

LETTERS

On two Jewish clades in mitochondrial DNA

European Journal of Human Genetics (2015) **23**, 993–994;
doi:10.1038/ejhg.2014.231; published online 5 November 2014

Recent research by Nogueiro *et al*¹ on mitochondrial DNA full genomic sequences from Northeast Portugal is a welcome addition to Sephardic genetics. We would like to clarify the published history of Sephardic T2e, a haplogroup found in three of their participants, in order to set the record straight. The researchers report on an exclusively Jewish signature within T2e that is characterized by a 9181 G mutation, apparently a new one based on their own observations of publically available sequences, their one new sequence, and PhyloTree. However, the data they draw from, as well as the conclusions they reach, were instead put forth by Bedford *et al*² in an article that is cited but for an unrelated reason.

Nogueiro *et al* state: 'Inside the T2e1 branch, a new sub-haplogroup T2e1b (Supplementary Figure 6), defined by the presence of variant m.9181 G, has been proposed in PhyloTree build 16.' The new branch and labels were not proposed by PhyloTree. They were proposed by Bedford *et al*, who report on 9181 G: 'The present 7 sequences along with the one on Genbank (Sephardic Bulgarian) and the pathological sequence from the medical literature were used to create a phylogenetic tree. The tree is presented in Figure 2. New labels for branches include T2e1b and T2e1b1 and are summarized for both this cluster and the previous 2308 G cluster in Table 2'. (As a historical note, Mannis van Oven, curator of PhyloTree, accepted our proposed branch assignments from Table 2 beginning with build 16; personal communication, email van Oven, 27 September 2013.)

Their article continues: 'T2e1b is supported by 11 complete mtDNA sequences, including the Bragança Jews. Except for two samples from mtDNA Community database, without information concerning their ethnicity, all the remaining nine individuals are Jews, Sephardim or Ashkenazim.' These ethnic observations, however, do not originate with Nogueiro *et al*, but instead from the work of Bedford *et al* who report in the Results section entitled 'The 9181 Cluster': 'For the total of 7 individuals with the defining mutation (kernel plus 6 matches), all 7 reported Jewish ancestry along the deep maternal line. This proved to be in sharp contrast to the 9 individuals without 9181 G, none of whom knew of any Jewish ancestry (Fisher exact test, $P < 0.0001$). Details of the Jewish ancestry among the 9181 G group were 2 Sephardic (The Netherlands, Romania), 4 Ashkenazi (Czech Republic, Lithuania, Poland, unknown) and 1 unknown Jewish.' And again in the Discussion: 'Through the present full genomic sequencing of seven new samples, along with one on Genbank, and one in the medical literature, several striking aspects became apparent 1) All of those harboring 9181 G report Jewish maternal ancestry.'

The 11 complete mtDNA sequences referred to by Nogueiro *et al* include the single Genbank entry available before the study by Bedford *et al*. This was deposited by Behar *et al*³ in connection with their finding that 9181 G within T2e (which they called 'T2f') was a founding lineage for Sephardic Bulgaria. It also includes the seven new

mitochondrial full genomic sequences added to Genbank by Bedford *et al* in connection with their investigation of the mutation, the additional sequence summarized by Bedford *et al* from the medical literature, and the new one by Nogueiro *et al* from their current results, with identical motif to our samples from Sephardic Netherlands, Ashkenazi Poland, and Ashkenazi Czech Republic. (The 11th sequence from a publicly available database was one of our sequences that had not yet appeared on Genbank.)

The explicit *raison d'être* of Bedford *et al* was to investigate two T2e clusters with Sephardic Jewish affiliation, those with the 9181 G mutation and the signature we found earlier, 41C@-16114 T-16192 T-2308 G-15499 T. Thus, the above quoted passages from Bedford *et al* were not incidental or hidden within a different purpose. Nogueiro *et al* were also well acquainted with the details of the article: In an extended email enquiry, they asked about every sequence from our Figure 2 tree of 9181 G sequences, including the matching of submitted Genbank numbers to each node and the geographic locations and ethnicity we reported (8 email messages from Nogueiro to Bedford, 12 replies, 18–27 November 2013).

Near the conclusion of their article, Nogueiro *et al* restate that T2e is found in multiple Jewish groups and offer two alternative explanations: 'Remarkably, for two of the founder lineages (T2e1b and U2e1a), defined by the complete mitochondrial genome, the shared sequences belong to both Sephardic as well as Ashkenazi Jews.' We agree it is remarkable, hence our use of the term 'striking' (for T2e1b) as quoted earlier. They explain as follows: 'Two possible scenarios could accommodate this finding: either the defining variants for each branch could have arisen before the separation between the two Jewish groups; or there may have been recent introgression of Sephardic lineages into Ashkenazim communities in the north of Europe.' Likewise, they were our alternatives: 'The mutation is found in both those of Ashkenazi and of Sephardic maternal origin. Taken together, these findings suggest that this is a surprisingly old clade that may well predate the split between Jewish groups' and 'The present study also leaves open the possibility that the appearance of this mutation in both Sephardic and Ashkenazi Jewish groups is due to a more recent admixture between the Jewish groups rather than predating their split. For instance...'

Concerning our other Sephardic signature, which we labeled T2e1a1a, the work was acknowledged by Nogueiro *et al*. Although one of their three T2e results is a part of the T2e1b branch, their other two sequences fall into this T2e1a1a branch. To clarify potentially ambiguous phrasing about this signature in Nogueiro *et al* concerning the discovery of the back mutation at position 41, observations concerning 41 T/C instability, inferences of the ancestry of the Mexican samples as Sephardic, and establishment of higher branches from which the clade derives, these were all explicit in Bedford⁴ and/or the Bedford *et al* article. For instance, Nogueiro *et al* state, 'Considering the growing number of complete mtDNA sequences available at this time, it was possible to define a new sub-haplogroup T2e1a1, based on m.15499C>T variant.' In case it is unclear, we were the ones who contributed the growing number of sequences as well as explicitly defined T2e1a1 as a nesting structure headed by 15499 T and T2e1a by 2308 G.

The three full genomic sequence additions to T2e by Nogueiro *et al* reveal exciting new aspects of these two previously established Jewish T2e branches. For T2e1a1, they find both the first coding region private mutations and the first nested structure to be reported within this subclade. This suggests to us that estimates of age of this cluster^{2,4}

be pushed back earlier. For T2e1b, they are the first to report its presence in Portugal. This suggests that, should the joint appearance in both Sephardim and Ashkenazim be due to a recent admixture, the direction of gene flow is more likely to have proceeded from Sepherad to Ashkenazi rather than the reverse. This was an issue concerning a recent admixture between the Jewish groups that was previously left unresolved.²

Finally and curiously, our perusal of the eight coding-region mutations described in the text of the article (Haplogroups HV0, N1, T2b, T2e, and U2) finds seven of eight of them to be nonsynonymous mutations, therefore altering the proteins manufactured from the DNA code and RNA translation. If the authors could clarify whether this is a coincidence, a notable finding, or reflects a special clinical population of subjects, it would be helpful for evaluating the representativeness of their results for Sephardim in Northeast Portugal and elsewhere.

Reply to letter from Felice L. Bedford and Doron Yacobi

European Journal of Human Genetics (2015) **23**, 994–995; doi:10.1038/ejhg.2014.232; published online 5 November 2014

The authors of Nogueiro *et al*¹ do thank Bedford and Yacobi comments on our paper providing the opportunity to clarify some issues that may have been put forward in an insufficiently clear or poorly phrased manner.

We must begin by restating the framework in which the paper was worked out; considering that the Iberian Peninsula constitutes the original geographic/historical source of Sephardic populations, the main goal was to characterize complete mitogenomes from self-designated Jews from north-eastern Portugal. The pertinence of the study lays on that the majority of previous studies (Belmonte² and Mallorca³ excluded) were based in descents from exiled Sephardic communities with a supposed Iberian origin rather than those who stayed in Iberia and constitute the remains of the original Sephardic population.

Accordingly, all previous works reporting maternal Sephardic lineages were addressed including those published by Bedford⁴ and Bedford *et al*.⁵

The clades of all haplotypes in Nogueiro *et al*¹ work were assigned according to PhyloTree built 16.⁶ This nomenclature was adopted during the revision process of the manuscript (see revision track record of Nogueiro *et al*¹ work), since the last version of the PhyloTree was launched on the 19 February 2014. At no moment, we intended to credit to us the authorship of the definition of any new branch of the mitochondrial DNA (mtDNA) phylogenetic tree, as clearly stated in the text.

The purpose of the phylogenetic tree of global human mtDNA variation and haplogroup nomenclature (www.phylotree.org)

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Felice L Bedford* and Doron Yacobi

Department of Psychology, Cognitive Science, University of Arizona, Tucson, AZ, USA

E-mail: Bedford@u.arizona.edu

1 Nogueiro I, Teixeira J, Amorim A, Gusmão L, Alvarez L. Echoes from Sepharad: signatures on the maternal gene pool of crypto-Jewish descendants. *Eur J Hum Genet* 2015; **23**: 693–699.

2 Bedford FL, Yacobi, Felix G, Garza FM. Clarifying mitochondrial DNA subclades of T2e from Mideast to Mexico. *J Phylogenetics Evol Biol* 2013; **2**: 1–8.

3 Behar DM, Metspalu E, Kivisild T *et al*: Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. *PLoS One* 2008; **3**: e2062.

4 Bedford FL. Sephardic signature in haplogroup T mitochondrial DNA. *Eur J Hum Genet* 2012; **20**: 441–448.

is to provide, as stated by the curators, a framework to the scientific community with update information from novel mitogenomes sequences.⁶ This tool has greatly improved the way researchers can transmit, compare and contribute with their results to the global mtDNA phylogeny knowledge. In this sense, as far as we understand, when a specific PhyloTree built is cited, it is credited to all contributors, as references of mitogenomes upon which a specific branch is based are always quoted.

In consequence, we do not understand the purpose of the Bedford and Yacobi⁷ letter concerning the credit of authorship attribution of the two Jewish clades. We do think to have followed the standard procedures in this matter but we apologize if, by any means that we have not intended, our text can be interpreted as implying an inappropriate authorship attribution.

Regarding the first specific branch addressed in Bedford and Yacobi,⁷ T2e1b (and T2e1b1), as referred in Nogueiro *et al*,¹ was newly included in the PhyloTree build 16 (van Oven and Kayser⁶) based on three quoted mitogenomes (GenBank accession numbers: KF048033, KF577586 and EF556188). In fact, Bedford *et al*⁵ submitted two of these mitogenomes, the remaining one being contributed by Behar *et al*.² Although the claims contained in Bedford and Yacobi⁷ could start a more transversal debate, we consider that in this specific issue no further acknowledgment should be expected as branches of the mitochondrial phylogenetic tree are not sponsored/personal domains. Otherwise, the detailed history of any (sub)haplogroup definition would have to be at least summarised, which in fact, for the one under discussion would require a lengthy and tedious description beginning with Torroni *et al*⁸ and including all the works having contributed for the definition of the branch.

Concerning the complete mitogenomes used to construct the most parsimonious tree of the analysed clades in Nogueiro *et al*¹ (Supplementary Figures 2), all the information regarding accession numbers (GenBank and EMPOP), sample IDs (mtDNA community) and/or direct citations is included. For example, following the corresponding links of any accession number, the information publicly available at GenBank includes: Authors, Title, Journal, Location and Ethnicity, among other details. Whenever ambiguous information was found, we have contacted by email the respective authors for clarification as it happened indeed with