



ELSEVIER

Contents lists available at ScienceDirect

## Data in Brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)



### Data Article

# Data on the time of integration of the human mitochondrial pseudogenes (NUMTs) into the nuclear genome

Konstantin Gunbin<sup>a,\*</sup>, Leonid Peshkin<sup>c</sup>, Konstantin Popadin<sup>b</sup>, Sofia Annis<sup>e</sup>, Rebecca R. Ackermann<sup>d</sup>, Konstantin Khrapko<sup>e,\*</sup>

<sup>a</sup> Institute of Cytology and Genetics SB RAS and Novosibirsk State University, Russia

<sup>b</sup> University of Lausanne, Center for Integrative Genomics, Switzerland

<sup>c</sup> Harvard Medical School, Boston, USA

<sup>d</sup> Department of Archaeology & Human Evolution Research Institute, University of Cape Town, South Africa

<sup>e</sup> Northeastern University, Boston, USA

### ARTICLE INFO

#### Article history:

Received 21 December 2016

Received in revised form

1 May 2017

Accepted 9 May 2017

Available online 17 May 2017

#### Keywords:

NUMT

Human evolution

Mitochondrial DNA

Pseudogene

Speciation

Punctuated evolution

### ABSTRACT

The data and methods presented in this article are supplementing the research article "Integration of mtDNA pseudogenes into the nuclear genome coincides with speciation of the human genus. A hypothesis", DOI: 10.1016/j.mito.2016.12.001 (Gunbin et al., 2017) [1]. Mitochondrial DNA is known to get inserted into nuclear DNA to form NUMTs, i.e. nuclear DNA pseudogenes of the mtDNA. We present here the sequences of selected NUMTs, in which time of integration can be determined with sufficient precision. We report their chromosomal positions, their position within the great ape mtDNA phylogeny, and their times of integration into the nuclear genome. The methods used to generate the data and to control their quality are also presented. The dataset is made publicly available to enable critical or extended analyzes.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

DOI of original article: <http://dx.doi.org/10.1016/j.mito.2016.12.001>

\* Corresponding authors.

E-mail addresses: [genkvg@gmail.com](mailto:genkvg@gmail.com) (K. Gunbin), [kkhrapko@gmail.com](mailto:kkhrapko@gmail.com) (K. Khrapko).

<http://dx.doi.org/10.1016/j.dib.2017.05.024>

2352-3409/© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## Specifications Table

Subject area	<i>Biology</i>
More specific subject area	<i>Human Evolution</i>
Type of data	<i>text file, table, graph</i>
How data was acquired	<i>Sequence data were acquired from the Genbank database using selection criteria for the sequences that allowed reliable determination of the time of NUMT insertion.</i>
Data format	<i>FASTA, graphics.</i>
Experimental factors	<i>N/A</i>
Experimental features	<i>N/A</i>
Data source location	<i>N/A</i>
Data accessibility	<i>The data are available with this article</i>

## Value of the data

- *The data presents the selected NUMTs in which time of integration can be determined with sufficient precision, their sequences and their chromosomal positions in the human nuclear genome.*
- *The data presents the position of the selected NUMTs within the great ape mtDNA phylogeny (Fig. S4 - available in the Supplementary Material), their times of integration into the nuclear genome.*
- *The dataset enables critical and extended analyzes of the human evolution features revealed by the analysis of these NUMTs [1].*

## 1. Data

The data presented in this article consists of the sequences of the selected NUMTs, which were chosen for being suitable for high resolution phylogenetic analysis (Supplementary Material, Appendix A). Also presented are their chromosomal positions and times of integration (Table S1 - Supplementary Material, Appendix A), the trees representing the relation of the selected NUMTs to the great ape mtDNA phylogeny are presented in Fig. S4 (Supplementary Material, Appendix A). The interrelationships between the NUMTs themselves are presented in Fig. S1. The times of integration of the selected NUMTs into the nuclear genome are presented in Table S1. Methods pertaining to the sources, generation and verification of the data are described in Materials and methods, and Fig. S2 and S3.

## 2. Materials and methods

### 2.1. Data sources (S.1)

78 nuclear mitochondrial DNAs (NUMTs) were obtained from Tsuji et al. [2] (hg18-numts.tsv file). We included NUMTs that were clustered with hominid branches of the primate tree (i.e., f, e, d branches in Fig. S1 of Supplementary Material of Tsuji et al. [2]). We used all 78 pseudogenes, including those flagged as “possible duplications”. 23 NUMTs were obtained from Dayama et al. [3] (sequenced/mapped NUMTs only, i.e. 23 out of 141 detected). Thus overall, we started with 101 human NUMT sequences.

Download English Version:

<https://daneshyari.com/en/article/4764986>

Download Persian Version:

<https://daneshyari.com/article/4764986>

[Daneshyari.com](https://daneshyari.com)