Biochemistry and Biophysics Reports 8 (2016) 114–119

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

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

Antonio Salas [a,b,*], Federico Martinón-Torres [b,c], Alberto Gómez-Carballa [a,b]

[a] Unidade de Xenética, Departamento de Anatomía Patolóxica e Ciencias Forenses, Instituto de Ciencias Forenses, Facultade de Medicina, Universidade de Santiago de Compostela, and GenPop Research Group, Instituto de Investigaciones Sanitarias (IDIS), Hospital Clínico Universitario de Santiago, Galicia, Spain
[b] Grupo de Investigación en Genética, Vacunas, Infecciones y Pediatría (GENVIP), Hospital Clínico Universitario and Universidade de Santiago de Compostela, Galicia, Spain
[c] Translational Pediatrics and Infectious Diseases, Department of Pediatrics, Hospital Clínico Universitario de Santiago de Compostela, Galicia, Spain

ARTICLE INFO

Article history:
Received 3 May 2016
Accepted 10 August 2016
Available online 11 August 2016

Keywords:
Infertility
Asthenozoospermic
Mtdna
Phylogeography
Haplogroup

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.phylotree.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

* Corresponding author at: Unidade de Xenética, Departamento de Anatomía Patolóxica e Ciencias Forenses, and Instituto de Ciencias Forenses, Grupo de Medicina Xenómica (GMX), Facultade de Medicina, Universidade de Santiago de Compostela, 15872 Galicia, Spain.
E-mail address: antonio.salas@usc.es (A. Salas).

http://dx.doi.org/10.1016/j.bbrep.2016.08.002

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 disconcerting 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 exceptionably 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 asthenospermic patients, 0 in normospermics, but 0 in asthenospermics + aormospermics). 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 *CytB* and *COXI*, 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, reported as novel 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 *COXII* 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 *Cytb* 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