Animal Behaviour 82 (2011) 601–606

Contents lists available at ScienceDirect

Animal Behaviour

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

Commentary

Molecular approaches in behavioural research: a cautionary note regarding mitochondrial transfers to the nucleus (numts)

Deborah A. Triant [a,*], Loren D. Hayes [b,1]

[a] Department of Biology, University of Virginia
[b] Department of Biology, University of Louisiana at Monroe

ARTICLE INFO

Article history:
Received 2 February 2011
Initial acceptance 22 March 2011
Final acceptance 27 May 2011
Available online 14 July 2011
MS. number: AS-11-00109R

Keywords
cytochrome *b*
Microtus
mitochondrial genome
numt
pseudogene

The focus of animal behaviour research has changed dramatically in recent years, rapidly moving towards integrative lines of questioning (Hughes 1998; Robinson 1999; Young & Wang 2004; Owens 2006; Blumstein et al. 2010). Within the last 25 years, behavioural ecologists have increasingly incorporated genetic methods into their research, broadening the scope of animal behaviour research and allowing researchers to test predictions in ways that were once not possible. Advances in genetic research have spurred transdisciplinary theory that has redefined behavioural ecology (e.g. sociogenomics, Robinson 1999). Modern molecular tools allow us to determine relatedness and parentage in natural populations, leading to advances in theory such as kin selection and sexual selection (Hughes 1998). Genetic techniques have been used to show that some birds and mammals engage in extrapair matings, a discovery that enabled researchers to focus their inquiry on both social and genetic mating systems (Birkhead et al. 1990; Hughes 1998). Molecular genetics have made it possible to directly measure a wide range of behaviours (e.g. dispersal: Girman et al. 1997; Winters & Waser 2003) and to determine fitness in natural populations (e.g. Griffin et al. 2003; Mabry et al. 2011), allowing behavioural research to move beyond indirect measures. Phylogenetic approaches that rely on genetic markers have been particularly useful and have altered our perception of the evolution of numerous behaviours (Harrison 1989; Hughes 1998).

Molecular approaches are now standard components in several lines of animal behaviour research. The next generation of animal behaviour researchers will be expected to develop a molecular 'toolkit', facilitating research that integrates proximate and ultimate levels of analysis (Robinson 1999; Owens 2006; Blumstein et al. 2010). It is important that these researchers are made aware of the potential pitfalls of genetic methods, testing the assumptions of these methods in a manner similar to the testing of assumptions of statistical tests. Otherwise, researchers may inadvertently set back rather than advance animal behaviour. We briefly review some of the exciting developments in behavioural research made possible by mitochondrial DNA (mtDNA) methods. We then discuss a risk to mtDNA work, the possibility for mtDNA to transfer to the nucleus. Given the paucity of behavioural papers that consider mtDNA transfers, we conclude this paper with recommendations for the consideration of mtDNA transfers in future animal behaviour research. Our aim is to highlight the potential for increasing the understanding of behaviour using mtDNA methods while making a cautionary note about treatment of mtDNA data. We hope that our discussion and recommendations will bring attention to the benefits and challenges of integrating behavioural and genetic

* Correspondence: D. A. Triant, Department of Biology, University of Virginia, Charlottesville, VA 22904, U.S.A.
E-mail address: dtriant@virginia.edu (D. A. Triant).
[1] L. D. Hayes is at the Department of Biology, University of Louisiana at Monroe, Monroe, LA 71209, U.S.A.

doi:10.1016/j.anbehav.2011.06.011

fields and ultimately, promote the strongest, integrative research possible.

mtDNA in Animal Behaviour Research

Mitochondrial DNA has been widely used in many animal taxa as a molecular marker because of its pattern of maternal inheritance, effective haploidy and rapid rate of nucleotide substitution relative to the nuclear genome (Harrison 1989; Avise 1991). Additionally, cells contain hundreds of copies of the mtDNA genome, making it relatively easy to extract from tissue. Behavioural researchers have used mtDNA methods to answer evolutionary questions about numerous behaviours including aggression (Thierry et al. 2008), brood parasitism (Payne et al. 2000; Peer et al. 2007; Langmore et al. 2009), burrowing (Weber & Hoekstra 2009), cooperation (Giray et al. 2000; Mitani et al. 2000), communication and recognition (e.g. Safi & Kerth 2003; Foitzik et al. 2007), dispersal (Fredsted et al. 2007), habitat exploration (Mettke-Hofmann et al. 2004) and mate choice (e.g. Wilkinson et al. 1998; Price & Lanyon 2004; Reudink et al. 2006; Hebets & Vink 2007; Guerra & Ron 2008; Ward & McLennan 2009). Mitochondrial DNA has also been used to determine patterns of hybridization (e.g. Tynkkynen et al. 2008), estimate parentage (e.g. Matsuura et al. 2002; Reudink et al. 2006) and establish social organization and kinship (e.g. Goodisman & Ross 1999; Mitani et al. 2000; Thierry et al. 2000; Yurk et al. 2002; Fabiani et al. 2006; Langmore et al. 2007). Additionally, mtDNA methods are often used in the construction of phylogenies for comparative analyses (e.g. Ron 2008), a method that increases the power to answer questions about speciation and historical origins of behaviour (e.g. Crespi et al. 1998). Finally, researchers have used mtDNA methods to determine geographical sources of behavioural variation and potential speciation events (e.g. Pröhl et al. 2007; Guerra & Ron 2008). Such phylogeographical studies will become more important components of behavioural ecology research as we learn more about nontraditional study organisms.

Numts

The applications of mtDNA to behavioural research will continue to grow as more researchers study geographical differences in behaviour and use comparative approaches to answer evolutionary questions about understudied taxa. Thus, it is crucial that researchers recognize the assumptions and challenges to mtDNA research, including the potential contamination of mtDNA data sets by nuclear copies of mtDNA sequences (e.g. Fredsted et al. 2007). The transfer of DNA from the mtDNA genome to the nucleus has been an ongoing process throughout the evolutionary history of eukaryotes, and in plants, can also involve the transfer of chloroplast DNA to the nucleus (Leister 2005). Nuclear copies of mtDNA that closely resemble mtDNA genes have been inadvertently isolated during molecular studies and can cause complications when included with legitimate mtDNA sequences.

The discovery of mtDNA sequences in the nucleus was first reported in 1967 (du Buy & Riley 1967). These transfers are now termed 'numts' (nuclear-mitochondrial sequences; Lopez et al. 1994) and have been reported in various taxa including invertebrates (Song et al. 2008; Hlaing et al. 2009; Viljakainen et al. 2010), birds (Allende et al. 2001; Cho et al. 2009), fish (Waters & Wallis 2001; Waters et al. 2010) and mammals (Liu & Zhao 2007; Dubey et al. 2009) as well as yeast (Jacques et al. 2010; Lenglez et al. 2010) and plants (Noutsos et al. 2005; Kleine et al. 2009). The availability of completely sequenced genomes has prompted whole-genome characterization of numts, but patterns of numt abundance appear random among different evolutionary lineages. For example, plants and humans harbour large numbers of numts (Woischnik & Moraes 2002; Sandoval et al. 2004), as do honeybees (Pamilo et al. 2007). Few numts have been reported in fish genomes (Venkatesh et al. 2006), and rodent distributions are inconsistent: mice have numerous numts while rats have few (Richly & Leister 2004; Triant & DeWoody 2007a). Hazkani-Covo et al. (2010) searched 85 sequenced eukaryotic genomes and reported that numt content is strongly correlated to genome size. However, the authors note that results can vary with different genome versions and search stringencies.

Numts can be transferred from any region of the mtDNA genome; therefore, the representation of different mtDNA regions within nuclear genomes is not consistent (Richly & Leister 2004; Triant & DeWoody 2007a). Most numts are less than 1 kb in length, but in humans a single numt was found to have originated from almost the entire mtDNA genome (14 654 nucleotides; Mourier et al. 2001), and Stupar et al. (2001) reported a 620 kb numt insertion in *Arapadopsis*. Most numts are presumed to be nonfunctional and transcriptionally inactive once integrated into the nucleus because of the differences in genetic codes between the mtDNA and nuclear genomes. The means by which mtDNA translocates to the nucleus is not well understood, but it has been proposed that double-strand chromosomal repair mechanisms facilitate their integration (Blanchard & Schmidt 1996; Hazkani-Covo & Covo 2008). Integration of numts typically occurs in noncoding regions of the genome, but it can sometimes occur in regions of high gene density (Erpenbeck et al. 2011) and has been associated with human disease when inserted into functional genes (Chen et al. 2005).

Numts as a Source of Errors

Because the transfer of numts is an ongoing evolutionary process, the degree to which numts correspond to their mtDNA counterparts can vary. Once integrated into the nuclear genome, numts are no longer under the selective constraints of the mtDNA genome and are free to accumulate mutations. Recent transfers may be easily recognizable as mtDNA fragments, while older copies could have acquired enough substitutions to make them difficult to identify. If nuclear copies are amplified along with true mtDNA sequences, the combination of different sequence types with different evolutionary histories will certainly confound downstream analyses. Numts can be particularly problematic for studies involving ancient DNA (den Tex et al. 2010), DNA barcoding (Song et al. 2008) and human disease diagnosis (Yao et al. 2008). Moreover, numts can combine with mtDNA sequences during PCR to produce mtDNA/numt recombinants (Thalmann et al. 2004). Phylogenetic and phylogeographical studies relying on mtDNA haplotypes can be compromised when numts are confused for distinct mtDNA lineages (Dubey et al. 2009). Studies utilizing data mining have found numts that had been deposited into public databases as mtDNA sequences but were later identified as nuclear copies (Hassanin et al. 2010).

Revealing Numts

There are many ways by which numts can be recognized within a mtDNA data set. One simple means of detection is to translate all sequences to ensure that the mtDNA open reading frame is intact. Copies from the nuclear genome can accumulate insertions, deletions, frame-shift mutations and mutations that cause stop codons, which would prematurely disrupt protein production in a functional mtDNA gene (Fig. 1). However, these preventative techniques will not be effective on noncoding portions of the mtDNA genome, such as the control region and ribosomal and transfer RNAs. When