Available online at www.sciencedirect.com
ScienceDirect
Gene 401 (2007) 61–70
GENE
www.elsevier.com/locate/gene

Extensive mitochondrial DNA transfer in a rapidly evolving rodent has been mediated by independent insertion events and by duplications

Deborah A. Triant *, J. Andrew DeWoody

Department of Forestry & Natural Resources, Purdue University, West Lafayette, IN 47907, USA

Received 12 January 2007; received in revised form 14 June 2007; accepted 2 July 2007
Available online 14 July 2007
Received by I. King Jordan

Abstract

Mitochondrial DNA translocations to the nucleus (numt pseudogenes) are pervasive among eukaryotes, but copy number within the nuclear genome varies widely among taxa. As an increasing number of genomes are sequenced in their entirety, the origins, transfer mechanisms and insertion sites of numts are slowly being characterized. We investigated mitochondrial transfers within a genetically diverse rodent lineage and here report 15 numts totaling 21.8 kb that are harbored within the nuclear genome of the vole *Microtus rossiaemeridionalis*. The 15 numts total 21.8 kb and range from 0.39 to over 3.0 kb in length. Phylogenetic analyses revealed that these numts resulted from three independent insertions to the nucleus, two of which were followed by subsequent nuclear duplication events. The dates of the two translocations that led to subsequent duplications were estimated at 1.97 and 1.19 MYA, which coincide with the origin and radiation of the genus *Microtus*. Numt sequence data from five *Microtus* species were used to estimate an average rate of nucleotide substitution as 2.6×10^{-8} subs/site/yr. This substitution rate is higher than in many other mammals, but is concordant with the elevated rate of mtDNA substitution in this lineage. Our data suggest that numt translocation in *Microtus* is more extensive than in either *Mus* or in *Rattus*, consistent with the elevated rate of speciation, karyotypic rearrangement, and mitochondrial DNA evolution in *Microtus*.

Keywords: Cytochrome *b*; *Microtus*; Mitochondrial genome; Substitution rate; Vole

1. Introduction

The evolution of eukaryotic genomes has been heavily influenced by the integration of DNA fragments such as mobile elements and pseudogenes into the nuclear genome (Mighell et al., 2000; Kazazian, 2004). While the mechanisms and possible functional roles for mobile elements and pseudogenes of nuclear origin have been investigated (Balakirev and Ayala, 2003; Deininger et al., 2003), less is known about pseudogenes that have been transferred to the nucleus from the mitochondrial genome (numts, Lopez et al., 1994). Numt pseudogenes have been reported in many plants and animals, typically found serendipitously during studies of mitochondrial DNA (mtDNA) (Bensasson et al., 2001). Various models of numt integration have been postulated (Ricchetti et al., 1999; Willett-Brozick et al., 2001) but because of differences in genetic codes, most mitochondrial transfers are thought to lose their original function immediately upon insertion into the nuclear genome. Numts are not uniformly distributed across eukaryotic chromosomes, and to date no general trends have been identified that predict their insertion sites.

Although it is not yet understood how mtDNA escapes from the inner confines of the mitochondria, several theories have been proposed. Foremost among these include the degradation of male mitochondria following fertilization and the subsequent

Abbreviations: A, adenine; C, cytosine; G, guanine; T, thymine; F84, Felsenstein's 1984 model; GTR, general time reversible; K2P, Kimura 2-parameter; mtDNA, mitochondrial DNA; numt, nuclear mitochondrial translocation; rRNA, ribosomal RNA; TBR, tree-bisection-reconnection; Ts, transition; Tv, transversion.

* Corresponding author. Current address: Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA. Tel.: +1 225 578 2688; fax: +1 225 578 2597.
E-mail address: datriant@hotmail.com (D.A. Triant).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.07.003

integration of paternal mtDNA fragments into the nascent embryonic nucleus (Woischnik and Moraes, 2002). Alternatively (or in addition to), transient breaches in the mitochondrial membrane during cell stress or variable environmental conditions may provide opportunities for mtDNA transfer (Thorsness and Weber, 1996). Because numts often span multiple mitochondrial genes and/or contain the non-coding control region, their integration into the nuclear genome is probably mediated by DNA as opposed to RNA (Zhang and Hewitt, 1996).

Analyses of human numts have revealed that the integration of mitochondrial fragments is a continuous evolutionary process. However, there have been conflicting reports as to whether most numt integrations are the result of independent insertion events or duplications of pre-existing numts (Mourier et al., 2001; Hazkani-Covo et al., 2003). Most numts contain partial fragments of mitochondrial genes, but there have been reports of numts that represent >50% of the mitochondrial genome such as those found in cats (Kim et al., 2006), plants (Stupar et al., 2001) and humans (Mourier et al., 2001). Richly and Leister (2004) assessed numts within 13 eukaryotic nuclear genomes and found high levels of interspecific variation in numt content. Although the mean size of numts across taxa was <600 bp, the overall proportion of the mitochondrial genome that had been transferred varied from <1% to 99%.

Many of these genome-wide analyses are likely to underestimate the number of numt pseudogenes because ancient numts that have been accumulating nucleotide changes over time should eventually become unrecognizable as mitochondrially-derived sequences. On the other hand, more recently derived numts can be virtually identical to mtDNA and thus are especially problematic in molecular studies because of their potential to contaminate mitochondrial datasets (Arctander, 1995; Triant and DeWoody, 2007). Genome-wide assessments of numts across diverse taxa may reveal broad trends in their evolution (e.g., preferences among insertion sites), but for non-model organisms without sequenced genomes, these types of assessments are intractable.

We have recently described a 4 kb numt (*Mr_numt*) found within a Eurasian vole, *Microtus rossiaemeridionalis* (Triant and DeWoody, in press) and other numts have been identified in related arvicoline taxa (DeWoody et al., 1999; Jaarola et al., 2004). Voles of the genus *Microtus* (rodent subfamily Arvicolinae) have undergone a rapid radiation resulting in more than 60 extant species in less than 2 million years (Repenning, 1980; Carleton and Musser, 2005). Their rates of karyotypic evolution are among the fastest known for mammals (Maruyama and Imai, 1981; Modi, 1987) and mtDNA nucleotide substitution rates are also elevated (Triant and DeWoody, 2006). To gain further insight into mitochondrial transfers that have occurred within this group, we have isolated and characterized more than a dozen numts in *M. rossiaemeridionalis*. Herein, we describe structural components of these numts, perform comparative sequence analyses of the numts with their mitochondrial counterparts, use phylogenetic analyses to examine their evolutionary history, assess rates of nucleotide substitution in the nuclear genome, and estimate dates of translocation from the mitochondrion to the nucleus.

2. Materials and methods

2.1. Numt isolation

For numt isolation we used three different methods: 1) primer-walking; 2) genomic library construction and 3) cloning of cytochrome *b* fragments. We used a Genome Walker kit (Clonetech) that employs a primer-walking approach to identify unknown sequences flanking a 4 kb numt (*Mr_numt*) in *M. rossiaemeridionalis* (Triant and DeWoody, in press). As we primer-walked along *Mr_numt*, we serendipitously isolated several additional numt fragments, each defined by a unique DNA sequence. These opportune discoveries prompted us to expand our search for numts within the same individual using tissue obtained from Texas Tech University (TK 44630). We did so by constructing a genomic library and screening it with *Mr_numt* as a probe, using the library protocols of Hamilton et al. (1999) as modified by Williams and DeWoody (2004). Briefly, genomic DNA from a single vole was digested with *NheI* and *XmnI*, then dephosphorylated and ligated to double-stranded SNX linkers (Hamilton et al., 1999). We used the biotinylated primer PcytbR and non-biotinylated primer PsdTotal_F (Table 1) to amplify a ~3 kb portion of the original 4 kb numt (Triant and DeWoody, in press). This amplicon was hybridized to the linker-ligated products according to the annealing program of Williams and DeWoody (2004). Kilobase Binder Dynabeads (Dynal) were used during the washes of the hybridized, numt-enriched DNA and positive templates were subsequently amplified using SNX primers according to Hamilton et al. (1999), substituting a 6 min extension time and a final extension of 30 min to help ensure complete amplification of long numts. PCR products were ligated into a pBSII SK(+) vector (Stratagene), cloned, and sequenced in both directions with Big Dye Terminator version 3.1.

Because cytochrome *b* is one of the most widely used mtDNA markers, we were especially interested in identifying numts that encompassed the cytochrome *b* gene. To this end, we used *Microtus* cytochrome *b* numt-specific primers PcytbF, PcytbF$_2$, PcytbR, PcytbR$_2$, Cleth-F$_2$, M.greg-F$_2$ with associated PCR and sequencing protocols (Triant and DeWoody, in press). PCR

Table 1
Primer sequences used in this study

Primer	Sequence (5′–3′)
PsdTotal_F	GATCGGCTGAGCCGCTGCAAACACAGCA
PcytbF	ATGACAATCATCTGGGGGGA
PcytbF$_2$	CTCTCTACTGGGCCTATGCT
PcytbR	GATTGGTATGAAGATTATGATAAT
PcytbR$_2$	TGATAATGGCGAAGTAGCCG
Pcytbseq1	TTCAGTAGACAAAGTCACTC
Pcytbseq2	GGAATAGTAGGAGAACTAAT
Cleth-F$_2$	TTATTCCTAGCTATACACTAT
M.greg-F$_2$	TACACTATACATCAGATACAATC
Ins_F	TCATCCGAAAAAATACCCAT
Ins_F$_2$	TCATCAATTACTCATTCATT
Ins_R	ACAATAATGACAAAATAGCT
Ins_R$_2$	CAATAGGTATGAAGATTACA
Ins_Up_Seq-R	CCAGAATCCTACTTTATTGG