'Infertile' studies on mitochondrial DNA variation in asthenozoospermic Tunisian men

Antonio Salas, Federico Martinón-Torres, Alberto Gómez-Carballa

Abstract

We reviewed five studies undertaken by the same research group on the possible links between mitochondrial DNA (mtDNA) variation and asthenozoospermia, all carried out on Tunisian men. A thorough assessment of these articles reveals that all five studies were carried out on virtually the same cohort of patients, although this information was concealed by the authors. Thus, the results were 'sliced' in order to unjustifiably maximize the number of publications. In addition, a phylogenetic analysis of their data indicates that the reported results are notably incomplete and deficient. Overall, contrary to the original claims, the association of mtDNA variants with asthenozoospermia finds no support on this saga on Tunisian infertile men.

1. Introduction

During the last few decades, variation in the mitochondrial DNA (mtDNA) molecule has been studied in the context of many complex multifactorial diseases [1–9]. In this regard, the search for mtDNA variation related to infertility has also received the attention of a vast body of literature [10–16,36]. Some of the positive findings were critically questioned by others [13,17], who pointed out problems of different nature, including methodological and theoretical misconceptions, as well as important statistical deficiencies [7,18–21]. For instance, the article by Bandelt [13] questioned the problematic findings of Holyoake et al. [22] who claimed that carriers of mtDNA mutations G9055A and G11719A could have their sperm mobility and/or quantity compromised. Bandelt [13] also pointed to population stratification problems in the study of Ruiz-Pesini et al. [23] where the authors found haplogroups H and T significantly more abundant in non-asthenozoospermic than and asthenozoospermic patients, respectively. Gómez-Carballa et al. [24] indicated a lack of evidence for mtDNA variation in infertility, but also showed important deficiencies in the studies of Montiel-Sosa et al. [25] and Feng et al. [26].

Here we reviewed a saga of problematic studies focused on Tunisian patients where several mutations on the mtDNA molecule have been proposed to be associated with asthenozoospermia.

2. Material and methods

We reviewed five articles published by the same research group from Tunisia. These articles were published in 2012 [27], 2013 [10,11], and 2014 [12,16]. Note that one of the 2014 papers [16] was indexed in PubMed using the first names of the authors instead of the family names, but the authors are virtually the same and following the same order in the five publications.

Maximum parsimony trees were built using procedures described previously [28,29]. Phylotree Build 17 (http://www.phylo-tree.org) [30] was used as a reference for haplogroup nomenclature and worldwide phylogeny. PhyloTree and Soares et al. [31] were also used as references for positional mutational rates.

3. Results

3.1. One or five different cohorts?

The articles published in 2012 [27] and 2013 [11] referred to the analysis of 66 patients, of which 32 were normospermic and 34 asthenozoospermic. In the 2013 article [11], the authors also mentioned an age range for patients of 23–57 years. In the other
three articles, the authors referred to a cohort of 64 patients, 33 normospermic and 31 asthenozoospermic, again in the range of 23–57 years of age (as stated in two of these three articles). All the patients in the five publications were from the same geographic origin (Tunisia).

The almost complete overlap in sample sizes, age ranges, and geographic origin, strongly support the hypothesis that the authors used always the same cohort (with a minor difference in sample size). Moreover, variants reported in the five articles overlap to a great extent (see below). It is however surprising that none of the articles mentioned this issue, therefore conveying the impression that in reality five different and independent cohorts were considered – one per article. It is also remarkable that only one article from 2013 [10] cited the 2012 article [27] (although not in regards to the cohorts), while there are no cross-references among the rest of the articles. This is most noteworthy if we consider that the aims of the five articles were exactly the same, and all of them were carried out on Tunisians.

Baklouti-Gargouri et al. also mention in the different publications the use of a control group of fertile men in order to investigate the incidence of their candidate mutations in healthy individuals. The sample size for this control group is 100 in the three initial publications [10,11,27]; it grew to 150 in Baklouti-Gargouri et al. [12], and for some unexplained reason, these controls were omitted in the last publication from 2014 [16]. Curiously, in this latter publication, the authors used their normospermic infertile men as controls. In all five publications, controls were used exclusively to claim that their best candidate causal mutation was not present in their controls; e.g. “This mutation [m.8021A > G] was absent in normospermic patients, suggesting that it could be associated to asthenospermia” [16]. Association tests were never carried out, nor estimates of risk for their candidate variants.

3.2. Five ‘salami slicing’ papers?

In the best cases, the reviewed articles reported only a list of mtDNA variations observed in the studied patients, but never the full list of haplotypes obtained. This is against best practice standards for mtDNA studies, which recommend to report the full results [32]. The information provided by the authors is therefore very limited so it is not possible to fully reconstruct the overlap existing between the different studies and the overall scenario. However, a number of inferences can be made from the available data.

Thus, there are some disconcutting issues in the five Tunisian studies. First, the articles published in 2012 [27] and 2013 [10] targeted COXI, the article from 2014 [12] only COXIII, while the other two articles targeted the COXI, COXII, COXIII, ATP6, ATP8, and CYTB genes (Table S1). For instance, COXI was sequenced in four out of the five studies. The study from 2012 [27] reported only the association of A6375G with asthenozoospermia. In contrast, the study from 2013 [10] reported 21 variants within COXI gene that do not include the previously reported A6375G (note also that their variant G7521A does not really fall within COXI, see below); 20 out of these 21 variants fully overlap with those reported in 2014 [16], when the authors added A6307G to the pool, claiming its novelty and association with the disease. In general, there is extensive overlap of variants between the different studies, even in those with low or very low mutation rate (i.e. A9425G and A9390G, with no hit either in PhyloTree or in Soares et al. [31]; see Table S2), adding definitive support to the hypothesis that the authors used virtually the same cohort. The few exceptions would be the novel and allegedly pathogenic mutations that were reported in exclusive in each paper.

3.3. ‘Infertile’ mitochondrial DNA data

In the five publications, the authors reported a total absence of their candidate mutations in controls, contrasting with the high frequency of these mutations in their patients, ranging from 8.8% to 100% (mutation G9588A in the 2014 article [16]). Although the authors never estimated the risk associated to these mutations, such risk would be so exceptionally high in the context of a multifactorial disease that it has to be called into question. If we accept the likely hypothesis that the authors used virtually the same cohort for the five publications, we can formulate further puzzling questions. For instance, if G9588A appears in all the patients in the 2014 paper [12], why did this mutation not show up in the other studies where the same region (COXIII) was analyzed [11,16]? The same question applies to other mutations reported in their papers. In an alternative, very unlikely scenario, assuming five different cohorts instead of one, it would still be very odd to have failed to observe these mutations in previous patients before finding them in 100% of the patients in the 2014 cohort.

In addition, the data reported by these authors contain many inconsistencies. To mention some examples, Table 2 in the 2014 study [16] shows the number of ‘polymorphisms SNPs in the asthenozoospermic and normozoospermic infertile patients’. The total numbers in this table make no sense (for instance, gene COXI: there are 25 SNPs in asthenozoospermic patients, 0 in normosperms, but 0 in asthenozoospermia and normozoosperms). In the 2013 article [11] polymorphism T9540C is reported twice out of the seven polymorphism listed within COXII; and the same occurs with transitions C14766T and T6221C of Cyb and COX1, respectively, in Table 3 of Siwar et al. [16]. In this table [16], transition T6221C was described as ‘reported in mitomap’ but also, in another row of the same table, as ‘novel’. Interestingly, this transversion had already been reported in two of the previous publications without assigning novelty status to it [10,11]. It should be noted that transition T6221C is a diagnostic position of haplogroup X and it appears in many other haplogroups of the worldwide mtDNA phylogeny (at least nine); this variant appeared dozens of times in the literature before the publication by Baklouti-Gargouri et al.; therefore, T6221C was definitely not novel at the time of these publications. Other examples of novelty misallocation are mutations A8413G, reported in Baklouti-Gargouri et al. [11] (e.g. two hits in previously published paper of Soares et al. [31]), or C6146T, C6296A, and T6614C, as reported in Siwar et al. [16] (e.g. all received one hit in Soares et al. [31]).

There are also a number of errors of misallocation of reference nucleotides in some of the reported polymorphisms. For example, in Fig. 1 of [11] they report a transition T7724C located in COXIIl gene (T-A amino acid change); however, the rCRS nucleotide at position 7724 should be an “A”; in the same Figure, the variant G8248A should have been reported as A8248G in order to be consistent with nomenclature.

Turning to other problematic issues, the primer pairs used by the authors to amplify the Cyb gene (primers m-22F and m-22R) were identical in Baklouti-Gargouri et al. [11] and in Siwar et al. [16] (see their Tables 1); these primers cover a sequence range from position 14856–15978 of the mtDNA molecule, and produce a PCR amplicon of size 1162 bp. Nevertheless, the authors reported two transitions out of this range, namely, C14766T and T14783C, in Siwar et al. [16] but not in Baklouti-Gargouri et al. [11]. Both tables are almost identical (including legend), with only one difference: Table 1 from Siwar et al. [16] does not show the start and end of the PCR products.

In the same line, the primer pairs used to amplify COXI gene in Baklouti-Gargouri et al. [10] cover a sequence range from 5855 to 7315; however, they indicated the presence of G7521A (which is out of the amplicon). Moreover, G7521A is not really located in
COXI gene, as they show in Table 1 of Baklouti-Gargouri et al. [10] and Table 3 of Siwar et al. [16], but in the mt-TD gene (tRNA-Asp). Additionally there are two mutations reported by the authors as novel in Table 3 of Baklouti-Gargouri et al. [11], namely, mutations T7572C and A8312G; these polymorphisms were again incorrectly allocated to a non-coding region, but in fact they fall in the genes that encode for tRNA-Asp and tRNA-Lys respectively.

Finally, based on the known mtDNA phylogeny, there are a number of polymorphisms that should have been found according to the amplicons analyzed by the authors, but which were however not reported in their lists of variants. Figs. 1 and 2 present and attempt at reconstructing the mtDNA phylogeny of the haplotypes inferred from the data reported in Baklouti-Gargouri et al. [11] and Siwar et al. [16], respectively, using the most parsimonious
approach. The phylogenies can only be inferred because the authors never published the complete haplotypes (but only lists of variants). There is no doubt that the existence of some branches can be inferred with full confidence: most of the variants along the branches have a very low mutation rates (75% of them fall below 5 hits in Phylotree); Table S2. Moreover, some mutations appear only once in Phylotree and therefore unambiguously identify one single haplogroup.

According to these phylogenies, the variants that were missed in Baklouti-Gargouri et al. [11] exceed by far the list of published variants (Fig. 1). To give an example, the four polymorphisms that determine the path from the rCRS to haplogroup R0 were altogether missed, although these variants were all reported in the next publication by Siwar et al. [16] (Fig. 2). Another clear example concerns haplogroup L3k1; it is signaled in the data of Baklouti-Gargouri [11] by diagnostic variants A8649G (one hit), T6620G (one hit) and G15314A (leading from L3e′i′k′x4L3k), and A9007T (L3k4L3k1); however, two other diagnostic variants in the same branch were missed, namely, T9467C (L3e′i′k′x4L3k), and G9329A (L3K4L3k1).

There are also numerous instances of missing data in Siwar et al. [16]. For instance, there is a series of unique variants that undeniably led to branch L1b. These variants are C8655T (L2′3′4′6), T6620G (one hit) and G15314A (leading from L3e′i′k′x > L3k), and A9007T (one hit; L3k > L3k1); however, two other diagnostic variants in the same branch were missed, namely, T9467C (L3e′i′k′x > L3k), and G9329A (L3k > L3k1).

In addition, there seem to be some methodological problems affecting not only particular variants but also relatively large sequence segments (even though these segments were supposedly covered by the indicated amplicons). For instance, in the article by Baklouti-Gargouri et al. [11], no variants were reported for the sequence segments 14856–15203 and 15488–15978 (totaling 837 bp; that is, the authors only reported variants within a segment of 285 bp out of an amplicon size of 1162 bp), while in the article from Siwar et al. [16], no variants were noted for the segment 8868–9300 (432 bp).

4. Discussion

Previous studies have already revealed important deficiencies in studies dealing with mtDNA variation and infertile men [13]. Here we examined a series of articles on male infertility that were published by the same research group. A thorough analysis of the articles indicates that the authors used always the same cohort of patients, although this aspect is not mentioned in their
publications. Their findings appear to have been split in order to maximize the number of publications without justification. The series of articles by Baklouti-Gargouri et al. might be interpreted as an example of “salamislicing” papers, a practice that is widely considered unethical by editors and overall by the scientific community.

Analysis of the literature on mtDNA variation in infertility has also revealed an important publication bias, which is not unusual in mtDNA disease studies [33]. The critical review by Bandelt [13] in regards to the Holyoake et al. [22] article revealed only 7 citations (Google Scholar, queried on March 2016), while the Holyoake et al. paper has received 89 citations. The article by Montiel-Sosa et al. [25], critically investigated by Gómez-Carbulla et al. [24], was cited 86 times. The only exception of a paper with negative findings that was highly cited (55 times) was Pereira et al. [34]; however, less than a quarter of these citations were related to infertility, while the rest were justified by the phylogenetic interest of their data in a human population context.

Overall, the few articles claiming an association between mtDNA variation and infertility do not rest of solid grounds. The present study has contributed to argue in this direction, while it constitutes an effort to reveal problems of different nature in the literature on infertility research, including scientific misconduct.

Acknowledgments

F.M-T received funding from the Grant ISCI/INT14/00245 (Cofinanciado FEDER).

Appendix A. Transparency document

Transparency document associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jbierep.2016.08.002.

References

[1] S.V. Baudouin, D. Saunders, W. Tiangyou, J.L. Elson, J. Poynter, A. Pyle, S. Keers, D.M. Turnbull, N. Howell, P.F. Chimney, Mitochondrial DNA and survival after sepsis: a prospective study, Lancet 366 (2005) 2118–2121.
[2] A. Achilli, A. Olivieri, M. Pala, B. Hooshvar Kashani, V. Carossa, U.A. Pereg, F. Gandini, A. Santoro, V. Battaglia, V. Grugni, L. Pereira, J. Goncalves, A. Goios, T. Rocha, A. Amorim, Human mtDNA haplogroups and reduced male fertility: real association or hidden population substructuring, Int. J. Androl. 28 (2005) 241–247.
[3] B.G. Siwar, G. Myriam, B.M. Aff, M.R. Emna, C. Nozha, S. Aifa, F. Fakhfakh, A. R. Leila, Two novel mutations in COII and tRNA(His) mitochondrial genes in azoospermic inferences men, Biochem. Biophys. Res. Commun. 450 (2014) 610–615.
[4] L. Pereira, J. Goncalves, H.-J. Bandelt, Mutation C11994T in the mitochondrial ND4 gene is not a cause of low sperm motility in Portugal, Fertil. Steril. 89 (2008) 738–741.
[5] A. Salas, M. García-Magariños, I. Logan, H.-J. Bandelt, The saga of the mitochondria, Mol. Biol. Rep. 40 (2013) 4705–4712.
[6] A. Salas, Á. Carracedo, V. Macaulay, M.B. Richards, A practical guide to mitochondrial DNA error prevention in clinical, forensic, and population genetics, Hum. Mutat. 30 (2009) 740–759.
[7] A. Salas, M. García-Magariños, I. Logan, H.-J. Bandelt, The exaggerated status of mtDNA variation and infertility do not rest on solid grounds. The present study has contributed to argue in this direction, while it constitutes an effort to reveal problems of different nature in the literature on infertility research, including scientific misconduct.

S. Baklouti-Gargouri, M. Ghorbel, A. Ben Mahmoud, E. Mkaouar-Rebai, M. Cherif, N. Chakroun, A. Sellami, F. Fakhfakh, L. Ammar-Keskes, A novel m.6307A > G mutation in the mitochondrial COXI gene in asthenozoospermic infertile men, Mol. Reprod. Dev. 80 (2013) 581–587.
[8] S. Baklouti-Gargouri, M. Ghorbel, A. Ben Mahmoud, E. Mkaouar-Rebai, M. Cherif, N. Chakroun, A. Sellami, F. Fakhfakh, L. Ammar-Keskes, Mitochondrial DNA mutations and polymorphisms in asthenospermic infertile men, Mol. Biol. Rep. 40 (2013) 4705–4712.
[9] S. Baklouti-Gargouri, M. Ghorbel, A. Ben Mahmoud, E. Mkaouar-Rebai, M. Cherif, N. Chakroun, A. Sellami, F. Fakhfakh, L. Ammar-Keskes, Identification of a novel m.9588G: A missense mutation in the mitochondrial COIII gene in asthenozoospermic Tunisian infertile men, J. Assist. Reprod. Genet. 31 (2014) 2115–2121.
[10] H.-J. Bandelt, Misanalysis gave false association of mtDNA mutations with infertility, Int. J. Androl. 31 (2008) 450–453.
[11] J.F. Montiel-Sosa, J.A. Enriguez, M.J. Lopez-Perez, Research of single mitochondrial nucleotide substitutions in male infertility should consider human mitochondrial haplogroups, Int. J. Androl. 25 (2002) 372–373 (author reply 374).
[12] L. Pereira, J. Goncalves, A. Goios, T. Rocha, A. Amorim, Human mtDNA haplogroups and reduced male fertility: real association or hidden population substructuring, Int. J. Androl. 28 (2005) 241–247.
[13] B.G. Siwar, G. Myriam, B.M. Aff, M.R. Emna, C. Nozha, S. Aifa, F. Fakhfakh, A. R. Leila, Two novel mutations in COII and tRNA(His) mitochondrial genes in azoospermic inferences men, Biochem. Biophys. Res. Commun. 450 (2014) 610–615.
[14] L. Pereira, J. Goncalves, H.-J. Bandelt, Mutation C11994T in the mitochondrial ND4 gene is not a cause of low sperm motility in Portugal, Fertil. Steril. 89 (2008) 738–741.
[15] A. Salas, J.-H. Bandelt, V. Macaulay, M.B. Richards, Phylogeographic investigations: the role of trees in forensic genetics, Forensic Sci. Int. 168 (2007) 1–13.
[16] Y.-G. Yao, Q.-P. Kong, A. Salas, H.-J. Bandelt, Pseudo-mitochondrial genome haunts disease studies, J. Med. Genet. 45 (2008) 769–772.
[17] Y.-G. Yao, A. Salas, I. Logan, H.-J. Bandelt, mtDNA data mining in GenBank needs surveying, Am. J. Hum. Genet. 85 (2009) 929–931 (author reply 933).
[18] A.J. Holyoake, P. McHugh, M. Wu, S. O’Carroll, P. Benn, I.Y. Sin, F.V. Sin. High incidence of single nucleotide substitutions in the mitochondrial genome is associated with poor semen parameters in men, Int. J. Androl. 24 (2001) 175–182.
[19] E. Ruiz-Pesini, A.C. Lapena, C. Diez-Sanchez, A. Perez-Martos, J. Montoya, E. Alvarez, M. Diaz, A. Urrues, I. Montoro, M.J. Lopez-Perez, J.A. Enriguez, Human mtDNA haplogroups associated with high or reduced spermatozoa motility, Ann. J. Hum. Genet. 67 (2000) 682–695.
[20] A. Gómez-Carbulla, J. Pardo-Deo, F. Martinón-Torres, A. Salas, Analyses of mtDNA variation in men do not support a role of mitochondrial DNA variation in male infertility, submitted for publication.
[21] E. Ruiz-Pesini, A.C. Lapena, C. Diez-Sanchez, A. Perez-Martos, J. Montoya, E. Alvarez, M. Diaz, A. Urrues, I. Montoro, M.J. Lopez-Perez, J.A. Enriguez, Human mtDNA haplogroups associated with high or reduced spermatozoa motility, Ann. J. Hum. Genet. 67 (2000) 682–695.
[22] A. Salas, J.-H. Bandelt, V. Macaulay, M.B. Richards, Phylogeographic investigations: the role of trees in forensic genetics, Forensic Sci. Int. 168 (2007) 1–13.
[23] Y.-G. Yao, Q.-P. Kong, A. Salas, H.-J. Bandelt, Pseudo-mitochondrial genome haunts disease studies, J. Med. Genet. 45 (2008) 769–772.
[24] Y.-G. Yao, A. Salas, I. Logan, H.-J. Bandelt, mtDNA data mining in GenBank needs surveying, Am. J. Hum. Genet. 85 (2009) 929–931 (author reply 933).
[25] A.J. Holyoake, P. McHugh, M. Wu, S. O’Carroll, P. Benn, I.Y. Sin, F.V. Sin. High incidence of single nucleotide substitutions in the mitochondrial genome is associated with poor semen parameters in men, Int. J. Androl. 24 (2001) 175–182.
[26] E. Ruiz-Pesini, A.C. Lapena, C. Diez-Sanchez, A. Perez-Martos, J. Montoya, E. Alvarez, M. Diaz, A. Urrues, I. Montoro, M.J. Lopez-Perez, J.A. Enriguez, Human mtDNA haplogroups associated with high or reduced spermatozoa motility, Ann. J. Hum. Genet. 67 (2000) 682–695.
[27] A. Gómez-Carbulla, J. Pardo-Deo, F. Martinón-Torres, A. Salas, Analyses of mtDNA variation in men do not support a role of mitochondrial DNA variation in male infertility, submitted for publication.
misleading studies falsely associating mitochondrial DNA with breast cancer, BMC Cancer 14 (2014) 659.

[34] L. Pereira, J. Goncalves, R. Franco-Duarte, J. Silva, T. Rocha, C. Arnold, M. Richards, V. Macaulay, No evidence for an mtDNA role in sperm motility: data from complete sequencing of asthenozoospermic males, Mol. Biol. Evol. 24 (2007) 868–874.

[35] R.M. Andrews, I. Kubacka, P.F. Chinnery, R.N. Lightowlers, D.M. Turnbull, N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA, Nat. Genet. 23 (1999) 147.

[36] A. Salas, J.L. Elson, Mitochondrial DNA as a risk factor for false positives in case-control association studies, J. Genet. Genomics 42 (4) (2015) 169–172.