Letters to the Editor

Am. J. Hum. Genet. 75:919–920, 2004

Etruscan Artifacts

To the Editor:

Recently, Vernesi et al. (2004) attempted to determine the variation in the first hypervariable segment (HVS-I) of mtDNA extracted from a number of Etruscan teeth. However, the rather unusual variation reported calls authenticity of the ancient mtDNA into question.

Multiple occurrences of the same tandem mutation in lineages from disjoint haplogroups distinctly signal artifacts (Bandelt et al. 2002). For instance, the "Etruscan" data harbor the transition pair 16193-16219 thrice, on quite different HVS-I backgrounds, even separated by a restriction site—once jointly with +14766 *Mse*I and twice with −14766 *Mse*I. The transition 16219, however, is quite a rare mutation, which is confined essentially to haplogroup U6ab (Maca-Meyer et al. 2003) and to a specific subclade of haplogroup H that bears the two characteristic mutations 16482 and 239. In the latter case, one indeed finds the motif 16193-16219-16362 in the worldwide database—for example, in Great Britain (Richards et al. 2000), in Germany (Pfeiffer et al. 1999), and in the United States (as recorded in the SWGDAM database [Monson et al. 2002]); however, in the study by Vernesi et al. (2004), the motif 16193-16219 occurs without 16362.

The mutation 16069, which almost always signals haplogroup J, is not seen in the "Etruscan" mtDNA data in combination with the 16126 transition or with +14766 *Mse*I but is recorded twice with −14766 *Mse*I. Could this surprising feature be explained by recurrent mutations at 16069? Hardly—the 16069 transition is not among the "speedy" transitions reported by Bandelt et al. (2002). The mtDNA lineages with motif 16126 (outside haplogroup H), 16126-16193, and 16126-16193-16278 are normally seen together with 16069 in the West Eurasian mtDNA pool (except in the evidently flawed HVS-I data of Fraumene et al. [2003]). Therefore, we have to expect at least four independent mutations at site 16069 in the "Etruscan" mtDNA data—if we do not want to invoke de novo mutations at 16193 and 16278 (and this without a single trace of any familiar haplogroup J lineage in this data set). The 16069 transition on non-J lineages has been observed earlier in the "Ladins" (Stenico et al. 1996), where it occurs three times (in tandem with the 16085 transition). The variation at 16069 in these data sets thus seems abnormal in comparison with the worldwide mtDNA database.

Vernesi et al. (2004) rejected the idea that postmortem damage could have been responsible for the assumed back mutations at 16069 (or 16294), with reference to the study of Gilbert et al. (2003), inasmuch as no such case was observed there. This argument is, however, invalid, since the regular haplogroup J (or T, respectively) nucleotide T at 16069 (or 16294, respectively) is less frequent by 1 order of magnitude than the corresponding majority nucleotide C (of the Cambridge reference sequence), so that no significant inference could be drawn from Gilbert et al. (2003) as it was for asymmetric transition probabilities (C→T vs. T→C) at these sites.

There are numerous technical flaws in figure 2 and table 1 of Vernesi et al. (2004)—in addition to the potential sequencing problems—thus additionally undermining confidence in their ancient mtDNA data. Haplotype 6AM is not well defined, because the corresponding HVS-I sequences (from Adria and Magliano/Marsiliana [Vernesi et al. 2004]) are observed with both +/−14766 *Mse*I, according to their table 1. Haplotype 13C is misplaced in figure 2, since it bears +14766 *Mse*I. This figure certainly does not present a reduced median network (as claimed), because the Adria 6AM haplotype should be a neighbor of haplotype 14CMT. It is not at all clear how this network was actually constructed, because, in the article, the reference is given to the median-joining algorithm (Bandelt et al. 1999), which is fundamentally different from the reduced-median algorithm (Bandelt et al. 2000). The median-joining algorithm, however, would instead reconstruct a triangle for site 16095, in this case. In the data set, the 16223 transition relative to rCRS is observed twice in connection with −14766 *Mse*I, which otherwise should be more the exception than the rule, since rCRS is the ancestral HVS-I motif of haplogroup HV (as well as of the superhaplogroup R). On the other hand, this restriction site has not been determined for haplotype 9A, yet it is reconstructed in the network, not most parsimoniously, as +14766 *Mse*I. Finally, the node sizes in figure 2 do not always correspond to the frequencies recorded in table 1 of the article.

The assertion that "all the strictest criteria for the val-

idation of ancient DNA sequences have been followed" (Vernesi et al. 2004 [p. 703]) is not quite correct, since one of the most important criteria of Cooper and Poinar (2000)—that is, that of independent replication in another lab—has not been followed for 25 of 28 of the reported HVS-I sequences or for any of the RFLP tests. Moreover, the 20 excluded sequences were not displayed. The claim that the "Etruscan" sequences "all belong to lineages that are still present in Europe" (Vernesi et al. 2004 [p. 702]) is not justified, in view of the unusual mutational pattern, especially as the basal haplogroup status (U, JT, pre-HV, N1, W, X, or other) was not determined in half of the data set. Under these circumstances, it is unclear to what extent the "Etruscan" data represent severely damaged or partly contaminated mtDNA sequences; therefore, any comparison with modern population data must be considered quite hazardous.

HANS-JÜRGEN BANDELT

Fachbereich Mathematik
Universität Hamburg
Hamburg

References

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Bandelt H-J, Macaulay V, Richards M (2000) Median networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA. Mol Phylogenet Evol 16:8–28

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

Cooper AR, Poinar H (2000) Ancient DNA: do it right or not at all. Science 289:1139

Fraumene C, Petretto E, Angius A, Pirastu M (2003) Striking differentiation of sub-populations within a genetically homogeneous isolate (Ogliastra) in Sardinia as revealed by mtDNA analysis. Hum Genet 114:1–10

Gilbert MTP, Willerslev E, Hansen AJ, Barnes I, Rudbeck L, Lynnerup N, Cooper A (2003) Distribution patterns of postmortem damage in human mitochondrial DNA. Am J Hum Genet 72:32–47 (erratum 72:779)

Maca-Meyer N, González AM, Pestano J, Flores C, Larruga JM, Cabrera VM (2003) Mitochondrial DNA transit between West Asia and North Africa inferred from U6 phylogeography. BMC Genetics 4:15

Monson KL, Miller KWP, Wilson MR, DiZinno JA, Budowle B (2002) The mtDNA population database: an integrated software and database resource for forensic comparison. Forensic Sci Comm 4:2

Pfeiffer H, Brinkmann B, Hühne J, Rolf B, Morris AA, Steighner R, Holland MM, Forster P (1999) Expanding the forensic German mitochondrial DNA control region database: genetic diversity as a function of sample size and microgeography. Int J Legal Med 112:291–298

Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, et al (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. Am J Hum Genet 67:1251–1276

Stenico M, Nigro L, Bertorelle G, Calafell F, Capitanio M, Corrain C, Barbujani G (1996) High mitochondrial sequence diversity in linguistic isolates of the Alps. Am J Hum Genet 59:1363–1375

Vernesi C, Caramelli D, Dupanloup I, Bertorelle G, Lari M, Cappellini E, Moggi-Cecchi J, Chiarelli B, Castrì L, Casoli A, Mallegni F, Lalueza-Fox C, Barbujani G (2004) The Etruscans: a population-genetic study. Am J Hum Genet 74: 694–704

Address for correspondence and reprints: Dr. Hans-Jürgen Bandelt, Fachbereich Mathematik, Universität Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany. E-mail: bandelt@math.uni-hamburg.de

0002-9297/2004/7505-0020$15.00

Am. J. Hum. Genet. 75:920–923, 2004

On the Etruscan Mitochondrial DNA Contribution to Modern Humans

To the Editor:

The growing number of ancient human mtDNA samples sequenced in recent years has given rise to the problem of correspondence between distributions of mutations in ancient and modern mtDNA sequences. It has been suggested that mtDNA nucleotide sequences obtained from human remains may include some artifacts, for multiple reasons, such as contamination with modern DNA; artifacts induced by cytosine deamination during multiple amplification of ancient DNA via PCR; and postmortem damage in DNA, occurring as hydrolytic deamination and depurination, double-strand breaks, and oxidative nucleotide modification (Hofreiter et al. 2001*a*). Therefore, to determine the nature of the DNA sequences amplified, each amplified product should be cloned, and the obtained clones should be sequenced (Pääbo 1989; Handt et al. 1996). The consensus sequence from each sample is determined from the sequences shared between all clones, and intraclone nucleotide differences represent the postmortem data set (Gilbert et al. 2003). Therefore, cloned sequences of ancient DNA samples may show a pattern of a shared consensus (haplotype), with many singleton substitutions corresponding to postmortem DNA changes. It has been suggested that the consensus sequence should be part of the original sequence (Hofreiter et al. 2001*b*).

In this study, we reanalyzed nucleotide sequences of the mtDNA HVS-I region in 575 clones derived from bone samples of 28 Etruscans (7th–3rd centuries B.C.),

recently published by Vernesi et al. (2004). To determine whether the Etruscan samples indeed represent mtDNA sequences similar to modern human ones, we compared mutational spectra derived from ancient mtDNAs and from mtDNA sequences characteristic of present-day Europeans, Asians, and Africans. Figure 1 shows the distribution of mutations found in both mtDNA data sets. For modern human populations, we analyzed the distribution of mutations found in ~8,000 HVS-I sequences belonging to 90 mtDNA haplogroups (Malyarchuk and Rogozin 2004). The number of parallel mutations that have independently arisen at certain nucleotide positions in different mtDNA haplogroups (monophyletic clusters) was used as a measure of DNA variability in modern populations (for details, see Malyarchuk et al. 2002). In the case of ancient mtDNA data, we assumed that mutations found in cloned sequences of each sample may be considered as having independently arisen in different DNA templates—that is, we do not suggest that cloned sequences are the products of a single template. Comparative analysis of data sets shows that, among 261 variable nucleotide positions seen in both spectra, 222 nucleotide positions are variable in the mutational spectrum of modern humans, 147 positions are variable in the Etruscan spectrum, and 108 positions appear to be shared between the two spectra. The frequency of variable positions in the Etruscan spectrum is 147/342 = 0.42 (the sequence length is 342 bases). The frequency of variable positions in the mutational spectrum of modern humans is 222/342 = 0.64. If we assume a random distribution of variable positions along the HVS-I sequence, the expected frequency of shared variable positions is 0.64 × 0.42 = 0.28. The observed frequency of shared variable positions is 0.31; this value is not significantly different from the randomly expected 0.28 ($\chi^2 = 2.3$, $P(\chi^2) = .13$).

This result suggested that the mutational spectrum of the Etruscan mtDNA is characterized by a large fraction of the Etruscan-specific mutations; moreover, the fraction of shared variable positions is not different from the random expectation. According to the database of HVS-I sequences combined with mtDNA coding region markers that was used for comparison (fig. 1), 26.5% (39 of 147) of variable positions observed in the cloned ancient mtDNA sequences contain unique mutations—that is, mutations that have not been found in modern humans. A similar value (27.7%) was found by comparison of the HVS-I region variation data in modern humans and the Cro-Magnon–type individuals, with a sample date of 23,000–25,000 years ago (Caramelli et al. 2003). In 165 cloned sequences of these two ancient individuals, 13 of 47 variable positions were found to be unique. Among them, only five variable positions (16057, 16059, 16112, 16139, and 16158) were shared by the Etruscan and Cro-Magnon HVS-I data sets. Meanwhile, other ancient and modern data sets on HVS-I variability show no or small deviation from the mutational spectrum of modern humans used in this study. For example, no differences were found between this spectrum and the HVS-I sequences of mtDNAs extracted from the skeletal remains of 44 specimens of the Xiongnu tribe (from the Egyin Gol necropolis, northern Mongolia, 3rd century B.C. to 2nd century A.D.) (Keyser-Tracqui et al. 2003), and small differences (<3%) were found in comparison with HVS-I sequence variation in modern populations of the Roma (1 of 64 variable positions) (Gresham et al. 2001), the Egyptians (2 of 71 variable positions) (Stevanovitch et al. 2004), and the Italians from Bologna (1 of 58 variable positions) (Bini et al. 2003).

Although the screening of the Mitomap database for polymorphic nucleotide positions in the HVS-I region has shown that 11 of 39 Etruscan-specific positions were previously found as variable in different individuals, the frequency of the variable positions specific to the mutational spectrum of the cloned sequences of the Etruscans (positions 16044, 16045, 16056, 16057, 16060, 16072, 16073, 16083, 16091, 16098, 16100, 16101, 16112, 16118, 16123, 16130, 16139, 16151, 16158, 16159, 16237, 16282, 16306, 16315, 16334, 16339, 16345, and 16348) remains very high (19%). It is likely that these nucleotide positions represent a mutational spectrum of the mtDNA molecules altered by postmortem damage. Comparison of these positions with the list of nucleotide positions suggested as sites with postmortem damage in the study of ancient DNA from northwestern European samples (Gilbert et al. 2003) shows that only position 16072 is shared between the two data sets. In addition, positions 16131, 16144, and 16325, which have an increased mutation rate of postmortem damage, according to Gilbert et al. (2003), were found as singleton mutations in the Etruscan cloned sequences.

It is important that some of the Etruscan mutations, which are rare or absent in modern humans, were found in multiple clones of ancient individuals and therefore were assigned by Vernesi et al. (2004) to the consensus haplotypes, suggesting that these mutations should represent the original mtDNA sequences of the Etruscans. The most noticeable position is 16334, which was found in 15 cloned Etruscan sequences belonging to two different haplotypes (3V and 22T). However, this position is invariable in almost 8,000 of the HVS-I sequences of modern humans. Nucleotide positions 16228 and 16229 are also among the most conservative positions in modern human data sets, but mutations at these positions were found frequently in the Etruscan nucleotide sequences—mutation C16228T was present as a consensus variant in haplotype 21T and as a singleton mutation in two cloned sequences belonging to another specimens, and mutation T16229C was observed as a consensus

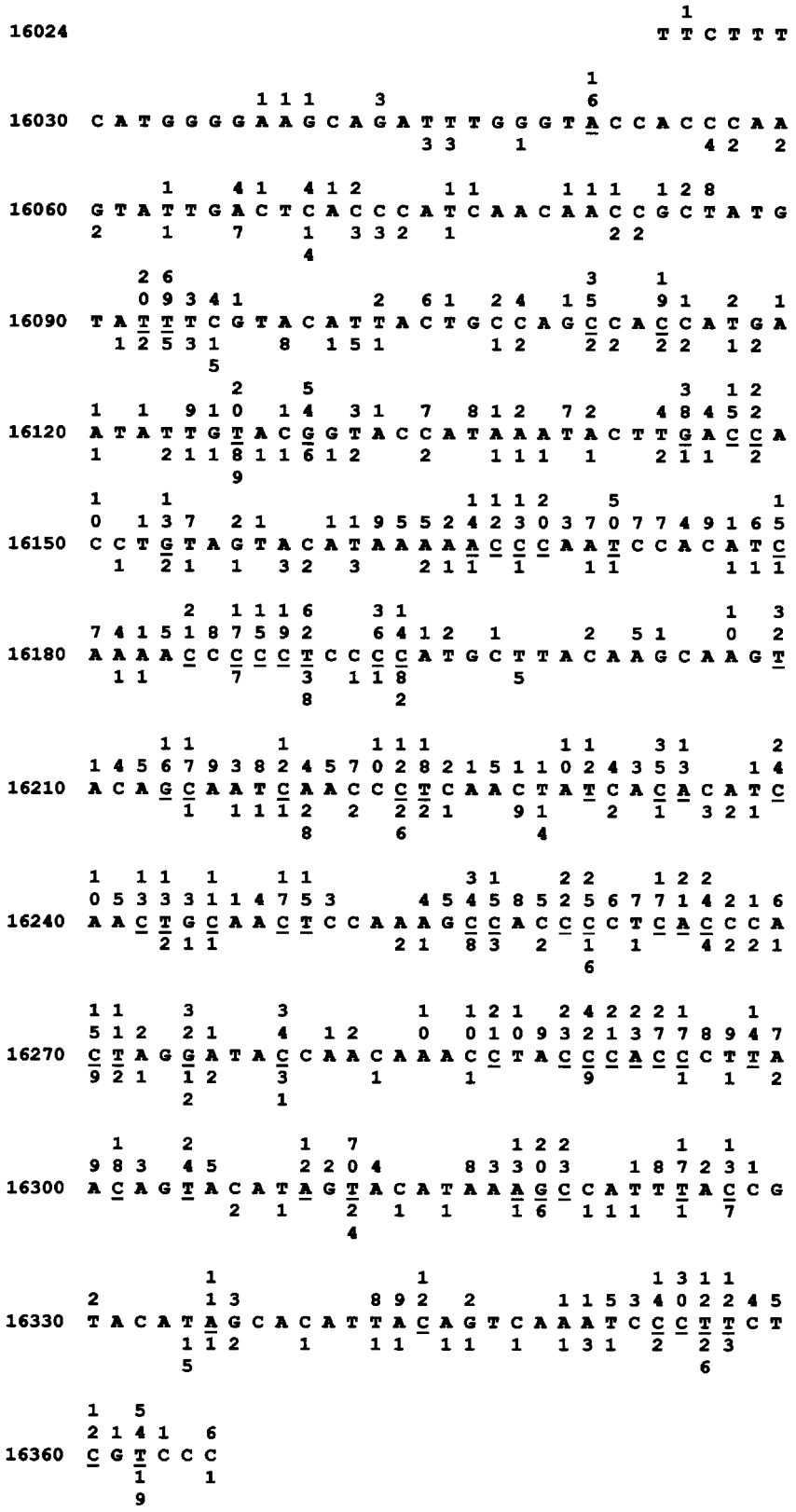

Figure 1 Comparison of HVS-I mutational spectra (between nucleotide positions 16024 and 16365) in ancient (Etruscan) and modern humans. Mutations are shown relative to the Cambridge reference sequence (Anderson et al. 1981). Numbers above the sequence are numbers of parallel mutations (both transitions and transversions) observed in mtDNA haplogroups from modern human populations; predicted mutational hotspots are underlined (for details, see Malyarchuk et al. [2002] and Malyarchuk and Rogozin [2004]). Numbers under the sequence are numbers of mutations found in cloned Etruscan HVS-I sequences.

variant in haplotypes 20T and 21T and as a singleton mutation in another sample. Haplotype 21T is represented by a combination of variants 16228T and 16229C, which has not been found in modern human HVS-I sequences.

Therefore, the mutational spectrum derived from the Etruscan mtDNA sequences shows some degree of similarity to modern human mtDNA sequences. However, many of the singleton mutations, as well as some consensus mutations found in the cloned sequences, represent substitutions that are very rare in living individuals or do not even exist. The possibility that these haplotypes underwent extinction (Vernesi et al. 2004) cannot be excluded. However, many of these mutations might be due to postmortem damage of mtDNA. The assignment of postmortem mutations in consensus variants of the haplotypes can lead to misidentification of mtDNA sequences. In addition, some phylogenetically informative nucleotide positions are highly susceptible to postmortem damage (Gilbert et al. 2003). These problems may lead to misassignment of mtDNA sequences to haplogroups and, consequently, to biased opinions about genetic history of human populations.

BORIS A. MALYARCHUK[1] AND IGOR B. ROGOZIN[2]
[1]*Institute of Biological Problems of the North, Far East Branch of the Russian Academy of Sciences, Magadan, Russia;* [2]*National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda*

Electronic-Database Information

The URL for data presented herein is as follows:

Mitomap: A Human Mitochondrial Genome Database, http://www.mitomap.org/

References

Anderson S, Bankier AT, Barrel BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJM, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Bini C, Ceccardi S, Luiselli D, Ferri G, Pelotti S, Colalongo C, Falconi M, Pappalardo G (2003) Different informativeness of the three hypervariable mitochondrial DNA regions in the population of Bologna (Italy). Forensic Sci Int 135:48–52

Caramelli D, Lalueza-Fox C, Vernesi C, Lari M, Casoli A, Mallegni F, Chiarelli B, Dupanloup I, Bertranpetit J, Barbujani G, Bertorelle G (2003) Evidence for a genetic discontinuity between Neandertals and 24,000-year-old anatomically modern Europeans. Proc Natl Acad Sci USA 100: 6593–6597

Gilbert MT, Willerslev E, Hansen AJ, Barnes I, Rudbeck L, Lynnerup N, Cooper A (2003) Distribution patterns of postmortem damage in human mitochondrial DNA. Am J Hum Genet 72:32–47

Gresham D, Morar B, Underhill PA, Passarino G, Lin AA, Wise C, Angelicheva D, Calafell F, Oefner PJ, Shen P, Tournev I, de Pablo R, Kucinskas V, Perez-Lezaun A, Marushiakova E, Popov V, Kalaydjieva L (2001) Origins and divergence of the Roma (Gypsies). Am J Hum Genet 69:1314–1331

Handt O, Krings M, Ward R, Pääbo S (1996) The retrieval of ancient human DNA sequences. Am J Hum Genet 59:368–376

Hofreiter M, Janicke V, Serre D, von Haeseler A, Pääbo S (2001*a*) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res 29:4793–4799

Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S (2001*b*) Ancient DNA. Nat Rev Genet 2:353–359

Keyser-Tracqui C, Crubezy E, Ludes B (2003) Nuclear and mitochondrial DNA analysis of a 2,000-year-old necropolis in the Egyin Gol Valley of Mongolia. Am J Hum Genet 73: 247–260

Malyarchuk BA, Rogozin IB (2004) Mutagenesis by transient misalignment in human mitochondrial DNA control region. Ann Hum Genet 68:324–339

Malyarchuk BA, Rogozin IB, Berikov VB, Derenko MV (2002) Analysis of phylogenetically reconstructed mutational spectra in human mitochondrial DNA control region. Hum Genet 111:46–53

Pääbo S (1989) Ancient DNA: extraction, characterization, molecular cloning and enzymatic amplification. Proc Natl Acad Sci USA 86:1939–1943

Stevanovitch A, Gilles A, Bouzaid E, Kefi R, Paris F, Gayraud RP, Spadoni JL, El-Chenawi F, Beraud-Colomb E (2004) DNA sequence diversity in a sedentary population from Egypt. Ann Hum Genet 68:23–39

Vernesi C, Caramelli D, Dupanlup I, Bertorelle G, Lari M, Cappellini E, Moggi-Cecchi J, Chiarelli B, Castri L, Casoli A, Mallegni F, Lalueza-Fox C, Barbujani G (2004) The Etruscans: a population-genetic study. Am J Hum Genet 74: 694–704

Address for correspondence and reprints: Dr. Boris A. Malyarchuk, Genetics Laboratory, Institute of Biological Problems of the North, Far East Branch of the Russian Academy of Sciences, Portovaya Street, 18, 685000, Magadan, Russia. E-mail: malyar@ibpn.kolyma.ru

0002-9297/2004/7505-0021$15.00

Am. J. Hum. Genet. 75:923–927, 2004

Etruscan Artifacts: Much Ado about Nothing

To the Editor:

Malyarchuk and Rogozin (2004 [in this issue]) and Bandelt (2004 [in this issue]) question the authenticity of

the sequences in the study by Vernesi et al. (2004), because these sequences contain substitution motifs that are seldom observed in modern samples. Before answering their criticisms in detail, we would like to make three general points: (1) in the study of ancient DNA, the presence of artifacts is inevitably hard to rule out with certainty, and this also applies to our Etruscan data; however, (2) the Etruscan sequences were determined using a very strict set of standards, and, hence, we are confident that they are rather reliable; and (3) many of Bandelt's speculations are based on sites with a high mutation rate, where substitutions have been observed in various associations. For instance, positions 16126, 16219, 16278, and 16362, found in Etruscan haplotypes that Bandelt (2004 [in this issue]) considers with suspicion, are described as mutational hotspots by Wakeley (1993), Hasegawa et al. (1993), Meyer et al. (1999), and even (three of them) Bandelt et al. (2002). Multiple occurrences of those mutations on different lineages does not prove, or suggest, sequencing artifacts.

To test the reproducibility of the results, in our Etruscan study, we cloned the PCR products and sequenced multiple clones from the same individuals (Handt et al. 1996). *Taq* polymerase errors may lead to the apparent occurrence of mutations unique to a single clone (singletons). These errors are, as completely as possible, eliminated by comparing clones and identifying a consensus sequence. Malyarchuk and Rogozin (2004 [in this issue]) compared variation across modern humans with that observed in the 575 clones derived from 28 Etruscan specimens. They noticed that the mutations in the two data sets overlap only in part and that mutations never observed in modern humans occur in the Etruscan clones. However, by making this comparison they were comparing incomparable quantities. On the one hand, they have the spectrum of mutations carried by modern individuals; on the other, they have a set of mutations comprising both those that were actually in the Etruscans' DNA and those that probably result from *Taq* errors. There is no reason to expect the two spectra to be similar. As mentioned on page 698 of our article (Vernesi et al. 2004), the misincorporation rate was 0.27% for the Etruscans—that is, less than observed, for instance, by Handt et al. (1994). Therefore, Malyarchuk and Rogozin (2004 [in this issue]) are wrong when they say that the clones of the Etruscan data set show an excess of singletons. On the contrary, if the appropriate comparisons are made, singleton sequences are relatively rare among them. Malyarchuk and Rogozin (2004 [in this issue]) also notice that two mutations (16229 and 16334) independently occurring twice in the Etruscan network are rare or absent in almost 8,000 modern samples. Their observation is correct, but, given the methodologies that we applied, at this stage, we do not think it implies that these mutations are necessarily laboratory artifacts.

Bandelt (2004 [in this issue]) makes five comments, which we list here along with our responses.

1. **The 16193-16219 motif occurs in association with different substitutions in modern samples and among the Etruscans.** Yes, we already wrote that in the second-to-last paragraph on page 698 (Vernesi et al. 2004). Bandelt notes that the mutation at position 16219 has been observed in two different haplogroups in modern people. We know of at least three (U6a, T*, and H) and, within H, of several different haplotypes with the 16219 mutation (H07, H09, H13, and H16). Meyer et al. (1999) showed that mutation rates at 16219 are fivefold higher than the average mutation rate in this region, and therefore 16219 seems to be a fast-mutating site, which can exist in association with a broad range of other mutations.
2. **The mutation at 16069 occurs in two individuals, 2V and 11C, and in neither is it associated with a mutation at 16126 or with a cut at −766 *Mse*I, which generally characterize haplogroup J.** We published sequence 2V for completeness of information, but we did not consider it in any numerical analyses, because it seemed suspicious to us as well (p. 698). As for sequence 11C, the absence of the restriction cut at 14766 *Mse*I was confirmed by sequencing the region around the restriction site. At any rate, a sequence, SCOT0492, with the 16069-16261 motif but without a mutation at 16126, was described by Helgason et al. (2001). Also, 16126 is considered by Bandelt et al. (2002) to be a "speedy site" and, hence, not a site informative for phylogenetic analysis. Finally, Meyer et al. (1999) showed that position 16069 has a higher-than-average mutation rate; hence, whether to consider it as a fast-mutating site seems largely a matter of taste.
3. **We rejected the hypothesis that postmortem changes may have affected the 16069 or 16294 sites.** No, we didn't. Instead, we wrote (p. 699) that we saw *no compelling reason* to think that postmortem changes had occurred there, because Gilbert et al. (2003) found no instance of changes of that kind in their study. Ruling out postmortem changes completely is probably impossible with the available methodologies.
4. **There are inconsistencies between table 1 and figure 2.** Yes, here Bandelt is right. Haplotype 6AM was erroneously reported in one case with +14766 *Mse*I, whereas it has been observed only in association with −14766 *Mse*I. Another error is a typo in the table, where 13C should be −14766 (as correctly represented in the network). On the other hand, no triangle should be reconstructed for site 95 (as asked by Ban-

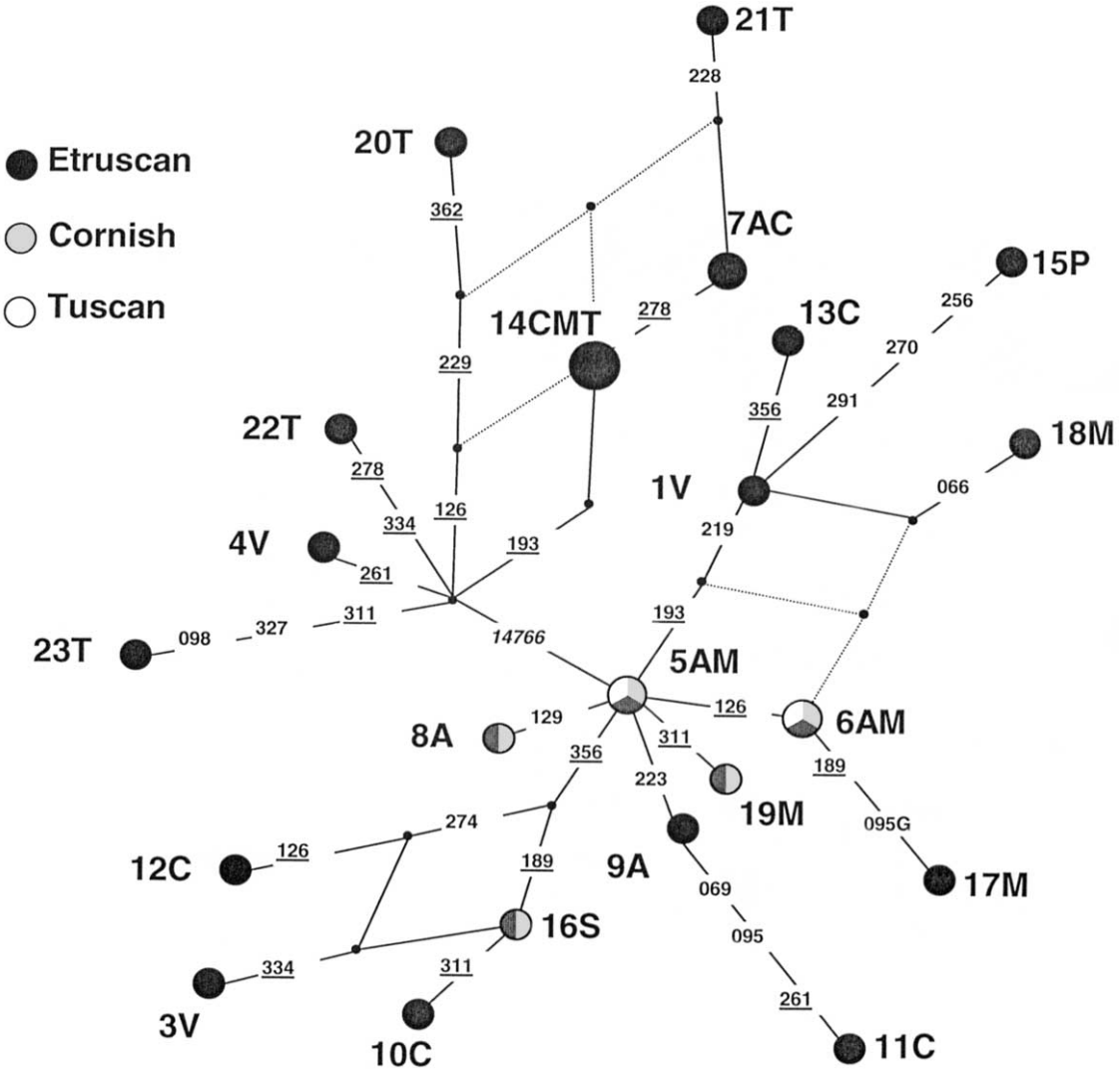

Figure 1 Reduced median network (Bandelt et al. 1995) of sequences identified in Etruscan individuals. The numbers are arbitrary codes; the letters refer to the sites where the sequences have been observed. Haplotype 5AM is the Cambridge Reference Sequence. Nodes are proportional to frequencies. The alleles shared with modern Tuscans and modern Cornish are highlighted.

delt), since the two mutations at this site are different (as correctly reported in the table and on p. 698; a gap in the label in the figure shows that the letter G, indicating a transversion, disappeared in print). He is also right when he points out that there is an error in the network method cited, but the error is in the reference list—namely, Bandelt et al. (1999) is given instead of (1995). However, Bandelt is wrong when he says that the node sizes in figure 2 do not correspond to the haplotype frequencies. Sequences 5AM, 6AM, and 7AC (two occurrences) are represented by big circles, and sequence 14CMT (three occurrences) by a very big circle. The network is a reduced median network (Bandelt et al. 1995, 2000), constructed by hand and checked with the program Network 4.1 for Windows (Fluxus Engineering Web site). Nucleotide positions were weighted using the list of Hasegawa et al. (1993) and following the weighting procedure of Richards et al. (1998). We further reduced the reticulation in the network by a procedure followed in other mitochondrial studies (Richards et al. 1996; Torroni et al. 2001) (see fig. 1). No conclusions of our article change on the basis of this slightly modified network.

5. **Not all of the strictest criteria for the validation of ancient sequences have been followed, because only three individuals were independently sequenced in two laboratories.** All right, then we shall reformulate our sentence as follows: We are not aware of any studies in which the criteria for the validation of an-

cient sequences were followed more strictly. See, for example, Keyser-Tracqui et al. (2003) and Maca-Meyer et al. (2004). Cooper and Poinar (2000) also recommend looking for human DNA in associated faunal material, and we could do that only for one individual, because animal bones were not retrieved from any other burial. We shall try to replicate greater numbers of results in future studies, but it must be understood that it is a very long process and one for which ancient material may not always be easily available.

In brief, we thank Bandelt for detecting some differences between the data reported in table 1 of our article and their graphical representation in the network. However, we suggest that Bandelt read our text more carefully and also consider the lists of fast-mutating sites compiled by authors other than himself. Bandelt is clearly right when he stresses the importance of quality in the data, and we showed that we share the same concern in three ways: (1) by excluding from the analysis 50 of the initial 80 samples, whenever any of eight biochemical tests did not suggest a good probability of obtaining reproducible results; (2) by further excluding two identical sequences from the same burial, to make sure that consanguinity did not bias our results; and (3) by excluding sequence 2V because it showed two sets of mutations previously observed in evolutionarily distant haplogroups, and that could conceivably result from the presence of multiple DNAs in the specimen. On the contrary, the problems Bandelt raises are due to individual substitutions that, he says, are "almost always" observed in different contexts. We suspect science would proceed very slowly, if at all, if scientists agreed to trust only the data they observe almost always. Mitochondrial data sets contain many homoplasies, and sites that were considered mutational hotspots have changed their status over time. Our understanding of mitochondrial variation is still evolving, and it seems bizarre to dismiss as implausible all the data that do not neatly conform to what we think we already know.

Whereas Malyarchuk, Rogozin, and Bandelt think that the Etruscan sequences are too different from modern ones to be good, Serre et al. (2004) argued that ancient mitochondrial sequences should be considered authentic only if they clearly differ from known modern ones. In other words, to make everybody happy, ancient mtDNA sequences should be at the same time *identical to and different from* the sequences of modern people. The famous lines of Joseph Heller's *Catch 22* come inevitably to mind: "Orr was crazy and could be grounded. All he had to do was ask; and as soon as he did, he would no longer be crazy and would have to fly more missions."

With this, we do not mean to deny that problems exist with the validation of ancient DNA sequences. What is necessary, however, is a set of standard criteria that everybody is reasonably happy about. Despite serious attempts to define these criteria (Cooper and Poinar 2000; Hofreiter et al. 2001), the present debate shows that a consensus has not been reached yet. As Helgason and Stefánsson (2003) remarked, errors can and do occur in large-scale DNA studies, either in the laboratory or in the construction of the databases. However, the impact of such errors, or of the possibility of such errors, should be evaluated critically before raising unjustified doubts about the conclusions of a study.

Guido Barbujani,[1] Cristiano Vernesi,[1] David Caramelli,[2] Loredana Castrì,[3] Carles Lalueza-Fox,[4] and Giorgio Bertorelle[1]
[1]*Dipartimento di Biologia, Università di Ferrara, Ferrara, Italy;* [2]*Dipartimento di Biologia Animale e Genetica, Laboratorio di Antropologia, Università di Firenze, Firenze, Italy;* [3]*Dipartimento di Biologia Evoluzionistica e Sperimentale, Università di Bologna, Bologna, Italy; and* [4]*Unitat d'Antropologia, Facultat de Biologia, Universitat de Barcelona, Barcelona*

Electronic-Database Information

The URL for data presented herein is as follows:

Fluxus Engineering, http://www.fluxus-engineering.com/sharenet.htm (for Network 4.1 for Windows)

References

Bandelt H-J (2004) Etruscan artifacts. Am J Hum Genet 75: 919–920 (in this issue)

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Bandelt HJ, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations using median networks. Genetics 141:743–753

Bandelt HJ, Macaulay V, Richards M (2000) Median networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA. Mol Phylogenet Evol 16:8–28

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

Cooper A, Poinar HN (2000) Ancient DNA: do it right or not at all. Science 289:1139

Gilbert MT, Hansen AJ, Willerslev E, Rudbeck L, Barnes I, Lynnerup N, Cooper A (2003) Characterization of genetic miscoding lesions caused by postmortem damage. Am J Hum Genet 72:48–61

Handt O, Krings M, Ward RH, Pääbo S (1996) The retrieval

of ancient human DNA sequences. Am J Hum Genet 59: 368–376

Handt O, Richards M, Trommsdorff M, Kilger C, Simanainen J, Georgiev O, Bauer K, Stone A, Hedges R, Schaffner W, Utermann G, Sykes B, Pääbo S (1994) Molecular genetic analyses of the Tyrolean Ice Man. Science 264:1775–1778

Hasegawa M, Di Rienzo A, Kocher TD, Wilson AC (1993) Toward a more accurate time scale for the human mitochondrial DNA tree. J Mol Evol 37:347–354

Helgason A, Hickey E, Goodacre S, Bosnes V, Stefansson K, Ward R, Sykes B (2001) mtDNA and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. Am J Hum Genet 68:723–737

Helgason A, Stefánsson K (2003) Erroneous claims about the impact of mitochondrial DNA sequence database errors. Am J Hum Genet 73:974–975

Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S (2001) Ancient DNA. Nat Rev Genet 2:353–359

Keyser-Tracqui C, Crubézy E, Ludes B (2003) Nuclear and mitochondrial DNA analysis of a 2,000-year-old necropolis in the Egyin Gol valley of Mongolia. Am J Hum Genet 73: 247–260

Maca-Meyer N, Arnay M, Rando JC, Flores C, Gonzalez AM, Cabrera VM, Larruga JM (2004) Ancient mtDNA analysis and the origin of Guanches. Eur J Hum Genet 12:155–162

Malyarchuk BA, Rogozin IB (2004) On the Etruscan mitochondrial DNA contribution to modern humans. Am J Hum Genet 75:920–923 (in this issue)

Meyer S, Weiss G, von Haeseler A (1999) Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. Genetics 152:1103–1110

Richards M, Côrte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Papiha S, Hedges R, Bandelt H-J, Sykes B (1996) Paleolithic and Neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 59:185–203

Richards MB, Macaulay VA, Bandelt HJ, Sykes BC (1998) Phylogeography of mitochondrial DNA in western Europe. Ann Hum Genet 62:241–260

Serre D, Langaney A, Chech M, Teschler-Nicola M, Paunovic M, Mennecier P, Hofreiter M, Possnert GG, Pääbo S (2004) No evidence of Neandertal mtDNA contribution to early modern humans. PLoS Biol 2:E57

Torroni A, Bandelt HJ, Macaulay V, Richards M, Cruciani F, Rengo C, Martinez-Cabrera V, et al (2001) A signal, from human mtDNA, of postglacial recolonization in Europe. Am J Hum Genet 69:844–852

Vernesi C, Caramelli D, Dupanloup I, Bertorelle G, Lari M, Cappellini E, Moggi-Cecchi J, Chiarelli B, Castri L, Casoli A, Mallegni F, Lalueza-Fox C, Barbujani G (2004) The Etruscans: a population-genetic study. Am J Hum Genet 74: 694–704

Wakeley J (1993) Substitution rate variation among sites in hypervariable region 1 of human mitochondrial DNA. J Mol Evol 37:613–623

Address for correspondence and reprints: Dr. Guido Barbujani, Dipartimento di Biologia, Università di Ferrara, via Borsari 46, I-44100 Ferrara, Italy. E-mail: bjg@dns.unife.it
0002-9297/2004/7505-0022$15.00

Am. J. Hum. Genet. 75:927–928, 2004

GDD1 Is Identical to TMEM16E, a Member of the TMEM16 Family

To the Editor:

In the June 2004 issue of *The American Journal of Human Genetics,* Tsutsumi et al. (2004) reported the identification and characterization of the *GDD1* gene, which is mutated in patients with gnathodiaphyseal dysplasia (MIM 166260). They claimed that human *GDD1* is a novel gene without any human homologs (Tsutsumi et al. 2004); however, we found that *GDD1* was identical to *TMEM16E* (MIM 608662), a member of the *TMEM16* gene family (Katoh and Katoh 2003, 2004*a*, 2004*b*).

In 2003, we identified and characterized the *TMEM16A* (*FLJ10261*) gene, which is located within the 11q13.3 amplicon (Katoh and Katoh 2003). The *CCND1-ORAOV1-FGF19-FGF4-FGF3-TMEM16A-FADD-PPFIA1-EMS1* amplicon at human chromosome 11q13.3 is one of the most frequently amplified regions in the human genome (Schwab 1998; Katoh and Katoh 2003). The *FLJ10261, C12orf3, C11orf25,* and *FLJ34272* genes, which encode mutually homologous eight-transmembrane proteins with N- and C-terminal tails facing the cytoplasm, were designated as "*TMEM16A,*" "*TMEM16B,*" "*TMEM16C,*" and "*TMEM16D,*" respectively, on the basis of our communication with the Human Gene Nomenclature Committee (see the HUGO Gene Nomenclature Committee Web site).

We then searched for novel members of the *TMEM16* gene family and identified the *TMEM16E, TMEM16F* (MIM 608663), and *TP53I5* genes (Katoh and Katoh 2004*a*, 2004*b*). TMEM16A, TMEM16B, TMEM16C, TMEM16D, TMEM16E, TMEM16F, and TP53I5 are eight-transmembrane proteins with TMEM16 homologous (TM16H1, TM16H2, and TM16H3) domains. Several Cys residues and Asn-linked glycosylation sites are included in the conserved residues (or the consensus sequence) of the TM16H1, TM16H2, and TM16H3 domains.

The *TMEM16E-NELL1* locus at human chromosome 11p15.1-p14.3 and the *TMEM16F-NELL2* locus at human chromosome 12q12 are paralogous regions (par-

alogons) within the human genome. Phylogenetic analysis revealed that TMEM16E and TMEM16F constitute a subfamily among TMEM16 family proteins. On the basis of these facts, we concluded that the *TMEM16E* and *TMEM16F* genes are paralogs within the human genome (Katoh and Katoh 2004*b*).

Tsutsumi et al. (2004) suggested that the human GDD1 protein showed no significant similarity to any other known proteins or protein classes except GDD1 orthologs in other species. They also reported that the human GDD1 protein showed 79%, 56%, 40%, and 41% identity with mouse, zebrafish, fruit fly, and mosquito orthologs (or homologs), respectively.

However, TMEM16E is identical to GDD1, as mentioned above, and BLAST programs reveal that TMEM16E (GDD1) is homologous to other members of the TMEM16 family, such as TMEM16F, TMEM16A, TMEM16B, TMEM16C, and TMEM16D. Human TMEM16E (GDD1) shows 50.3% total amino acid identity with human TMEM16F (Katoh and Katoh 2004*b*). Human TMEM16E (GDD1) is more homologous to human TMEM16F than to fruit fly or mosquito Tmem16e homologs.

Cys 356 of TMEM16E (GDD1) is substituted with Arg or Gly in patients with gnathodiaphyseal dysplasia (Tsutsumi et al. 2004). Because Cys 356 is conserved among members of the TMEM16 family (Katoh and Katoh 2004*b*), we can now predict that Cys residues might also be essential for the biological function of members of the TMEM16 family. In the postgenome era, comprehensive identification of related genes within the human genome is important for the progression of genome science and medical science.

MASUKO KATOH[1] AND MASARU KATOH[2]
[1]*M & M Medical BioInformatics, Narashino, Japan; and* [2]*Genetics and Cell Biology Section, Genetics Division, National Cancer Center Research Institute, Tokyo*

Electronic-Database Information

The URLs for data presented herein are as follows:

HUGO Gene Nomenclature Committee, http://www.gene.ucl.ac.uk/nomenclature/

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for gnathodiaphyseal dysplasia, *TMEM16E,* and *TMEM16F*)

References

Katoh M, Katoh M (2003) *FLJ10261* gene, located within the *CCND1-EMS1* locus on human chromosome 11q13, encodes the eight-transmembrane protein homologous to *C12orf3, C11orf25* and *FLJ34272* gene products. Int J Oncol 22:1375–1381

——— (2004*a*) Identification and characterization of human *TP53I5* and mouse *Tp53i5* genes *in silico*. Int J Oncol 25: 225–230

——— (2004*b*) Identification and characterization of *TMEM16E* and *TMEM16F* genes *in silico*. Int J Oncol 24: 1345–1349

Schwab M (1998) Amplification of oncogenes in human cancer cells. Bioessays 20:473–479

Tsutsumi S, Kamata N, Vokes TJ, Maruoka Y, Nakakuki K, Enomoto S, Omura K, Amagasa T, Nagayama M, Saito-Ohara F, Inazawa J, Moritani M, Yamaoka T, Inoue H, Itakura M (2004) The novel gene encoding a putative transmembrane protein is mutated in gnathodiaphyseal dysplasia (GDD). Am J Hum Genet 74:1255–1261

Address for correspondence and reprints: Dr. Masaru Katoh, Genetics and Cell Biology Section, National Cancer Center Research Institute, Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan. E-mail: mkatoh@seaple.ne.jp

0002-9297/2004/7505-0023$15.00

Am. J. Hum. Genet. 75:928–929, 2004

Response to Katoh and Katoh

To the Editor:

Gnathodiaphyseal dysplasia (GDD [MIM 166260]) is a syndrome characterized by bone fragility, sclerosis of tubular bone, and cemento-osseous lesions of the jawbone. We have mapped the GDD locus to an 8.7-cM interval on chromosome 11p14.3-15.1 by linkage analysis of a Japanese family with GDD (Tsutsumi et al. 2003). We studied a cDNA (GenBank accession number AL833271) in the candidate region, and, early in 2003, we detected a missense mutation (C356R [MIM 608662.0001]) in the affected members of the family. The gene was named "*GDD1*" (MIM 608662) and was submitted to the National Center for Biotechnology Information (NCBI) database on October 28, 2003 (GenBank accession number AB125267). The mouse homolog was also cloned and was submitted to the NCBI database on November 4, 2003 (GenBank accession number AB125740). We found cellular localization of the GDD1 protein to the endoplasmic reticulum, as well as another missense mutation (C356G [MIM 608662.0002]), in the affected members of an African American family with GDD. Overexpression of *GDD1* genes with both of the mutations found in the patients with GDD dramatically changed the cellular characteristics. Our study containing these results was electronically published in *The American Journal of Human Genetics* on April 29, 2004 (Tsutsumi et al. 2004).

On the other hand, Katoh and Katoh independently reported that they had found, through an *in silico* anal-

ysis, that *FLJ10261, C12orf3, C11orf25,* and *FLJ34272* are genes that share a structural homology (Katoh and Katoh 2003). They then reported the other members of the gene family (GenBank accession numbers AL833271 and AL832340) in the May 2004 issue of *The International Journal of Oncology* (Katoh and Katoh 2004*b*). In that study, they named these genes "*TMEM16A*"–"*TMEM16F*" (MIM 608663). They claimed to name AL833271 "*TMEM16E,*" whereas we designated it "*GDD1.*"

As they point out in their letter to the editor (Katoh and Katoh 2004*a* [in this issue]), it is now clear that *GDD1* belongs to a structurally related family of eight-transmembrane proteins, although cellular localizations of these genes are not known, except in the human GDD1 protein, which is localized to the endoplasmic reticulum. Functional analysis of these proteins will reveal the mechanisms of pathogenesis of GDD.

NOBUYUKI KAMATA
Department of Oral and Maxillofacial Surgery
Division of Cervico-Gnathostomatology
Graduate School of Biomedical Sciences
Hiroshima University
Hiroshima

Electronic-Database Information

Accession numbers and URLs for data presented herein are as follows:

GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for the cDNA of *GDD1* [accession numbers AL833271 and AB125267], the GDD1 ortholog amino acid sequence of mouse [accession number AB125740], and another member of the gene family [accession number AL832340])

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for GDD, *GDD1* C356R, *GDD1, GDD1* C356G, and *TMEM16A–TMEM16F*)

References

Katoh M, Katoh M (2003) *FLJ10261* gene, located within the *CCND1-EMS1* locus on human chromosome 11q13, encodes the eight-transmembrane protein homologous to *C12orf3, C11orf25* and *FLJ34272* gene products. Int J Oncol 22:1375–1381

——— (2004*a*) GDD1 is identical to TMEM16E, a member of the TMEM16 family. Am J Hum Genet 75:927–928 (in this issue)

——— (2004*b*) Identification and characterization of *TMEM16E* and *TMEM16F* genes *in silico*. Int J Oncol 24: 1345–1349

Tsutsumi S, Kamata N, Maruoka Y, Ando M, Tezuka O, Enomoto S, Omura K, Nagayama M, Kudo E, Moritani M, Yamaoka T, Itakura M (2003) Autosomal dominant gnathodiaphyseal dysplasia maps to chromosome 11p14.3-15.1. J Bone Miner Res 18:413–418

Tsutsumi S, Kamata N, Vokes TJ, Maruoka Y, Nakakuki K, Enomoto S, Omura K, Amagasa T, Nagayama M, Saito-Ohara F, Inazawa J, Moritani M, Yamaoka T, Inoue H, Itakura M (2004) The novel gene encoding a putative transmembrane protein is mutated in gnathodiaphyseal dysplasia (GDD). Am J Hum Genet 74:1255–1261

Address for correspondence and reprints: Dr. Nobuyuki Kamata, Department of Oral and Maxillofacial Surgery, Division of Cervico-Gnathostomatology, Graduate School of Biomedical Sciences, Hiroshima University, Hiroshima 734-8553, Japan. E-mail: nokam@hiroshima-u.ac.jp

0002-9297/2004/7505-0024$15.00