Mitochondrion 66 (2022) 1–6

Contents lists available at ScienceDirect

Mitochondrion

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

Nuclear mitochondrial DNA sequences in the rabbit genome

Bálint Biró [a], Zoltán Gál [a], Giuseppina Schiavo [b], Anisa Ribari [b], Valerio Joe Utzeri [b], Michael Brookman [c], Luca Fontanesi [b], Orsolya Ivett Hoffmann [a,*]

[a] Hungarian University of Agricultural and Life Sciences, Institute of Genetics and Biotechnology, Szent-Györgyi Albert Str. 4, H-2100, Gödöllö, Hungary
[b] University of Bologna, Department of Agricultural and Food Sciences, Division of Animal Sciences, Viale Fanin 46, 40127 Bologna, Italy
[c] Hanze University of Applied Sciences, Department for Biology and Medical Laboratory Research, Zernikeplein 7, 9747 AS Groningen, Netherlands

ARTICLE INFO

Keywords:
Genomes
Mitochondrion
Numt
Numtogenesis
Rabbit

ABSTRACT

Numtogenesis is observable in the mammalian genomes resulting in the integration of mitochondrial segments into the nuclear genomes (numts). To identify numts in rabbit, we aligned mitochondrial and nuclear genomes. Alignment significance threshold was calculated and individual characteristics of numts were analysed. We found 153 numts in the nuclear genome. The GC content of numts were significantly lower than the GC content of their genomic flanking regions or the genome itself. The frequency of three mammalian-wide interspersed repeats were increased in the proximity of numts. The decreased GC content around numts strengthen the theory which supposes a link between DNA structural instability and numt integration.

1. Introduction

1.1. Numts and their significance

Mitochondrial DNA (mtDNA) sequences that are integrated into the nuclear genome (gDNA) are called "nuclear mitochondrial sequences" (numt) and the integration process itself is called numtogenesis (Lopez et al., 1994). The role of numtogenesis was verified in several pathophysiological conditions in which the activation of oncogenes and/or the deactivation of tumour suppressors are the results of numt insertions (Singh et al., 2017; Srinivasainagendra et al., 2017; Palodhi et al., 2020). Beside their cancer research involvement, numts are also important in forensic and phylogenetic studies.

Mitostress (ionization radiation, endotoxins, ROS etc.) is the prerequisite of numtogenesis (Srinivasainagendra et al., 2017). mtDNA fragments can escape from the damaged mitochondrion during not perfect mitophagy, fusion, fission events etc (Hazkani-Covo et al., 2010; Puertas & González-Sánchez, 2020). In the cytoplasm, mtDNA sequences are protected from nucleases due to a vacuole mediated mechanism and/or due to complex formation with DNA-binding-histone-like proteins. These fragments enter the nucleus due to membrane fusion (Puertas & González-Sánchez, 2020). In most of the cases, the integration of mtDNA fragments is facilitated by non-homologous end joining (NHEJ) or microhomology at double stranded breaks (DSB).

1.2. Evolutionary background of numtogenesis

During the early evolution of multicellular organisms, an intracellular cooperation took place between an alphaproteobacterium and an Archea (Roger et al., 2017; Martin et al., 2015). This partnership proved to be beneficial for both participants and thus the evolution of eukaryotes has started. Through the evolution, different subcellular compartments were formed to carry out complex tasks. One of these organelles is the mitochondrion which is responsible mainly for the aerobic respiration. Besides its principal function, the mitochondrion plays important roles in different intracellular pathways (Roger et al., 2017). A very important phenotypic feature of the mitochondrion is that it has its own genome (mtDNA) which was one of the strongest evidences of its endosymbiotic origin (Martin et al., 2015). During the coevolution of symbiont and host, as a phenomenon called endosymbiotic gene transfer (EGT), the symbiont's genome (mtDNA) has been reduced and a decent amount of its genetic material has been localised into the host genome (genomic DNA, gDNA) (Kelly, 2020). The molecular driving mechanism of EGT is explained by Muller's ratchet theory (Muller, 1964). This hypothesis states that in an asexually isolated population or genome (mtDNA in the case of numtogenesis), deleterious mutations are going to occur more frequently and so a steady loss of the genetic material is going to happen. By stochastic fluctuation, an irreversible loss of individuals with the smallest number of deletions (the fittest individuals) can be observed. In short term, this will cause genomic erosion and in

* Corresponding author.
E-mail address: hoffmann.orsolya.ivett@uni-mate.hu (O.I. Hoffmann).

https://doi.org/10.1016/j.mito.2022.07.003
Received 24 March 2022; Received in revised form 23 May 2022; Accepted 10 July 2022
Available online 13 July 2022

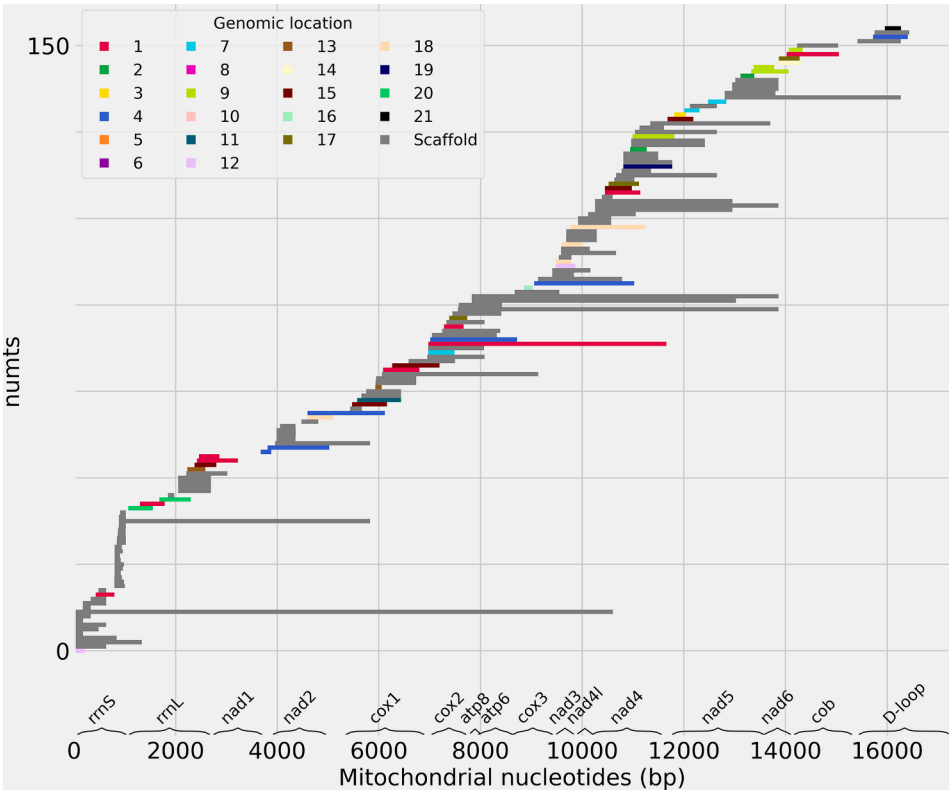

Fig. 1. Distribution of rabbit numts with their genomic locations. Grey horizontal bars represent scaffold localised numts, while the colourful horizontal bars represent chromosome localised numts colour coded by their corresponding chromosomes.

longer term the population or genome will disappear (Metzger & Eule, 2013; Naito & Pawlowska, 2016). mtDNA is in asexual isolation since in majority of the eukaryotes it has uniparental inheritance, lacking recombination (Breton & Stewart, 2015). Therefore, mtDNA is possibly exposed to the Muller's ratchet (Howe & Denver, 2008). Hence EGT is functioning to prevent mtDNA from Muller's ratchet (Martin & Hermann, 1998; Kelly, 2020) by inserting numts from mtDNA into the gDNA.

1.3. Objective

Rabbit became one of the most frequently used model organisms for studying a wide range of human diseases (Esteves et al., 2018). Some studies were published about rabbit numtogenesis in general (Calabrese et al., 2017; Hazkani-Covo et al., 2010) but these papers did not investigate numts deeply. Therefore, in this study our first goal is to characterise the numts of the rabbit genome which can be interesting for the scientific community for modelling cancerous diseases or when dealing with phylogenetic studies.

2. Materials and Methods

Rabbit gDNA (OryCun2.0) and mtDNA were acquired from the Ensembl genome browser (http://ftp.ensembl.org/pub/release-104/fasta/oryctolagus_cuniculus/dna/). Mitochondrial annotation was obtained by using MITOS server (http://mitos.bioinf.uni-leipzig.de/index.py) provided by the University of Leipzig (Bernt et al., 2013) where the Genetic Code: 02 - Vertebrate setting was used. An alignment threshold was calculated to decrease the number of random hits. This threshold was the lowest *e* value of a sequence alignment between gDNA and a simple reversed (but not complemented) mtDNA by using the LASTAL software (v1219) (Kiełbasa et al., 2011). Considering that nucleic acid sequences do not evolve by simple reversal, it is safe to consider any hits between gDNA and a simple reversed mtDNA to occur by chance. This method was proved to be efficient in reducing the number of false positives (Tsuji et al., 2012; Schiavo et al., 2017). Once the threshold was computed, two mtDNA sequences were aligned to the rabbit genome to identify numts even if they contain the linearization point of the mtDNA. The scoring scheme of + 1 for matches, −1 for mismatches, 7 for gap-open penalty and 1 for gap-extension penalty was used for sequence alignment as proposed by Tsuji et al. (2012). These alignments were then filtered based on the threshold calculated previously. Any alignment with e value under the threshold can be considered as numt. RepeatMasker (RM) at the USCS server (https://genome.ucsc.edu/cgi-bin/hgTables) was used to investigate repetitive elements (Smit et al., 1996; Bedell et al., 2000) within 5 kb flanking region of the numts. The settings of RM were the followings, clade Mammalian, genome Rabbit, assembly Apr. 2009 (Broad/oryCun2), group Variation and Repeats, track RepeatMasker. For the genomic examinations each chromosome was randomly sampled as many times as the number of numts per a given chromosome in a length of a given numt. Anderson Darling test was performed to examine normality at 0.05*P* value. Based on the result of normality testing, *t*-test or Wilcoxon signed rank test was performed. Results with lower *P* value than 0.05 were considered as significant results. Statistical calculations were performed in Python Scipy (v1.6.2). To reduce the false positive rate, the relationship between chromosomes and scaffolds were investigated in a preliminary study. Chromosomes and scaffolds were aligned with LASTAL using the same scoring scheme as described above. All the alignments that did not overlap with numts were discarded. Flanking regions were acquired by using SAMTOOLS's (v1.6) faidx function in a length corresponding to the given numt. Then pairwise alignments were performed between chromosome numt and scaffold numt and the corresponding flanking regions with BioPython (v1.78) with the scoring scheme as described above. Repetitive elements were also investigated in the flanking regions of chromosome and scaffold numts with RM. For this preliminary

Table 1
Descriptive statistics of numts in the rabbit genome on the chromosomal level.

Chr	Count	Size mean (bp)	Size SD (bp)	Size median (bp)
1	9	1047.4	1300.5	658
2	2	299	26	299
3	1	–	–	–
4	8	1133.8	537.3	1085.5
7	3	389.3	93.7	363
9	4	553	225.3	558
11	1	–	–	–
12	3	334.3	86.2	368
13	2	238.5	112.5	238.5
41	1	–	–	–
15	5	599.4	162.9	506
16	1	–	–	–
17	3	451.3	104	407
18	4	655.3	458.1	444.5
19	1	–	–	–
20	2	545	75	545
21	1	–	–	–

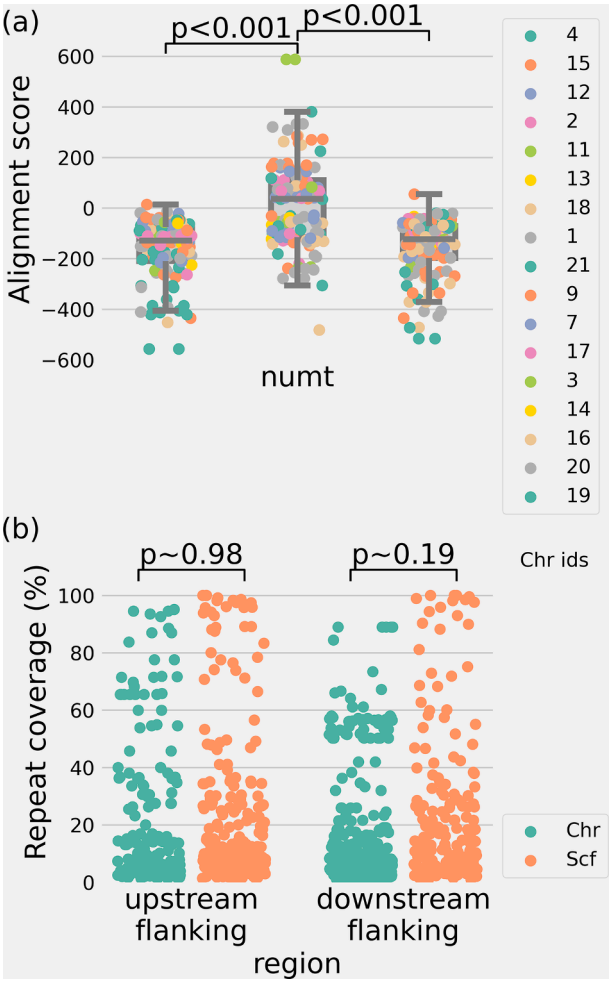

Fig. 2. Investigation of false positives. (a) Distribution of alignment scores of chromosome and scaffold localised numts and their corresponding upstream and downstream flanking regions. Chromosomes are color-coded. (b) The portion of repetitive elements in the upstream and downstream flanking regions of numts relative to the length of the given flanking region. Turquoise dots in the stripplot are representing chromosomes (Chr) while orange dots are representing scaffolds (Scf). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

study RM (v4.1.2-p1) was run locally. All the codes are publicly available at https://github.com/balintbiro/Oryctolagus_cuniculus_numts.

3. Results

3.1. Localisation of numts

50 numts were identified as chromosomal inserts (Fig. 1, Fig. 3, Table 1) while the rest of them were located on scaffolds (Fig. 1-Fig. 2.). No numts were found on the chr5, chr6, chr8. chr10 and chr21. chr3 contained only one numt while chr1 contained 9 of them.

When comparing chromosomes and scaffolds in a preliminary study to reduce false positive rate, it turned out that several numts are located both on chromosomes and scaffolds (Fig. 2a). However, their flanking regions differed significantly (upstream flanking regions vs numts: $p < 0.001$, downstream flanking regions vs numts: $p < 0.001$). No numt with it's corresponding flanking regions was found that would have been the same on chromosomes and scaffolds. The portion of repetitive elements in the flanking regions of numts were also investigated (Fig. 2b). It gave the result that there was no difference between chromosome and scaffold flanking regions in terms of repetitive elements neither in upstream ($p \sim 0.98$) nor in downstream flanking regions ($p \sim 0.19$).

Numt sequence length varied from104 bp to 10608 bp in a total of 194 317 bp. chr1 showed the highest cumulative numt length, whereas chr16 had only 162 bp of numt (Table 1).

Interestingly, when the sum of the numts bp to the length of the actual chromosome were compared, chr4 showed the highest value (~0.009 %), while chr3 had the lowest value with ~ 0.0001 %. No correlation was found between chromosome length and number of numts nor between chromosome lengths and proportion of numts (*P* value > 0.05). 48 intragenic numts were identified. From these 48 genes 36 were novel, yet uncharacterised genes (Fig. 3).

3.2. Repetitive elements around numts

The frequency of repetitive elements was evaluated amongst upstream flanking region, downstream flanking region and the genome (Fig. 4). 7 common repeat classes were found through upstream, downstream flanking and genomic samples namely DNA, SINE, LINE, Low complexity, LTR, Simple repeat and tRNA. From the common repeat classes, three (SINE, LINE and Simple repeat) contained repetitive elements that showed significantly different (*P* value < 0.05) frequency amongst the samples. In SINE repeat class, MIRb (*P* value ~ 0.002) had different frequency in the upstream flanking regions than in the downstream flanking regions of numts. MIR (*P* value ~ 0.007) and MIR3 (*P* value ~ 0.006) frequencies were different when upstream flanking regions with the genome was compared. MIR3 (*P* value ~ 0.003) had also different frequency when downstream flanking regions with the genome was examined. In LINE repeat class, L1M5 (*P* value ~ 0.04) frequency differed in the downstream flanking regions from the genome. In Simple repeat class, the frequency of (TG)n oligomers (*P* value ~ 0.04) were different when the flanking regions with the genome was compared.

3.3. GC content of numts and their flanking regions

The GC contents of the genomic samples and numts were compared (Fig. 5). Numts have significantly lower GC content (median ~ 0.35) than the genome (median ~ 0.43). The GC content of numt's flanking regions and the GC content of the genomic samples were also compared. The flanking regions have significantly lower GC content (median ~ 0.4) than the genome (median ~ 0.45).

4. Conclusions

In this paper we described the patterns of numts in the rabbit genome. Not surprisingly we found that numts are also present in the

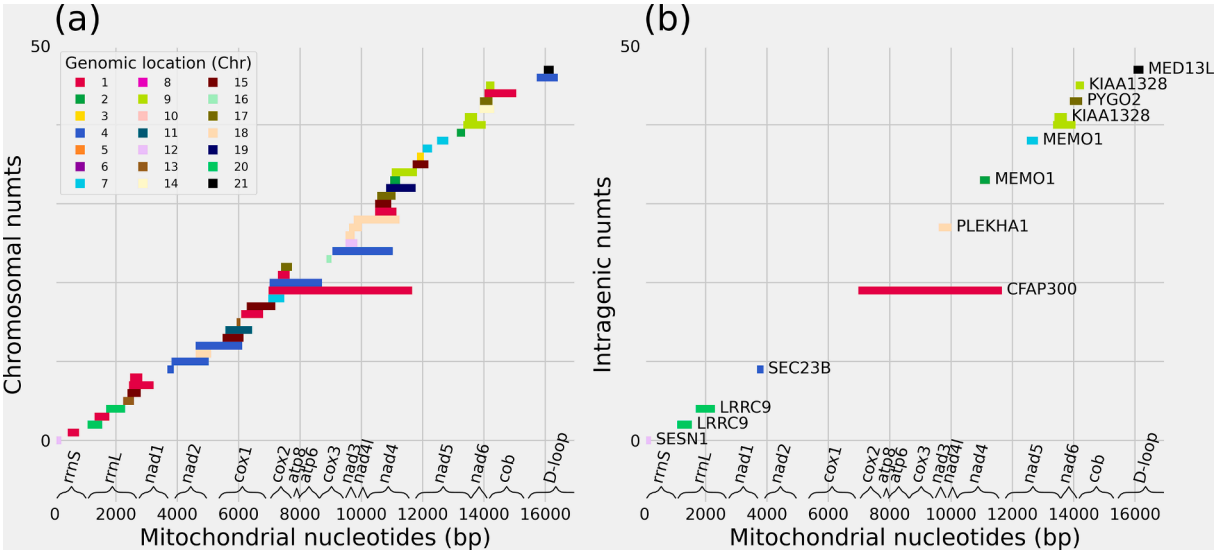

Fig. 3. Chromosomal and intragenic numts. (a) Chromosome located numts with their corresponding chromosomes coded by colours. (b) Chromosome located, intragenic numts with their corresponding chromosomes coded by colours and corresponding external gene names.

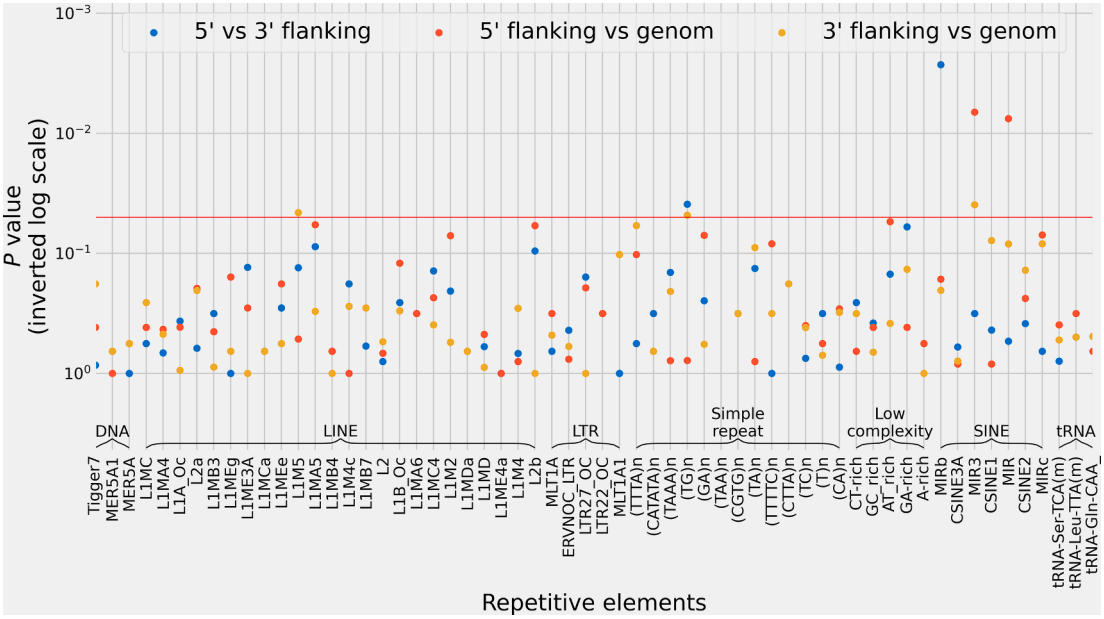

Fig. 4. Analysis of repetitive elements in the flanking regions of numts. Horizontal red line represents the 0.05P value significance threshold, points above it are considered as significant results. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

rabbit nuclear genome just like in many other eukaryotic genomes (Calabrese et al., 2017). Our preliminary analysis showed that in several cases the same numt is present on a chromosome and on a scaffold too. This can cause problems if the given scaffold is a part of a given chromosome since it will increase false positive rate. However, we were able to clarify that our findings are not false positives since the flanking regions of numts on chromosomes and scaffolds were not aligned with each other. The reason for the existence of scaffolds and for poor genome assembly sometimes is the strong presence of repetitive elements and so we also investigated this pattern in the flanking regions of the numts on chromosomes and scaffolds to reduce false positives. Our analysis showed that the portion of repetitive elements in the flanking regions did not differ significantly between chromosomes and scaffolds. These two phenomena (flanking regions are not aligned and there is no difference in terms of repetitive elements) are probably the result of a numt copy and paste mechanism. There is no consensus amongst the scientific community about the treatment of scaffolds in numt related research. For example, Grau et al. (2020) discarded all scaffold localised results from the bovine genome. However, Krampis et al. (2006), Shi et al. (2017) and Wang et al. (2020) included scaffold localised numts into their further analysis. Krampis et al. (2006) also explain the high density of numts on scaffolds with numt propagation. The number of the identified numts is in accordance with the high throughput results of Hazkani-Covo et al. (2010) who found ~ 180 numts with a total length of ~ 183 000 bp in the rabbit genome. Our analysis shows that some of the chromosomes do not contain numts (chr5, chr6, chr8. chr10 and chr21) while there are chromosomes (chr1, chr9, chr20) that contain numerous. The uneven distribution of numts on the chromosome level was also reported among others in the European honey bee genome (Behura, 2007), in the bovine genome (Grau et al., 2020) and in several bat genomes (Zhang et al., 2021). Chromosomes without numt integration were identified in the bovine genome (Grau et al., 2020) and in

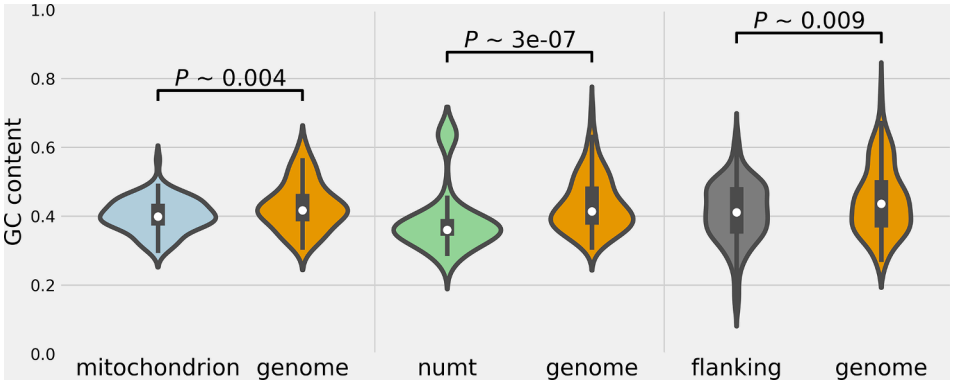

Fig. 5. Analysis of GC contents in the genomic samples and in the flanking regions of numts. Nuclear genome was sampled as it is described in the Materials and Methods section.

the genome of pale spear-nosed bat (Zhang et al., 2021). However, the different structures of the genomes make it quite challenging to compare the exact locations of the numts between species. The majority (two third) of the identified numts are located on scaffolds. We found out that the longest chromosome contains the biggest number of numts, but we found no correlation between the actual chromosome length and the number of numts. This result is in agreement with Behura (2007) who reported ~ 0.01 Spearman correlation coefficient between chromosome length and the number of numts in the genome of the European honey bee, Apis mellifera. As Singh et al. (2017) have pointed out, insertional mutagenesis caused by numt integration can affect the development of different diseases. Many target genes products of the identified intragenic numts are involved in signalling pathways whose malfunctions are associated with developmental abnormalities and malignus transformations (Li et al., 2018; Schotanus & Van Otterloo, 2020; Aprea et al., 2021). The results indicate that the genomic flanking regions of numts and the numts themselves tend to have different (in our cases lower) GC contents than the rest of the genome. Mishmar et al. (2004) also reported this kind of imbalance in the GC content corresponding to human numts. The altered GC content of numts has been also described by Porter & Hajibabaei (2021), Srinivasainagendra et al. (2017), Calderon (2012) and Behura et al. (2011). The preferred numt integration at sites with lower GC content is because these sites are generally poor gene-containing regions and so selection does not clean out numts from these parts of the genome as pointed out by Lascaro et al. (2008) in human. When investigating the phenomenon of lower GC content in numt flanking regions there is a speculative explanation that has to be considered. Based on our findings, it is possible that the chromosomal structure has an influence on numt integrations. GC rich regions tend to have higher stability since GC base pairs have one additional hydrogen bond compared to AT base pairs (Chen & Skylaris, 2021). Therefore, lower GC content is a factor that can indicate destabilised DNA part which is more susceptible to numt integration. Our findings are in accordance with the results of Wang et al. (2020) when it comes to repetitive elements. They found out that repetitive elements are present in a higher concentration around numt insertions. Mishmar et al. (2004) also pointed out that the presence of repetitive elements has a strong influence on numt integration. We observed that from the SINE group, MIR, MIRb and MIR3 had significantly different frequencies around numts. SINEs are non-autonomous Short Interspersed Elements which are non-long terminal repeat (non-LTR) type retrotransposons (Richardson et al., 2015), while MIRs are Mammalian-wide interspersed repeats, which can be found in every mammalian species (Choi et al., 2020). From the LINE RM group, we found L1M5 to have significantly different frequency around numts. LINEs are Long Interspersed Nuclear Elements that are also part of the non-LTR retrotransposons (Bourque et al., 2018). Krampis et al. (2006) and Black & Bernhardt (2009) also reported that retrotransposons can propagate numts. However, Song et al. (2013) goes one step further by suggesting that transposable elements are not just associated with the dispersion of numts but also with the actual integration events themselves.

Further research possibility is to establish a uniformed numt mining pipeline and investigate the nuclear sequences with mitochondrial origins in different organisms.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

B.B. received NTP-NFTÖ-21-B-0127 scholarship. O.I.H. was funded by NKFIH OTKA FK 124708, János Bolyai Research Scholarship of the Hungarian Academy of Sciences and New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund (ÚNKP-21-5).

References

Aprea, I., Raidt, J., Höben, I.M., Loges, N.T., Nöthe-Menchen, T., Pennekamp, P., Olbrich, H., Kaiser, T., Biebach, L., Tüttelmann, F., Horvath, J., Schubert, M., Krallmann, C., Kliesch, S., Omran, H., Dutcher, S.K., 2021. Defects in the cytoplasmic assembly of axonemal dynein arms cause morphological abnormalities and dysmotility in sperm cells leading to male infertility. PLoS Genet. 17 (2), e1009306.
Bedell, J.A., Korf, I., Gish, W., 2000. MaskerAid: a performance enhancement to RepeatMasker. Bioinformatics 16 (11), 1040–1041.
Behura, S.K., 2007. Analysis of nuclear copies of mitochondrial sequences in honeybee (Apis mellifera) genome. Mol. Biol. Evol. 24 (7), 1492–1505.
Behura, S.K., Lobo, N.F., Haas, B., deBruyn, B., Lovin, D.D., Shumway, M.F., Puiu, D., Romero-Severson, J., Nene, V., Severson, D.W., 2011. Complete sequences of mitochondria genomes of Aedes aegypti and Culex quinquefasciatus and comparative analysis of mitochondrial DNA fragments inserted in the nuclear genomes. Insect Biochem. Mol. Biol. 41 (10), 770–777.
Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., Stadler, P.F., 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. 69 (2), 313–319.
Black IV, W.C., Bernhardt, S.A., 2009. Abundant nuclear copies of mitochondrial origin (NUMTs) in the Aedes aegypti genome. Insect Mol. Biol. 18 (6), 705–713.
Bourque, G., Burns, K.H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin, H.L., Macfarlan, T.S., Mager, D.L., Feschotte, C., 2018. Ten things you should know about transposable elements. Genome Biol. 19 (1).
Breton, S., Stewart, D.T., Bonen, L., 2015. Atypical mitochondrial inheritance patterns in eukaryotes. Genome 58 (10), 423–431.
Calabrese, F.M., Balacco, D.L., Preste, R., Diroma, M.A., Forino, R., Ventura, M., Attimonelli, M., 2017. NumtS colonization in mammalian genomes. Sci. Rep. 7 (1), 1–10.
Calderon, I.D.S., 2012. Evolution of nuclear integrations of the mitochondrial genome in Great Apes and their potential as molecular markers. University of New Orleans. Doctoral dissertation.

Chen, H., Skylaris, C.K., 2021. Analysis of DNA interactions and GC content with energy decomposition in large-scale quantum mechanical calculations. PCCP 23 (14), 8891–8899.
Choi, J.D., Del Pinto, L.A., Sutter, N.B., 2020. SINE Retrotransposons Import Polyadenylation Signals to 3'UTRs in Dog (Canis familiaris). bioRxiv.
Esteves, P.J., Abrantes, J., Baldauf, H.-M., BenMohamed, L., Chen, Y., Christensen, N., González-Gallego, J., Giacani, L., Hu, J., Kaplan, G., Keppler, O.T., Knight, K.L., Kong, X.-P., Lanning, D.K., Le Pendu, J., de Matos, A.L., Liu, J., Liu, S., Lopes, A.M., Lu, S., Lukehart, S., Manabe, Y.C., Neves, F., McFadden, G., Pan, R., Peng, X., de Sousa-Pereira, P., Pinheiro, A., Rahman, M., Ruvoën-Clouet, N., Subbian, S., Tuñón, M.J., van der Loo, W., Vaine, M., Via, L.E., Wang, S., Mage, R., 2018. The wide utility of rabbits as models of human diseases. Exp. Mol. Med. 50 (5), 1–10.
Grau, E.T., Charles, M., Féménia, M., Rebours, E., Vaiman, A., Rocha, D., 2020. Survey of mitochondrial sequences integrated into the bovine nuclear genome. Sci. Rep. 10 (1), 1–11.
Hazkani-Covo, E., Zeller, R.M., Martin, W., Malik, H.S., 2010. Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. PLoS Genet. 6 (2), e1000834.
Howe, D.K., Denver, D.R., 2008. Muller's Ratchet and compensatory mutation in Caenorhabditis briggsae mitochondrial genome evolution. BMC Evol. Biol. 8 (1), 1–13.
Kelly, S., 2020. The economics of endosymbiotic gene transfer and the evolution of organellar genomes. bioRxiv.
Kiełbasa, S.M., Wan, R., Sato, K., Horton, P., Frith, M.C., 2011. Adaptive seeds tame genomic sequence comparison. Genome Res. 21 (3), 487–493.
Krampis, K., Tyler, B.M., Boore, J.L., 2006. Extensive variation in nuclear mitochondrial DNA content between the genomes of Phytophthora sojae and Phytophthora ramorum. Mol. Plant Microbe Interact. 19 (12), 1329–1336.
Lascaro, D., Castellana, S., Gasparre, G., Romeo, G., Saccone, C., Attimonelli, M., 2008. The RHNumtS compilation: features and bioinformatics approaches to locate and quantify Human NumtS. BMC Genomics 9 (1), 1–13.
Li, B., Yu, L., Liu, D., Yang, X., Zheng, Y., Gui, Y., Wang, H., 2018. MIB1 mutations reduce Notch signaling activation and contribute to congenital heart disease. Clin. Sci. 132 (23), 2483–2491.
Lopez, J.V., Yuhki, N., Masuda, R., Modi, W., O'Brien, S.J., 1994. Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. J. Mol. Evol. 39 (2), 174–190.
Martin, W.F., Garg, S., Zimorski, V., 2015. Endosymbiotic theories for eukaryote origin. Philos. Trans. R. Soc. B: Biol. Sci. 370 (1678), 20140330.
Martin, W., Herrmann, R.G., 1998. Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol. 118 (1), 9–17.
Metzger, J.J., Eule, S., Wilke, C.O., 2013. Distribution of the fittest individuals and the rate of Muller's ratchet in a model with overlapping generations. PLoS Comput. Biol. 9 (11), e1003303.
Mishmar, D., Ruiz-Pesini, E., Brandon, M., Wallace, D.C., 2004. Mitochondrial DNA-like sequences in the nucleus (NUMTs): Insights into our African origins and the mechanism of foreign DNA integration. Hum. Mutat. 23 (2), 125–133.
Muller, H.J., 1964. The relation of recombination to mutational advance. Mutation Res./Fundam. Mol. Mech. Mutagenesis 1 (1), 2–9.
Naito, M., Pawlowska, T.E., Dubilier, N., Taylor, J.W., 2016. Defying Muller's Ratchet: Ancient heritable endobacteria escape extinction through retention of recombination and genome plasticity. MBio 7 (3).
Palodhi, A., Singla, T., Maitra, A., 2020. Profiling of NUMTs in Gingivobuccal Oral Cancer. bioRxiv.
Porter, T.M., Hajibabaei, M., 2021. Profile hidden Markov model sequence analysis can help remove putative pseudogenes from DNA barcoding and metabarcoding datasets. BMC Bioinf. 22 (1), 1–20.
Puertas, M.J., González-Sánchez, M., 2020. Insertions of mitochondrial DNA into the nucleus—effects and role in cell evolution. Genome 63 (8), 365–374.
Richardson, S.R., Doucet, A.J., Kopera, H.C., Moldovan, J.B., Garcia-Perez, J.L., Moran, J.V., Lambowitz, A., Craig, N., 2015. The influence of LINE-1 and SINE retrotransposons on mammalian genomes. Microbiol. Spectrum 3 (2).
Roger, A.J., Muñoz-Gómez, S.A., Kamikawa, R., 2017. The origin and diversification of mitochondria. Curr. Biol. 27 (21). R1177-R1192.
Schiavo, G., Hoffmann, O. I., Ribani, A., Utzeri, V. J., Ghionda, M. C., Bertolini, F., ... & Fontanesi, L. (2017). A genomic landscape of mitochondrial DNA insertions in the pig nuclear genome provides evolutionary signatures of interspecies admixture. DNA Res., 24(5), 487-498.
Schotanus, M.D., Van Otterloo, E., 2020. Finding MEMO—Emerging Evidence for MEMO1′ s Function in Development and Disease. Genes 11 (11), 1316.
Shi, H., Xing, Y., Mao, X., 2017. The little brown bat nuclear genome contains an entire mitochondrial genome: Real or artifact? Gene 629, 64–67.
Singh, K. K., Choudhury, A. R., & Tiwari, H. K. (2017, December). Numtogenesis as a mechanism for development of cancer. In Seminars in cancer biology (Vol. 47, pp. 101-109). Academic Press.
Smit, AFA, Hubley, R & Green, P. (1996) RepeatMasker Open-4.0. (repeatmasker.org).
Song, S., Jiang, F., Yuan, J., Guo, W., Miao, Y., 2013. Exceptionally high cumulative percentage of NUMTs originating from linear mitochondrial DNA molecules in the Hydra magnipapillata genome. BMC Genomics 14 (1), 1–13.
Srinivasainagendra, V., Sandel, M.W., Singh, B., Sundaresan, A., Mooga, V.P., Bajpai, P., Singh, K.K., 2017. Migration of mitochondrial DNA in the nuclear genome of colorectal adenocarcinoma. Genome Med. 9 (1), 1–15.
Tsuji, J., Frith, M.C., Tomii, K., Horton, P., 2012. Mammalian NUMT insertion is non-random. Nucleic Acids Res. 40 (18), 9073–9088.
Wang, J.X., Liu, J., Miao, Y.H., Huang, D.W., Xiao, J.H., 2020. Tracking the distribution and burst of nuclear mitochondrial DNA sequences (NUMTs) in Fig Wasp Genomes. Insects 11 (10), 680.
Zhang, G., Geng, D., Guo, Q., Liu, W., Li, S., Gao, W., ... & Niu, H., 2021. Genomic landscape of mitochondrial DNA insertions in 23 bat genomes: characteristics, loci, phylogeny, and polymorphism. Integrative Zoology.