

Warm-Blooded Isochore Structure in Nile Crocodile and Turtle

Sandrine Hughes,* Dominique Zelus,† and Dominique Mouchiroud*

*Laboratoire de Biométrie et Biologie Evolutive, Unité Mixte de Recherche-Centre National de la Recherche Scientifique 5558, Université Claude Bernard, Villeurbanne, France; and †UMR-Centre National de la Recherche Scientifique 49, École Normale Supérieure de Lyon, Lyon, France

The genomes of warm-blooded vertebrates are characterized by a strong heterogeneity in base composition, with GC-rich and GC-poor isochores. The GC content of sequences, especially in third codon positions, is highly correlated with that of the isochore they are embedded in. In amphibian and fish genomes, GC-rich isochores are nearly absent. Thus, it has been proposed that the GC increase in a part of mammalian and avian genomes represents an adaptation to homeothermy. To test this selective hypothesis, we sequenced marker protein genes in two cold-blooded vertebrates, the Nile crocodile *Crocodylus niloticus* (10 genes) and the red-eared slider *Trachemys scripta elegans* (6 genes). The analysis of base composition in third codon position of this original data set shows that the Nile crocodile and the turtle also exhibit GC-rich isochores, which rules out the homeothermy hypothesis. Instead, we propose that the GC increase results from a mutational bias that took place earlier than the adaptation to homeothermy in birds and before the turtle/crocodile divergence. Surprisingly, the isochore structure appears very similar between the red-eared slider and the Nile crocodile than between the chicken and the Nile crocodile. This point questions the phylogenetic position of turtles as a basal lineage of extant reptiles. We also observed a regular molecular clock in the Archosauria, which enables us, by using a more extended data set, to confirm Kumar and Hedges's dating of the bird-crocodile split.

Introduction

Vertebrate genomes are heterogeneous in base composition. They can be described as a mosaic of isochores, i.e., very long DNA segments (200–1,000 kb) in which base composition is homogeneous, that belong to a small number of families characterized by different GC levels (see Bernardi [1995] for a recent review). The GC content of protein-coding genes, and, more specifically, that in third codon positions, is highly correlated with that of the isochore into which they are embedded for cold-blooded (Bernardi and Bernardi 1991) and warm-blooded (Kadi et al. 1993; Clay et al. 1996) vertebrates. Thus, the analysis of base composition in third codon positions (GC3) from orthologous coding sequences can be used to study the structure (Zoubak, Clay, and Bernardi 1996) and the evolution (Bernardi, Hughes, and Mouchiroud 1997; Robinson, Gautier, and Mouchiroud 1997; Galtier and Mouchiroud 1998) of the isochores in vertebrate genomes.

The isochore patterns appear to be very different in amphibians and fishes than in birds and mammals, as assessed by analytical DNA ultracentrifugation (Bernardi and Bernardi 1990a, 1990b) and sequence analysis (Bernardi et al. 1988; Bernardi and Bernardi 1991). The amphibian and fish genomes, more precisely, Actinopterygii, exhibit much lower compositional heterogeneity. Whereas GC-poor isochores are compositionally conserved among all vertebrates, GC-rich isochores are nearly absent in amphibian and fish genomes. These results indicate that the separation of the genomes of birds and mammals and those of amphibians and fishes re-

sulted from a “major compositional transition” in which a sizeable proportion of the cold-blooded vertebrate genomes underwent GC increases, leading to heterogeneous warm-blooded genomes.

Debate rages on the question of isochore evolution in vertebrate genomes. The GC heterogeneity has been considered a convergently derived state that independently occurred twice, in association with the development of homeothermy in both mammalian and avian lineages (Bernardi 1993). An increase in GC content in some genomic regions could protect DNA, RNA, and proteins from degradation by heat. Sequence data available thus far have not clearly distinguished a “warm-blooded” pattern from an “amniote” pattern, since only mammal and bird sequences were representative of amniotes in the database. Thus, the isochore composition of the genomes of turtles, lizards, snakes, and crocodiles, all “cold-blooded” amniotes, is highly relevant to this question.

The purpose of the present work was to test the homeothermy hypothesis by investigating the isochore patterns in two cold-blooded reptilian groups, crocodylians and turtles. Ten protein-coding genes selected as markers of the GC transition were sequenced for *Crocodylus niloticus*, and six were sequenced for the turtle *Trachemys scripta elegans*. We compared their base compositions with those of orthologous genes of other vertebrates, cold-blooded (*Xenopus*) and warm-blooded (chicken and human). Our results demonstrate that genes with high GC contents are already present in the crocodylian and turtle genomes, indicating the presence of some GC-rich isochores and, consequently, refuting the homeothermy hypothesis.

Materials and Methods

Marker Sequences

Marker genes were chosen according to the following criteria: (1) orthologous sequences known in a mam-

Key words: isochore evolution, homeothermy, reptilian genome, relative rates, *Crocodylus niloticus*, *Trachemys scripta elegans*.

Address for correspondence and reprints: Sandrine Hughes, Laboratoire de Biométrie et Biologie Evolutive, Unité Mixte de Recherche-Centre National de la Recherche Scientifique 5558, Université Claude Bernard, Lyon 1, 43, boulevard du 11 Novembre 1918, 69622 Villeurbanne cedex, France. E-mail: hughes@biomserv.univ-lyon1.fr.

Mol. Biol. Evol. 16(11):1521–1527. 1999

© 1999 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

Table 1
GC Content in Third Codon Positions of Aligned Sites for the 10 Genes Sequenced for *Crocodylus niloticus* and for *Trachemys scripta elegans* Compared with Human, Chicken, and *Xenopus* Orthologous Genes

GENE NAME	%GC3					LENGTH ^a (bp)
	Human	Chicken	<i>Xenopus</i>	Crocodile	Turtle	
C-ski protein ^b (<i>c-ski</i>)	91.0	64.7	79.0	75.8	77.0	756
Hepatocyte growth factor ^c (<i>hgf</i>)	73.3	80.7	47.9	67.2	—	933
Pyruvate kinase ^c (<i>pk</i>)	67.4	65.1	41.9	54.9	56.3	645
PPAR Alpha ^d (<i>ppara</i>)	59.9	41.3	68.8	38.2	—	981
PPAR Beta ^d (<i>pparb</i>)	80.1	78.5	48.0	76.5	—	906
PPAR Gamma ^d (<i>pparg</i>)	55.8	40.4	69.6	40.4	44.5	957
Retinol binding protein ^b (<i>rbp</i>)	80.9	54.8	48.8	57.1	63.1	252
Ribosomal protein S6 ^e (<i>rps6</i>)	42.2	55.6	46.1	45.6	46.1	540
SPARC protein ^c (<i>sparc</i>)	82.7	84.8	56.9	92.9	85.8	591
Vitamine D receptor ^c (<i>vdr</i>)	74.4	85.5	34.3	70.1	—	621

^a Length of orthologous sequences between the four or five species after alignment and gap removal.

^b Genes selected with high GC differences between humans and the two others species, *Xenopus* and chickens.

^c Genes selected with high GC differences between the two warm-blooded vertebrates and *Xenopus*.

^d Genes sequenced in collaboration with another group. The sequences of the chicken PPAR alpha, beta, and gamma were obtained from I. Takada and K. Umesono (personal communication).

^e Genes selected with low GC differences between the three species.

mal (*Homo sapiens*), a bird (*Gallus gallus*), and an amphibian (*Xenopus laevis*); (2) very large GC differences in third codon positions (ΔGC3) between at least two of these orthologous genes (table 1) (We also sequenced a GC-poor gene [*rps6*] having similar GC contents in humans, chickens, and *Xenopus*); (3) sufficient base conservation for the design of multiple PCR primers (see below) and a resulting fragment size that is not too small compared to complete gene size; and (4) some knowledge of tissue expression for reverse transcription–polymerase chain reaction (RT-PCR) experiments (see below).

For the first criterion, orthologous genes were selected using the HOVERGEN software (release 31; Duret, Mouchiroud, and Gouy 1994), which allowed us to check the orthology status by analyzing phylogenetic trees. Multigenic families were disregarded in most cases to avoid mistakes. After this step, genes were extracted from GenBank with QUERY_WIN (Perriere, Gouy, and Gojobori 1994) and aligned using CLUSTAL W (Thompson, Higgins, and Gibson 1994). Two pairs of 20mer-degenerated primers (sequences available on request) were constructed for each gene based on the most conserved zones according to Escriva et al. (1997).

On the other hand, three genes of the peroxisome-proliferator-activated receptor family (*ppar*) were sequenced in collaboration with another group (I. Takada,

K. Umesono, and V. Laudet) and introduced in our data set. Other crocodylian or turtle genes, already available in GenBank but not markers of the transition, were added in the sequence analysis. These sequences are available on request. The GenBank/EMBL accession numbers for the 16 genes sequenced are AJ011391–AJ011397, AJ011515, AJ2433131–AJ243136, AJ2433138, and AJ243139.

RT-PCR Method

Biological material used in this study was taken from an adult female Nile crocodile (*C. niloticus*) from the Crocodile Farm of Pierrelatte (France) and from an adult female red-eared slider (*T. s. elegans*). Different tissue samples are conserved in liquid nitrogen or ethanol at the Centre d'Analyse Moléculaire de la Biodiversité (CAMB), Lyon, France.

Total RNAs were extracted from frozen liver tissues with Qiagen RNeasy kit. For each gene, a first touchdown RT-PCR was performed under various conditions with the Perkin-Elmer Amplitaq Gold Taq polymerase, followed by a second run with nested primers to enhance specificity of the products, according to Escriva et al. (1997). Resulting fragments were purified (Qiagen Qiaquick system), ligated into pUC18 (Amersham SureClone ligation kit) and cloned into *Escherichia coli* strain DH5- α . Fluorescence sequencing (Amersham kit) of some clones was performed using a Pharmacia automatic sequencer ALF express.

Sequence Analysis

Sequences of *C. niloticus* and *T. s. elegans* were aligned by hand with the already-known vertebrate sequences using SEAVIEW (Galtier, Gouy, and Gautier 1996). Phylogenetic reconstructions were made with PHYLO-WIN (Galtier, Gouy, and Gautier 1996) by neighbor joining applied to nonsynonymous evolutionary distances (K_a) (Li 1993). K_a distances between crocodiles ($K_{a(\text{cr})}$) or chickens ($K_{a(\text{ch})}$) and their most recent

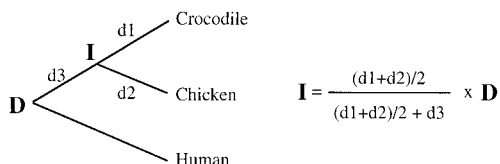


FIG. 1.—Divergence time estimate (I) between birds and crocodylians by the lineage-specific method (Kumar and Hedges 1998). D = divergence time between birds and mammals fixed at 310 Myr, $d1 = K_{a(\text{cr})}$, and $d2 = K_{a(\text{ch})}$.

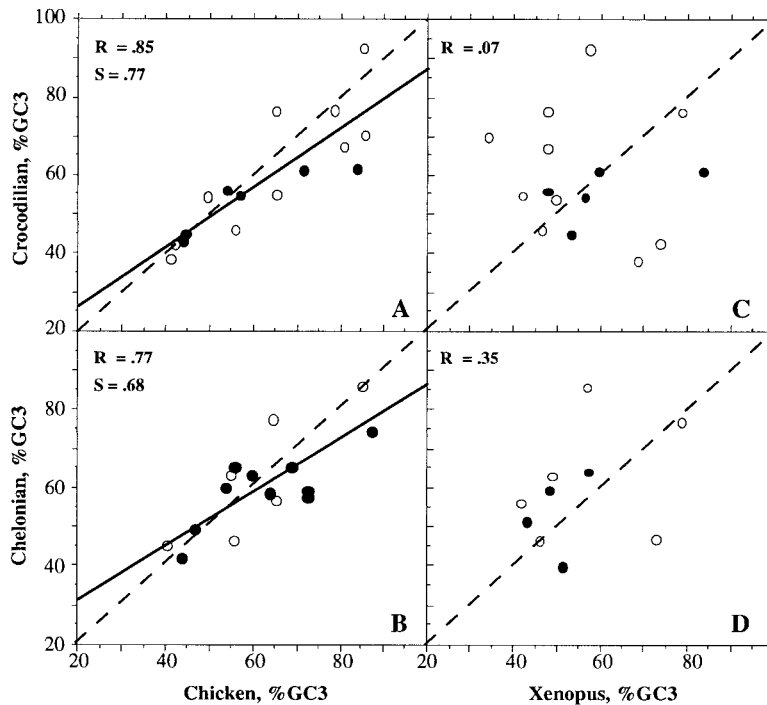


FIG. 2.—Correlation between GC3 values of orthologous genes from (A) chickens and crocodilians, (B) chickens and chelonians, (C) *Xenopus* and crocodilians, and (D) *Xenopus* and chelonians. White circles represent partial genes sequenced for *Crocodylus niloticus* or *Trachemys scripta elegans*. Black circles represent previously known genes from GenBank. Broken lines indicate diagonals. Solid lines indicate the linear regression line. *R* is the correlation coefficient, and *S* is the slope of linear regression.

common ancestor (fig. 1) were calculated for each gene as follows, with humans as outgroup:

$$K_{a(ch)} = \frac{1}{2}(K_{a(ch-hu)} - K_{a(cr-hu)} + K_{a(ch-cr)})$$

$$K_{a(cr)} = K_{a(ch-cr)} - K_{a(ch)}.$$

The relative-rate tests to compare substitution rates between two lineages or two groups of sequences were done using the program RRTree (Robinson et al. 1998). For dating the bird-crocodile divergence time, we applied the lineage-specific method (fig. 1) developed by Kumar and Hedges (1998). This method is not biased by high or low evolutionary rate in the outgroup lineage.

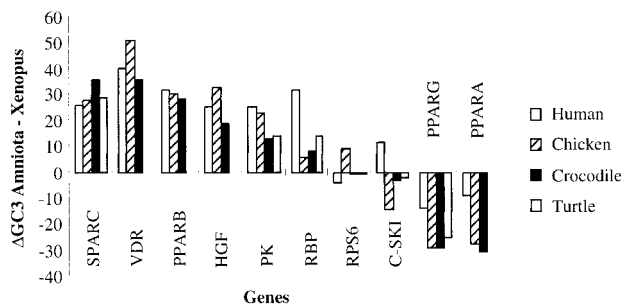


FIG. 3.—Differences in GC3 content between orthologous genes from humans, chickens, crocodiles, turtles, and *Xenopus*. Positive values of $\Delta GC3$ represent increases in GC3 content relative to the *Xenopus* sequence.

Results

GC Content in Crocodilian and Chelonian Coding Genes

The RT-PCR method used allowed only partial sequencing of coding genes. Thus, fragments of various sizes (ranging from 252 to 981 bp) were obtained, representing 7,182 bp of the nuclear genome of *C. niloticus* and 3,741 bp of that of *T. s. elegans* (table 1). The GC levels in third codon positions in orthologous sequences from humans, chickens, and *Xenopus* and in gene fragments from crocodiles and turtles are given in table 1. The GC3 differences in orthologous genes between *Xenopus* and the other species are plotted in figure 2.

Globally, GC3 levels in the crocodile and turtle coding sequences appear much more similar to those of humans and/or chickens than to those of *Xenopus* (table 1 and fig. 2). Indeed, the GC3 levels of nearly all coding sequences are statistically different between the two cold-blooded reptiles and the amphibian (fig. 2C and D, white points). Six genes (*sparc*, *vdr*, *pparb*, *hgf*, *pparg*, and *ppara*) show more than 20% difference in GC3 between *Xenopus* and the Nile crocodile (fig. 3). As expected, the low GC3 level of *rps6* in the three known species also stands in the Nile crocodile and turtle genes. For with large GC3% differences between humans and chickens (*c-ski*, *rbp*, *ppara*, *pparg*, and *rps6*), the GC3 level of homologous crocodilian genes is more similar to that in chicken sequences (Student's *t*-test on paired data; $P = 0.98$) than to that in human sequences ($P = 0.0387$). In addition, data in table 1 indicate that the

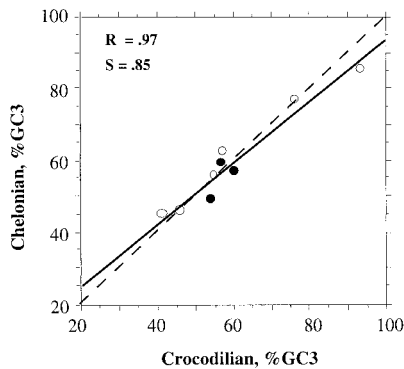


FIG. 4.—Correlation between GC3 values of orthologous genes from crocodilians and chelonians. White circles represent partial genes in *Crocodylus niloticus* or *Trachemys scripta elegans*. Black circles represent previously known genes from GenBank. Broken lines indicate diagonals. Solid lines indicate the linear regression line. *R* is the correlation coefficient, and *S* is the slope of linear regression.

GC3 levels of the six turtle coding genes are similar to those in crocodiles ($P = 0.6028$).

Increasing our data set with previously known sequences in GenBank enabled us to corroborate the above conclusion with a strong positive correlation between GC3 values of orthologous genes from chickens and crocodiles ($r = 0.85$, $P < 0.0001$) (fig. 2A) or from chickens and turtles ($r = 0.77$, $P = 0.003$) (fig. 2B). No such correlations were observed for comparison between crocodile and *Xenopus* ($r = 0.07$, $P = 0.8061$) (fig. 2C) or between turtle and *Xenopus* GC3 values ($r = 0.35$, $P = 0.3360$) (fig. 2D). The most surprising result is the extremely high correlation coefficient obtained for the turtle/crocodile comparison ($r = 0.97$, $P < 0.0001$) (fig. 4), although it was supported by a small gene sample. The original evolution of base composition

in *ppara* and *pparg* genes, which show higher GC levels in *Xenopus*, could be explained by a secondary translocation of these genes (Bernardi, Hughes, and Mouchiroud 1997).

Divergence Time Between Birds and Crocodilians

Before dating the bird/crocodile split, we tested potential evolutionary-rate differences between the avian and crocodilian lineages. The relative-rate tests were computed on nonsynonymous substitution rates (K_a) using human sequences as outgroup (fig. 1). For 13 of the 16 genes tested, we did not detect any significant difference between chicken and crocodilian nonsynonymous rates. Three genes presented significant rate differences: *c-ski*, *eno*, and *znf*. Paralogy or functional changes could explain the significant K_a differences for these three genes. Thus, these genes were disregarded for divergence time estimation.

The divergence of birds and mammals (310 MYA; Benton 1990) provides a reliable point at which to anchor the molecular clock in amniotes. For each gene, the divergence time between the bird and crocodilian lineages was estimated by the lineage method (Kumar and Hedges 1998), with humans as outgroup. The bird/crocodilian divergence times range from 80 Myr for the *pparg* gene to 280 Myr for the *rhd* gene (table 2), with an average of 180 ± 62 Myr. This value increases to 196 ± 50 Myr if the two genes with the shortest divergence times (*pparb* and *pparg*) are disregarded. Moreover, the divergence time obtained with a combined analysis of the 12 protein-coding genes, 200 Myr, increases to 215 Myr without the two *ppar* genes.

Discussion

Crocodile/Bird Divergence

As expected, phylogenetic analysis of the 10 presently sequenced genes from the Nile crocodile strongly

Table 2
Relative-Rate Tests for Nonsynonymous Substitution Rates (K_a) Between Chicken and Crocodilian Lineages with Humans as Outgroup and Estimation of Divergence Time

Gene Name	No. of Sites Used	$K_{a(ch)}$ (%) Chicken	$K_{a(cr)}$ (%) Crocodile	Relative-Rate Test	Divergence Time (Myr)
<i>c-ski</i> ^a	504	0.19	1.59	$-1.40 \pm 0.6^*$	—
<i>hgf</i> ^a	655	8.97	6.47	2.50 ± 1.9	216
<i>pk</i> ^a	432	2.14	2.08	0.06 ± 1.0	226
<i>ppara</i> ^a	642	0.53	1.43	-0.90 ± 0.6	151
<i>pparb</i> ^a	601	0.55	0.56	-0.01 ± 0.4	90
<i>pparg</i> ^a	749	0.23	0.28	-0.05 ± 0.3	80
<i>rhp</i> ^a	224	2.32	3.12	-0.80 ± 1.6	147
<i>rps6</i> ^a	339	0.22	0.30	-0.08 ± 0.4	141
<i>sparc</i> ^a	403	1.29	1.15	0.14 ± 0.8	162
<i>vdr</i> ^a	436	6.23	6.24	-0.01 ± 1.9	271
<i>c-mos</i> ^b	242	7.46	7.63	-0.17 ± 3.0	152
<i>eno</i> ^b	808	0.80	2.40	$-1.60 \pm 0.6^*$	—
<i>rhd</i> ^b	707	2.94	2.43	0.51 ± 0.9	280
<i>wtp</i> ^b	503	1.20	1.15	0.05 ± 0.7	221
<i>pr</i> ^c	705	1.87	2.23	-0.36 ± 0.8	194
<i>znf</i> ^c	343	0.56	2.65	$-2.09 \pm 0.9^*$	—

^a Genes sequenced for *Crocodylus niloticus* and known for humans, chickens, and *Xenopus*.

^b Genes previously known for humans, chickens, *Xenopus*, and another crocodilian species, *c-mos* = *c-mos* proto-oncogene (*Crocodylus porosus*); *eno* = alpha enolase (*Alligator mississippiensis*); *rhd* = rhodopsin (*A. mississippiensis*); *wtp* = Wilm's tumor (*A. mississippiensis*).

^c Genes previously known only for humans, chickens, and another crocodilian species, *pr* = progesterone receptor (*Crocodylus siamensis*); *znf* = ZFY-related zinc finger gene (*A. mississippiensis*).

* Significant difference ($P < 0.05$).

supports the Archosauria, which groups birds and crocodilians in a sister group with bootstrap values higher than 85%, except for *vdr* (59%) and *pk* (71%). As a whole, these data definitively rule out the “Haemothermia” hypothesis, which groups birds and mammals in a sister group (Hedges, Moberg, and Maxon 1990).

For the divergence time between birds and crocodilians, a recent work has proposed a mean value of 222 ± 40 Myr (Kumar and Hedges 1998), in agreement with the 240-Myr value suggested by fossil analyses (Benton 1990). The new divergence time (215 Myr) that we estimated on 10 concatenated nuclear coding genes (i.e., 2.5 more genes than in previous studies) is in accordance with Kumar’s findings. Analysis of the avian/crocodilian split, performed on the complete mitochondrial genome of *Alligator mississippiensis*, gives a divergence time of 254 Myr (Janke and Arnason 1997), close to our estimate for the nuclear genome.

The Phylogenetic Position of the Turtle Among Vertebrates

Among reptiles, the phylogenetic position of turtles is presently questioned, since recent works disproved that turtles were first reptiles (Rieppel and DeBraga 1996; Zardoya and Meyer 1998). Our results are presently too scarce to provide valuable information on this question. Nevertheless, the extremely high correlation coefficient observed between GC3 levels of orthologous genes from crocodiles and turtles (fig. 4) suggests that turtles and crocodiles might be near parents, as proposed by Hedges and Poling (1999). In order to test this hypothesis, complete sequencing of marker coding genes studied here is in progress for a squamate.

GC Variability in Crocodilian and Turtle Genomes

In light of results obtained with Nile crocodile coding genes and confirmed with those of the turtle, it appears that the GC increase observed for warm-blooded vertebrates cannot be explained by homeothermy, considering that crocodiles and turtles are ectotherms.

The discrepancy between the GC3 heterogeneity observed here for *C. niloticus* and *T. s. elegans* coding genes and the compositional homogeneity obtained through analytical ultracentrifugation (Bernardi and Bernardi 1990a, 1990b) could result from the sensitivity of the latter approach. One might expect such a result if the isochore structure in the reptilian genome were intermediate between isochore structures of other cold-blooded and warm-blooded vertebrates. In fact, many molecular characters place reptilian genomes as intermediate between cold-blooded and warm-blooded vertebrates. Indeed, genome sizes are smaller and more uniform in reptilian genomes than in amphibian and fish genomes (Olmo 1991). The methylation level in reptiles is roughly only half that in fishes and amphibians (Jabbari et al. 1997) but the same as that in mammals and birds. R-banding patterns corresponding to GC-rich DNA were demonstrated in the chromosomes of turtles, snakes, and some lizards but were absent in fishes and amphibia (Schmid and Guttenbach 1988).

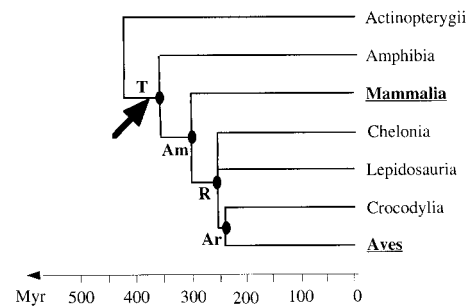


FIG. 5.—Simplified phylogeny of vertebrates. Homeotherms are boldface and underlined. T = Tetrapoda, Am = Amniota, R = Reptilia, Ar = Archosauria. The arrow indicates the point at which the GC-increase of a part of genome could have taken place.

However, our data do not allow us to claim that the GC variability in the Nile crocodile or in the red-eared slider genome is as high as that in birds. A larger gene sample is necessary to confirm a birdlike isochore structure in cold-blooded reptilian genomes. Nevertheless, the high correlations observed between GC3 values for chickens and crocodilians (fig. 2A) and between those for chickens and chelonians (fig. 2B) argue in favor of such a hypothesis.

Isochore Evolution in Vertebrates

The present results suggest that the GC increase observed in part of avian genomes certainly occurred before the crocodile/turtle divergence. Furthermore, although without squamate data, the GC variability in the crocodilian and chelonian genome indicates that the common reptilian ancestor of mammals and birds would already have undergone an increase in GC content.

Thus, we have some arguments for positioning the GC increase in the common ancestor of tetrapods (fig. 5). A comparison of orthologous genes from *Xenopus*, chickens, and humans has shown that the majority of increasingly GC-rich genes were the same in the two warm-blooded vertebrates and, moreover, corresponded to GC3-rich genes from *Xenopus* (Bernardi, Hughes, and Mouchiroud 1997). The finding that DNA from the most GC-rich isochores of humans hybridizes with the most GC-rich isochores of other vertebrates, including those in amphibia (Bernardi 1995), also supported this result. These observations indicate that the same genes and, consequently, the same DNA regions, are implicated in the GC increase in all tetrapods and could result from a unique event in the ancestral genome. In other respects, sequence analysis has shown that *Xenopus* and other amphibian genomes exhibit a slightly more heterogeneous genome structure than fish genomes and, notably, the Actinopterygii (data not shown). Moreover, this slight GC heterogeneity in amphibian genomes was not detected with compositional profile. This suggests that the GC increase may have taken place before the separation of amniota, perhaps in the common ancestor of tetrapods, between 450 and 350 MYA (fig. 5).

GC Increase and Mutational Pattern

Since the present results rule out the adaptive hypothesis explaining GC increase by higher stability of

DNA in homeotherm genomes, we propose another hypothesis. The origin of the regional GC increases could be explained by the variation in mutational pattern related to nonuniform gene distribution along DNA sequences (Bernardi et al. 1988; Mouchiroud et al. 1991). Indeed, gene density is 10 to 20 times as great in the most GC-rich isochores than in the most GC-poor isochores. Moreover, the nonuniform gene distribution is likely to be common to cold-blooded vertebrates (Duret, Mouchiroud, and Gautier 1995; Andersson et al. 1996). One argument in favor of the mutational bias hypothesis is that birds showing the strongest information compaction (McQueen, Siriaco, and Bird 1998) exhibit the highest GC increases (Kadi et al. 1993). Thus, the high gene density associated with various mutational biases linked to high recombination rate (Eyre-Walker 1993), early replication timing (Wolfe, Sharp, and Li 1989), or more efficient repair (Filipski 1990) could by itself explain the regional GC increases.

In conclusion, we propose that the establishment of GC-rich isochores, initially observed in mammalian and avian genomes, took place only once in vertebrate evolution (in the common ancestor of tetrapods) and probably resulted from a mutational bias linked to the nonuniform gene distribution. Nevertheless, if turtles and crocodiles are grouping in the same clade, the unique origin of the GC increase in amniotes cannot be advanced without squamate data. Therefore, we have undertaken the sequencing of the genes studied here for a squamate species to provide additional information about the phylogenetic position of turtles among reptiles and the origin of GC increases in some amniote genomes.

Acknowledgments

We thank L. Duret, M. Robinson, V. Laudet, R. Grantham, M. Gouy, and C. Gautier for useful comments and, more especially, V. Laudet and J. Briolay (CAMB, Lyon I) for technical advice. We are also grateful to I. Takada and K. Umesono for providing us with the sequences of the chicken PPAR alpha, beta, and gamma. We also thank L. Fougeirol ("Ferme aux crocodiles," Pierrelatte, France) for providing us with tissues of the Nile crocodile and to P. Joly and A. Cadi for tissues of the red-eared slider.

LITERATURE CITED

- ANDERSSON, L., A. ARCHIBALD, M. ASHUBURNER et al. (52 co-authors). 1996. Comparative genome organization of vertebrates. The First International Workshop on Comparative Genome Organization. *Mamm. Genome* 7:717–734.
- BENTON, M. J. 1990. Phylogeny of the major tetrapod groups, morphological data and divergence dates. *J. Mol. Evol.* 30:409–424.
- BERNARDI, G. 1993. The vertebrate genome: isochores and evolution. *Mol. Biol. Evol.* 10:186–204.
- . 1995. The human genome: organization and evolutionary history. *Ann. Rev. Genet.* 29:445–476.
- BERNARDI, G., and G. BERNARDI. 1990a. Compositional patterns in the nuclear genome of cold-blooded vertebrates. *J. Mol. Evol.* 31:265–281.
- . 1990b. Compositional transitions in the nuclear genomes of cold-blooded vertebrates. *J. Mol. Evol.* 31:282–293.
- . 1991. Compositional properties of nuclear genes from cold-blooded vertebrates. *J. Mol. Evol.* 33:57–67.
- BERNARDI, G., S. HUGHES, and D. MOUCHIROUD. 1997. The major compositional transition in the vertebrate genome. *J. Mol. Evol.* 44:44–51.
- BERNARDI, G., D. MOUCHIROUD, C. GAUTIER, and G. BERNARDI. 1988. Compositional patterns in vertebrate genomes: conservation and change in evolution. *J. Mol. Evol.* 28:7–18.
- CLAY, O., S. CACCIO, S. ZOUBAK, D. MOUCHIROUD, and G. BERNARDI. 1996. Human coding and noncoding DNA, compositional correlations. *Mol. Phylogenet. Evol.* 5:2–12.
- DURET, L., D. MOUCHIROUD, and C. GAUTIER. 1995. Statistical analysis of vertebrate sequences reveals that long genes are scarce in GC-rich isochores. *J. Mol. Evol.* 40:308–317.
- DURET, L., D. MOUCHIROUD, and M. GOUY. 1994. HOVERGEN, a database of homologous vertebrate genes. *Nucleic Acids Res.* 22:2360–2365.
- ESCRIVA, H., R. SAFI, C. HANNI, M. C. LANGLOIS, P. SAUMITOU-LAPRADE, D. STEHELIN, A. CAPRON, R. PIERCE, and V. LAUDET. 1997. Ligand binding was acquired during evolution of nuclear receptors. *Proc. Natl. Acad. Sci. USA* 94:6803–6808.
- EYRE-WALKER, A. 1993. Recombination and mammalian genome evolution. *Proc. R. Soc. Lond. B Biol. Sci.* 252:237–243.
- FILIPSKI, J. 1990. Evolution of DNA sequence contributions of mutational bias and selection to the origin of chromosomal compartments. *Adv. Mutagenesis Res.* 2:1–54.
- GALTIER, N., M. GOUY, and C. GAUTIER. 1996. SEAVIEW and PHYLO-WIN, two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.* 12:543–548.
- GALTIER, N., and D. MOUCHIROUD. 1998. Isochore evolution in mammals: a human-like ancestral structure. *Genetics* 150:1577–1584.
- HEDGES, S. B., K. D. MOBERG, and L. R. MAXSON. 1990. Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and a review of the evidence for amniote relationships. *Mol. Biol. Evol.* 7:607–633.
- HEDGES, S. B., and L. L. POLING. 1999. A molecular phylogeny of reptiles. *Science* 283:998–1001.
- JABBARI, K., S., CACCIO, J. P. PAIS DE BARROS, J. DESGRES, and G. BERNARDI. 1997. Evolutionary changes in CpG and methylation levels in the genome of vertebrates. *Gene* 205:119–124.
- JANKE, A., and U. ARNASON. 1997. The complete mitochondrial genome of *Alligator mississippiensis* and the separation between recent Archosauria birds and crocodiles. *Mol. Biol. Evol.* 14:1266–1272.
- KADI, F., D. MOUCHIROUD, G. SABEUR, and G. BERNARDI. 1993. The compositional patterns of the avian genomes and their evolutionary implications. *J. Mol. Evol.* 37:544–551.
- KUMAR, S., and S. B. HEDGES. 1998. A molecular timescale for vertebrate evolution. *Nature* 392:917–920.
- LI, W. H. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J. Mol. Evol.* 36:96–99.
- MCQUEEN, H. A., G. SIRIACO, and A. P. BIRD. 1998. Chicken microchromosomes are hyperacetylated, early replicating, and gene rich. *Genome Res.* 8:621–630.
- MOUCHIROUD, D., G. D'ONOFRIO, B. AISSANI, G. MACAYA, C. GAUTIER, and G. BERNARDI. 1991. The distribution of genes in the human genome. *Gene* 100:181–187.

- OLMO, E. 1991. Genome variations in the transition from amphibians to reptiles. *J. Mol. Evol.* **33**:68–75.
- PERRIERE, G., M. GOUY, and T. GOJOBORI. 1994. NRSUB, a non-redundant data base for the *Bacillus subtilis* genome. *Nucleic Acids Res.* **22**:5525–5529.
- RIEPEL, O., and M. DEBRAGA. 1996. Turtles as diapsid reptiles. *Nature* **384**:453–455.
- ROBINSON, M., C. GAUTIER, and D. MOUCHIROUD. 1997. Evolution of isochores in rodents. *Mol. Biol. Evol.* **14**:823–828.
- ROBINSON, M., M. GOUY, C. GAUTIER, and D. MOUCHIROUD. 1998. Sensitivity of the relative-rate test to taxonomic sampling. *Mol. Biol. Evol.* **15**:1091–1098.
- SCHMID, M., and M. GUTTENBACH. 1988. Evolutionary diversity of reverse (R) fluorescent chromosome bands in vertebrates. *Chromosoma* **97**:101–114.
- THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 1994. CLUSTAL W, improving the sensitivity of progressive multiple alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
- WOLFE, K. H., P. M. SHARP, and W. H. LI. 1989. Mutation rates differ among regions of the mammalian genome. *Nature* **337**:283–285.
- ZARDOYA, R., and A. MEYER. 1998. Complete mitochondrial genome suggests diapsid affinities of turtles. *Proc. Natl. Acad. Sci. USA* **95**:14226–14231.
- ZOUBAK, S., O. CLAY, and G. BERNARDI. 1996. The gene distribution of the human genome. *Gene* **174**:95–102.

DAN GRAUR, reviewing editor

Accepted July 26 1999