7	05	05
1	101 17	0.5	0.5
4 Manleon	101.47 Nomo	0.5	0.5
	100 7F	10/1	0 5
1	109.75	0.5	0.5
	191./5	0.5	0.66
Marker	Name	17	o =
1	197.81	0.5	0.5
2	201.81	0.5	0.5
Marker	Name	2538	
1	208.38	0.5	0.75
2	212.38	0.5	0.75
Marker	Name	1644	
1	220.2	0.5	0.5
2	222.2	0.5	0.5

Marker	Name	3854		
1	55.4	0.5	0.5	
2	59.21	0.59	0.5	
Marker	Name	2275		
1	69.3	0.5	0.5	
2	76.09	0.5	0.5	
2 Markor	Name	3072	0.5	
1	104.07	0 E	0 5	
1	104.07	0.5	0.5	
۲ ۸	111.U/	0.5	0.5	
Marker	Name	//2	0 5	
1	116.6	0.5	0.5	
2	119.6	0.5	0.5	
Marker	Name	2313		
1	126.01	0.5	0.5	
2	135.01	0.5	0.5	
Marker	Name	397		
1	161.99	0.5	0.5	
2	165.99	0.5	0.5	
Marker	Name	1636		
1	172.86	0.5	0.5	
2	174.86	0.5	0.5	
_ Marker	Name	51		
1	182 11	05	05	
2	187 11	0.5	0.5	
2 Markor	Namo	2421	0.5	
1	207 47	05	0.75	
1	207.47	0.5	0.75	
۲ ۸	211.4/	0.5	0.75	
Marker	Name	2264	0 5	
1	222.8	0.5	0.5	
3	225.8	0.4	0.4	mutant
2	226.8	0.4	0.66	
Marker	Name	2256		
1	57.38	0.5	0.5	
2	60.38	0.5	0.5	
Marker	Name	128		
1	66.67	0.5	0.5	
2	69.67	0.5	0.5	
Marker	Name	15		
1	76.1	0.5	0.5	
2	87.0	0.5	0.64	
– Marker	Name	2241	0101	
1	106.0	0.64	05	
2	114.01	0.51	0.5	
2 Markor	Namo	0.5 110	0.5	
1	110 0 <i>1</i>	419 0 E	065	
1	110.04	0.5	0.05	
	125.84	0.5	0.5	
Marker	Name	943	0 5	
1	154.51	0.56	0.5	
2	158.51	0.5	0.5	
Marker	Name	159		
1	166.85	0.61	0.5	
2	171.9	0.5	0.5	

Marker	Name	2005		
1	182.28	0.5	0.55	
2	188.28	0.64	0.5	
Marker	Name	250		
1	198.85	0.5	0.5	
2	200.85	0.5	0.5	
Marker	Name	1802		
1	211.13	0.5	0.5	
2	214.13	0.5	0.5	
Marker	Name	1607		
1	219.02	0.5	0.5	
2	220.95	0.62	0.5	
Marker	Name	406		
1	65.39	0.5	0.5	
2	67.39	0.5	0.5	
Marker	Name	1386		
1	70.71	0.5	0.71	
2	91.57	0.5	0.5	
Marker	Name	1726		
1	103.82	0.5	0.5	
2	116.5	0.5	0.5	
Marker	Name	3626		
1	140.36	0.5	0.5	
2	156.25	0.5	0.5	
Marker	Name	360		
1	169.42	0.4	0.4	
3	170.42	0.4	0.4	mutant
2	171.42	0.4	0.4	
Marker	Name	1603		
1	212.05	0.5	0.56	
2	215.97	0.5	0.5	
Marker	Name	2719		
1	225.38	0.5	0.5	
2	229.38	0.5	0.5	
Marker	Name	1734		
1	57.11	0.5	0.5	
2	60.79	0.5	0.5	
Marker	Name	94		
1	89.14	0.5	0.5	
2	92.14	0.5	0.56	

Version	n GM v 3.	0					
Kit type	e:	MICROSATELLITE					
Chemis	try Kit	AIM-indelplex none					
Panel	AIM-ind	delplex none					
1470	blue	59.498930343000005 65.8182525690	00001	-	9	0.0	none
777	blue	68.55843072 73.031835852 -	9	0.0	none		
196	blue	77.588872041 82.088429436 -	9	0.0	none		
881	blue	85.617309136 91.15076970400001	-	9	0.0	none	
3122	blue	95.320690594 100.900499185-	9	0.0	none		
548	blue	105.057912973108.532867596-	9	0.0	none		
659	blue	113.556264981117.080076701-	9	0.0	none		
2011	blue	124.498164706131.041303657-	9	0.0	none		
2929	blue	146.955519673150.824584052-	9	0.0	none		
593	blue	153.83022589700002 157.143132983	399998	-	9	0.0	none
798	blue	165.32897286899998 173.501119818	3-	9	0.0	none	
1193	blue	178.658800433182.45524446299999	-	9	0.0	none	
1871	blue	188.823452509192.72466145 -	9	0.0	none		
17	blue	197.084416798202.680731973-	9	0.0	none		
2538	blue	206.773228342213.827474758-	9	0.0	none		
1644	blue	219.483987195223.06873612500002	-	9	0.0	none	
3854	Green	54.61523001 59.995735289 -	9	0.0	none		
2275	Green	68.507204408 76.863106493 -	9	0.0	none		
3072	Green	103.268780394111.960791133-	9	0.0	none		
772	Green	115.766532729120.46989852499999	-	9	0.0	none	
2313	Green	125.024081548136.274250832-	9	0.0	none		
397	Green	161.189615282166.818202663-	9	0.0	none		
1636	Green	172.08629923200002 175.600073663	3-	9	0.0	none	
51	Green	181.211398852188.025739155-	9	0.0	none		
2431	Green	206.293870962214.07491917299998	-	9	0.0	none	
2264	Green	221.926816896228.05683674 -	9	0.0	none		
2256	Yellow	56.680080642 61.154696174 -	9	0.0	none		
128	Yellow	65.9389478170000170.400997293	-	9	0.0	none	
15	Yellow	75.5 87.975809403 - 9	0.0	none			
2241	Yellow	104.973806014114.835216709-	9	0.0	none		
419	Yellow	118.07762833899999 126.850206732	7-	9	0.0	none	
943	Yellow	153.696434559159.341659618-	9	0.0	none		
159	Yellow	166.02409311300002 172.70691387	-	9	0.0	none	
2005	Yellow	181.501371715189.06208221400001	-	9	0.0	none	
250	Yellow	198.065914853201.572590476-	9	0.0	none		
1802	Yellow	210.217316797215.228136625-	9	0.0	none		
1607	Yellow	218.273123059221.87207210900002	-	9	0.0	none	
406	Red	64.678876207 68.176704628 -	9	0.0	none		
1386	Red	69.56966599 92.812677351 -	9	0.0	none		
1726	Red	102.977336125117.29484620400001	-	9	0.0	none	
3626	Red	139.43282453700002 156.750843119	)-	9	0.0	none	
360	Red	168.792202855172.14353350399998	-	9	0.0	none	
1603	Red	211.384383599216.862882022-	9	0.0	none		
2719	Red	224.388680709230.34325627 -	9	0.0	none		
1734	Red	56.285110841 61.814691114000006	-	9	0.0	none	
94	Green	88.363699841 92.99304501099999	-	9	0.0	none	

Version Gl	M v 3.0	)	
Chemistry	v Kit	34plex_	POP4
BinSet Na	me	34plex	POP4
Panel Nan	ne	34plex	POP4
Marker Na	ame	P01 T	
T 85	5.75	0.4	0.4
Marker Na	ame	A07 G	
G 26	5.92	0.53	0.48
Marker Na	ame	A07 A	
A 29	9.26	0.46	0.42
Marker Na	ame	P03 C	-
C 27	7.22	0.6	0.47
Marker Na	ame	P03 T	-
T 28	3.98	0.62	0.46
Marker Na	ame	P04 C	
C 88	3.76	0.4	0.4
Marker Na	ame	P04 T	011
T 80	9 56	04	04
Marker Na	ame	A29 G	0.1
G 28	R 44	0.46	0 45
Marker N:	ome	Δ29 Δ	0.45
$\Delta = 20$	111C	ΠΔ ) Π Π ΛΩ	0.47
Markor N	).T)	DOS C	0.77
$C$ $2^{\prime}$	anne 1 70	052	0.45
Markor N	1.7 J	0.55 DOS T	0.45
		0 4 2	0.41
I J. Markor M	).J Jmo	0.42 A21 C	0.41
	anne 1 1 0	AZIG	0.41
U 34 Morkor M	t.10	0.4	0.41
			0.47
A 30 Montron M	0.40	0.51 D06a C	0.47
		PU0a L	0.47
L 3/	/.39	0.48 D0( - T	0.47
Marker Na	ame	P06a I	0.40
I 30	3.65	0.45 D00 C	0.42
Marker Na	ame	PU8 G	0.41
G 40	).37	0.45	0.41
Marker Na	ame	PU8 A	0.40
A 4.	1.01	0.41	0.42
Marker Na	ame		0 5
	9.23	0.5	0.5
Marker Na	ame	P07 T	0.45
T 39	9.67	0.43	0.47
Marker Na	ame	A40 G	
G 43	3.22	0.42	0.42
Marker Na	ame	A40 A	
A 44	1.34	0.51	0.41
Marker Na	ame	P09a C	
C 44	1.64	0.5	0.5
Marker Na	ame	P09a T	_
T 45	5.67	0.46	0.44
Marker Na	ame	P10 G	
G 46	5.52	0.42	0.42

Marker	Name	P10 C	
С	46.94	0.46	0.46
Marker	Name	P11 A	
А	49.0	0.5	0.47
Marker	Name	P11 T	
Т	49.63	0.4	0.4
Marker	Name	P12 C	0.1
C	49 34	0.43	046
Markor	Namo	0.45 D12 T	0.10
		0.4.4	0.41
I Maultau	50.97 Nama	0.44 D12 C	0.41
Marker	Name	P13G	0 5
G	49.85	0.5	0.5
Marker	Name	P02 A	
A	88.17	0.5	0.5
Marker	Name	P02 C	
С	87.87	0.46	0.45
Marker	Name	P01 G	
G	84.55	0.4	0.4
Marker	Name	P13 A	
А	51.5	0.46	0.45
Marker	Name	P14 C	
С	53.37	0.43	0.46
Marker	Name	р14 Т	0.10
Т	54 34	04	0 4 1
1 Markor	Namo	D15 C	0.71
		0.47	0.45
U Maulaan	55.29 Nome		0.45
Marker	Name	P15 A	0.40
A	56.1	0.4/	0.43
Marker	Name	P16a G	
G	56.79	0.42	0.45
Marker	Name	P16a A	
А	57.64	0.4	0.4
Marker	Name	P17 C	
С	58.84	0.44	0.42
Marker	Name	P17 T	
Т	59.56	0.44	0.43
Marker	Name	P18 G	
G	61.26	0.41	0.42
Marker	Name	P18 A	0.12
Δ	62 14	0.47	0 4 1
Markor	Name	D10C	0.71
C	62 57	0.46	0 / 1
U Manlton	02.37 Nomo	0.40 D10 T	0.41
магкег	Name	P191	0.4
	63.82	0.4	0.4
Marker	Name	P20 G	
G	63.87	0.45	0.42
Marker	Name	P20 T	
Т		0.40	0 1 5
1	65.82	0.43	0.45
Marker	65.82 Name	0.43 P21 A	0.45
Marker A	65.82 Name 66.17	0.43 P21 A 0.47	0.45
Marker A Marker	65.82 Name 66.17 Name	0.43 P21 A 0.47 P21 C	0.45

Marker	Name	P22a C	
С	69.27	0.4	0.4
Marker	Name	P22a T	
Т	69.98	0.44	0.43
Marker	Name	P23 G	
G	70.23	0.45	0.45
Marker	Name	P23 A	
А	70.24	0.5	0.5
Marker	Name	P24 A	
А	73.53	0.47	0.4
Marker	Name	P24 C	
С	73.08	0.4	0.4
Marker	Name	P24 T	
Т	73.83	0.46	0.4
Marker	Name	A52 A	
А	78.77	0.43	0.48
Marker	Name	A52 T	
Т	79.01	0.4	0.49
Marker	Name	P25a G	
G	76.83	0.51	0.52
Marker	Name	P25a C	
С	77.18	0.48	0.46
Marker	Name	P26 C	
С	80.5	0.4	0.47
Marker	Name	P26 T	
Т	81.07	0.44	0.43
Marker	Name	A13 G	
G	79.24	0.43	0.4
Marker	Name	A13 A	
А	80.07	0.4	0.47
Marker	Name	P27 G	
G	83.88	0.45	0.42
Marker	Name	P27 A	
А	85.12	0.42	0.42
Marker	Name	P27 C	
С	84.71	0.42	0.43
Marker	Name	P28 G	
G	89.59	0.4	0.4
Marker	Name	P28 A	
А	90.5	0.43	0.45

Version	GM v 3.	0						
Kit type	2:	MICROS	SATELLI	TE				
Chemis	try Kit	34plex_	POP4	none				
Panel	34plex_	POP4	none					
P01 T	Red	85.0	86.5	-	2	0.0	rs2304925	-
A07 G	Blue	25.95	27.5	-	2	0.0	rs917118	-
A07 A	Green	28.5	29.88	-	2	0.0	rs917118	-
P03 C	Yellow	26.2	28.0	-	2	0.0	rs1321333	-
P03 T	Red	28.2	29.6	-	2	0.0	rs1321333	-
P04 C	Yellow	88.1	89.5	-	2	0.0	rs2814778	-
P04 T	Red	88.7	90.5	-	2	0.0	rs2814778	-
A29 G	Blue	27.8	29.17	-	2	0.0	rs1024116	-
A29 A	Green	29.8	31.1	-	2	0.0	rs1024116	-
P05 C	Yellow	31.0	32.4	-	2	0.0	rs7897550	-
P05 T	Red	32.7	33.9	-	2	0.0	rs7897550	-
A21 G	Blue	33.6	34.71	-	2	0.0	rs722098	-
A21 A	Green	35.75	37.1	-	2	0.0	rs722098	-
P06a C	Yellow	36.8	38.0	-	2	0.0	rs10843344	-
P06a T	Red	38.0	39.25	-	2	0.0	rs10843344	-
P08 G	Blue	39.8	41.01	-	2	0.0	rs12913832	-
P08 A	Green	40.4	41.65	-	2	0.0	rs12913832	-
P07 C	Yellow	38.5	40.0	-	2	0.0	rs239031	-
P07 T	Red	39.0	40.4	-	2	0.0	rs239031	-
A40 G	Blue	42.6	43.85	-	2	0.0	rs2040411	-
A40 A	Green	43.6	44.93	-	2	0.0	rs2040411	-
P09a C	Yellow	43.9	45.35	-	2	0.0	rs1978806	-
P09a T	Red	45.0	46.3	-	2	0.0	rs1978806	-
P10 G	Blue	45.9	47.2	_	2	0.0	rs773658	-
P10 C	Yellow	46.25	47.6	_	2	0.0	rs773658	-
P11 A	Green	48.3	49.7	_	2	0.0	rs10141763	-
P11 T	Red	48.8	50.5	_	2	0.0	rs10141763	-
P12 C	Yellow	48.75	50.0	_	2	0.0	rs182549	-
P12 T	Red	50.3	51.6	-	2	0.0	rs182549	-
P13 G	Blue	49.1	50.7	-	2	0.0	rs1573020	-
P02 A	Green	87.5	88.85	-	2	0.0	rs5997008	-
P02 C	Yellow	87.19	88.5	-	2	0.0	rs5997008	-
P01 G	Blue	84.1	85.3	-	2	0.0	rs2304925	-
P13 A	Green	50.85	52.2	-	2	0.0	rs1573020	-
P14 C	Yellow	52.75	54.0	-	2	0.0	rs896788	-
P14 T	Red	53.75	54.95	-	2	0.0	rs896788	-
P15 G	Rlue	54.65	559	-	2	0.0	rs2065160	-
P15 A	Green	554	56 75	-	2	0.0	rs2065160	-
P16a G	Blue	56.2	57 45	-	2	0.0	rs2572307	-
P16a A	Green	56.7	58.2	_	2	0.0	rs2572307	_
P17 C	Yellow	58.2	59.4	_	2	0.0	rs2303798	_
Р17 Т	Red	59.0	60.2	_	2	0.0	rs2303798	-
P18 G	Rlue	60.65	61.9	_	2	0.0	rs2065982	-
P18 A	Green	61 5	62 75	_	2	0.0	rs2065982	_
P197	Yellow	62.0	63.2	-	2	0.0	rs3785181	-
Р19 Т	Red	63.22	64.4	-	2	0.0	rs3785181	-
P20 C	Rhie	63.22	64 55	-	2	0.0	rs881979	-
P20 T	Red	65.22	66 5	-	2	0.0	rs881979	_
1 20 1	neu	05.4	00.5		4	0.0	13001727	-

P21 A	Green	65.5	66.8	-	2	0.0	rs1498444	-
P21 C	Yellow	65.45	66.6	-	2	0.0	rs1498444	-
P22a C	Yellow	68.65	69.85	-	2	0.0	rs1426654	-
P22a T	Red	69.35	70.6	-	2	0.0	rs1426654	-
P23 G	Blue	69.55	70.85	-	2	0.0	rs2026721	-
P23 A	Green	69.55	70.95	-	2	0.0	rs2026721	-
P24 A	Green	72.85	74.15	-	2	0.0	rs4540055	-
P24 C	Yellow	72.5	73.65	-	2	0.0	rs4540055	-
P24 T	Red	73.17	74.4	-	2	0.0	rs4540055	-
A52 A	Green	78.15	79.4	-	2	0.0	rs1335873	-
A52 T	Red	78.4	79.7	-	2	0.0	rs1335873	-
P25a G	Blue	76.2	77.5	-	2	0.0	rs16891982	-
P25a C	Yellow	76.5	77.8	-	2	0.0	rs16891982	-
P26 C	Yellow	79.9	81.15	-	2	0.0	rs730570	-
P26 T	Red	80.4	81.7	-	2	0.0	rs730570	-
A13 G	Blue	78.6	79.8	-	2	0.0	rs1886510	-
A13 A	Green	79.5	80.75	-	2	0.0	rs1886510	-
P27 G	Blue	83.25	84.5	-	2	0.0	rs5030240	-
P27 A	Green	84.5	85.75	-	2	0.0	rs5030240	-
P27 C	Yellow	84.1	85.35	-	2	0.0	rs5030240	-
P28 G	Blue	89.0	90.5	-	2	0.0	rs3827760	-
P28 A	Green	89.9	91.15	-	2	0.0	rs3827760	-

Version GM v 3.0 Chemistry Kit 34-PLEX BinSet Name 34-PLEX Panel Name 34-Plex Electrophoretic Shift Marker Name 01rs1321333 ASR 34.77 37.61 С 34.77 35.9 Yellow Т 36.39 37.6100000000001 Red Marker Name 02rs917118 ASR 32.6 35.67 33.809999999999995 Blue G 32.6 А 34.87000000000005 35.67 Green Marker Name 03rs1024116 33.35 36.39 ASR G 33.35 34.54 Blue А 35.17 36.39 Green Marker Name 04rs7897550 ASR 37.53 40.05 С 37.53 38.36 Yellow Т 38.96 40.05000000000004 Red 05rs722098 Marker Name ASR 37.7 40.78 37.7 38.5 G Blue 39.98000000000004 40.78 Green А Marker Name 06rs10843344 ASR 41.68 44.0 С 41.68 42.48 Yellow Т 43.2 44.0 Red Marker Name 07rs239031 ASR 42.35 45.17 С 42.35 43.15 Yellow Т 44.36 45.17 Red Marker Name 08rs12913832 44.16 45.64 ASR 44.1600000000004 44.96 Blue G А 44.84 45.64 Green Marker Name 09rs2040411 ASR 46.48 48.49 46.48000000000004 47.78 Blue G A 47.6900000000005 48.49 Green Marker Name 10rs1978806 ASR 48.63 50.43 С 48.62999999999995 49.75 Yellow Т 49.63 50.43 Red Marker Name 11rs773658 ASR 49.49 51.02 G 49.49 50.29 Blue 50.22 51.01999999999999 Yellow С Marker Name 12rs10141763 ASR 52.01 54.21 А 52.0100000000005 53.08 Green Т 52.66999999999995 54.21 Red

Marker Name 13rs182549 ASR 51.8 53.81 52.92 Yellow 51.8 С Т 53.01 53.8099999999999 Red Marker Name 14rs1573020 ASR 54.94 53.09 G 53.09 54.59 Blue 53.69 54.94 Green А Marker Name 15rs896788 ASR 55.66 57.75 С 55.66 57.03 Yellow Т 56.9000000000006 57.75 Red Marker Name 16rs2065160 56.85 58.04 ASR G 56.84999999999994 57.75 Blue А 57.24 58.04 Green 17rs2572307 Marker Name ASR 58.03 59.66 G 58.03 58.83 Blue А 58.86 59.66 Green Marker Name 18rs2303798 ASR 59.78 61.12 С 59.78 60.58 Yellow Т 60.32 61.12 Red Marker Name 19rs2065982 ASR 62.31 63.86 G 62.31 63.11 Blue А 63.06 63.86 Green 20rs3785181 Marker Name ASR 66.12 63.89 С 63.89 64.79 Yellow Т 65.32 66.12 Red Marker Name 21rs881929 64.29 67.39 ASR 64.29 65.26 Blue G Т 66.589999999999999 67.39 Red Marker Name 22rs1498444 ASR 67.19 68.17 67.19 67.99000000000001 Α Green С 67.36999999999999 68.17 Yellow Marker Name 23rs1426654 ASR 70.48 72.13 С 70.479999999999999 71.28 Yellow Т 71.33 72.13000000000001 Red Marker Name 24rs2026721 71.28 72.65 ASR G 71.28 72.08000000000001 Blue А 71.85 72.65 Green Marker Name 25rs4540055 ASR 74.18 76.08 А 74.5 75.3000000000001 Green С 74.1799999999999 74.98 Yellow Т 75.28 76.0800000000001 Red Marker Name 26rs16891982 79.03 ASR 77.66 78.4600000000001 G 77.66 Blue С 78.229999999999999 79.03 Yellow Marker Name 27rs1335873 ASR 79.89 81.1 79.89 80.6900000000001 Α Green Т 80.3 81.10000000000001 Red Marker Name 28rs1886510 ASR 80.52 82.01 G 80.5200000000001 81.31 Blue 81.179999999999999 82.01 А Green Marker Name 29rs730570 ASR 82.12 83.68 С 82.12 83.06 Yellow Т 82.67 83.679999999999999 Red Marker Name 30rs5030240 ASR 85.05 86.71 G 85.05 85.67 Blue 86.05 A 86.71 Green 85.95 С 86.58 Yellow 31rs2304925 Marker Name ASR 87.25 89.32 G 87.25 87.789999999999999 Blue Т 88.55 89.3200000000001 Red Marker Name 32rs5997008 ASR 88.14 88.85 88.14 88.83 Green А С 88.23 88.85000000000001 Yellow Marker Name 33rs3827760 ASR 90.46 91.43 G 90.46 91.02 Blue 90.73 91.429999999999999 А Green Marker Name 34rs2814778 ASR 91.78 93.27 С 91.78 92.7700000000001 Yellow Т 92.39 93.27 Red 34-PLEX Panel Name Marker Name 01rs1321333 ASR 32.62 35.85 32.6200000000005 33.75 С Yellow Т 34.629999999999995 35.85 Red Marker Name 02rs917118 ASR 31.54 34.26 G 31.5400000000003 32.75 Blue А 33.46 34.26 Green Marker Name 03rs1024116 ASR 34.99 31.87 G 31.8700000000005 33.06 Blue 33.77 34.99 Green А Marker Name 04rs7897550

ASR 36.44 39.3 36.4400000000005 37.27 Yellow С Т 38.21 39.30000000000004 Red Marker Name 05rs722098 ASR 36.94 39.87 G 36.94000000000005 37.74 Blue 39.07 39.87 Green А Marker Name 06rs10843344 40.55 43.09 ASR 40.55000000000004 41.35 Yellow С Т 42.29 43.08999999999996 Red Marker Name 07rs239031 41.21 44.01 ASR С 41.21 42.01 Yellow Т 43.19999999999996 44.01 Red Marker Name 08rs12913832 43.28 ASR 44.56 G 43.28 44.08 Blue А 43.76 44.55999999999999 Green Marker Name 09rs2040411 45.69 47.51 ASR 45.6900000000005 46.99 Blue G А 46.71 47.51 Green Marker Name 10rs1978806 47.22 49.46 ASR С 47.22 48.34 Yellow Т 48.66000000000004 49.46 Red Marker Name 11rs773658 48.36 49.88 ASR G 48.36 49.16 Blue С 49.08 49.87999999999995 Yellow Marker Name 12rs10141763 ASR 50.6 53.45 51.66999999999995 Green А 50.6 Т 51.91 53.45 Red Marker Name 13rs182549 ASR 51.13 53.81 С 51.12999999999995 52.25 Yellow Т 53.01 53.80999999999999 Red Marker Name 14rs1573020 ASR 52.36 54.19 G 52.36 53.86 Blue 52.94 54.19 Green А Marker Name 15rs896788 ASR 55.02 57.07 С 55.01999999999996 56.39 Yellow Т 56.2200000000006 57.07 Red Marker Name 16rs2065160 ASR 56.85 58.04 G 56.849999999999994 57.75 Blue 57.24 58.04 Green А Marker Name 17rs2572307

ASR 58.03 59.66 G 58.03 58.83 Blue 59.66 Green А 58.86 Marker Name 18rs2303798 ASR 59.78 61.12 С 59.78 60.58 Yellow Т 60.32 61.12 Red 19rs2065982 Marker Name ASR 62.31 63.86 G 62.31 63.11 Blue А 63.06 63.86 Green Marker Name 20rs3785181 ASR 63.89 66.12 С 64.79 Yellow 63.89 Т 65.32 66.12 Red Marker Name 21rs881929 67.39 ASR 64.29 G 64.29 65.26 Blue Т 66.589999999999999 67.39 Red Marker Name 22rs1498444 67.19 ASR 68.17 67.19 67.9900000000000 Α Green 67.36999999999999 С 68.17 Yellow Marker Name 23rs1426654 70.48 72.13 ASR С 70.479999999999999 71.28 Yellow Т 71.33 72.13000000000001 Red Marker Name 24rs2026721 71.28 72.65 ASR G 71.28 72.0800000000001 Blue 71.85 72.65 Green А Marker Name 25rs4540055 ASR 74.18 76.08 74.5 75.3000000000001 А Green С 74.1799999999999 74.98 Yellow 75.28 76.0800000000001 Т Red Marker Name 26rs16891982 ASR 77.66 79.03 G 77.66 78.46000000000001 Blue С 78.229999999999999 79.03 Yellow Marker Name 27rs1335873 ASR 79.89 81.1 А 79.89 80.6900000000001 Green Т 80.3 81.10000000000001 Red Marker Name 28rs1886510 80.52 82.01 ASR G 80.5200000000001 81.31 Blue 81.179999999999999 82.01 Green А Marker Name 29rs730570 ASR 82.12 83.68 С 82.12 83.06 Yellow Т 83.679999999999999 Red 82.67

Marker Name 30rs5030240 85.05 ASR 86.71 85.05 G 85.67 Blue 86.05 86.71 Green А С 85.95 86.58 Yellow 31rs2304925 Marker Name 87.25 89.32 ASR 87.25 87.789999999999999 Blue G Т 88.55 89.3200000000001 Red Marker Name 32rs5997008 ASR 88.14 88.85 А 88.14 88.83 Green С 88.23 88.8500000000001 Yellow 33rs3827760 Marker Name ASR 90.46 91.43 G 90.46 91.02 Blue 90.73 91.429999999999999 Green А 34rs2814778 Marker Name ASR 91.78 93.27 С 91.78 92.77000000000001 Yellow Т 92.39 93.27 Red 34-PLEX Extra Panel Name Marker Name 02rs917118 ASR 69.9 71.5 G 69.899999999999999 70.55 Blue А 70.8000000000001 71.5 Green Marker Name 03rs1024116 32.4 34.99 ASR G 32.4000000000006 33.24 Blue А 33.77 34.99 Green Marker Name 05rs722098 ASR 36.67 39.4 G 36.67 37.47 Blue 38.6 39.4 Green А 06rs10843344 Marker Name ASR 39.25 41.42 С 39.25 40.05 Yellow Т 40.6200000000005 41.42 Red Marker Name 07rs239031 ASR 43.27 45.87 С 43.27 44.07 Yellow Т 45.06 45.87000000000005 Red Marker Name 13rs182549 ASR 51.13 53.81 С 51.12999999999995 52.25 Yellow Т 53.01 53.80999999999999 Red Marker Name 28rs1886510 ASR 80.52 82.01 G 80.52000000000001 81.31 Blue 81.179999999999999 82.01 Green А Panel Name 34-PLEX Extra Mod Marker Name 02rs917118

ASR 69.9 71.5 69.899999999999999 70.55 Blue G 71.5 А 70.80000000000001 Green Marker Name 03rs1024116 ASR 32.4 34.99 G 32.40000000000006 33.24 Blue 33.77 34.99 Green А Marker Name 05rs722098 36.67 ASR 39.4 36.67 37.47 Blue G А 38.6 39.4 Green Marker Name 06rs10843344 39.25 41.42 ASR С 39.25 40.05 Yellow Т 40.62000000000005 41.42 Red Marker Name 07rs239031 43.27 45.87 ASR 43.27 44.07 Yellow С Т 45.06 45.8700000000000 Red Marker Name 13rs182549 51.13 53.81 ASR С 51.12999999999995 52.25 Yellow 53.01 53.80999999999995 Red Т Marker Name 10rs1978806 ASR 47.0 49.9 С 47.0 48.0 Yellow Т 48.9 49.9 Red 34-Plex Electrophoretic Shift 2 Panel Name Marker Name 01rs1321333 ASR 32.86 36.83 С 32.86 33.98999999999995 Yellow Т 35.61 36.8300000000005 Red Marker Name 02rs917118 ASR 32.6 35.67 G 32.6 33.809999999999995 Blue 34.87000000000005 35.67 Green А Marker Name 03rs1024116 ASR 33.35 36.39 G 33.35 34.54 Blue А 35.17 36.39 Green Marker Name 04rs7897550 ASR 36.47 39.46 С 36.47 37.3 Yellow Т 38.37 39.46 Red Marker Name 05rs722098 36.72 39.59 ASR G 36.72 37.5199999999999 Blue 38.79 39.589999999999996 Green А Marker Name 06rs10843344 ASR 41.68 44.0 С 41.68 42.48 Yellow Т 43.2 44.0 Red

Marker Name 07rs239031 42.35 45.17 ASR С 42.35 43.15 Yellow Т 44.36 45.17 Red Marker Name 08rs12913832 ASR 44.16 45.64 G 44.1600000000004 44.96 Blue 44.84 45.64 Green А Marker Name 09rs2040411 46.48 48.49 ASR 46.4800000000004 47.78 Blue G 47.6900000000005 48.49 Green А Marker Name 10rs1978806 ASR 47.82 49.26 С 47.82 48.94000000000005 Yellow Т 48.46 49.26 Red Marker Name 11rs773658 ASR 48.55 50.04 G 48.5500000000004 49.35 Blue С 49.24 50.04 Yellow Marker Name 12rs10141763 50.32 52.76 ASR 50.32 51.3899999999999 Green А Т 51.58 52.76 Red Marker Name 13rs182549 ASR 51.0 53.29 С 51.0 52.1200000000005 Yellow Т 52.49 53.29 Red Marker Name 14rs1573020 ASR 51.23 53.38 G 51.23 52.73 Blue 52.13 53.38 Green А Marker Name 15rs896788 54.01 56.51 ASR С 54.01 55.38 Yellow Т 55.66000000000004 56.51 Red Marker Name 16rs2065160 ASR 56.36 57.68 56.36 57.2600000000000 Blue G А 56.88 57.68 Green Marker Name 17rs2572307 ASR 58.03 59.66 G 58.03 58.83 Blue 58.86 59.66 Green А Marker Name 18rs2303798 59.78 61.12 ASR С 59.78 60.58 Yellow Т 60.32 61.12 Red Marker Name 19rs2065982 ASR 62.31 63.86 G 62.31 63.11 Blue А 63.06 63.86 Green

Marker Name 20rs3785181 ASR 63.89 66.12 63.89 64.79 Yellow С Т 65.32 66.12 Red Marker Name 21rs881929 67.39 ASR 64.29 G 64.29 65.26 Blue Т 66.589999999999999 67.39 Red Marker Name 22rs1498444 ASR 67.19 68.17 67.99000000000001 А 67.19 Green С 67.36999999999999 68.17 Yellow Marker Name 23rs1426654 70.48 72.13 ASR С 70.479999999999999 71.28 Yellow Т 71.33 72.1300000000001 Red 24rs2026721 Marker Name ASR 71.28 72.65 G 71.28 72.0800000000001 Blue А 71.85 72.65 Green Marker Name 25rs4540055 ASR 74.18 76.08 75.3000000000001 А 74.5 Green 74.179999999999999 С 74.98 Yellow Т 75.28 76.0800000000001 Red Marker Name 26rs16891982 ASR 77.66 79.03 G 77.66 78.46000000000001 Blue С 78.229999999999999 79.03 Yellow Marker Name 27rs1335873 79.89 ASR 81.1 79.89 80.69000000000001 Green А Т 80.3 81.10000000000001 Red Marker Name 28rs1886510 ASR 80.52 82.01 G 80.5200000000001 81.31 Blue А 81.179999999999999 82.01 Green Marker Name 29rs730570 ASR 82.12 83.68 С 82.12 83.06 Yellow Т 82.67 83.679999999999999 Red Marker Name 30rs5030240 ASR 85.05 86.71 G 85.05 85.67 Blue A 86.05 86.71 Green 85.95 С 86.58 Yellow Marker Name 31rs2304925 ASR 86.42 88.52 G 86.42 86.96 Blue Т 87.75 88.5200000000001 Red 32rs5997008 Marker Name ASR 88.14 88.85

88.14 88.83 Green А С 88.23 88.8500000000001 Yellow 33rs3827760 Marker Name 90.46 91.43 ASR G 90.46 91.02 Blue 91.42999999999999 Green 90.73 А Marker Name 34rs2814778 92.14 ASR 90.16 С 90.16 91.15 Yellow Т 91.26 92.14 Red 34-plex Elec Feb 2014 Panel Name Marker Name 01rs1321333 32.62 35.85 ASR С 32.6200000000005 33.75 Yellow 34.629999999999995 35.85 Red Т Marker Name 02rs917118 31.54 34.26 ASR G 31.5400000000003 32.75 Blue А 33.46 34.26 Green Marker Name 03rs1024116 31.87 34.99 ASR 31.87000000000005 33.06 Blue G А 33.77 34.99 Green Marker Name 04rs7897550 36.44 39.3 ASR С 36.44000000000005 37.27 Yellow Т 38.21 39.30000000000004 Red Marker Name 05rs722098 36.94 39.87 ASR G 36.9400000000005 37.74 Blue 39.07 39.87 Green А Marker Name 06rs10843344 ASR 40.55 43.09 С 40.55000000000004 41.35 Yellow Т 42.29 43.08999999999996 Red Marker Name 07rs239031 ASR 41.21 44.01 С 41.21 42.01 Yellow Т 43.199999999999996 44.01 Red Marker Name 08rs12913832 ASR 43.28 44.56 G 43.28 44.08 Blue 43.76 44.55999999999999 Green А Marker Name 09rs2040411 ASR 45.69 47.51 G 45.6900000000005 46.99 Blue А 46.71 47.51 Green Marker Name 10rs1978806 ASR 47.22 49.46 С 47.22 48.34 Yellow Т 48.6600000000004 49.46 Red Marker Name 11rs773658

ASR 48.36 49.88 G 48.36 49.16 Blue 49.08 49.879999999999995 Yellow С Marker Name 12rs10141763 ASR 50.6 53.45 50.6 51.66999999999995 Green Α Т 51.91 53.45 Red Marker Name 13rs182549 ASR 51.13 53.81 51.12999999999995 52.25 Yellow С Т 53.01 53.80999999999999 Red Marker Name 14rs1573020 52.36 54.19 ASR G 53.86 Blue 52.36 А 52.94 54.19 Green Marker Name 15rs896788 55.02 57.07 ASR С 55.01999999999996 56.39 Yellow Т 56.2200000000006 57.07 Red Marker Name 16rs2065160 56.85 58.04 ASR 56.84999999999994 57.75 Blue G А 57.24 58.04 Green Marker Name 17rs2572307 58.03 59.66 ASR G 58.03 58.83 Blue 58.86 А 59.66 Green 18rs2303798 Marker Name 59.78 61.12 ASR С 59.78 60.58 Yellow Т 60.32 61.12 Red Marker Name 19rs2065982 ASR 62.31 63.86 62.31 63.11 Blue G А 63.06 63.86 Green Marker Name 20rs3785181 ASR 63.89 66.12 С 63.89 64.79 Yellow Т 65.32 66.12 Red Marker Name 21rs881929 ASR 64.29 67.39 G 64.29 65.26 Blue Т 66.5899999999999 67.39 Red Marker Name 22rs1498444 ASR 67.19 68.17 67.19 67.9900000000001 А Green С 67.36999999999999 68.17 Yellow Marker Name 23rs1426654 ASR 70.48 72.13 С 70.47999999999999 71.28 Yellow Т 71.33 72.13000000000001 Red Marker Name 24rs2026721

ASR 71.28 72.65 72.08000000000001 G 71.28 Blue 72.65 Green А 71.85 Marker Name 25rs4540055 ASR 74.18 76.08 А 74.5 75.3000000000001 Green С 74.1799999999999 74.98 Yellow Т 75.28 76.0800000000001 Red Marker Name 26rs16891982 ASR 77.66 79.03 78.46000000000001 G 77.66 Blue С 78.229999999999999 79.03 Yellow Marker Name 27rs1335873 79.89 ASR 81.1 А 79.89 80.69000000000001 Green Т 80.3 81.1000000000001 Red 28rs1886510 Marker Name ASR 80.52 82.01 G 80.5200000000001 81.31 Blue А 81.179999999999999 82.01 Green Marker Name 29rs730570 ASR 82.12 83.68 С 82.12 83.06 Yellow Т 82.67 83.679999999999999 Red Marker Name 30rs5030240 ASR 85.05 86.71 G 85.05 85.67 Blue 86.05 А 86.71 Green С 85.95 86.58 Yellow Marker Name 31rs2304925 86.12 87.95 ASR G 86.12 86.66 Blue Т 87.179999999999999 87.95 Red Marker Name 32rs5997008 87.74 ASR 88.78 88.429999999999999 А 87.74 Green С 88.16 88.78 Yellow Marker Name 33rs3827760 ASR 90.46 91.29 91.02 Blue G 90.46 А 90.59 91.289999999999999 Green Marker Name 34rs2814778 ASR 89.64 91.85 С 89.64 90.6300000000001 Yellow Т 90.97 91.85 Red

Version GM v 3.	0		
Kit type:	SNP		
Chemistry Kit	34-PLE	X none	
Panel 34-Plex	Electro	phoretic Shift	none
01rs1321333	-	none	
02rs917118	-	none	
03rs1024116	-	none	
04rs7897550	-	none	
05rs722098	-	none	
06rs10843344	-	none	
07rs239031	-	none	
08rs12913832	-	none	
09rs2040411	-	none	
10rs1978806	-	none	
11rs773658	-	none	
12rs10141763	-	none	
13rs182549	-	none	
14rs1573020	-	none	
15rs896788	-	none	
16rs2065160	-	none	
17rs2572307	-	none	
18rs2303798	-	none	
19rs2065982	-	none	
20rs3785181	-	none	
21rs881929	-	none	
22rs1498444	-	none	
23rs1426654	-	none	
24rs2026721	-	none	
25rs4540055	-	none	
26rs16891982	-	none	
27rs1335873	-	none	
28rs1886510	-	none	
29rs730570	-	none	
30rs5030240	-	none	
31rs2304925	-	none	
32rs5997008	-	none	
33rs3827760	-	none	
34rs2814778	-	none	
Panel 34-PLE	x	none	
01rs1321333	-	none	
02rs917118	_	none	
02rs1024116	_	none	
04rs7897550	_	none	
05rs722098	_	none	
0513722070 06rs10843344	_	none	
001310043344	_	none	
0.8rs12012822	_	none	
09rs2010002	_	none	
10rs1978806	-	none	
11rs772658	_	none	
12rs10141762	_	none	
13rs182549	-	none	
1010102017			

14rs1573020 -	none
15rs896788 -	none
16rs2065160 -	none
17rs2572307 -	none
18rs2303798 -	none
19rs2065982 -	none
20rs3785181 -	none
21rs881929 -	none
22rs1498444 -	none
23rs1426654 -	none
24rs2026721 -	none
25rs4540055 -	none
26rs16891982 -	none
201310051502 27rs1335873	none
271313335073 28rs1886510	none
20131000310 = 20rc720570 =	none
2913730370 - $30rc5030240$ -	none
$21rc^{2}20/025$	none
31132304923 - 22m = 0.07000	none
32185997008 -	none
33r\$382/760 -	none
54152014//0 -	
Panel 34-PLEX	Extra none
021391/110 - 02m = 1024116	none
05151024110 - 05m = 722000	none
0515722090 - 06m = 10042244	none
001310043344 - 07m 220021	none
0/rs239031 -	none
13r\$182549 -	none
28r\$1886510 -	none
Panel 34-PLEX	Lintua Mad nama
02 017110	Extra Mod none
02rs917118 -	Extra Mod none none
02rs917118 - 03rs1024116 -	Extra Mod none none none
02rs917118 - 03rs1024116 - 05rs722098 -	Extra Mod none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 -	Extra Mod none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 -	Extra Mod none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 -	Extra Mod none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 -	Extra Mod none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex F	Extra Mod none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex H 01rs1321333 -	Extra Mod none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex F 01rs1321333 - 02rs917118 -	Extra Mod none none none none none none none Electrophoretic Shift 2 none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex F 01rs1321333 - 02rs917118 - 03rs1024116 -	Extra Mod none none none none none none none Electrophoretic Shift 2 none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex H 01rs1321333 - 02rs917118 - 03rs1024116 - 04rs7897550 -	Extra Mod none none none none none none none Electrophoretic Shift 2 none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex F 01rs1321333 - 02rs917118 - 03rs1024116 - 04rs7897550 - 05rs722098 -	Extra Mod none none none none none none none Electrophoretic Shift 2 none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex F 01rs1321333 - 02rs917118 - 03rs1024116 - 04rs7897550 - 05rs722098 - 06rs10843344 -	Extra Mod none none none none none none none Electrophoretic Shift 2 none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex F 01rs1321333 - 02rs917118 - 03rs1024116 - 04rs7897550 - 05rs722098 - 06rs10843344 - 07rs239031 -	Extra Mod none none none none none none none Electrophoretic Shift 2 none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex F 01rs1321333 - 02rs917118 - 03rs1024116 - 04rs7897550 - 05rs722098 - 06rs10843344 - 07rs239031 - 08rs12913832 -	Extra Mod none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex H 01rs1321333 - 02rs917118 - 03rs1024116 - 04rs7897550 - 05rs722098 - 06rs10843344 - 07rs239031 - 08rs12913832 - 09rs2040411 -	Extra Mod none none none none none none none enone Electrophoretic Shift 2 none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex F 01rs1321333 - 02rs917118 - 03rs1024116 - 04rs7897550 - 05rs722098 - 06rs10843344 - 07rs239031 - 08rs12913832 - 09rs2040411 - 10rs1978806 -	Extra Mod none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex F 01rs1321333 - 02rs917118 - 03rs1024116 - 04rs7897550 - 05rs722098 - 06rs10843344 - 07rs239031 - 08rs12913832 - 09rs2040411 - 10rs1978806 - 11rs773658 -	Extra Mod none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex F 01rs1321333 - 02rs917118 - 03rs1024116 - 04rs7897550 - 05rs722098 - 05rs722098 - 05rs722098 - 06rs10843344 - 07rs239031 - 08rs12913832 - 09rs2040411 - 10rs1978806 - 11rs773658 - 12rs10141763 -	Extra Mod none none none none none none none none
02rs917118 - 03rs1024116 - 05rs722098 - 06rs10843344 - 07rs239031 - 13rs182549 - 10rs1978806 - Panel 34-Plex F 01rs1321333 - 02rs917118 - 03rs1024116 - 04rs7897550 - 05rs722098 - 05rs722098 - 06rs10843344 - 07rs239031 - 08rs12913832 - 09rs2040411 - 10rs1978806 - 11rs773658 - 12rs10141763 - 13rs182549 -	Extra Mod none none none none none none none none

15rs896788	-	none	
16rs2065160	-	none	
17rs2572307	-	none	
18rs2303798	-	none	
19rs2065982	-	none	
20rs3785181	-	none	
21rs881929	-	none	
22rs1498444	-	none	
23rs1426654	-	none	
24rs2026721	-	none	
25rs4540055	-	none	
26rs16891982	-	none	
27rs1335873	-	none	
28rs1886510	-	none	
29rs730570	-	none	
30rs5030240	-	none	
31rs2304925	-	none	
32rs5997008	_	none	
33rs3827760	_	none	
34rs2814778	_	none	
Panel 34-nlev	Flec	Feb 2014	none
01rc1321333	-	none	none
01131321333 02rc017118	_	none	
0213717110 03rs1024116	_	none	
0.3131024110 0.4rc7897550	_	none	
04137077330	_	none	
0.513722090 0.6rc10843344	-	nono	
001310043344	-	nono	
0713239031 08rc12012822	-	nono	
001312713032 00rc201011	_	nono	
10rc1078806	-	nono	
10131970000 11rc772658	-	nono	
1115773030 12ma10141762	-	none	
121510141703 12ma102540	-	none	
1315102349 14ma1572020	-	none	
141515/3020 15m006700	-	none	
1515090700 16ro2065160	-	none	
10152005100 17mc2572207	-	none	
1/1525/250/	-	none	
10152303790	-	none	
19rs2065982	-	none	
20rs3/85181	-	none	
21rs881929	-	none	
22rs1498444	-	none	
23rs1426654	-	none	
24rs2026721	-	none	
25rs4540055	-	none	
26rs16891982	-	none	
27rs1335873	-	none	
28rs1886510	-	none	
29rs730570	-	none	
30rs5030240	-	none	
31rs2304925	-	none	

32rs5997008	-	none
33rs3827760	-	none
34rs2814778	-	none

Version	GM v 3.	0		
Chemist	try Kit	AIM-ind	lelplex	
BinSet I	Vame	AIM-ind	lelplex	
Panel N	ame	AIM-ind	lelplex	
Marker	Name	1470		
1	61.28	0.5	0.76	
2	66.54	0.5	0.5	
Marker	Name	777		
1	70.9	0.5	0.5	
2	73.85	0.5	0.5	
Marker	Name	196		
1	81.19	0.5	0.5	
2	84.3	0.5	0.5	
Marker	Name	881		
1	89.41	0.5	0.5	
2	93.46	0.5	0.5	
– Marker	Name	3122	010	
1	98.52	0.5	0.5	
2	103.05	0.5	0.5	
- Marker	Name	548	010	
1	107 97	05	05	
2	109.91	0.5	0.5	
2	104.91	0.5	0.5	mutant
J Marker	Name	659	0.5	mutant
1	116 64	05	0.63	
2	118.07	0.5	0.05	
2 Markor	Name	2011	0.5	
1	127 A5	05	05	
2	127.45	0.5	0.5	
4 Markor	152.01 Namo	0.J 2020	0.5	
1	150 72	05	05	
1 2	152.05	0.5	0.5	
4 Markor	152.95 Namo	0.J 502	0.5	
1	156 67	05	05	
1	150.07	0.5	0.5	
4 Markor	150.05 Namo	0.5 700	0.5	
1	160 67	790	05	
1	174 57	0.5	0.5	
4 Markor	1/4.5/ Namo	0.5	0.5	
	101 1C	1193	0 5	
1	101.40 102 E	0.5	0.5	
2 Markor	105.5 Nomo	0.5	0.5	
	101 2	10/1	0 5	
1	191.5 102.2F	0.5	0.5	
Z Maulaau	193.25 Name	0.5	0.66	
Marker	Name	1/	0 5	
1	200.28	0.5	0.5	
۷ Marsler	204.Z	0.5	0.5	
Marker	Name	2538	0.75	
1	210.23	0.5	0.75	
۲ ا	Z13.79	0.5	0.75	
Marker	Name	1644	0 5	
1	222.95	0.5	0.5	

2	224.81	0.5	0.5	
Marker	Name	3854		
1	56.54	0.97	0.5	
2	60.12	0.59	0.5	
Marker	Name	2275		
1	70.9	0.5	0.5	
2	77.92	0.5	0.5	
– Marker	Name	3072	010	
1	106 57	05	05	
1 2	1125	0.5	0.5	
4 Markar	Namo	0.5	0.5	
1 Mai Kei	110 A	//Z	0 5	
1	110.4	0.5	0.5	
	121.6	0.5	0.5	
Marker	Name	2313	~ -	
1	128.4	0.5	0.5	
2	137.67	0.5	0.5	
Marker	Name	397		
1	164.01	0.5	0.5	
2	168.03	0.5	0.5	
Marker	Name	1636		
1	174.2	0.5	0.5	
2	175.88	0.5	0.5	
Marker	Name	51		
1	184.47	0.5	0.5	
2	189.56	0.5	0.5	
– Marker	Name	2431	010	
1	209.32	05	0.75	
1 2	207.52	0.5	205	0.75
2 Markor	215.09- Namo	226A	0.5	0.75
1	110111C	2204 0 E	0 5	
1	225.25	0.5	0.5	
3	228.0	0.4	0.4	mutant
	229.12	0.4	0.66	
Marker	Name	2256	~ <del>-</del>	
1	58.21	0.5	0.5	
2	61.3746	596174	0.5	0.5
Marker	Name	128		
1	68.29	0.5	0.5	
2	71.41	0.5	0.5	
Marker	Name	15		
1	78.8	0.5	0.5	
2	89.72	0.5	0.64	
Marker	Name	2241		
1	107.93	0.64	0.5	
2	115.84	0.5	0.5	
Marker	Name	419		
1	120.77	0.5	0.65	
2	127.68	0.5	0.5	
- Marker	Name	943	0.0	
1	156.89	0.56	05	
- 2	160.77	0.50	0.5	
4 Markov	Name	150	0.5	
	160 75	139	0 5	
T	100./2	0.01	0.5	

2	173.8	0.5	0.5	
Marker	Name	2005		
1	184.78	0.5	0.55	
2	190.57	0.64	0.5	
Marker	Name	250		
1	201.03	0.5	0.5	
2	203.1	0.5	0.5	
Marker	Name	1802		
1	213.67	0.5	0.5	
2	216.64	0.5	0.5	
Marker	Name	1607		
1	221.42	0.5	0.5	
2	223.03	0.62	0.5	
Marker	Name	406		
1	66.96	0.5	0.5	
2	68.77	0.5	0.5	
Marker	Name	1386		
1	72.66	0.5	0.71	
2	94.14	0.5	0.5	
Marker	Name	1726		
1	105.6	0.5	0.5	
2	118.28	0.5	0.5	
Marker	Name	3626		
1	142.49	0.5	0.5	
2	158.53	0.5	0.5	
Marker	Name	360		
1	170.34	0.4	0.4	
3	171.2	0.4	0.4	mutant
2	172.15	0.4	0.4	
Marker	Name	1603		
1	214.33	0.5	0.56	
2	218.0	0.5	0.5	
Marker	Name	2719		
1	227.71	0.5	0.5	
2	231.5	0.5	0.5	
Marker	Name	1734		
1	57.94	0.5	0.5	
2	61.77	0.5	0.5	
Marker	Name	94		
1	91.31	0.5	0.5	
2	94.25	0.5	0.56	

Version GM v 3.0												
Kit type	e:	MICROSATELLITE										
Chemis	try Kit	AIM-indelplex none										
Panel	AIM-ind	delplex none										
1470	blue	59.498930343000005	68.0	-	9	0.0	none					
777	blue	68.55843072 76.0	-	9	0.0	none						
196	blue	77.588872041 85.0	-	9	0.0	none						
881	blue	85.617309136 94.0	-	9	0.0	none						
3122	blue	95.320690594 103.8	-	9	0.0	none						
548	blue	104.2 110.8 -	9	0.0	none							
659	blue	113.556264981120.0	-	9	0.0	none						
2011	blue	124.498164706135.0	-	9	0.0	none						
2929	blue	146.955519673153.8	-	9	0.0	none						
593	blue	154.7 160.0 -	9	0.0	none							
798	blue	165.32897286899998	176.0	-	9	0.0	none					
1193	blue	178.658800433186.0	-	9	0.0	none						
1871	blue	188.823452509195.5	-	9	0.0	none						
17	blue	197.084416798205.0	-	9	0.0	none						
2538	blue	206.773228342216.0	-	9	0.0	none						
1644	blue	219.483987195226.0	-	9	0.0	none						
3854	Green	54.61523001 64.0	-	9	0.0	none						
2275	Green	68.507204408 80.0	-	9	0.0	none						
3072	Green	103.268780394114.0	-	9	0.0	none						
772	Green	115.766532729123.0	-	9	0.0	none						
2313	Green	125.024081548139.5	-	9	0.0	none						
397	Green	161.189615282170.0	-	9	0.0	none						
1636	Green	172.08629923200002	177.5	-	9	0.0	none					
51	Green	181.211398852191.0	-	9	0.0	none						
2431	Green	206.293870962216.0	-	9	0.0	none						
2264	Green	221.926816896231.0	-	9	0.0	none						
2256	Yellow	56.680080642 64.0	-	9	0.0	none						
128	Yellow	66.0 72.0 -	9	0.0	none							
15	Yellow	76.5 91.0 -	9	0.0	none							
2241	Yellow	104.973806014116.5	-	9	0.0	none						
419	Yellow	118.07762833899999	129.0	-	9	0.0	none					
943	Yellow	153.696434559162.0	-	9	0.0	none						
159	Yellow	166.02409311300002	175.0	-	9	0.0	none					
2005	Yellow	181.501371715192.0	-	9	0.0	none						
250	Yellow	198.065914853204.0	-	9	0.0	none						
1802	Yellow	210.217316797218.0	-	9	0.0	none						
1607	Yellow	218.9 224.0 -	9	0.0	none							
406	Red	64.678876207 69.5	-	9	0.0	none						
1386	Red	70.5 96.0 -	9	0.0	none							
1726	Red	104.0 121.0 -	9	0.0	none							
3626	Red	141.0 160.0 -	9	0.0	none							
360	Red	168.792202855174.0	-	9	0.0	none						
1603	Red	213.0 219.0 -	9	0.0	none							
2719	Red	226.0 232.5 -	9	0.0	none							
1734	Red	56.285110841 63.5	-	9	0.0	none						
94	Green	88.363699841 96.0	-	9	0.0	none						

Supplementary File S3. SNP and Indel genotypes used in the exercise as reference poj

- File S3.1. PCA input file formatted for use with *Snipper* at: <u>http://math</u>
- File S3.2. Training set file applicable to custom analyses (and used as fixed reference
- File S3.3. Likelihood ratios from cross-validation of PCA input file genotypes (compa

Note: Worksheets S3.1 and S3.2 need to be placed in 'position 1' to use make the genoty;

LRs obtained with Snipper from complete profile data in Supplementary File S3.1

Inference	34-plex, 3-group
European	9947A is 2,118,840,589,047,061,020,672 times more likely EUROPE than E
East Asian	A is 361,148,635,069,545,024 times more likely EAST ASIA than EL
European	B is 64,191,487,284,485,608 times more likely EUROPE than EAST ASIA
East Asian	C is 13,115,706 times more likely EAST ASIA than AFRICA
East Asian	D is 248,539,593,557 times more likely EAST ASIA than EUROPE
African	E is 556,454,701,312,037,054,117,314,560 times more likely AFRIC
	46-plex, 4-group
European	9947A is 1,937,432,967,198 times more likely EUROPE than EAST ASIA
East Asian	A is 6,993,957 times more likely EAST ASIA than AMERICA
European	B is 143,659,679,122 times more likely EUROPE than EAST ASIA
East Asian	C is 131 times more likely EAST ASIA than EUROPE
American	D is 944,698,134 times more likely AMERICA than EAST ASIA
African	E is 3,229,841,442,838,053,650,432 times more likely AFRICA than
	80 Markers, 5-group
Oceanian	C is 153,747,536,542,653 times more likely OCEANIA than EAST ASIA



#### **Supplementary File S4**

A. Numbers of heterozygotes (bars) and PHR values (points) plotted for all participant's AIM-Indel data. Samples A-E are average values, sample F is per participant and includes laboratories #15, #20, #1 and #18 removed from the PHR comparisons made with the Kruskal-Wallis rank sum test summarized in C.

**B.** Grid of *p*-values for the pairwise comparison of numbers of heterozygotes in A-E (average number from 19 laboratories) and individual numbers per laboratory for sample F, applying a unilateral 2-sample test for equality-of-proportions (with continuity correction). Grey cells mark significant *p*-values.

В

-	Sample-A	Sample-B	Sample-C	Sample-D	Sample-E	Lab-15	Lab-20	Lab-1	Lab-18	Lab-7	Lab-6	Lab-9	Lab-11	Lab-21	Lab-2	Lab-17	Lab-4	Lab-19	Lab-12	Lab-16	Lab-8	Lab-13	Lab-14
Sample-B	0.00018																						
Sample-C	0.01582	0.90390																					
Sample-D	0.04519	0.96120	0.59380																				
Sample-E	0.04519	0.96120	0.59380	0.50000																			
Lab-15	0.00144	0.66470	0.25230	0.12860	0.12860																		
Lab-20	0.00144	0.66470	0.25230	0.12860	0.12860	0.50000																	
Lab-1	0.00074	0.58380	0.18880	0.08909	0.08909	0.50000	0.50000																
Lab-18	8.85e-05	0.50000	0.06527	0.02441	0.02441	0.26260	0.26260	0.33650															
Lab-7	1.88e-05	0.33800	0.02726	0.00875	0.00875	0.14650	0.14650	0.20070	0.41730														
Lab-6	1.50e-06	0.14860	0.00568	0.00146	0.00146	0.04729	0.04729	0.07201	0.20210	0.33800													
Lab-9	1.50e-06	0.14860	0.00568	0.00146	0.00146	0.04729	0.04729	0.07201	0.20210	0.33800	0.50000												
Lab-11	1.50e-06	0.14860	0.00568	0.00146	0.00146	0.04729	0.04729	0.07201	0.20210	0.33800	0.50000	0.50000											
Lab-21	1.50e-06	0.14860	0.00568	0.00146	0.00146	0.04729	0.04729	0.07201	0.20210	0.33800	0.50000	0.50000	0.50000										
Lab-2	1.50e-06	0.14860	0.00568	0.00146	0.00146	0.04729	0.04729	0.07201	0.20210	0.33800	0.50000	0.50000	0.50000	0.50000									
Lab-17	6.09e-07	0.10540	0.00313	0.00075	0.00075	0.03016	0.03016	0.04761	0.14830	0.33800	0.50000	0.50000	0.50000	0.50000	0.50000								
Lab-4	6.09e-07	0.10540	0.00313	0.00075	0.00075	0.03016	0.03016	0.04761	0.14830	0.33800	0.50000	0.50000	0.50000	0.50000	0.50000	0.50000							
Lab-19	2.39e-07	0.07201	0.00166	0.00037	0.00037	0.01851	0.01851	0.03028	0.10500	0.20070	0.41620	0.41620	0.41620	0.41620	0.41620	0.50000	0.50000						
Lab-12	2.39e-07	0.07201	0.00166	0.00037	0.00037	0.01851	0.01851	0.03028	0.10500	0.20070	0.41620	0.41620	0.41620	0.41620	0.41620	0.50000	0.50000	0.50000					
Lab-16	2.39e-07	0.07201	0.00166	0.00037	0.00037	0.01851	0.01851	0.03028	0.10500	0.20070	0.41620	0.41620	0.41620	0.41620	0.41620	0.50000	0.50000	0.50000	0.50000				
Lab-8	9.09e-08	0.04729	0.00085	0.00017	0.00017	0.01091	0.01091	0.01851	0.07144	0.14650	0.33530	0.33530	0.33530	0.33530	0.33530	0.41540	0.41540	0.50000	0.50000	0.50000			
Lab-13	3.34e-08	0.02980	0.00041	7.98e-05	7.98e-05	0.00616	0.00616	0.01084	0.04667	0.10280	0.26080	0.26080	0.26080	0.26080	0.26080	0.33380	0.33380	0.41450	0.41450	0.41450	0.50000		
Lab-14	3.34e-08	0.02980	0.00041	7.98e-05	7.98e-05	0.00616	0.00616	0.01084	0.04667	0.10280	0.26080	0.26080	0.26080	0.26080	0.26080	0.33380	0.33380	0.41450	0.41450	0.41450	0.50000	0.50000	
Lab-5	3.335e-08	0.02980	0.00041	7.98e-05	7.98e-05	0.006161	0.006161	0.01084	0.04667	0.10280	0.26080	0.26080	0.26080	0.26080	0.26080	0.33380	0.33380	0.41450	0.41450	0.41450	0.50000	0.50000	0.50000

# **C**. Grid of *p*-values for pairwise comparisons of PHR values applying a Kruskal-Wallis rank sum test (grey cells: significant values)

	Sample-A	Sample-B	Sample-C	Sample-D	Sample-E	Lab-7	Lab-6	Lab-9	Lab-11	Lab-21	Lab-2	Lab-17	Lab-4	Lab-19	Lab-12	Lab-16	Lab-8	Lab-13	Lab-14
Sample-B	0.24530																		
Sample-C	0.36500	0.05085																	
Sample-D	0.60150	0.26500	0.40090																
Sample-E	0.36080	0.86880	0.31060	0.41180															
Lab-7	0.00777	1.61e-06	3.58e-05	0.00033	5.56e-05														
Lab-6	0.01237	2.36e-06	9.51e-05	0.00033	8.82e-05	0.33620													
Lab-9	0.00341	1.92e-07	3.36e-06	5.80e-05	1.15e-05	0.37810	0.89810												
Lab-11	0.00727	6.92e-07	2.04e-05	0.00017	3.78e-05	0.17310	0.88360	0.67380											
Lab-21	0.01237	2.93e-06	0.00011	0.00026	8.82e-05	0.05445	0.38970	0.30540	0.36020										
Lab-2	0.00341	1.05e-07	3.36e-06	5.80e-05	1.15e-05	0.13820	0.71430	0.49830	0.78370	0.54590									
Lab-17	0.00672	5.98e-07	1.58e-05	0.00018	3.07e-05	0.29770	0.90090	1.00000	0.73530	0.24030	0.60590								
Lab-4	0.01128	1.29e-06	3.01e-05	0.00023	8.05e-05	0.28880	0.97160	0.88680	0.76230	0.34570	0.70870	0.86950							
Lab-19	0.00524	2.40e-07	1.24e-05	8.41e-05	2.52e-05	0.12990	0.87620	0.65260	0.86260	0.44620	1.00000	0.56700	0.74910						
Lab-12	0.01670	1.23e-06	2.64e-05	0.00020	8.41e-05	0.29780	0.86260	0.71620	0.97240	0.57960	0.88980	0.78760	0.89290	0.98040					
Lab-16	0.01034	8.91e-07	2.06e-05	0.00012	6.48e-05	0.08159	0.43600	0.38670	0.42580	0.98620	0.60350	0.32880	0.41900	0.55520	0.62300				
Lab-8	0.00582	3.67e-07	9.82e-06	9.93e-05	2.08e-05	0.07849	0.43810	0.31990	0.50010	0.94620	0.64900	0.32120	0.45560	0.58730	0.62070	0.97450			
Lab-13	0.00882	6.68e-07	1.14e-05	0.00017	4.31e-05	0.55370	0.81810	0.84370	0.64550	0.30840	0.46980	0.81050	0.71320	0.57540	0.58600	0.37510	0.28170		
Lab-14	0.00276	3.69e-08	1.49e-06	2.58e-05	5.86e-06	0.13170	0.53240	0.41140	0.63380	0.78000	0.88250	0.47200	0.63160	0.68580	0.72040	0.80340	0.87950	0.34400	
Lab-5	0.00544	2.62e-07	7.86e-06	8.05e-05	1.74e-05	0.11850	0.59910	0.45970	0.74250	0.69340	0.96070	0.51230	0.64300	0.75560	0.68580	0.82750	0.91550	0.38310	0.92930

**Supplementary File S5** 

Click here to download e-component: Supplementary File S5.pdf

Supplementary File S5 Next generation sequencing experiments using exercise PCR multiplexes.

## 1. Post-PCR DNA processing for SNP analysis with the MiSeq

Libraries were prepared directly from PCR products using the Illumina TruSeq ChIP sample preparation kit. Libraries were then run on the MiSeq with the 300 cycle version 2 reagent kit and sequences aligned to a custom 'genome' containing the reference sequences for all 34 SNPs in a single unified strand using Burrows-Wheeler alignment. SNP genotypes were called using GATK.

### 2. Post-PCR DNA processing for Indel and SNP analysis with the TFS-LT Ion PGM<sup>™</sup>

Libraries were prepared directly from PCR products using the TFS-LT Ion Xpress<sup>™</sup> Plus gDNA fragment library preparation protocol applying the Ion Xpress<sup>™</sup> Plus Fragment Library Kit. This kit processes DNA not amplified using AmpliSeq primers by enabling ligation and nick repair reactions. Thereafter sequencing followed standard protocols using kits: Ion OneTouch<sup>™</sup> 200 Template v2 and Ion PGM<sup>™</sup> Sequencing 200 v2. Sequences were aligned to custom BED files and genotypes called from human genome build hg19 using TFS-LT Torrent Suite<sup>™</sup> 4.0.2.

#### **3.** Genotyping performance for samples A-E

Both NGS systems gave comparable high levels of genotyping performance for SNP analysis. Ion PGM<sup>TM</sup> had just singleton no-calls or missing data (no sequences detected carrying expected SNP sites), while MiSeq gave the only miscalled genotype in rs5030240, where a sequence ratio of A=5810 / G=9866 was recorded as a GG, although it was detected as atypical.

Indel genotyping performance with the Ion PGM<sup>™</sup> was slightly lower, but this could be due to alignment issues. The Indel rs60612424 (MID-3854) was not detected in any samples, while C had a disproportionately high number of no-calls that might be the result of population-specific flanking indels blocking secure alignment to the reference sequence.

**4.** Summary tables of genotyping performance from SNP analysis with both NGS systems and Indel analysis with Ion PGM<sup>™</sup>. Matches count concordant genotype calls made in both CE and NGS.

	lo	Ion PGM <sup>™</sup> sequence data for 34 SNPs												
	A	В	С	D	E	F								
NGS no-calls	1	0	0	1	1	1								
SNaPshot no-calls	1	0	0	0	0	0								
Ion PGM miscalls	0	0	0	0	0	4								
Missing data	0	1	1	1	0	1								
NGS-CE genotype matches *	32	33	33	32	33	28								
Heterozygote number	10	5	10	9	10	14								

MiSeq sequence	data for	34 SNPs
----------------	----------	---------

A	В	С	D	E	F
0	0	0	0	0	0
1	0	0	0	0	0
0	0	1	0	0	4
0	0	0	0	0	0
33	34	33	34	34	30
10	5	9	10	10	17

#### Ion PGM<sup>™</sup> sequence data for 46 Indels

	А	В	С	D	E	F
Ion PGM no-calls	4	1	12	3	1	3
PCR-to-CE no-calls	0	0	0	0	0	0
Ion PGM miscalls	0	3	3	0	0	6
Missing data	1	2	2	1	1	1
NGS-CE genotype matches *	41	40	29	42	44	36
Heterozygote number	4	17	11	10	10	21

\* Matching genotypes denote calls concordant with conventional CE analysis

**5.** NGS analysis of mixed sample F indicated a higher number of displaced sequence ratios positioned outside ranges around an ideal midline ratio of 0.5:0.5, although Indel genotyping showed several irregular ratios for A-E likely due to alignment issues. SNP genotyping with both platforms was very sensitive to imbalanced sequence ratios in F. In samples A-E, Ion PGM<sup>™</sup> detected **4/44** peak pairs outside nominal 0.4:0.6-0.5:0.5 sequence ratio ranges, and MiSeq **2/44** (discounting tri-allelics). In sample F by contrast, Ion PGM<sup>™</sup> detected **8/14** and MiSeq **14/17** peak pairs outside these ratios, providing clear indications of a mixture that is largely absent from SNaPshot data. Both systems detected displaced ratios in the two tri-allelic SNPs, as well as a remarkably well matched pattern of sample F sequence ratios across the 34 SNPs.

A

B

🔺 C

D



Indels in order of decreasing sequence coverage (left to right)
