bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Genome size variation and species diversity in salamanders

Bianca Sclavi [1*] and John Herrick [2]**

Author affiliations:

1. LBPA, UMR 8113, ENS Paris-Saclay, 95235 Cachan, France
2. Department of Physics, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia VSA 1S6, Canada

* Corresponding author

** Corresponding author at 3, rue des Jeûneurs, 75002 Paris, France.

E-mail addresses: jhenryherrick@yahoo.fr (J. Herrick), sclavi@lbpa.ens-cachan.fr (B. Sclavi).

Running head: Salamander genome size and species diversity

Keywords: C-value, species diversity, body size, Urodela, geographical area

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Abstract: Salamanders (Urodela) have among the largest vertebrate genomes, ranging in size from 10 to 120 pg. Although changes in genome size often occur randomly and in the absence of selection pressure, non-random patterns of genome size variation are evident among specific vertebrate lineages. Here we report i) an unexpected relationship between genome size (C-value) and species richness at the family taxonomic level in salamander clades; ii) C-value and species richness both correlate with clade crown-age but not with diversification rates; iii) C-value is strongly associated with geographical area. Consistent with this observation we found that iv) genome size in salamanders is also strongly related to climatic niche rate. Finally, we report a relationship between C-value diversity and family-level species diversity as measured by the coefficient of variation of genome size within clades. Based on these results we propose that variation in C-value and genomic organization (heterochromatin content and gene synteny) might underlie variation in species richness in agreement with the Geographical Area Hypothesis. These observations suggest that changes in C-value, and hence changes in the amount and/or organization of non-coding DNA in the vertebrate genome, are associated with the allelic incompatibility believed to drive reproductive isolation and speciation in salamanders.

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Introduction

Genome size in vertebrates varies more than three hundred fold from 0,4 picograms (pg) in pufferfish to over 120 picograms (pg) in lungfish (Gregory 2015). Most of the variation in vertebrate genome size corresponds to differences in non-coding DNA such as transposable elements, microsatellites and other types of repetitive and intergenic DNA (Metcalfe and Casane 2013). The DNA accounting for differences in genome size between related species has previously been considered devoid of any universal function such as gene regulation, structural maintenance and protection against mutagens (Palazzo and Gregory 2014). Genome size, however, is known to have a direct impact on important physiological parameters such as cell size, cell cycle duration and developmental time (Cavalier-Smith 1978, Francis, Davies et al. 2008). C-values, which measure genome size, are exceptionally high and vary more widely in salamanders compared to most other vertebrates (Sessions 2008), making the Urodela an attractive model for studying genome size variation and its biological effects between and within taxonomically related vertebrate clades.

In addition to its impact on cell physiology, genome size frequently correlates with a number of ecological and physiological parameters including climate, metabolic rate, extinction risk and species diversity (Sessions and Larson 1987 , Gregory 2004, Mank and Avise 2006). A comparative study across vertebrates, for example, has found some evidence for a negative correlation between species diversity and genome size above the class-level of clades (Olmo 2006, Kraaijeveld 2010). The strongest association between genome size and species diversity was observed at C-values greater than 5 pg. Consistent with reduced species diversity in vertebrates with large genomes, at higher taxonomic levels extinction risk was found to increase with genome size (Vinogradov 2004). Consequently, species number might correlate negatively with the proportion of repetitive DNA (Olmo 2006), suggesting that

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

71 either evolvability (propensity to speciate) or extinction risk varies according to the amount of non-
72 coding DNA present in the genome.
73
74 The large genomes of salamanders containing disproportionate amounts of non-coding DNA therefore
75 represent a potentially important factor that is influencing species diversity across the Urodela. While
76 some Urodela clades, such as the Plethodontidae and Salamandridae, are exceptionally rich in species
77 number, other family-level clades are extremely species poor. Why species diversity varies so widely
78 among salamander clades remains unclear, though many physiological and environmental factors are
79 known to contribute to differences in Urodela species diversity (Nevo and Beiles 1991, Pyron and Wiens
80 2013). Recently, for example, it has been shown in salamanders and songbirds that species richness is
81 strongly associated with climatic niche evolution (Kozak and Wiens 2010, Kozak and Wiens 2016, Title
82 and Burns 2015, Cooney 2016).
83
84 The large Urodela genomes have been attributed to slow losses of DNA in salamanders compared to
85 other vertebrates (Sun, López Arriaza et al. 2012). The mechanisms underlying the slower loss rates
86 remain to be elucidated, but cell cycle checkpoints and DNA repair systems likely play an important role
87 in the evolution of large genomes and mutation rate variation (Herrick 2011, Sun and Mueller 2014).
88 Variations in the proportion of non-coding DNA in genomes (eg. indels), for example, correspond to *de*
89 *facto* mutations that have occurred during the course of genome evolution (Böhne, Brunet et al. 2008,
90 Jurka, Bao et al. 2012).
91
92 Consistent with low mutation rates in organisms with large genomes, earlier studies have revealed
93 exceptionally low levels of genetic diversity in salamanders (Pierce and Mitton 1980, Nevo and Beiles

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

1991), yet the sources of the low genetic variation remain unclear (Larson 1981, Parker and Kreitman 1982). Genetic variation is a determining factor in both evolvability and extinction risk (Frentiu, Ovenden et al. 2001, Vinogradov 2004, Matsui, Tominaga et al. 2008, Hughes, Schmidt et al. 2015), and is therefore potentially associated with species diversity. Hence, the low levels of genetic variation in gene-expressing regions in salamander genomes (Pierce and Mitton 1980, Karlin and Means 1994), suggests that genome size might also influence species diversity in salamanders. A potential relationship between genome size and species diversity, however, remains to be established across salamander clades. In the following, we investigate the relationship between genome size and species diversity in the Urodela.

Materials and methods

Genome sizes were obtained from the Animal Genome Size Database (Gregory 2015). C-value refers to haploid nuclear DNA content (1C). Reported polyploids, when indicated in the Animal Genome Size Database, were removed from the analyses. Average C-values were determined for each genus belonging to a given family-level clade. These values were then used to calculate the average C-value of each family-level clade. The C-value for each species in turn is determined as the average of the available values for that species when more than one C-value is recorded in the database. Each average of averages is weighted either by the number of entries in the database for a given species, or by the number of species in a genus. The weighting is independent of branch lengths.

$$Average\ C\ value = \frac{\sum_{i=1}^{n} \mu_i N_i}{\sum_{i=1}^{n} N_i}$$

To measure the average C-value for a genus of n species, each species has an average C-value μ calculated from N entries. The distributions in genome size for each salamander family and among the genera of Plethodontidae have been published previously (Herrick and Sclavi 2014).

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

The data on crown age, stem age, species diversity and geographic area were obtained from Pyron and Wiens (Pyron and Wiens 2013). Both the maximum likelihood (ML) and the time-calibrated trees used here were obtained from those of Pyron and Wiens (Pyron and Wiens 2011, Pyron 2011). They obtained the time-calibrated tree by determining divergence times from a set of fossil constraints using treePL developed by S.A. Smith (Smith and O'Meara 2012), and applied to the ML phylogeny determined previously for 2871 species using data from 3 mitochondrial and 9 nuclear genes (Pyron and Wiens 2011). They determined species diversity from the assignment of all known amphibian species to clades as classified in their phylogeny (Pyron and Wiens 2011).

The Pyrons and Wiens ML and time-calibrated trees were used to create a tree at the family level by first automatically assigning the species to families using the taxize package in R (Chamberlain and Szöcs 2013). These were manually verified against the taxonomy of the Pyrons and Wiens tree. The HighLevelTree function in the EvobiR package in R by Heath Blackmon was used to obtain the family level trees (evobiR: evolutionary biology in R. R package version 1.0. http://CRAN.R-project.org/package=evobiR).

Regression analysis: The pgls analysis was carried out in R with the *caper* package. The maximum likelihood value of lambda was allowed to vary while kappa and delta were set to 1 as in Kozak and Wiens (Ecology and Evolution 2016). The time-calibrated family tree was obtained from the Amphibia tree of Pyron and Wiens (Pyron and Wiens 2013), or the phylogenetic tree from Kozak and Weins (Kozak and Wiens 2010) for the Plethodontidae dataset.

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Results

Species diversity, body size and C-value variation at the family-level of urodeles clades.

The oldest known urodeles date from 166 to 168 Mya (Marjanovic and Laurin 2014, Laurin, Canoville et al. 2015). The urodeles inhabit a wide variety of ecological niches and exhibit a large diversity of life-history traits, including small and large body sizes, paedomorphy, metamorphosis and direct development (Wake 2009). In an earlier study of Amphibia, Pyron and Wiens revealed a number of ecological correlates between species diversity and variables such as geographical latitude, geographic range and environmental energy (Pyron and Wiens 2013). Species diversity in frogs, salamanders and caecilians also varies according to abiotic factors such as humidity and temperature and biotic factors such as productivity and rates of diversification (extinction and speciation). Figure 1 shows the phylogenetic tree derived from Pyron and Wiens that was used here to investigate the relationship between salamander genome size and species diversity (Pyron and Wiens 2011). Substantial variations in species diversity, body size and C-value among and within clades is apparent in the phylogenetic tree of the Urodela.

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

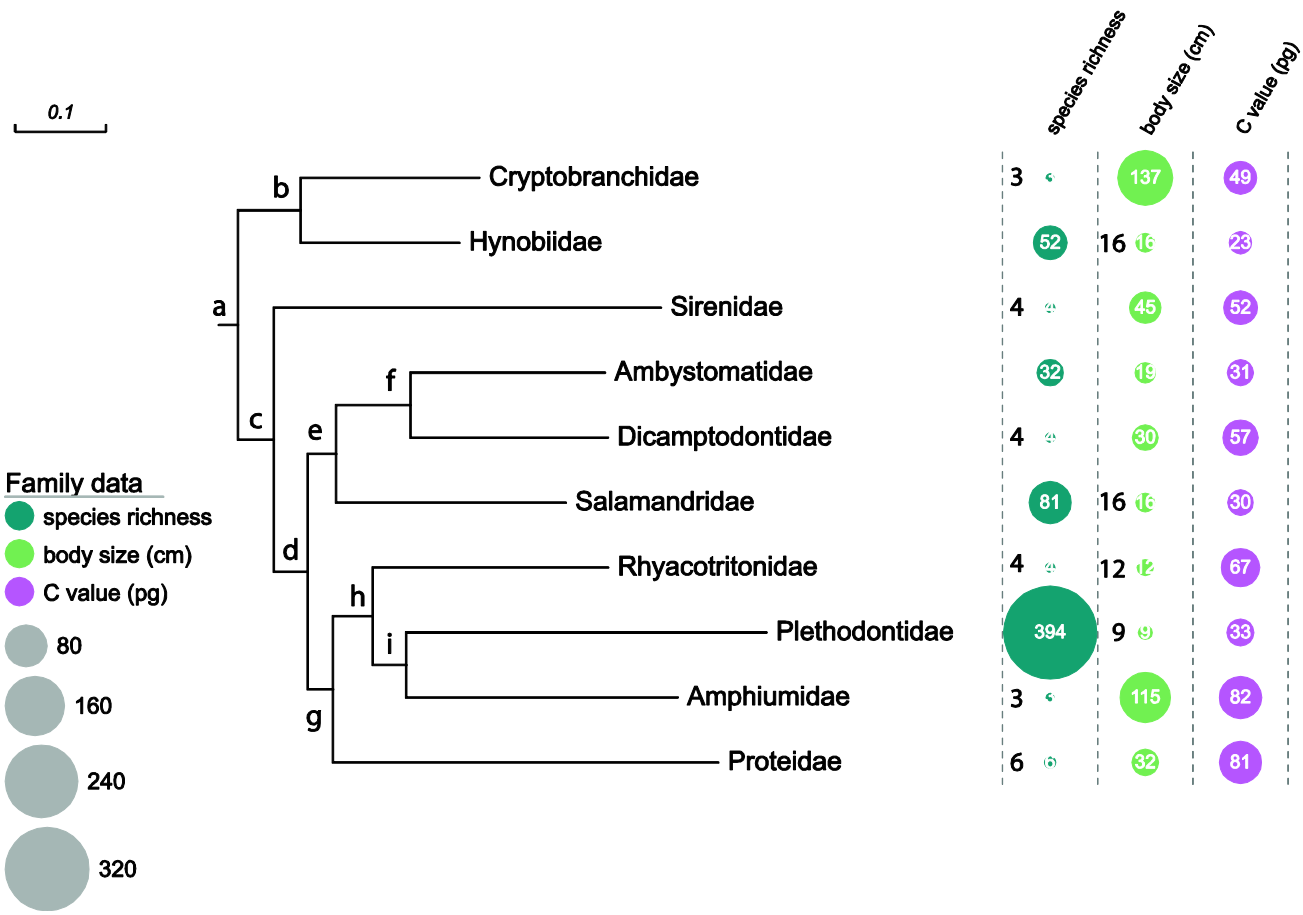


Figure 1. Maximum likelihood phylogenetic tree of Caudata derived from Pyron and Wiens 2011. Terminal branch lengths shown are the distances measured as substitutions per site up to the crown origin of each family. Species diversity, body size and average C-value are shown next to the family name in circles proportional to size. The branch length is indicated by the scale bar. Six families form three sister-pair taxa: nodes i, f and b, which correspond to the most recent common ancestor (MRCA) of the subtending families. Figure created using EvolView (Zhang, Gao et al. 2012).

Crown age, but not diversification rate, correlates with clade diversity

Among animal taxa, it has been reported that clade age rather than diversification rate explains species richness (McPeek and Brown 2007). A more recent study of amphibians, birds and mammals supports the finding that time (older lineages), and not diversification rates, explains extant species richness (Marin and Blair Hedges 2016). A significant positive correlation between crown group age at the family level and species diversity has also been reported in salamanders (Eastman and Storfer 2011). In

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

contrast to these reports, our initial ordinary least squares (OLS) analysis at the family-level of Urodela did not confirm a significant relationship between crown age and species richness (Figure 1S).

We determined, however, that the family Proteidae represents an outlier in the regression analyses (studentized residual: 2.35; studentized deleted residual: 3.94). We further determined that the outlier status of the Proteidae is due to a single monotypic genus: the European *Proteus anguinus,* the only genus in the Proteidae clade found outside of North America. Monotypic groups complicate estimates of crown age (Wiens 2017). Excluding this monotypic genus from the regression analysis changes the Proteidae family-level crown age of 121 Mya to that of the *Necturus* genus-level crown age of 13 Mya (www.timetree.org). We repeated the OLS in the absence of *Proteus anguinus*, and found a significant relationship between crown age and species richness (Figure 2A; Table 1). Phylogenetic generalized least squares (PGLS) analysis confirmed that species diversity in salamanders increases with clade age: older clades tend to have higher species diversity than comparatively younger clades (Table 1). In contrast to crown age, we found no correlation between stem age and clade diversity at the family level ($R^2 = 0.1$; $P = 0.8$; Table 1), in agreement with earlier reports on other taxa (Rabosky, Slater et al. 2012, Stadler, Rabosky et al. 2014, Wiens 2017). These observations underscore the outlier status of the single species *Proteus anguinus* in the regression analyses performed here. We have therefore conducted the following analyses with and without this species to examine the relationship between C-value and other physiological and ecological parameters (see Table S1 and Table 1).

Diversification rates have also been associated with extant species richness (Wiens 2017). Initial OLS regression analysis between species richness and diversification rates using the Pyron and Wiens dataset revealed a weak correlation between these two variables in the Urodela clades examined here ($R^2 =$

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

191 0,22; P=0.11; Supplementary Figure 2). PGLS regression analysis confirmed a non-significant
192 relationship between diversification rate and species richness (R^2 = 0.18; P = 0.12), suggesting that clade
193 ages rather than diversification rates better explain species richness patterns at the family-level of
194 Urodela clades.
195
196 *Average C-value correlates with crown age but not diversification rate.*
197 The ancestral Amphibia genome (C-value ~ 3 pg) has experienced massive amplification during the
198 evolution of the Urodela (Organ, Canoville et al. 2011, Organ, Struble et al. 2016). An early ancestor of
199 salamanders, for example, has been estimated to have a large genome size of approximately 36.7 pg,
200 which is similar to the reconstructed ancestral genome size of ~32 pg (Laurin, Canoville et al. 2015). An
201 earlier study suggested that genome size in salamanders has increased with time, although the exact rate
202 of increase remains unclear (Martin and Gordon 1995, Sessions 2008). We therefore examined the
203 relationship between phylogenetic stem age and genome size, but found no significant relationship
204 between stem age and C-value (Table 1) (Supplementary Figure 3). Examination of the relationship
205 between crown age and C-value using the Pyron and Wiens dataset, in contrast, revealed a significant
206 negative relationship between C-value and crown age (R^2 = 0.66; P = 0.025; Figure 2B). PGLS
207 regression analysis confirmed a strong negative correlation between C-value and crown age at the
208 family-level of Urodela clades (R^2 = 0,7 P = 003; Table 1).
209
210 We next investigated the relationship between C-value and diversification rate using the data on
211 diversification rates from Pyron and Wiens. Using OLS, C-value is not significantly correlated with
212 diversification rate (R^2 = 0.13; P = 0.16; Table 1). Species richness normalized by crown age
213 corresponds to the net diversification rate of a clade (Kozak and Wiens 2016). We therefore examined

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

214 the relationship between net diversification rates and C-value, and found that the relationship between
215 C-value and net diversification rate is also not significant ($R^2 = 0.18$; $P = 0.12$; Table 1).
216
217 *C-value is related to species richness*.
218 Figure 1 suggests that C-value is negatively associated with both species diversity and body size. Our
219 analysis of body size did not reveal a significant relationship between C-value and average body size
220 (Supplementary Figure 4 AB; Table 1). To further characterize the apparent trend between C-value and
221 species diversity revealed in the tree, we plotted species richness values against C-values obtained from
222 the Animal Genome Size Database (Gregory 2015). Figure 2C reveals a clear trend between average C-
223 value of a clade at the family level and its corresponding species diversity ($R^2 = 0.46$; $P = 0.02$). PGLS
224 analysis confirmed the association between C-value and species richness (Table 1): as C-value decreases
225 below 50 pg, clade diversity increases more than 10X. The sister clades Hynobidae and
226 Cryptobranchidae, for example, differ over 2X with respect to average genome size and over 10X with
227 respect to species diversity. The other two sister clades apparent in Figure 1 exhibit a similar trend
228 (Amphiumidae: Plethodontidae; Dicamptodontidae: Ambystomatidae). Together, these observations
229 suggest that the relationship between C-value and species richness is independent of phylogenetic
230 relatedness.
231

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

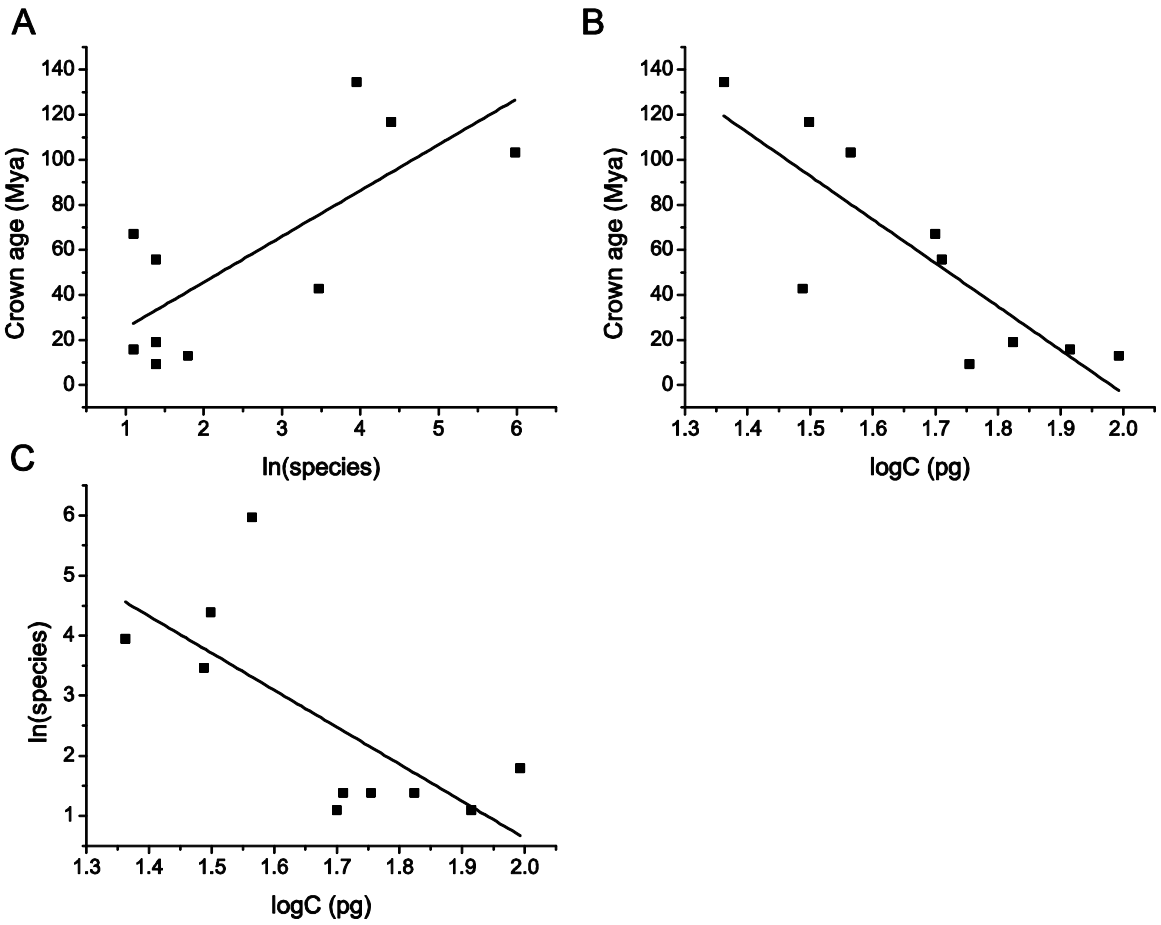


Figure 2

Figure 2: Species diversity is negatively associated with C-value. Correlations between crown age, species richness and genome size in Urodele. The regression line corresponds to ordinary least squares. The results of the ordinary least squares fit are shown in Supplementary Table 1. The results from the pgls analysis are shown in Table 1. The data on crown age and species richness is from (Pyron and Wiens 2013).

Geographic area correlates with species richness and C-value.

Older clades (crown age) have had more time to disperse over larger areas, suggesting that geographic area might correlate positively with species richness. The OLS results shown in Figure 3A reveal a positive relationship between clade diversity and geographic area in urodeles. PGLS analysis confirmed a strong relationship between species richness and geographic area (R^2 = 0.6, P = 0.006; Table 1). In

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

contrast, a strong negative correlation was found between C-value and geographic area: clades with smaller average genome sizes occupy larger geographic areas (Figure 3B; $R^2 = 0.66$; $P = 0.002$). PGLS analysis confirmed the strong association between average genome size and geographic area ($R^2 = 0.74$; $P = 0.0008$; Table 1).

Climatic niche evolution is associated with both area and C-value.

Larger geographic areas are expected to comprise higher levels of both habitat and niche diversity, suggesting that geographic range at the family-level of Urodela clades is associated with the rate of climatic-niche evolution (Kozak and Wiens 2010, Pyron and Wiens 2013). OLS analysis revealed a positive relationship between niche rate and area ($R^2 = 0.54$; $P = 0.009$; Figure 3C), as well as a negative relationship between niche rate and average C-value ($R^2 = 0.48$; 0.015; Figure 3D). PGLS analyses confirmed that niche rate is significantly related to area, C-value and species richness (Table 1).

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

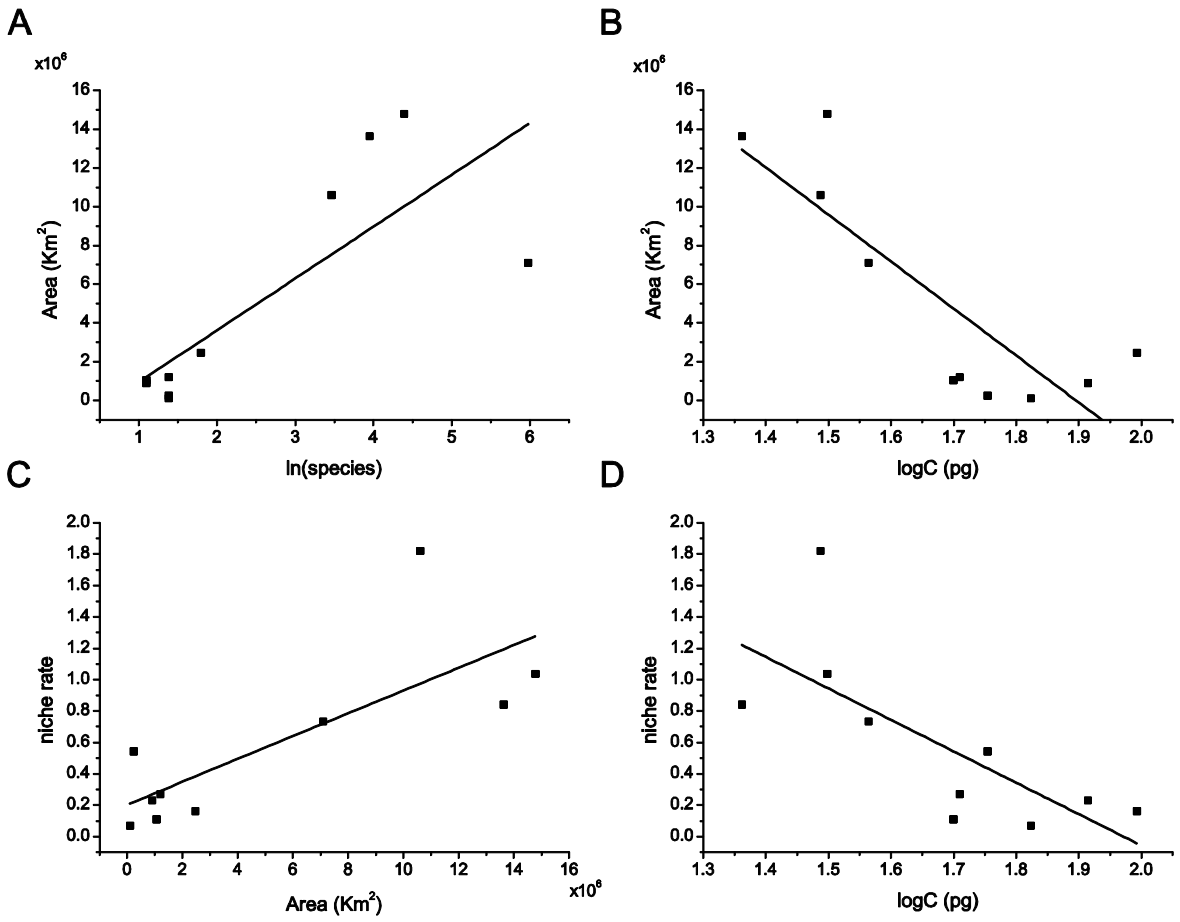


Figure 3

Figure 3: Correlations between area, species richness, genome size and niche rate. The regression line corresponds to ordinary least squares. The results of the ordinary least squares fit are shown in Supplementary Table 1. The results from the pgls analysis are shown in Table 1. Area, species richness and niche rate data were obtained from (Pyron and Wiens 2013).

Plethodontids

A significant relationship has previously been reported between diversification rate and niche rate in 15 plethodontid clades (Kozak and Wiens 2010). Our analysis reveals that niche rate and diversification rate are similarly related at the family-level of Urodela clades (R^2 = 0.44; P = 0.02; Supplementary Figure 6A; Table 1). Using the values of niche rate for plethodontids reported in Kozak and Wiens 2016, we also found a strong association between C-value and niche rate at this lower taxonomic level (R^2 =

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

0.48; P = 0.025; Supplementary Figure 6B; Table 2). In contrast, C-value in plethodontids is not associated with either geographic area, diversification rates, or species diversity (Table 2).

Genome size diversity and species diversity

Together, our results suggest a potential relationship between species diversity and genome size diversity: clades with higher levels of species diversity are expected to have correspondingly higher levels of genome size diversity. Genome size is expected to evolve in a manner that is proportional to C-value, suggesting that larger genomes are changing faster in size compared to smaller genomes (Oliver, Petrov et al. 2007). Taxonomic groups with larger genomes sizes should therefore exhibit higher genome size diversity. The coefficient of variation (CV) of the log of C-value was used here to assess genome size diversity. We found a negative relationship between average C-value in a clade and its corresponding CV; however, the relationship was not significant despite the apparent trend (Figure 4A; $R^2 = 0.25$; P = 0.14).

We next examined the relationships between CV of C-value and the other diversity variables. Figure 4B reveals a very strong relationship between CV of C-value and species diversity (Table 1), as expected if genome size has been evolving in parallel with species diversity since the emergence of the different family-level salamander clades. Consistent with an increase in species diversity with time, we also found a significant relationship between crown age and the CV of C-value ($R^2 = 0.53$; P = 0.04; Figure 4C). CV of C-value, however, did not correlate significantly with area, niche rate or diversification rate (Table 1.).

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

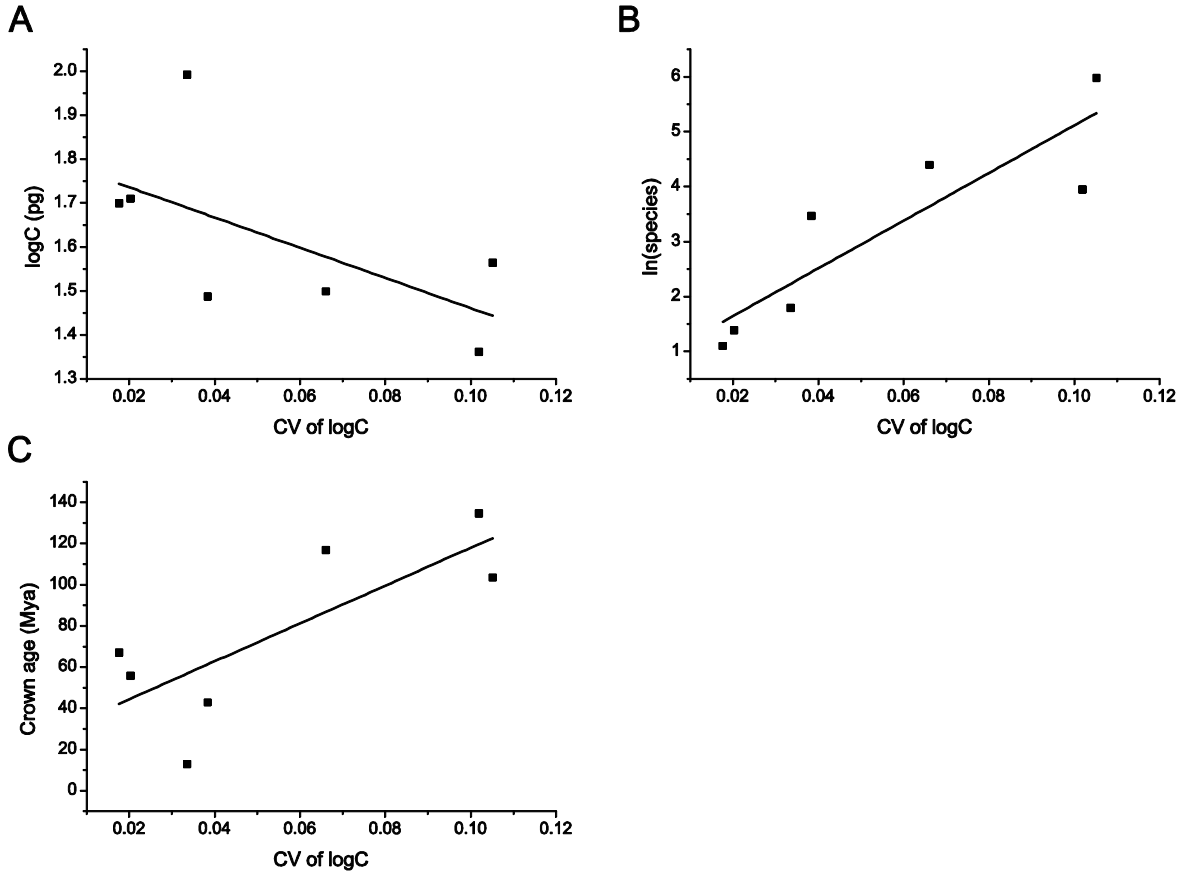


Figure 4

Figure 4: Correlations between the CV of log(C-value), species richness and crown age. The regression line corresponds to ordinary least squares. The results of the ordinary least squares fit are shown in Supplementary Table 1. The results from the pgls analysis are shown in Table 1. Crown age and species richness data were obtained from (Pyron and Wiens 2013).

Discussion

We report here a relationship between average genome size and species diversity at the family level of Urodela clades (Figure 2B). Although we cannot, based on these analyses, establish a causal relation between these two variables, our findings nevertheless provide evidence that C-value constitutes an additional trait associated with species diversity, at least in urodeles. We also acknowledge that our

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

sample size is limited (ten clades); however, the different and strongly significant correlations found here are remarkably consistent: C-value correlates with time, area, niche rate and species richness. Moreover, our findings on the correlation between C-value and niche rate as well as the correlation between diversification rate and niche rate are well supported by our additional analyses on the larger plethodontid dataset of Kozac and Wiens (n = 15). Examining other factors such as plethodontid clade age, geographic area and rate of climatic-niche evolution suggests that the relationship between genome size and species diversity is mediated directly or indirectly through a number of different variables. Our findings suggest that time and geographic area rather than diversification rate account for the relationship found here between species richness and C-value.

Some questions might be raised concerning the phylogenetic status of *P. anguinus*. The monophyly of the Proteidae has been controversial for several decades (Larsen and Guthrie 1974, Hecht and Edwards 1976). Only relatively recently has this species been established as a bona fide member of the family Proteidae (Wiens, Bonett et al. 2005). In our analyses, however, this species constitutes a consistent outlier among the Proteidae suggesting that its relationship to the other Proteidae species is substantially more distant than between any of the other salamader species at the family taxonomic level (Weisrock, Harmon et al. 2005). Including *P. anguinus* in the same taxonomic group as Necturus, for example, results in anomalously high levels of Proteidae heterozygosity (Nevo and Beiles 1991), genome and body size variance and other genetic and physiological discrepancies that render the Proteidae a biometric outlier in the Urodela group. Here we have controlled for the outlier status of *P. anguinus* by performing our analyses with and without this species. We find that, although the correlations with C-value are significantly weakened, including *P. anguinus* in the analyses does not change the overall

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

conclusion: genome size and genome size diversity are closely associated with species richness, climatic niche rate and clade crown age (Table S1).

What is the nature of the relationship between average C-value and these different ecological variables in the Urodela? Sessions reported that the range in genome size in family-level Urodela clades tends to increase as their average C-values decrease (Sessions 2008), suggesting a potential relationship between genome size diversity and average genome size in a clade. Our analysis does not support a significant relationship between CV of genome size and C-value. We did find, however, a significant relationship between CV of genome size and both species richness and crown age, suggesting that diversification of genome size over time coincided with diversification of species in a clade. In contrast, genome size diversity is not significantly related to geographic range (Table 1 and Supplementary Table 1). These observations suggest that changes in genome size underlie speciation events in urodeles, consistent with findings in plants that rates of genome size evolution correlate with rates of speciation (Puttick, Clark et al. 2015).

We note that the observations made here concerning average C-value and species diversity apply predominantly to the family-level of Urodela clades. At lower taxonomic levels such as the plethodontids, the relationship between genome size and species richness shows no consistent pattern. The Bolittoglossinae, for example, tend to have larger clade-average genome sizes but higher species diversity than other genera in the plethodontids. Indeed, the Plethodontidae as a group exhibit the highest levels of species diversity among family-level Urodela clades (Wake 2009), and have correspondingly elevated levels of genome size diversity and substitution rates (Herrick and Sclavi 2014). Genome sizes range from ~10 pg in the genus *Grynophilus*, for example, to over 70 pg in the

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

346 genus *Hyrdromantes*. C-values in *Plethodon* alone, one of the most species rich genera of
347 Plethodontidae, range from 18 pg in *P. Shenandoah* to 69 pg in *P. vandykei*, suggesting a more general
348 relationship between genome size evolution, mutation rate and speciation. We suggest that genome size
349 diversity, rather than genome size itself, reflects the correspondingly higher substitution rates previously
350 reported for the plethodontidae as a group (Herrick and Sclavi 2014).
351
352 Together, our analyses support the *Geographic Range Hypothesis*, according to which taxa with wider
353 geographic distributions have higher probabilities that genetic changes such as indels, chromosomal
354 inversions and transposon-mediated modifications in genome size will become fixed in a population
355 (Feder, Gejji et al. 2011, Martinez, Jacobina et al. 2017). If habitat and niche availability both increase
356 with geographic area, then our observations suggest that changes in genome size in urodeles might have
357 occurred in parallel with adaptations that made available habitats and niches more accessible to
358 dispersing ancestral populations. Since standardized contrasts can be interpreted as representing
359 evolutionary rates (Oliver, Petrov et al. 2007), the strong correlation and negative slope between
360 independent contrasts of clade area and C-value indicate that the rate at which C-value changes
361 coincides with the rate at which geographic range changes (Supplementary Figure 6; adjusted $R^2 = 0.74$;
362 $P = 0.0009$).
363
364 In support of that proposal, a recent study on karyotypic diversification rates at the family level of
365 mammalian clades suggests that large, past geographic distributions in heterogenous environments
366 might have favored higher levels of chromosomal diversity; or conversely, higher rates of chromosomal
367 diversification might have promoted colonization of new habitats and expanding geographic ranges
368 (Martinez, Jacobina et al. 2017). These results are consistent with earlier results on the rate of karyotype

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

diversification in mammals, frogs and salamanders (Wilson, Sarich et al. 1974, Wilson, Bush et al. 1975, Bush, Case et al. 1977, Bengtsson 1980). These authors found that salamander genomes have slower rates of karyotype diversification than frogs, while frogs exhibit slower rates than mammals, suggesting that taxonomic groups with larger genomes on average have slower rates of genome evolution (Hooper and Price 2015, Leaché, Banbury et al. 2016). Our findings therefore support the proposal that genome turnover (eg. chromosomal inversions, translocations and rearrangements), in addition to changes in genome size, is a factor underlying speciation rates and extant species richness in vertebrates. We are currently investigating the hypothesis that genome diversification rates and corresponding levels of genome size and karyotype diversity, rather than absolute C-value, explain rates of molecular evolution and rates of speciation in Amphibia and other eukaryotes.

Acknowledgements

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.

References

Bengtsson, B. O. (1980). "Rates of karyotype evolution in placental mammals." Hereditas **92**(1): 37-47.

Bush, G. L., S. M. Case, A. C. Wilson and J. L. Patton (1977). "Rapid speciation and chromosomal evolution in mammals." Proc Natl Acad Sci U S A **74**(9): 3942-3946.

Böhne, A., F. Brunet, D. Galiana-Arnoux, C. Schultheis and J. N. Volff (2008). "Transposable elements as drivers of genomic and biological diversity in vertebrates." Chromosome Res **16**(1): 203-215.

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.

Chamberlain, S. A. and E. Szöcs (2013). "taxize: taxonomic search and retrieval in R." F1000Res **2**: 191.

Eastman, J. M. and A. Storfer (2011). "Correlations of life-history and distributional-range variation with salamander diversification rates: evidence for species selection." Syst Biol **60**(4): 503-518.

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

396 Feder, J. L., R. Gejji, T. H. Powell and P. Nosil (2011). "Adaptive chromosomal divergence driven by
397 mixed geographic mode of evolution." Evolution **65**(8): 2157-2170.

398 Francis, D., M. S. Davies and P. W. Barlow (2008). "A strong nucleotypic effect on the cell cycle
399 regardless of ploidy level." Ann Bot **101**(6): 747-757.

400 Frentiu, F., J. Ovenden and R. Street (2001). "Australian lungfish (Neoceratodus forsteri:Dipnoi) have
401 low genetic variation at allozyme and mitochondrial DNA loci: A conservation alert?" Conservation
402 Genetics **2**: 63-67.

403 Gregory, T. (2004). "Macroevolution, Hierarchy Theory, and the C-value Enigma." Paleobiology **30**(2):
404 179-202.

405 Gregory, T. R. (2015). Animal Genome Size Database. http://www.genomesize.com.

406 Hecht, M. K. and J. L. Edwards (1976). "The determination of parallel or monophyletic relationships:
407 The Proteidae salamanders-- a test case." American Naturalist **110**: 653-677.

408 Herrick, J. (2011). "Genetic variation and replication timing, or why is there late replicating DNA?"
409 Evolution **65**(11): 3031-3047.

410 Herrick, J. and B. Sclavi (2014). "A new look at genome size, evolutionary duration and genetic
411 variation in salamanders." Comptes Rendus Palevol **13**(7): 611-615.

412 Hooper, D. M. and T. D. Price (2015). "Rates of karyotypic evolution in Estrildid finches differ between
413 island and continental clades." Evolution **69**(4): 890-903.

414 Hughes, J. M., D. J. Schmidt, J. A. Huey, K. M. Real, T. Espinoza, A. McDougall, P. K. Kind, S.
415 Brooks and D. T. Roberts (2015). "Extremely low microsatellite diversity but distinct population
416 structure in a long-lived threatened species, the Australian lungfish Neoceratodus forsteri (Dipnoi)."
417 PLoS One **10**(4): e0121858.

418 Jurka, J., W. Bao, K. K. Kojima, O. Kohany and M. G. Yurka (2012). "Distinct groups of repetitive
419 families preserved in mammals correspond to different periods of regulatory innovations in vertebrates."
420 Biol Direct **7**: 36.

421 Karlin, A. A. and D. B. Means (1994). "Genetic variation in the aquatic salamander genus Amphiuma."
422 American Midland Naturalist **132**(1): 1-9.

423 Kozak, K. H. and J. J. Wiens (2010). "Accelerated rates of climatic-niche evolution underlie rapid
424 species diversification." Ecol Lett **13**(11): 1378-1389.

425 Kozak, K. H. and J. J. Wiens (2016). "What explains patterns of species richness? The relative
426 importance of climatic-niche evolution, morphological evolution, and ecological limits in salamanders."
427 Ecol Evol **6**(16): 5940-5949.

428 Kraaijeveld, K. (2010). "Genome Size and Species Diversification." Evolutionary Biology **37**: 227-233.

429 Larsen, J. H. and D. J. Guthrie (1974). "Parallelism in the Proteidae reconsidered." Copeia: 635-643.

430 Larson, A. (1981). "A reevaluation of the relationship between genome size and genetic variation." The
431 American Naturalist **118**(1): 119-125.

432 Laurin, M., A. Canoville, M. Struble and V. Buffrénil (2015). "Early genome size increase in urodeles."
433 Comptes Rendus Palevol **15**(1-2): 74-82.

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Leaché, A. D., B. L. Banbury, C. W. Linkem and A. N. de Oca (2016). "Phylogenomics of a rapid radiation: is chromosomal evolution linked to increased diversification in north american spiny lizards (Genus Sceloporus)?" BMC Evol Biol **16**: 63.

Mank, J. E. and J. C. Avise (2006). "Cladogenetic correlates of genomic expansions in the recent evolution of actinopterygiian fishes." Proc Biol Sci **273**(1582): 33-38.

Marin, J. and S. Blair Hedges (2016). "Time best explains global variation in species richness of amphibians, birds and mammals " Journal of Biogeography **43**: 1069-1079.

Marjanovic, D. and M. Laurin (2014). "An updated paleontological timetree of lissamphibians with comments on the anatomy of Jurassic crown-group salamanders (Urodela)." Historical Biology: An International Journal of Paleobiology **26**(4): 535-550.

Martin, C. C. and R. Gordon (1995). "Differentiation trees, a junk DNA molecular clock, and the evolution of neoteny in salamanders." J. Evol. Biol. **8**(3): 339-354.

Martinez, P. A., U. P. Jacobina, R. V. Fernandes, C. Brito, C. Penone, T. F. Amado, C. R. Fonseca and C. J. Bidau (2017). "A comparative study on karyotypic diversification rate in mammals." Heredity (Edinb) **118**(4): 366-373.

Matsui, M., A. Tominaga, W. Z. Liu and T. Tanaka-Ueno (2008). "Reduced genetic variation in the Japanese giant salamander, Andrias japonicus (Amphibia: Caudata)." Mol Phylogenet Evol **49**(1): 318-326.

McPeek, M. A. and J. M. Brown (2007). "Clade age and not diversification rate explains species richness among animal taxa." Am Nat **169**(4): E97-106.

Metcalfe, C. J. and D. Casane (2013). "Accommodating the load: The transposable element content of very large genomes." Mob Genet Elements **3**(2): e24775.

Nevo, E. and A. Beiles (1991). "Genetic diversity and ecological heterogeneity in amphibian evolution." Copeia **19**(3): 565-592.

Oliver, M. J., D. Petrov, D. Ackerly, P. Falkowski and O. M. Schofield (2007). "The mode and tempo of genome size evolution in eukaryotes." Genome Res **17**(5): 594-601.

Olmo, E. (2006). "Genome size and evolutionary diversification in vertebrates." Italian Journal of Zoology **73**(2): 167 - 171.

Organ, C., M. Struble, A. Canoville, V. Bruffénil and M. Laurin (2016). "Macroevolution of genome size in sarcopterygians during the water–land transition." Comptes Rendus Palevol **15**(1-2): 65-73.

Organ, C. L., A. Canoville, R. R. Reisz and M. Laurin (2011). "Paleogenomic data suggest mammal-like genome size in the ancestral amniote and derived large genome size in amphibians." J Evol Biol **24**(2): 372-380.

Palazzo, A. F. and T. R. Gregory (2014). "The case for junk DNA." PLoS Genet **10**(5): e1004351.

Parker, E. and M. Kreitman (1982). "On the relationship between heterozygosity and DNA content." The American Naturalist **119**(5): 749-752.

Pierce, A. and B. Mitton (1980). "The relationship between genome size and genetic variation." The American Naturalist **116**(6): 850-861.

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Puttick, M. N., J. Clark and P. C. Donoghue (2015). "Size is not everything: rates of genome size evolution, not C-value, correlate with speciation in angiosperms." Proc Biol Sci **282**(1820): 20152289.

Pyron, R. A. and J. J. Wiens (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**(2): 543-583.

Pyron, R. A. and J. J. Wiens (2013). "Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity." Proc Biol Sci **280**(1770): 20131622.

Pyron, R. A. a. W. J. J. (2011). "A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians." Molecular phylogenetics and evolution **61**(2): 543--583.

Rabosky, D. L., G. J. Slater and M. E. Alfaro (2012). "Clade age and species richness are decoupled across the eukaryotic tree of life." PLoS Biol **10**(8): e1001381.

Sessions, S. (2008). "Evolutionary cytogenetics in salamanders." Chromosome Res. **16**(1): 183-201.

Sessions, S. K. and A. Larson (1987). "Developmental correlates of genome size in plethodontid salamanders and their implications for genome size evolution." Evolution **41** (6): 1239-1251

Smith, S. A. and B. C. O'Meara (2012). "treePL: divergence time estimation using penalized likelihood for large phylogenies." Bioinformatics **28**(20): 2689-2690.

Stadler, T., D. L. Rabosky, R. E. Ricklefs and F. Bokma (2014). "On age and species richness of higher taxa." Am Nat **184**(4): 447-455.

Sun, C., J. R. López Arriaza and R. L. Mueller (2012). "Slow DNA loss in the gigantic genomes of salamanders." Genome Biol Evol **4**(12): 1340-1348.

Sun, C. and R. L. Mueller (2014). "Hellbender genome sequences shed light on genomic expansion at the base of crown salamanders." Genome Biol Evol **6**(7): 1818-1829.

Vinogradov, A. E. (2004). "Genome size and extinction risk in vertebrates." Proc Biol Sci **271**(1549): 1701-1705.

Wake, D. (2009). "What Salamanders have Taught Us about Evolution." Ann. Rev. Ecol. Evol. Syst. **40**: 333-352.

Weisrock, D. W., L. J. Harmon and A. Larson (2005). "Resolving deep phylogenetic relationships in salamanders: analyses of mitochondrial and nuclear genomic data." Syst Biol **54**(5): 758-777.

Wiens, J., R. Bonett and P. Chippindale (2005). "Ontogeny discombobulates phylogeny: paedomorphosis and higher-level salamander relationships." Syst Biol **54**(1): 91-110.

Wiens, J. J. (2017). "What explains patterns of biodiversity across the Tree of Life?: New research is revealing the causes of the dramatic variation in species numbers across branches of the Tree of Life." Bioessays **39**(3).

Wilson, A. C., G. L. Bush, S. M. Case and M. C. King (1975). "Social structuring of mammalian populations and rate of chromosomal evolution." Proc Natl Acad Sci U S A **72**(12): 5061-5065.

Wilson, A. C., V. M. Sarich and L. R. Maxson (1974). "The importance of gene rearrangement in evolution: evidence from studies on rates of chromosomal, protein, and anatomical evolution." Proc Natl Acad Sci U S A **71**(8): 3028-3030.

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

511 Zhang, H., S. Gao, M. J. Lercher, S. Hu and W. H. Chen (2012). "EvolView, an online tool for
512 visualizing, annotating and managing phylogenetic trees." Nucleic Acids Res **40**(Web Server issue):
513 W569-572.

514

515

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

516

	Adj R^2	p	AICc	Pagel's λ
Area vs logC	**0.70**	**0.001**	**332**	**0**
Area vs species richness	**0.67**	**0.002**	**334**	**0**
Area vs niche rate	**0.53**	**0.01**	**337**	**0**
Area vs CV of logC	0.47	0.05	239	0.001
Area vs crown age	**0.65**	**0.003**	**335**	**0.568**
LogC vs area	**0.74**	**0.0008**	**-12.5**	**0**
LogC vs species richness	**0.69**	**0.003**	**-9**	**1**
LogC vs sr/crown	0.16	0.13	-1.35	0
LogC vs niche rate	**0.58**	**0.006**	**-7.4**	**1**
LogC vs crown age	**0.7**	**0.0015**	**-11.6**	**0**
LogC vs div. rate	0.26	0.07	-1.8	1
CV of logC vs area	0.47	0.05	-25	0
CV of logC vs logC	0.38	0.08	-24	0
CV of logC vs niche rate	-0.04	0.428	-21	0
CV of logC vs crown age	**0.58**	**0.03**	**-27**	**0**
CV of logC vs species richness	**0.78**	**0.005**	**-30**	**1**
CV of logC vs div rate	0.12	0.23	-22	0
Species richness vs niche rate	**0.39**	**0.03**	**37**	**0**
Species richness vs div rate	0.18	0.12	40	0
Species richness vs crown age	**0.67**	**0.002**	**34**	**1**

517
518
519 **Table 1. PGLS analysis for Salamander family-level clades of correlations of C-values with data**
520 **from Pyrons and Wiens (Pyron and Wiens 2013) without *Proteus anguinus*.**
521

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

	Adj R^2	*p*	*AICc*	*Pagel's λ*
LogC vs area	0.14	0.09	-12.65	0.313
LogC vs species richness	-0.07	0.8	-11	1
LogC vs sr/crown	-0.07	0.76	-11	1
LogC vs niche rate	**0.48**	**0.002**	**-19**	**0**
LogC vs crown age	-0.07	0.76	-11	1
LogC vs div. rate	-0.02	0.45	-11	1
Species richness vs area	**0.24**	**0.04**	**33**	**0.976**
Species richness vs div rate	**0.47**	**0.003**	**25**	**0**
Species richness vs niche rate	**0.33**	**0.015**	**29**	**0**
Species richness vs crown age	-0.07	0.96	36	0.54
Div rate vs crown	**0.36**	**0.01**	**-55**	**0**
Div rate vs niche rate	**0.33**	**0.015**	**-55**	**0**
Crown age vs area	-0.07	0.96	473	1

Table 2. PGLS analysis of Plethodontidae data from Kozak and Wiens (Kozak and Wiens 2016).

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Supplementary material

	Adj R^2	p	AICc	Pagel's λ
Area vs logC	**0.77**	**0.0005**	**329**	**0**
Area vs species richness	**0.69**	**0.002**	**332**	**0**
Area vs niche rate		*No change*		
Area vs CV of logC	0.27	0.13	238	0
Area vs crown age	**0.45**	**0.02**	**337**	**0**
LogC vs area	**0.80**	**0.0003**	**-17**	**1**
LogC vs species richness	**0.69**	**0.002**	**-13**	**1**
LogC vs sr/crown	0.12	0.17	-4	0
LogC vs niche rate	**0.60**	**0.005**	**-11**	**1**
LogC vs crown age	0.31	0.056		0.001
LogC vs div. rate	0.22	0.09	-4	1
CV of logC vs area	0.27	0.13	-26	0
CV of logC vs logC	0.07	0.28	-25	0
CV of logC vs niche rate	-0.14	0.62		0
CV of logC vs crown age	**0.7**	**0.01**	**-32**	**0**
CV of logC vs species richness	0.49	0.05	-29	0
CV of logC vs div rate	-0.05	0.44	-24	0
Species richness vs niche rate	**0.38**	**0.03**	**35**	**0**
Species richness vs div rate	0.18	0.12	38	0
Species richness vs crown age	**0.36**	**0.04**	**36**	**0.001**

Supplementary Table 1. PGLS analysis for Salamander family-level clades of correlations of C-values with data from Pyrons and Wiens (Pyron and Wiens 2013) including *Proteus anguinus*.

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

	Adj R^2	*p*
LogC vs area	**0.66**	**0.002**
LogC vs species richness	**0.46**	**0.02**
LogC vs sr/crown	0.18	0.12
LogC vs niche rate	**0.48**	**0.015**
LogC vs crown age	**0.66**	**0.025**
LogC vs stem	0.1	0.8
LogC vs div. rate	0.13	0.16
Species richness vs crown age	**0.52**	**0.01**
Species richness vs stem age	-0.09	0.64
Species richness vs div rate	0.2	0.11
Species richness vs niche rate	0.32	0.05
Species richness vs CV of logC	**0.75**	**0.007**
Area vs species richness	**0.6**	**0.006**
Area vs niche rate	**0.54**	**0.009**
Area vs CV of logC	0.33	0.1
Area vs crown age	**0.58**	**0.006**
CV of logC vs logC	0.25	0.14
CV of logC vs niche rate	-0.08	0.49
CV of logC vs crown age	**0.53**	**0.04**
CV of logC vs div rate	0.16	0.2
Niche rate vs div rate	**0.42**	**0.02**

Supplementary Table 2: results of OLS analysis for family-level clades of Salamanders family-level clades data from Pyrons and Wiens (Pyron and Wiens 2013).

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

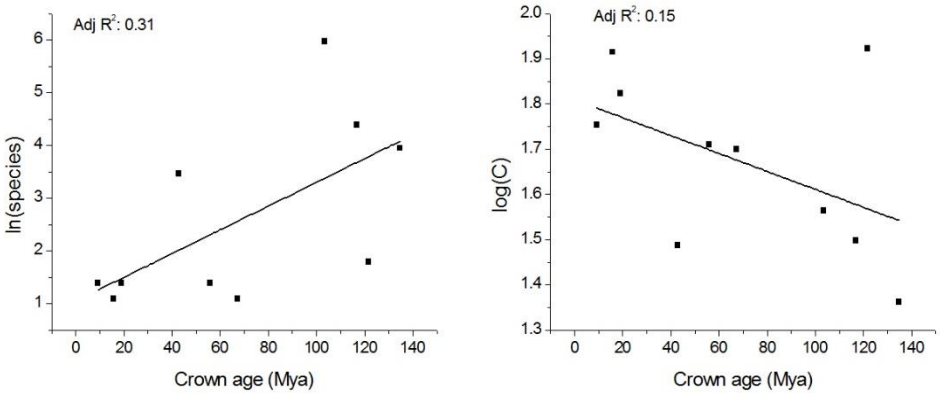


Figure 1S: OLS Regression analysis of crown age versus species richness and crown age vs genome size with *P. anguinus* included in the Proteidae family-level clade.

A B

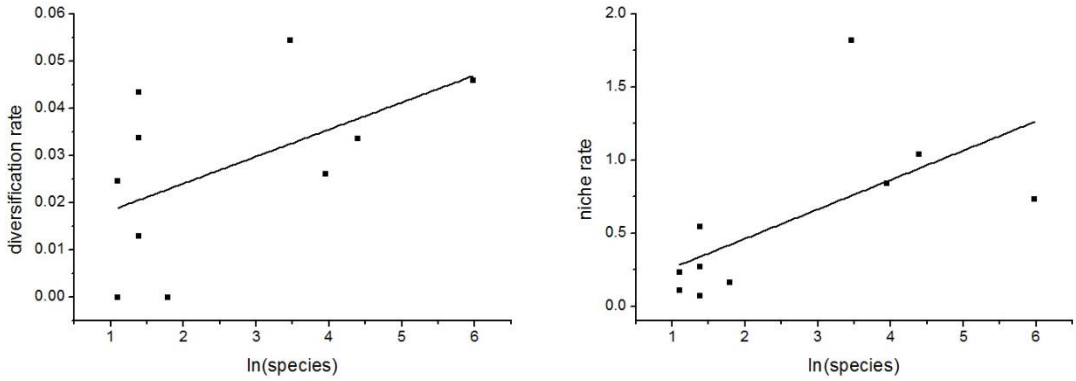


Figure 2S: OLS regression of diversification rate (A) and niche rate (B) versus species richness.

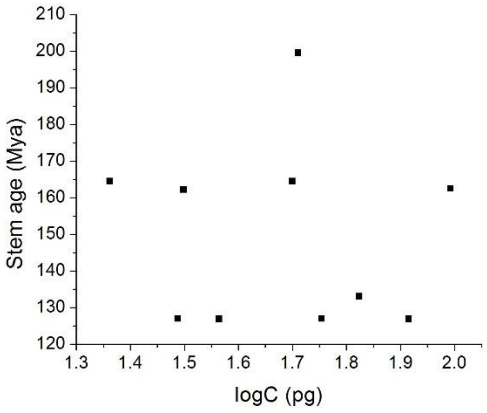
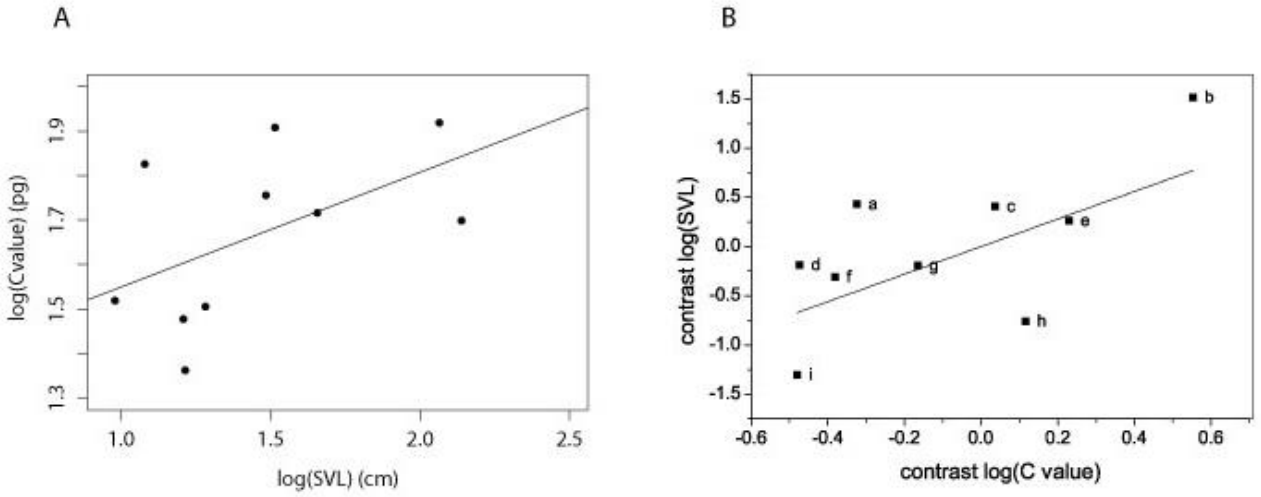


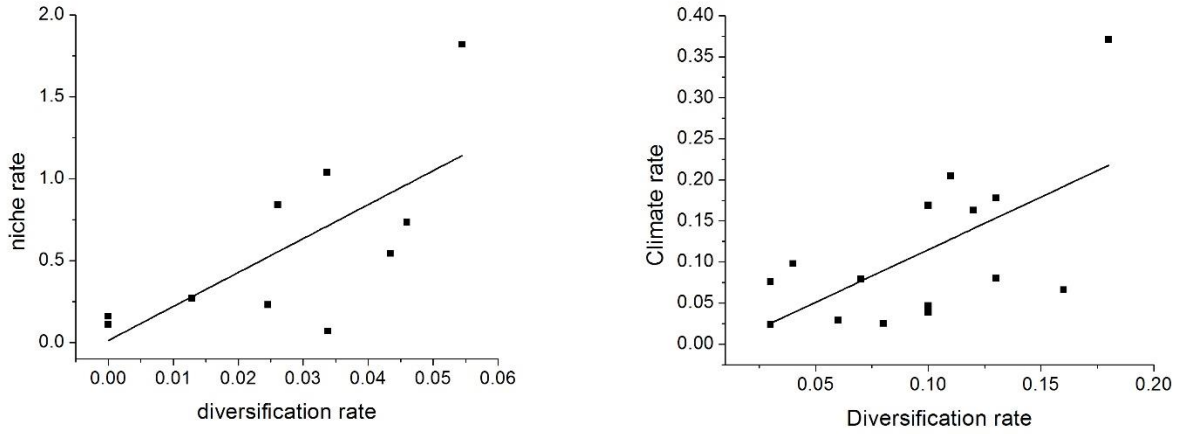
Figure 3S: Stem age versus C-value

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



552
553 Figure 4S: A) SVL (body size) versus C-value. B) Independent contrasts of SVL (body size) vs.
554 independent contrasts of C-value. The letters correspond to the nodes of the phylogenetic tree shown in
555 Figure 1 in the main text.
556
557
558

A B



559
560
561 Figure 5S: Diversification rate versus niche rate for Salamanders family-level clades (A) and
562 Plethodontidae from Kozak and Wiens (B) (Kozak and Wiens 2016)
563

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

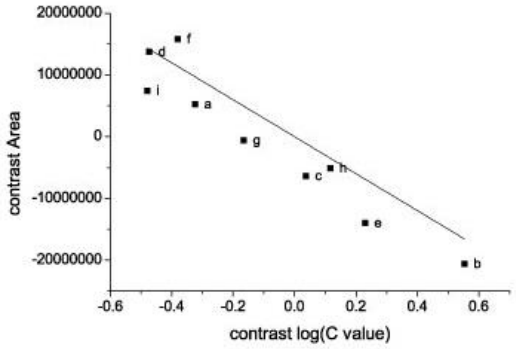


Figure 6S: Independent contrasts of C-value versus independent contrasts of area. Letters refer to nodes in the phylogenetic tree (see Figure 1).

Supplementary Methods:

Data on snout to vent length (SVL) for the designated species were obtained from AmphibiaWeb (http://www.amphibiaweb.org) using the Rafaelli account. Only those values were used for which the genome size is represented in the Animal Genome Size Database. SVL is commonly used as a proxy for body size, and is measured as the distance from the tip of the snout to the anterior posterior to the cloaca. In some cases, only either male or female body size was available. When both male and female body size were reported, the average of the two was used. Average SVL for each clade was calculated as described for C-values in the main text.

Independent contrasts, pics, were carried out in R using the *ape* library based on the branch lengths of the ML tree shown in Figure 1. The regression of the independent contrasts was forced through the origin. The residuals of the regression for the independent contrasts (log C-value vs. log body size and log C-value vs. area, respectively) were controlled for normality using the Shapiro-Wilk test, confirming a normal distribution.

bioRxiv preprint doi: https://doi.org/10.1101/065425; this version posted November 6, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

585

586