Martina Wiesbauer[1]
David Meierhofer[1]
Johannes A. Mayr[1]
Wolfgang Sperl[1]
Bernhard Paulweber[2]
Barbara Kofler[1]

[1]Department of Paediatrics, Paracelsus Private Medical University Salzburg, Salzburg, Austria
[2]Department of Internal Medicine I, Paracelsus Private Medical University Salzburg, Salzburg, Austria

Received February 17, 2006
Revised June 8, 2006
Accepted June 23, 2006

Short Communication

Multiplex primer extension analysis for rapid detection of major European mitochondrial haplogroups

The evolution of the human mitochondrial genome is reflected in the existence of ethnically distinct lineages or haplogroups. Alterations of mitochondrial DNA (mtDNA) have been instrumental in studies of human phylogeny, in population genetics, and in molecular medicine to link pathological mutations to a variety of human diseases of complex etiology. For each of these applications, rapid and cost effective assays for mtDNA haplogrouping are invaluable. Here we describe a hierarchical system for mtDNA haplogrouping that combines multiplex PCR amplifications, multiplex single-base primer extensions, and CE for analyzing ten haplogroup-diagnostic mitochondrial single nucleotide polymorphisms. Using this rapid and cost-effective mtDNA genotyping method, we were able to show that within a large, randomly selected cohort of healthy Austrians (n = 1172), mtDNAs could be assigned to all nine major European haplogroups. Forty-four percent belonged to haplogroup H, the most frequent haplogroup in European Caucasian populations. The other major haplogroups identified were U (15.4%), J (11.8%), T (8.2%) and K (5.1%). The frequencies of haplogroups in Austria is within the range observed for other European countries. Our method may be suitable for mitochondrial genotyping of samples from large-scale epidemiology studies and for identifying markers of genetic susceptibility.

Keywords: European Caucasian / Haplogroup / mtDNA / Multiplex primer extension
DOI 10.1002/elps.200600086

Studies of mitochondrial DNAs (mtDNAs) from a wide range of human populations have revealed a number of stable polymorphic sites. These define related groups of mtDNAs called haplogroups. Nine European, seven Asian (including Native American), and three African mtDNA haplogroups have been identified [1–4].

Interestingly, mtDNA haplogroups show striking differences in the ratio of synonymous and nonsynonymous nucleotide changes between geographic regions at different latitudes [5]. Evidence has accumulated that different human mtDNA lineages are functionally different, and that such differences may have health implications. For example, positive and negative associations of mtDNA haplogroups have been found in asthenozoospermia [6], multiple sclerosis [7–10], optical stroke [11], survival after sepsis [12] and longevity [13], suggesting that mtDNA haplogroups, at first thought to be phenotypically indistinguishable, may be important determinants of human disorders of complex etiology [14]. However, mtDNA haplogroup-association analyses suffer in several regards. Many of the studies used only subpopulations from geographical areas and had small numbers in each cohort. To be meaningful, such studies require patient and control cohorts numbering in the several hundreds, if not thousands [14]. For this purpose a reliable, rapid and economic assay for mitochondrial haplogrouping is desirable.

RFLP studies of a wide range of human populations have revealed a number of stable polymorphic sites in mtDNA. These sites define related groups of mtDNAs called haplogroups. Accordingly, most haplogroup studies have relied on RFLP analysis, using up to 14 endonucleases to determine mtDNA haplogroups [1, 15, 16]. Other mtDNA genotyping methods employ TaqMan assays or melting curve

Correspondence: Professor Barbara Kofler, Department of Pediatrics, Paracelsus Private Medical University Salzburg, Muellner-Hauptstrasse 48, A-5020 Salzburg, Austria
E-mail: b.kofler@salk.at
Fax: +43-662-4482-4765

Abbreviations: mtDNA, mitochondrial DNA; **mtSNP**, mitochondrial single nucleotide polymorphism; **SNP**, single nucleotide polymorphism

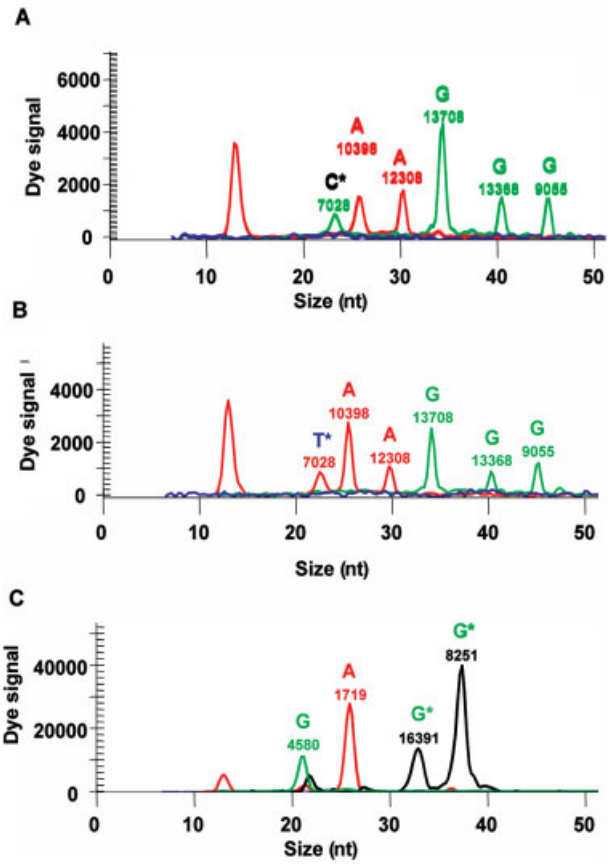

Figure 1. Primer extension assays for two samples that exhibit different haplogroups. (A) Multiplex assay A of a sample with haplogroup H. (B) Multiplex assay A of a sample, which needs to be further identified in multiplex B. (C) Multiplex assay B of the sample in (B) now exhibits a SNP combination diagnostic for haplogroup X. Asterisks indicate bases complementary to the ones indicated in the chromatogram due to the use of reverse primer. The first peak corresponds to the internal size standard of 13 bp.

analyses with allele-specific T_m detection probes [20, 21]. However, both these assays require a separate PCR amplification for each single nucoeltide polymorphism (SNP) site comprising the haplogroup, using costly chemicals.

More recently, primer extension assays (Fig. 1) have been employed to examine forensically informative mitochondrial (mt)SNPs [17–19]. The technique of single-base primer extension entails the annealing of an ordinary DNA oligonucleotide to a DNA template such that the 3'-end of the primer falls one base short of the SNP site present on the template. In the presence of all four dideoxynucleotides, each one labelled with a different fluorescent dye, the DNA polymerase will add to the primer whichever dye terminator is complementary to the base on the template that defines the SNP site.

The aim of the present study was to establish a standardized multiplex primer extension assay for rapid mtDNA based haplogrouping and to determine the haplogroup distribution in a population from central Austria.

Specific primers were designed to amplify the mtDNA regions containing the polymorphic sites that characterize each of the major European haplogroups (H, I, J, K, T, U, V, W, X) in two multiplex PCRs, A and B (Tables 1 and 2). MtSNPs for European haplogroup typing were selected as previously reported [20, 21]. Assay A was designed to determine the most common European haplogroups, H, J, K, T, U, and included primers to amplify five PCR fragments encompassing six informative mtSNPs (Tables 1 and 2). For haplogroups not identified by assay A, a second assay, B, covering another four mtSNPs was used to allow identification of haplogroups V, W, I, X. Primers for PCR amplification and primer extension analysis are listed in Table 1. PCR mixtures contained 10 × PCR buffer B (Solis Biodyne), 180 pM of each primer, 140 μM of each deoxynucleotide triphosphate (Invitrogen) and 0.064 U of Hot Fire Polymerase (Solis Biodyne) in a volume of 30 μL. Pre-aliquoted reaction mixes were stored at −20°C. After thawing, 60 ng of genomic DNA was added and all fragments were amplified under the following conditions: 95°C for 15 min, 43 cycles at 95°C for 60 s, 54°C for 60 s and 72°C for 90 s, and a final extension at 72°C for 5 min. Freezing of preformulated and aliquoted PCR mixes did not affect the performance of the amplification or subsequent primer extension analysis, but saved time and enhanced reproducibility [22, 23]. Reactions for the multiplex PCR of one individual can be prepared within seconds simply by adding the template DNA. For large-scale screening of samples, mixes can be supplied in convenient 96-well plates.

To remove primers and unincorporated deoxynucleotides, PCR products (10 μL) were treated with Quick Clean according to the manufacturer's protocol (Bioline). The purified PCR product was resuspended in 20 μL water and further treated with 0.34 U shrimp alkaline phosphatase (Promega) for 1 h at 37°C, followed by incubation at 80°C for 15 min to inactivate the enzyme.

Multiplex primer extension reactions were carried out in a total volume of 10 μL containing 2 μL primer extension mix (CEQ™ SNP primer extension kit, Beckman Coulter), 2 μL purified PCR product, and the primer mix as listed in Table 1. To enable electrophoretic resolution of the different primer extension products, poly-adenosine tails 5–25 nucleotides in length were included at the 5'-end of each SNP primer (see Table 1). The choice of orientation and concentration of the SNP primers in the multiplex assays was determined empirically. Thermal cycling conditions for single base extension reactions were performed with: 25 cycles of denaturation at 96°C for 10 s, annealing at

Table 1. Primer sequences used to amplify mtDNA fragments containing haplogroup specific SNPs and extension primers

Position	Sequence PCR primer	bp	Position[a)]	Sequence SNP primer	T_m	pmol
Multiplex A						
6 954– 7 491	GGTGGCCTGACTGGCATTG GTTGGCTTGAAACCAGCTTTGG	538	7 028 r	TATTGATAGGACATAGTGGAAGTG	57.8	0.5
8 854– 9 335	GCGGGCGCAGTGATTATAGG GAGGAGCGTTATGGAGTGGAAG	482	9 055 f	$(A)_{25}$CTACTCATGCACCTAATTGGAAGC	65.2	1.0
10 279–10 634	CCCTACCATGAGCCCTACAAAC TAAGAGGGAGTGGGTGTTGAGG	356	10 398 f	$(A)_5$TGACTACAAAAAGGATTAGACTGA	58.2	0.5
12 009–12 462	GCTCACTCACCCACCACAT GGATGCGACAATGGATTTTA	454	12 308 f	$(A)_{10}$CTATCCATTGGTCTTAGGCCCCAA	64.7	0.6
13 249–13 786	TCCACTTCAAGTCAACTAGGAC GGGGATTGTTGTTTGGAAGGG	537	13 708 f 13 368 f	$(A)_{15}$CTACTAAACCCCATTAAACGCCTG $(A)_{21}$GCCATACTATTTATGTGCTCCGG	64.2 64.8	0.69 2.0
Multiplex B						
1 485– 1 950	GCCCGTCACCCTCCTCAAG ACGGGTGTGCTCTTTTAGCTG	466	1 719 f	$(A)_5$ACCTTACTACCAGACAACCTTA	62.4	1.0
4 506– 5 004	CATCTTTGCAGGCACACTCATC GATTTTGCGTAGCTGGGTTTGG	499	4 580 f	CTGAGTAGGCCTAGAAATAAACAT	57.6	0.5
8 011– 8 560	AGTACTCCCGATTGAAGCCCC GGGCAATGAATGAAGCGAACAG	550	8 251 r	$(A)_{15}$GTGCTATAGGGTAAATACGGG	66.3	0.3
16 344– 276	CAGTCAAATCCCTTCTCGTCCC TCTGTGTGGAAAGTGGCTGTG	502	16 391 r	AG $(A)_{10}$GGAGGATGGTGGTCAAGGGA	67.7	0.5

a) The numbering of site positions follows the revised Cambridge reference sequence genebank no. J01415.1. r: primer in reverse orientation; f: primer in forward orientation.

50°C for 20 s and extension at 72°C for 50 s. After the primer extension reaction, excess fluorescence labelled dideoxynucleotides were digested by adding 0.34 U shrimp alkaline phosphatase (Promega) to 4 μL of the reaction product. Digestion was performed at 37°C for 1 h followed by enzyme inactivation at 80°C for 15 min. The primer extension products (1 μL) were mixed with 39 μL Sample Loading Solution (Beckman Coulter) and 0.125 μL internal standard (GenomeLab™ DNA standard kit 80, Beckman Coulter). After a denaturation step of 90°C for 60 s, samples were separated by CE on a Beckman Coulter CEQ™ 8000 Genetic Analysis System. The system is based on a capillary filled with urea in buffered polyacrylamide sieving matrix of 33 cm length and 75 μm internal diameter. The injection was for 30 s at a field strength of 2 kV and the separation electrical field strength was 6 kV. Separation time of primer extension products was 16 min. Using the CEQ software it is possible to assign the locus to the peaks and to call the genotypes automatically. The results of our multiplex analysis were confirmed by direct sequencing of a representative sample of each haplogroup.

Our study population consisted of 1172 unrelated Caucasians of the Salzburg region in central Austria. The cohort included 930 men between 40 and 55 years of age and 527 women between 50 and 65 years of age. The study was conducted according to the Austrian Gene Technology Act and was approved by the ethics committee of the medical association of Salzburg. All subjects gave written, informed consent before entering the study. Venous blood was collected in 5 mL EDTA coated tubes; genomic DNA was isolated with a Nucleospin Blood kit (Macherey-Nagel) and stored at 4°C.

Running multiplex A alone allowed 84% of the sample to be assigned to five haplogroups: H, J, K, T, U (Table 2). The remaining specimens were further analyzed by multiplex reaction B to identify haplogroups V, W, I, X (Table 2). All nine major European haplogroups were observed in our cohort. The haplogroup frequencies were 44% H (n = 511), 15.4% U (n = 181), 11.8% J (n = 138), 8.2% T (n = 96), 5.1% K (n = 60), 1.9% V (n = 23), 1.7% W (n = 20), 1.1% X (n = 13) and 1.1% I (n = 13). Failure of primer extension analysis of a successful PCR amplification occurred in 1% of cases. In addition, 10% of individuals did not fit into these nine predefined haplogroups and were classified as "others". A similar percentage (8.2%) of unclassified haplogroups was observed in a previous study using the same SNPs in a TaqMan Allelic Discrimination Assay [20]. If further typing of these samples is desired, additional mtSNPs could be included in multiplex assay B. A combination of our method and sequencing of the hypervariable regions of the D-loop should allow determination of all mtDNA haplogroup subtypes.

Table 2. MtDNA polymorphic sites used to determine major European haplogroups in multiplex analysis A and B[a)]

Multiplex A										Haplogroup
7028	9055	10398	12308	13708	13368	1719	4580	8251	16391	
C	G	A	A	G	G	–	–	–	–	H
T	G	G	A	A	G	–	–	–	–	J
T	G	A	G	G	G	–	–	–	–	U
T	G	A	A	G	A	–	–	–	–	T
T	A	G	G	G	G	–	–	–	–	K
T	G	G	A	G	G	–	–	–	–	Multiplex B
T	G	A	A	G	G	–	–	–	–	Multiplex B
						Multiplex B				
T	G	G	A	G	G	A	G	A	A	I
T	G	A	A	G	G	G	A	G	G	V
T	G	A	A	G	G	G	G	A	G	W
T	G	A	A	G	G	A	G	G	G	X

a) The numbering of site positions follows the revised Cambridge reference sequence genebank no. J01415.1.

A previous study also employed multiplex SNP assays for mtDNA haplogrouping for forensic applications, analyzing 16 SNPs in two reactions [18]. Our method is a rapid and cost effective alternative to the methods mentioned above, since 84% of the samples can be analyzed with one multiplex assay. We used the second multiplex assay only for the remaining 16% of samples that could not be assigned to a haplogroup by the first reaction. Sequencing becomes more cost effective. However, to determine a haplogroup for 84% of the samples, six fragments would have to be sequenced. Here only one reaction using in principle the same chemicals (same price) is necessary. There is an additional purification step of the PCR product in the present method necessary compared to a sequencing reaction, but this is compensated for by less primer extension reactions and subsequent separations on the sequencer. In addition, automatization of the analysis of several sequence reactions in parallel to determine a haplogroup is not available.

The distribution of haplogroups we observed falls within the range for European Caucasians reported in other studies [1]. However, it differs slightly from a previous study of 277 Caucasians from western Austria [18]. The differences in frequencies of haplogroups J, K, T, U may simply be a function of the different sample sizes (277 versus 1172). Alternatively, the differences in haplogroup frequencies may be real, possibly reflecting the geographic separation of the populations of western and central Austria imposed by the Alps.

With our multiplex primer extension assays detailed here we could readily assign major European haplogroups in the Austrian population. Our method will be useful for association studies aimed at identifying mtDNA genotypes that predispose individuals to various multifactorial diseases. These assays can be extended to other panels of mtSNPs for typing the haplogroups prevalent in different countries or continents.

This work was supported by a grant of the "Medizinische Forschungsgesellschaft Salzburg" to B. Kofler and of the "Kamillo Eisner Stiftung" to B. Paulweber and the "Vereinigung zur paediatrischen Forschung und Fortbildung Salzburg".

References

[1] Torroni, A., Huoponen, K., Francalacci, P., Petrozzi, M. *et al.*, *Genetics* 1996, *144*, 1835–1850.

[2] Torroni, A., Schurr, T. G., Yang, C. C., Szathmary, E. J. *et al.*, *Genetics* 1992, *130*, 153–162.

[3] Silva, W. A., Jr., Bonatto, S. L., Holanda, A. J., Ribeiro-Dos-Santos, A. K. *et al.*, *Am. J. Hum. Genet.* 2002, *71*, 187–192.

[4] Schurr, T. G., Wallace, D. C., *Hum. Biol.* 2002, *74*, 431–452.

[5] Mishmar, D., Ruiz-Pesini, E., Golik, P., Macaulay, V. *et al.*, *Proc. Natl. Acad. Sci. USA* 2003, *100*, 171–176.

[6] Ruiz-Pesini, E., Lapena, A. C., Diez-Sanchez, C., Perez-Martos, A. *et al.*, *Am. J. Hum. Genet.* 2000, *67*, 682–696.

[7] Kalman, B., Lublin, F. D., Alder, H., *Mult. Scler.* 1995, *1*, 32–36.

[8] Seyrantepe, V., Topaloglu, H., Simsek, E., Ozguc, M., Yordam, N., *Lancet* 1996, *347*, 695–696.

[9] Hofmann, S., Bezold, R., Jaksch, M., Kaufhold, P. *et al.*, *Mol. Cell. Biochem.* 1997, *174*, 209–213.

[10] Kalman, B., Li, S., Chatterjee, D., O'Connor, J. *et al.*, *Acta Neurol. Scand.* 1999, *99*, 16–25.

[11] Majamaa, K., Finnila, S., Turkka, J., Hassinen, I. E., *Lancet* 1998, *352*, 455–456.

[12] Baudouin, S. V., Saunders, D., Tiangyou, W., Elson, J. L. *et al.*, *Lancet* 2005, *366*, 2118–2121.

[13] De Benedictis, G., Rose, G., Carrieri, G., De Luca, M. *et al.*, *FASEB J.* 1999, *13*, 1532–1536.

[14] Herrnstadt, C., Howell, N., *Mitochondrion* 2004, *4*, 791–798.

[15] Santos, C., Montiel, R., Angles, N., Lima, M. *et al.*, *Hum. Biol.* 2004, *76*, 431–453.

[16] Huerta, C., Castro, M. G., Coto, E., Blazquez, M. *et al.*, *J. Neurol. Sci.* 2005, *236*, 49–54.

[17] Vallone, P. M., Just, R. S., Coble, M. D., Butler, J. M., Parsons, T. J., *Int. J. Legal Med.* 2004, *118*, 147–157.

[18] Brandstatter, A., Parsons, T. J., Parson, W., *Int. J. Legal Med.* 2003, *117*, 291–298.

[19] Quintans, B., Alvarez-Iglesias, V., Salas, A., Phillips, C. *et al.*, *Forensic Sci. Int.* 2004, *140*, 251–257.

[20] van der Walt, J. M., Nicodemus, K. K., Martin, E. R., Scott, W. K. *et al.*, *Am. J. Hum. Genet.* 2003, *72*, 804–811.

[21] Jiang, Y., Ellis, T., Greenlee, A. R., *Clin. Med. Res.* 2004, *2*, 99–106.

[22] Klausegger, A., Hell, M., Berger, A., Zinober, K. *et al.*, *J. Clin. Microbiol.* 1999, *37*, 464–466.

[23] Meierhofer, D., Mayr, J. A., Ebner, S., Sperl, W., Kofler, B., *Mitochondrion* 2005, *5*, 282–296.

