J Mol Evol (1999) 49:376–384

JOURNAL OF MOLECULAR EVOLUTION

Intron–Genome Size Relationship on a Large Evolutionary Scale

Alexander E. Vinogradov

Institute of Cytology, Russian Academy of Sciences, St. Petersburg 194064, Russia

Received: 23 October 1997 / Accepted: 14 April 1999

Abstract. The intron–genome size relationship was studied across a wide evolutionary range (from slime mold and yeast to human and maize), as well as the relationship between genome size and the ratio of intervening/coding sequence size. The average intron size is scaled to genome size with a slope of about one-fourth for the log-transformed values; i.e., on the global scale its increase in evolution is lower than the increase in genome size by four orders of magnitude. There are exceptions to the general trend. In baker's yeast introns are extraordinarily long for its genome size. Tetrapods also have longer introns than expected for their genome sizes. In teleost fish the mean intron size does not differ significantly, notwithstanding the differences in genome size. In contrast to previous reports, avian introns were not found to be significantly shorter than introns of mammals, although avian genomes are smaller than genomes of mammals on average by about a factor of 2.5. The extra-/intragenic ratio of noncoding DNA can be higher in fungi than in animals, notwithstanding the smaller fungal genomes. In vertebrates and invertebrates taken separately, this ratio is increasing as the increase in genome size. Two hypotheses are proposed to explain the variation in the extra-/intragenic ratio of noncoding DNA in organisms with similar numbers of genes: transition (dynamic) and equilibrium (static). According to the transition model, this variation arises with the rapid shift of genome size because the bulk of extragenic DNA can be changed more rapidly than the finely interspersed intron sequences. The equilibrium model assumes that this variation is a result of selective adjustment of genome size with constraints imposed on the intron size due to its putative link to chromatin structure (and constraints of the splicing machinery).

Key words: Genome evolution — Junk DNA — Noncoding DNA — C-value paradox — Intervening sequence — Chromatin structure — Metabolic rate — Mammals — Birds — Fungi — Baker's yeast

Introduction

The biological significance of redundant (apparently noncoding) DNA, which constitutes the greater part of the eukaryotic genome, remains puzzling. Introns inside eukaryotic genes, which are transcribed together with coding sequences and thereby incur a double metabolic cost, are especially intriguing. It was reported that introns are generally smaller in the chicken than in the human, which was related to the smaller avian genomes (Duret et al. 1995; Hughes and Hughes 1995; Oliver and Marin 1996). However, the general relationship between intron and genome size on a wide evolutionary scale remains unknown. This relationship is interesting because it reflects the evolution of two types of noncoding DNA, residing inside and outside of genes, which may be subjected to different selection constraints. The intron–genome size relationship, as well as the relationship between genome size and the ratio of intervening to coding sequence size (IVS/CDS), is studied here across a wide evolutionary range of eukaryotes.

Materials and Methods

Genome size values were taken from the literature as indicated in Table 1. Sequencing data on genes of 27 species were extracted from GenBank (total, 7584 introns and 2009 genes). The positions of introns were taken from the sequence annotations. Intron sizes were calculated

Table 1. Summary of the data set

Species	Genome size (1c; pg) and reference[a]	Intron size (bp; log)		Number of introns (genes) studied	Mean IVS size (bp; log)	Mean IVS/CDS ratio (log)
		Mean	SE			
Bos taurus (cow)	3.70 (a)	2.5123	0.0442	94 (35)	2.9044	0.2251
Canis familiaris (dog)	2.90 (a)	2.5132	0.0470	88 (32)	3.1970	0.3538
Cavia porcellus (quinea pig)	3.90 (a)	2.7336	0.0888	24 (14)	2.7525	0.1389
Equus caballus (horse)	3.15 (b)	2.5252	0.0683	36 (14)	2.8742	0.1865
Homo sapiens (human)	3.50 (a)	2.5794	0.0195	708 (164)	3.1837	0.3487
Mus musculus (mouse)	3.25 (a)	2.4501	0.0258	219 (62)	2.9380	0.1846
Oryctolagus cuniculus (rabbit)	3.25 (a)	2.6088	0.0421	109 (35)	3.1221	0.2893
Rattus norvegicus (rat)	3.05 (a)	2.3904	0.0314	165 (50)	2.8681	0.1356
Sus scrofa (pig)	3.20 (a)	2.5339	0.0288	252 (59)	3.1829	0.3358
Gallus gallus (chicken)	1.25 (a)	2.4924	0.0239	295 (64)	3.1587	0.3117
Xenopus laevis (clawed frog)	3.55 (a)	2.6888	0.0351	152 (56)	3.1845	0.3823
Brachydanio rerio (zebrafish)	1.80 (a)	2.3121	0.0534	50 (25)	2.7721	−0.1911
Cyprinus carpio (carp)	1.70 (b)	2.1677	0.0342	76 (25)	2.6791	−0.0833
Fugu rubripes (pufferfish)	0.415 (c)	2.2150	0.0176	430 (44)	3.1760	0.1105
Oncorhynchus mykiss (rainbow trout)	2.55 (a)	2.2523	0.0450	61 (15)	2.9833	0.1063
Drosophyla melanogaster (fruitfly)	0.17 (c)	2.1275	0.0217	556 (125)	2.7583	−0.3519
Caenorhabditis elegans (nematode)	0.10 (c)	2.0923	0.0169	840 (150)	2.8458	−0.2470
Arabidopsis thaliana (thale cress)	0.15 (d)	2.0528	0.0075	955 (138)	2.8437	−0.2115
Glycine max (soybean)	1.25 (d)	2.3535	0.0271	172 (60)	2.8325	−0.1723
Lycopersicon esculentum (tomato)	1.00 (d)	2.3175	0.0279	238 (68)	2.8853	−0.1969
Oryza sativa (rice)	0.48 (d)	2.2625	0.0193	392 (96)	2.8722	−0.1406
Zea mays (maize)	2.75 (d)	2.2334	0.0201	388 (80)	2.8858	−0.1578
Aspergillus nidulans (mold)	0.032 (e)	1.7945	0.0075	413 (140)	2.2040	−0.9610
Neurospora crassa (fungus)	0.048 (f)	1.9285	0.0121	302 (140)	2.1983	−0.8900
Saccharomyces cerevisiae (baker's yeast)	0.0137 (g)	2.3312	0.0333	95 (92)	2.3464	−0.3912
Schizosaccharomyces pombe (fission yeast)	0.0153 (g)	1.8562	0.0161	310 (128)	2.2472	−0.7714
Dictyostelium discoideum (slime mold)	0.048 (h)	2.0639	0.0164	164 (98)	2.2797	−0.8177

[a] a, Vinogradov (1998a); b, Tiersch et al. (1989); c, Brenner et al. (1993); d, Bennett and Leicht (1995); e, Brody and Carbon (1989); f, Orbach et al. (1988); g, Goffeau et al. (1996); h, Firtel and Bonner (1972).

as average log-transformed lengths. The intervening sequence sizes were calculated as sums of all the internal introns of a given gene (only for the complete gene sequences) and were log-transformed afterward. The analyses were done with StatGraphics software package (plotting, with SigmaPlot).

Results and Discussion

The average intron size increases as the increase in genome size with a slope of about one-fourth for log-transformed values (Fig. 1A). The intercept is indistinguishable from zero (which indicates that prokaryotic genomes without introns do not conform to this regression line). There is one spectacular exception to the whole trend, baker's yeast (*Saccharomyces cerevisiae*), which lies outside the prediction limits of the regression (performed with the inclusion of this point). When this outlier was excluded, the slope was slightly higher, with the intercept still indistinguishable from zero. The ratio of intervening/coding sequence size increases more rapidly with the increase in genome size (Fig. 1B) because of the increase in the ratio of intron number to CDS size (Fig. 2A).

The larger introns of the baker's yeast genes, compared to other fungi, were noted previously (Zhang and Marr 1994). It is interesting that only 4% of the baker's yeast genes contain introns, whereas about 40% of the fission yeast genes are interrupted (Goffeau et al. 1996). Furthermore, the baker's yeast genes usually contain only one intron, whereas on average there are more than two introns in the fission yeast gene. Therefore *S. cerevisiae* is not as exceptional in the IVS/CDS ratio-vs-genome size plot (Fig. 1B) as in the intron-vs-genome size plot (Fig. 1A). The coding sequences of genes containing introns are comparatively short in baker's yeast, and this organism conforms with the general trend in the intron number/CDS-vs-genome size plot (Fig. 2A). As a result of the low total number of introns (however long), the baker's yeast genome shows an exceptionally high ratio of extra-/intragenic noncoding DNA (Table 2). Although *S. cerevisiae* is supposed to have diverged from *S. pombe* as early as from animals (Russell and Nurse 1986), its exceptional position on the intron-vs-genome size plot cannot be explained by this early divergence. The slime mold (*Dictyostelium*), a sister group to all the so-called crown eukaryotes—animals, plants, and fungi

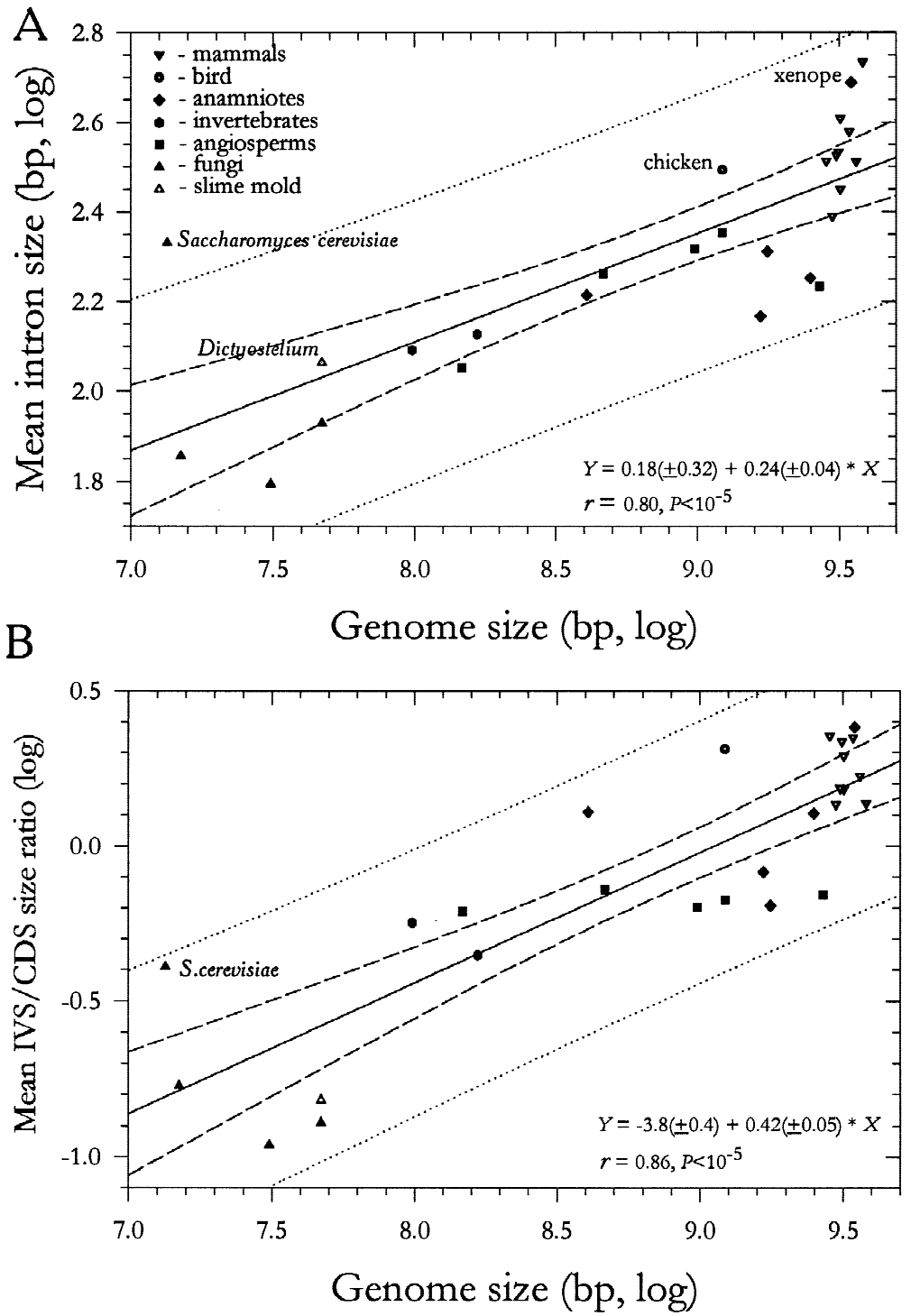

Fig. 1. **A** Regression of mean intron size on genome size for the species studied. When the outlier (baker's yeast) was excluded, the regression-line equation was $Y = -0.31\ (\pm 0.28) + 0.29\ (\pm 0.03) * X$ ($r = 0.89$, $p < 10^{-5}$). **B** Regression of the ratio of mean intervening/coding sequence size on genome size for the species studied (for completely sequenced genes). Dashed lines, confidence limits; dotted lines, prediction limits (both for $p = 0.95$).

(Knoll 1992; Sogin and Patterson 1995), is the most remote organism in the data set but it conforms well with the general trend (Fig. 1A).

The low total amount of intragenic noncoding DNA in the baker's yeast genome may be related to the fact that chromosomes of this organism do not undergo condensation in the course of the cell cycle (Russell and Nurse 1986). It was supposed that intervening sequences are involved in DNA–protein interaction securing the structure of chromatin since exons are under selection pres-

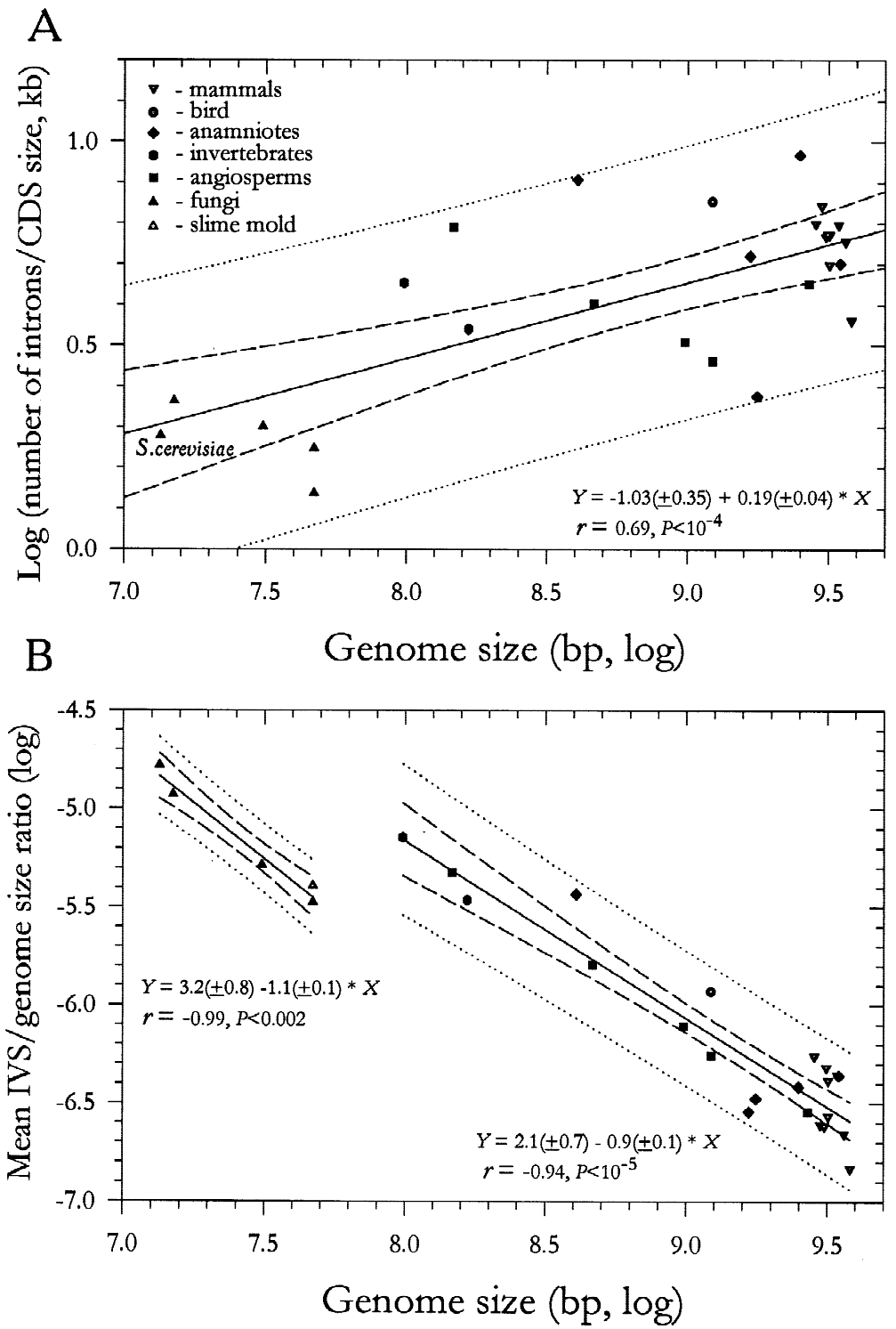

Fig. 2. **A** Regression of the ratio of mean number of introns per gene/coding sequence size on the genome size for the species studied (for completely sequenced genes). **B** Regression of the ratio of mean intervening sequence/genome size on the genome size (for completely sequenced genes). Dashed and dotted lines are as in the legend to Fig. 1.

sure for informational content and their adaptation to the structure of chromatin should be limited (Zuckerkandl 1981, 1997; Zuckerkandl and Hennig 1995). The longer intervening sequences may indicate the tighter chromatin condensation by analogy with the longer nucleosome spacer (linker), which occurs in more compact (and less actively transcribed) chromatin (Villeponteau et al. 1992; Blank and Becker 1996). It is noteworthy that the maximum chromosome condensation is fivefold lower in fission yeast than in mammals (Russell and Nurse 1986). The overall amount of intervening sequences in the genome may characterize the general degree of chromatin

Table 2. Fractions of coding (CDS), intervening (IVS), and extragenic DNA in genomes of different species[a]

Species	Total CDS	Total IVS	Extragenic DNA	Extra-/intragenic ratio of noncoding DNA
Mammals	0.02	0.05	0.93	20
Chicken	0.04	0.09	0.87	10
Xenope	0.02	0.05	0.93	20
Pufferfish	0.22	0.28	0.50	2
Fruitfly	0.20	0.09	0.71	8
Nematode	0.31	0.18	0.51	3
Fission yeast	0.61	0.03	0.36	12
Baker's yeast	0.70	0.004	0.30	75

[a] These rough estimates were calculated on the assumptions that there are about 75,000 genes in vertebrates and 25,000 genes in invertebrates (Bird 1995), 6000 genes (4% of which contain introns) comprise 70% of the genome in *S. cerevisiae*, and the density of genes (40% of which contain introns) is 1.15-fold lower in *S. pombe* (Goffeau et al. 1996).

condensation, whereas the local IVS/CDS ratio may be related to chromatin compactness around a given gene. The longer introns occur in GC-poor isochores of vertebrate genomes (Duret et al. 1995). These isochores are known to be gene-poor, later replicated, less actively transcribed, and located in more condensed chromatin (Mouchiroud et al. 1987; Holmquist 1989; Bernardi 1993; Saccone et al. 1993). Parenthetically, these isochores also contain the long interspersed repeats of noncoding DNA, in contrast to the short repeats in the GC-rich isochores (Holmquist 1989). Thus, intron size may play a role in regulation of gene activity via its relation to chromatin condensation.

The scaling of intron size to genome size with a slope of about one-fourth is intriguing. This is a slope of allometric scaling of many organismal-level traits to body size over a wide evolutionary range—homeothermic and poikilothermic animals and unicellular eukaryotes (Kleiber 1961; Hemmingsen 1960; Schmidt-Nielsen 1984; Elgar and Harvey 1987; Wootton 1987; Charnov 1991). Therefore it seems to be one of the fundamental biological constants. The theoretical basis for the "quarter-power law" was proposed recently on the grounds of fractal geometry (West et al. 1997; see also Williams 1997). Exactly the same slope was observed in the regression of cell doubling time on genome size in plant cells and unicellular eukaryotes (Shuter et al. 1983). Significantly, the cell doubling time of prokaryotes, which do not have introns (nor chromatin structure), is not correlated with genome size [$r = 0.27$, $p = 0.4$, calculated on the grounds of data of Shuter et al. (1983)]. It is tempting to speculate that intron size may be related to cell doubling time through its link to chromatin condensation.

The ratio of average intervening sequence size to genome size decreases with an increase in genome size (Fig. 2B), which suggests that extragenic DNA accumulates more rapidly than intervening DNA. Two regression lines can be seen in the plot, one for fungi and slime mold and the other for animals and plants. The gap between them is probably due to the greater number of genes in genomes of the latter group. Owing to the difference in gene numbers, the extra-/intragenic ratio of noncoding DNA can be even higher in fungi than in animals, notwithstanding the much smaller fungal genomes (Table 2). However, in vertebrates and invertebrates taken separately (presumably with the same number of genes within a group), this ratio is increases with an increase in genome size (Table 2). The same can be suggested for angiosperm plants (Fig. 2B).

After the hierarchical averaging of mean intron and genome sizes over taxonomic levels up to the phylum (with the exception of chordates), the fungi turn out to conform with the general trend, whereas tetrapods show comparatively longer introns for their genome sizes (Fig. 3). Fish seem to conform to the general trend (therefore they were not pooled with tetrapods). If introns do participate in regulation of gene activity via their link to chromatin condensation, the nonproportionally large introns of tetrapods may suggest a more complex genetic machinery with a tighter compaction of chromatin. It is unlikely to be related to the higher sectorial heterogeneity in GC percentage (isochore organization), revealed in genomes of warm-blooded vertebrates (Bernardi et al. 1985, 1997; Bernardi 1993), since the xenope has introns similar in size to those of warm-blooded vertebrates (Fig. 1A). On the other hand, the relationship between genome size and the total GC percentage in tetrapods is distinct from that in (teleost) fish: both parameters are directly correlated in tetrapods (with the xenope being in conformity with the trend), whereas they are linked inversely in fish (Vinogradov 1998a). This contrast may be related to the discordance between fish and tetrapods observed in the intron–genome size plot (Fig. 3), since coding, intervening, and extragenic sequences usually differ in their GC percentages (Clay et al. 1996; Bernardi et al. 1997; Carels et al. 1998).

Mammals show a positive correlation between intron and genome size ($r = 0.66$, $p < 0.05$). However, its significance is determined mainly by the guinea pig point, and this conclusion should be considered as preliminary. Vertebrates as a whole also show a positive correlation ($r = 0.65$, $p < 0.01$). In fish, pufferfish introns do not differ significantly from introns of the other species, which have larger genomes (Table 1), so there is no correlation between intron and genome size. For rainbow trout and carp, the absence of intron enlargement compared to pufferfish can be partly explained by their ancestral polyploidy (Johnson et al. 1987; Larhammar and Risinger 1994; Allendorf and Danzmann 1997), but the zebrafish does not seem to have undergone polyploidization after the fish–tetrapod divergence (Daga et al. 1996; Postlethwait et al. 1998).

The chicken seems to have introns similar in size to

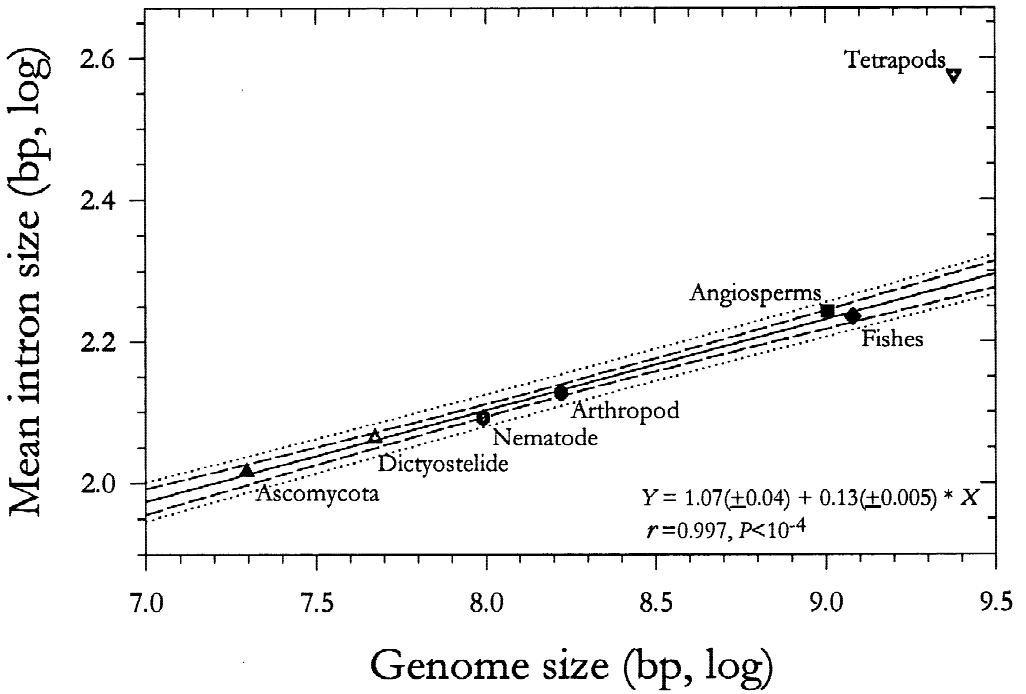

Fig. 3. Plot of mean intron size vs genome size for the phyla studied. The regression was established with the exclusion of tetrapods. When the tetrapods were included, the regression-line equation was $Y = -0.45\ (\pm 0.44) + 0.21\ (\pm 0.05) * X$ ($r = 0.87$, $p < 0.01$). Dashed and dotted lines are as in the legend to Fig. 1.

introns of mammals, notwithstanding the difference in genome size (Fig. 1A). Since it was reported previously on the grounds of pairwise analysis of homologous genes that introns are smaller in chicken than in human (Hughes and Hughes 1995), this difference was specifically analyzed here. Because introns of only a tiny fraction of homologous genes are available so far [<0.1% of ca. 75,000 genes in the vertebrate genome (e.g., Bird 1995)], the validity of statistical tests should be addressed briefly. The parametric *t* test uses the assumption of a Gaussian (normal) distribution, and Fig. 4 shows that this assumption is not correct for non-log-transformed intron sizes. The corresponding non-log-transformed contrasts (pairwise differences) are distributed more symmetrically (not shown) but these distributions are still unlikely to be normal ($p < 0.01$, by K-S test). This leaves us *t* tests with log-transformed values and nonparametric (insensitive to distribution parameters but less powerful) tests.

For the chicken–mammal comparison, the (statistically significant) intron differences were observed in a *t* test with log-transformed values and the nonparametric Mann–Whitney test only for human, guinea pig, and rabbit. (These analyses were done with the whole data set of introns for each species, i.e., not only with homologous introns). The mean intron size was found to be smaller in the rat than in the chicken ($p < 0.02$). The conclusion on the human–chicken difference was drawn in a previous report by Hughes and Hughes (1995) on the grounds of *t* test of pairwise contrasts of homologous introns with non-log-transformed values (which seems not quite justified). The nonparametric tests with pairwise contrasts, applied here to the chicken–human comparison (with 176 introns of 55 homologous genes, vs 111 introns of 31 genes in the previous work) and to the chicken–mouse comparison, failed to reveal significant differences (Fig. 4). In any case, notwithstanding the outcome of separate comparisons, the mean size of chicken introns seems to lie within the range of the mammalian scatter, although it may be somewhat below the mammalian average (Fig. 5).

The chicken–human and human–mouse intron size differences were also noted by Duret et al. (1995). It is noteworthy that in all studies where the chicken–mammal difference was addressed (Duret et al. 1995; Hughes and Hughes 1995; Oliver and Marin 1996), chicken introns were compared with those of the human, which are the longest among the mammals studied. Even if the chicken–human difference proved to be significant, it could not justify the conclusion about the chicken–mammal difference. It is interesting that introns of GC-poor and GC-medium genes are longer in mammals than in chickens, whereas introns of GC-rich genes are shorter (Oliver and Marin 1996, Table 4). Since a large portion of genes in warmblooded vertebrates is GC-rich (Mouchiroud et al. 1987; Bernardi 1993), chicken introns on average may not be smaller than those of mammals. On the other hand, genes from GC-poor isochores are

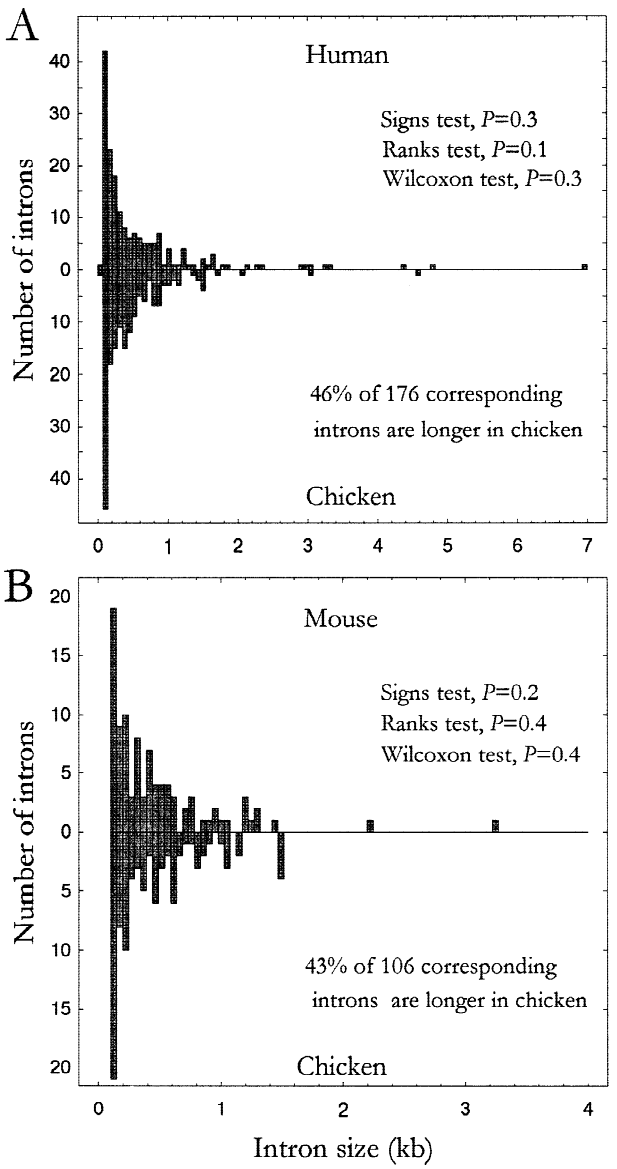

Fig. 4. Histograms of non-log-transformed lengths of homologous introns for the human–chicken **A** and the mouse–chicken **B** comparisons.

underrepresented in current databases (Duret et al. 1995), and the recent estimate may be biased in favor of longer chicken introns, residing in GC-rich isochores. Along the lines of a possible relation between intron size and chromatin condensation, it is tempting to suggest that late-replicating chromatin, which harbors GC-poor and GC-medium genes (Mouchiroud et al. 1987; Bernardi 1993; Saccone et al. 1993), is more condensed in mammals (compared to birds), whereas chromatin containing GC-rich genes is relatively more compact in birds.

Neither the IVS/CDS ratio nor the intron number/CDS ratio is lower in the chicken than in mammals taken as a group (Figs. 1B and 2A). The IVS/genome size ratio is higher in the chicken compared to mammals (Fig. 2B), which suggests that the chicken–mammal difference in genome size is determined mainly by extragenic DNA. The calculation of the extra-/intragenic ratio of noncoding DNA supports this suggestion (Table 2).

It was assumed that the smaller size of chicken introns compared to human introns was related to the smaller genome and higher metabolic rate of birds (Hughes and Hughes 1995). The significant chicken–mammal difference in intron size was not found here. Generally, the body mass-independent metabolic rates of birds also overlap with those of mammals, whereas their genomes are consistently smaller (Vinogradov 1995, 1997). It is interesting that the chicken–mammal relation in metabolic rate is similar to their intron–size but not their genome–size relation, whereas for mammals taken separately there is a negative correlation between genome size and body mass-independent metabolic rate (Vinogradov 1995). For the mammals studied here, the chicken has a rather low metabolic rate, and its point on the intron size-vs-metabolic rate plot overlaps with the prediction limits of regression for mammals (Fig. 5). (On the genome size-vs-metabolic rate plot, the chicken point would lie very far off the prediction limits due to the 2.5-fold difference in genome size between the chicken and mammals.) The inverse correlation between intron size and (body mass-independent) metabolic rate in mammals, as well as the conformity of the chicken with this trend, may have been related to the putative link between intron size and chromatin compaction, which suppresses gene activity.

If the assumption of intron involvement in chromatin condensation is correct, the general intron enlargement in evolution may be connected to the higher compactness of chromatin in higher organisms, with their more complex genetic machinery. The departure from the general trend in organisms with similar numbers of genes occurs because of variation in the extra-/intragenic ratio of noncoding DNA. Two hypotheses, transition (dynamic) and equilibrium (static), can be proposed for this variation. According to the transition model, this variation may arise from a rapid shift in genome size, since the bulk of extragenic DNA can be changed more rapidly than the finely interspersed intron sequences, requiring multiple mutation–fixation events. The equilibrium model assumes that this variation is a result of genome size adjustment by some selection vector with constraints imposed on the intron size because of its link to chromatin structure. There are also constraints imposed by the splicing machinery, which requires a certain minimum intron size (Goodall and Filipowicz 1990; Guo and Mount 1995). However, if the existence of introns themselves (and thus of the splicing machinery) is explained by the requirements of chromatin structure, the splicing constraints should be secondary. The adjustment of the total genome size may also be related to a similar end, namely, to the condensation state of chromatin. For in-

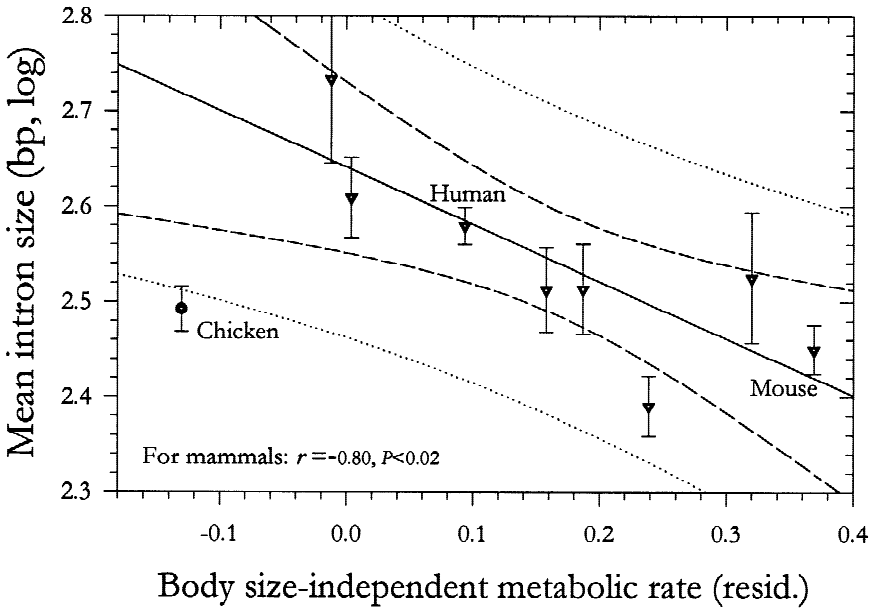

Fig. 5. Plot of mean intron size (with SE bars) vs. residuals of regression of metabolic rate on body mass. The residuals are from Vinogradov (1995), with the data for the cow added from Elgar and Harvey (1987) and those for the rabbit from Altman and Ditmer (1974). Dashed and dotted lines are as in the legend to Fig. 1.

stance, it was supposed that noncoding DNA is involved in sectorial gene repression (Zuckerkandl 1997) or buffers the effect of fluctuations in the intracellular solute composition on the chromatin condensation state and on binding of DNA-tropic proteins (Vinogradov 1998b).

The baker's yeast deviation from the general intron–genome size correlation can probably be regarded as the equilibrium case because of the very early divergence of this organism from the other fungi in this data set. The same can be supposed for birds, if they have introns similar in size to introns of their reptilian ancestor, notwithstanding the difference in genome size (as the chicken–mammal comparison suggests). Parenthetically, the average genome size is smaller in extant reptiles than in mammals (Vinogradov 1998a, Fig. 4A), therefore even if birds do have somewhat smaller introns than mammals, their intron size may correspond to the reptilian genome size. The main avian (and mammalian) orders are now thought to have diverged about 100 million years ago (Hedges et al. 1996; Cooper and Penny 1997; see also Gibbons 1997). Even if the compaction of the ancestral avian genome was rapid, since then there has been time for equilibration of the extra-/intragenic amounts of noncoding DNA. The case of mammals, whose intron size seems to correlate with genome size, supports this suggestion. In mammals and birds variation in genome size is small compared to that in fish and amphibians. Generally, in amniotes the size of the genome is more evolutionarily conservative (a greater part of the variance resides at higher taxonomic levels) than the GC percentage, whereas the reverse is true for anamniotes (Vinogradov 1998a). This suggests that in mammals and birds a change in the amount of extragenic DNA can be slow enough to be followed by a change in the intervening sequence size. The transition (dynamic) model may be relevant for fish, whose introns seem to be similar in size notwithstanding the differences in genome size (although polyploidy can explain part of the differences). This model is plausible for amphibians (especially anurans), where even closely related species always differ in genome size (Vinogradov 1998a).

Acknowledgment. The author is grateful to Laurent Duret for encouragement and helpful comments. This work was supported by a grant from the Russian Foundation for Basic Research (RFBR).

References

Allendorf FW, Danzmann RG (1997) Secondary tetrasomic segregation of MDH-B and preferential pairing of homeologues in rainbow trout. Genetics 145:1083–1092

Altman PL, Ditmer DS (1974) Biology data book, Vol III. Federation of American Societies for Experimental Biology, Bethesda, MD, pp 1613–1617

Bennett MD, Leicht IJ (1995) Nuclear DNA amount in angiosperms. Ann Bot 76:113–176

Bernardi G (1993) The vertebrate genome: isochores and evolution. Mol Biol Evol 10:186–204

Bernardi G, Hughes S, Mouchiroud D (1997) The major compositional transitions in the vertebrate genome. J Mol Evol 44 (Suppl 1):S44–S51

Bernardi G, Olofsson B, Filipski J, et al. (1985) The mosaic genome of warm-blooded vertebrates. Science 228:953–958

Bird AP (1995) Gene number, noise reduction and biological complexity. Trends Genet 11:94–100

Blank TA, Becker PB (1996) The effect of nucleosome phasing sequences and DNA topology on nucleosome spacing. J Mol Biol 260:1–8

Brenner S, Elgar G, Sanford R, Macrae A, Venkatesh B, Aparicio S (1993) Characterization of the pufferfish (Fugu) genome as a compact model vertebrate genome. Nature 366:265–268

Brody H, Carbon J (1989) Electrophoretic karyotype of Aspergillus nidulans. Proc Natl Acad Sci USA 86:6260–6263

Carels N, Hatey P, Jabbari K, Bernardi G (1998) Compositional prop-

erties of homologous coding sequences from plants. J Mol Evol 46:45–53

Clay O, Caccio S, Zoubak S, Mouchiroud D, Bernardi G (1996) Human coding and noncoding DNA: Compositional correlations. Mol Phylogenet Evol 5:2–12

Charnov EL (1991) Evolution of life history variation among female mammals. Proc Natl Acad Sci USA 88:1134–1137

Cooper A, Penny D (1997) Mass survival of birds across the Cretaceous–Tertiary boundary: Molecular evidence. Science 275:1109–1113

Daga RR, Thode G, Amores A (1996) Chromosome complement, C-banding, Ag-NOR and replication banding in the zebrafish Danio rerio. Chromosome Res 4:29–32

Duret L, Mouchiroud D, Gautier C (1995) Statistical analysis of vertebrate sequences reveals that long genes are scarce in GC-rich isochores. J Mol Evol 40:308–317

Elgar MA, Harvey PH (1987) Basal metabolic rates in mammals: Allometry, phylogeny and ecology. Funct Ecol 1:25–36

Firtel RA, Bonner J (1972) Characterization of the genome of the cellulose slime mold *Dictyostelium discoideum.* J Mol Biol 66:339–361

Gibbons A (1997) Did birds sail through the K-T extinction with flying colors? Science 275:1068

Goodall GJ, Filipowicz W (1990) The minimum functional length of pre-mRNA introns in monocots and dicots. Plant Mol Biol 14:727–733

Goffeau A, Barrell BG, Bussey H, et al. (1996) Life with 6000 genes. Science 274:546–567

Guo M, Mount SM (1995) Localization of sequences required for size-specific splicing of a small Drosophila intron in vitro. J Mol Biol 253:426–437

Hedges SB, Parker PH, Sibley CG, Kumar S (1996) Continental breakup and the ordinal diversification of birds and mammals. Nature 381:226–229

Hemmingsen AM (1960) Energy metabolism as related to body size and respiratory surfaces, and its evolution. Rep Steno Mem Hosp 9:1–110

Holmquist GP (1989) Evolution of chromosome bands: Molecular ecology of noncoding DNA. J Mol Evol 28:469–486

Hughes AL, Hughes MK (1995) Small genomes for better flyers. Nature 377:391

Johnson KR, Wright JE Jr, May B (1987) Linkage relationships reflecting ancestral tetraploidy in salmonid fish. Genetics 116:579–591

Kleiber M (1961) The fire of life: An introduction to animal energetics. Wiley & Sons, New York

Knoll AH (1992) The early evolution of eukaryotes: A geological perspective. Science 256:622–627

Larhammar D, Risinger C (1994) Molecular genetic aspects of tetraploidy in the common carp Cyprinus carpio. Mol Phylogenet Evol 3:59–68

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

Oliver JL, Marin A (1996) A relation between GC-content and coding-sequence length. J Mol Evol 43:216–223

Orbach MJ, Vollrath D, Davis RW, Yanofsky C (1988) An electrophoretic karyotype of *Neurospora crassa.* Mol Cell Biol 8:1469–1473

Postlethwait JH, Yan YL, Gates MA, et al. (1998) Vertebrate genome evolution and the zebrafish gene map. Nature Genet 18:345–349

Russell P, Nurse P (1986) *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*: A look at yeasts divided. Cell 45:781–782

Saccone S, De Sario A, Wiegant J, Raap AK, Della Valle G, Bernardi G (1993) Correlations between isochores and chromosomal bands in the human genome. Proc Natl Acad Sci USA 90:11929–11933

Schmidt-Nielsen K (1984) Scaling: Why is animal size so important? Cambridge University Press, Cambridge

Shuter BJ, Thomas JE, Taylor WD, Zimmerman AM (1983) Phenotypic correlates of genomic DNA content in unicellular eukaryotes and other cells. Am Nat 122:26–44

Sogin ML, Patterson DJ (1995 onwards) Tree of life. Mitochondrial eukaryotes. http://phylogeny.arizona.edu/tree/eukaryotes/mitochondrial_eukaryotes.html

Tiersch TR, Chandler RW, Wachtel SS, Elias S (1989) Reference standards for flow cytometry and application in comparative studies of nuclear DNA content. Cytometry 10:706–710

Villeponteau B, Brawley J, Martinson HG (1992) Nucleosome spacing is compressed in active chromatin domains of chick erythroid cells. Biochemistry 31:1554–1563

Vinogradov AE (1995) Nucleotypic effect in homeotherms: Body mass-corrected basal metabolic rate of mammals is related to genome size. Evolution 49:1249–1259

Vinogradov AE (1997) Nucleotypic effect in homeotherms: Body mass-independent resting metabolic rate of passerine birds is related to genome size. Evolution 51:220–225

Vinogradov AE (1998a) Genome size and GC-percent in vertebrates as determined by flow cytometry: The triangular relationship. Cytometry 31:100–109

Vinogradov AE (1998b) Buffering: A possible passive-homeostasis role for redundant DNA. J Theor Biol 193:197–199

West GB, Brown JH, Enquist BJ (1997) A general model for the origin of allometric scaling laws in biology. Science 276:122–126

Williams N (1997) Fractal geometry gets the measure of life's scales. Science 276:34

Wootton JT (1987) The effects of body mass, phylogeny, habitat, and trophic level on mammalian age at first reproduction. Evolution 41:732–749

Zhang MQ, Marr TG (1994) Fission yeast gene structure and recognition. Nucleic Acids Res 22:1750–1759

Zuckerkandl E (1981) A general function of noncoding polynucleotide sequences. Mass binding of transconformational proteins. Mol Biol Rep 7:149–158

Zuckerkandl E (1997) Junk DNA and sectorial gene repression. Gene 205:323–343

Zuckerkandl E, Hennig W (1995) Tracking heterochromatin. Chromosoma 104:75–83