Dawnay, L, Flamson, R and Dawnay, N

Expanded and mega-plex STR panels as a tool for presumptive population assignment

http://researchonline.ljmu.ac.uk/id/eprint/8788/

Article

Citation (please note it is advisable to refer to the publisher’s version if you intend to cite from this work)

Dawnay, L, Flamson, R and Dawnay, N (2018) Expanded and mega-plex STR panels as a tool for presumptive population assignment. Forensic Science and Criminology, 3 (2). ISSN 2513-8677

LJMU has developed LJMU Research Online for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk
Expanded and mega-plex STR panels as a tool for presumptive population assignment

Louise Dawnay¹, Robyn Flamson² and Nick Dawnay*²

¹Scientific Consultant, Chester, UK
²School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, UK

Abstract

Historically, forensic STR panels have been unsuccessful for population assignment due to the limited ancestry information that can be derived from the non-coding STR loci and the low number of loci included in the panel. However, given the recent adoption of expanded (16+ loci) and ‘mega-plex’ (23+ loci) STR panels, the ability to identify source population groups may be improved. This study assessed the impact of increasing locus number on population assignment under different analysis conditions using a published US population dataset comprised of individuals from the African American, Caucasian, Hispanic and Asian populations. The Bayesian clustering programme STRUCTURE was used to assess first, whether increasing the number of loci and the inclusion of known sample population data enabled greater resolution between the four populations in the dataset, and second, the utility for population assignment using criteria based on inferred ancestry scores. Results suggest that increasing the number of loci and including population of origin data allowed the identification of more distinct populations, with three primary populations being observed; African American, Asian, and Caucasian/Hispanic. The close grouping of the Caucasian and Hispanic populations is supported by their recently common ancestry from Western Europe. The ability of the programme to support population assignment to each of the four existing populations was assessed through the application of population and panel specific assignment thresholds based on the inferred ancestry scores obtained from the analysis programme. Predictive accuracy based on a training dataset of 984 individuals suggest that assignment accuracy is > 96% across the four populations and can reach 100% under some test conditions. The accuracy was > 90% when blind testing was performed on 40 ‘unknown’ individuals. As such, the approach described is considered within the acceptable range for a presumptive test and can be performed using data already collected as part of routine forensic investigations.

Introduction

Inference of the ethnic origin of a suspect from their DNA recovered from a crime scene sample can act as ‘investigative intelligence’ and help enforcement agencies concentrate their resources in the absence of any other suspect specific information. This idea is not new and has been explored extensively in the literature through the development and application of assignment approaches that use genetic markers to identify unique genetic groups or populations [1-3]. The specific identification of Ancestry Informative Markers (AIMs) panels that are particularly powerful at inferring ethnic origin has also been the focus of much research and the assessment of different classes of molecular marker has slowly moved from mitochondrial DNA (mtDNA) sequence variation, through autosomal and Y linked Short Tandem Repeat (STRs) markers to Single Nucleotide Polymorphisms (SNPs) [4-8]. SNP markers are currently considered the most applicable for use due to deficiencies displayed by other markers; uniparental markers (mtDNA and Y-STRs) typically require large datasets to be useful while autosomal STRs are not considered informative by some researchers in numbers less than 50 [9].

A number of biogeographically informative SNP marker panels have been developed [9,10], and with the adoption of Massively Parallel Sequencing (MPS) as a resource in forensic genetics, commercial MPS forensic panels are available that amplify both traditional STR, Y-STR phenotypic and AIM markers in a single reaction run [11-13]. However, despite the promise of MPS and the development of AIM panels, the forensic genetics community continues to use traditional multiplex-PCR kits for the amplification and size separation of STRs through capillary electrophoresis (CE). Reasons for the slow adoption of MPS technologies include high per sample cost, increased processing time and uncertainty around data handling and ethics [13,14]. As such, laboratories continue to use CE approaches and have seen the number of loci included in commercial panels virtually double in the last five years. For example, the European standard set (ESS) now comprises 17 STR loci and the American CODIS system comprises 20 core loci [15-18], while both the commercially available GloablFiler and PowerPlex Fusion kits boast an impressive 22+ STR loci [19,20]. Consequently, the increasing number of STR loci may now enable better resolution between populations, thus making expanded and mega-plex STR kits suitable for genetic differentiation between populations [21] and the inference of ethnic origin. Such an approach would be useful if labs continue to use CE methods for the foreseeable future.

Population assignment requires the use of a mathematical model that groups an unknown individual to a putative population and can be used to detect dispersal, hybridization, genetic mixture, origin of specific individuals, population delineation and structure [22]. Common population assignment models include Bayesian assignment, frequency-based, and Bayesian clustering approaches. The Bayesian

Accepted: March 05, 2018; Accepted: March 26, 2018; Published: March 29, 2018
The assignment approach developed by Rannala and Mountain [23] calculates the posterior probabilities that a genotype is observed at a locus when the individual belongs to each putative population. The probability is then determined for each locus (assuming no linkage) and multiplied, and results are provided as the posterior probability with lower values indicating rarer events. This approach has been used in the detection of poaching hot-spots [24], differentiation between closely related species [25] and the identification of illegally translocated deer [26]. An alternative, frequency-based method developed by Paeckel et al., [27] calculates genotype likelihood ratios and determines the probability that the genotype groups with each population using Monte Carlo resampling. This approach has been used to assign individual dogs to their population of origin [28], identify livestock predators [29], and to detect fishing competition fraud [30]. These two approaches, popular in molecular ecology, have seen little application in human population assignment, where research has concentrated on the development of bespoke models [7,10,31]. Perhaps one of the most common approaches to investigate human population genetic differentiation is the Bayesian clustering method developed by Pritchard et al., [32] which uses multi-locus genotype data to infer the number of distinct genetic clusters (populations) based on the allele frequencies observed in each population. Individuals across the dataset are assigned to single populations, or to multiple populations if admixture is detected. This approach has been successfully used to map clines in human population genetic structure with geography [33-35].

This proof of concept research aims to establish whether the increased number of loci used in expanded and mega-plex STR panels improves assignment accuracy and asks whether there is scope for a ‘presumptive’ population assignment test for forensic laboratories continuing to use CE based systems.

**Methods**

**Sample data**

STR profiles from 1036 individuals previously reported in ref. [36,37] were downloaded from Promega (https://www.promega.com/products/pm/genetic-identity/population-statistics/allele-frequencies/). Samples with genotyping errors as highlighted by Steffen et al., [38] were removed from the analysis resulting in a final sample set of 1024 individuals. This final dataset represents four common American population groups; African American (AA; n = 338), Caucasian (Ca; n = 358), Hispanic (His; n = 232) and Asian (As; n = 96). Genotype data were reformatted with new allele nomenclature (see supplemental Table 1) to allow for software analysis. Y-linked markers were removed from the dataset due to software input criteria, resulting in a panel comprising 30 autosomal STR loci. Separate input files were created to represent commonly used commercial STR profiling kits, each with a different number of STR loci; PowerPlex 16 (16 autosomal loci), AmpFISTR NGM Select (17 autosomal loci), AmpFISTR GlobalFiler™ (22 autosomal loci, and excluding the DYS391 Y-linked locus) and PowerPlex Fusion (23 autosomal loci, and excluding the YINDEL Y-linked locus).

**Population structuring**

The Bayesian clustering method STRUCTURE [32] was first used to identify the likely number of distinct genetic clusters (populations; K) existing in the data for each of the STR profiling kits. Two different analysis parameters (1 and 2) were initially tested to explore population structuring with and without the inclusion of known sample population data (Table 1). Each parameter set underwent five analysis iterations at each possible K (1-5). The optimal K was identified using three different approaches, avoiding the use of a single ad-hoc approach [32,39]: first, the highest mean log-likelihood value (lnPD) method outlined in ref. [32] was used; second, the ΔK method detailed in ref. [40] was calculated using the web-based STRUCTURE HARVESTER programme [41]; and third, the point at which the lnPD values begin to plateau as outlined in ref. [33]. CLUMPAK [42] was used to visualise the data. The use of the LOCPRIOR setting in parameter set two was shown to identify fine scale population differences more effectively and was selected for use when assessing population assignment.

**Population assignment**

Assignment accuracy was assessed under the expected number of populations (K=4) using analysis parameter three (Table 1) for each of the five STR profiling kits under study. This parameter set included the use of the POPFLAG feature in STRUCTURE that allows the assignment of individuals of unknown origin to a dataset containing individuals of known origin. Ten random individuals from each population were labelled as ‘unknowns’ resulting in a total of 984 known and 40 ‘unknown’ samples run using analysis parameter

| Analysis Parameter | Purpose | STRUCTURE Software Conditions | Estimated K using 3 methods |
|--------------------|---------|-------------------------------|---------------------------|
| Run Conditions     | Population Model | LOCPRIOR | POPFLAG | STR Panel and Locus number (n) | Highest lnPD | AK | Plateau lnPD |
| 1                  | Structure Identification | 100k burnin, 100k MCMC reps | Admixed | No | No | PowerPlex (16) | 3 | 2 | 2 |
|                    |                      |                  |          | NGS Select (17) | 3 | 2 | 3 |
|                    |                      |                  |          | Global Filer (22) | 3 | 2 | 3 |
|                    |                      |                  |          | PowerPlex Fusion (23) | 4 | 2 | 3 |
|                    |                      |                  |          | Combined Panel (30) | 5 | 2 | 4 |
| 2                  | Structure Identification | 100k burnin, 100k MCMC reps | Admixed | Yes | No | PowerPlex (16) | 3 | 2 | 2 |
|                    |                      |                  |          | NGS Select (17) | 4 | 2 | 3 |
|                    |                      |                  |          | Global Filer (22) | 5 | 2 | 3 |
|                    |                      |                  |          | PowerPlex Fusion (23) | 4 | 2 | 3 |
|                    |                      |                  |          | Combined Panel (30) | 4 | 2 | 3 |
| 3                  | Setting Population Assignment Criteria and Assessing | 100k burnin, 100k MCMC reps | Admixed | Yes | Yes | PowerPlex (16) | 5 | 2 | 3 |
|                    |                      |                  |          | NGS Select (17) | 5 | 2 | 3 |
|                    |                      |                  |          | Global Filer (22) | 5 | 2 | 3 |
|                    |                      |                  |          | PowerPlex Fusion (23) | 5 | 2 | 3 |
|                    |                      |                  |          | Combined Panel (30) | 5 | 2 | 3 |
set three. Predicted assignment accuracy was calculated based on the reduced data set of 984 individuals, which was used as a 'training' sample set to develop suitable acceptance criteria for assignment. To do this, the mean inferred ancestry scores for each individual from the five iterations when K=4 were calculated in CLUMPAK and the distributions for each population in each of the four clusters plotted. Assignment criteria for each population for each STR kit were then determined by setting a threshold for inferred ancestry score for each of the four genetic clusters. This approach is analogous to the setting of an analytical threshold to differentiate signal (true contribution) and noise (false contribution). Across the training dataset of 984 samples the number of individuals that were assigned to a single group using the defined criteria was calculated. An individual that satisfied all four criteria was given a score of four and categorised as either True Positives (TP) or False Positives (FP), while individuals that were given a scores of less than four were classified as True Negatives or False Negatives following the definitions supplied in Supplemental Table 2. Once categorised, these values were used to determine the predicted test sensitivity and specificity following the binary classification system outlined in ref. [43]. The ten randomly selected 'unknown' individuals from each population were then assessed to see how well they were assigned to the populations based on the defined criteria.

Results and Discussion

The STRUCTURE analysis shows that the number of distinct genetic groups (K) identified varies depending on which method is used in the estimation (Table 1). The recommended method for interpreting the correct K using STRUCTURE is to not use a single add-hoc approach as they each have limitations [32,39]. The ΔK method outlined in ref. [40] is predicted to underestimate the number of distinct clusters, as they each have limitations [32,39]. The ΔK method outlined in ref. [43]. The ten randomly selected 'unknown' individuals from each population were then assessed to see how well they were assigned to the populations based on the defined criteria. There are slight differences in the returned K value for analysis parameter sets 1 and 2, with a higher K obtained with less loci for parameter set 2 using both the plateau and highest lnPD. Parameter set 2 used the LOCPRIOR setting in STRUCTURE which allows the software to use information associated with the samples such as phenotype, in this instance the sample population of African American, Caucasian, Hispanic or Asian, to support the resolution of fine scale clusters [32]. Analysis under these parameters provided greater resolution to the inferred ancestry scores leading to more confidence in the population clusters (Figure 1), while not having a substantial impact on the number of clusters observed. For application in a forensic setting, the samples can be considered as a database containing samples of known origin allowing the use of the LOCPRIOR setting for population assignment.

Table 2: Predicted accuracy of population assignment test for different STR panels

| Population     | Accuracy | PowerPlex (16) | AmpFISTR NGM Select (17) | Global Filer (22) | PowerPlex Fusion (23) | Combined Panel (30) |
|----------------|----------|---------------|--------------------------|------------------|-----------------------|---------------------|
| African American | TP       | 328           | 328                      | 328              | 328                   | 328                 |
|                 | TN       | 656           | 656                      | 656              | 656                   | 656                 |
|                 | FP       | 0             | 0                        | 0                | 0                     | 0                   |
|                 | FN       | 0             | 0                        | 0                | 0                     | 0                   |
|                 | Sensitivity | 100%          | 100%                     | 100%             | 100%                  | 100%                |
|                 | Specificity | 100%          | 100%                     | 100%             | 100%                  | 100%                |
| Caucasian       | TP       | 348           | 348                      | 348              | 348                   | 348                 |
|                 | TN       | 636           | 636                      | 632              | 631                   | 635                 |
|                 | FP       | 0             | 0                        | 4                | 5                     | 1                   |
|                 | FN       | 0             | 0                        | 0                | 0                     | 0                   |
|                 | Sensitivity | 100%          | 100%                     | 100%             | 100%                  | 100%                |
|                 | Specificity | 100%          | 100%                     | 99.3%            | 99%                   | 99.8%               |
| Hispanic        | TP       | 222           | 222                      | 222              | 222                   | 222                 |
|                 | TN       | 676           | 762                      | 728              | 738                   | 760                 |
|                 | FP       | 0             | 0                        | 34               | 24                    | 2                   |
|                 | FN       | 0             | 0                        | 0                | 0                     | 0                   |
|                 | Sensitivity | 100%          | 100%                     | 100%             | 100%                  | 100%                |
|                 | Specificity | 100%          | 100%                     | 96%              | 97%                   | 99.7%               |
| Asian           | TP       | 86            | 86                       | 86               | 86                    | 85                  |
|                 | TN       | 898           | 898                      | 898              | 898                   | 898                 |
|                 | FP       | 0             | 0                        | 0                | 0                     | 0                   |
|                 | FN       | 0             | 0                        | 0                | 0                     | 0                   |
|                 | Sensitivity | 100%          | 100%                     | 100%             | 100%                  | 100%                |
|                 | Specificity | 100%          | 100%                     | 100%             | 100%                  | 100%                |

TP: True Positives, TN: True Negatives, FP: False Positives, FN: False Negatives. Sensitivity = TP/(TP+FN) and Specificity = TN/(TN+FP) calculated according to Ref. [43].

Forensic Sci Criminol, 2018 doi: 10.15761/FSC.1000126 Volume 3(1): 3-7
other populations group as a single cluster (Figure 1). The next cluster to appear when \( K = 3 \) is the separation of the Asian population from the Caucasian and Hispanic populations. This pattern is expected as human populations have their geographic origins in Africa, with dispersal first east through the Asian continent and then again later west through the European continent [46]. In addition, both the Asian and European populations are thought to have undergone population bottlenecks in their life histories [47] that will have led to variation in their allele frequencies. The grouping of Caucasian and Hispanic populations when \( K = 3 \) can also be explained due to the admixed nature of the Hispanic population in America, which is derived from the influx of Europeans into the native population, and so shares recent common ancestry with the Caucasian population [48]. Differences between the Hispanic and Caucasian populations begin to emerge when \( K = 4 \) and 5 and is apparent for the 30-locus panel shown in Figure 1.

Calculating the predictive accuracy of population assignment was performed on the 984 individuals analysed under parameter 3 using data for \( K = 4 \) to enable assignment to the four known populations. While the STRUCTURE results provide most support for three clusters (although \( K = 5 \) for highest lnPD), the distribution of inferred ancestry scores for the 30-locus panel shows there is some clear pattern of differentiation between the four populations over four clusters (Figure 2). Threshold values, as shown for the African American population in Figure 2, were determined from these distributions and formed the four criteria for assignment to a population.

Thresholds set for each population for each kit were used to assess the predicative accuracy of the assignment test. Individuals satisfying all criteria were considered positive (true or false) and any single criteria not satisfied led to the individual being considered negative (true or false). The sensitivity and specificity for the test was calculated based on the number of true/false positive and true/false negative assignments across the 984 individuals (Table 2). For the 30-locus data set the thresholding mechanism described was able to correctly assign 99% of all individuals to the correct population. A high degree of accuracy (>96%) was also shown for the two commercial mega-plex panels and those with expanded core loci, suggesting that the approach detailed here is robust and repeatable across panels, with only small fluctuations in accuracy.

Using the same thresholds and assignment criteria described above, the 40 'unknown' individuals removed from the 'training' dataset were assigned to each of the populations (Figure 3). The sensitivity of the assignment was > 90% for all the STR panels tested and with only 8 false negatives observed; three for 30-locus panel, 1 for PowerPlex Fusion, 1 for GlobalFiler, 2 for NGM Select, 1 for PowerPlex 16. Specificity was greater than 99% with only 2 false positives observed; 1 for PowerPlex Fusion, 1 for GlobalFiler. These observed levels align closely with the predicted values based on the training data. The few samples that were not correctly assigned were outliers that did not

**Figure 1.** Stacked bar graphs for 1024 individuals for 30-locus panel showing inferred ancestry scores for population genetic clusters identified by different colours. IAS1 = Orange; IAS2 = Light Blue; IAS3 = Green; IAS4 = Purple; IAS5 = Pink. Population sample data grouped showing African American (AA), Caucasian (Ca), Hispanic (His) and Asian (As). Parameter 2 shows better differentiation between the genetic clusters and was used for setting threshold criteria for population assignment.

**Figure 2.** Box and Whisker plots showing distribution of Inferred Ancestry Scores (IAS) to each of the four genetic clusters identified by STRUCTURE for the 30-locus panel. Where there is overlap in one distribution there is usually clear separation at another which allows the setting of a contingent threshold for each population for each cluster. The thresholds and examples values for the AA population are shown in red.
The time [53]. The results of this ethnicity bias study underlines why eye
individuals found that the correct ethnicity was assigned only 44.2% of
between 'Anglo' (non-Hispanic white), 'black', and Mexican American
[52]. The first study to assess the accuracy of cross-racial identification
correct identification of race when the individual is of the same race
as 'own race/ethnicity bias' which results in a larger percentage of
providing eyewitness testimony's can adhere to a phenomenon known
to conscious and unconscious bias [49-51]. For example, individuals
of suspect skin colour and race are commonly addressed through
origin using less reliable methods than that described here. Questions
to consider that enforcement agencies already attempt to infer ethnic
populations are represented in the database. As such it is the authors
opinion that this data be viewed as a test case for assignment rather
than a usable approach for the identification of African American,
Caucasian, Hispanic and Asian American individuals. Only with the
addition of more population groups will the utility of the approach be
truly understood. However, the high accuracy achieved would suggest
there is merit in exploring the described application in more detail.

The utility of a presumptive population assignment test such as that
described here is important to consider in relation to other approaches.
While there are better characterised panels of ancestry informative
markers available, they have not all been commercialised into a quality-
controlled product and are not used as part of routine analysis by most
forensic laboratories. Additionally, the adoption of MPS as an approach
is still some years away for many laboratories. It therefore seems
prudent to understand whether there is any further information that
can be derived from the use of existing STR panels. It is also important
to consider that enforcement agencies already attempt to infer ethnic
origin using less reliable methods than that described here. Questions
of suspect skin colour and race are commonly addressed through
witness statements and CCTV, both with low accuracy and subject
to conscious and unconscious bias [49-51]. For example, individuals
providing eyewitness testimony's can adhere to a phenomenon known as
'own race/ethnicity bias' which results in a larger percentage of
correct identification of race when the individual is of the same race
[52]. The first study to assess the accuracy of cross-racial identification
between 'Anglo' (non-Hispanic white), 'black', and Mexican American
individuals found that the correct ethnicity was assigned only 44.2% of
the time [53]. The results of this ethnicity bias study underlines why eye
witness testimonies are not viewed by jurors as highly reliable evidence
[54], but also recognises there is a balance to be struck between accuracy
and evidential strength. Novel genetic assays and tools developed for
forensic science are typically expected to be extremely accurate, an
expectation derived from the high discriminatory power observed when
performing DNA profiling. However, the development and adoption of
presumptive DNA tests for 'rapid-intelligence' [55,56] suggest that the
reality is likely to be more nuanced. As such, the authors feel that there
may be some merit to the consideration of a 'presumptive' population
assignment approach similar to that described here. Other non-forensic
applications of this approach may include the inference of ancestry for
supporting biogeographic research [1]. Given that STR data for human
populations are routinely collected and published online it seems likely
that the approach described here can be initially used to support wider
research into human life history.

Summary
This preliminary study demonstrates that population assignment
is possible using expanded and mega-plex STR panel with >90%
accuracy. This level of accuracy is in the range of a presumptive test
and the analysis can be performed using existing STR data collected
as part of routine criminal casework. The use of presumptive tests
to inform investigative leads is common in forensic laboratories and
enforcement authorities understand the limitations of such tests. As
such it is considered that this approach may provide some useful
insight into the ethnic origin of unknown individuals based on their
crime scene profile. Before adoption, further research should look to
include a greater range of population data to understand the true utility
of this approach and to assess the reproducibility of the approach. To
facilitate this, the authors have included the relevant allele coding in
the supplemental material for other interested groups to perform their
own analysis. Once optimised the approach can be compared to other
existing population assignment methods described in the literature to
compare performance.

Acknowledgements
The authors would like to thank the reviewers whose advice has helped
shape this version of the manuscript. Further thanks go to Stephanie
Johnson, Candice Carabini and Robert Measor who were involved in the
evolution of this paper.
References

1. Elhai E, Tatarinova T, Chebotarev D, Piras IS, Caló CM, et al. (2014) Geographic population structure analysis of worldwide human populations infers their biogeographical origins. Nat Commun 29: 3513. [Crossref]

2. Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, et al. (2002) Genetic structure of human populations. Science 298: 2381-2385. [Crossref]

3. Ekins JE, Ekins JB, Layton L, Hutchison LA, Myres NM, et al. (2006) Inference of ancestry: Constructing hierarchical reference populations and assigning unknown individuals. Hum genomics 2: 212. [Crossref]

4. Hwa HL, Lin CP, Huang TY, Kuo PH, Hsieh WH, et al. (2017) A panel of 130 autosomal single-nucleotide polymorphisms for ancestry assignment in five Asian populations and in Caucasians. Forensic Sci Med Pathol 13: 177-187. [Crossref]

5. Egeland T, Bøvelstad HM, Storvik GO, Salas A (2004) Inference the most likely geographical origin of mtDNA sequence profiles. Ann Hum Genet 68: 461-473. [Crossref]

6. Jorde LB, Watkins WS, Bamshad MJ, Dixon ME, Ricker CE, et al. (2000) The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. Am J Hum Genet 66: 979-988. [Crossref]

7. Lowe AL, Urquhart A, Foreman LA, Evett IW (2001) Inferring ethnic origin by means of an STR profile. Forensic Sci Int 119: 17-22. [Crossref]

8. Rosenberg NA, Li LM, Ward R, Pritchard JK (2003) Informativeness of genetic markers for inference of ancestry. Am J Hum Genet 73: 1402-1422. [Crossref]

9. Phillips C, Salas A, Sanchez JJ, Fondevila M, Gomez-Tato A, et al. (2007) Infering ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. Forensic Sci Int Genet 1: 273-280. [Crossref]

10. Zeng X, Warshauer DH, King JL, Churchill JD, Chakraborty R, et al. (2016) Empirical testing of a 23-AIMS panel of SNPs for ancestry evaluations in four major US populations. Int J Leg Med 130: 891-896. [Crossref]

11. Zeng X, King JL, Stojiljovar M, Warshauer DH, LaRue BL, et al. (2015) High sensitivity multiplex short tandem repeat loci analyses with massively parallel sequencing. Forensic Sci Int Genet 14: 38-47. [Crossref]

12. Xavier C, Parson W (2017) Evaluation of the Illumina ForenSeq™ DNA Signature Prep Kit—MPS forensic application for the MiSeq FGx™ benchtop sequencer. Forensic Sci Int Genet 28: 188-194. [Crossref]

13. Scudder N, McNevin D, Kelty SF, Walsh SJ, Robertson J (2017) Massively parallel sequencing and the emergence of forensic genomics: Defining the policy and legal issues for law enforcement. Sci Justice. 58: 153-158. [Crossref]

14. Butler JM (2015) The future of forensic DNA analysis. Philos Trans R Soc Lond B Biol Sci 370. [Crossref]

15. Cotton EA, Allopoe RF, Guest JL, Frazier RR, Koumi P, et al. (2000) Validation of the AmpFlSTR® NGM Select™ PCR Amplification Kit: a next-generation STR multiplex with the SE33 locus. Forensic Sci Int Genet 4: 359-361. [Crossref]

16. Hares DR (2012) Expanding the CODIS core loci in the United States. Forensic Sci Int Genet 16: 38-47. [Crossref]

17. Altman DG, Bland JM (1994) Diagnostic tests. 1: Sensitivity and specificity. BMJ 308: 145-147. [Crossref]

18. Bryc K, Velez C, Karafet T, Moreno-Estrada A, Reynolds A, et al. (2010) Genome-wide surveys of ancestry in diverse human populations. Proc Natl Acad Sci U S A 107: 9166-9170. [Crossref]

19. Li H, Durbin R (2011) Inference of human population history from individual whole-genome sequences. Nature 478: 493-496. [Crossref]

20. Gallagnier R, Fabbrini E, Mastrogiuseppe L, Randi E (2013) Who is who? Identification of livestock predators using forensic genetic approaches. Forensic Sci Int Genet 7: 397-404. [Crossref]

21. Primmer CR, Koskinen MT, Pironen J (2000) The one that did not get away: individual assignment using microsatellite data detects a case of fishing competition fraud. Proc Biol Sci 267: 1699-1704. [Crossref]

22. Frantz AC, Pourtsis JT, Heuertz M, Shlewy L, Flmand MC, et al. (2006) Genetic structure and assignment tests demonstrate illegal translocation of red deer (Cervus elaphus) into a continuous population. Molecular Ecology. 15: 3191-3203. [Crossref]

23. Paetzel D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. Mol Ecol 4: 347-354. [Crossref]

24. Hill CR, Duweer DL, Kline MC, Coble MD, Butler JM (2013) US population data for 29 autosomal STR loci. Forensic Sci Int Genet 7: e82-e83. [Crossref]

25. Hill CR, Butler JM, Coble MD (2012) Variability of new STR loci and kits in US population groups. Profiles in DNA. Available at http://www.promega.com/resources/articles/profiles-in-dna-2012 variabilty-of-new-STR-loci-and-kits-in-us-population-groups/. [Crossref]

26. Elhai E, Tatarinova T, Chebotarev D, Piras IS, Caló CM, et al. (2014) Geographic population structure analysis of worldwide human populations infers their biogeographical origins. Nat commun 5: 3513. [Crossref]

27. Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorrell JC, et al. (2017) The K-2 comanul. Mol Ecol 475: 493-496. [Crossref]

28. Daves DG, Bland JM (1994) Diagnostic tests 1: Sensitivity and specificity. BMJ 308: 1552. [Crossref]

29. Silva NM, Pereira L, Poloni ES, Currat M (2012) Human neutral genetic variation and forensic STR data. PLoS One 7: e603-5. [Crossref]

30. Slatkin M (1985) Rare alleles as indicators of genetic flow. Evolution 39: 53-65. [Crossref]

31. Hellman JR, Poloni ES, Currat M, Poncelet M, Lyonhein J (2006) Why did modern human populations disperse from Africa ca. 60,000 years ago? A new model. Proc Natl Acad Sci U S A 103: 9381-9386. [Crossref]

32. Li H, Durbin R (2011) Inference of human population history from individual whole-genome sequences. Nature 475: 493-496. [Crossref]

33. Brys K, Velcz C, Karafet T, Moreno-Estrada A, Reynolds A, et al. (2010) Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. Proc Natl Acad Sci U S A 11: 8954-8961. [Crossref]
49. Meissner CA, Brigham JC (2001) Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychol Pub Poli and Law* 7: 3-35.

50. Wells GL, Olson EA (2001) The other-race effect in eyewitness identification: What do we do about it? *Psychol Pub Poli and Law* 7: 230-402.

51. Brigham JC, Maass A, Snyder LD, Spaulding K (1982) Accuracy of eyewitness identification in a field setting. *Journal of Personality and Social Psychology* 42: 673.

52. Wilson J, Hugenberg K, Bernstein M (2013) The Cross-Race Effect and Eyewitness Identification: How to Improve Recognition and Reduce Decision Errors in Eyewitness Situations. *Social Issues and Policy Review* 7: 83-113.

53. Platz S, Hosch H (1988) Cross-Racial/Ethnic Eyewitness Identification: A Field Study. *Journal of Applied Social Psychology* 18: 972-984.

54. Wixted JT, Wells GL (2017) The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychol Sci Public Interest* 18: 10-65. [Crossref]

55. Dawnay N, Stafford-Allen B, Moore D, Blackman S, Rendell P, et al. (2014) Developmental Validation of the ParaDNA® Screening System—A presumptive test for the detection of DNA on forensic evidence items. *Forensic Sci Int Genet* 11: 73-9. [Crossref]

56. Blackman S, Dawnay N, Ball G, Stafford-Allen B, Tribble N, et al. (2015) Developmental validation of the ParaDNA® Intelligence System—A novel approach to DNA profiling. *Forensic Sci Int Genet* 17: 137-48. [Crossref]

Copyright: ©2018 Dawnay L. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.