ELSEVIER

Gene 261 (2000) 71–83

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

The compositional transition between the genomes of cold- and warm-blooded vertebrates: codon frequencies in orthologous genes

Stéphane Cruveiller[a,b], Giuseppe D'Onofrio[a], Giorgio Bernardi[a,*]

[a]*Laboratorio di Evoluzione Molecolare, Stazione Zoologica Anton Dohrn, Villa Comunale, I-80121, Naples, Italy*
[b]*Laboratoire de Génétique Moléculaire, Institut Jacques Monod, 2 Place Jussieu, 75005 Paris, France*

Accepted 30 October 2000
Received by T. Gojobori

Abstract

The genomes of the ancestors of mammals and birds underwent a compositional change in which the gene-richest regions increased their GC levels. Here we investigated this compositional transition by analyzing the levels of G and C in third codon positions, as well as the codon frequencies of orthologous genes from human, chicken and *Xenopus*. The results may be summed up as follows: (i) GC-poor genes, that did not undergo the compositional transition, showed only minor differences in orthologous sets from *Xenopus*, human and chicken; this is remarkable in view of the very many nucleotide substitutions that occurred over the long evolutionary times separating these species; (ii) GC-rich genes, that underwent the compositional transition, showed large differences between *Xenopus* and warm-blooded vertebrates, but not between chicken and human. In other words, the independent changes that occurred in avian and mammalian genes, on the average, were the same.

Keywords: Birds; Genes; Isochores; Mammals; *Xenopus*

1. Introduction

DNAs from warm-blooded vertebrates (mammals and birds) exhibit high compositional heterogeneities and strongly asymmetrical bands in analytical density gradients of CsCl, whereas DNAs from cold-blooded vertebrates (fishes, amphibians and reptiles) are generally characterized by low compositional heterogeneities and by only slightly asymmetrical CsCl bands (Thiery et al., 1976; Bernardi and Bernardi, 1990a,b; satellite mitochondrial and ribosomal DNAs are neglected here). Both differences, in heterogeneity and asymmetry, are due to the presence in the genomes of warm-blooded vertebrates of a small percentage (about 15%) of GC-rich DNA molecules (GC is the molar ratio of guanine + cytosine in DNA) that are absent, or scarcely represented, in the genomes of most cold-blooded vertebrates. The findings just mentioned indicated that two major compositional shifts occurred in the distinct ancestral lines leading to warm-blooded vertebrates (see Bernardi, 2000a,b, for recent reviews, and discussion).

The two major shifts were then detected at the coding sequence level (Bernardi et al., 1985, 1988; Perrin and Bernardi, 1987; Mouchiroud et al., 1987, 1988; Bernardi and Bernardi, 1991). Indeed, coding sequences from cold-blooded vertebrates are compositionally relatively homogeneous and generally characterized by low GC levels, whereas coding sequences from warm-blooded vertebrates are compositionally much more heterogeneous and reach very high GC levels, up to 100% GC in the third codon positions of genes.

In other words, the compositional distributions of coding sequences showed patterns that mimicked those first seen at the DNA level. This can be understood since linear correlations hold between the GC levels of coding sequences and those of the DNA molecules carrying the sequences (Bernardi et al., 1985; Clay et al., 1996).

The best evidence for the major shifts was obtained by comparing GC_3 values (GC_3 is the average GC level of third codon positions of a gene) of orthologous genes from human and *Xenopus* (Perrin and Bernardi, 1987; Mouchiroud et al., 1987, 1988; Bernardi and Bernardi, 1991; D'Onofrio et al., 1999; Cruveiller et al., 1999). These GC_3 plots show that while in the case of GC-poor genes, human and *Xenopus* values lie around the diagonal, in the case of GC-rich genes, human genes tend to have increasingly higher values

* Corresponding author. Tel.: +39-081-5833215/3402; fax: +39-081-2455807.
E-mail address: bernardi@alpha.szn.it (G. Bernardi).

0378-1119/00/$ - see front matter
PII: S0378-1119(00)00520-5

compared to their *Xenopus* orthologs. At the protein level, proteins encoded by GC-poor genes show similar amino acid frequencies, whereas those encoded by GC-rich genes show significant changes of amino acids leading to an increased hydrophobicity of proteins (Cruveiller et al., 1999).

It is important to stress that the 15% or so of the genome affected by the compositional transitions was made up by the gene-richest regions of the genome which comprise about 50% of the genes. These regions replicate early in the cell cycle, have a high transcriptional and recombinogenic activity and are endowed with an open chromatin structures (see Bernardi, 2000a,b, for reviews).

The present investigation concerns a detailed analysis of the compositional transition, which was performed by comparing: (i) three sets of orthologous genes, from human/*Xenopus*, chicken/*Xenopus*, and human/chicken, respectively; (ii) two large non-redundant sets of human and chicken genes (see section 2). The comparison concerned the levels of individual nucleotides in third codon positions, the amino acid composition of the encoded proteins and codon usage. Data from ultracentrifugation analyses of DNAs, as well as the compositional properties of genes, indicate that the genomes of *Xenopus,*on one hand, and of human and chicken on the other, are good representatives of the genomes of cold- and warm-blooded vertebrates, respectively (Bernardi, 2000a). Moreover, a large number of sequences from these genomes are available, making it possible to perform reliable statistical analyses on orthologous genes.

2. Materials and methods

Homologous sequences were retrieved from the HOVERGEN Database (Duret et al.,1994) using a computer program written in C language. Datasets of homologous genes were checked in order to remove paralogous and partial genes. Orthologous sequences were identified using the PROTDIST software (Felsenstein, 1989) on the basis of the Kimura (1983) Kimura's (1983) distance $D = -ln(1 - p - 0.2p^2)$, where p is the proportion of amino acids that differ between the two sequences. When more than one couple of homologous genes were identified, only that showing the lowest D-value was retained and the other discarded as paralogous. Only proteins having a size difference lower than, or equal to, 30 residues were used, creating three datasets of orthologous genes which comprised (i) 453 gene pairs from *Xenopus* and human (255 gene pairs >60% GC_3); (ii) 273 gene pairs from *Xenopus* and chicken (168 gene pairs >60% GC_3); and (iii) 643 pairs of genes from human and chicken (343 gene pairs >60% GC_3). The three gene sets were different from each other, the percentage of non-overlapping pairs being equal to 61% for the human/*Xenopus*, 67% for the human/chicken and 15% for the chicken/*Xenopus* comparison (see Fig. 9). The average base compositions of third codon positions, as well as codon frequencies, were determined by using the program ANALSEQ (Gautier and Jacobzone, 1989). The same analyses were performed on two large sets of non-redundant human (7771, 3973 of which >60% GC_3) and chicken (1262, 674 of which >60% GC_3) genes. Redundancies were removed by CLEANUP (Grillo et al., 1996).

3. Results

3.1. Compositional correlations between orthologous genes

The frequencies of C_3 and G_3 (the molar ratio of cytosine and guanine at third codon positions) in the *Xenopus* genes that were orthologous to human and chicken genes, respectively, showed positive and statistically significant correlations, with R values equal to 0.52 for C_3 and ranging from 0.46 to 0.55 for G_3, P-values being lower than 10^{-3}–10^{-5} and slopes being equal to 2.0–2.4 (Fig. 1). Data covered comparable ranges in the two sets of genes, and showed very close averages and standard deviations (Table 1). The corresponding plots for human and chicken genes (Fig. 2), showed high correlation coefficients, 0.65 and 0.62 ($P < 10^{-5}$), as well as slopes equal to 0.90 and 0.92, for C_3 and G_3, respectively.

3.2. Codon frequencies

The results of Fig. 1 corroborate the observation that two

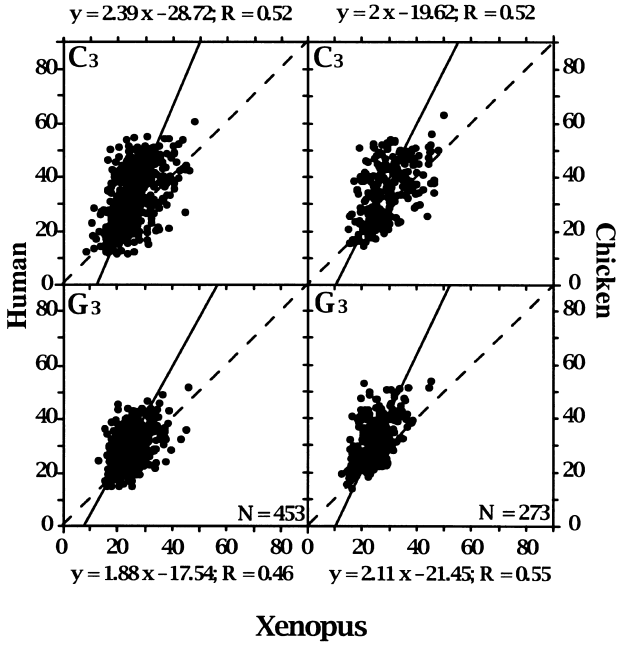

Fig. 1. C_3 and G_3 levels of genes from human and chicken were plotted against those of orthologous genes from *Xenopus*. N is the number of gene pairs analyzed. The correlation coefficients and the equations of the orthogonal regression lines are reported. The broken line is the diagonal.

Table 1
The frequencies of C_3 and G_3 in the *Xenopus* genes that were orthologous to human and chicken genes

Species	Pairs of orthologues	GC_3		C_3		C_3	
		Mean	SD	Mean	SD	Mean	SD
H. sapiens	453	0.63	0.16	0.33	0.11	0.30	0.07
X. Iaevis		0.51	0.10	0.26	0.07	0.25	0.05
G.gallus	273	0.66	0.17	0.35	0.10	0.31	0.84
X. Iaevis		0.52	0.10	0.27	0.07	0.25	0.05
G.gallus	643	0.64	0.17	0.33	0.11	0.31	0.08
H. sapiens		0.64	0.16	0.34	0.11	0.30	0.08

classes of genes can be distinguished in the human genome. The former one is made of GC_3-poor genes that are compositionally close to those of *Xenopus* genes, the latter one, the GC_3-rich genes, shows an increasing GC_3 level in human compared to *Xenopus* (see Bernardi, 2000a, for a review). The threshold between GC_3-poor and GC_3-rich genes falls around 60% GC_3. Indeed, two clear-cut results have been reported: (i) orthologous proteins from *Xenopus* and human higher than 60% GC_3 show large differences at the amino acid level, affecting protein hydrophobicity (Cruveiller et al., 1999); and (ii) the compositional distribution of human genes shows a discontinuity at 60% GC_3 (see Bernardi, 2000a, for a review).

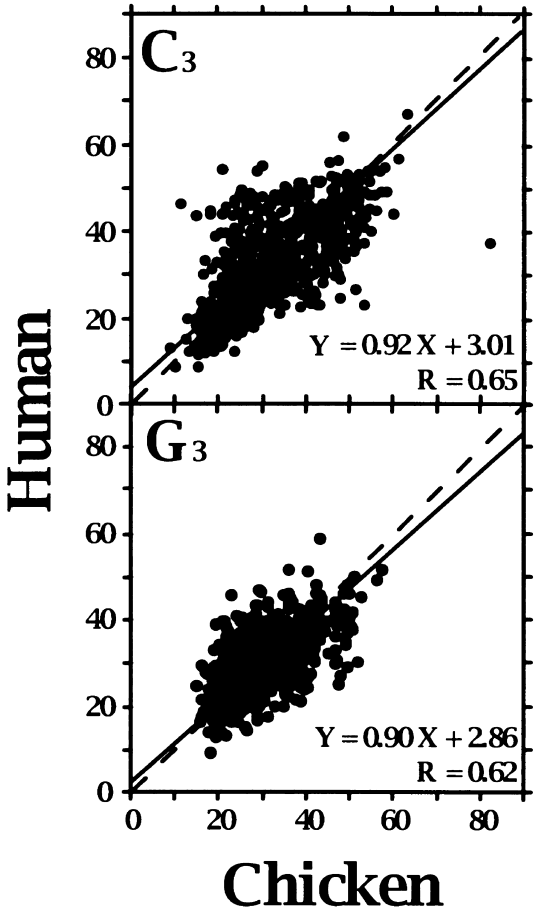

Fig. 2. C_3 and G_3 levels of human genes are plotted against those of orthologous chicken genes. Other indications as in Fig. 1.

In view of the similarity of results involving human and chicken in Figs. 1 and 2, this separation at 60% GC_3 was also applied to the chicken/*Xenopus* and human/chicken comparison.

The sets of orthologous gene pairs from human/*Xenopus* and chicken/*Xenopus* were split into two sub-sets according to the GC_3 levels of the GC-richer organism: the 'GC-poor genes' (0–60% GC_3), and the 'GC-rich genes' (60–100% GC_3). In each sub-set, the average frequencies of all codons, as well as their differences, were calculated for both species (Figs. 3–6).

The relative frequencies of codons of the GC_3-poor genes were very similar in both human/*Xenopus* and chicken/*Xenopus* comparisons. Indeed, differences were very small and their averages were 0.08%(±0.06) and 0.1%(±0.09), respectively.

Also the histograms of codon frequencies of the GC_3-rich genes were very similar in both comparisons, averages being 0.75%(±0.4) and 0.73%(±0.4). However, some differences were observed, mainly among the four-fold degenerate codons. Indeed, in the human/*Xenopus* comparison, the differences of the C-ending codons was higher than that of the G-ending codons, whereas in the chicken/*Xenopus* comparison, C- and G-ending codons showed much closer frequencies.

Finally, the histograms of GC-poor and GC-rich genes from human and chicken were very similar (Figs. 7, 8) and, indeed, the average and standard deviations were 0.21%(±0.15) and 0.18%(±0.12), respectively.

Identical results were obtained comparing non-homologous sets of human and chicken genes (data not shown).

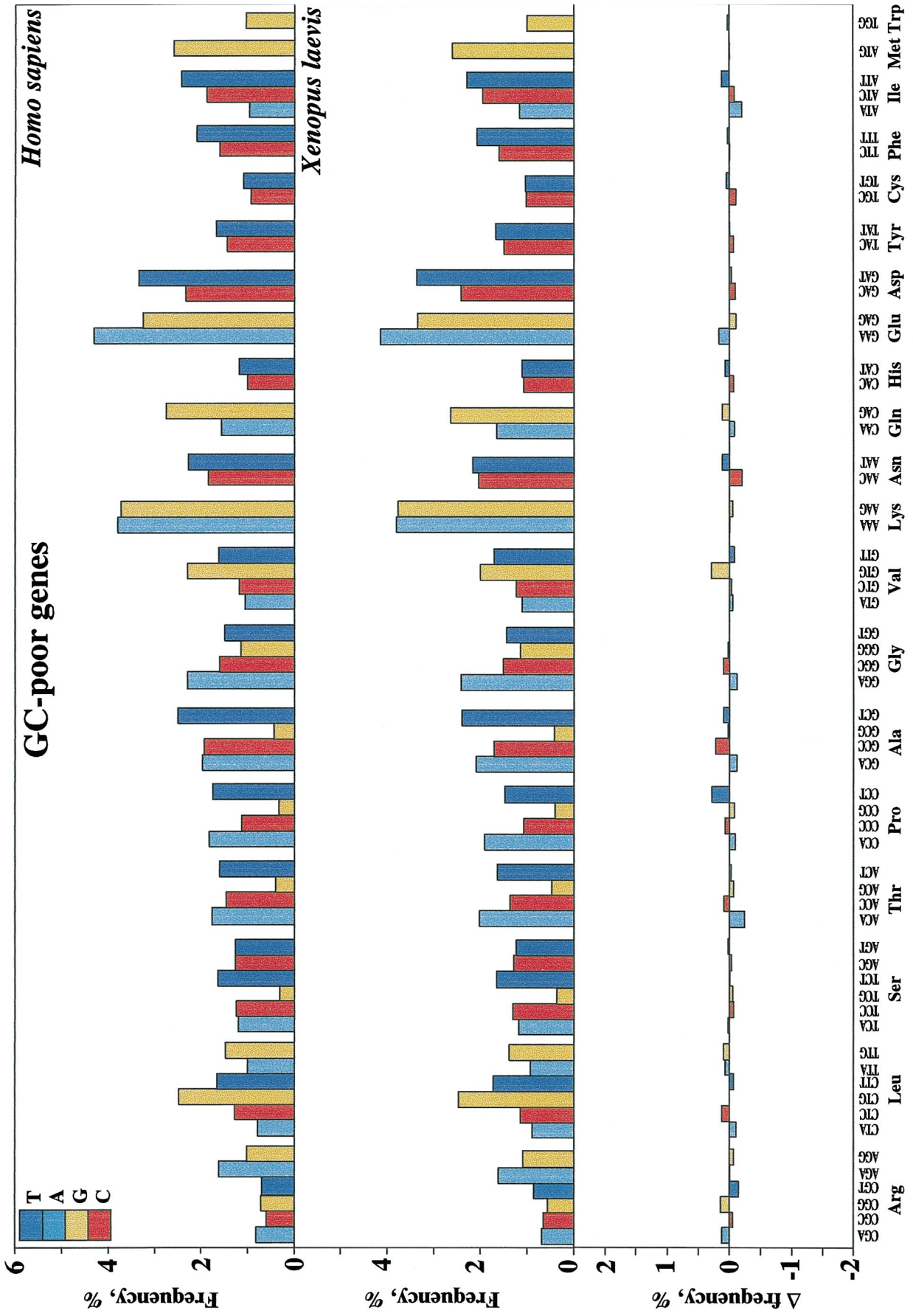

Fig. 3. Codon frequencies of human (top) and *Xenopus* (middle) GC-poor (0–60% GC_3) orthologous genes. The bottom histogram shows the differences in codon frequencies.

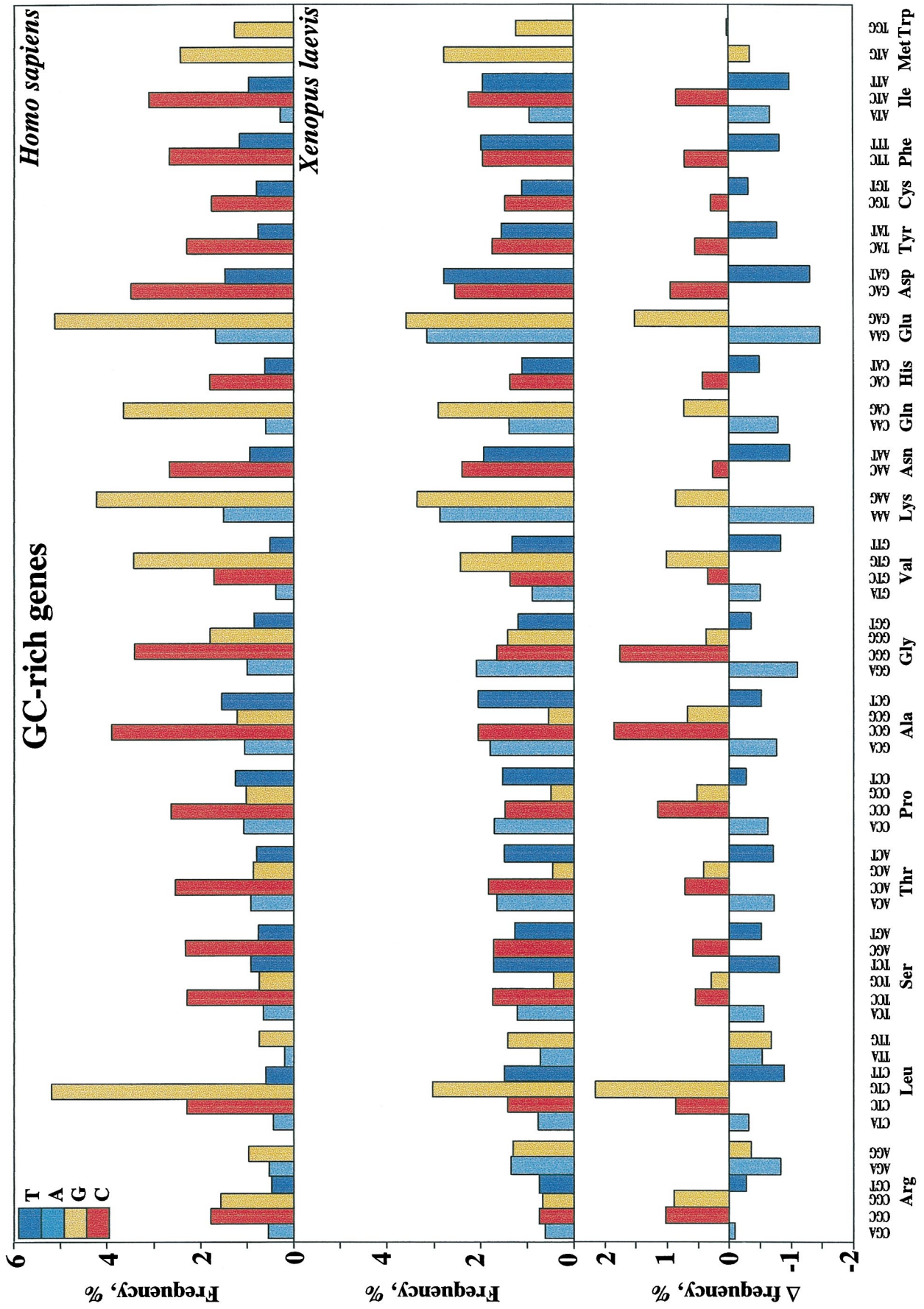

Fig. 4. Codon frequencies of human (top) and *Xenopus* (middle) GC-rich (60–100% GC_3) orthologous genes. The bottom histogram shows the differences in codon frequencies.

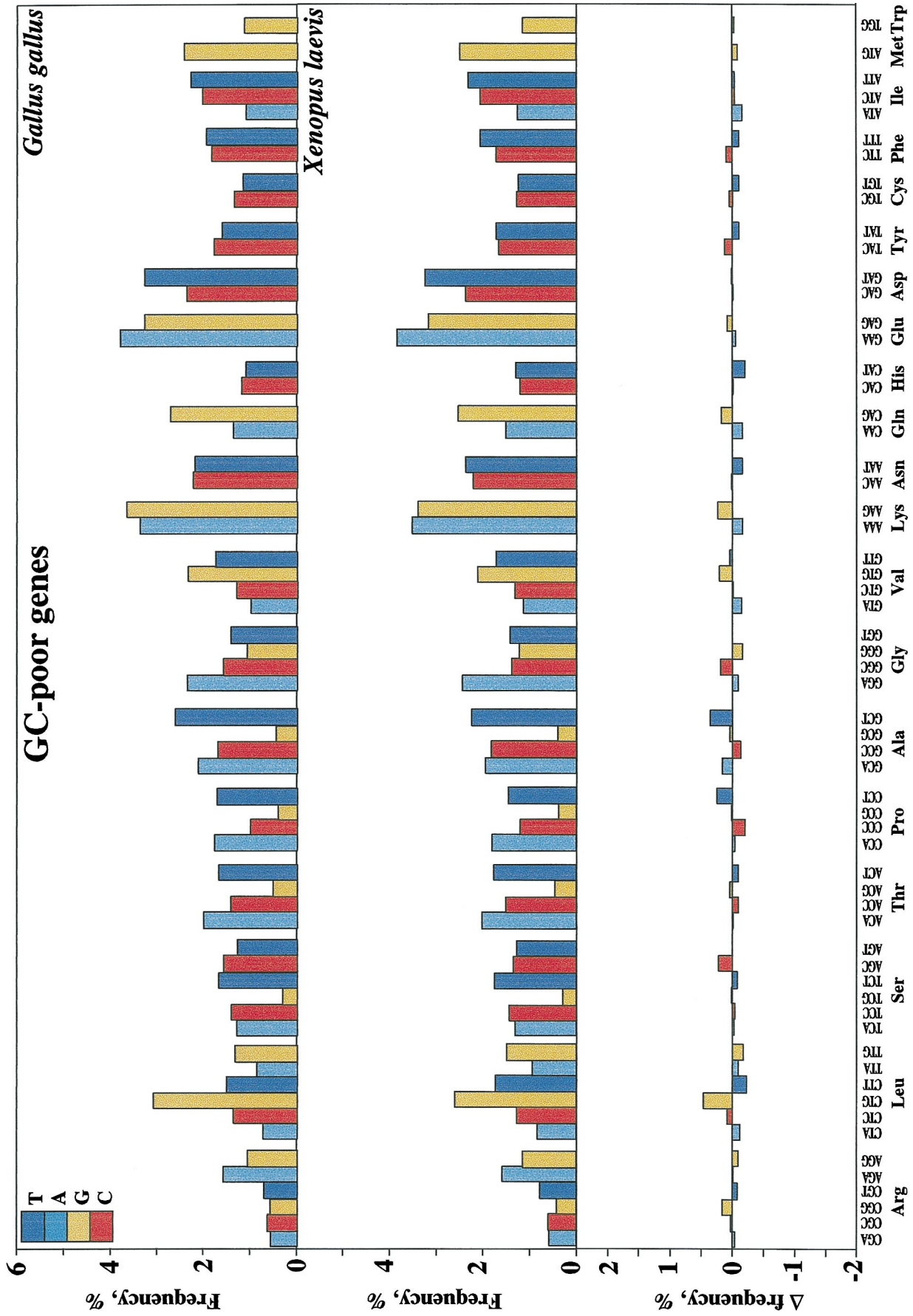

Fig. 5. Codon frequencies of chicken (top) and *Xenopus* (middle) GC-poor (0–60% GC_3) orthologous genes. The bottom histogram shows the differences in codon frequencies.

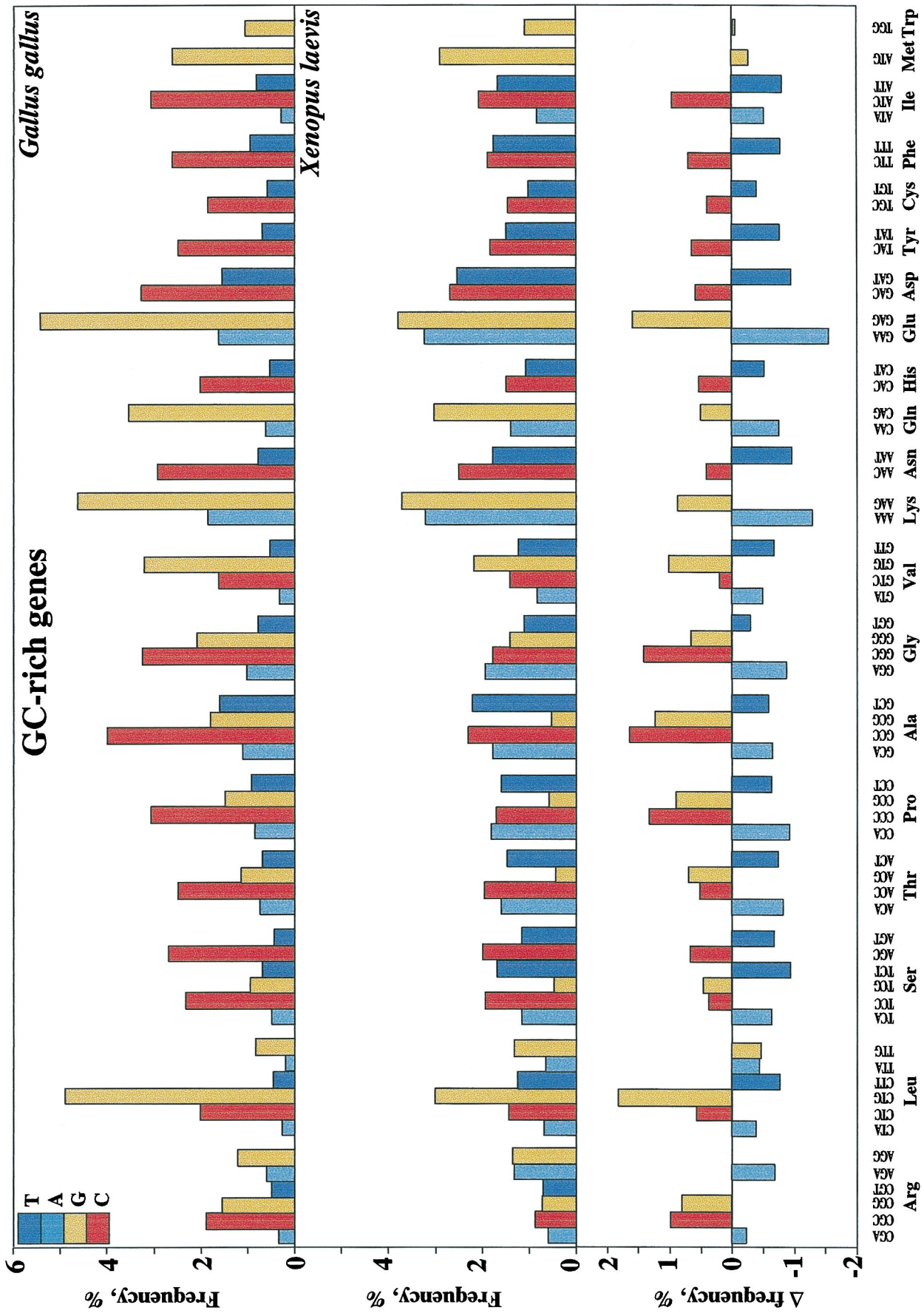

Fig. 6. Codon frequencies of chicken (top) and *Xenopus* (middle) GC-rich (60–100% GC_3) orthologous genes. The bottom histogram shows the differences in codon frequencies.

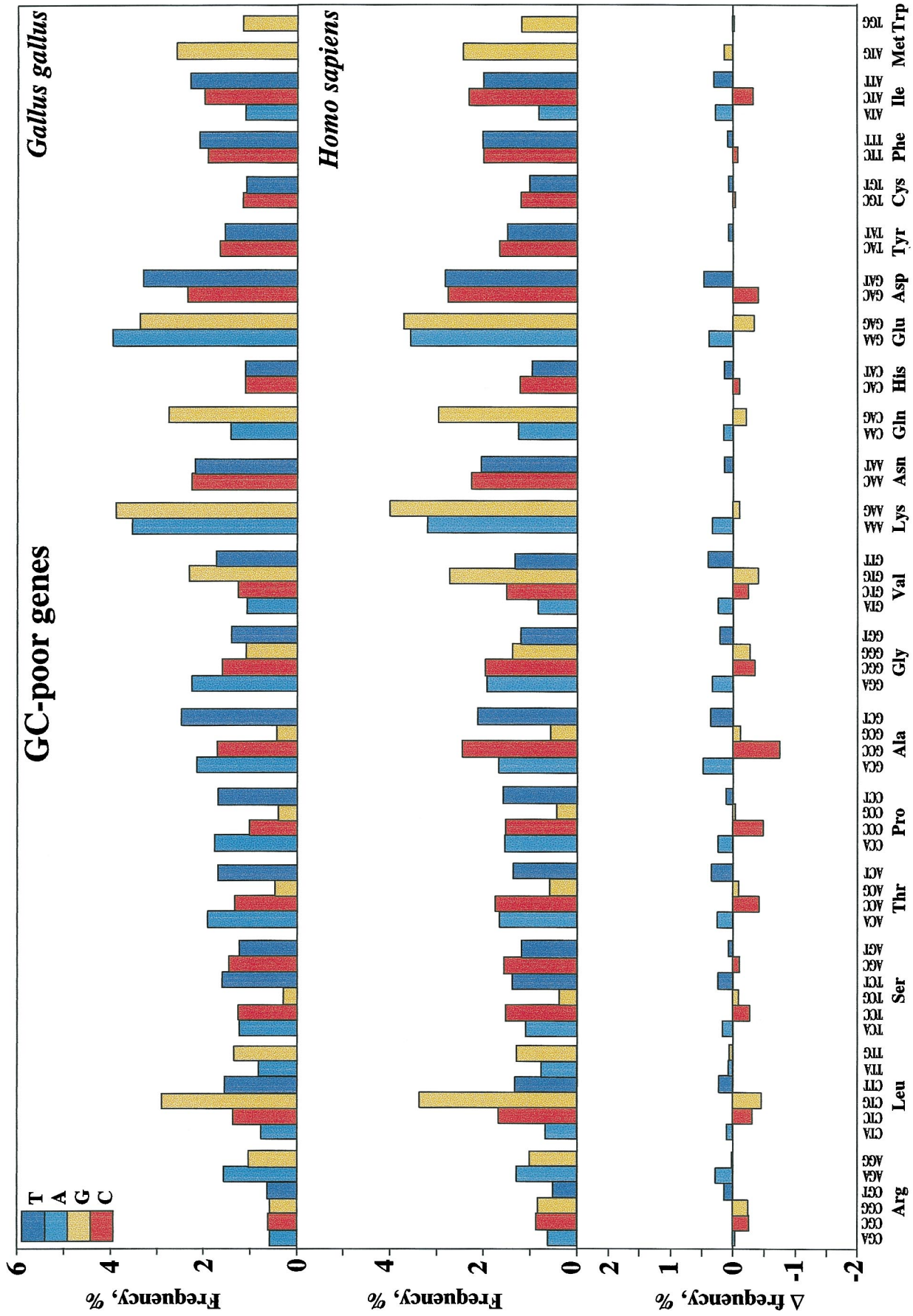

Fig. 7. Codon frequencies of human (top) and chicken (middle) GC-poor (0–60% GC_3) orthologous genes. The bottom histogram shows differences in codon frequencies.

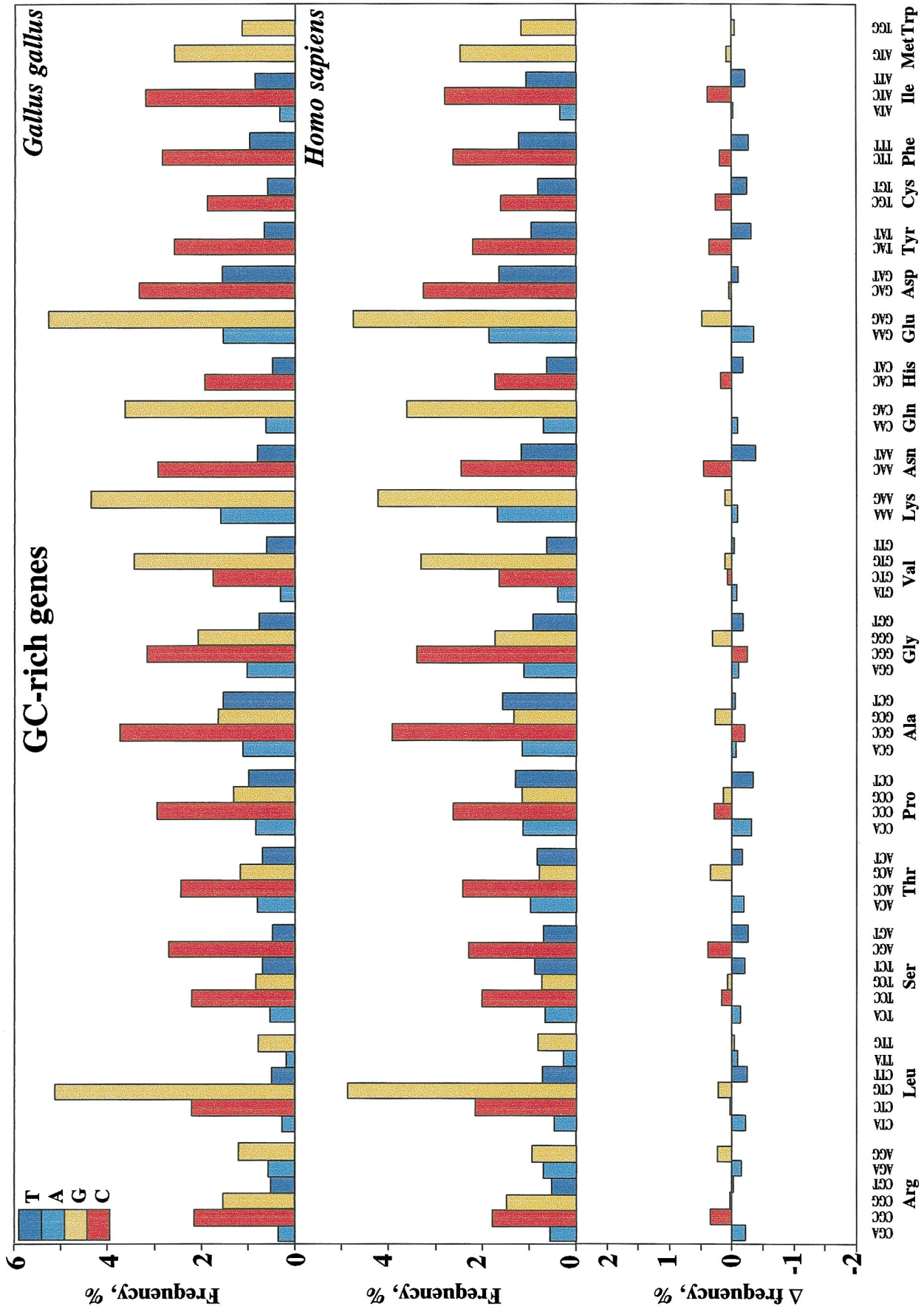

Fig. 8. Codon frequencies of human (middle) and chicken (top) GC-rich (60–100% GC_3) orthologous genes. The bottom histogram shows differences in codon frequencies.

4. Discussion

4.1. Third codon positions

The high correlation coefficients and similar high slopes of the linear regressions of C_3 and G_3 plots of human or chicken versus *Xenopus* (Fig. 1) considerably extend previous data (see Bernardi, 2000a), since (a) the sample sizes were larger; (b) the correlations involving C_3 and G_3 were better than those concerning GC_3; (c) the correlations were very similar for the two plots, in spite of the fact that the samples of orthologous genes were different in the two cases, with a large percentage of non-overlapping genes (see Fig. 9).

In both cases, at low C_3 or G_3, values of warm-blooded vertebrates are similar to values in orthologous genes from *Xenopus*. In contrast, at high C_3 or G_3, values of warm-blooded vertebrates are higher than those of orthologous *Xenopus* genes.

In the case of the human/chicken comparison, which are based on yet another set of orthologous genes (see Fig. 9), correlation coefficients are higher than in the plots of Fig. 1 and slopes are close to unity. In other words, points from both human and chicken orthologous genes tend to be compositionally similar in their third codon positions. Moreover the same result was also obtained from sets of non-homologous genes.

The conclusions that can be drawn from Figs. 1 and 2 confirm and reinforce those previously obtained by comparing GC_3 values of orthologous genes. Indeed, the compositional transitions concerned the GC-rich genes from the gene-dense regions of the genomes. Moreover, they were similar in the independent ancestral lines of mammals and birds, and stopped at least at the time of appearance of present-day mammals and birds.

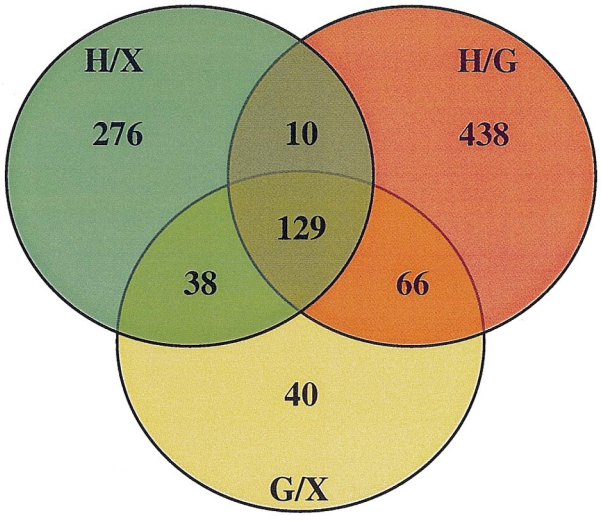

Fig. 9. Degree of overlap of the sets of orthologous genes investigated in this work. Figs. indicate numbers of genes, H/X, H/G and G/X the orthologous sets of human/*Xenopus*, human/chicken and chicken/*Xenopus* genes.

4.2. Amino acid and codon frequencies

As far as GC_3-poor genes are concerned, the results are very clear-cut in that differences in amino acids and codon usage between *Xenopus* and human or chicken are extremely small, confirming the data of Cruveiller et al. (1999). In other words, the average levels of amino acids and the average codon frequencies of the two sets of orthologous data are very similar and the hydropathy of proteins was the same (Cruveiller et al., 1999). Not surprisingly, this also applies to the comparison of GC_3-poor genes from human and chicken.

In contrast, GC_3-rich genes of human and chicken show remarkable differences in the levels of encoded amino acids and codon frequencies relative to *Xenopus* and an increase in the hydrophobicity of proteins. Very interestingly, such differences are, however, very small when chicken and human data are compared.

It is remarkable that codon choices are the same even for quartet codons. Indeed, it could be claimed that the similarity in codon preferences between human and chicken for each set of genes is nothing but the predictable result of the fact that genes with similar GC_3 levels are compared. Therefore one might expect for the case of GC_3 data set, that G- and C-ending codons exhibit higher values in both species. However, an examination of Figs. 4 and 6 show that codon frequencies in quartets is the same for both species. For instance in the leucine codon group, CTG was by far the most frequent codon in both species, while CTC is the second in the order of frequency, while in the glycine codon group, GGC is the most frequent codon in both species. The same comparison for the remaining codon groups and also for the case of GC_3-poor data set clearly indicates that the similarity in codon frequencies go much beyond what would be expected from the similarity in GC_3 levels.

It is important to stress that CsCl profiles and GC_3 plots of avian genomes indicate an extremely high level of similarity among different orders (Kadi et al., 1993; Mouchiroud and Bernardi, 1993). This similarity also applies to the mammalian genomes belonging in the ‘general pattern’ (Sabeur et al., 1993; Mouchiroud and Bernardi, 1993), namely to most mammalian genomes (except for the genomes of murids). In other words, what has been shown here for human and chicken genes is generally valid for avian genomes and for the genomes of most mammals. Unfortunately, the scarcity of sequence data prevents us from making a similar statement for the genomes of cold-blooded vertebrates. Preliminary indications suggest, however, that this might be the case. If so, the comparisons analyzed here should apply to cold- and warm-blooded vertebrates, in general.

4.3. One or two transitions leading to the genomes of warm-blooded vertebrates?

The simplest explanation for the striking similarity of mammalian/avian data is that this is the consequence of a

compositional transition in the common ancestor of mammals and birds. In fact, this was claimed to be the case (Hughes et al., 1999) on the basis of the similarity of GC_3 levels of genes from turtle and crocodile with their chicken orthologs. Unfortunately, the data of Hughes et al. (1999) are not acceptable, because they show no correlation with the orthologous *Xenopus* genes, as it is expected. This suggests that they either concern outliers, or, that the plots are simply unreliable because of the sample being extremely small and the coding sequences only partial. Ironically, what the results of Hughes et al. (1999) failed to prove, had already been demonstrated by Aïssani and Bernardi (1991a,b; a report apparently overlooked by Hughes et al., 1999). Indeed, these authors provided evidence for crocodiles and turtles displaying a genome heterogeneity that mimicks that of warm-blooded vertebrates, yet genes from turtles and crocodiles lack such diagnostic features of the genomes of mammals and birds as CpG islands.

This heterogeneity cannot, however, be taken as an evidence for a single transition, because of the very homogeneous genomes of squamates (see below).

Incidentally, the proposal of Hughes et al. (1999) that this single transition took place before the separation of amniota, perhaps in the common ancestor of tetrapods, between 450 and 350 Mya, present extremely serious problems. Indeed, there is no evidence of a compositional transition between fishes and the common ancestor of tetrapods (see Bernardi and Bernardi, 1990a,b). But even if there were one, this would not account for the transition under discussion here, which concerns the genomes of *Xenopus*, as a representative of cold-blooded vertebrates, and of warm-blooded vertebrates.

The ancient *Haemothermia* hypothesis, which puts mammals and birds in sister groups (a hypothesis revived by Hedges et al., 1990), whose common ancestor might have undergone the transition, does not deserve any discussion here since this hypothesis has been disposed of by the phylogenies of Figs. 10 and 11.

Indeed, if we consider the 'classical' phylogenetic tree of amniotes (Fig. 10), the problem with the single transition is that the squamates (snakes and lizards) are endowed with genomes that are typically 'cold-blooded', namely remarkably homogeneous in terms of base composition. Since, according to Fig. 10, the ancestors of squamates were endowed with 'warm-blooded', heterogeneous genomes, the squamate genomes should be the result of a process of extreme compositional homogeneization for which there is no example among vertebrate genomes. Indeed, only one case (the case of murids) is known in which an increased mutation rate has led to a decreased genome heterogeneity compared to other mammals, yet this homogeneization did not reach by far, the very high homogeneity of the genomes of squamates (see Bernardi, 2000a,b). If squamates are at the root of the reptilian tree (Fig. 11), as strongly suggested by recent data (Hedges and Poling, 1999), a major difficulty with the single transition is that squamates, as already mentioned, exhibit very homogeneous genomes, the amounts of GC-rich components being comparable to those of the *Xenopus* genome. In fact the tree strongly supports two independent compositional transitions, to mammals and birds.

A last point about the single compositional transition should be stressed. Indeed, the similarity of third codon positions, of average amino acid composition of proteins and of codon frequencies would have lasted not over 100 million years (the radiation time of mammalian orders), but over more than 300 million years (since the time of the common ancestor of warm-blooded vertebrates; see Benton, 1990).

A last-ditch attempt to save the regional mutational bias hypothesis (see Bernardi, 2000b) as an explanation of the findings reported here would be to consider: (i) that the orthologous genes investigated are syntenic; and (ii) that the same mutational biases in the replication machinery

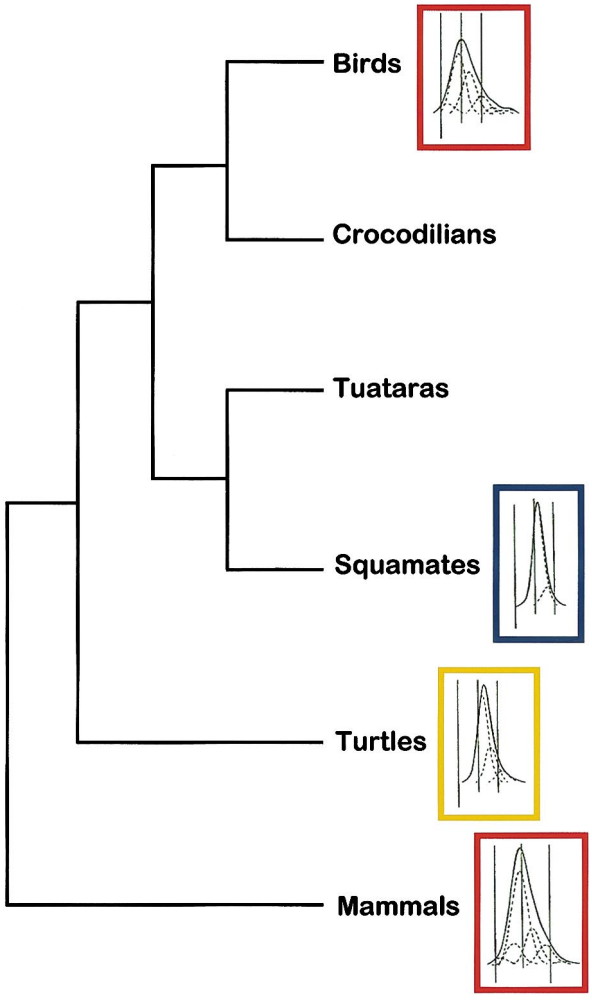

Fig. 10. The classical phylogeny of living reptiles based on morphology and fossil reports (from Hedges and Poling, 1999) is compared with the CsCl profiles (from Thiery et al., 1976) of representative species *T. graeca, G. gallus, I. iguana* and *H. sapiens* (the three vertical bars correspond to buoyant densities of 1.690, 1.700 and 1.710 g/cm^3, respectively).

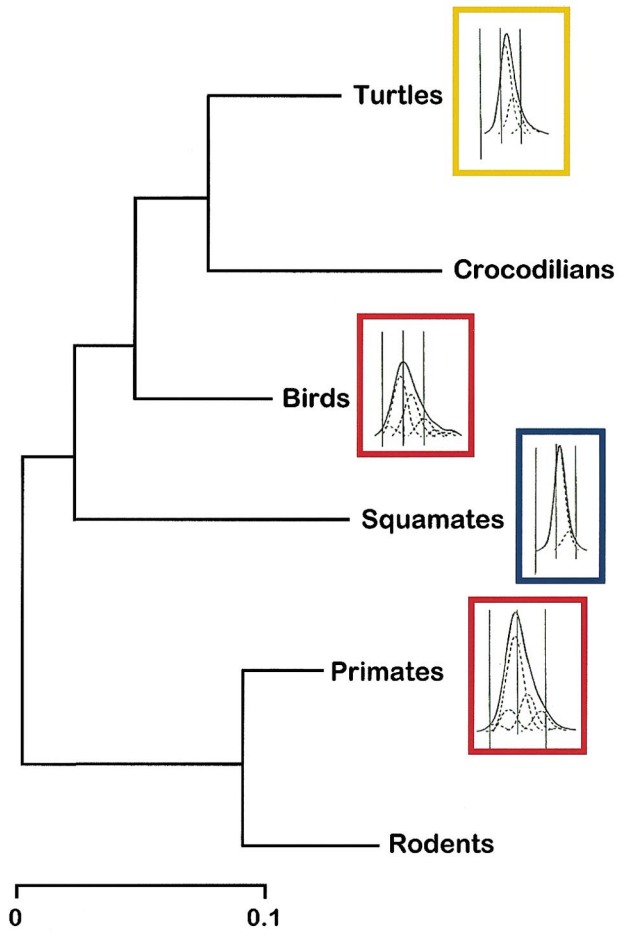

Fig. 11. The maximun likelihood phylogeny of combined sequences from 11 nuclear proteins, 1.943 amino acids, (from Hedges and Poling, 1999) is compared with the CsCl profiles (from Thiery et al., 1976). Scale bar indicates amino acids substitutions per site. See also legend of Fig. 10.

caused the same effects, and maintained them in very similar chromatin environments. However, since the same results were also obtained from non-homologous sets of genes, synteny can be excluded as an explanation of the present results. Moreover, the regional mutation bias hypothesis has several problems (see Bernardi, 2000b), the major one being the preference for CTG vs. CTC and of GTG vs. GTC in the codons of leucine and valine respectively, as opposed to the preference for C-endings in all other codons (see Figs. 4 and 6). In other words, the same selection pressures would have been operating over the vertebrates genomes not only over such an extremely long time, but over genomes as different as those of reptiles on the one hand, and mammals and birds on the other.

We do not see, therefore, any valid alternative hypothesis to that of two independent compositional transitions. The convergence of the transitions and their maintenance are really striking in view of the present results. In fact, those finding cannot be reconciled at all with the mutational bias hypothesis, adding one more difficulty to the long series of problems associated with that proposal (see Bernardi, 2000a,b, for a more detailed discussion).

In contrast, the results can be explained by natural selection. The proposal that a major selective advantage accounting for the compositional transition is increased body temperatures, may possibly also account for the increased compositional heterogeneity of the genomes of some reptiles (*T. graeca, C. niloticus, C. cerastes*; Bernardi and Bernardi 1990a,b; Aïssani and Bernardi, 1991a,b).

Acknowledgements

We wish to thank our colleague Fernando Alvarez and Adam Eyre-Walker for comments on this paper.

References

Aïssani, B., Bernardi, G., 1991a. CpG islands: features and distribution in the genome of vertebrates. Gene 106, 173–183.

Aïssani, B., Bernardi, G., 1991b. CpG islands, genes and isochores in the genome of vertebrates. Gene 106, 185–195.

Benton, M.J., 1990. Phylogeny of the major tetrapod groups, morphological data and divergence dates. J. Mol. Evol. 30, 409–424.

Bernardi, G., 2000a. Isochores and the evolutionary genomics of vertebrates. Gene 241, 3–17.

Bernardi, G., 2000b. The compositional evolution of vertebrate genomes. Gene (in press).

Bernardi, G., Bernardi, G., 1990a. Compositional patterns in the nuclear genomes of cold-blooded vertebrates. J. Mol. Evol. 31, 265–281.

Bernardi, G., Bernardi, G., 1990b. Compositional transitions in the nuclear genomes of cold-blooded vertebrates. J. Mol. Evol. 31, 282–293.

Bernardi, G., Bernardi, G., 1991. Compositional properties of nuclear genes from cold-blooded vertebrates. J. Mol. Evol. 33, 57–67.

Bernardi, G., Olofsson, B., Filipski, J., Zerial, M., Salinas, J., Cuny, G., Meunier-Rotival, M., Rodier, F., 1985. The mosaic genome of warm-blooded vertebrates. Science 228, 953–958.

Bernardi, G., Mouchiroud, D., Gautier, C., Bernardi, G., 1988. Compositional patterns in vertebrate genomes: conservation and change in evolution. J. Mol. Evol. 28, 7–18.

Cruveiller, S., D'Onofrio, G., Jabbari, K., Bernardi, G., 1999. Different hydrophobicities of orthologous proteins from *Xenopus* and man. Gene 238, 15–21.

D'Onofrio, G., Jabbari, K., Musto, H., Alvarez-Valin, F., Cruveiller, S., Bernardi, G., 1999. Evolutionary genomics of vertebrates and its implications. Ann. New York Acad. Sci. 18, 81–94.

Duret, L., Mouchiroud, D., Gouy, M., 1994. HOVERGEN: Homologous Vertebrate Genes data base. Nucleic Acids Res. 22, 2360–2363.

Felsenstein, J., 1989. Phylogeny Inference Package (Version 3.527). Cladistics 5, 164–166.

Gautier, C., Jacobzone, M., 1989. <http://biom3.univlyon.fr:8080/doclogi/docanals/manuel.html>, Publication interne, UMR CNRS 5558 Biométrie, Génétique et Biologie des Populations, Universite Claude Bernard - Lyon I, France.

Grillo, G., Attimonelli, M., Liuni, S., Pesole, G., 1996. CLEANUP: a fast computer programme for removing redudancies from nucleotide sequence databank. CABIOS 12, 1–8.

Hedges, S.B., Moberg, K.D., Maxson, L.R., 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., Poling, L.L., 1999. A molecular phylogeny of reptiles. Science 283, 998–1001.

Hughes, S., Zelus, D., Mouchiroud, D., 1999. Warm-blooded isochores structure in nile crocodile and turtle. Mol. Biol. Evol. 16, 1521–1527.

Kadi, F., Mouchiroud, D., Sabeur, G., Bernardi, G., 1993. The compositional patterns of the avian genomes and their evolutionary implications. J. Mol. Evol. 37, 544–551.

Kimura, M., 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge, U.K.

Mouchiroud, D., Fichant, G., Bernardi, G., 1987. Compositional compartmentalization and gene composition in the genome of vertebrates. J. Mol. Evol. 26, 198–204.

Mouchiroud, D., Gautier, C., Bernardi, G., 1988. The compositional distribution of coding sequences and DNA molecules in humans and murids. J. Mol. Evol. 27, 311–320.

Perrin, P., Bernardi, G., 1987. Directional fixation of mutations in vertebrate evolution. J. Mol. Evol. 26, 301–310.

Sabeur, G., Macaya, G., Kadi, F., Bernardi, G., 1993. The isochore patterns of mammalian genomes and their phylogenetic implications. J. Mol. Evol. 37, 93–108.

Thiery, J.P., Macaya, G., Bernardi, G., 1976. An analysis of eukaryotic genomes by density gradient centrifugation. J. Mol. Biol. 108, 219–235.