

Letter to the Editor

An Extra Nucleotide Is Not Translated in Mitochondrial ND3 of Some Birds and Turtles

David P. Mindell, Michael D. Sorenson,¹ and Derek E. Dimcheff

Department of Biology and Museum of Zoology, University of Michigan

We have discovered a single “extra” nucleotide in the mitochondrial (mt) NADH dehydrogenase subunit III (ND3) gene in 46 species of birds (fig. 1) at position 174 relative to the published ND3 sequence of the chicken (*Gallus gallus*) (Desjardins and Morais 1990). We used standard methods of DNA isolation, PCR amplification, and automated sequencing (Mindell et al. 1997), and the extra base has been identified in each species based on sequencing of both light and heavy strands. We took requisite precautions against amplification of potential former mitochondrial sequences currently residing in the nuclear genome by examining all electropherograms for (1) double peaks resulting from coamplification of mtDNA and nuclear DNA sequences, (2) unexpected stop codons, (3) mismatches in overlapping sequence for a given taxon in contiguous amplification products using different primers, and (4) frameshifts (see Sorenson and Quinn 1998). The frameshift for mt ND3 reported here was discovered in the course of such examinations; however, we were able to eliminate the possibility that it was occurring in a potential nuclear copy, as none of the other indicators mentioned were present. Sequences with the extra base were unambiguous and well conserved relative to mt ND3 sequences from species without the extra base. To amplify mt ND3, we used primers (L10647: 5'-TTYGAAGCMGCMGCMTGATACTG-3'; H11100: 5'-TCTGCYCAYTCTARKCCTCCYTG-3') targeted to conserved sequence regions and having degenerate sites at variable positions among birds to reduce the possibility that mismatches between the primer and mtDNA sequence would result in the preferential amplification of nuclear sequences potentially present in low copy numbers. Furthermore, overlapping sequence determined using different mt primer sets consistently matched the ND3 region in question. We found this extra base insertion to be missing from the ND3 sequence of 15 additional bird species (fig. 1B). The extra base is not present in any previously published vertebrate ND3 sequences, with the exception of that of an ostrich (see below). Although not reported in chicken mt ND3 sequenced by Desjardins and Morais (1990), the extra base was present in chicken mt ND3 sequenced by us.

If the extra base we found in ND3 were decoded during translation, the reading frame would be shifted one position forward, yielding multiple stop codons in

the downstream portion of ND3 and, in turn, a truncated gene 207 bp (68 amino acids) long, compared with the usual length of 354 bp (117 amino acids) in birds without the extra base. Premature stopping of translation would result in a relaxation of selective constraints on the downstream portion of the gene and the accumulation of nonsynonymous changes and indels in the manner of a pseudogene. If we omit the extra base from those species in which it occurs, however, the original reading frame is maintained without stop codons in all 46 bird sequences. In addition, the amino acid sequence is highly conserved among the diverse bird taxa sampled, regardless of the presence or absence of the extra base, and sequences following the extra base also show the usual pattern of increasing conservation at codon positions 3, 1, and 2, respectively (fig. 2). Indeed, position-by-position nucleotide diversity values for taxa with and without the extra base are highly correlated both before and after the extra base, suggesting similar constraints on sequence evolution in the two groups of taxa and the maintenance of a functional reading frame following the extra base (fig. 2). This comparison of ND3 sequences from birds with and without the extra base strongly suggests that the entire gene remains functional in taxa with the extra base, contrary to the conclusion of Härlid, Janke, and Arnason (1997) that ND3 is prematurely terminated in an ostrich (*Struthio camelus*).

Translational frameshifting denotes a change in reading frame during the translation of mRNA into amino acids. Mechanisms responsible for such changes include both missense changes in which an “incorrect” tRNA is recruited to read a codon and processivity changes involving tRNA slippage or hopping. These changes stem from the conflicting requirements for speed and accuracy during translation (Parker 1989). Translational frameshifting can lead to new functions for existing genes, such as the expression of alternative proteins, alternative enzymatic activities, and autogenous control factors (reviewed in Farabaugh 1996). Translational frameshifting is known from various virus, retrotransposon, bacteria, and yeast genes, as well as from mammalian antizyme genes (Kastelein et al. 1982; Matsufuji et al. 1995; Farabaugh 1996). There have been no reports, however, of these functionally and evolutionarily important mechanisms in the mt genes of plants or animals. Experiments elucidating the mechanisms of translational frameshift have focused on a few model organisms, such as *Escherichia coli* and *Saccharomyces cerevisiae*, and the phylogenetic distribution and evolutionary history of translational frameshifting remain little known.

+1 frameshifting in *S. cerevisiae* Ty1 retrotransposons occurs when a tRNA reads transcript by decoding only two of the three nucleotides within a codon in

¹ Present address: Department of Biology, Boston University.

Key words: translational frameshifting, RNA editing, mitochondrial DNA translation, ND3, molecular evolution in birds.

Address for correspondence and reprints: David P. Mindell, Department of Biology and Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109. E-mail: mindell@umich.edu.

Mol. Biol. Evol. 15(11):1568–1571. 1998

© 1998 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

A) Mitochondrial ND3 for 3 species with and 3 without the extra base

	163	171	*	181	191	201
1 Gallus	CGA UUC UUC CUC	<u>A</u> GUA GCC AUC CUA UUC	C	CUU CUA UUC GAC		
2 Falco	CGA UUC UUC CUC	<u>A</u> GUA GCA AUU UUA UUC	C	CUC CUA UUU GAC		
3 Aythya	CGA UUC UUC CUC	<u>A</u> GUA GCC AUC CUG UUU	C	CUA CUA UUC GAC		
1 Vidua	CGA UUC UUC CU	<u>A</u> GUA GCC AUC CUA UUC	C	CUC CUA UUC GAC		
6 Buteo	CGA UUC UUC CU	<u>C</u> GUA GCC AUC CUA UUC	C	CUA CUA UUC GAU		
7 Scolopax	CGC UUU UUC CU	<u>C</u> GUG GCA AUU UUA UUC	C	CUA CUA UUU GAC		

B) Potential stem-and-loop structures following extra base position

	species with extra base (# bird species/% of birds with extra base)	species without (# bird species/% of birds without)		species with extra base (# bird species/% of birds with extra base)	species without (# bird species/% of birds without)		
1	AGUAG • UUAUC	1-33 (33/72)	34-41 (6/40)	7	CGUGG • UUAUU	-	61 (1/7)
2	AGUAG • • UUAUU	42-51 (9/20)	52 (1/7)	8	AGUAG • UUUUC	-	62 (1/7)
3	AGUAG •• UUGUC	53-54 (2/4)	-	9	AGUAG • UUCUC	63-64 (2/4)	-
4	AGUGG • • UUAUC	-	55-57 (3/20)	10	AGUCG • UUAUC	-	65-66 (1/7)
5	GGUAG •• UUAUC	-	58-59 (1/7)	11	Other species with 1-3 mismatches in stem pairs 3-5	-	67-72 (0/0)
6	CGUAG • UUAUC	-	60 (1/7)				

FIG. 1.—A, Alignment of the mitochondrial ND3 gene region spanning the extra base site (designated by *) for six representative birds (full species names given below). Site numbers correspond to ND3 for *Gallus gallus* (Desjardins and Morais 1990). Numbers preceding species names and the boxed sequences correspond to the numbered stem-and-loop hypotheses in B. B, Sequence and hypothesized stem-and-loop structure for mitochondrial ND3 immediately downstream from the extra base site (starting at position 175 in A) for a set of vertebrate animals with and without the extra base. Structure numbers and species (names of nonbirds are underlined) with nucleotide sequences shown in the diagram are: (1, with extra base) 1–33: *Acryllium vulturinum*, *Burhinus oedicnemus*, *Cacatua goffini*, *Cathartes aura*, *Cathartea melambrotus*, *Chordeiles minor*, *Coccyzus erythrophthalmus*, *Colaptes auratus*, *Colius striatus*, *Coracias caudata*, *Diomedea nigripes*, *Fulica americana*, *Gallus gallus*, *Gavia immer*, *Gavia pacifica*, *Grus canadensis*, *Gyps fulvus*, *Meleagris gallopavo*, *Mycteria americana*, *Neophema elegans*, *Nyctanassa violacea*, *Opisthocomus hoazin*, *Otus asio*, *Pandion haliaetus*, *Phalacrocorax pelagicus*, *Phasianus colchicus*, *Phoenicoptera rubra*, *Podiceps nigricollis*, *Podilymbus podiceps*, *Rhea americana*, *Struthio camelus*, *Tauraco hartlaubi*, *Trogon curucui*; (1, without extra base) 34–41: *Alligator mississippiensis*, *Certhia familiaris*, *Corvus brachyrhynchos*, *Cyprinus carpio* (GenBank accession number X61010), *Dryoscopus gambensis*, *Lamprotornis caudatus*, *Smithornis sharpei*, *Vidua chalybeata*; (2, with extra base) 42–51: *Alcedo cristata*, *Aptenodytes patagonicus*, *Bonasa umbellus*, *Caprimulgus vociferus*, *Cephus columba*, *Chrysemys picta*, *Falco peregrinus*, *Sphyrapicus varius*, *Tockus erythrorhynchus*, *Urocolius macrourus*; (2, without extra base) 52: *Geococcyx californianus*; (3, with extra base) 53–54: *Anseranas semipalmata*, *Aythya americana*; (4, without extra base) 55–57: *Chaetura cinereiventris*, *Dendrocygna arcuata*, *Elminia longicauda*; (5, without extra base) 58–59: *Sphenodon punctatus*, *Sturnus vulgaris*; (6, without extra base) 60: *Buteo jamaicensis*; (7, without extra base) 61: *Scolopax minor*; (8, without extra base) 62: *Lanius collurio*; (9, with extra base) 63–64: *Eudromia elegans*, *Zenaida macroura*; (10, without extra base) 65–66: *Crossostoma lacustre* (M91245), *Sayornis phoebe*; (11, without extra base) 67–72: *Crocodylus porosus*, *Homo sapiens* (J01415), *Mus musculus* (J01420), *Ornithorhynchus anatinus* (X83427), *Didelphis virginiana* (Z29573), *Xenopus laevis* (M10217). Numbers in parentheses below the species numbers (1–72) indicate the percentages of bird species sampled, with or without the extra base, showing the indicated stem-and-loop structure. Previously unpublished sequences have GenBank accession numbers AF069422–AF069428, AF069430, AF076294–AF076308, and AF076338–AF076379.

which there is no wobble position pairing (Lagerkvist 1978), and this may provide a reasonable model for avian mt ND3 translation. tRNA^{Leu} (CUN) binds ND3 mRNA CUC just upstream of the avian ND3 extra base (fig. 1, positions 172–174), with a mismatch at the wobble position. Potential +1 slippage, decoding UCA with anticodon UAG from tRNA^{Leu} (CUN), would also yield a single mismatch, but at the second position and a

transversion. Decoding based on only two of three codon positions appears to be common in animal mitochondrial translation (see Gillham 1994, p. 320), and this might yield a level of translational frameshifting sufficient for expression of the frameshifted product.

Frameshifting occurs in competition with in-frame translation and can be facilitated by a pause in translation provided, in some cases, by an mRNA secondary

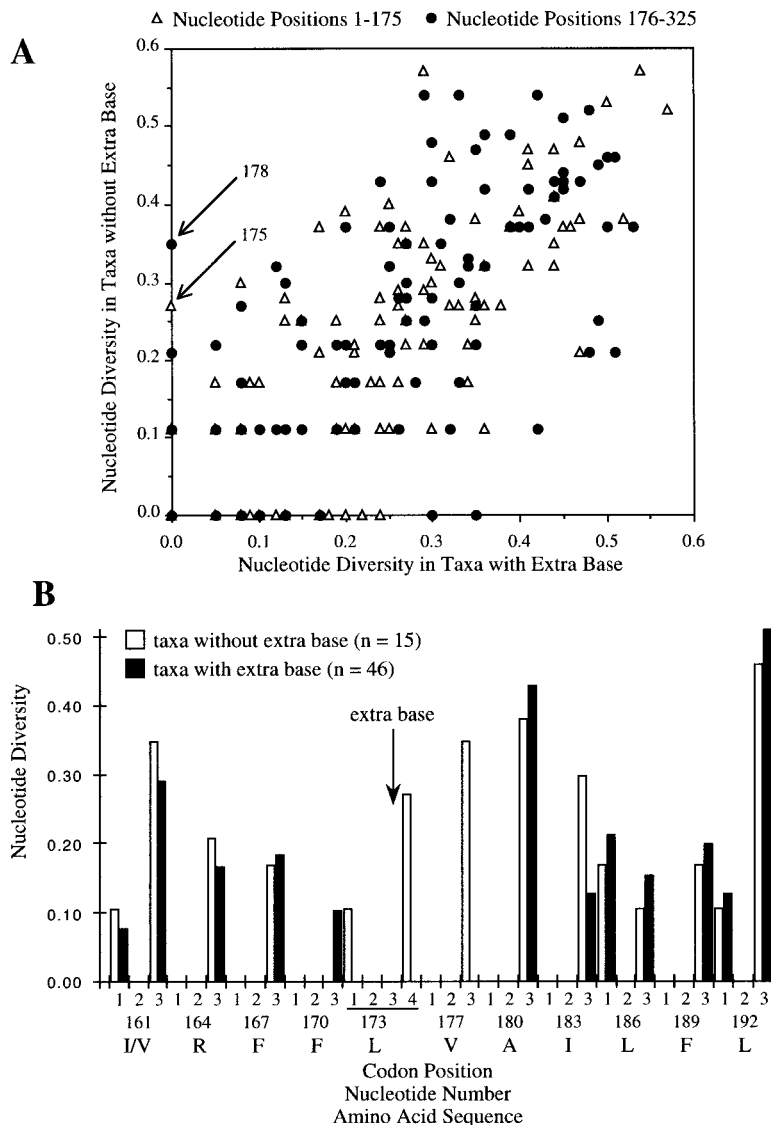


FIG. 2.—*A*, Comparison of mitochondrial ND3 position-by-position nucleotide diversity values ($= -\sum p_i \times \log(p_i)$ for $i = A, C, G, T$) for avian taxa, named in the caption to figure 1*B*, with and without the extra base. Values of nucleotide diversity for taxa with and without the extra base are highly correlated ($r^2 = 0.72$, $P \ll 0.00001$) across the entire length of the gene, suggesting similar constraints on sequence evolution for the whole gene in the two groups of taxa and the maintenance of a functional reading frame following the extra base. *B*, The same diversity value for the ND3 region spanning 11 codons immediately before and after the extra base. Two outliers (positions 175 and 178) from the overall relationship are in the immediate vicinity of the extra base. At these two positions, taxa with the extra base show no variation, while taxa without the extra base are variable. This pattern is consistent with the existence of additional constraints on sequence evolution for taxa with the extra base and a relaxation of those constraints for taxa that have lost the extra base.

structure (stem-and-loop) impediment just below the frameshift site (Matsufuji et al. 1995). We have identified a possible secondary structure in ND3 just downstream of the extra base site (fig. 1*B*). In this hypothetical structure, the A at position 175 for *G. gallus* is the first of 15 bases involved in a 5-bp stem plus 5-bp loop. If this structure is necessary for a frameshift to occur and a functional ND3 protein to be produced, birds and other taxa without the extra base might be expected to show greater sequence variation in this region and to have sequences inconsistent with the structure. Although the nucleotide sequence in all taxa is also constrained by a conserved amino acid sequence in this region, when potential variation is limited primarily to third co-

don positions, our data suggest greater sequence variation and a tendency toward fewer paired bases in birds lacking the extra ND3 base. Only 2 of the 46 bird species (4.4%) having the extra base have a single stem mismatch, whereas 4 of the 15 bird species (26.7%) lacking the extra base have a stem mismatch (fig. 1*B*). In addition, two positions involved in the stem structure showed much greater variation (as measured by nucleotide diversity) in birds without the extra base (fig. 2*B*, positions 175 and 178), whereas those with the extra base are invariant. This pattern is consistent with the existence of additional constraints on sequence evolution in taxa with the extra base, and a relaxation of those constraints in taxa that have lost the extra base. We sug-

gest that these additional constraints may be associated with the maintenance of a secondary structure in the mRNA that is involved in a translational frameshifting mechanism. Furthermore, we note that without frameshifting, the extra base observed in avian ND3 yields an underutilized codon for serine, AGU. In *G. gallus* mitochondrial DNA, for example, AGC, which also encodes serine, outnumbers AGU by 68 to 3. A bias against efficient decoding of AGU may provide an additional kinetic stimulus for frameshifting.

RNA editing involving deletion of the extra base is an alternative explanation for maintenance of the reading frame and functionality of ND3 in birds having the insertion, and we cannot rule out this possibility. Neither translational frameshifting nor RNA editing has been previously reported for animal mitochondrial protein-coding genes, although RNA editing for some animal mitochondrial tRNA genes has been documented (e.g., Yokobori and Pääbo 1997). To date, no single adenine or cytosine deletions via RNA editing have been reported for any organism. The RNA editing deletion type most similar to single adenine or cytosine deletion would be deletion of numerous uracil residues in mitochondria of kinetoplastid protozoans (Smith, Gott, and Hanson 1997).

Translational frameshifting and RNA editing can provide important mechanisms for evolutionary change at the molecular level, and these mechanisms might be expected to be conserved among taxa in which they have evolved. For our avian study taxa, we found that the extra base insertion, and possible translational frameshifting or RNA editing, is not conserved within groups of related species. For example, in the avian orders Anseriformes (*Aythya*, *Anseranas*, *Dendrocygna*), Falconiformes (*Pandion*, *Falco*, *Buteo*), Charadriiformes (*Burhinus*, *Cephus*, *Scolopax*), and Cuculiformes (*Coccyzus*, *Geococcyzus*), we found some species with the extra base and other species without it (fig. 1B), an unexpected result presuming monophyly of the orders mentioned and a tendency toward conservation of the mechanism.

We also sequenced the relevant portion of ND3 in *Chrysemys picta* (Eastern painted turtle) and *Alligator mississippiensis* (American alligator). *Alligator* and published sequences for fish, amphibians, and mammals all lack the extra base, whereas *Chrysemys* has an extra base in the same position as do birds (fig. 1B). Given a sister relationship for birds and crocodylians relative to turtles, it is possible that frameshifting is a relatively old mechanism within Reptilia, being present in a turtle and many birds and being lost in a crocodylian and various other birds. Further sampling across diverse vertebrates will help in determining the evolutionary history of this extra base and its processing during translation.

We have reported on an extra base found in the mitochondrial ND3 genes of a diverse set of birds and one turtle. The extra base appears not to be processed during translation, as the downstream reading frame and amino acid sequence of the gene are conserved across a diverse set of vertebrate animals. Translational frameshifting and RNA editing are alternative mechanisms ca-

pable of explaining how the extra base is processed during or prior to translation. Regardless of the processing mechanism in effect, single extra bases that might be considered anomalous sequence errors or insertions yielding stop codons and truncation of gene product may instead be first indications of an alternative reading of mitochondrial mRNA, and worthy of further scrutiny by molecular evolutionists.

Acknowledgments

We thank Philip Farabaugh and two anonymous reviewers for valuable insights and commentary. We thank Frederick Sheldon and Van Remsen at the Louisiana State University Museum of Natural Science, Les Christidis, and Nathan Rice for providing tissue samples or DNA. This work was supported by grants from the NSF (BSR 9496343) and the University of Michigan Museum of Zoology Bird Division Swales and Fargo funds to D.P.M. M.D.S. was supported by NSF grant IBN-9412399 to Robert B. Payne.

LITERATURE CITED

- DESJARDINS, P., and R. MORAIS. 1990. Sequence and gene organization of the chicken mitochondrial genome. *J. Mol. Biol.* **212**:599–634.
- FARABAUGH, P. J. 1996. Programmed translational frameshifting. *Microbiol. Rev.* **60**:103–134.
- GILLHAM, N. W. 1994. *Organelle genes and genomes*. Oxford University Press, New York.
- HÄRLID, A., A. JANKE, and U. ARNASON. 1997. The mtDNA sequence of the ostrich and the divergence between paleognathous and neognathous birds. *Mol. Biol. Evol.* **14**:754–761.
- KASTELEIN, R. A., E. RAMAUT, W. FIERS, and J. VANDUIN. 1982. Lysis gene expression of RNA phage MS2 depends on a frameshift event during translation of the overlapping coat protein gene. *Nature* **295**:35–41.
- LAGERKVIST, U. 1978. "Two out of three": an alternative method for codon reading. *Proc. Natl. Acad. Sci. USA* **75**:1759–1762.
- MATSUFUJI, S., T. MATSUFUJI, Y. MIYAZAKI, Y. MURAKAMI, J. F. ATKINS, R. F. GESTELAND, and S. HAYASHI. 1995. Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme. *Cell* **80**:51–60.
- MINDELL, D. P., M. D. SORENSON, C. J. HUDDLESTON, H. MIRANDA, A. KNIGHT, S. J. SAWCHUK, and T. YURI. 1997. Phylogenetic relationships among and within select avian orders based on mitochondrial DNA. Pp. 211–245 in D. P. MINDELL, ed. *Avian molecular evolution and systematics*. Academic Press, New York.
- PARKER, J. 1989. Errors and alternatives in reading the universal genetic code. *Microbiol. Rev.* **53**:273–289.
- SMITH, H. C., J. M. GOTT, and M. R. HANSON. 1997. A guide to RNA editing. *RNA* **3**:1105–1123.
- SOORENSON, M. D., and T. W. QUINN. 1998. Numts: a challenge for avian systematics and population biology. *Auk* **115**:214–221.
- YOKOBORI, S., and S. PÄÄBO. 1997. Polyadenylation creates the discriminator nucleotide of chicken mitochondrial tRNA (Tyr). *J. Mol. Biol.* **265**:95–99.

ROSS H. CROZIER, reviewing editor

Accepted July 28, 1998