Available online at www.sciencedirect.com
SCIENCE DIRECT®
Analytical Biochemistry 325 (2004) 171–173
ANALYTICAL BIOCHEMISTRY
www.elsevier.com/locate/yabio

Notes & Tips

Resolution of mixed human DNA samples using mitochondrial DNA sequence variants

Jerilyn A. Walker, Randall K. Garber, Dale J. Hedges, Gail E. Kilroy, Jinchuan Xing, and Mark A. Batzer*

Department of Biological Sciences, Biological Computation and Visualization Center, Louisiana State University, 202 Life Sciences Building, Baton Rouge, LA 70803, USA

Received 15 September 2003

Forensic casework samples routinely contain DNA from multiple contributors [1], posing a challenge to investigators attempting to resolve the components of complex DNA mixtures. Analysis of human mitochondrial DNA (mtDNA) is commonly used in forensic investigations to match evidentiary samples to potential suspects [2,3]. Unique DNA sequences from the hypervariable sequence (HVS)[1] region of the mitochondrial genome (nucleotides 16,024–16,576) from different samples can be compared with known samples to determine if there is a "match." However, this comparison strategy becomes exponentially complicated with the presence of additional contributors within human DNA samples. Recently, a technique based on denaturing high-performance liquid chromatography was reported for the rapid screening of mtDNA for resolution of mtDNA mixtures and the determination of the number of contributors [4]. Using this approach, "identity versus nonidentity" was accurately determined in less than 7 min per sample for 106 pairwise comparisons. Although this approach demonstrates the ability to detect multiple sequences in a mixed sample, and subsequent pairwise comparisons can be used to identify potential contributors, the sequence data themselves are not obtained and it relies on specialized instrumentation that is not routinely available in most forensic laboratories.

In addition to forensic applications, sequence analyses of the human mtDNA hypervariable control region have been performed by many investigators as a means of studying the demographic expansion and migration patterns of various population groups [5,6]. As a result, phylogenetic and demographic associated mtDNA HVS variants have previously been identified [7–11]. Here we present a mtDNA HVS-based approach designed to determine the number of contributors within a mixed human DNA sample.

Sequence-specific oligonucleotide primers were used for polymerase chain reaction (PCR) amplification of a portion of the human mtDNA hypervariable control region (HV/F16144: 5′-TGACCACCTGTAGTACA TAA-3′; HV/R16410: 5′-GAGGATGGTGGTCAAG GGAC-3′) from three individuals of different origin. DNA samples from one Japanese, one Southeast Asian, and one South American individual were used in these experiments. All were PCR amplified individually and in combination from a mixed template constructed with equal concentrations of DNA from each individual. Amplicons were resolved on a 2% agarose gel (Fig. 1A), excised, and gel purified using the Wizard SV gel and PCR cleanup system (Promega, Inc.). Gel purified products were cloned using a TOPO-TA cloning kit (Invitrogen) and sequenced using an ABI Prism 3100 genetic analyzer and BigDye v3.0 (Applied Biosystems, Inc.). Two separate clones were sequenced from each individual DNA sample to obtain the reference sequence from each of the known samples. Thirty-eight clones derived from the mixed template amplicons containing all three individuals were sequenced. On review of each electropherogram, sequences were stored and later aligned using MegAlign with the ClustalW algorithm and the default settings (DNAstar Version 5.0 for Windows) followed by manual refinement. Unique mtDNA HVSs for each of the individuals were detected from the mixed DNA sample (Fig. 1B). Sequence data from 29 clones matched the Japanese individual, 8

* Corresponding author. Fax: +1-225-578-7113.
E-mail address: mbatzer@lsu.edu (M.A. Batzer).

[1] *Abbreviations used:* HVS, hypervariable sequence.

0003-2697/$ - see front matter
doi:10.1016/j.ab.2003.10.003

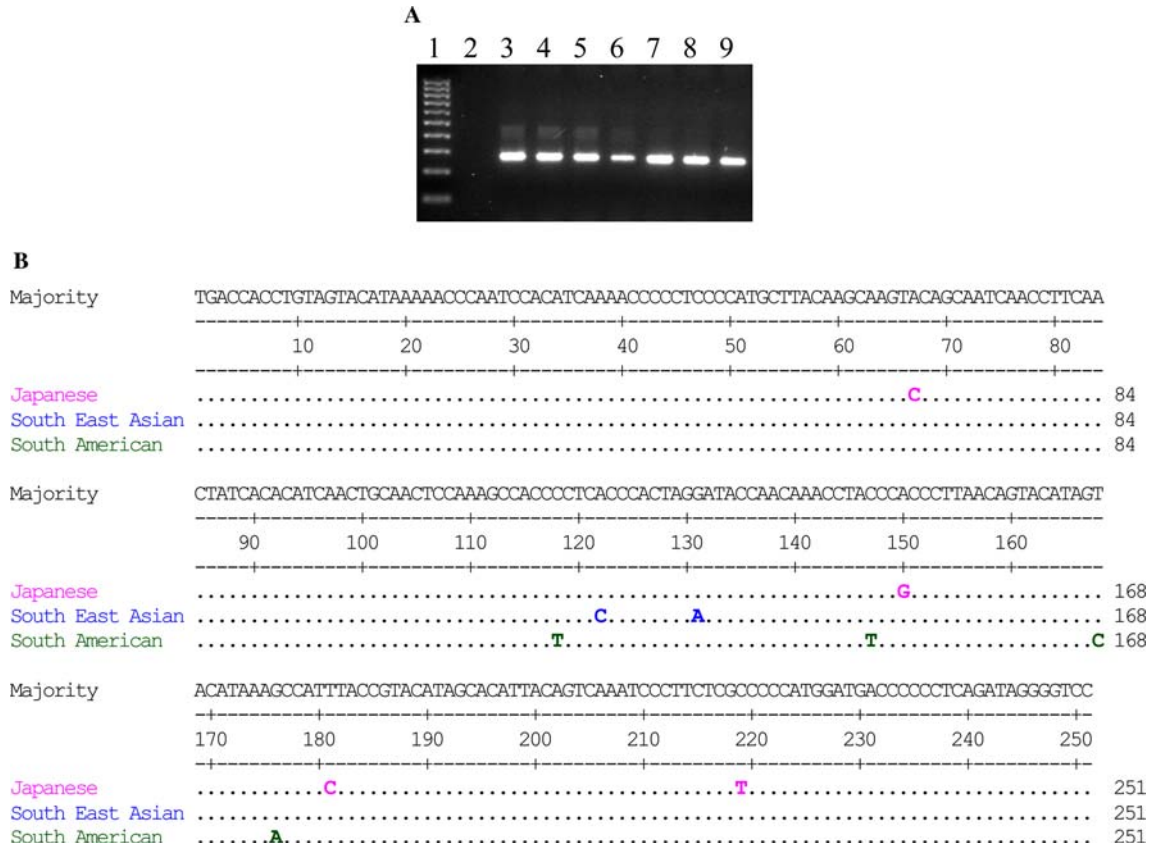

Fig. 1. Amplification of the mtDNA HVS and resolution of the number of human contributors. (A) An agarose gel chromatograph of PCR amplicons from the mtDNA HVS region. Lane 1: 100-bp DNA ladder; Lane 2: negative control; Lanes 3–6: mixture of DNAs from individuals in lanes 7–9; Lane 7: DNA from a Japanese individual; Lane 8: DNA from a Southeast Asian individual; Lane 9: DNA from a South American individual. (B) mtDNA HVS alignment of multiple contributors in a mixed human DNA sample. Bases matching the majority consensus sequence are represented by dashes and unique nucleotide substitutions are shown in boldface. The sequence matching the Japanese individual and the sequence matching the South American individual each contained four unique nucleotide substitutions different from the majority consensus sequence. The sequence matching the Southeast Asian individual contained two diagnostic substitutions.

clones matched the South American sequence, while only one matched the Southeast Asian. In this experiment clones were sequenced until the unique sequences of all three known contributors were detected. The application of this approach to the analysis of complex DNA mixtures with an unknown number of contributors requires expanded sequencing of different clones until no new variants are detected. Assuming equal representation of individuals in the mixture, the number of sequenced clones, x, required to detect n contributors at a 95% frequency can be estimated by $[1 - ((n-1)/n)^x \geqslant 0.95]$. For a mixture of three sequences, this is approximately 8 clones. However, complicating factors in forensic settings will result in individual cases deviating significantly from the ideal situation. As a general guideline, we recommend sequencing clones in sets of 10 until no new sequence variants are detected.

We have demonstrated the utility of this sequence-based approach for determination of the number of contributors in a complex DNA mixture. By retaining the unique composite sequence data for each individual, inferences about the number of contributors and their geographic origin can be made by comparison with existing mtDNA sequence databases [5,7,12]. The more individuals from each population that have been characterized and are present in the database, the more accurate the inferences about sample geographic affiliation will be.

Potential pitfalls using this approach are the possibility of over-estimating the number of contributors due to the occurrence of heteroplasmic mitochondria, and the possibility of under-estimating the number due to individuals with identical haplogroups. Although heteroplasmy is believed to be rare and difficult to account for computationally, the probability that a mixture could contain individuals with identical haplogroups can be estimated. The probability of this occurring is contingent on both the actual number of individuals in the mixture and the frequencies of the haplogroups in the population. In general, the probability of two individuals in a mixture having the same sequence can be represented by $[1 - (1 - \sum(X_i)^2)^M]$, where X_i represents

the frequency of each haplogroup in the population and M is the number of possible pairings in the sample of interest. Pairings increase with the number of contributors, n, in the following manner: $n!/2!(n-2)!$. Therefore, the probability of two individuals in a mixture having the same mtDNA sequence using this approach is low given a sample with few contributors from diverse haplogroups, and increases with haplogroup isolation and increased number of contributors.

Acknowledgments

This research was supported by the Louisiana Board of Regents Governor's Biotechnology Initiative GBI (2002-005) to M.A.B.

References

[1] C. Ladd, H.C. Lee, N. Yang, F.R. Bieber, Interpretation of complex forensic DNA mixtures, Croat. Med. J. 42 (2001) 244–246.
[2] M.R. Wilson, J.A. DiZinno, D. Polanskey, J. Replogle, B. Budowle, Validation of mitochondrial DNA sequencing for forensic casework analysis, Int. J. Legal Med. 108 (1995) 68–74.
[3] M.M. Holland, T.J. Parsons, Mitochondrial DNA sequence analysis-validation and use for forensic casework, Forensic Sci. Rev. 11 (1999) 21–50.
[4] G.S. LaBerge, R.J. Shelton, P.B. Danielson, Forensic utility of mitochondrial DNA analysis based on denaturing high-performance liquid chromatography, Croat. Med. J. 44 (2003) 281–288.
[5] M. Stoneking, D. Hedgecock, R.G. Higuchi, L. Vigilant, H.A. Erlich, Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes, Am. J. Hum. Genet. 48 (1991) 370–382.
[6] B. Budowle, M.R. Wilson, J.A. DiZinno, C. Stauffer, M.A. Fasano, M.M. Holland, K.L. Monson, Mitochondrial DNA regions HVI and HVII population data, Forensic Sci. Int. 103 (1999) 23–35.
[7] L.B. Jorde, W.S. Watkins, M.J. Bamshad, M.E. Dixon, C.E. Ricker, M.T. Seielstad, M.A. Batzer, The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data, Am. J. Hum. Genet. 66 (2000) 979–988.
[8] T. Melton, S. Clifford, M. Kayser, I. Nasidze, M. Batzer, M. Stoneking, Diversity and heterogeneity in mitochondrial DNA of North American populations, J. Forensic Sci. 46 (2001) 46–52.
[9] T. Melton, M. Stoneking, Extent of heterogeneity in mitochondrial DNA of ethnic Asian populations, J. Forensic Sci. 41 (1996) 591–602.
[10] T. Melton, C. Ginther, G. Sensabaugh, H. Soodyall, M. Stoneking, Extent of heterogeneity in mitochondrial DNA of sub-Saharan African populations, J. Forensic Sci. 42 (1997) 582–592.
[11] T. Melton, M. Wilson, M. Batzer, M. Stoneking, Extent of heterogeneity in mitochondrial DNA of European populations, J. Forensic Sci. 42 (1997) 437–446.
[12] C. Romualdi, D. Balding, I.S. Nasidze, G. Risch, M. Robichaux, S.T. Sherry, M. Stoneking, M.A. Batzer, G. Barbujani, Patterns of human diversity, within and among continents, inferred from biallelic DNA polymorphisms, Genome Res. 12 (2002) 602–612.