Letters to the Editor 1341

*LIT1* distinguish patients with Beckwith-Wiedemann syndrome with cancer and birth defects. Am J Hum Genet 70: 604–611

- Engel J, Smallwood A, Harper A, Higgins M, Oshimura M, Reik W, Schofield P, Maher E (2000) Epigenotype-phenotype correlations in Beckwith-Wiedemann syndrome. J Med Genet 37:921–926
- Fitzpatrick GV, Soloway PD, Higgins MJ (2002) Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. Nat Genet 32:426–431
- Gaston V, Le Bouc Y, Soupre V, Burglen L, Donadieu J, Oro H, Audry G, Vazquez MP, Gicquel C (2001) Analysis of the methylation status of the KCNQ1OT and H19 genes in leukocyte DNA for the diagnosis and prognosis of Beckwith-Wiedemann syndrome. Eur J Hum Genet 9:409–418
- Gaston V, Le Bouc Y, Soupre V, Vazquez MP, Gicquel C (2000) Assessment of p57(KIP2) gene mutation in Beckwith-Wiedemann syndrome. Horm Res 54:1–5
- Humpherys D, Eggan K, Akutsu H, Hochedlinger K, Rideout WM 3rd, Biniszkiewicz D, Yanagimachi R, Jaenisch R (2001) Epigenetic instability in ES cells and cloned mice. Science 293:95–97
- Lee M, Debaun M, Mitsuya K, Galonek H, Branderburg S, Oshimura M, Feinberg A (1999) Loss of imprinting of a paternally expressed transcript, with antisense orientation to KvLQT1, occurs frequently in Beckwith-Wiedemann syndrome and is independent of insulin-like growth factor II imprinting. Proc Natl Acad Sci USA 96:5203–5208
- Li E (2002) Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet 3:662–673
- Maher ER, Brueton LA, Bowdin SC, Luharia A, Cooper W, Cole TR, Macdonald F, Sampson JR, Barratt CL, Reik W, Hawkins MM (2003) Beckwith-Wiedemann syndrome and assisted reproduction technology (ART). J Med Genet 40: 62–64
- Maher ER, Reik W (2000) Beckwith-Wiedemann syndrome: imprinting in clusters revisited. J Clin Invest 105:247–252
- Manning M, Lissens W, Bonduelle M, Camus M, De Rijcke M, Liebaers I, Van Steirteghem A (2000) Study of DNAmethylation patterns at chromosome 15q11-q13 in children born after ICSI reveals no imprinting defects. Mol Hum Reprod 6:1049–1053
- Mitsuya K, Meguro M, Lee M, Katoh M, Schulz T, Kugoh H, Yoshida M, Niikawa N, Feinberg A, Oshimura M (1999) LIT1, an imprinted antisense RNA in the human KvLQT1 locus identified by screening for differentially expressed transcripts using monochromosomal hybrids. Hum Mol Genet 8:1209–1217
- Olivennes F, Mannaerts B, Struijs M, Bonduelle M, Devroey P (2001) Perinatal outcome of pregnancy after GnRH antagonist (ganirelix) treatment during ovarian stimulation for conventional IVF or ICSI: a preliminary report. Hum Reprod 16: 1588–1591
- Ørstavik KH, Eiklid K, van der Hagen CB, Spetalen S, Kierulf K, Skjeldal O, Buiting K (2003) Another case of imprinting defect in a girl with Angelman syndrome who was conceived by intracytoplasmic sperm injection. Am J Hum Genet 72: 218–219
- Reik W, Walter J (2001) Genomic imprinting: parental influence on the genome. Nat Rev Genet 2:21–32
- Rideout WM 3rd, Eggan K, Jaenisch R (2001) Nuclear cloning and epigenetic reprogramming of the genome. Science 293: 1093–1098
- Smilinich NJ, Day CD, Fitzpatrick GV, Caldwell GM, Lossie AC, Cooper PR, Smallwood AC, Joyce JA, Schofield PN, Reik W, Nicholls RD, Weksberg R, Driscoll DJ, Maher ER, Shows TB, Higgins MJ (1999) A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. Proc Natl Acad Sci USA 96:8064–8069
- Weksberg R, Nishikawa J, Caluseriu O, Fei YL, Shuman C, Wei C, Steele L, Cameron J, Smith A, Ambus I, Li M, Ray PN, Sadowski P, Squire J (2001) Tumor development in the Beckwith-Wiedemann syndrome is associated with a variety of constitutional molecular 11p15 alterations including imprinting defects of KCNQ1OT1. Hum Mol Genet 10:2989– 3000
- Young L, Fernandes K, McEvoy T, Butterwith S, Gutierrez C, Carolan C, Broadbent P, Robinson J, Wilmut I, Sinclair K (2001) Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture. Nat Genet 27:153– 154
- Young LE, Sinclair KD, Wilmut I (1998) Large offspring syndrome in cattle and sheep. Rev Reprod 3:155–163

Address for correspondence and reprints: Dr. Christine Gicquel, Laboratoire d'Explorations Fonctionnelles Endocriniennes, Hôpital Trousseau, 26 Avenue Arnold Netter, 75012 Paris, France. E-mail: christine.gicquel@trs.ap-hop-paris.fr 2003 by The American Society of Human Genetics. All rights reserved. 0002-9297/2003/7205-0029\$15.00

*Am. J. Hum. Genet. 72:1341–1346, 2003*

## **To Trust or Not to Trust an Idiosyncratic Mitochondrial Data Set**

## *To the Editor:*

In a recent report, Silva et al. (2002) provided partial (8.8 kb) information on the mtDNA coding region (within the region 7148–15946, in the numbering of the Cambridge reference sequence [CRS]; Anderson et al. [1981]) in 40 individuals from Brazil. On the basis of the similarity in nucleotide diversity and age estimates of the four founder haplogroups A, B, C, and D, they claimed to have added new evidence for a single early entry of the founder populations into America. However, a site-by-site audit of the data reveals that their sequences are not of high enough quality to justify such statements. The authors failed to realize that a large number of mutations associated with basal branches of the worldwide mtDNA phylogeny (Finnilä et al. 2001; Maca-Meyer et al. 2001; Torroni et al. 2001; Derbeneva et al. 2002; Herrnstadt et al. 2002; Kivisild et al. 2002) were not correctly scored in their data set.



Sequence Variation in 40 Samples Reported by Silva et al. (2002) **Sequence Variation in 40 Samples Reported by Silva et al. (2002) Table 1**



underlined.<br><sup>a</sup> All bear 14766 in addition.<br><sup>b</sup> Basal polymorphisms that were undetected or omitted by Silva et al. (2002), including 11719 and the two rare mutations (8860 and 15326) in the CRS.

In the case of the hypervariable segments of the mtDNA control region, Bandelt et al. (2001, 2002) have highlighted lab-specific idiosyncrasies through comparative phylogenetic analysis. For the coding region, the task of identifying anomalies and reconstructing their potential causes is somewhat easier because the vast majority of sites there do not appear to undergo frequent mutations. The coding region well supports a basal nesting of (monophyletic) haplogroups, many of which had already been identified through RFLP analysis and sequencing of the hypervariable segments (Richards and Macaulay 2001). For example, the basal division of Eurasian mtDNAs into macrohaplogroups M and N is amazingly clear cut. The Eurasian mtDNA phylogeny that emerges from the phylogenetic analysis of the complete mtDNA database is detailed (for east Asia) in figure 1 of Kivisild et al. (2002), which attempts a reconstruction of the mutational history. The African mtDNA phylogeny has also been well documented in recent papers (Maca-Meyer et al. 2001; Torroni et al. 2001; Herrnstadt et al. 2002).

Silva et al. (2002) reported 40 mtDNAs, of which they assigned 31 to the Native American haplogroups A, B, C, and D (according to their fig. 1). The remaining nine mtDNAs can be assigned unambiguously to the Asian haplogroups B4 and D4, the Eurasian haplogroup U, and the African haplogroup L2a (table 1), as we will argue below. Figure 1 displays the truncation (relative to the 8.8-kb fragment under study) of the rooted phylogeny that is relevant for assigning these 40 mtDNAs to their respective haplogroups. This phylogeny is unanimously supported by the earlier publications. (However, note that mutations at 15301 and 11944 were not reconstructed most parsimoniously along the African mtDNA tree shown in fig. 1 of Herrnstadt et al. [2002]). The only instances of recurrent mutations (real or not) for the mutations and haplogroups highlighted in figure 1 are then as follows: the transversion 15487T is missing in the single haplogroup C lineage of Maca-Meyer et al. (2001); in the data of Herrnstadt et al. (2002), the B4b lineage 375 has experienced a transition at 14766, the L2a lineage 223 lacks the 7521 transition, and the 14566 transition is missing in the L2a lineage 165, which is closely related to another L2a lineage (bearing the 14566 mutation) from Torroni et al. (2001) in that they both share additional mutations at 3010 and 6663.

It is conspicuous that in all five haplogroup L2a mtDNAs of Silva et al. (2002), two of the basal transitions, 8206 and 14566, characteristic of L2 and L2a, respectively, are missed. Further L2a-diagnostic mutations, such as 7175, 7771, 13803, and 15784, are not always reported in the sequences (table 1). Moreover, the five L2a lineages have a total of only 11 other (private) mutations, comprising as many as five transversions, four deletions, and only two transitions. This pattern of private mutations differs from that in the three

L2a lineages (nine transitions and no other mutations) of Ingman et al. (2000) and Torroni et al. (2001) in the same mtDNA region. It thus looks as though most of the real private mutations in the L2a mtDNAs were missed and that, instead, phantom mutations were scored.

The basal mutation 15487T of haplogroup M8 (which embraces haplogroups C and Z) is omitted in all seven C lineages of Silva et al.'s data (table 1). Other basal mutations for haplogroup C lineages are missing at sites 7196A, 8584, and 14318, in different combinations. It is remarkable that even deep mutations, such as 10400, 10873, and 15301 that distinguish macrohaplogroups M and N, were overlooked in six of the seven C lineages.

Among the seven D lineages in Silva et al. (2002), three sequences share mutations or motifs with D sequences reported elsewhere (Ingman et al. 2000; Derbeneva et al. 2002). The sequence JAP1045 (from an individual of Japanese origin) shares 8964, 9296, and 9824A with a Japanese mtDNA sequence from Ingman et al. (2000) and, therefore, definitely belongs to haplogroup D4, although the two characteristic D4 transitions (8414 and 14668) are not reported in the entire data set, except for one occurrence of 14668 in an L2a sequence! Similarly, the Japanese mtDNA sequence JAP1043 bears one of the mutations, 11215, found in Siberian mtDNAs of haplogroup D4 (Ingman et al. 2000; Derbeneva et al. 2002). The Guarani sequence GRC0131 of Silva et al. (2002) shares a rare transversion 10816T and a rare transition 13059 with the Guarani sequence of Ingman et al. (2000), but only the latter one has 8414 and 14668 and is thus confirmed as belonging to D4. These cases provide strong evidence for the systematic oversight of the basal mutations 8414 and 14668 in all haplogroup D lineages from Silva et al. (2002). Just as in the case of haplogroup C, several of the basal mutations that separate M and N are also missing in most of the D lineages.

Anomalies are also found in the nine sequences belonging to haplogroup A, although it was claimed by Silva et al. (2002) to be "the most homogeneous and best characterized" cluster in figure 1. Sample KCR0029 contains basal mutations 10398 and 10400 for haplogroup M. Sample KPO0013 has the 14566 mutation that is characteristic of haplogroup L2a. Sample PTJ0003 bears the L2abc-specific mutation 11944. Moreover, site 8027 is found mutated in only one A lineage, whereas this mutation was present in all the A sequences in Herrnstadt et al. (2002) and in one Chukchi sequence reported by Ingman et al. (2000).

In the 11 B lineages, only sample KPO0001 has the 9-bp deletion in the COII/tRNA<sup>Lys</sup> intergenic region, characteristic of haplogroup B. One or both of the basal mutations of B4b, 13590 and 15535, occur in all the samples (with the exception of JAP1044) and hint that they belong to B4b. It should be noted that in Herrnstadt



**Figure 1** Skeleton of the basal mtDNA phylogeny for the haplogroups identified in the data of Silva et al. (2002). "CRS" and "rCRS" refer to the reference sequence of Anderson et al. (1981) and the revised reference sequence of Andrews et al. (1999), respectively. The suffixes A, G, C, and T indicate transversions, and "del" indicates a deletion. Parallel mutations in different branches are underlined.

et al. (2002), mutations 9950 and 11177 further defined a subhaplogroup of B4b that was baptized "B2." We suggest that the 11177 mutation could have been omitted by Silva et al. (2002) as well. The Japanese B lineage JAP1044 could belong to haplogroup B4c or, alternatively, to B4a, as judged by the 15346 mutation or the 10238 transition, respectively (if the latter was simply misreported as a deletion). Two samples, KRC0033 and QUE1880, bear the 10400 mutation of haplogroup M, whereas sample QUE1881 harbors the 15043 mutation of M.

The U sequence in Silva et al. (2002) contains the full motif of haplogroup U, plus two transversions and three transitions not previously found in the published U sequences (Ingman et al. 2000; Finnilä et al. 2001; Maca-Meyer et al. 2001; Herrnstadt et al. 2002).

Rare deletions are found in two L2a and one B lineage of Silva et al. (2002). The 15802delA and 15848delA

in the cytochrome *b* gene of sample NGR0522, 8047delT in the COII gene of sample NGR0524, and 10238delT in the ND3 gene of sample JAP1044 generate premature stop codons in these genes. These rare deletions all occur at a 2-bp repeat of the deleted base and might be generated by the Sequencer reading program. It is clear that the sequences of Silva et al. (2002) harbor more rare transversions and fewer private transitions than other reported sequences (Ingman et al. 2000; Finnila¨ et al. 2001; Maca-Mayer et al. 2001; Torroni et al. 2001; Herrnstadt et al. 2002). One cannot exclude the possibility that true transitions were erroneously scored as transversions or deletions by Silva et al. (2002). The two rare mutations 8860 and 15326 of the CRS are also missed in most of the sequences. The mutation 11335 in the CRS, which was found to be a sequencing error (Andrews et al. 1999), was present in 16 mtDNAs.

Processes that could account for these anomalies include the following:

- 1. Only one strand of mtDNA was sequenced;
- 2. Sequences were aligned with some variant of the CRS (a likely source of problems in the past; see Macaulay et al. [1999]);
- 3. Sequences from different samples, especially those belonging to different haplogroups, were aligned together during the editing process (In this way, one might easily "borrow" a fragment of one sample into another when the sequences of the latter were not overlapping and, thus, introduce basal polymorphisms of one mtDNA lineage into another);
- 4. Possible sample crossover or contamination during data collection;
- 5. Relying just on the sequence scored by the Sequencer reading program without further manual checking of the chromatogram, especially relevant in the case of the rare deletions; and/or
- 6. PCR errors during amplification.

In summary, we have every reason to mistrust the mtDNA sequences published by Silva et al. (2002). One cannot escape the conclusion that these data are seriously flawed or, at least, are not mtDNA as we know it.

> YONG-GANG YAO,<sup>1</sup> VINCENT MACAULAY,<sup>3</sup> TOOMAS KIVISILD,<sup>4</sup> YA-PING ZHANG,<sup>1,2</sup> AND HANS-JÜRGEN BANDELT<sup>5</sup>

1 *Kunming Institute of Zoology, Chinese Academy of Sciences, and* <sup>2</sup> *Laboratory for Conservation and Utilization of Bio-Resource, Yunnan University, Kunming, Yunnan, China;* <sup>3</sup> *Department of Statistics, University of Oxford, Oxford, United Kingdom;* 4 *Institute of Molecular and Cell Biology, Tartu University, Tartu, Estonia; and* <sup>5</sup> *Fachbereich Mathematik, Universita¨t Hamburg, Hamburg*

## **References**

- Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465
- Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet 23:147
- Bandelt H-J, Lahermo P, Richards M, Macaulay V (2001) Detecting errors in mtDNA data by phylogenetic analysis. Int J Legal Med 115:64–69
- Bandelt H-J, Quintana-Murci L, Salas A, Macaulay V (2002) The fingerprint of phantom mutations in mitochondrial DNA data. Am J Hum Genet 71:1150–1160
- Derbeneva OA, Sukernik RI, Volodko NV, Hosseini SH, Lott MT, Wallace DC (2002) Analysis of mitochondrial DNA diversity in the Aleuts of the Commander Islands and its implications for the genetic history of Beringia. Am J Hum Genet 71:415–421
- Finnila¨ S, Lehtonen MS, Majamaa K (2001) Phylogenetic network for European mtDNA. Am J Hum Genet 68:1475–1484
- Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, Howell N (2002) Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. Am J Hum Genet 70:1152–1171; 71:448–449 (erratum)
- Ingman M, Kaessmann H, Pääbo S, Gyllensten U (2000) Mitochondrial genome variation and the origin of modern humans. Nature 408:708–713
- Kivisild T, Tolk H-V, Parik J, Wang Y, Papiha SS, Bandelt H-J, Villems R (2002) The emerging limbs and twigs of the East Asian mtDNA tree. Mol Biol Evol 19:1737–1751
- Maca-Meyer N, González AM, Larruga JM, Flores C, Cabrera VC (2001) Major genomic mitochondrial lineages delineate early human expansions. BMC Genetics 2:13
- Macaulay V, Richards M, Sykes B (1999) Mitochondrial DNA recombination: no need to panic. Proc R Soc Lond B 266: 2037–2039
- Richards M, Macaulay V (2001) The mitochondrial gene tree comes of age. Am J Hum Genet 68:1315–1320
- Silva WA Jr, Bonatto SL, Holanda AJ, Ribeiro-dos-Santos AK, Paixão BM, Goldman GH, Abe-Sandes K, Rodriguez-Delfin L, Barbosa M, Paçó-Larson ML, Petzl-Erler ML, Valente V, Santos SEB, Zago MA (2002) Mitochondrial genome diversity of Native Americans supports a single early entry of founder populations into America. Am J Hum Genet 71:187– 192
- Torroni A, Rengo C, Guida V, Cruciani F, Sellitto D, Coppa A, Luna Calderon F, Simionati B, Valle G, Richards M, Macaulay V, Scozzari R (2001) Do the four clades of the mtDNA haplogroup L2 evolve at different rates? Am J Hum Genet 69:1348–1356

Address for correspondence and reprints: Dr. Yong-Gang Yao, Kunming Institute of Zoology, Chinese Academy of Sciences, 32 Jiaochang Donglu, Kunming, Yunnan, 650223, China. E-mail: ygyaozh@yahoo.com

 2003 by The American Society of Human Genetics. All rights reserved. 0002-9297/2003/7205-0030\$15.00

*Am. J. Hum. Genet. 72:1346–1348, 2003*

## **Correction: Mitochondrial DNA Variation in Amerindians**

*To the Editor:*

We thank Yao et al. (2003 [in this issue]) for calling our attention to inconsistencies in our data reporting mitochondrial DNA variations in Amerindians (Silva et al. 2002). We reviewed the original chromatograms and resequenced all the samples (forward and reverse). On the basis of the reanalysis of the initial data and sequencing that has been repeated, we conclude that most criticisms of Yao et al. are correct. We identified two sources of problems: (*a*) alignment with a variant CRS (Macaulay et al. 1999) and (*b*) mutations missed at regions of lowquality chromatograms in one (forward or reverse) of the first sequencing. Elimination of these two problems, by a second (and, in a few cases, a third) sequencing, careful manual checking of the chromatograms, and use of the correct rCRS reference sequence (MITOMAP) eliminated the discrepancies. A summary of all 40 corrected sequences is presented in figure 1, and the general pattern is similar to that recently reported by Herrnstadt et al. (2002). The presence of a private mutation in more than one individual or the absence of a basal mutation probably represent examples of homoplasy or of reverse mutations. Extensive homoplasy within the coding region of mtDNA has been documented (Eyre-Walker et al. 1999; Herrnstadt et al. 2002) and will probably be found more often as the number of mtDNA samples sequenced increases. For instance, the group C basal mutation 9545G was found in one individual from the haplogroup A, whereas private mutation 14460G was found in two individuals who belong to haplogroups A and D, and 15670C is present in one individual who belongs to haplogroup A and two who belong to haplogroup C (Herrnstadt et al. 2002). The finding of two similar private mutations (12406A) in two individuals of the same tribe (TYR0004 and TYR0016) is probably the consequence of a single mutational event, as is the occurrence of the reverse mutation 8584 in two individuals of another tribe (YAN0669 and YAN0650).

Recalculation of the age estimates for the four founder haplogroups on the basis of the reviewed data continues