Review Article

Cytogenetic and Genome Research

Cytogenet Genome Res 2015;147:217–239
DOI: 10.1159/000444429

Accepted: December 3, 2015
by M. Schmid
Published online: March 12, 2016

Transposons, Genome Size, and Evolutionary Insights in Animals

Adriana Canapa Marco Barucca Maria A. Biscotti Mariko Forconi
Ettore Olmo

Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, Ancona, Italy

Key Words
Comparative genomics · C-value · Deuterostomes · Mobile elements · Protostomes · Repetitive DNA

Abstract
The relationship between genome size and the percentage of transposons in 161 animal species evidenced that variations in genome size are linked to the amplification or the contraction of transposable elements. The activity of transposable elements could represent a response to environmental stressors. Indeed, although with different trends in protostomes and deuterostomes, comprehensive changes in genome size were recorded in concomitance with particular periods of evolutionary history or adaptations to specific environments. During evolution, genome size and the presence of transposable elements have influenced structural and functional parameters of genomes and cells. Changes of these parameters have had an impact on morphological and functional characteristics of the organism on which natural selection directly acts. Therefore, the current situation represents a balance between insertion and amplification of transposons and the mechanisms responsible for their deletion or for decreasing their activity. Among the latter, methylation and the silencing action of small RNAs likely represent the most frequent mechanisms.

Despite many studies, the C-value enigma sensu Gregory [2005], that is the expansion of the genome size mainly due to the accumulation of non-coding DNA not related to the genetic complexity and to the evolutionary position, remains to date one of the most fascinating problems which is still not fully understood in the organization and evolution of the genome.

One of the peculiarities of these changes in genome size is that they seem to influence, regardless of their sequences, various parameters of the cell which have an impact on morphological and functional characteristics of the organism. These characteristics are subject to natural selection, therefore influencing major evolutionary processes.

To explain the variability in genome size, several theories have been discussed. According to some of these theories, there would be no causal relationship between the amount of DNA and cellular parameters, and the accumulation of DNA would depend simply on different levels of tolerance of each species towards insertion, amplification, and storage of repetitive sequences [Pagel and Johnstone, 1992; Charlesworth et al., 1994]. Other hypotheses imply that variations in genome size would always have a causal role in changes in cellular parameters and in those traits of the organism that are subject to natural selection and therefore to stabilizing selection for the

1424–8581/16/1474–0217$39.50/0

E-Mail karger@karger.com
www.karger.com/cgr

Ettore Olmo
Dipartimento di Scienze della Vita e dell'Ambiente
Università Politecnica delle Marche
Via Brecce Bianche, IT–60131 Ancona (Italy)
E-Mail e.olmo@univpm.it

Table 1. Minimum and maximum genome sizes in all animal taxa

Taxon	Genome size (pg/N)	
	minimum	maximum
Basal taxa	0.04	1.84
Protostomes	0.02	64.62
Flatworms	0.03	20.57
Nematodes	0.02	2.50
Molluscs	0.43	7.85
Annelids	0.06	7.64
Arthropods		
Cheliceratans	0.08	7.50
Crustaceans	0.16	64.62
Polar species	5.10	64.62
Deep sea species	10.58	38.47
Insects	0.09	16.93
Other protostomes	0.05	19.87
Deuterostomes	0.06	123.9
Chordates	0.06	123.9
Urochordates	0.06	0.20
Cephalochordates	0.59	–
Vertebrates	0.35	123.9
Jawless	1.29	4.59
Cartilagineous fishes	1.51	17.05
Ray-finned fishes	0.35	7.17
Teleosts	0.35	4.4
Lobe-finned fishes	3.00	123.9
Amphibians	0.95	120
Apodans	3.70	13.95
Frogs	0.95	12.4
Salamanders	10.12	120
Reptiles	1.05	5.44
Birds	0.92	2.16
Mammals	1.63	8.40
Other deuterostomes	0.38	4.40

Data are according to the Animal Genome Size Database (www.genomesize.com).

optimum genome size [Bennett, 1971, 1973; Cavalier-Smith, 1982, 1985]. Finally, other hypotheses speculate that the genome size of a species evolves until the loss of non-coding self-replicating DNA through small deletions has an equal rate of DNA gain through long insertions [Petrov, 2002a].

The most recent studies on genome sequencing have allowed a significant progress in the knowledge of the organization and composition of the genome. Moreover, studies on mobile DNA and their propensity to independently amplify and to specifically insert in the host genome have changed the perspective of the C-value enigma.

Many works have shown a general positive correlation among different parameters of the genome, especially between repeated sequences including transposons and the genome size within numerous eukaryotes [Elliot and Gregory, 2015a] and mainly in vertebrates [Chalopin et al., 2015]. However, this correlation appears to especially affect the quantitative aspects of transposons, since a limited correlation between the diversity of transposons and genome size was noted only for genomes <500 Mb [Elliot and Gregory, 2015b].

In this regard, changes in the percentage of transposons and genome size in 161 species of animals, whose entire genome has been sequenced until now, including 88 deuterostomes, 68 protostomes, and 2 cnidarians, 2 ctenophores, and 1 placozoan species, were examined in the light of major evolutionary transitions in order to verify the influence that these sequences had on evolutionary processes such as speciation and adaptation to the environment. Data on genome sizes were obtained from the Animal Genome Size Database (www.genomesize.com), and minimum and maximum values for each taxon are shown in table 1.

Genome Size Landscape

Most of the studies on genome size involved vertebrates, while those focusing on invertebrates are very limited (no more than 1% of all living invertebrate species have been studied). The situation regarding genome sequencing and transposon analysis is more balanced. In fact, 54.7% of these studies focused on chordates, especially vertebrates, and 42.2% on protostomes and 3.1% on primitive metazoans (table 2).

Primitive Metazoans

The information on genome size and composition of primitive metazoans is very limited. The genome size is generally low, ranging from 0.04 to 1.84 pg/N (table 1). The percentage of transposons was studied only in 5 species and seems not to be correlated with the genome size. Indeed, in *Trichoplax* the percentage of transposons is only 0.13%, the lowest among animals, while in ctenophores it is <10%. In cnidarians, despite the small genome size, the percentage of transposons shows high values comparable with those identified in many protostomes and deuterostomes. Although these data are very scarce, they might suggest that at the origin of metazoans genome sizes and perhaps even the percentage of transposons were generally low [Gregory, 2005] and that trans-

Table 2. Genome size and percentage of transposons in 161 animal species

Species	Genome size, pg/N	Transposons, %	More frequent transposable element	Reference(s)
Placozoans				
Trichoplax adhaerens	0.04	0.13	DNA	Wang et al. [2010]
Cnidarians				
Hydra magnipapillata	1.4	57.64	non-LTR	Chapman et al. [2010]
Nematostella vectensis	0.23	26.2	DNA	Putnam et al. [2007]
Ctenophores				
Pleurobrachia bachei	0.18[a]	8.5	DNA	Moroz et al. [2014]
Mnemiopsis leidyi	0.31	5.57	DNA	Ryan et al. [2013]
Chordates				
Urochordates				
Ciona intestinalis	0.20	18.8	non-LTR	Dehal et al. [2002]
Ciona savignyi	0.19[a]	16.7	DNA	Small et al. [2007]
Oikopleura dioica	0.07	19.5	LTR	Chalopin et al. [2015]
Cephalochordates				
Branchiostoma floridae	0.54[a]	22.7	DNA	Chalopin et al. [2015]
Vertebrates				
Jawless				
Petromyzon marinus	2.2	34.7	non-LTR	Smith et al. [2013]
Cartilagineous fishes				
Callorhinchus milii	1.94	42.72	non-LTR	Chalopin et al. [2015]
Ray-finned fishes				
Lepisosteus osseus	0.98[a]	19.77	non-LTR	Chalopin et al. [2015]
Anguilla europaea	1.125[a]	13.64	DNA	Chalopin et al. [2015]
Austrolebias charrua	3.07	45		Garcia et al. [2015]
Cynoglossus semilaevis	0.56[a]	5.85	DNA	Chen et al. [2014]
Cynopoecilus melanotaenia	1.36	25		Garcia et al. [2015]
Danio rerio	1.75	54.94	DNA	Howe et al. [2013]; Chalopin et al. [2015]
Esox lucius	1.15	18.1	DNA	Rondeau et al. [2014]
Gadus morhua	0.92	25.4	DNA	Star et al. [2011]; Chalopin et al. [2015]
Gasterosteus aculeatus	0.58	13.91	non-LTR	Chalopin et al. [2015]
Oncorhynchus mykiss	2.78	29	non-LTR	Berthelot et al. [2014]
Oryzias latipes	0.89[a]	28.09	DNA	Kasahara et al. [2007]; Chalopin et al. [2015]
Astatotilapia burtoni	0.95[a]	16.56	DNA	Brawand et al. [2014]
Metriaclima zebra	0.97[a]	17.38	DNA	Brawand et al. [2014]
Neolamprologus brichardi	1.01[a]	15.96	DNA	Brawand et al. [2014]
Oreochromis niloticus	0.99	18.75	DNA	Brawand et al. [2014]
		24.06		Chalopin et al. [2015]
Pundamilia nyererei	1.02[a]	16.97	DNA	Brawand et al. [2014]
Takifugu rubripes	0.4[a]	6.72	non-LTR	Chalopin et al. [2015]
Tetraodon nigroviridis	0.35	5.85	non-LTR	Chalopin et al. [2015]
Xiphophorus maculatus	0.99[a]	21.22	DNA	Schartl et al. [2013]; Chalopin et al. [2015]
Lobe-finned fishes				
Latimeria chalumnae	3.61	22	non-LTR	Amemiya et al. [2013]; Chalopin et al. [2015]
Neoceratodus forsteri	52.75	39.4	non-LTR	Metcalfe et al. [2012]
Amphibians				
Xenopus (Silurana) tropicalis	1.5	34.5	DNA	Hellstein et al. [2010]; Sun et al. [2015]
		42.6		
Nanorana parkeri	2.64[a]	41.4	LTR	Sun et al. [2015]
Aneides flavipunctatus	42.9	47.52	LTR	Sun et al. [2012a]
Batrachoseps nigriventris	26.06	39.39	LTR	Sun et al. [2012a]
Bolitoglossa occidentalis	43.5	33.19	LTR	Sun et al. [2012a]

Table 2 (continued)

Species	Genome size, pg/N	Transposons, %	More frequent transposable element	Reference(s)
Bolitoglossa rostrata	47.8	30.18	LTR	Sun et al. [2012a]
Cryptobranchus alleganiesis	55	49.56	LTR	Sun and Mueller [2014]
Desmognathus ochrophaeus	15.70	39.69	LTR	Sun et al. [2012a]
Eurycea tynerensis	25.0[b]	25.18	LTR	Sun et al. [2012a]
Reptiles				
Chelonia mydas	2.64	37.35	non-LTR	Wang et al. [2013]
Chrysemys picta	2.97	9.11	non-LTR	Shaffer et al. [2013]
Pelodiscus sinensis	2.25[a]	42.47	non-LTR	Wang et al. [2013]
Alligator mississippiensis	2.66	23.44	non-LTR	St John et al. [2012]
		36.96		Green et al. [2014]
Crocodylus porosus	2.78[a]	27.22	non-LTR	St John et al. [2012]
		36.83		Green et al. [2014]
Gavialis gangeticus	2.9[a]	35.49	non-LTR	Green et al. [2014]
Anolis carolinensis	2.2	34.4	non-LTR	Alfoldi et al. [2011]
Ophisaurus gracilis	1.82[a]	49.63	non-LTR	Song et al. [2015]
Agkistrodon contortrix	1.37	42.4	non-LTR	Castoe et al. [2013]
Boa constrictor	2.3	25.49	non-LTR	Castoe et al. [2013]
Crotalus atrox	1.52	35.59	non-LTR	Castoe et al. [2013]
Micrurus fulvius	1.74[a]	36.7	non-LTR	Castoe et al. [2013]
Ophiophagus hannah	1.36[a]	31.28	non-LTR	Castoe et al. [2013]
Pantherophis guttatus	1.8[a]	39.14	non-LTR	Ullate-Agote et al. [2014]
Python molurus	1.67[a]	24.59	non-LTR	Castoe et al. [2013]
Thamnophis sirtalis	1.91	39.26	non-LTR	Castoe et al. [2013]
Typhlops reticulatus	1.9[a]	24.89	non-LTR	Castoe et al. [2013]
Birds				
Anas platyrhynchos	1.49	5.85	non-LTR	Huang et al. [2013]
Balearica regulorum	1.44	6.08	non-LTR	Zhang et al. [2014]
Calypte anna	1.14	8.05	non-LTR	Zhang et al. [2014]
Cariama cristata	1.5	5.49	non-LTR	Zhang et al. [2014]
Corvus brachyrhynchos	1.25	7.37	non-LTR	Zhang et al. [2014]
Columba livia	1.59	7.25	non-LTR	Zhang et al. [2014]
Ficedula albicollis	1.17[a]	10.68	non-LTR	Ellegren et al. [2012]
Falco peregrinus	1.45	5.5	non-LTR	Zhan et al. [2013]
Gallus gallus	1.25	9.4	non-LTR	International Chicken Sequencing Consortium [2004]; Chalopin et al. [2015]
Haliaeetus leucocephalus	1.43	6.89	non-LTR	Zhang et al. [2014]
Meleagris gallopavo	1.31	5.74	non-LTR	Dalloul et al. [2010]
Melopsittacus undulatus	1.33	7.9	non-LTR	Zhang et al. [2014]
Phoenicopterus ruber	1.52	5.6	non-LTR	Zhang et al. [2014]
Struthio camelus	2.16	4.11	non-LTR	Zhang et al. [2014]
Taeniopygia guttata	1.25	7.7	non-LTR	Warren WC et al. [2010]; Chalopin et al. [2015]
Tetrao tetrix	1.02[a]	6.34	non-LTR	Wang et al. [2014]
Tyto alba	1.73	5.49	non-LTR	Zhang et al. [2014]
Mammals				
Prototherians				
Ornithorhynchus anatinus	3.06	44.6	non-LTR	Warren et al. [2008]; Chalopin et al. [2015]
Metatherians				
Macropus eugenii	2.75[a]	52.8	non-LTR	Renfree et al. [2011]
Monodelphis virginiana	4.15	53.84	non-LTR	Mikkelsen et al. [2007]; Chalopin et al. [2015]
Sarcophilus harrisii	3.63	52	non-LTR	Gallus et al. [2015]

Table 2 (continued)

Species	Genome size, pg/N	Transposons, %	More frequent transposable element	Reference(s)
Eutherians				
Ailuropoda melanoleuca	2.46[a]	34.7	non-LTR	Li et al. [2010]
Bos taurus	3.65	46.5	non-LTR	Bovine Genome Sequencing and Analysis Consortium [2009]
Canis familiaris	2.88	35.15	non-LTR	Lindblad-Toh et al. [2005]
Eidolon helvum	2.03	29.7	non-LTR	Parker et al. [2013]
Pteronotus parnellii	2.67	29.22	non-LTR	Parker et al. [2013]
Rhinolophus ferrumequinum	2.68	29.16	non-LTR	Parker et al. [2013]
Equus caballus	3.15	46	non-LTR	Wade et al. [2009]
Felis catus	2.91	32.1	non-LTR	Pontius et al. [2007]
Panthera tigris	2.90	37.5	non-LTR	Cho et al. [2013]
Homo sapiens	3.5	42.85	non-LTR	Chalopin et al. [2015]
Macaca mulatta	3.59	50	non-LTR	Rhesus Macaque Genome Sequencing and Analysis Consortium [2007]
Mus musculus	3.25	38.55	non-LTR	Mouse Genome Sequencing Consortium [2002]; Chalopin et al. [2015]
Rattus norvegicus	3.05	40.3	non-LTR	Gibbs et al. [2004]
Sus scrofa	2.81	40	non-LTR	Groenen et al. [2012]
Arthropods				
Insects				
Coleopterans				
Dendroctonus ponderosae	0.21	17.21	n.a.	Keeling et al. [2013]
Tribolium castaneum	0.21	27	non-LTR	*Tribolium* Genome Sequencing Consortium [2008]
Dipterans				
Aedes aegypti	0.83	47	non-LTR	Nene et al. [2007]
Aedes albopictus	1.40[c]	71	non-LTR	Chen et al. [2015]
Anopheles albimanus	0.17[a]	1.98	non-LTR	Nafsey et al. [2015]
Anopheles arabiensis	0.25[a]	9.38	non-LTR	Nafsey et al. [2015]
Anopheles christyi	0.18[a]	2.81	non-LTR	Nafsey et al. [2015]
Anopheles dirus	0.29[a]	5.09	non-LTR	Nafsey et al. [2015]
Anopheles epiroticus	0.22[a]	6.27	non-LTR	Nafsey et al. [2015]
Anopheles funestus	0.26[a]	4.03	non-LTR	Nafsey et al. [2015]
Anopheles gambiae	0.27	17.78	non-LTR	Nafsey et al. [2015]
Anopheles melas	0.23[a]	7.29	non-LTR	Nafsey et al. [2015]
Anopheles merus	0.26[a]	11.43	non-LTR	Nafsey et al. [2015]
Anopheles quadrannulatus	0.29[a]	7.69	non-LTR	Nafsey et al. [2015]
Anopheles stephensi	0.24	5.04	non-LTR	Nafsey et al. [2015]
Culex quinquefasciatus	0.54	29	DNA	Arensburger et al. [2010]
Drosophila ananassae	0.19	25	LTR	*Drosophila* 12 Genome Consortium [2007]
Drosophila erecta	0.16	6.9	LTR	*Drosophila* 12 Genome Consortium [2007]
Drosophila grimshawi	0.24	2.84	LTR	*Drosophila* 12 Genome Consortium [2007]
Drosophila melanogaster	0.16	5.35	LTR	*Drosophila* 12 Genome Consortium [2007]
Drosophila mojavensis	0.17	8.92	LTR	*Drosophila* 12 Genome Consortium [2007]
Drosophila persimilis	0.18	8.47	LTR	*Drosophila* 12 Genome Consortium [2007]
Drosophila pseudoobscura	0.16	2.76	LTR	*Drosophila* 12 Genome Consortium [2007]
Drosophila sechellia	0.17	3.67	LTR	*Drosophila* 12 Genome Consortium [2007]
Drosophila simulans	0.15	2.7	LTR	*Drosophila* 12 Genome Consortium [2007]
Drosophila virilis	0.34	13.96	LTR	*Drosophila* 12 Genome Consortium [2007]
Drosophila willistoni	0.21	15.57	LTR	*Drosophila* 12 Genome Consortium [2007]
Drosophila yakuba	0.17	12.04	LTR	*Drosophila* 12 Genome Consortium [2007]
Musca domestica	0.92	52	DNA	Scott et al. [2014]

Table 2 (continued)

Species	Genome size, pg/N	Transposons, %	More frequent transposable element	Reference(s)
Hemipterans				
Acyrthosiphon pisum	0.31	38	n.a.	International Aphid Genome Sequencing Consortium [2010]
Diaci psyllid	0.49[a]	34	DNA	Scott et al. [2014]
Anoplurans				
Pediculus humanus	0.11	7	non-LTR	Kirkness et al. [2010]
Hymenopterans				
Apis mellifera	0.24	1	DNA	Honeybee Genome Sequencing Consortium [2006]
Atta cephalotes	0.31	21.9	DNA	Suen et al. [2011]
Camponotus floridanus	0.31	15	DNA	Bonasio et al. [2010]
Harpegnathos saltator	0.3[a]	27	DNA	Bonasio et al. [2010]
Linepithema humile	0.26	23.5	DNA	Smith CD et al. [2011]
Microplitis demolitor	0.23[a]	42	n.a.	Scott et al. [2014]
Nasonia vitripennis	0.34	43.4	non-LTR	Warren JH et al. [2010]
Pogonomyrmex barbatus	0.27[a]	8	DNA	Smith CR et al. [2011]
Lepidopterans				
Bombyx mori	0.52	45	non-LTR	Mita et al. [2004]
Danaus plexippus	0.29	13.1	n.a.	Zhan et al. [2011]
Heliconius melpomene	0.3	25	non-LTR	Lavoie et al. [2013]
Plutella xylostella	0.4[a]	33.97	non-LTR	You et al. [2013]
Orthopterans				
Locusta migratoria	6.35	58.86	non-LTR	Wang et al. [2014]
Schistocerca gregaria	8.71	50	n.a.	Camacho et al. [2015]
Cheliceratans				
Strigamia maritima	0.3	42.44	LTR	Chipman et al. [2014]
Mesobuthus martensii	1.35	13	LTR	Cao et al. [2013]
Tetranychus urticae	0.08	10.01	LTR	Grbić et al. [2011]
Crustaceans				
Daphnia pulex	0.23	9.4	LTR	Colbourne et al. [2011]
Molluscs				
Crassostrea gigas	0.91	36	DNA	Zhang et al. [2012]
Lottia gigantea	0.43	21	non-LTR	Simakov et al. [2013]
Flatworms				
Clonorchis sinensis	0.66[a]	25.96	non-LTR	Wang et al. [2011]
Echinococcus granulosus	0.45	30.25	n.a.	Zheng et al. [2013]
Fasciola hepatica	1.32[a]	32	n.a.	Cwiklinski et al. [2015]
Schistosoma haematobium	0.39[a]	47	non-LTR	Young et al. [2012]
Schistosoma japonicum	0.41[a]	40.1	non-LTR	*Schistosoma japonicum* Genome Sequencing and Functional Analysis Consortium [2009]
Schistosoma mansoni	0.26	45	non-LTR	Berriman et al. [2009]
Nematodes				
Ascaris suum	0.25	4.4	non-LTR	Jex et al. [2011]
Brugia malayi	0.11	15	n.a.	Ghedin et al. [2007]
Caenorhabditis briggsae	0.11	22.4	n.a.	Stein et al. [2003]
Caenorhabditis elegans	0.10	18.3	n.a.	*Caenorhabditis elegans* Sequence Consortium [1998]
Meloidogyne hapla	0.05	1	DNA	Opperman et al. [2008]
Pristionchus pacificus	0.17	17	non-LTR	Dieterich et al. [2008]
Trichinella spiralis	0.26	18	n.a.	Mitreva et al. [2011]
Annelids				
Capitella teleta	0.24	31	non-LTR	Simakov et al. [2013]
Helobdella robusta	0.31	33	non-LTR	Simakov et al. [2013]

Table 2 (continued)

Species	Genome size, pg/N	Transposons, %	More frequent transposable element	Reference(s)
Rotifers				
Adineta vaga	0.25	3.03	n.a.	Flot et al. [2013]

The genome size values of the 161 animals (1 placozoan, 2 cnidarians, 2 ctenophores, 88 deuterostomes, and 68 protostomes) are according to the Animal Genome Size Database (www.genomesize.com). When more than one record was reported, the value obtained through flow cytometry or the more recent one was chosen.

[a] Data originally reported in bp were converted into pg/N, taking into account that 978 Mb correspond to 1 pg/N.

[b] Represents an average of 9 other *Eurycea* species.

[c] Data reported by different authors.

posons would have amplified independently in various phyla in the early phases of evolution.

Protostomes

Among protostomes, arthropods and especially insects have been extensively investigated. Within insects genome sizes are limited, ranging from 0.09 to ~4 pg/N, with most of the species abutting values <1.5 pg/N. Exceptions are represented by orthopterans with values reaching up to 16.93 pg/N. An interesting aspect concerns the differences between the holometabolous (with average values <2 pg/N) and hemimetabolous (possessing values significantly higher) insects [Gregory, 2005; Hanrahan and Johnston, 2011]. A comparison of the percentage of transposons indicates that these differences depend predominantly on their expansion (fig. 1), as shown by hemipterans and orthopterans having on average a high percentage of transposons compared to other orders (table 2).

In arachnids, genome sizes vary 8×, from 0.74 to 5.7 pg/N with an average of 2 pg/N. These arthropods do not metamorphose but show various molts during growth, and, similarly to hemimetabolous insects, they do not have a genome size limit of 2 pg/N [Gregory and Shorthouse, 2003].

Higher and more variable values can be observed among crustaceans, ranging from 0.16 to over 50 pg/N, with an average of about 3 pg/N and with most of the species not exceeding 6 pg/N. Genome size values >20 pg/N are extremely rare among invertebrates and crustaceans, and they are limited to species living in extreme environments, such as polar regions, especially the Antarctic, and the deep seas (mainly in hydrothermal vents) [Gregory, 2005; Rees et al., 2007, 2008; Bonnivard et al., 2009; Dufresne and Jeffery, 2011].

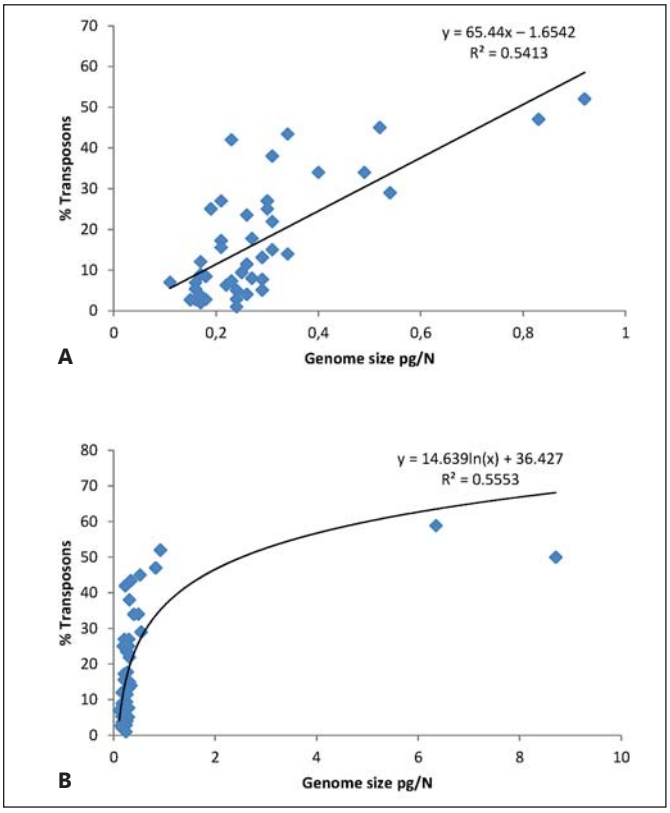

Fig. 1. A Relationship between genome size and percentage of transposons in insects without data from orthopterans. **B** Relationship between genome size and percentage of transposons in insects including orthopterans.

The other arthropod groups harbor a fairly limited genome size that only rarely exceeds 5 pg/N [Gregory and Shorthouse, 2003; Gregory, 2005; Hanrahan and Johnston, 2011].

Data collected so far in arthropods suggest that a small genome size was a widespread common ancestral condition within this phylum and that during evolution increases in genome size independently occurred in different lineages [Hanrahan and Johnston, 2011].

In mollusks the occurrence of a whole genome duplication has been speculated on [Yoshida et al., 2011]. The genome size has been investigated mainly in bivalves, and in this group, values range from a minimum of 0.43 pg/N to a maximum of 7.85 pg/N with an average of 1.8 pg/N. One of the highest values has been found in the Antarctic bivalve *Neobuccinum eatoni* [Libertini et al., 1993]. Among gastropods, it was noted that terrestrial species display a genome ~2× larger than their freshwater relatives [Hinegardner, 1974; Vinogradov, 2000; Gregory, 2005]. Hinegardner [1976] hypothesized an increase in genome size during evolution from more generalized to more specialized species. However, the latest results contrast with this hypothesis and suggest instead that more generalized mollusk species possess larger genomes [Rodriguez-Juiz et al., 1996].

In almost all other protostomes, except in very rare cases like the flatworm *Otomesostoma*, whose genome reaches 20 pg/N, a quite similar situation has been described with a range that goes from a minimum of about 0.1 pg/N to a maximum of 5 pg/N [Rodriguez-Juiz et al., 1996; Gregory, 2005]. Despite many species of annelids, nematodes, and flatworms being parasites and generally having small genome sizes [Sundberg and Pulkkinen, 2015], no difference in the percentage of transposons between free-living and parasitic species was noticed [Zheng et al., 2013].

Examining the presence and the proportion of repetitive DNA in various protostome phyla it is evident that in all cases the expansion of the genome depends on the expansion of various classes of transposons (table 2; online suppl. fig. 1; for all online suppl. material, see www.karger.com/doi/10.1159/000444429). In insect species harboring genomes <2 pg/N there is a linear and direct relationship between the percentage of transposons and the genome size (fig. 1A), which indicates that the expansion of the genome is caused mostly, if not completely, by the amplification of transposons. However, taking into account the orthopterans displaying genomes >6 pg/N, the above relationship shows a logarithmic development (fig. 1B). This trend is also confirmed when looking at all arthropods and protostomes (online suppl. figs. 1, 2). These data suggest that exceeding certain values, the increase in genome size is not fully justified by the amount of transposons, but it could depend on other sequences or on the preservation of a certain percentage of transposons without any selective pressure causing the accumulation of mutations over time and masking its repetitiveness. Concerning the classes of transposons in various protostomes, and especially within the group of insects, certain homogeneity was noticed at the level of genus or family, while there is a considerable diversity, both among the various phyla and within each phylum or subphylum (table 2).

Deuterostomes

While studying genome sizes in deuterostomes, the most evident aspect is the clear difference found between vertebrates and other deuterostomes including primitive chordates. Indeed, in the latter, genome sizes are small, while in vertebrates they are on average higher, more variable, and can reach very high values of >100 pg/N [Gregory, 2005]. This could be explained by the hypothesis that origin and some important steps in the evolution of vertebrates would have been characterized by a duplication of the entire genome. Such duplications would have occurred at the origin of vertebrates, at the separation of gnathostomes and agnates, and at the origin of teleosts after the separation of actinopterygians and sarcopterygians. These duplications coincide with a burst of new character appearances and with the acquisition and the increase in phenotypic complexity [Meyer and Schartl, 1999; Panopoulou et al., 2003; Donoghue and Purnell, 2005; Panopoulou and Poustka, 2005; Volff, 2005]. However, the significant differences identifiable within certain classes cannot be explained only by genome duplication events, suggesting on one hand that they were not the only causes of changes in genome size and that genomic dimensions have been affected by many factors on the other [Chalopin et al., 2015]. A very significant increase in genome size not due to duplication of the entire genome is known in lungfishes and salamanders, whose origin and early stages of evolution have occurred in conjunction with the transition from aquatic vertebrates to terrestrial ones.

Although some hypotheses suggest that in vertebrates the most primitive species possess small genomes [Gregory, 2005], during evolution the genome size does not appear to have followed just one trend. For each class of the subphylum different trends are observed, leading to a genome expansion in some classes and to a contraction in others.

Beside lungfishes and amphibians, the cartilaginous fishes display the largest genome size ranging from 1.5 pg/N in the chimaeras to 17 pg/N in sharks with an aver-

age of 5.7 pg/N [Stingo et al., 1980; Gregory, 2005]. Although the lowest values are found in the chimaeras, which are considered to be the most primitive group among chondrichthyes, it is not possible to identify a clear trend in genome size evolution within this class. Larger cells and nuclei were observed in cartilaginous fish living in cold temperatures. Although a direct relationship between genome size and cell and nucleus size is well known, it has been noticed that cells of selachian species living in cold water are larger than those of species living in warm water even if they have the same genome size [Hardie and Hebert, 2003]. Higher average genome sizes were also observed in some deep-sea selachians, although this correlation is not significant [Sion et al., 2004].

The percentage of transposons has been studied so far only in the chimaera *Callorhinchus milii* [Chalopin et al., 2015]. However, C_0t analyses, although based on a limited number of species, would suggest that the increase in genome size in this fish is correlated to an increase in moderately repetitive DNA [Morescalchi and Olmo, 1982; Olmo et al., 1982; Stingo et al., 1989] (online suppl. fig. 3). Since it is known that transposons belong to the above-mentioned fraction [Krebs et al., 2013], it is presumed that even in this class the expansion of the genome could depend on the amplification of transposons. Reassociation kinetics have also shown that in some species the increase of DNA would also be accompanied by a doubling of the so-called single-copy fraction that is largely made up of structural gene sequences [Krebs et al., 2013]. This could be a remnant of a primitive whole genome duplication [Olmo et al., 1982].

From an evolutionary point of view the ray-finned fishes and especially teleosts are the most successful group among the vertebrates. They comprise 99% of the 30,000 species of extant fishes and, along with mammals, exhibit the highest rate of diversification in the course of evolution [Benton, 2000; Olmo, 2006].

The origin of teleosts would have been characterized by a specific whole genome duplication that would coincide with a burst of character acquisitions and with an increased phenotypic complexity [Vanderpoele et al., 2004; Donoghue and Purnell, 2005; Volff, 2005]. Loss or sub-function partitioning of duplicated genes would have been involved in the generation of phenotypic variability of these fishes [Volff, 2005], which actually experienced more frequent gene linkage disruptions than other vertebrates [Ravi and Venkatesh, 2008]. Despite their great evolutionary success, the teleost fishes show a remarkable uniformity in genome size, with values ranging from a minimum of 0.4 to a maximum of 4.4 pg/N with an average of 1.2 pg/N, one of the lowest among all the deuterostomes. A correlation between genome size and extreme environments has been noted also in bony fish species: mesopelagic and bathypelagic species display larger genomes than surface water fishes [Ebeling et al., 1971], and the cold-water species are larger in genome size than warm-water species [Hardie and Hebert, 2003].

Also in fish a clear and direct correlation between the percentage of transposons and the increase in genome size was identified (table 2). Moreover, it has also been hypothesized that speciation events could be associated with retrotranspositional bursts [Volff, 2005].

The typical tendency of teleosts to preserve small genomes is evident in pufferfishes, which possess the smallest and most compact genomes of all vertebrates. They have a lower percentage of repeated sequences (especially DNA transposons) and shorter intronic sequences. This situation would depend both on a high rate of intron and transposon loss and on a higher level of indels (insertions/deletions) [Imai et al., 2007; Loh et al., 2008; Noleto et al., 2009; Guo et al., 2012].

One of the most controversial steps in the evolution of vertebrate genome sizes is the transition from aquatic to terrestrial environments involving lobe-finned fishes and amphibians.

The extant lobe-finned fishes include the coelacanths, very popular in the Devonian, and are represented today only by the *Latimeria* genus with 2 species dwelling in the deep waters of Africa and Indonesia and the lungfishes, which are shown by molecular studies to be the direct ancestors of the tetrapods [Amemiya et al., 2013; Biscotti et al., 2016].

The 2 species of *Latimeria* have a moderate genome size of ~3 pg/N of which 20% consist of non-LTR transposons [Amemiya et al., 2013; Chalopin et al., 2015].

The lungfishes are freshwater fish that date back to the Devonian. Widely spread in the early Carboniferous, they began to decline in the Mesozoic. Currently, there are 6 living species of lungfish belonging to 3 genera: *Neoceratodus*, *Protopterus*, and *Lepidosiren*. Lungfishes have huge genomes, the largest among animals, exceeding 100 pg/N in *Lepidosiren*.

By studying the size of the bone lacunae (an indirect measure of genome size) in fossil and living lungfishes, Thomson [1972] noted that cell sizes (and presumably genome sizes) were uniformly small in the Devonian and that a progressive and significant increase took place independently in *Neoceratodus* and *Lepidosiren* lineages since the Carboniferous, reaching its climax at the beginning of the Mesozoic. This increase accompanied a pro-

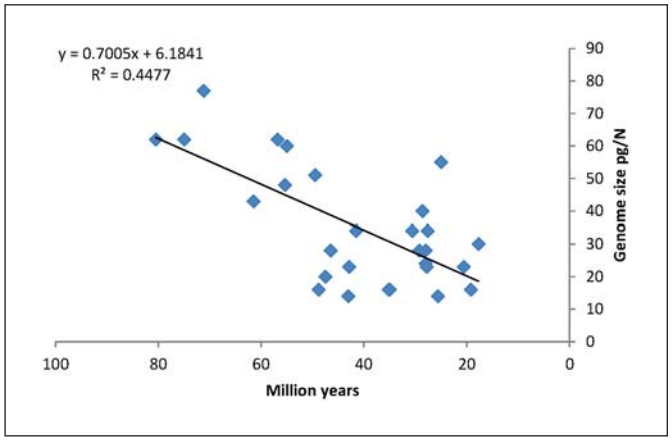

Fig. 2. Reduction of genome size during the evolution of salamanders.

gressive evolutionary decline. The only lungfish that has been studied for the composition of the genome is *N. forsteri*, in which ~40% of the DNA consists of non-LTR transposons, mainly CR1 and L2 [Sirijovski et al., 2005; Metcalfe et al., 2012]. A similar percentage of transposons (between 35 and 40%) was also inferred in *Lepidosiren* [Metcalfe et al., 2012].

A similar study on the size of the bone lacunae in fossil and living amphibians [Thomson and Muraszko, 1978] suggested that cell and genome sizes were relatively small at the origin of this class and that any increase occurred secondarily and separately in different class lineages. Furthermore, it is speculated that a genome size of 2.5–5 pg/N would have been also the baseline for coelacanths, lungfishes, and for all living tetrapods. Similar values of genome and cell size are indeed common in several species of frogs, lepospondyl amphibians, living and extinct non-avian reptiles, and mammals [Organ et al., 2007, 2011]. In this regard it has been speculated that a genome size included in the above range would represent the ancestral and characteristic value of the entire sarcopterygian lineage from which the large genomes of Dipnoi and salamanders and the small genomes of birds would be secondarily derived [Organ et al., 2011].

Amphibians include 3 extant orders which differ in genome size: the frog and Apoda genomes are moderate, ranging from 0.95 to about 13 pg/N, urodeles possess instead very high DNA values between 13.5 and 60 pg/N [Gregory, 2005]. Increases in amphibian genome sizes mainly depend on the increase in repetitive DNA, especially in the moderately repetitive C_0t analysis fraction [Morescalchi and Olmo, 1982] and on a lengthening of the introns, which in *Ambystoma mexicanum* and in other salamanders are longer than in humans, chickens, and frogs [Smith et al., 2009; Eo et al., 2012; Sun et al., 2012a; Voss et al., 2013]. The presence of transposons has been studied only in 2 frog species and in 7 species of salamanders: the primitive *Cryptobranchus* and 6 species of plethodontids, one of the most advanced family of the suborder. In the 2 frog species the percentage of transposons is ~40%, with a prevalence of DNA transposons in *Xenopus* and LTR retrotransposons in *Nanorana* (table 2) [Sun et al., 2015]. In urodeles, transposons range from 25 to 50%, and almost all belong to Gypsy LTR, DIRS, and ERV 1, which indicates that salamander transposons all have the same origin (table 2) [Sun et al., 2012a, b; Sun and Mueller, 2014].

Some authors speculate that Lissamphibia have originated in the Permian, others in the early Triassic [Marjanovic and Laurin, 2007]. The oldest urodelian fossils date back to the middle Jurassic [Gao and Shubin, 2003]. Therefore, it is not easy to imagine the scenario in which the very large genomes found in extant salamanders would have originated. By comparing the genome sizes to the time of speciation in 28 species of salamanders, it is possible to infer that the largest genomes are found in the oldest species and that a gradual contraction happened up to the lowest values found in more recent species (fig. 2; online suppl. table 1). Considering the uniformity of transposon percentages, it can be assumed that the increase in the size of the salamander genomes is derived from a burst of a single group of LTR retrotransposons during the Paleocene era [Sun and Mueller, 2014] when the primitive cryptobranchids appeared. Subsequently, due to the loss of transposon sequences, there would have been a progressive reduction in genome size until the Miocene.

The non-avian reptiles own moderate genome sizes ranging from 1.1 to 5.4 pg/N, values similar to those assumed for the basal-most lobe-finned fishes and tetrapods. Transposons account for on average 30% of the entire genome, and they all are non-LTR retrotransposons. Birds are characterized by some of the smallest genomes among vertebrates, ranging from 1 to 2.2 pg/N with an average value of 1.4 pg/N. The percentage in transposons is also very low with an average of 6.8% (table 2). It has been suggested that such small genomes have evolved to acquire high metabolism levels, essential for the ability to fly [Hughes and Hughes, 1995]. However, the study of fossil cell sizes in dinosaurs has shown that the reduction of cell and genome sizes occurred between 230 and 250 Mya in saurischians (the lineage from which birds have originated), long before the appearance of the first birds,

and this would have depended on a drastic reduction of non-LTR transposons in this lineage [Organ et al., 2007].

The genome sizes of mammals have a range similar to that of non-avian reptiles, from a minimum of 1.7 to a maximum of 8.4 pg/N and an average of 3.5 pg/N, close to the value for man. Regarding transposable elements, mammals present an average of 40%, the highest of all amniotes, if compared to 29% in non-avian reptiles and 6% in birds (table 2). Also in this class, transposons have caused an elongation of the introns [Wang et al., 2012].

A particular situation was described in the vespertilionids: since 36 Mya this group has experienced a rapid adaptive radiation, leading to the onset of the mammal family most rich in species (>400). This radiation coincided with an initial burst of DNA transposons that would have facilitated the rapid diversification of these bats [Platt et al., 2014].

Analyzing the trends in genome sizes with the percentage of transposons in chordates, there is a situation similar to that seen in protostomes. In fact, even in this phylum, for relatively low values (<5 pg/N) there is a significant linear and direct correlation between the increase in genome size and the increase in transposon percentage which suggests that the expansion of the genome size is caused primarily, if not only, by the expansion of transposable sequences (fig. 3A). Vice versa, if we consider species with genomes >5 pg/N, the correlation shows a logarithmic pattern which suggests that the repetitive DNA, and in particular transposons, are not the only cause of the increase in the genome size (fig. 3B). Alternatively, it is assumed that a certain percentage of transposons are stored without being subjected to any selective pressure and are thereby free to accumulate mutations that in time would mask the original repetitiveness.

Similarly to what is known in protostomes, even in deuterostomes there is certain variability with regard to the different classes of transposons. In general in the ray-finned fish DNA transposons are more frequent, while in tetrapods, except for the salamanders, the non-LTR retrotransposon subclass is more frequent (table 2).

In animals, as in all eukaryotes, the contribution of transposons to the changes in genome size would only be quantitative, since a limited correlation between the diversity of transposons and genome size was found restricted to species harboring genomes <500 Mb and since trends in this correlation are not the same in animals and in plants [Elliott and Gregory, 2015b].

Although the range of variation in genome size is different in protostomes and deuterostomes, there are common characteristics. There is a correlation between genome size and intronic length [Moriyama et al., 1998; Vinogradov, 1999a; Wang et al., 2012; Zhang and Edwards, 2012]. Moreover, the rate of change in genome sizes is mainly dependent on variations in the proportion of repetitive DNA, especially transposons, which is in turn proportional to the initial genome size [Oliver et al., 2007]. Finally, up to certain values, the ratio of the increase in genome size to the increase in the percentage of transposons is linear, while beyond certain limits the trend shows a logarithmic pattern, indicating that further expansions of the genome depend on DNA sequences whose degree of repetitiveness is not detectable (online suppl. fig. 4).

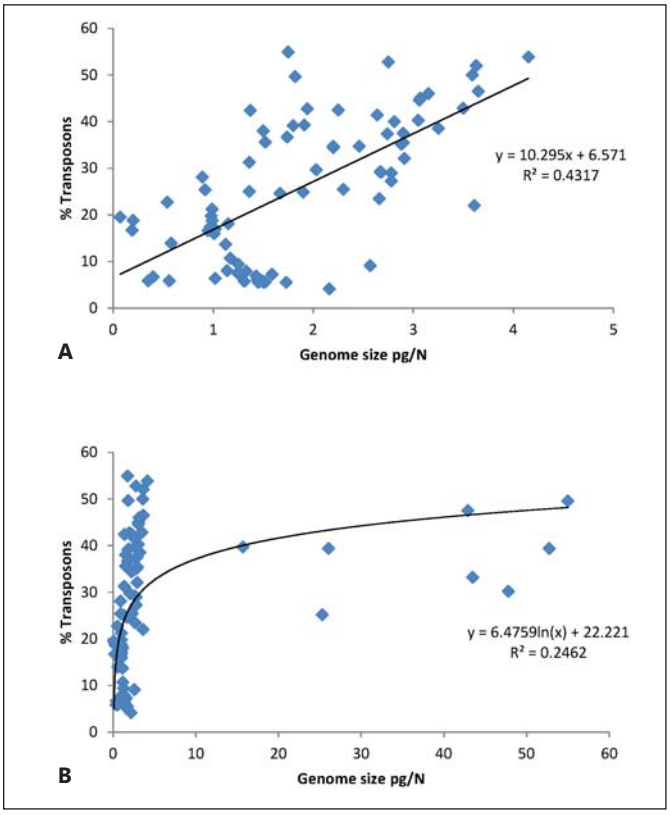

Fig. 3. A Relationship between genome size and percentage of transposons in chordates excluding dipnoans and salamanders. **B** Relationships between genome size and percentage of transposons in chordates including dipnoans and salamanders.

Mechanisms and Causes

There are many factors influencing changes in genome size: whole genome duplication, segmental duplication, DNA repeat proliferation, polyploidy, etc. However, most

of the results obtained so far and mentioned in the previous section point out that the most important cause of genome expansion is represented by the amplification of transposable elements [Kidwell, 2002; Chalopin et al., 2015]. Given their ability to rapidly replicate, transposable elements represent one of the best mechanisms, if not the very best, able to relatively quickly determine changes in genome size [Dufresne and Jeffery, 2011].

In this regard there are some important issues that require in-depth analysis:

- the causes of amplification of transposable elements;
- the mechanisms interacting with transposon activity impacting the expansion and the contraction of the genome;
- the influence of genome size and transposon percentage on structural and functional characteristics of the genome and the cells;
- the effect of changes in genome size and transposon percentage on evolutionary processes and in particular on the adaptation to changing environmental conditions.

Stimulation of Transposon Activity

Barbara McClintock [1984] formerly suggested that the activity of mobile elements in the genome represented a reaction to environmental stressors. This hypothesis was supported by several other authors [Capy et al., 2000; Kidwell, 2002; Chénais et al., 2012; Piacentini et al., 2014], and numerous cases of transposon activation due to environmental stress conditions have been observed in plants, where one of the most influential environmental factors is represented by temperature [Vitte and Panaud, 2005; Kelly and Leitch, 2011; Chénais et al., 2012; Ito, 2013; Wheeler, 2013; Ishiguru et al., 2014; Kim et al., 2014]. Examples of correlations between environmental stressors and transposon activity are rarer in the animal kingdom. However, similar connections were noted in *Drosophila melanogaster*, where differences in the rate of transposition were related to the development of temperature [Capy et al., 2000; Kim et al., 2014; Piacentini et al., 2014] and to the development of a resistance to pesticides [Chénais et al., 2012]. In human cells a reorganization of the transcriptome after thermal shock was described, putatively involving SINE transposon sequences [Wheeler, 2013]. Changes in the proportion of repetitive DNA in relation to different latitudes and altitudes were observed in some invertebrates [Fielman and Marsh, 2005; Dufresne and Jeffery, 2011] and vertebrates [Litvinchuck et al., 2007]. Moreover, an indirect indication of the relation between transposon activity and environmental stresses would be the large genome size observed in some fish and crustaceans living in the deep waters [Rees et al., 2007, 2008; Bonnivard et al., 2009].

Transposon activation induced by environmental stressors causes deleterious effects in certain cases but favors an increase in genetic variability in others. Therefore, it represents an effective adaptive response to drastic environmental changes [Piacentini et al., 2014].

One of the most extreme changes in genome size and in the percentage of transposons regards vertebrates involved in the water-to-land transition characterized by the appearance of huge genomes such as those of lungfishes and salamanders. The activation of transposons as a result of environmental stress may be the best explanation for the rapid emergence of these huge genomes, although there is currently no direct evidence.

The transition from lungfishes to tetrapods and the early stages of the amphibian evolution began in the late Devonian and continued until the Carboniferous, a very long period with extreme climate changes.

During the Carboniferous, in lungfishes there was a progressive increase in genome size starting from values of 2.5–5 pg/N and leading to values >100 pg/N found in extant species. This increase, independently taking place in different lineages, was likely accompanied by a dramatic reduction of morphological and taxonomic evolution [Thomson, 1972; Stanley, 1975]. This evolutionary decline is certainly due to drastic environmental changes, since in the second half of the Carboniferous (Pennsylvanian) severe climate changes with temperature drops and aridification phenomena took place, also provoking mass extinction [Sahney et al., 2010]. Lungfishes, living in shallow freshwater, are particularly sensitive to temperature changes and drying. Therefore, increases in genome size were very likely the consequence of a burst of transposon amplification (probably non-LTR) induced by stressful situations.

A similar increase in genome size was observed in amphibians whose fossil forms appeared in the late Devonian and achieved the highest evolutionary success in the Carboniferous. In some amphibian fossils, especially temnospondyls, a genome expansion would have taken place, leading to high values similar to those of some living salamanders [Thomson and Muraszko, 1978]. Again, this expansion would have occurred during the latter stages of the amphibian radiation in the late Permian and early Triassic period, in which ~70% of the terrestrial amphibian families probably died [Erwin, 1994; Roelants et al., 2007]. Also in this class it is therefore likely that the increases in genome size depended on a burst of transposons stimulated by environmental stress.

The presence of the enormous genomes possessed by living salamanders is less clear to define, also because genome sizes are much smaller in frogs and Apoda [Gregory, 2005]. The oldest urodelian fossils date from the middle Jurassic [Gao and Shubin, 2003], long after the disappearance of temnospondyls [Marjanovic and Laurin, 2007, 2013]. It is therefore not possible to determine whether the genome expansion of the living salamanders is a direct result of the expansion observed in temnospondyls or whether it is a new and independent one.

As mentioned in the previous section, the large genomes of extant urodeles derived from an amplification of LTR transposons dated back to the end of the late Cretaceous, another period characterized by mass extinctions. Therefore, it is possible that even this LTR transposon burst would have been stimulated by severe conditions of environmental stress.

Regulation of Transposon Activity

In some cases transposons can provide evolutionary advantages to the host as for the molecular domestication that leads to the appearance of new functional genes or regulatory sequences or the ability to increase the genetic variability through their mutagenic action [Kidwell, 2002; Schmidt and Anderson, 2006; Volff, 2006; Boehne et al., 2008; Chalopin et al., 2012; Lee and Kim, 2014; Piacentini et al., 2014; Platt et al., 2014; Chalopin et al., 2015]. In most cases, however, their activity has a deleterious effect: they can alter gene expression by fitting into regulatory sequences, their insertion can result in deleterious mutations or even in extensive chromosomal rearrangements [Petrov et al., 2003], and they can impose an exaggerated functional load to the host due to an increased replication of foreign DNA [Kidwell, 2002; Johnson, 2007; Dufresne and Jeffery, 2011; Chenais et al., 2012; Sun et al., 2012a; Lee and Kim, 2014]. In response to this situation the organisms have triggered various mechanisms to curb the activity of transposons and keep the genome size within certain limits.

Comparing genome size and the percentage of transposons in different species of plants and animals, it seems clear that the dimensions of the genomes are the result of a complex balance between gain and loss of repeated sequences [Petrov et al., 2003]. The propensity of transposon insertion would not be random, but it would depend on specific compositional and functional characteristics, including the initial genome size in some organisms [Dufresne and Jeffery, 2011]. Lynch and Conery [2003] have hypothesized that below a minimum genome size the insertion of a transposon would not be possible; at intermediate sizes only some species allow insertions, while above a specific value all species would be contaminated.

The control of the genome size is implemented through 2 phases: the inactivation of transposons with a consequent limitation of their insertion and the deletion of transposons and other repetitive DNAs.

Transposon inactivation depends largely on 2 mechanisms: methylation [Johnson, 2007; Kelly and Leitch, 2011; Wheeler, 2013] and interference of small RNAs such as plant small interfering RNAs (siRNAs) and animal PIWI-interacting RNAs (piRNAs, probably evolved as a defense mechanism against viruses) [Arensburger et al., 2010; Ito, 2013]. Initially, methylation was thought to be the only responsible mechanism for transposon control in plants, and only later was it noted that transposon silencing depends on the interference of siRNAs [Ito, 2013]. The prevailing mechanism in animals is piRNAs [Wheeler, 2013; Klenov et al., 2014], which have a very ancient origin given their presence in the sponge *Amphimedon* and in the cnidarian *Nematostella* [Grimson et al., 2008].

Transposon silencing through piRNAs entails the recognition of aberrant double-helix RNAs derived from transposons and their subsequent processing to small RNAs (sRNAs) [Slotkin and Martienssen, 2007]. Small single-stranded RNAs derived from an sRNA filament may drive the degradation on transposon sequences or they may affect DNA methylation and histone modification to prevent the transcriptional activity of transposons [Kelly and Leitch, 2011].

The role of piRNAs in transposon inactivation and the resulting control over expansion of the genome has been demonstrated by a genome comparison between *D. melanogaster* and *Aedes aegypti*. The latter harbors a genome 8× larger than the former and with a much higher percentage of transposons (table 2). The diversity of transposon sequences is comparable in both species; however, only 19% of *Aedes* piRNAs map on mobile elements compared to 51% in *Drosophila*. As a result the action of piRNAs against transposons is much smaller in the first species, easily allowing their insertion and a subsequent increase in genome size [Arensburger et al., 2010]. A similar situation could be observed in a comparison between teleosts and mammals. In the former, where genomes are on average smaller and exhibit the highest diversity of transposons among all vertebrates [Chalopin et al., 2015], the piRNA genes have a faster evolution and this may represent an adaptive mechanism to counteract the greater variability in transposons. Moreover, a greater variability in teleost piRNAs has been suggested to be linked to the external fertilization

where gametes are subjected to a greater risk of transposon invasion [Yi et al., 2014].

Methylation is another effective mechanism for the restriction of transposon activity. This is supported by a direct correlation observed in the mouse between the demethylation of the intracisternal A particles (IAPs, a family of LTR retrotransposons) and an increase in their expression, often causing diseases through their insertion into genes [Barbot et al., 2002]. A similar correlation was also observed in mammals between the hypomethylation of retrotransposons in germ cells in the early stages of development, in which these transposons are active, and their hypermethylation in somatic cells, in which they cannot be mobilized [Kazazian, 2004]. Also interesting is the increase in DNA methylation of repeats observed in several metazoans with the increase in genome size [Jabbari et al., 1997; Lechner et al., 2013] and a significant level of methylation found in another deuterostome, the echinoderm *Strongylocentrotus purpuratus* [Regev et al., 1998].

This mechanism, which in some cases may be influenced by sRNAs [Wheeler, 2013; Yi et al., 2014], appears to be less generalized, especially in animals. Indeed, repeated elements are generally highly methylated in plants, yeast, and vertebrates but are less methylated in protostomes, especially in insects [Albalat et al., 2012]. In this class the landscape is very variable: a clear correlation between genome size and methylation levels was not found, and species with a large genome and higher levels of genomic parasites, such as *Culex*, display a low level of methylation [Regev et al., 1998]. In *Crassostrea* it has been observed that only certain classes of repetitive elements represent methylation targets [Gonzalez and Petrov, 2009] and, unlike in vertebrates, methylation would not be the prevalent mechanism in silencing of genes located within the transposons, but it would act mainly at an intergenic level [Riviere, 2014].

Although interference of sRNAs and methylation are the main mechanisms that regulate transposon activity, other processes have been identified, such as the inhibition of transposition (especially retrotransposition) by cytosine deaminase [Dutko et al., 2005; Stenglein and Harris, 2006] or by factors involved in DNA repair [Curcio and Garfinkel, 1999; Bryk et al., 2001], even if their influence seems to be less relevant.

However, in some instances transposons could be able to activate mechanisms to neutralize the action of sRNAs and methylation, causing an increase in and expansion of the genome. This would be done mainly through the action of environmental stressors [Vitte and Panaud, 2005; Kelly and Leitch, 2011; Nosaka et al., 2012; Wheeler, 2013; Piacentini et al., 2014]. An example was observed in *Antirrhinum majus* where a cold-induced hypomethylation favored the transposition of Tm3 transposons [Ito, 2013].

As previously mentioned, the genome size is the result of a balance between insertion and loss of DNA sequences, especially transposons. As well as the mechanisms of transposon amplification, deletion of sequences is not accidental. A key mechanism in the control of genome size would be the so-called indel bias [Petrov, 2001; Gregory, 2004], that is the ratio of the levels of insertions to the removal of the ectopic sequences. The indels are divided into small (1–30 bp) and large (involving thousands of bases) [Sun et al., 2012b]. The mechanisms of insertion and DNA loss are different. In the genome, gene sequences are always spaced by stretches of more or less long non-genic sequences (introns and intergenic sequences). Therefore, very long deletions would bear the risk of eliminating gene sequences. This favors small deletions, which would then be the main cause of DNA loss and genome size decrease [Petrov, 2002a]. A clear negative correlation between the rate of DNA loss through small deletions and genome size is observed when comparing insects with differently sized genomes: *D. melanogaster* (0.16 pg/N), the Luapala cricket (1.93 pg/N), and the locust *Podisma* (16.99 pg/N) have a rate of DNA deletion inversely related to their respective genome sizes: in *Podisma* the loss of DNA is much slower than even that of man (3.5 pg/N) [Petrov et al., 2000; Bensasson et al., 2001; Petrov, 2002b]. Moreover, in some mosquito species the genome shows an inverse correlation with the DNA loss rate [Chen et al., 2015]. Similar situations have been recorded also in vertebrates. In archosaurs the indel rate is higher in birds than in turtles and crocodiles, whose transposon percentage is higher [Green et al., 2014], and in mammals the deletion rate of pseudogenes is lower in humans than in rodents [Graur et al., 1989; Ophir and Graur, 1997]. The indel bias might also explain the huge genomes of salamanders and perhaps also of lungfishes. Indeed, a study of the DNA loss in 5 species of salamanders has shown that the rate in these species is much lower than that found in species belonging to 5 classes of non-urodelian vertebrates [Sun et al. , 2012b; Frahry et al., 2015]. In all these cases the different levels of deletion would depend more on differences in the size of the elided sequences than on their frequency [Bensasson et al., 2001; Sun et al., 2012b].

From a comparison between 2 species of pufferfish and between mouse and man it was noted that in the 2 tetraodontids the loss of DNA is higher than in the other

2 species harboring larger genomes, and that this loss mainly concerns sequences at the level of introns [Loh et al., 2008]. Even in chicken, indels primarily affect introns, while they appear highly reduced (if not absent) in intergenic regions [Rao et al., 2010].

The gain of sequences may take place either through small insertions or through large insertions, and the latter occurrence is in line with the expansion of the genome due to rapid insertions and amplifications, typical for transposons [Petrov, 2002a; Dufresne and Jeffery, 2011].

These mechanisms that control the activity of transposons and the balance between insertion and loss of ectopic sequences clearly explain the changes in genome size within low to intermediate values, but they do not completely explain the appearance and preservation of huge genomes such as those of lungfishes and salamanders. As previously mentioned, some events of transposon amplification and the resulting huge genomes may represent a mechanism stimulated by environmental stress that would provide evolutionary advantages [Piacentini et al., 2014]. However, unlike what we see in species harboring smaller genomes, in the genomes that exceed certain high values a clear and direct relationship between the increase in transposon percentage and the increase in DNA content does not seem to exist. Moreover, the increase in DNA content also depends on an amount of non-coding sequences whose repetition is no longer detectable [Kidwell, 2002; Metcalfe and Casane, 2013]. A possible reason for this occurrence could depend on the different insertion sites of the transposons into the host genome [Zhang et al., 2011]. Transposons inserting within or near genes have a mutagenic effect or may affect gene expression, being detrimental to the host genome, and therefore are subject to purifying selection and more frequent deletions [Metcalfe and Casane, 2013; Shen et al., 2013; Lee and Kim, 2014]. This occurrence is supported by the fact that in many organisms the level of indels is higher in introns than in intergenic sequences [Loh et al., 2008; Rao et al., 2010; Lee and Kim, 2014]. Conversely, transposons inserted in regions far from the genes and with low recombination, like intergenic regions and telomeric or centromeric heterochromatin, are less subject to deletions [Kidwell, 2002; Rao et al., 2010] and therefore tend to be conserved and can progressively accumulate mutations losing their repetitivity. Indeed, for example, in *Neoceratodus*, where most of the transposons belong to non-LTR L2 and CR1 families (present in all classes of vertebrates) [Chalopin et al., 2015], traces of CR1 and other ancient transposons that have largely changed over time were recognizable in the single-copy fraction of DNA [Sirijovski et al., 2005]. A similar situation is found also in some species of salamanders [Metcalfe et al., 2012; Metcalfe and Casane, 2013; Sun and Mueller, 2014] and in the orthopteran *Podisma*, a species displaying a genome much bigger than *Drosophila* and where a large proportion of the DNA excess derives from the accumulation of mutations in older pseudogenes [Bensasson et al., 2001].

Interaction of Transposons and Genome Size with Evolutionary Processes

Although transposon activity and changes in genome size cannot be considered the main drivers of evolution, they may have various relevant effects on certain evolutionary processes. Some of those effects depend on their direct action, while others are mediated by alterations carried out on structural and functional parameters of cells and organisms. One of the direct effects is the influence that transposons may have on genetic variability. Some results suggest that transposons can be considered important factors for the reorganization of the genome through chromosomal rearrangements such as duplications, inversions, and translocations (which have consequences on adaptive phenomena), and also through molecular domestication, a phenomenon giving rise to new coding genes and regulatory elements such as enhancers [Bejerano et al., 2006; Matveev and Okada, 2009; Nakanishi et al., 2012; Piacentini et al., 2014]. The amplification of transposons, especially that resulting from environmental stresses, enables a rapid accumulation of mutations that cause an increase in variability and create the basis for speciation [Piacentini et al., 2014]. An example for this situation is the adaptive radiation of vespertilionids, which corresponds to a burst of activity of transposons [Platt et al., 2014].

However, as it has been previously noted, although transposons can provide evolutionary advantages to the host, in most cases their activity has detrimental consequences [Caceres et al., 2001; Kidwell, 2002; Johnson, 2007; Dufresne and Jeffery, 2011; Chénais et al., 2012; Sun et al., 2012b; Lee and Kim, 2014], and various experimental data suggest that their effect on the rise of genetic variation and speciation might be limited to species with smaller genomes. These correlations between genome size and genetic variability appear to have generally negative effects on the evolutionary processes such as taxonomic diversity, speciation, and extinction. In many animal taxa an inverse correlation was noted between the

level of speciation and genome size [Vinogradov, 2004; Kraaijeveld, 2010] and among taxonomic levels of variability, genome size, and proportion of repetitive DNA [Olmo, 2006; Kraaijeveld, 2010]. In this regard, an inverse correlation between genome size and the level of heterozygosity was observed in teleosts [Yi and Streelman, 2005] and urodeles [Pierce and Mitton, 1980]. A similar correlation was also observed between genome size and chiasma frequency in several vertebrates [Olmo et al., 1989; Peterson et al., 1994] and between chromosome changing rate and genome size in reptiles [Olmo, 2005]. Adaptive radiations observed in various vertebrate taxa, such as saurischians [Organ et al., 2007], hummingbirds [Gregory et al., 2009], pufferfishes [Volff et al., 2003], and plethodontids [Kozak et al., 2006], coincided with significant reductions in genome size followed by bursts of morphological diversification. Moreover, freshwater teleost species harboring larger genomes have a lower level of variability compared to marine species with smaller genomes [Hardie and Hebert, 2004]. Furthermore, in insects it was also noted that only those with a genome <2 pg/N have a large increase in taxonomic diversity, while in clades with larger genomes the variability is much lower [Kraaijeveld, 2010]. A final negative consequence of the increase in genome size seems to be also an increase in endangered species and the related risk of extinction [Vinogradov, 2004].

In the light of these observations the main problem to solve is whether the appearance of large genomes only depends on the tolerance of some species to the accumulation of non-coding DNA, especially transposons, or whether large genomes can provide some evolutionary advantages.

A possible answer to this question may come from an analysis of the consequences of genome amplification on some structural and functional parameters of the cell, which in turn affect the morphological and functional characteristics of the organism exposed to natural selection.

Two important cellular parameters affected by genome size are the duration of the cell cycle and the size of the nucleus and the cell.

Although studies on the relationship between the amount of DNA and the cell cycle length are limited both in plants and in animals, a strong direct correlation has been found between the genome size and the S phase and to a lesser extent with the length of mitosis, while there seems to be no correlation with the duration of the G1 and G2 phases [Nagl, 1974a, b; Grosset and Odartchenko, 1975a, b; Horner and MacGregor, 1983; Vinogradov, 1999b; Simova and Harben, 2012].

Correlations between genome size, nucleus, and cell sizes were noted in early cellular studies that led to the formulation of the concept of the nucleoplasmic ratio [Gregory, 2005]. Overall, although the size of the nucleus and the cell are determined by genetic factors [Cavalier-Smith, 2005] and are also influenced by certain phases of the cell cycle, in different species a direct and positive correlation between genome size and various nucleus and cell morphometric parameters (volume and surface) [Szarski, 1968, 1970, 1976; Olmo and Morescalchi, 1975, 1978; Kuramoto, 1981; Olmo and Odierna, 1982; Olmo, 1983; Gregory, 2005; Mueller et al., 2008; Simova and Harben, 2012] and an inverse correlation between genome size and the surface/volume ratio of the cell were noted [Olmo and Morescalchi, 1975, 1978; Olmo and Odierna, 1982; Olmo, 1983].

An important evolutionary interaction is the influence that the genome size has on the duration of embryonic and larval development through the dimensions of the cell and the duration of the cell cycle, especially of the S phase, the extent of which is 50× shorter in embryonic cells compared to somatic cells [Callan, 1972]. Moreover, it should be remembered that during early development the G1 and G2 phases are extremely short and should not affect the total cycle duration [Tang, 2010].

An inverse correlation between genome size and duration of the development has been noted in some invertebrates and anamniote vertebrates, especially amphibians [Goin et al., 1968; Oeldorf et al., 1978; Horner and MacGregor, 1983; Jockusch, 1997; Gregory, 2002a], in which also an inverse correlation between genome size, growth rate, and differentiation rate in the process of regeneration has been noted [Sessions and Larson, 1987]. Conversely, in amniote vertebrates any correlation between genome size and parameters of development was not noticed [Olmo, 2003; Gregory, 2002a, b, 2005].

The evolutionary importance of the developmental and larval period is evident in frogs and salamanders, where species, reproducing in temporary pools and having a very rapid embryonic and larval development that allows them to adapt to arid environments, have very small genome and cell sizes and a very short cell cycle. Conversely, large genomes and cell sizes are present in species living in water-rich but colder environments and having longer embryonic and larval periods, hindering metamorphosis during unfavorable periods of the year [Goin et al., 1968; Oeldorf et al., 1978].

Increasing the genome size as well as the duration of the development could have an impact also on the developmental complexity [Gregory, 2002a, 2005]. In insects

the increase in DNA content determined the gradual transition from species with complete metamorphosis (holometabolous, limited to species with <2 pg/N) to species with incomplete (hemimatabolous) or absent metamorphosis (ametabolous). In urodeles the increase in the genome allowed the transition from a biphasic life cycle to increasingly frequent cases of paedomorphosis or even to obliged neoteny [Gregory, 2002a, 2005]. In plethodontids it was observed that the ancestral forms probably had a shorter larval period than current forms [Bonett et al., 2014], implying a co-evolution between genome increase and paedomorphosis/neoteny. The progressive development of paedomorphosis and neoteny characterizes also the evolution of the current lungfishes harboring very large genomes [Joss, 2006].

Another important evolutionary parameter related to genome and cell size is the metabolic rate. In many animals an inverse relationship between genome/cell size and metabolic rate was noted [Licht and Lowcock, 1991; Vinogradov, 1995, 1997; Gregory and Hebert, 1999; Gregory, 2003, 2005], even if a precise relationship was determined between cell size and metabolic rate [Monnickendam and Balls, 1973; Szarski, 1976; Starostova et al., 2009], and in some cases no precise correlation was indicated with the genome size [Starostovà et al., 2009]. A particularly influential parameter of the metabolic rate, especially in larger cells, is the cell surface/volume ratio, which limits the exchange of nutrients and gases between the cell and the surrounding environment [Szarski, 1976; Olmo, 1983, 2003]. Also important are the parameters that govern the exchanges between the nucleus and cytoplasm such as the nucleoplasmic ratio [Szarski, 1976], the nuclear surface/volume ratio, and the frequency of the nuclear pores [Olmo, 1983].

A particular consequence of the metabolic level has been hypothesized to be the ability to fly [Hughes and Hughes, 1995]. In birds stronger flyers have genomes significantly smaller than weak flyers and flightless birds. It has been also hypothesized that the reduction in genome size was necessary to acquire the high metabolic levels essential for the flight [Hughes and Hughes, 1995; Hughes, 1999], a reduction depending on the shortening of intron lengths and of other sequences such as transposons [Hughes and Hughes, 1995]. A similar situation was also described in bats, which possess the smallest genomes among mammals and also the shortest introns [Van den Bussche et al., 1995; Gregory, 2005; Zhang and Edwards, 2012]. However, the correlations between small genomes, small cells, and high metabolic rates are not related only to the ability to fly, because it was noted that a reduction in cell size and length of introns had occurred within the archosaurian lineage from which birds originated long before the first flying birds [Waltari and Edwards, 2002; Organ et al., 2007].

Flying insects achieve the highest known mass-specific rates of O_2 consumption in the animal kingdom [Suarez, 2000]. Except for the orthopterans, they have the lowest genome sizes among animals and similarly show a reduction in the length of introns within smaller genomes [Wang et al., 2014]. However, there is no difference in the correlation between the genome size and the metabolic level among orthopterans and groups harboring smaller genomes.

Other correlations were assumed between genome size and other evolutionarily important functional parameters such as the control of the extra/intra-cellular solute composition [Vinogradov, 1998], which, however, does not seem sufficiently supported by the experimental point of view. It is important to note that all the above-mentioned correlations indicate that variations in the dimensions and in the percentage of genomic non-coding DNA are also indirectly subject to selective pressures through their effects on the cell and on the morphometric and functional parameters in the organism.

Conclusions

Although causes and consequences of the interaction between transposon activity and expansion of the genome are demonstrated, their direct or indirect influence on the evolutionary processes, as well as the causes of the so-called C-value enigma are not fully elucidated, but some points seem to be quite defined.

Given their ability to autonomously expand, transposons certainly represent the main cause of the increase in the genome. Their influence would, however, be only quantitative, since it is not possible to notice an equal correlation between genome size and variety of transposons in all eukaryotes.

The current levels of genome size found in many eukaryotes are the results of a balance between the activity of transposons and selective pressures acting on various levels:

- a first pressure on a genomic level due to the defense mechanisms put into practice by the cells to limit any harmful action of transposons;
- a selection pressure aimed at maintaining optimal levels of structural and functional parameters within cells and the organism;

- a selective pressure acting on the effects that changes at the genomic and cytological level have had on morphological and functional characteristics which are important for the adaptation to different environmental conditions.

Relevant genome increasing occurred at particular times of evolution, such as the conquest of the land or adaptation to extreme environments, e.g. the deep sea or very cold areas, due to burst amplification of transposons stimulated by environmental stresses.

References

Albalat R, Martì-Solans J, Canestro C: DNA methylation in amphioxus; from ancestral functions to new roles in vertebrates. Brief Funct Genomics 2:142–155 (2012).

Alfoldi J, Di Palma F, Grabherr M, Williams C, Kong L, et al: The genome of the green anole lizards and a comparative analysis with birds and mammals. Nature 477:587–591 (2011).

Amemiya CT, Alfoldi J, Lee AP, Fan S, Philippe H, et al: The African coelacanth genome provides insight into tetrapod evolution. Nature 496:311–316 (2013).

Arensburger P, Megy K, Waterhouse RM, Abrudan J, Amedeo P, et al: Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. Science 330:86–88 (2010).

Barbot W, Dupressoir A, Lazar V, Heidmann T: Epigenetic regulation of an IAP retrotransposon in the aging mouse: progressive demethylation and de-silencing of the element by its repetitive induction. Nucleic Acid Res 30: 2365–2373 (2002).

Bejerano G, Lowe CB, Ahituv N, King B, Siepel A, et al: A distal enhancer and an ultraconserved exon are derived from a novel retroposon. Nature 441:87–90 (2006).

Bennet MD: The duration of meiosis. Proc R Soc Lond B 178:277–299 (1971).

Bennet MD: Nuclear characters in plants. Brookhaven Symp Biol 25:344–366 (1973).

Bensasson D, Petrov DA, Zhang DX, Hartl DL, Hewitt GM: Genomic gigantism: DNA loss is slow in mountain grasshoppers. Mol Biol Evol 18:246–253 (2001).

Benton MJ: Vertebrate Paleontology, ed 2 (Blackwell Science, Oxford 2000).

Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, et al: The genome of the blood fluke *Schistosoma mansoni*. Nature 460:352–358 (2009).

Berthelot C, Brunet F, Chalopin D, Juanchich A, Bernard M, et al: The rainbow trout genome provides novel insight into evolution after whole-genome duplication in vertebrates. Nat Commun 5:3657 (2014).

Biscotti MA, Gerdol M, Canapa A, Forconi M, Olmo E, et al: The lungfish transcriptome: a glimpse into molecular evolution events at the transition from water to land. Sci Rep 6:21571 (2016).

Boehne A, Brunet F, Galiana-Arnoux D, Schulteis C, Volff JN: Transposable elements as drivers of genomic and biological diversity in vertebrates. Chromosome Res 16:203–215 (2008).

Bonasio R, Zhang G, Ye C, Mutti NS, Fang X, et al: Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Science 329:1068–1071 (2010).

Bonett RM, Steffen MA, Robinson GA: Heterochrony repolarized: a phylogenetic analysis of developmental timing in plethodontid salamanders. EvoDevo 5:27 (2014).

Bonnivard E, Catrice O, Ravaux J, Brown SC, Higuet D: Survey of genome size in 28 hydrothermal vent species covering 10 families. Genome 52:524–536 (2009).

Bovine Genome Sequencing and Analysis Consortium, Elsik CG, Tellam RL, Worley KC, Gibbs RA, et al: The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science 324:522–528 (2009).

Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, et al: The genomic substrate for adaptive radiation in African cichlid fish. Nature 513:375–381 (2014).

Bryk M, Banerjee M, Conte D Jr, Curcio MJ: The Sgs 1 helicase of *Saccharomyces cerevisiae* inhibits retrotransposition of Ty1 multimeric arrays. Mol Cell Biol 21:5374–5388 (2001).

Caceres M, Puig M, Ruiz A: Molecular characterization of two natural hotspots in the *Drosophila buzzatii* genome induced by transposon insertions. Genome Res 11:1353–1364 (2001).

Caenorhabditis elegans Sequencing Consortium: Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science 282:2012–2018 (1998).

Callan HG: Replication of DNA in the chromosomes of eukaryotes. Proc R Soc Lond B 181: 19–41 (1972).

Camacho JP, Ruiz-Ruano FJ, Martin-Blàzquez R, Lòpez-Leòn MD, Cabrero J, et al: A step to the gigantic genome of the desert locust: chromosome sizes and repeated DNAs. Chromosoma 124:263–275 (2015).

Cao Z, Yu Y, Hao P, Di Z, He Y, et al: The genome of *Mesobuthus mertensii* reveals a unique adaptation model of arthropods. Nat Commun 4:2602 (2013).

Capy P, Gasperi G, Biémont C, Bazin C: Stress and transposable elements: co-evolution or useful parasites. Heredity 85:101–106 (2000).

Castoe TA, Jason de Koning AP, Hall KT, Card DC, Schield DR, et al: The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. Proc Natl Acad Sci USA 110:20645–20650 (2013).

Cavalier-Smith T: Skeletal DNA and evolution of genome size. Annu Rev Biophys Bioeng 11: 273–302 (1982).

Cavalier-Smith T: Cell volume and the evolution of genome size, in Cavalier-Smith (ed): The Evolution of Genome Size, pp 105–184 (Wiley, Chichester 1985).

Cavalier-Smith T: Economy, speed and size matter: evolutionary forces driving nuclear genome miniaturization and expansion. Ann Bot 95:147–175 (2005).

Chalopin D, Galiana D, Volff JN: Genetic innovation in vertebrates: Gypsy integrase genes and other genes derived from transposable elements. Int J Evol Biol 2012:724519 (2012).

Chalopin D, Naville M, Plard F, Galiana D, Volff JN: Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. Genome Biol Evol 7: 567–580 (2015).

Chapman JA, Kirkness EF, Simakov O, Hampson SE, Mitros T, et al: The dynamic genome of Hydra. Nature 464:592–596 (2010).

Charlesworth B, Sniegowski P, Stephan W: The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371:215–220 (1994).

Chen L, Zhang G, Shao C, Huang Q, Liu G, et al: Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. Nat Genet 46:253–260 (2014).

Chen XG, Jiang X, Gu J, Xu M, Wu Y, et al: Genome sequence of the Asian Tiger *Aedes albopictus* reveals insights into its biology, genetics and evolution. Proc Natl Acad Sci USA 112:5907–5915 (2015).

Chénais B, Caruso A, Hiard S, Casse N: The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. Gene 509:7–15 (2012).

Chipman AD, Ferrier DE, Brena C, Qu J, Hughes DS, et al: The first myriapod genome sequence reveals conservative arthropod gene content and genome organisation in the centipede *Strigama maritime*. PLoS Biol 12:e1002005 (2014).

Cho YS, Hu L, Hou H, Lee H, Xu J, et al: The tiger genome and comparative analysis with lion and snow leopard genomes. Nat Commun 4: 2433 (2013).

Colbourne JK, Pfrender ME, Gilbert D, Thomas WK, Tucker A, et al: The ecoresponsive genome of *Daphina pulex*. Science 331:555–561 (2011).

Curcio MJ, Garfinkel DJ: New lines of host defense: inhibition of Ty1 retrotransposition by Fus3p and NER/TFIIH. Trends Genet 15:43–45 (1999).

Cwiklinski K, Dalton JP, Dufresne PJ, La Course J, Williams DJ, et al: The *Fasciola hepatica* genome: gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution. Genome Biol 16:71 (2015).

Dalloul RA, Long JA; Zimin AV, Aslam L, Beal K, et al: Multiplatform next-generation sequencing of the domestic turkey *(Meleagris gallopavo)* genome assembly and analysis. PLoS Biol 8:e10000475 (2010).

Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, et al: The draft genome of *Ciona intestinalis:* insight into chordate and vertebrate origins. Science 298:2157–2167 (2002).

Dieterich C, Clifton SW, Schuster LN, Chinwalla A, Delehaunty K, et al: The *Pristionichis pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. Nat Genet 40:1193–1198 (2008).

Donoghue PC, Purnell MA: Genome duplication, extinction and vertebrate evolution. Trends Ecol Evol 20:312–319 (2005).

Drosophila 12 Genome Consortium, Clark AG, Eisen MB, Smith DR, Bergman CM, et al: Evolution of genes and genomes on the *Drosophila* phylogeny. Nature 450:203–218 (2007).

Dufresne F, Jeffery N: A guided tour of large genome size in animals: what we know and where we are heading. Chromosome Res 19: 925–938 (2011).

Dutko JA, Schaefer A, Kenny AE, Cullen BR, Curcio MJ: Inhibition of a yeast LTR retrotransposon by human APOBEC3 cytidine deaminases. Curr Biol 15:661–666 (2005).

Ebeling MA, Atkin NB, Setzer PY: Genome sizes in teleostean fishes: increase in some deep-sea species. Am Nat 105:549–561 (1971).

Ellegren HH, Smeds L, Burri R, Olason PI, Backström N, et al: The genomic landscape of species divergence in *Ficedula flycatchers*. Nature 491:756–760 (2012).

Elliot TA, Gregory TR: What's in a genome? The C-value enigma and the evolution of eukaryotic genome content. Phil Trans R Soc B 370: 20140331 (2015a).

Elliot TA, Gregory TR: Do larger genomes contain more diverse transposable elements? BMC Evol Biol 15:69 (2015b).

Eo SH, Doyle JM, Hale MC, Marra NJ, Ruhl JD, et al: Comparative transcriptomics and gene expression in larval tiger salamanders *(Ambystoma tigrinum)* gill and lung tissues as revealed by pyrosequencing. Gene 492:328–338 (2012).

Erwin DH: The Permo-Triassic extinction. Nature 367:231–235 (1994).

Fielman KT, Marsh AG: Genome complexity and repetitive DNA in metazoans from extreme marine environments. Gene 5:362–398 (2005).

Flot JF, Hespeels B, Li X, Noel B, Arkhipova I, et al: Genome evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. Nature 500: 453–457 (2013).

Frahry MB, Sun C, Chong RA, Mueller RL: Low levels of LTR retrotransposon deletion by ectopic recombination in the gigantic genomes of salamanders. J Mol Evol 80:120–129 (2015).

Gallus S, Hallström BM, Kumar V, Dodt WG, Janke A, et al: Evolutionary histories of transposable elements in the genome of the largest living marsupial carnivore, the Tasmanian devil. Mol Biol Evol 32:1268–1283 (2015).

Gao KQ, Shubin NH: Earliest known crown-group salamanders. Nature 422:424–428 (2003).

Garcia G, Rios N, Gutierrez V: Next-generation sequencing detects repetitive elements expansion in giant genomes of annual killifish genus *Austrolebias* (Cyprinodontidae, Rivulidae). Genetica 14:353–360 (2015).

Ghedin E, Wang S, Spiro D, Caler E, Zhao Q, et al: Draft genome of the filarial nematode parasite *Brugia malayi*. Science 317:1756–1760 (2007).

Gibbs RA, Weinstock GM, Metzker ML, Muzni DM, Sodergren EJ, et al: Genome sequence of the brown Norway rat yields insights into mammalian evolution. Nature 428:493–521 (2004).

Goin OB, Goin CJ, Bachmann K: DNA and amphibian life history. Copeia 3:532–540 (1968).

Gonzalez J, Petrov DA: The adaptive role of transposable elements in the *Drosophila* genome. Gene 448:124–133 (2009).

Graur D, Shuali Y, Li WH: Deletion in processed pseudogenes accumulate faster in rodents than in humans. J Mol Evol 28:279–285 (1989).

Grbić M, Van Leeuwen T, Clark RM, Rombauts S, Rouzé P, et al: The genome of *Tetranychus urticae* reveals herbivorous pest adaptation. Nature 479:487–492 (2011).

Green RE, Braun EL, Armstrong J, Earl D, Nguyen N, et al: Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. Science 346:1254449 (2014).

Gregory TR: Genome size and developmental complexity. Genetica 115:131–146 (2002a).

Gregory TR: Genome size and developmental parameters in the homeothermic vertebrates. Genome 45:833–838 (2002b).

Gregory TR: Variation across amphibian species in the size of the nuclear genome supports a pluralistic, hierarchical approach to the C-value enigma. Biol J Linnean Soc 79:329–339 (2003).

Gregory TR: Insertion-deletion biases and the evolution of genome size. Gene 324:15–34 (2004).

Gregory TR: Genome size and evolution in animals, in Gregory TR (ed): The Evolution of the Genome, pp 3–87 (Elsevier, Burlington 2005).

Gregory TR, Hebert PD: The modulation of DNA content: proximate causes and ultimate consequences. Genome Res 9:317–324 (1999).

Gregory TR, Shorthouse DP: Genome sizes in spiders. J Hered 94:285–290 (2003).

Gregory TR, Andrews CB, McGuire JA, Witt CC: The smallest avian genomes are found in hummingbirds. Proc Biol Sci 276:3753–3757 (2009).

Grimson A, Srivastava M, Fahey B, Woodcroft BJ, Chiang HR, et al: Early origin and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature 455:1193–1197 (2008).

Groenen MA, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, et al: Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491:393–398 (2012).

Grosset L, Odartchenko N: Relationships between cell cycle duration, S-period and nuclear DNA content in erythroblasts of four vertebrate species. Cell Tissue Kinet 8:81–90 (1975a).

Grosset L, Odartchenko N: Duration of mitosis and separate mitotic phases compared to nuclear DNA content in erythroblasts of four vertebrates. Cell Tissue Kinet 8:91–96 (1975b).

Guo B, Zou M, Wagner A: Pervasive indels and their evolutionary dynamics after the fish-specific genome duplication. Mol Biol Evol 29:3005–3022 (2012).

Hanrahan SJ, Johnston JS: New genome size estimates of 134 species of arthropods. Chromosome Res 19:809–823 (2011).

Hardie DC, Hebert PD: The nucleotypic effects of cellular DNA content in cartilaginous and ray-finned fishes. Genome 46:683–706 (2003).

Hardie DC, Hebert PD: Genome-size evolution in fishes. J Fish Aquat Sci 61:1636–1646 (2004).

Hellsten U, Harland RM, Gilchrist MJ, Hendrix D, Jurka J, et al: The genome of the Western clawed frog *Xenopus tropicalis*. Science 328: 633–636 (2010).

Hinegardner R: Cellular DNA content of the Mollusca. Comp Biochem Physiol 47:447–460 (1974).

Hinegardner R: Evolution of genome size, in Ayala FJ (ed): Molecular Evolution, pp 179–199 (Sinauer Associates Inc., Sunderland 1976).

Honeybee Genome Sequencing Consortium: Insight into social insects from the genome of the honeybee *Apis mellifera*. Nature 443:931–949 (2006).

Horner HA, MacGregor HC: C value and cell volume: their significance in the evolution and development of amphibians. J Cell Sci 63: 135–146 (1983).

Howe K, Clarck MD, Torroja CF, Torrance J, Berthelot C, et al: The zebrafish reference genome sequence and its relationship to the human genome. Nature 496:498–503 (2013).

Huang Y, Li Y, Burt DW, Hualan C, Zhang Y, et al: The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. Nature Genet 45:776–783 (2013).

Hughes AL: Adaptive Evolution of Genes and Genomes (Oxford University Press, Oxford 1999).

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

Imai S, Sasaki T, Shimizu A, Asakawa S, Hozi H, et al: The genome size evolution of medaka *(Oryzias latipes)* and fugu *(Takifugu rubripes)*. Gene Gent Syst 82:135–144 (2007).

International Aphid Genome Consortium, Richards S, Gibbs RA, Gerardo NM, Moran N, et al: Genome sequence of the pea aphid *Acyrthosiphon pisum*. PLoS Biol 8:e1000313 (2010).

International Chicken Sequencing Consortium, Hillier LW, Miller W, Birney E, Warren W, et al: Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695–716 (2004).

Ishiguru S, Ogasawara K, Fujino K, Sato Y, Kishima Y: Low temperature-responsive changes in the anther transcriptome's repeat sequences are indicative of stress sensitivity and pollen sterility. Plant Physiol 164:671–682 (2014).

Ito H: Small RNAs and regulation of transposons in plants. Genes Genet Syst 88:3–7 (2013).

Jabbari K, Cacciò S, Pais de Barros JP, Desgrès J, Bernardi G: Evolutionary changes in CpG and methylation levels in the genome of vertebrates. Gene 205:109–118 (1997).

Jex AR, Liu S, Li B, Young ND, Hall RS, et al: *Ascaris suum* draft genome. Nature 479:529–533 (2011).

Jockusch EJ: An evolutionary correlate of genome size change in plethodontid salamanders. Proc R Soc Lond B 264:597–604 (1997).

Johnson LJ: The genome strikes back: the evolutionary importance of defense against mobile elements. Evol Biol 34:121–129 (2007).

Joss JM: Lungfish evolution and development. Gen Comp Endocrinol 148:285–289 (2006).

Kasahara M, Naruse K, Sasaki S, Nakatani Y, Qu W, et al: The medaka draft genome and insights into vertebrate genome. Nature 447:714–719 (2007).

Kazazian HH Jr: Mobile elements: drivers of genome evolution. Science 303:1626–1632 (2004).

Keeling CJ, Yuen MM, Liao NY, Docking TR, Chan SK, et al: Draft genome of the mountain pine beetle *Dendroctonus ponderosae* Hopkins, a major forest pest. Genome Biol 14:R27 (2013).

Kelly LJ, Leitch IJ: Exploring giant plant genomes with next-generation sequencing technology. Chromosome Res 19:939–953 (2011).

Kidwell MG: Transposable elements and the evolution of genome size in eukaryotes. Genetica 115:49–63 (2002).

Kim YB, Oh JH, Mcliver LJ, Rashkovetsky E, Michalak K, et al: Divergence of *Drosophila melanogaster* repeatomes in response to a sharp microclimate contrast in Evolution Canyon, Israel. Proc Natl Acad Sci USA 111:10630–10635 (2014).

Kirkness EF, Haas BJ, Sun W, Braig HR, Perotti MA, et al: Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. Proc Natl Acad Sci USA 107:12168–12173 (2010).

Klenov MS, Lavrov SA, Korbut AP, Stolyarenko AD, Yakushev EY, et al: Impact of nuclear Piwi elimination on chromatin state in *Drosophila melanogaster* ovaries. Nucleic Acid Res 42:6208–6218 (2014).

Kozak KH, Blaine RA, Larson A: Gene lineages and eastern North American palaeodrainage basins: phylogeography and speciation in salamanders of the *Eurycea bislineata* species. Mol Ecol 15:191–207 (2006).

Kraaijeveld K: Genome size and species diversification. Evol Biol 37:227–233 (2010).

Krebs JE, Goldstein ES, Kilpatrick ST: Lewin's Essential Genes, ed 3 (Jones and Bartlett Publishers, Inc., Sudbury 2013).

Kuramoto M: Relationships between number, size and shape of red blood cells in amphibians. Comp Biochem Physiol 69A:771–775 (1981).

Lavoie CA, Platt RN, Novick PA, Counterman BA, Ray DA: Transposable element evolution in *Heliconius* suggests genome diversity within Lepidoptera. Mob DNA 4:21 (2013).

Lechner M, Marz M, Ihling C, Sinz A, Stadler PF, et al: The correlation of genome size and DNA methylation rate in metazoans. Theory Biosci 132:47–60 (2013).

Lee SI, Kim NS: Transposable elements and genome size variations in plants. Genomics Inform 12:87–97 (2014).

Li R, Fan W, Tian G, Zhu H, He L, et al: The sequence and de novo assembly of the giant panda genome. Nature 463:311–317 (2010).

Libertini A, Panozzo M, Zenere R: Chromosome number and genome size in the Antarctic whelk, *Neobuccinum eatoni* (Smith) (Prosobranchia, Neogastropoda). Chrom Inform Service 55:5–6 (1993).

Licht LE, Lowcock LA: Genome size and metabolic rate in salamanders. Comp Biochem Physiol 100B:83–92 (1991).

Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, et al: Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438:803–819 (2005).

Litvinchuk SN, Rosanov JM, Borkin LJ: Correlation of geographic distribution and temperature of embryonic development with the nuclear DNA content in the Salamandridae (Urodela, Amphibia). Genome 50:333–342 (2007).

Loh YH, Brenner S, Venkatesh B: Investigation of loss and gain of introns in the compact genomes of pufferfishes *(Fugu* and *Tetraodon)*. Mol Biol Evol 25:526–535 (2008).

Lynch M, Conery JS: The origins of genome complexity. Science 302:1401–1404 (2003).

Marjanovic D, Laurin M: Fossils, molecules, divergence times, and the origin of Lissamphibians. Syst Biol 56:369–388 (2007).

Marjanovic D, Laurin M: The origin(s) of extant amphibians: a review with emphasis on the 'lepospondyl hypothesis'. Geodiversitas 35:207–272 (2013).

Matveev V, Okada N: Retroposons of salmonid fishes (Actinopteygii: Salmonoidei) and their evolution. Gene 434:16–28 (2009).

McClintock B: The significance of responses of the genome to challenge. Science 226:792–801 (1984).

Metcalfe CJ, Casane D: Accommodating the load: the transposable element content of very large genomes. Mob Genet Elements 3:e24775 (2013).

Metcalfe CJ, Filée J, Germon I, Joss J, Casane D: Evolution of the Australian lungfish *(Neoceratodus forsteri)* genome: a major role for CR1 and L2 LINE elements. Mol Biol Evol 29:3529–3539 (2012).

Meyer A, Schartl M: Gene and genome duplications in vertebrates: the one-to-four (-to eight in fish) rule and the evolution of novel gene functions. Curr Opin Cell Biol 11:699–704 (1999).

Mikkelsen TS, Wakefield MJ, Aken B, Amemiya CT, Chang JL, et al: Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177 (2007).

Mita K, Kasahara M, Sasaki S, Nagayasu Y, Yamada T, et al: The genome sequence of silkworm, *Bombyx mori*. DNA Res 11:27–35 (2004).

Mitreva M, Jasmer DP, Zarlenga DS, Wang Z, Abubucker S, et al: The draft genome of the parasitic nematode *Trichinella spiralis*. Nat Genet 43:228–235 (2011).

Monnickendam MA, Balls M: The relationship between cell size, respiration rates and survival of amphibian tissues in long-term organ cultures. Comp Biochem Physiol 44A:871–880 (1973).

Morescalchi A, Olmo E: Single-copy DNA and vertebrate phylogeny. Cytogenet Cell Genet 34:93–101 (1982).

Moriyama EN, Petrov DA, Hartl DL: Genome size and intron size in *Drosophila*. Mol Biol Evol 15:770–773 (1998).

Moroz LL, Kocot KM, Citarella MR, Dosung S, Norekian TP, et al: The ctenophore genome and the evolutionary origin of neural systems. Nature 510:109–114 (2014).

Mouse Genome Sequencing Consortium, Chinwalla AT, Cook LL, Delehaunty KD, Ginger A: Initial sequencing and comparative analysis of the mouse genome. Nature 420:520–562 (2002).

Mueller RL, Gregory TR, Gregory SM, Hsich A, Boore JL: Genome size, cell size, and the evolution of enucleated erythrocytes in attenuate salamanders. Zoology 111:218–230 (2008).

Nafsey DE, Waterhouse RM, Abai MR, Aganezov SS, Alekseyev MA, et al: Mosquito genomics. Highly evolvable malaria vectors: the genome of 16 *Anopheles* mosquitoes. Science 347: 1258522 (2015).

Nagl W: Roles of heterochromatin in the control of cell cycle duration. Nature 249:53–54 (1974a).

Nagl W: Mitotic cycle time in perennial and annual plants with various amounts of DNA and heterochromatin. Dev Biol 39:342–346 (1974b).

Nakanishi A, Kobayashi N, Suzuki-Hirano A, Nishihara H, Sasaki T, et al: A SINE-derived element constitutes a unique modular enhancer for mammalian diencephalic *Fgf 8*. PLoS One 7:e43785 (2012).

Nene V, Wortman JR, Lawson D, Haas B, et al: Genome sequence of *Aedes aegypti*, a major arbovirus vector. Science 316:1718–1723 (2007).

Noleto RB, de Souza Fonseca Guimaraes F, Paludo KS, Vicari MR, Antoni RF, et al: Genome size evolution in Tetraodontiformes fishes from the Neotropical region. Mar Biotechnol 11:680–685 (2009).

Nosaka M, Itoh J, Nagato Y, Ono A, Ishiwara A, et al: Role of transposon-derived small RNAs in the interplay between genomes and parasitic DNA in rice. PLoS Genet 8:e1002953 (2012).

Oeldorf E, Nishioka M, Bachmann K: Nuclear DNA amounts and developmental rate in holarctic anura. J Zool Syst Evol Res 16:216–224 (1978).

Oliver MJ, Petrov D, Ackerly D, Falkowski P, Schofield OM: The mode and tempo of genome size evolution in eukaryotes. Genome Res 17:594–601 (2007).

Olmo E: Nucleotype and cell size in vertebrates: a review. Bas Appl Histochem 27:227–256 (1983).

Olmo E: Reptiles: a group of transition in the evolution of genome size and of the nucleotypic effect. Cytogenet Genome Res 101:166–171 (2003).

Olmo E: Rate of chromosome changes and speciation in reptiles. Genetica 15:185–203 (2005).

Olmo E: Genome size and evolutionary diversification in vertebrates. Ital J Zool 73:167–171 (2006).

Olmo E, Morescalchi A: Evolution of the genome and cell size in salamanders. Experientia 31: 804–806 (1975).

Olmo E, Morescalchi A: Genome and cell sizes in frogs: a comparison with salamanders. Experientia 34:44–46 (1978).

Olmo E, Odierna G: Relatioships between DNA content and cell morphometric parameters in reptiles. Bas Appl Histochem 26:27–34 (1982).

Olmo E, Stingo V, Cobror O, Capriglione T, Odierna G: Repetitive DNA and polyploidy in selachians. Comp Biochem Physiol 73 B:739–745 (1982).

Olmo E, Capriglione T, Odierna G: Genome size evolution in vertebrates: trends and constraints. Comp Biochem Physiol 92B:447–453 (1989).

Ophir R, Graur D: Patterns and rate of indel evolution in processed pseudogenes from humans and murids. Gene 205:191–202 (1997).

Opperman C, Bird DM, Burke M, Cohn J, Rokhsar DS, Burke M, et al: Sequence and genetic map of *Meloidogyne hapla*: a compact nematode genome for plant parasitism. Proc Natl Acad Sci USA 105:14802–14807 (2008).

Organ CL, Shedlock AM, Meade A, Pagel M, Edwards SV: Origin of avian genome size and structure in non-avian dinosaurs. Nature 446: 180–184 (2007).

Organ CL, Canoville A, Reisz RR, Laurin M: Paleogenomic data suggest mammal-like genome size in the ancestral amniote and derived large genome size in amphibians. J Evol Biol 24:372–380 (2011).

Pagel M, Johnstone RA: Variation across species in the size of the nuclear genome supports the junk-DNA explanation for the C-value paradox. Proc R Soc Lond B 249:119–124 (1992).

Panopoulou G, Poustka AJ: Timing and mechanism of ancient vertebrate genome duplications-the adventure of a hypothesis. Trends Genet 21:559–567 (2005).

Panopoulou G, Hennig S, Groth D, Krause A, Poustka AJ, et al: New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes. Genome Res 13:1056–1066 (2003).

Parker J, Tsagkogeorga G, Cotton JA, Liu Y, Provero P, et al: Genome-wide signatures of convergent evolution in echolocating mammals. Nature 502:228–231 (2013).

Peterson DG, Stack SM, Healy JL, Donohoe BS, Anderson LK: The relationship between synaptonemal complex length and genome size in four vertebrate classes (Osteicthyes, Reptilia, Aves, Mammalia). Chromosome Res 2: 153–162 (1994).

Petrov DA: Evolution of genome size: new approaches to an old problem. Trends Genet 17: 23–28 (2001).

Petrov DA: Mutational equilibrium model of genome size evolution. Theor Popul Biol 61: 533–546 (2002a)

Petrov DA: DNA loss and evolution of genome size in *Drosophila*. Genetica 115:81–91 (2002b).

Petrov DA, Sangster TA, Johnston JS, Hartl DL, Shaw KL: Evidence for DNA loss as a determinant of genome size. Science 287:1060–1062 (2000).

Petrov DA, Aminetzach YT, Davis JC, Bensasson D, Hirsh AE: Size matters: non-LTR retrotransposable elements and ectopic recombination in *Drosophila*. Mol Biol Evol 20:880–892 (2003).

Piacentini L, Fanti L, Specchia V, Bozzetti MP, Berloco M, et al: Transposons, environmental changes, and heritable induced phenotypic variability. Chromosoma 123:345–354 (2014).

Pierce BA, Mitton JB: The relationship between genome size and genetic variation. Am Nat 116:850–861 (1980).

Platt RN, Vanderwege MW, Kern C, Schmidt CJ, Hoffman FG, et al: Large numbers of novel miRNAs originate from DNA transposons and are coincident with a large species radiation in bats. Mol Biol Evol 31:1536–1545 (2014).

Pontius JU, Mullikin JC, Smith DR; Agencourt Sequencing Team, Lindblad-Toh K, et al: Initial sequence and comparative analysis of the cat genome. Genome Res 17:1675–1689 (2007).

Putnam NH, Srivastava M, Hellstein U, Dirks B, Chapman J, et al: Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science 317:86–94 (2007).

Rao YS, Wang ZF, Chai XW, Wu GZ, Nie QH, et al: Indel segregating within introns in the chicken genome are positively correlated with the recombination rates. Hereditas 147:53–57 (2010).

Ravi V, Venkatesh B: Rapidly evolving fish genomes and teleost diversity. Curr Opin Genet Dev 18:544–560 (2008).

Rees DJ, Dufresne F, Glémet H, Belzile C: Amphipod genome sizes: first estimates for Arctic species reveal genomic giants. Genome 50: 151–158 (2007).

Rees DJ, Belzile C, Glémet H, Dufresne F: Large genomes among caridean shrimp. Genome 51:159–163 (2008).

Regev A, Lamb MJ, Jablonka E: The role of DNA methylation in invertebrates: developmental regulation or genome defense? Mol Biol Evol 15:880–891 (1998).

Renfree MB, Papenfuss AT, Deakin JE, Lindsay J, Heider T, et al: Genome sequence of an Australian kangaroo, *Macropus eugenii*, provides insight into the evolution of mammalian reproduction and development. Genome Biol 12:R81 (2011).

Rhesus Macaque Genome Sequencing and Analysis Consortium, Gibbs RA, Rogers J, Katze MG, Bumgarner R, et al: Evolutionary and biomedical insights from the rhesus macaque genome. Science 316:222–234 (2007).

Riviere G: Epigenetic features in the oyster *Crassostrea gigas* suggestive of functionally relevant promoter DNA methylation in invertebrates. Front Physiol 5:1–7 (2014).

Rodriguez-Juiz AM, Torrado M, Mèndez J: Genome-size variation in bivalve mollusks determined by flow cytometry. Marine Biol 126: 489–497 (1996).

Roelants K, Gower DJ, Wilkinson M, Loader SP, Biju SD, et al: Global patterns of diversification in the history of modern amphibians. Proc Natl Acad Sci USA 104:887–892 (2007).

Rondeau EB, Minkley DR, Leong JS, Messmer AM, Jantzen JR, et al: The genome and linkage map of the northern pike *(Esox lucius)*: conserved synteny revealed between the salmonid sister group and the Neoteleostei. PLoS One 9:e102089 (2014).

Ryan JF, Pang K, Schnitzel CE, Nguyen AD, Moreland RT, et al: The genome of the ctenophore *Mnenopsis leidyi* and its implications for cell type evolution. Science 342:1242592 (2013).

Sahney S, Benton MJ, Falcon-Lang HJ: Rainforest collapse triggered Carboniferous tetrapod diversification in Euroamerica. Geology 38: 1079–1082 (2010).

Schartl M, Walter RB, Shen Y, et al: The genome of the platyfish, *Xiphophorus maculatus,* provides insight into evolutionary adaptation and several complex traits. Nat Genet 45:567–572 (2013).

Schistosoma japonicum Genome Sequencing and Functional Analysis Consortium, Zhou Y, Zheng H, Chen Y, Zhang L, et al: The *Schistosoma japonicum* genome reveals features of host-parasite interplay. Nature 460:345–351 (2009).

Schmidt AL, Anderson LM: Repetitive DNA elements as mediators of genomic changes in response to environmental cues. Biol Rev 81: 531–543 (2006).

Scott JG, Warren WC, Beukeboom LW, Bopp D, Clark AG, et al: Genome of the house fly *Musca domestica* L., a global vector of diseases with adaptations to a septic environment. Genome Biol 15:466 (2014).

Sessions SK, Larson A: Developmental correlates of genome size in plethodontid salamanders and their implications for genome evolution. Evolution 41:1239–1251 (1987).

Shaffer HB, Minx P, Warren DE, Shedlock AM, Thomson RC, et al: The western painted turtle genome, a model for the evolution of extreme physiological adaptations in a slowly evolving lineage. Genome Biol 14:R28 (2013).

Shen JJ, Dushoff J, Bewick AJ, Chain FJJ, Evans BJ: Genomic dynamics of transposable elements in the western clawed frog *(Silurana tropicalis)*. Genome Biol Evol 5:998–1009 (2013).

Simakov O, Marletaz F, Cho SJ, Edsinger-Gonzales E, Havlak P, et al: Insight into bilaterian evolution from three spiralian genomes. Nature 493:526–531 (2013).

Simova I, Harben T: Geometrical constraints in the scaling relationships between genome size, cell size and cell cycle length in herbaceous plants. Proc Biol Sci 279:867–875 (2012).

Sion L, Bozzano A, D'Onghia G, Capezzuto F, Panza M: Chondrichthyes species in deep waters of the Mediterranean sea. Scientia Marina 68:153–162 (2004).

Sirijovski N, Woolnough C, Rock J, Joss JM: NfCR1, the first non-LTR retrotransposon characterized in the Australian lungfish genome, *Neoceratodus forsteri,* shows similarities to CR1-like elements. J Exp Zool Mol Dev Evol 304B:40–49 (2005).

Slotkin RK, Martienssen R: Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 8:272–285 (2007).

Small KS, Brudno M, Hill MM, Sidow A: A haplome alignment and reference sequence of the high polymorphic *Ciona savigny* genome. Genome Biol 8:R41 (2007).

Smith CD, Zimin A, Holt C, Abouheif E, Benton R, et al: Draft genome of the globally widespread and invasive Argentine ant (*Linepithema humile*). Proc Natl Acad Sci USA 108: 5673–5678 (2011).

Smith CR, Smith CD, Robertson HM, Helmkampf M, Zimin A, et al: Draft genome of the red harvester ant *Pogonomyrmex barbatus*. Proc Natl Acad Sci USA 108:5667–5672 (2011).

Smith JJ, Putta S, Zhu W, Pao GM, Verma IM, et al: Genic regions of a large salamander genome contain long introns and novel genes. BMC Genomics 10:19 (2009).

Smith JJ, Kuraku S, Holt C, Sauka-Spengler T, Jiang N, et al: Sequencing of the sea lamprey *(Petromyzon marinus)* genome provides insight into vertebrate evolution. Nat Genet 45: 415–421 (2013).

Song B, Cheng S, Sun Y, Zhong X, Jin J, et al: A genome draft of the legless anguid lizard *Ophisaurus gracilis*. Gigascience 4:17 (2015).

Stanley SM: A theory of evolution above the species level. Proc Natl Acad Sci USA 72:646–650 (1975).

Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmstrøm M, et al: The genome sequence of Atlantic cod reveals a unique immune system. Nature 477:207–210 (2011).

Starostovà Z, Kubicka L, Konarzewski, Kozlowski, Kratochvil L: Cell size but not genome size affects scaling of metabolic rate in eyelid geckos. Am Nat 174:E101–105 (2009).

Stein LD, Bao Z, Blasair D, Blumenthal T, Brent MR, et al: The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. PLoS Biol 1:e45 (2003).

Stenglein MD, Harris RS: APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. J Biol Chem 281:16837–16841 (2006).

Stingo V, Du Buit MH, Odierna G: Genome size of some selachian fishes. Boll Zool 47:129–137 (1980).

Stingo V, Rocco L, Improta R: Chromosome markers and karyology of selachians. J Exp Zool Suppl 2:175–185 (1989).

St John JA, Braun EL, Isberg SR, Miles LG, Chong AY, et al: Sequencing three crocodilian genomes to illuminate the evolution of archosaurs and amniotes. Genome Biol 13:415 (2012).

Suarez RK: Energy metabolism during insect flight: biochemical design and physiological performance. Physiol Biochem Zool 73:765–771 (2000).

Suen G, Teiling C, Li L, Holt C, Abouheif E, et al: The genome sequence of the leaf-cutter ant *Atta caphalotes* reveals insights into its obligate symbiotic lifestyle. PLoS Genet 7: e1002007 (2011).

Sun C, Mueller RL: Hellbender genome sequences shed light on genomic expansion at the base of crown salamanders. Genome Biol Evol 6: 1818–1829 (2014).

Sun C, Lopez Arriaza JR, Mueller RL: Slow DNA loss in the gigantic genomes of salamanders. Genome Biol Evol 4:1340–1348 (2012a).

Sun C, Shepard DB, Chong RA, Lopez Arriaza J, Hall K, et al: LTR transposons contribute to genomic gigantism in plethodontid salamanders. Gen Biol Evol 4:168–183 (2012b).

Sun YB, Xiong ZJ, Xiang XY, Liu SP, Zhou WW, et al: Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative evolution of tetrapod genome. Proc Natl Acad Sci USA 112:1257–1262 (2015).

Sundberg LR, Pulkkinen K: Genome size evolution in macroparasites. Int J Parasitol 45:285–288 (2015).

Szarski H: Changes in the amount of DNA in cell nuclei during vertebrate evolution, in Orvig T (ed): Current Problems of Lower Vertebrate Phylogeny, pp 445–453 (Nobel Symp No 4 Almqvist & Wiksell, Stockholm 1968).

Szarski H: Changes in the amount of DNA in cell nuclei during vertebrate evolution. Nature 226:651–652 (1970).

Szarski H: Cell size and nuclear DNA content in vertebrates. Int Rev Cytol 44:93–111 (1976).

Tang ZI: The domino and clock models of cell cycle regulation. Nat Educ 3:56 (2010).

Thomson KS: An attempt to reconstruct evolutionary changes in the cellular DNA content of lungfish. J Exp Zool 180:363–372 (1972).

Thomson KS, Muraszko K: Estimation of cell size and DNA content in fossil fishes and amphibians. J Exp Zool 205:315–320 (1978).

Tribolium Genome Sequencing Consortium, Richards S, Gibbs RA, Weinstock GM, Brown SJ, et al: The genome of the model beetle and pest *Tribolium castaneum*. Nature 452:949–955 (2008).

Ullate-Agote A, Milinkovitch MC, Tzika A: The genome sequence of the corn snake *(Pantherophis guttatus)*, a valuable resource for EvoDevo studies in squamates. Int J Dev Biol 58:881–888 (2014).

Van den Bussche RA, Longmire JL, Baker RJ: How bats achieve a small C-value: frequency of repetitive DNA in *Macrotus*. Mamm Genome 6:521–525 (1995).

Vanderpoele K, De Vos W, Taylor JS, Meyer A, Van de Peer Y: Major events in the genome evolution of vertebrates: paranome age and size differ considerably between ray-finned fishes and land vertebrates. Proc Natl Acad Sci USA 101:1638–1643 (2004).

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

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

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

Vinogradov AE: Intron-genome size relationship on a large evolutionary scale. J Mol Evol 49: 376–384 (1999a).

Vinogradov AE: Genome in toto. Genome 42: 361–362 (1999b).

Vinogradov AE: Larger genomes for molluskan land pioneers. Genome 43:211–212 (2000).

Vinogradov AE: Genome size and extinction risk in vertebrates. Proc R Soc Lond B 271:1701–1705 (2004).

Vitte C, Panaud OL: TR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. Cytogenet Genome Res 110:91–107 (2005).

Volff JN: Genome evolution and biodiversity in teleost fish. Heredity 94:280–294 (2005).

Volff JN: Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. Bioessays 218:913–922 (2006).

Volff JN, Bouneau L, Ozouf-Costaz C, Fisher C: Diversity of retrotransposable elements in compact pufferfish genomes. Trends Genet 19:674–678 (2003).

Voss SR, Putta S, Walker JA, Smith JJ, Maki N, et al: Salamander *Hox* clusters contain repetitive DNA and expanded non-coding regions: a typical *Hox* structure for non-mammalian tetrapod vertebrates? Hum Genomics 7:9 (2013).

Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, et al: Genome sequence and comparative analysis, and population genetics of the domestic horse. Science 326:865–867 (2009).

Waltari E, Edwards SV: Evolutionary dynamics of intron size and physiological correlates in archosaurs. Am Nat 160:539–552 (2002).

Wang D, Su Y, Wang X, Lei H, Yu J: Transposon-derived repetitive sequences play distinct functional roles in Mammalian intron size expansion. Evol Bioinform Online 8:301–319 (2012).

Wang S, Zhang L, Meyer E, Bao Z: Genome-wide analysis of transposable elements and tandem repeats in the compact placozoan genome. Biol Direct 5:18 (2010).

Wang X, Chen W, Huang Y, Sun J, Men J, et al: The draft genome of the carcinogenic human liver fluke *Clonorchis sinensis*. Genome Biol 12:R107 (2011).

Wang X, Fang X, Yang P, Jiang X, Jiang F, et al: The locust genome provides insight into swarm formation and long-distance flight. Nat Commun 5:2957 (2014).

Wang Z, Pascual-Anaya J, Zadissa A: The draft genome of the soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. Nature 45:701–706 (2013).

Warren JH, Richards S, Desjardins CA, Niehuis O, Gadau J, et al: Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. Science 327:343–348 (2010).

Warren WC, Hillier LW, Marshall Graves JA, et al: Genome analysis of the platypus reveals unique signature of evolution. Nature 453: 175–183 (2008).

Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LDW, et al: The genome of a songbird. Nature 464:757–762 (2010).

Wheeler BS: Small RNAs, big impact: small RNA pathways in transposon control and their effect on the host stress response. Chromosome Res 21:587–600 (2013).

Yi M, Chen F, Luo M, Cheng Y, Zhao H, et al: Rapid evolution of piRNA pathway in the teleost fish; implication for an adaptation to transposon diversity. Genome Biol Evol 6: 1393–1407 (2014).

Yi S, Streelman JT: Genome size is negatively correlated with effective population size in ray-finned fish. Trends Genet 21:643–646 (2005).

Yoshida M, Ishikura Y, Moritaki T, Shoguchi E, Shmizu KK, et al: Genome structure analysis of mollusks revealed whole genome duplication and lineage repeat variation. Gene 483: 63–71 (2011).

You M, Yue Z, He W, Yang X, Yang G, et al: A heterozygous moth genome provides insights into herbivory and detoxification. Nat Genet 45:220–225 (2013).

Young ND, Jex AR, Li B, Liu S, Yang L, et al: Whole-genome sequences of *Schistosoma haematobium*. Nat Genet 44:221–225 (2012).

Zhan S, Merlin C, Boore JL: The monarch butterfly genome yields insights into long-distance migration. Cell 147:1171–1185 (2011).

Zhan X, Pan S, Wang J, Dixon A, He J, et al: Peregrine and saker falcon genome sequences provide insight into evolution of a predatory lifestyle. Nat Genet 45:563–566 (2013).

Zhang G, Fang X, Guo X, Li L, Luo R, et al: The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490: 49–54 (2012).

Zhang G, Li B, Li C, Gilbert T, Jarvis ED, et al: Comparative genomic data of the Avian phylogenomics project. Gigascience 3:26 (2014).

Zhang J, Yu C, Krishnaswarmy L, Peterson T: Transposable elements as catalysts for chromosome rearrangements. Methods Mol Biol 701:315–326 (2011).

Zhang Q, Edwards S: The evolution of intron size in amniotes: a role for powered flight? Genome Biol Evol 4:1033–1043 (2012).

Zheng H, Zhang W, Zhang I, et al: The genome of the hydatid tapeworm *Echinococcus granulosus*. Nat Genet 45:1168–1175 (2013).