C. R. Palevol 13 (2014) 611–621

Contents lists available at ScienceDirect

Comptes Rendus Palevol

www.sciencedirect.com

General palaeontology, systematics and evolution (Evolutionary processes and microevolution)

A new look at genome size, evolutionary duration and genetic variation in salamanders

Un autre regard sur la taille du génome, la durée de l'évolution et la variation génétique chez les salamandres

John Herrick[a,*], Bianca Sclavi[b,**]

[a] *Department of Physics, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia VSA 1S6, Canada*
[b] *LBPA, UMR 8113 du CNRS, ENS Cachan, 94235 Cachan, France*

ARTICLE INFO

Article history:
Received 21 September 2013
Accepted after revision 18 June 2014
Available online 24 July 2014

Handled by Michel Laurin

Keywords:
Molecular evolution
C-value
Urodel
Mutation rates
Evolutionary duration

ABSTRACT

Salamanders have some of the largest genomes among vertebrates, and also some of the lowest reported levels of genetic diversity. Paedomorphs, in particular, have the largest genomes on average among urodela, and display exceptionally low levels of nucleotide and protein variation. Here, we address the question of genetic variation in relation to genome size in eight different salamander families. Using the *rag1* gene as a probe for evolutionary rates, we found that rates of substitution are exceptionally low in obligate paedomorphs (neotenes) and other salamander species. Substitution rates in some cases are as low as those reported for cartilaginous fish, which have the slowest mutation rates recorded so far in vertebrates. Confirming and extending an earlier study, we also found that genome size is correlated with phylogenetic age in Plethodontidae, indicating a more general trend in genome size evolution in urodela. The Plethodontidae, furthermore, display much higher levels of genetic variance than the obligate neotene families, consistent with greater habitat heterogeneity in terrestrial salamanders. Finally, we present the first direct evidence of a gene, *rag1*, whose substitution rate is negatively associated with genome size. Based on these and other observations, we propose a hypothesis according to which mutation rates in nuclear genes tend to increase as genome size decreases during the course of vertebrate evolution.

Mots clés :
Évolution moléculaire
Valeur de *C*
Urodèle
Vitesse de mutation
Durée d'évolution

RÉSUMÉ

Les salamandres ont les génomes parmi les plus grands au sein des Vertébrés, et aussi les niveaux de diversité génétique les plus bas. Les pédomorphes, en particulier, ont en moyenne les plus grands génomes parmi les urodèles et révèlent des niveaux exceptionnellement bas de variation de nucléotides et de protéines. Dans cet article, la question de la variation génétique en relation avec la taille du génome est examinée dans huit différentes familles de salamandres. En utilisant le gène *rag 1* comme indicateur de vitesse d'évolution, les auteurs trouvent que les taux de substitution semblent exceptionnellement bas chez les pédomorphes (néotènes) obligatoires et chez d'autres espèces de salamandres. Dans

* Corresponding author at: 3, rue des Jeûneurs, 75002 Paris, France.
** Corresponding author.
E-mail addresses: jhenryherrick@yahoo.fr (J. Herrick), sclavi@lbpa.ens-cachan.fr (B. Sclavi).

http://dx.doi.org/10.1016/j.crpv.2014.06.002
1631-0683/

certains cas, les taux de substitution sont aussi bas que ceux répertoriés chez les poissons cartilagineux, qui ont les vitesses de mutation les plus lentes enregistrées jusqu'à présent chez les Vertébrés. En conformité avec une étude antérieure, les auteurs observent que la taille du génome est corrélée avec l'âge phylogénétique chez les Plethodontidae, ce qui indique une tendance plus générale de l'évolution en taille du génome chez les urodèles. En outre, les Plethodontidae montrent des niveaux beaucoup plus élevés de variance génétique que les familles à néoténie obligatoire, ce qui est compatible avec une plus grande hétérogénéité de l'habitat chez les salamandres terrestres. Finalement, les auteurs fournissent la première preuve directe que le taux de substitution d'un gène, *rag 1*, est négativement associé à la taille du génome. Sur la base de ces observations et d'autres, une hypothèse est proposée, selon laquelle la vitesse de mutation dans les gènes nucléaires tend à augmenter lorsque la taille du génome décroît au cours de l'évolution chez les Vertébrés.

1. Introduction

Genome size in vertebrates varies over 200-fold (Gregory, 2001; Olmo, 2006). Salamander genomes, which are some of the largest among vertebrate lineages, range in size from 10 to over 100 picograms (pg) (Sessions, 2008). The mechanisms underlying genome size variation and stability during the course of evolution remain largely unknown (Dufresne and Jeffrey, 2011), but size variation is closely associated with the amount of non-coding DNA in the genome (Gregory, 2005; Metcalfe and Casane, 2014). A substantial proportion of non-coding nuclear DNA in the salamander genome, for example, is composed of transposable elements (TE) (Sun et al., 2012b), which have a major impact on both genome stability and genome size (Feschotte and Pritham, 2007; Izsvak et al., 2009; Kapitonov and Jurka, 2005; Metcalfe and Casane, 2014).

A number of physiological features are associated with genome size, including nuclear volume, cell size and the length of the DNA synthetic, or S-phase, of the cell cycle (Cavalier-Smith, 1978; Francis et al., 2008; Hardie and Hebert, 2003). Consequently, the duration of the cell cycle frequently varies with genome size, because larger genomes in general require more time to replicate (Francis et al., 2008; Hardie and Hebert, 2003). At the organismal level, genome size influences a number of life-history traits including generation time (GT) and developmental time (Bromham, 2011; Kozłowski et al., 2003; Krieger and Fuerst, 2002; Sessions and Larson, 1987). In Plethodontid salamanders, embryonic developmental time is directly correlated with genome size (Jockusch, 1997).

C-values (genome size) have also been associated with extinction rates. Studies in plants, fish and animals have revealed a genome size correlation with species richness (Knight et al., 2005; Olmo, 2006). In angiosperm families, a negative correlation exists between mean genome size and first appearance in geological time (Vinogradov, 2003). Hence, less speciose plant families of more recent evolutionary origin have larger genomes on average. This has been interpreted as indicating that the proliferation of TEs in plant species with large genomes is maladaptive (Knight et al., 2005; Vinogradov, 2003).

In contrast, large genomes do not appear to be associated with higher rates of extinction in vertebrates (Metcalfe and Casane, 2014; Vinogradov, 2004b). Paedomorphic salamander families of older phylogenetic age, for example, exhibit larger genomes than species that have appeared more recently in the geological record (Martin and Gordon, 1995). The association between genome size and evolutionary duration suggests a constant and gradual accumulation of transposable elements in the urodel genome (Sun et al., 2012b). The passive increase in *C*-value during salamander evolution has recently been attributed to a genomic bias in favor of retaining DNA versus deleting it (Sun et al., 2012a).

The relationship between genome size evolution, genome stability and mechanisms of speciation has been the focus of growing interest. Mounting evidence suggests that adaptive radiations are frequently associated with DNA deletions and contractions in *C*-value (Kraaijeveld, 2010; Kuo and Ochman, 2009). Other studies suggest that massive bursts of transposition events have driven the diversification of a number of invertebrate and vertebrate lineages (Biémont and Vieira, 2005; Britten, 2010; Xing et al., 2006). Hence, both expansions and contractions in *C*-value accompany speciation and diversification.

More recently, a model was proposed according to which speciation events are associated with periods of rapid genome expansion followed by longer periods of slow genome contraction and evolutionary stasis (Wolf and Koonin, 2013). Similar models involving genome expansions and contractions have been proposed to explain genome evolution (Bennetzen et al., 2005; Petrov, 2002; Vinogradov, 2004a). The underlying molecular mechanisms involved in genome dynamics, however, remain poorly understood, although some evidence implicates DNA repair pathways in the evolution, of intron density (Farlow et al., 2011).

Currently, genome size and mutation rates are believed to be biologically independent of each other, and to have co-evolved as consequences of the balance between the forces of genetic drift and natural selection (Lynch and Conery, 2003; Sung et al., 2012). Low mutation rates are required for the evolution of large genomes (Hurst, 1995), but the influence of genome size on mutation rates and rates of evolution have been largely overlooked (Hinegardner and Rosen, 1972). Indeed, little or no evidence currently exists of an unambiguous association between genome size, mutation rate and evolutionary rates despite tentative findings that in salamanders the rDNA

locus is evolving eight times more slowly than in mammals (*C*-value: typically 2 to 6 pg) (Larson and Wilson, 1989). This observation and others have motivated the following investigation into the relationship between *C*-values and substitution rates in salamanders.

2. Material and methods

The nucleotide and amino acid sequences of five orthologous urodel genes (*rag1, pomc, crcx4, ncx1* and *slc8A3)* were obtained from GENBANK. Lineages were then selected according to the availability of their *C*-values in the Animal Genome Size Database (www.genomesize.com) (Gregory et al., 2007). The *rag1* and *pomc* genes nucleotide sequence datasets were the ones for which the largest number of species have an entry in both databases and were thus chosen for further analysis.

Nucleotide sequences were codon aligned with the amino acid sequence using Pal2Nal (Mikita et al., 2006) (www.bork.embl.de/pal2nal/) and CLUSTAL W in Mega5 (Tamura et al., 2011), and then refined by hand. The tree topology was obtained from (Pyron and Wiens, 2011) and used in HyPhy (http://hyphy.org/w/index.php/Main_Page) to calculate the likelihood of synonymous and non-synonymous substitutions per site in each branch. In HyPhy, we used the codon model MG94 × 3X4 with local parameters and partition frequency estimation.

Phylogenetically independent sister pairs were identified in the tree (Fig. 3; supplementary material, Table S1). Since each sister pair diverged from a common ancestor, each member of the pair has had an equal time to accumulate substitutions in their respective branches. In order to test the hypothesis that there is a correlation between genome size and mutation rates in nuclear genes, we carried out a Wilcoxon sign rank test between the numbers of substitutions per site in the larger genome versus the smaller genome. The sister pairs examined had differences in genome size of at least 4 pg, and differences in substitution above three significant figures.

In order to obtain an estimate of the absolute synonymous substitution rate, the number of substitutions per site was divided by the time of divergence of the specific branch. This value was used to compare evolutionary rates between different species in the tree. A MatLab script was used to search for the mean divergence times of species pairs with known *C*-values from the TimeTree website (www.timetree.org) (Hedges et al., 2006).

Saturation tests were carried out on the *pomc* and the *rag1* genes. Saturation at synonymous sites was tested by plotting the number of transitions per site in the third codon position against sequence divergence for pairs of taxa. Nucleotide distances and the number of transitions and transversions per site were estimated in Mega5 using the Nei-Gojobori and Kimura 2-parameter model with a gamma distribution value of 2 for each gene. In contrast to the *pomc* gene, synonymous nucleotide sites in the *rag1* gene are not saturated over the evolutionary distances considered here (supplementary material, Fig. 1S).

Eight families of salamander were identified with *C*-values ranging from 10 pg (*Gyrinophilus porphyriticus*) to 119 pg (*Necturus punctatus*). We initially determined molecular evolutionary rates for six different salamander families: Plethodontidae, Amphiumidae, Ambystomatidae, Sirenidae, Proteidae and Cryptobranchidae. The Ambystomatidae and Cryptobranchidae families include metamorphic or partially metamorphic species, whereas the Sirenidae, Amphiumidae and Proeteidae are composed of entirely paedomorphic species. Rates of evolution were measured as the number of silent site substitutions per million years (dS/Mya) since the time of divergence (molecular evolutionary rates). Values of dN were frequently either too low or equal to zero to be included in the analyses.

3. Results

3.1. Genome size variation in salamanders

Earlier studies in plants, and animals revealed associations between genome size, extinction rates and species richness (Knight et al., 2005; Kraaijeveld, 2010; Olmo, 2006; Vinogradov, 2004b). The association between genome size and species richness becomes especially apparent in groups with genome sizes larger than 5 pg in amniotes and 14 pg in plants, suggesting that large genomes above these sizes are maladaptive (Knight et al., 2005; Vinogradov, 2003). Among vertebrates, salamander genome sizes are exceeded only by lungfish. Salamanders and lungfish, however, appear to be an exception to the hypothesis that large genomes increase the risk of extinction (Metcalfe and Casane, 2014).

Here, we examined genome size variation in direct developing, metamorphic and paedomorphic salamanders. Fig. 1A shows the distribution of genome sizes for each salamander family examined, indicating that each family has a restricted and characteristic range of genome sizes. Consistent with earlier reports, families comprising either obligate paedomorphic or partially metamorphic salamanders (Cryptobranchids, Sirenids, Amphiumids, Proteids) typically have larger genomes than other non-obligate neotenes and terrestrial salamanders in the Plethodontidae family. With the exception of the Proteidae, families including paedomorphic and partially metamorphic species in general display a relatively smaller variation in genome size that decreases with increasing *C*-value (Fig. 1A; supplementary material, Fig. 2S).

Fig. 1B reveals that the Plethodontidae, in contrast, display a large variance in genome size compared to other salamander families. Genome size in Plethodontidae varies 5X, from an average of 14 pg in *Desmognathus* to an average of almost 60 pg in the *Hydromantes* lineage. The genus *Plethodon*, which are among the most species-rich of salamander genera, display the largest variation in genome size among Plethodontidae. Similarly, the *Bolitoglossa* genus, which is also species-rich, displays a wider range of genome size than other less speciose families. The larger *C*-value variance in the direct developing *Plethodon* and *Bolitoglossa* supports the niche-width hypothesis, according to which genetic variance increases in lineages with greater habitat heterogeneity. The results are summarized in Table 1.

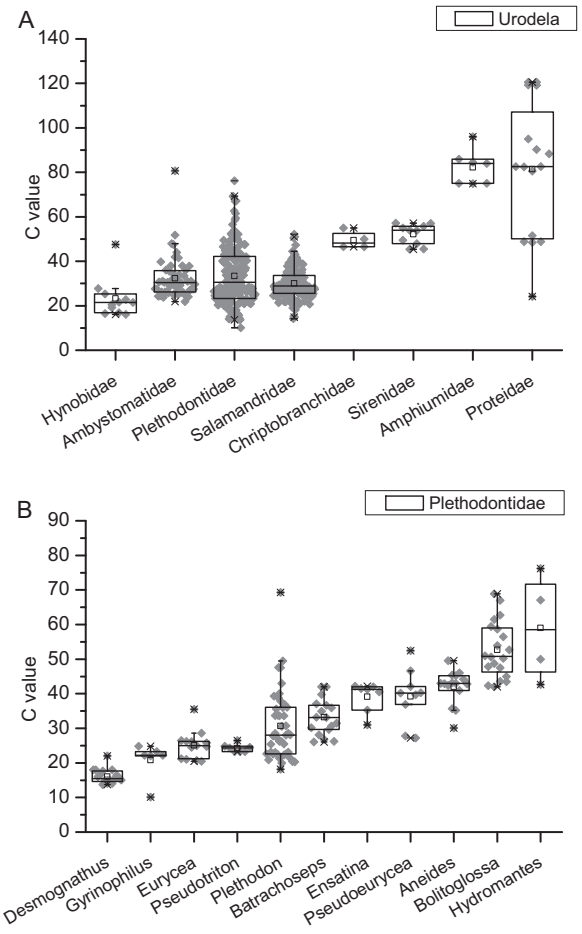

Fig. 1. *C*-value distributions for the different urodel families (A) and for the Plethodontidae (B). All urodels in the Animal Genome Size Database (Gregory et al., 2007) were grouped according to their respective families. The median is shown by a horizontal line, the mean by a square, the box ranges from the 25th to the 75th percentile of the data the whiskers correspond to the outliers with an Interquartile range (IQR) of 1.5.
Fig. 1. Distributions des valeurs de *C* pour diverses familles d'urodèles (A) et pour les Plethodontidae (B). Dans la base de données relatives à la taille du génome chez les animaux (Gregory et al., 2007), tous les urodèles ont été regroupés en fonction de leurs familles respectives. La médiane est représentée par une ligne horizontale, la moyenne par un carré, la boîte s'étend du 25e au 75e percentile des données ; les trichites correspondent aux points avec un IQR de 1,5.

3.2. Smaller genome sizes in urodela of more recent evolutionary origin

Fig. 2 shows that the average genome size across salamanders increases linearly with the phylogenetic age (evolutionary duration) of the respective family. The time of origin was obtained from Marjanović and Laurin, 2007. An earlier report found that DNA content per nucleus increased with the geological duration of neoteny in obligate paedomorphs (Martin and Gordon, 1995). This trend is reproduced here and extended to include the Plethodontidae (facultative paedomorphs, metamorphic and direct developing salamanders). *Pleurodeles* and Desmognathinae, for example, are of the most recent origin among the salamanders examined here, and they exhibit correspondingly smaller genome sizes.

Interestingly, *Eurycea*, comprising species with some of the smallest genome sizes among paedomorphs, is also of more recent evolutionary origin. Previously, it was suggested that neotene salamanders have large genomes because of their fluctuating aquatic environments, which are expected to result in frequent genetic bottlenecks and small effective population sizes (Larson, 1981; Parker and Kreitman, 1982; Shaffer and Breden, 1989). The smaller *Eurycea* genomes, however, suggest that life-history, although important, is not predominantly responsible for the evolution of large genomes in paedomorphic salamanders. Preliminary evidence also suggests that the *Eurycea*, which include facultative paedomorphs, have higher amounts of genetic variation and significantly faster rates of substitution than do most other paedomorphs with larger genomes (data not shown; Table 1).

These observations indicate that either genome sizes have expanded at a constant rate since the origin of a family (Martin and Gordon, 1995), thus resulting in larger genomes in older families; or, conversely, the appearance of younger families in the geological record have coincided with decreases in genome size (Kraaijeveld, 2010); or both processes might be impacting the mode and tempo of genome size evolution (Petrov, 2002; Petrov et al., 2000). We conclude that smaller genome sizes are associated with families of more recent evolutionary divergence across all salamander taxa examined so far, and propose that similar trends between genome size and phylogenetic age might apply to other vertebrate lineages.

3.3. Low but heterogeneous substitution rates in the rag1 gene

In order to determine whether there might be a correlation between the difference in genome size and the difference in nucleotide substitution rate, we chose an approach employed by Duchene and Bromham (Duchene and Bromham, 2013), which involves identifying phylogenetically independent sister pairs that share a common ancestor. The number of synonymous and non-synonymous substitutions is then determined from the respective branch lengths leading to each species in the sister pair. Because each species in the pair had exactly the same amount of time to accumulate the substitutions, the difference between branch lengths can be used to determine relative rates of substitution since the two species diverged. This approach has the advantages that phylogenetic non-independence is taken into account and divergence times are not required to assess differences in evolutionary rates.

We found that of the six genes available for this type of phylogenetic analysis only two, *rag1* and *pomc*, were adequately represented in the relevant databases (including both gene sequence and genome size). The other genes examined did not yield a statistically reliable number of sister pairs for species with known *C*-values. Checking for phylogenetic saturation, we found that the *pomc* gene is saturated in this data set while *rag1* is not (supplementary

Table 1
Statistics of the distribution in *C*-values and substitution rate (dS/Mya) in the clades from Fig. 3.
Tableau 1
Statistiques de distribution des valeurs de *C* et des taux de substitution (dS/Mya) dans les clades de la Fig. 3.

	Average *C* (pg)	Standard deviation	C variance	Average rate (dS/Mya)	Standard deviation	Rate variance
Plethodontidae	34	15	219	0.004	0.003	9.08E–06
Others	51	19	351	0.002	0.001	1.27E–06
Bolitoglossini	52	7	45	0.003	0.003	6.63E–06
Spelerpini	24	3	10	0.007	0.004	1.45E–05
W. Plethodons	43	14	206	0.003	0.001	7.05E–07
E. Plethodons	27	5	26	0.003	0.002	4.73E–06
Desmognathus and Aneides	27	13	177	0.004	0.001	7.78E–07

pg: picograms; dS: branch lengths; Mya: million years ago.

material, Fig. 2S). Hence, *rag1*, which is a very slowly evolving gene compared to the others, constituted the most reliable candidate gene to probe and assess the effect of genome size on substitution rate in the different sister pairs.

Using a Wilcoxon rank sign test to compare branch lengths (dS) between the larger and smaller genomes in the sister pairs (Fig. 3), we found a significant association between genome size and synonymous substitution rates in *rag1* (sample size: 13; W: 21; *P*-value: 0.032), suggesting that in larger genomes the *rag1* gene tends to have slower rates of synonymous substitution. The small number of sister pairs with a dN that was greater than zero (or where the difference was greater than two significant figures) did not allow us to measure the significance of a non-synonymous substitution effect.

A similar analysis was performed on the concatenated set of genes; but no significant correlation was found, suggesting that different genes have different sensitivities to genome size. This result is expected given that mutation/substitution rates vary considerably over the genome, and depend on chromosome context and chromatin organization, e.g., heterochromatin versus euchromatin (Herrick, 2011; Schuster-Bockler and Lehner, 2012). To our knowledge, this observation represents the first evidence of a gene that exhibits some mutational sensitivity to genome size.

3.4. Average C-value and rate of evolution are independent at the genus level in the family Plethodontidae

We next examined the relationship between average branch length and genome size at the genus level within Plethodontidae. Fig. 3 reveals phylogenetic relationships between species and genera belonging to the

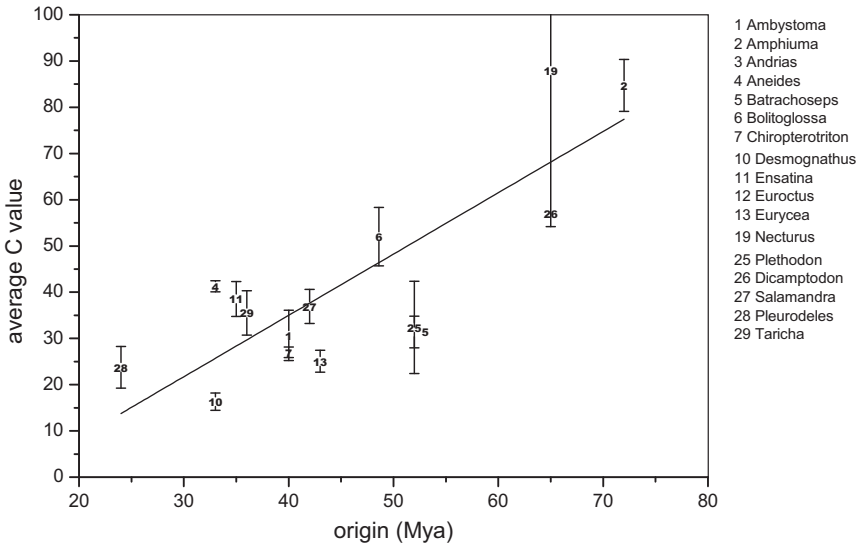

Fig. 2. Negative correlation between average *C*-value and evolutionary duration (phylogenetic age). The Animal Genome Size Database (Gregory et al., 2007) was searched for those families with published time of origin (data from Marjanović and Laurin, 2007). The slope is approximately 1 picogram/million years ago.
Fig. 2. Corrélation négative entre la valeur-*C* moyenne et la durée d'évolution (âge phylogénétique). La base de données relatives à la taille du génome chez les animaux (Gregory et al., 2007) a été recherchée pour les familles dont la date d'origine est publiée (âges d'après Marjanović et Laurin, 2007). La pente est approximativement de 1 pg/Mya.

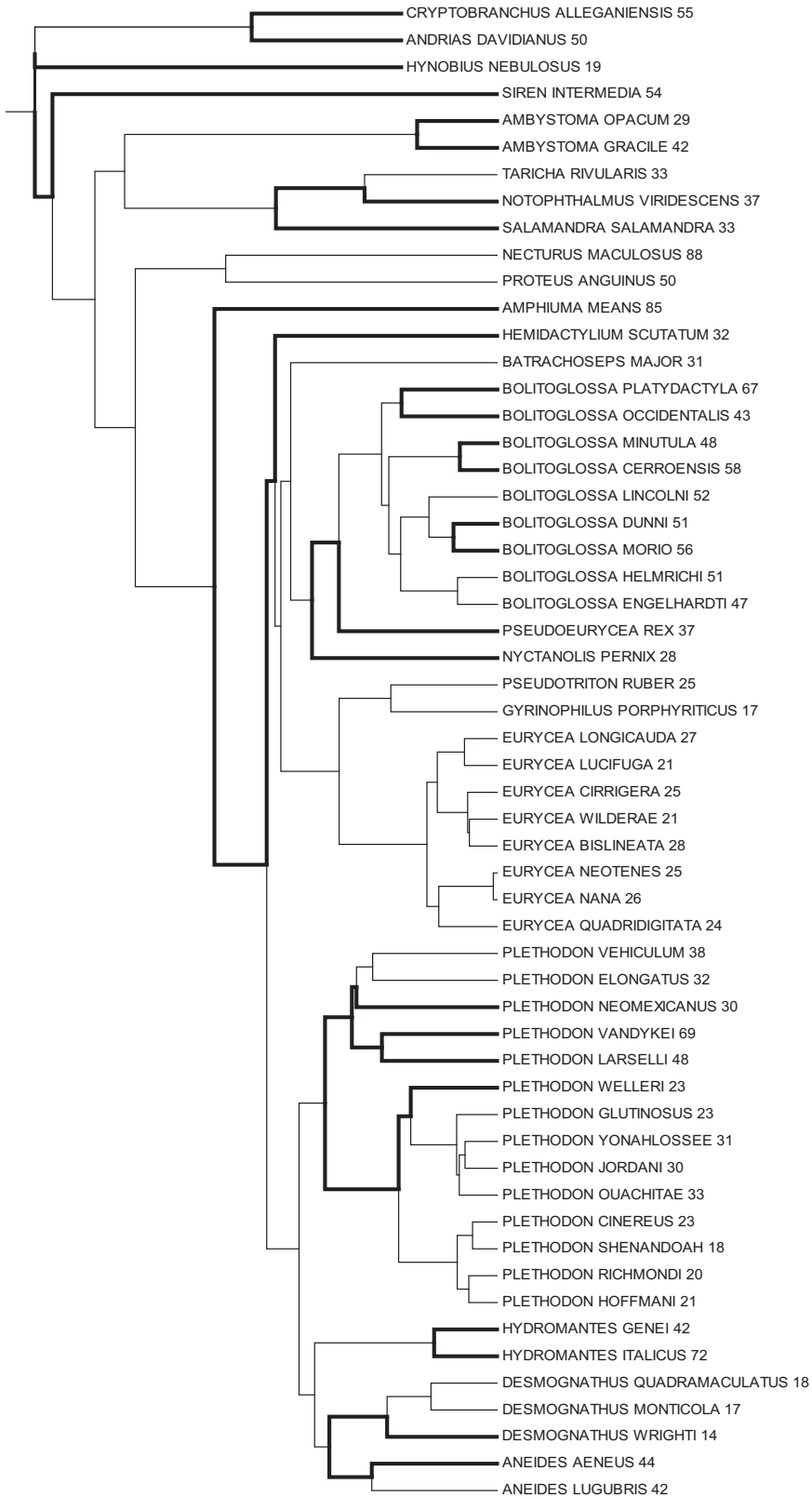

Fig. 3. Phylogenetic tree of the eight salamander families. All species for which the *rag1* sequence is available were chosen according to their presence in the Animal Genome Size database. The average recorded *C*-values in picograms are shown next to the species name. Sister pairs analyzed appear in bold. The branch lengths are proportional to the number of substitutions per site. The tree topology was obtained from Pyron and Wiens, 2011.

major tribes that comprise the Plethodontidae minus the tribe Ensatinini (Vieites et al., 2011). We compared four clades corresponding to the genera *Bolitoglossa*, *Eurycea*, *Plethodon*, and *Hydromantes*, *Desmognathus* and *Aneides*. The *Plethodon* are further divided into two clades corresponding to Eastern and Western *Plethodon*. For each clade, we determined the average *C*-value and the average branch length (dS) leading to each species from the root node of the respective clade (Table 1).

The results of this analysis revealed no association between average *C*-value and average branch length among the Plethodontidae, indicating that *C*-value and rate of evolution are independent, on average, in this family of salamander (Table 1). At the species level, in contrast, sister pairs in the genera *Bolitoglossa* and *Hydromantes*, both of which are composed of species with large genomes (> 40 pg), exhibit an unambiguous negative association between genome size and rates of evolution (Fig. 3; Table 1; supplementary material, Fig. 3S).

Similarly, sister pair analysis between the Western and Eastern *Plethodon* reveals that species with larger genomes in the Western clade tend to have slower substitution rates when paired with species from the Eastern clade (supplementary material, Fig. 4S). Together, these observations suggest that either substitution rates tend to decrease as genome size increases within a sister pair; or, conversely, they increase with decreasing genome size.

3.5. *Very slow rates of evolution in some species of the Plethodontidae and obligate neotene salamander families*

The phylogenetic tree based on the species examined here confirms a deep split between the Plethodontidae and the older urodela families, which diverged from each other about 130 million years ago (Mya) (Marjanović and Laurin, 2007, 2014; Mueller, 2006). A recent hypothesis proposes that the dominant mode of evolution involves two regimes:

- bursts of innovation are accompanied by an expansion in *C*-values;
- more prolonged periods of evolutionary stasis, which follow periods of innovation, are accompanied by a clock-like contraction in genome size (Wolf and Koonin, 2013).

Genera of Plethodontidae appear in the geological record about 38 Mya (Elmer et al., 2013; Marjanović and Laurin 2014; Martin and Gordon, 1995), and have undergone a number of well documented adaptive radiations, for example the genera *Plethodon* in North America and *Bolitoglossa* in Central and South America (Elmer et al., 2013; Vieites et al., 2007). In contrast, the obligate neotene families we examined are of older evolutionary origin (80 to 150 Mya) (Marjanović and Laurin, 2007, 2014; Martin and Gordon, 1995), and appear to have undergone prolonged periods of evolutionary stasis (Gao and Shubin, 2003). This suggests that the Plethodontidae have undergone a mode of speciation and genome size evolution that is distinct from the other salamander families.

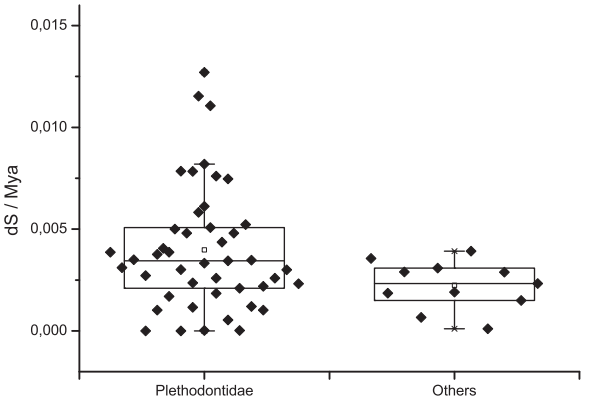

Fig. 4. Distribution of synonymous substitution rates (dS/Mya) in Plethodontidae compared to other salamander families (paedomorphs). The median is shown by a horizontal line, the mean by a square, the box ranges from the 25th to the 75th percentile of the data the whiskers correspond to the outliers with an IQR of 1.5.
Fig. 4. Distribution des taux de substitution synonymes (dS/Mya) chez les Plethodontidae, comparés à ceux des autres familles de salamandres (paedomorphes). La médiane est représentée par une ligne horizontale, la moyenne par un carré ; la boîte s'étend du 25e au 75e percentile des données ; les trichites correspondent aux points avec un IQR de 1,5.

To assess any differences in the mode of evolution between the Plethodontidae and the obligate/facultative neotene families, we measured evolutionary rates as determined by dS/Mya, and found that salamander species include some of the slowest evolving vertebrates (supplementary material, Fig. 5S). Interestingly, no significant difference in the average evolutionary rate was found between the Plethodontidae and the older obligate neotenes (Fig. 4), suggesting that urodela might be approaching a minimum permissible rate of evolution in vertebrates (Sung et al., 2012). Evolutionary rates of the Plethodontidae, for example, "bottom out" at a dS/Mya of 0.001—the lowest observed rates in this study (Fig. 4).

Remarkably, the Plethodontidae exhibit a larger variance in their rates of evolution, on the order of ten-fold compared to less than two-fold found in the aquatic paedomorphs (Table 1). This observation is consistent with the wide variation in genome sizes in Plethodontidae (5X; Fig. 1B). Hence, substitution rates and genome sizes in obligate neotene families appear to be more evolutionarily conserved than in the Plethodontidae. Together, these observations suggest that within-family variations in genome size are closely associated with variations in rates of substitution and evolution, and thus salamander families with higher levels of genome size diversity tend to have a wider range of evolutionary rates.

Fig. 3. Arbre phylogénétique pour les huit familles de salamandres. Toutes les espèces pour lesquelles la séquence de *rag 1* est disponible ont été choisies dans la base de données relatives à la taille du génome chez les animaux. Les valeurs moyennes de *C* en picogrammes sont présentées à côté du nom de l'espèce. Les paires analysées apparaissent en gras. Les longueurs des branches sont proportionnelles au nombre de substitutions par site. L'arbre est issu de Pyron et Wiens (2011).

4. Discussion

4.1. Conclusion

Here, we have investigated the relationship between genome size, genetic variation and evolutionary duration. Our principal findings concern:

- the positive correlation between genome size and phylogenetic age is now extended to genera inside the family Plethodontidae;
- some, but not all, of the salamander taxa examined here exhibit exceptionally low rates of evolution in the *rag1* gene.

Other studies have also provided tentative evidence that large genomes in both plants and animals are associated with lower levels of genetic variation and rates of evolution (Buschiazzo et al., 2012; Pierce and Mitton, 1980). In plants, for example, gymnosperms, which have significantly larger genomes than angiosperms, are evolving at much slower rates (Buschiazzo et al., 2012). Population genetic effects independent of cell physiology and other molecular traits, however, can apparently account for many of these observations (Larson, 1981; Shaffer and Breden, 1989).

The results obtained from this study also revealed a weak but significant negative association between genome size and phylogenetic branch lengths in salamanders: shorter branch lengths in sister pairs tend to be associated with the species having the larger genome. Our analysis, moreover, provided preliminary evidence that changes in genome size between sister pairs are positively associated with differences in substitution rates: to the extent which genomes in a sister pair diverge from each other in size, they also tend to diverge in their respective rates of substitution (supplementary material, Fig. 5S).

Our findings, however, appear to be restricted to sister pairs alone, since a comparison between clade averaged rates of substitution in *rag1* did not reveal an association between genome size and substitution rate (Table 1). Indeed, we found that despite widely varying life-history traits, different salamander families are on average evolving at similar rates (Fig. 4; Table 1), a finding that contrasts with the comparatively lower levels of heterozygosity reported in obligate paedomorphs (Larson, 1981; Parker and Kreitman, 1982; Pierce and Mitton, 1980; Shaffer and Breden, 1989). An overall trend nevertheless emerges from the sister pair analysis presented here: rates of evolution (dS/Mya) tend to decrease as genome size increases (supplementary material, Fig. 6S). Additional studies, however, are needed to confirm this observation in urodels and to extend it to other amphibians and vertebrates.

4.2. Heterogeneous evolutionary rates in the Plethodontidae

Our analysis revealed a number of interesting findings concerning the relationship between genome size and substitution rates, notably among the Plethodontidae. We found, for example, that the clade including members of the genus *Bolitoglossa*, which have among the largest genomes within the Plethodontidae, are among the fastest evolving salamander species in this family. *Bolitoglossa* species with the larger genomes in sister pairs were found, however, to be evolving more slowly, suggesting that slower rates of substitution at the species level are associated with larger genomes within this salamander genus (supplementary material, Fig. 3S). The species *B. platydactyla* is a notable exception, and is evolving faster than the other *Bolitoglossa* species examined here, although it has the largest genome (supplementary material, Fig. 3S).

The clade including species belonging to the genus *Eurycea* revealed a similarly complex relationship between genome size and substitution rate in the Plethodontidae. These species all have similar genome sizes and are evolving at similar rates (data not shown). The species *G. porphyriticus* and *Pseudotriton ruber*, however, are evolving more slowly compared to the *Eurycea* congeners (not shown). When sister pairs are averaged over all species of the clade comprising *Eurycea*, we found that genome size and branch length nevertheless tend to be negatively associated within this genus (supplementary material, Fig. 3S).

An examination of the genus *Plethodon* revealed that, independently of genome size, these species are evolving more slowly than species in either the *Eurycea* or *Bolitoglossa* genera (supplementary material, Fig. 3S). We also found that species in the clade comprising the Western *Plethodon* (average *C*-value: 43.4 pg) are evolving more slowly than the clade comprising the Eastern *Plethodon* (average *C*-value: 24.6 pg), again supporting a potential negative association between genome size and substitution rates (supplementary material, Fig. 4S). The sister pair *P. vandykei* (69 pg) and *P. larselli* (48 pg), however, represent another exception to the general trend, and underscores the importance of other lineage-specific influences on substitution rates.

Finally, the clade including the genera *Aneides* (42–44 pg), *Desmognathus* (14–18 pg) and *Hydromantes* (42–72 pg) likewise reveals a highly heterogeneous relationship between genome size and *rag1* evolutionary rates in Plethodontidae (supplementary material, Fig. 3S). In this clade, *H. genei* and *H. italicus* are the fastest evolving species while the Desmognathinae, which have the smallest genomes, are the slowest evolving. The *Hydromantes* sister pair, in contrast, exhibits a clear negative association between genome size and substitution rate in *rag1* that is not found in either the *Desmognathus* or *Aneides* sister pairs. Indeed, sister pairs within the latter two genera do not display as large a difference in genome size compared to the *Hydromantes* sister pair, suggesting that differences in genome size must exceed a certain threshold before an unambiguous association between *C*-value and substitution rate can be detected (Fig. 3).

4.3. Genome size and phylogenetic age

Earlier studies reported an apparent correlation between nuclear DNA content and evolutionary duration (phylogenetic age) in salamanders (Fig. 2) (Martin and Gordon, 1995). The authors interpreted the trend as evidence that genome size increases depending on how long

a species has been an obligate neotene, and proposed that the rate of junk DNA accumulation could be used as a possible second molecular clock (accumulating 0.63 pg/Mya). This interpretation is consistent with observations that the ancestral urodel genome was much smaller and comparable in size to extant mammalian genomes (approximately 3 pg), indicating a massive expansion in salamander *C*-values since their time of origin (Organ et al., 2011).

Alternatively, the trend reproduced and extended here might suggest that, subsequent to long periods of expansion, reductions in genome size accompany evolution and speciation in salamanders (Kraaijeveld, 2010). We note, for example, the higher levels of diversity in evolutionary rates in the Plethodontidae (supplementary material, Figs. 3S and 5S), which is consistent with the corresponding tribes and genera radiating more recently into more heterogeneous terrestrial niches (Nevo and Beiles, 1991). Additionally, the Plethodontidae have the widest and most diverse range of genome sizes among salamander families, and include the smallest existing urodela genomes (Fig. 1B). Based on these and other observations, we propose that elevated levels of molecular diversity, both in terms of genome size and mutation rates, tend to be associated with smaller genomes and more recent adaptive radiations.

4.4. Hypothesis and final remarks

Due to a variety of evolutionary constraints, organisms with large genomes appear to be under selective pressure to evolve more efficient DNA replication and repair processes that minimize mutation rates to a level set by genetic drift (Massey, 2008; Sung et al., 2013). How organisms achieve an optimal balance between mutation (predominantly DNA replication errors) and DNA repair rates is unclear at the molecular level, but compartmentalizing mutation rates within the genome between early and late replicating DNA provides one plausible, though not complete, explanation [for a more detailed discussion see (Herrick, 2011)] (Wintersberger, 2000).

Based on the considerations discussed above and preliminary observations on the effect of genome size on nuclear substitution rates, we propose a hypothesis according to which speciation events in salamanders are associated with contractions in genome size that are indirectly yet mechanistically related to increases in mutation and substitution rates in nuclear genes (Herrick, 2011). Accordingly, genome reduction (expansion) and reorganization during speciation might result in changes in the DNA replication and repair programs that impact rates of mutation and evolution. Eukaryotes with larger genomes, for example, rely more heavily on error prone DNA repair pathways, which operate preferentially in late S and G2/M phases, compared to other eukaryotes with smaller genomes (Farlow et al., 2011).

A corollary to that proposal implies that species with larger genomes should display correspondingly higher mutation and divergence rates in the latest replicating sequences (such as microsatellite DNA and TEs) compared to species with smaller genomes. Conversely, sequences replicating earliest in S-phase (such as housekeeping genes) should display correspondingly lower mutation rates in species with larger genomes compared to those with smaller genomes (Herrick, 2011). Some evidence suggests that this indeed is the case (Chen et al., 2010; Lang and Murray, 2011; Stamatoyannopoulos et al., 2009; Weber et al., 2012). TE rich regions, tend to replicate late in the vertebrate S-phase, and therefore they experience higher rates of mutation and genetic erosion (Chen et al., 2010; Stamatoyannopoulos et al., 2009; Weber et al., 2012).

Higher rates of mutational erosion and genetic extinction can explain why in lungfish and salamanders, which both have exceptionally low levels of genetic variation, TE content appears to be under-represented in the respective species' genomes compared to other species with lower *C*-values (Metcalfe and Casane, 2014). The lower than expected proportion of TEs in large genomes can be explained by their inactivation and decay over time (Metcalfe and Casane, 2014) and/or through a genome size dependent reliance on error prone DNA repair pathways operating in late replicating DNA (Diamant et al., 2012; Mao et al., 2008). Consequently, a large proportion of the lungfish and salamander genomes are expected to be comprised of late replicating "fossilized" TEs that have experienced correspondingly higher rates of mutation and substitution. Whether or not mutation rates between early and late replicating DNA are anti-correlated in a genome size dependent manner remains, however, to be demonstrated.

The findings presented here suggest that the relationship between mutation rate, rates of evolution and genome size in vertebrates and other eukaryotes warrants further investigation. Genome size, is a proxy variable for a number of different physiological, molecular and life-history traits such as cell cycle duration, metabolic rate and developmental time. It will be interesting to investigate how these variables interact with each other dynamically to modify and shape the overall architecture of the genome in different species during the course of evolution.

Acknowledgements

The authors would like to thank Michel Laurin and four anonymous reviewers for their helpful comments on the manuscript. The authors would also like to thank Alex Pyron and Peter Midford for additional assistance with the sequence datasets and tree topology data. BS is supported by a grant from Human Frontier Science Program (RGY0079). JH benefited from support from John Bechhoefer's lab, Physics Department, Simon Fraser University.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.crpv.2014.06.002.

References

Bennetzen, J.L., Ma, J., Devos, K.M., 2005. Mechanisms of recent genome size variation in flowering plants. Ann. Bot. 95, 127–132.

Biémont, C., Vieira, C., 2005. What transposable elements tell us about genome organization and evolution: the case of Drosophila. Cytogenet. Genome Res. 110, 25–34.

Britten, R.J., 2010. Transposable element insertions have strongly affected human evolution. Proc. Natl. Acad. Sci. U S A 107, 19945–19948.
Bromham, L., 2011. The genome as a life-history character: why rate of molecular evolution varies between mammal species. Philos. Trans. R. Soc. London, Ser. B, Biol. Sci. 366, 2503–2513.
Buschiazzo, E., Ritland, C., Bohlmann, J., Ritland, K., 2012. Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms. BMC Evol. Biol. 12, 8.
Cavalier-Smith, T., 1978. Nuclear volume control by nucleoskeletal DNA, selection for cell volume and cell growth rate, and the solution of the DNA *C*-value paradox. J. Cell. Sci. 34, 247–278.
Chen, C.L., Rappailles, A., Duquenne, L., Huvet, M., Guilbaud, G., Farinelli, L., Audit, B., d'Aubenton-Carafa, Y., Arneodo, A., Hyrien, O., Thermes, C., 2010. Impact of replication timing on non-CpG and CpG substitution rates in mammalian genomes. Genome Res. 20, 447–457.
Diamant, N., Hendel, A., Vered, I., Carell, T., Reissner, T., de Wind, N., Geacinov, N., Livneh, Z., 2012. DNA damage bypass operates in the S and G2 phases of the cell cycle and exhibits differential mutagenicity. Nucleic Acids Res. 40, 170–180.
Duchene, D., Bromham, L., 2013. Rates of molecular evolution and diversification in plants: chloroplast substitution rates correlate with species richness in the Proteaceae. BMC Evol. Biol. 13, 65.
Dufresne, F., Jeffrey, N., 2011. A guided tour of large genome size in animals: what we know and where we are heading. Chromosome Res. 19, 925–938.
Elmer, K.R., Bonett, R.M., Wake, D.B., Lougheed, S.C., 2013. Early Miocene origin and cryptic diversification of South American salamanders. BMC Evol. Biol. 13, 59.
Farlow, A., Meduri, E., Schlötterer, C., 2011. DNA double-strand break repair and the evolution of intron density. Trends Genet. 27, 1–6.
Feschotte, C., Pritham, E.J., 2007. DNA transposons and the evolution of eukaryotic genomes. Annu. Rev. Genet. 41, 331–368.
Francis, D., Davies, M.S., Barlow, P.W., 2008. A strong nucleotypic effect on the cell cycle regardless of ploidy level. Ann. Bot. 101, 747–757.
Gao, K.Q., Shubin, N.H., 2003. Earliest known crown-group salamanders. Nature 422, 424–428.
Gregory, T.R., 2001. Coincidence, coevolution, or causation? DNA content, cell size, and the *C*-value enigma. Biol. Rev. Cam. Philo. Soc. 76, 65–101.
Gregory, T.R., 2005. Synergy between sequence and size in large-scale genomics. Nat. Rev. Genet. 6, 699–708.
Gregory, T.R., Nicol, J.A., Tamm, H., Kullman, B., Kullman, K., Leitch, I.J., Murray, B.G., Kapraun, D.F., Greilhuber, J., Bennett, M.D., 2007. Eukaryotic genome size databases. Nucleic Acids Res. 35, D332–D338.
Hardie, D.C., Hebert, P.D., 2003. The nucleotypic effects of cellular DNA content in cartilaginous and ray-finned fishes. Genome 46, 683–706.
Hedges, S.B., Dudley, J., Kumar, S., 2006. TimeTree: a public knowledge-base of divergence times among organisms. Bioinformatics 22, 2971–2972.
Herrick, J., 2011. Genetic variation and replication timing, or why is there late replicating DNA? Evolution 65 (11), 3031–3047.
Hinegardner, R., Rosen, D.E., 1972. Cellular content and the evolution of teleostean fishes. Am. Nat. 106, 621–644.
Hurst, L.D., 1995. Evolutionary genetics. The silence of the genes. Curr. Biol. 5, 459–461.
Izsvak, Z., Wang, Y., Ivics, Z., 2009. Interactions of transposons with the cellular DNA repair machinery. In: Lankenau, D.-H., Volff, J.-N. (Eds.), Genome Dynamics and Stability. Springer, Berlin Heidelberg, pp. 133–176.
Jockusch, E., 1997. An evolutionary correlate of genome size change in plethodontid salamanders. Proc. R. Soc. Land. B 264, 597–604.
Kapitonov, V.V., Jurka, J., 2005. RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. PLoS Biol. 3, e181.
Knight, C.A., Molinari, N.A., Petrov, D.A., 2005. The large genome constraint hypothesis: evolution, ecology and phenotype. Ann. Bot. 95, 177–190.
Kozłowski, J., Konarzewski, M., Gawelczyk, A.T., 2003. Cell size as a link between non-coding DNA and metabolic rate scaling. Proc. Natl. Acad. Sci. U S A 100, 14080–14085.
Kraaijeveld, K., 2010. Genome size and species diversification. Evol. Biol. 37, 227–233.
Krieger, J., Fuerst, P.A., 2002. Evidence for a slowed rate of molecular evolution in the order acipenseriformes. Mol. Biol. Evol. 19, 891–897.
Kuo, C.H., Ochman, H., 2009. Deletional bias across the three domains of life. Genome Biol. Evol. 1, 145–152.
Lang, G.I., Murray, A.W., 2011. Mutation rates across budding yeast Chromosome VI are correlated with replication timing. Genome Biol. Evol. 3, 799–811.
Larson, A., 1981. A reevaluation of the relationship between genome size and genetic variation. Am. Nat. 118, 119–125.
Larson, A., Wilson, A.C., 1989. Patterns of ribosomal RNA evolution in salamanders. Mol. Biol. Evol. 6, 131–154.
Lynch, M., Conery, J.S., 2003. The origins of genome complexity. Science 302, 1401–1404.
Mao, Z., Bozzella, M., Seluanov, A., Gorbunova, V., 2008. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. Cell Cycle 7, 2902–2906.
Marjanović, D., Laurin, M., 2007. Fossils, molecules, divergence times, and the origin of lissamphibians. Syst. Biol. 56, 369–388.
Marjanović, D., Laurin, M., 2014. An updated paleontological timetree of lissamphibians with comments on the anatomy of Jurassic crown-group salamanders (Urodela). Hist. Biol. 26 (4).
Martin, C.C., Gordon, R., 1995. Differentiation trees, a junk DNA molecular clock, and the evolution of neoteny in salamanders. J. Evol. Biol. 8, 339–354.
Massey, S.E., 2008. The proteomic constraint and its role in molecular evolution. Mol. Biol. Evol. 25, 2557–2565.
Metcalfe, C.J., Casane, D., 2014. Accommodating the load: the transposable element content of very large genonmes. Mobile Genet. Elem. 3, 1–8.
Mikita, S., Torrents, D., Bork, P., 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34, W609–W612.
Mueller, R.L., 2006. Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. Syst. Biol. 55, 289–300.
Nevo, E., Beiles, A., 1991. Genetic diversity and ecological heterogeneity in amphibian evolution. Copeia 19, 565–592.
Olmo, E., 2006. Genome size and evolutionary diversification in vertebrates. Ital. J. Zool. 73, 167–171.
Organ, C.L., Canoville, A., Reisz, R.R., Laurin, M., 2011. Paleogenomic data suggest mammal-like genome size in the ancestral amniote and derived large genome size in amphibians. J. Evol. Biol. 24, 372–380.
Parker, E., Kreitman, M., 1982. On the relationship between heterozygosity and DNA content. Am. Nat. 119, 749–752.
Petrov, D.A., 2002. Mutational equilibrium model of genome size evolution. Theor. Popul. Biol. 61, 531–544.
Petrov, D.A., Sangster, T.A., Johnston, J.S., Hartl, D.L., Shaw, K.L., 2000. Evidence for DNA loss as a determinant of genome size. Science 287, 1060–1062.
Pierce, A., Mitton, B., 1980. The relationship between genome size and genetic variation. Am. Nat. 116, 850–861.
Pyron, R.A., Wiens, J.J., 2011. A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. Mol. Phylogenet. Evol. 61, 543–583.
Schuster-Bockler, B., Lehner, B., 2012. Chromatin organization is a major influence on regional mutation rates in human cancer cells. Nature 488, 504–507.
Sessions, S.K., 2008. Evolutionary cytogenetics in salamanders. Chromosome Res. 16, 183–201.
Sessions, S.K., Larson, A., 1987. Developmental correlates of genome size in plethodontid salamanders and their implications for genome size evolution. Evolution 41, 1239–1251.
Shaffer, B.H., Breden, F., 1989. The relationship between allozyme variation and life-history: non-transforming salamanders are less variable. Copeia 4, 1016–1023.
Stamatoyannopoulos, J., Adzhubei, I., Thurman, R., Kryukov, G., Mirkin, S., Sunyaev, S., 2009. Human mutation rate associated with DNA replication timing. Nat. Genet. 41, 393–395.
Sun, C., López Arriaza, J.R., Mueller, R.L., 2012a. Slow DNA loss in the gigantic genomes of salamanders. Genome Biol. Evol 4, 1340–1348.
Sun, C., Shepard, D.B., Chong, R.A., López Arriaza, J., Hall, K., Castoe, T.A., Feschotte, C., Pollock, D.D., Mueller, R.L., 2012b. LTR retrotransposons contribute to genomic gigantism in plethodontid salamanders. Genome Biol. Evol. 4, 168–183.
Sung, W., Ackerman, M.S., Miller, S.F., Doak, T.G., Lynch, M., 2012. Drift-barrier hypothesis and mutation rate evolution. Proc. Natl. Acad. Sci. U S A 109, 18488–18492.
Sung, W., Ackerman, M.S., Miller, S.F., Doak, T.G., Lynch, M., 2013. Reply to Massey: drift does influence mutation rate evolution. Proc. Natl. Acad. Sci. U S A 110, E860.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28 (10), 2731–2739.
Vieites, D.R., Min, M.S., Wake, D.B., 2007. Rapid diversification and dispersal during periods of global warming by plethodontid salamanders. Proc. Natl. Acad. Sci. U S A 104, 19903–19907.

Vieites, D.R., Román, S.N., Wake, M.H., Wake, D.B., 2011. A multigenic perspective on phylogenetic relationships in the largest family of salamanders, the Plethodontidae. Mol. Phylogenet. Evol. 59, 623–635.
Vinogradov, A.E., 2003. Selfish DNA is maladaptive: evidence from the plant Red List. Trends Genet. 19, 609–614.
Vinogradov, A.E., 2004a. Evolution of genome size: multilevel selection, mutation bias or dynamical chaos? Curr. Opin. Genet. Dev. 14, 620–626.
Vinogradov, A.E., 2004b. Genome size and extinction risk in vertebrates. Proc. Biol. Sci. 271, 1701–1705.
Weber, C.C., Pink, C.J., Hurst, L.D., 2012. Late replicating domains have higher divergence and diversity in Drosophila melanogaster. Mol. Biol. Evol. 29, 873–882.
Wintersberger, E., 2000. Why is there late replication? Chromosoma 109, 300–307.
Wolf, Y.I., Koonin, E.V., 2013. Genome reduction as the dominant mode of evolution. Bioessays 35, 829–837.
Xing, J., Wang, H., Belancio, V.P., Cordaux, R., Deininger, P.L., Batzer, M.A., 2006. Emergence of primate genes by retrotransposon-mediated sequence transduction. Proc. Natl. Acad. Sci. U S A 103, 17608–17613.