Nuclear DNA content correlates with depth, body size, and diversification rate in amphipod crustaceans from ancient Lake Baikal, Russia

Journal:	<i>Genome</i>
Manuscript ID	gen-2016-0128.R1
Manuscript Type:	Article
Date Submitted by the Author:	16-Aug-2016
Complete List of Authors:	Jeffery, Nicholas; Bedford Institute of Oceanography, Fisheries and Oceans Canada Yampolsky, Lev; East Tennessee State University, Department of Biological Sciences Gregory, Ryan; University of Guelph,
Keyword:	ancient lakes, crustaceans, C-value, genome size, speciation

SCHOLARONE™
Manuscripts

Nuclear DNA content correlates with depth, body size, and diversification rate in
amphipod crustaceans from ancient Lake Baikal, Russia

Nicholas W. Jeffery^{1,2*}, Lev Yampolsky³, T. Ryan Gregory¹

¹ Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada, N1G 2W1

² Current address: Fisheries and Oceans Canada, Bedford Institute of Oceanography, 1
Challenger Drive, Dartmouth, Nova Scotia, Canada

³ Department of Biological Sciences, East Tennessee State University, Tennessee, USA

*Corresponding Author: nick.w.jeffery3@gmail.com

Abstract

Lake Baikal in Russia is a large, ancient lake that has been the site of a major radiation of amphipod crustaceans. Nearly 400 named species are known in this single lake, and it is thought that many more await description. The size and depth of Lake Baikal, in particular, may have contributed to the radiation of endemic amphipods by providing a large number of microhabitats for species to invade and subsequently experience reproductive isolation. Here we investigate the possibility that large-scale genomic changes have also accompanied diversification in these crustaceans. Specifically, we report genome size estimates for 36 species of Baikal amphipods, and examine the relationship between genome size, body size, and the maximum depths at which the amphipods are found in the lake. Genome sizes ranged nearly 8-fold in this sample of amphipod species, from 2.15 to 16.63pg, and there were significant, positive, phylogenetically-corrected relationships between genome size, body size, maximum depth, and diversification rate among these species. Our results suggest that major genomic changes, including transposable element proliferation, have accompanied speciation that was driven by selection for differences in body size and habitat preference in Lake Baikal amphipods.

Keywords: ancient lakes, crustaceans, C-value, flow cytometry, genome size, speciation

Introduction

Lake Baikal in Eastern Siberia is the largest freshwater lake in the world by volume – containing about 20% of the world's liquid fresh water -- as well as being the deepest (>1.5km at maximum) and oldest (~70 My) lake on Earth. A World Heritage Site since 1997, Lake Baikal is known as the “Galapagos of Russia” as it is home to a large number of endemic taxa. The unique fauna of the lake includes 350 described endemic species of amphipod in seven families. These likely represent radiations following two separate invasions into the lake from the surrounding watershed, with close phylogenetic affinities to gammarids found across the Holarctic (MacDonald et al. 2005; Väinölä et al. 2008). As with other groups showing high levels of endemism in Lake Baikal (e.g., gastropod molluscs, ostracod crustaceans, flatworms, and sculpins), the diversification of amphipods is thought to be associated with the physical isolation, old age, and extreme depth of the lake (MacDonald et al. 2005).

The Lake Baikal amphipod assemblage is remarkably diverse morphologically, genetically, and ecologically, including not only benthic species, but also the specialized pelagic *Macrohectopus branickii*, and several parasitic species (Bazikalova 1945; Melnik et al. 1995; MacDonald et al. 2005). Different species of amphipods are found at various depths of the lake, with evidence suggesting that expansion into deep-water habitats likely occurred independently several times in different lineages (Sherbakov et al. 1998). Some acanthogammarid amphipods in Lake Baikal exhibit spiny armour and gigantism, traits that are also seen in amphipods in the ancient Lake Ohrid and Lake Titicaca, and in the distantly-related epimeriid amphipods of Antarctica (Takhteev 2000; Chapelle, 2002; Chapelle and Peck, 2004).

Among amphipods, spiny bodies are thought to be associated with cold, low-energy environments (Martens 1997; Chapelle 2002), a manifestation of the generally observed pattern of larger body size associated with colder climates in a wide range of taxa known as Bergmann's Rule (Meiri and Dayan 2003; Walters and Hassall 2006). At the genomic level, there is evidence from various animal taxa, including amphipods, that species inhabiting polar regions possess larger genomes than their temperate counterparts (Rees et al. 2007; Dufresne and Jeffery 2011). Rees et al. (2007) suggested that the large genomes of marine arctic amphipods are the result of a-(adverse) selection, an extension of K-selection, and it is possible that large genomes exist in Baikal amphipods due to environmental similarities between deep lakes and the polar oceans (Martens 1997).

With such morphological and genetic disparity, changes in genome size itself may have accompanied speciation in some groups, and have likely accompanied changes in life history. Selective pressures for larger body size or more rapid development can lead to changes in genome size through the relationship between a larger genome requiring a larger cell to contain it, known as nucleotypic selection (Commoner 1964; Gregory 2001). Genome size has been found to correlate positively with body size in copepods (Gregory et al. 2000; Wyngaard et al. 2005) and amphipods (Hessen and Persson 2009), and this relationship exists due to larger genomes being contained within larger cells, which will lead to larger body sizes in organisms with determinate growth. Due to high variation in microhabitat (especially depth) in Lake Baikal, different selective pressures at different depths may have allowed for expansion or contraction of the genome over evolutionary time.

Here we conduct the first investigation of genome size diversity for nearly 10% of Lake Baikal's known amphipod species, and use phylogenetically independent contrasts to examine

the relationships between genome size, body size, and maximum depth. We also examine the relationship between genome size and diversification rate across a current phylogeny.

Methods

Ethics statement

All necessary permits for collecting and handling amphipods were obtained from the respective agencies. All animals were cared for according to the Guide for the Care and Use of Experimental Animals.

Specimen collection

Specimens were collected in the 2012 and 2013 field seasons by dip net sweeps at depths 0-0.5 m, by shallow (15-100 m) and deep (400-500 m) traps baited with fish (*Coregonus* sp.) and chicken bones, by scuba diving (5-25 m), and by plankton net tows from 100m to surface for the planktonic species *Macrohectopus branickii*. Specimens were preserved either by flash freezing in liquid nitrogen or by storing in 95% ethanol for short periods of time. Jeffery and Gregory (2014) have shown that crustacean tissues preserved in ethanol for up to one year at room temperature still provide accurate genome size estimates relative to frozen samples. All specimens were identified according to Bazikalova (1945).

Genome size estimation

Genome sizes were estimated using Feulgen image analysis densitometry (FIAD) for all species in this study. For FIAD, gill tissue was removed from each specimen and disrupted mechanically in 40% (v/v) acetic acid on a glass microscope slide until no large pieces of tissue remained. The suspension was then covered with a coverslip and compressed with clothespins, and then each slide was frozen on dry ice with the clothespins still attached. After several

minutes, the coverslip was flipped off the slide using a razorblade, and the slides were then immersed in 95% ethanol for 1 minute before being allowed to air dry. All slides were then stained according to Hardie et al. (2002) using freshly prepared Schiff reagent and a 2-hour hydrolysis in 5N HCl at room temperature. Slides of chicken blood (1C=1.25pg) and rainbow trout blood (1C=2.40pg) were used as internal standards in each staining run. A minimum of 30 nuclei per slide was measured, and the mean integrated optical density (IOD) was calculated and compared to the IOD of each standard to calculate genome size. In most cases only one individual was available per amphipod species, though three pseudoreplicate slides were made per individual (Table 1).

As an additional check, genome size was also estimated using flow cytometry (FCM) for five species. FCM was only used for frozen specimens, where legs were removed and crushed using a glass pestle in 500µl ice-cold LB01 buffer (Doležal et al. 1989). Blood from chicken or rainbow trout was added to each sample tube and the cell suspensions were filtered through a 30µm nylon filter, and then treated with 3µl (6µg/ml) of RNase and 12 µl (24µg/ml) propidium iodide. Each sample was stained in the dark on ice for 1 hour. Samples were analyzed on a FC500 flow cytometer (Beckman-Coulter) using a 488nm blue laser operating at 22mW. A minimum of 1500 nuclei per sample was obtained and all CVs were <10% (Supplementary Figure 1). Three to five pseudoreplicates were run per individual.

Phylogeny construction and statistical analyses

The maximum body length for each species was measured from the largest collected individuals or taken from the literature. Maximum depths (in metres) that each species inhabits were taken from Bazikalova (1945). Linear regressions of genome size versus body size, and

genome size versus maximum depth were performed in R v.3.1.2 (R Core Team 2014). Shapiro tests for normality showed that genome size, body size, and maximum depth values deviated from a normal distribution ($p < 0.0001$, $p = 0.004$, $p < 0.0001$ respectively) and so the data were log-transformed. Phylogenetically independent contrasts (PICs) (Felsenstein 1985) were calculated for genome size, body size, and maximum depth in R using the *ape* package (Paradis et al. 2004).

The phylogeny of 66 Baikal species for which transcriptome data are available (Naumenko et al., in press) was constructed by using alignments of 175 clusters of orthologs with 100% occupancy (i.e., with a single ortholog represented in all 66 species) by means of maximum likelihood (RAxML; Stamatakis 2014) with 1000 bootstrap replicates. To obtain these alignments, paired-end Illumina reads were filtered to remove mitochondrial sequences, and assembled de novo using Trinity (Grabherr et al 2011), retaining all isoforms of a given locus for the further analysis. Assembled contigs were further checked for contamination and loci having isoforms with top hits against the NT database outside of Arthropoda were excluded. The longest protein coding sequence for each locus was used for orthology inference using OrthoMCL pipeline (Li et al 2003) and orthologous protein-coding sequences were aligned with mafft (Kato and Standley 2013). The best fit substitution model was determined by jmodeltest2 (Darriba et al. 2012; Guindon and Gascuel 2003) and GTR+ Γ +I, the best scoring model, was used for phylogeny reconstruction. For the purpose of this study the resulting tree was trimmed to include only the species for which the genome size is available using drop.tip() function of *ape* R package (Supplementary Figure 2). No polytomies existed in the tree and so 32 independent contrasts were calculated. A linear regression was then plotted using the PICs for

genome size versus body size, and genome size versus maximum depth. PICs were run on both non-transformed and log-transformed data.

We then aimed to investigate a relationship between genome size and diversification rate. The R packages *phytools* (Revell 2012) and *geiger* (Harmon et al. 2008) were used to test for phylogenetic signal of genome size using Pagel's lambda (Pagel 1999) across the phylogeny. We then divided the phylogeny into 12 clades (Supplementary Table 1) and estimated diversification rate as $\text{rate} = \log_2(N)/T$, where N is the number of described species (taken from Bazikalova 1945) and T is time since the clade's origin, which was measured as the clades' average branch lengths. We then regressed average genome size against estimated diversification rate.

Results

We estimated genome size for 36 species of amphipods from 19 genera and 7 families using FIAD and verified five of these estimates using FCM. Genome sizes estimated by FIAD ranged from 2.15 to 16.63pg (~8-fold variation). The FCM estimates were consistently higher (by ~9%) than the FIAD values (Table 1). The average genome size for the entire sample was $5.67 \pm 2.52\text{pg}$ (mean $\pm$ standard deviation) which is lower than the average value of 8.72pg for the 27 species of amphipods previously included in the Animal Genome Size Database (Gregory 2016).

A non-phylogenetically corrected linear regression of FIAD-estimated genome size versus maximum body length yielded a significant positive relationship between the two parameters ($r^2=0.406$, $p<0.0005$) and a regression of genome size and maximum depth yielded a significant positive relationship ($r^2=0.33$, $p=0.0003$). PICs revealed a positive, significant correlation between genome size and body size ($r^2=0.38$, $p=0.0001$ non-transformed; $r^2=0.29$,

$p=0.001$ log-transformed) (Figure 1) and between genome size and maximum depth ($r^2=0.31$, $p=0.0007$ non-transformed; $r^2=0.38$, $p=0.0001$ log-transformed) (Figure 2). Both relationships remained significantly positive even after removing the smallest contrast point which appears to anchor the regression.

No phylogenetic structuring of genome size was evident across the phylogeny ($\lambda=6.61^{-5}$, log likelihood=-80.60, $p=0.99$). This suggests that the large variation in genome size evolved independently multiple times and across all clades within the phylogeny. However, genome size was significantly correlated with diversification rate for the 12 clades for which we were able to obtain genome size and clade richness data ($r^2=0.44$, $p=0.018$) (Figure 3).

Discussion

Here we conducted the first survey of genome size diversity of the amphipods of Lake Baikal, Russia and in the process more than doubled the number of amphipod estimates available in the Animal Genome Size Database (Gregory 2016). We revealed nearly 8-fold variation in genome size across the 36 species of amphipods from Lake Baikal, representing approximately 10% of the known diversity of amphipods in this large, ancient lake. The mean genome size (5.67pg) is lower than the published average value (8.72pg) for amphipods in the Animal Genome Size Database, though this average is driven higher by the “genomic giants” discovered in the Arctic, as seen by an enormous standard deviation of 15.43pg (see Rees et al. 2007).

No intraspecific variation in genome size was observed aside from a 9% increase in genome size observed when using flow cytometry relative to Feulgen image analysis. This difference may simply be due to experimental error and slight differences in the fluorescence emissions and light absorbance for flow cytometry and Feulgen image analysis respectively

rather than real biological variation in genome size. However, our estimate for *Eulimnogammarus verrucosus* ranged from 6.10 to 7.13pg based on the method utilized (FIAD and FCM, respectively), which is much smaller than the value of nearly 10pg estimated by Rivarola-Duarte et al. (2014) using k-mer estimation from paired-end sequencing. While FCM and k-mer based methods have sometimes yielded similar results (e.g. Nygaard et al. 2011; Wang et al. 2014), the k-mer method is not widely accepted and has yet to be established as a reliable means of estimating genome size. Irrespective of the method used, it is also possible that there are cryptic species within Lake Baikal that differ in genome size (e.g. Väinölä and Kamaltynov 1999).

The greatest variability in genome size occurred within the family *Eulimnogammaridae*, with values ranging from $1C = 3.77$ to 16.63pg. This large amount of variability could be indicative of large-scale chromosome or genome duplication events, including polyploidization, but it bears noting that Salemaa and Kamaltynov (1994) found little variation in chromosome number ($n=26$ or $n=32$ chromosomes for 13% and 39% of the described Baikal species and genera respectively, including 12 of the genera in the present study), suggesting that differences in genome size may be the result of some other process. This is likely to be the proliferation of repetitive DNA such as transposable elements, which are more abundant in larger genomes (Elliott et al. 2015a, 2015b). Investigation of the other 87% of described species for which chromosome data does not exist will provide further evidence to either support or reject polyploidy as a mechanism which explains the differences in genome size observed here.

We found significant positive relationships between genome size and body size, and between genome size and the maximum depths where the species are found, including when controlling for phylogenetic relationships among species. For example, the morphologically

unique, small-bodied and shallow-dwelling species *Baikalogammarus pullis*, which is the only extant species in its lineage, has the smallest genome size at 2.15pg, whereas all species considered as 'abyssal' (>500m maximum depth; Takhteev (2000)) had genome sizes greater than 7pg. The single planktonic species, *Macrohectopus branickii*, has an average genome size for the Baikal flock, though no meaningful comparison can be made between this planktonic species and the other benthic species. Genome size can correlate with body size due to the positive relationship between genome size and cell size, a pattern that has been observed in several invertebrate taxa, such as polychaete annelids (Gambi et al.1997) and other crustaceans including copepods (Gregory et al. 2000; Wyngaard et al. 2005) and amphipods (Hessen and Persson 2009). A relationship between genome size and depth could exist due to differences in developmental rate, as species with larger genomes tend to develop more slowly due to longer cell division (Gregory 2002). In plants, negative relationships have been found between altitude and genome size (Reeves et al. 1998) and latitude (Grime and Mowforth 1982; Rayburn et al. 1985) due to nucleotypic selection for faster development in these harsh environments. However, in crustaceans, larger genomes appear to be more common at higher latitudes in marine species due to the high stability and low-energy of the polar oceans which could eliminate the pressure for rapid development (Rees et al. 2007). Polar oceans share some similarities, including temperature regime and oxygen content, with deep, ancient lakes such as Baikal (Martens 1997), and so a lack of selective pressure for rapid development could lead to expansion of the genome through transposable element proliferation in species at greater depths until this becomes too costly to replicate. For example, evidence in several species of gammarid amphipods suggests that the developmental time of a species is exponentially longer in colder temperatures (Steele and Steele 1973). This is, in effect, the opposite pattern found in plants at high latitudes and

altitudes as a result of nucleotypic selection. However, further testing of this hypothesis is required.

A significant, positive correlation between genome size and diversification rate was observed among 12 clades from a recently constructed phylogeny based on assembled transcriptomes. While genome size is typically thought of as a physical property affecting phenotypic characters such as cell size, genomes that become larger through transposable element and repetitive DNA proliferation, gene duplications, and polyploidy may provide the raw genetic material to lead to speciation (see Kidwell and Lisch 2001; Kraaijeveld 2010). While we provide preliminary evidence of a relationship between genome size and diversification rate, an in-depth analysis of this correlation is required across taxa to determine the actual role of genome size in clade diversification.

Overall, selective pressures for larger body sizes or to transition to different habitats at different depths may have corresponded with changes in genome size, though it is difficult to say whether changes in genome size occurred before or after these changes in organismal life history occurred. Future genome size surveys of diverse faunae such as ancient lake crustaceans will be helpful in understanding both the evolutionary history of these groups, and how changes in genome size accompany changes in phenotypes, ecological lifestyles, and biodiversity.

Acknowledgements

This work was supported by an NSERC Canada Graduate Scholarship to NWJ, an NSERC Discovery Grant to TRG and NSF grant 1136706 to LYY. We thank Anna Etingova and the crew of RV “Professor Treskov” (Baikal Museum, Russian Academy of Sciences) for help with

sampling as well as Sergey Naumenko and colleagues (Moscow State University) for providing the phylogeny used in this manuscript.

Draft

References

- Bazikalova, A. 1945. Amphipods of Lake Baikal. *Trudy Baikal'skoj Limnologicheskoy Stantsii* v. XI.
- Chapelle, G. 2002. Antarctic and Baikal amphipods: a key for understanding polar gigantism. Diss. Dr. Sci., Institut royal des Sciences naturelles de Belgique Université Catholique de Louvain.
- Chapelle, G., and Peck, L.S. 2004. Amphipod crustacean size spectra: new insights in the relationship between size and oxygen. *Oikos*, **106**: 167-175.
- Commoner, B. 1964. Roles of deoxyribonucleic acid in inheritance. *Nature* **202**: 960-968.
- Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772.
- Doležel, J., Binarová, P., and Lucretti, S. 1989. Analysis of nuclear DNA content in plant cells by flow cytometry. *Biol. Plantarum* **31**: 113-120.
- Dufresne, F., and Jeffery, N. 2011. A guided tour of large genome size in animals: what we know and where we are heading. *Chromosome Res.* **19**: 925-938.
- Elliott, T.A., and Gregory, T.R. 2015a. Do larger genomes contain more diverse transposable elements? *BMC Evol. Biol.* **15**: 69.
- Elliott, T.A., and Gregory, T.R. 2015b. What's in a genome? The C-value enigma and the evolution of eukaryotic genome content. *Philos. T. Roy. Soc. B* **370**: 20140331.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.* **125**: 1–15.

- Gambi, M.C., Ramella, L., Sella, G., Protto, P., and Aldieri, E. 1997. Variation in genome size in benthic polychaetes: systematic and ecological relationships. *J. Mar. Biol. Assoc. UK* **77**: 1045-1057.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., et al. 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **29**: 644-652. doi: 10.1038/nbt.1883.
- Gregory, T.R. 2001. Coincidence, coevolution, or causation? DNA content, cellsize, and the C-value enigma. *Biol. Rev.* **76**: 65-101.
- Gregory, T.R. 2002. Genome size and developmental complexity. *Genetica*, **115**: 131-146
- Gregory, T.R. 2016. Animal Genome Size Database. <http://www.genomesize.com>.
- Gregory, T.R., Hebert, P. D., and Kolasa, J. 2000. Evolutionary implications of the relationship between genome size and body size in flatworms and copepods. *Heredity*, **84**: 201-208.
- Grime, J.P., and Mowforth, M.A. 1982. Variation in genome size – an ecological interpretation. *Nature*, **299**: 151-153.
- Guindon, S., and Gascuel, O. 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Syst. Biol.* **52**: 696-704.
- Hardie, D.C., Gregory, T.R., and Hebert, P.D. 2002. From pixels to picograms: a beginners' guide to genome quantification by Feulgen image analysis densitometry. *J. Histochem. Cytochem.* **50**: 735-749.

- Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E., and Challenger, W. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics*, **24**: 129-131.
- Hessen, D.O., and Persson, J. 2009. Genome size as a determinant of growth and life-history traits in crustaceans. *Biol. J. Linn. Soc.* **98**: 393-399.
- Jeffery, N.W., and Gregory, T.R. 2014. Genome size estimates for crustaceans using Feulgen image analysis densitometry of ethanol-preserved tissues. *Cytometry Part A*, **85A**: 862-868.
- Kamaltynov, R.M. 1999. On the evolution of Lake Baikal amphipods. *Crustaceana*, **72**: 921-931.
- Katoh, K., and Standley, D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* **30**: 772-780.
- Kidwell, M.G., and Lisch, D.R. 2001. Perspective: transposable elements, parasitic DNA, and genome evolution. *Evolution*, **55**: 1-24.
- Kraaijeveld, K. 2010. Genome size and species diversification. *Evol. Biol.* **37**: 227-233.
- MacDonald III, K.S., Yampolsky, L., and Duffy, J.E. 2005. Molecular and morphological evolution of the amphipod radiation of Lake Baikal. *Mol. Phylogenet. Evol.* **35**: 323-243.
- Li, L., Stoeckert, C.J.Jr., and Roos, D.S. 2003. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**: 2178-2189.
- Martens, K. 1997. Speciation in ancient lakes. *Trends Ecol. Evol.* **12**: 177-182.
- Meiri, S., and Dayan, T. 2003. On the validity of Bergmann's rule. *J. Biogeogr.* **30**: 331-351.

- Melnik, N.G., Timoshkin, O.A., and Sideleva, V.G. 1995. Distribution of *M. branickii* and some characteristics of its ecology. *Guide and Key to Pelagic Animals of Baikal (with Ecological Notes)*, Nauka, Novosibirsk, 511-522. (In Russian)
- Naumenko, S.A., Logacheva, M.D., Popova, N.V., Klepikova, A.V., Penin, A.A., Bazykin, G.A., Etingova, A.E., Mugue, N.S., Kondrashov, A.S., and Yampolsky, L.Y. Transcriptome-based phylogeny of endemic Lake Baikal amphipod species flock: fast speciation accompanied by frequent episodes of positive selection. *Mol. Ecol.* In press.
- Nygaard, S., Zhang, G., Schiøtt, M., Li, C., Wurm, Y., Hu, H, Zhou J, Ji L, Qiu F, Rasmussen M, Pan H, Hauser F, Krogh A, Grimmelikhuijzen CJ, Wang J, and Boomsma JJ 2011. The genome of the leaf-cutting ant *Acromyrmex echinator* suggests key adaptations to advanced social life and fungus farming. *Genome Res.* **21**: 1339-1348.
- Pagel, M. 1999. Inferring the historical patterns of biological evolution. *Nature*, **401**: 877-884.
- Rayburn, A.L., Price, H.J., Smith, J.D., and Gold, J.R 1985. C-band heterochromatin and DNA content in *Zea mays*. *Am. J. Bot.* **72**: 1610-1617.
- Rees, D.J., Dufresne, F., Glemet, H., and Belzile, C. 2007. Amphipod genome sizes: first estimates for Arctic species reveal genomic giants. *Genome*, **50**, 151-158.
- Reeves, G., Francis, D., Davies, M.S., Rogers, H.J., and Hodgkinson, T.R. 1998. Genome size is negatively correlated with altitude in natural populations of *Dactylis glomerata*. *Ann. Bot. London*, **82**: 99-105.
- Revell, L.J. 2012. phytools: an R package for phylogenetic comparative biology and other things. *Methods Ecol. Evol.* **3**: 217-223.

Rivarola-Duarte, L., Otto, C., Jühling, F., Schreiber, S., Bedulina, D., Jakob, L., Gurkov, A., Axenov-Gribanov, D., Sahyoun, A.H., Lucassen, M., Hackermüller, J., Hoffmann, S., Sartoris, F., Pörtner, H.O., Timofeyev, M., Luckenbach, T., and Stadler, P.F. 2014. A first glimpse at the genome of the Baikalian amphipod *Eulimnogammarus verrucosus*. J. Exp. Zool. **322**: 177-189.

Salemaa, H., and Kamaltynov, R. 1994. The chromosome numbers of endemic amphipoda and isopoda: An evolutionary paradox in the ancient lakes Ohrid and Baikal. Ergebnisse der Limnologie **44**: 247-256.

Sherbakov, D.Y., Kamaltynov, R.M., Ogarkov, O.B., and Verheyen, E. 1998. Patterns of evolutionary change in Baikalian gammarids inferred from DNA sequences (Crustacea, Amphipoda). Mol. Phylogenet. Evol. **10**:160-167.

Stamatakis, A. 2014. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics **30**: 1312-1313. doi: 10.1093/bioinformatics/btu033

Steele, D.H., and Steele, V.J. 1973. The biology of *Gammarus* (Crustacea, Amphipoda) in the northwestern Atlantic. VII. The duration of embryonic development in five species at various temperatures. Can. J. Zool. **51**: 995-999.

Takhteev, V.V. 2000. Trends in the evolution of Baikal amphipods and evolutionary parallels with some marine Malacostracan fauna. Adv. Ecol. Res. **31**: 197-220.

Väinölä, R., and Kamaltynov, R.M. 1999. Species diversity and speciation in the endemic amphipods of Lake Baikal: molecular evidence. Crustaceana, **72**: 945-956.

Väinölä, R., Witt, J.D.S., Grabowski, M., Bradbury, J.H., Jazdzewski, K., and Sket, B. 2008. Global diversity of amphipods Amphipoda; Crustacea. in freshwater. *Hydrobiologia*, **595**: 241-255.

Walters, R.J., and Hassall, M. 2006. The temperature-size rule in ectotherms: may a general explanation exist after all? *Am. Nat.* **167**: 510-523.

Wang, W., Haberer, G., Gundlach, H., Gläßer, C., Nussbaumer, T., Luo, M.C., Lomsadze, A., Borodovsky, M., Kerstetter, R.A., Shanklin, J., Bryant, D.W., Mockler, T.C., Appenroth, K.J., Grimwood, J., Jenkins, J., Chow, J., Choi, C., Adam, C., Cao, X.-H., Fuchs, J., Schubert, I., Rokhsar, D., Schmutz, T.P., Michael, T.P., Mayer, K.F.X., and Messing, J. 2014. The *Spirodela polyrhiza* genome reveals insights into its neotenuous reduction fast growth and aquatic lifestyle. *Nature Comm.* **5**: 3311.

Wyngaard, G.A., Rasch, E.M., Manning, N.M., Gasser, K., and Domangue, R. 2005. The relationship between genome size, development rate, and body size in copepods. *Hydrobiologia*, **532**: 123-137.

Figure Captions

Figure 1. Log-transformed phylogenetic independent contrasts of body size versus genome size for 32 species of amphipods from Lake Baikal, showing a significant, positive relationship ($r^2=0.29$, $p=0.001$).

Figure 2. Log-transformed phylogenetic independent contrasts of maximum depth and genome size for 32 species of amphipods from Lake Baikal, showing a significant, positive relationship ($r^2=0.38$, $p=0.0001$).

Figure 3. A positive relationship ($r^2=0.44$, $p=0.018$) exists between average genome size and broad speciation rate, measured as $\text{rate}=\log_2(N)/T$, where N is the number of species within the clade, and T is the time since the clade's origin.

Table 1. A list of all species collected and their genome size estimates from FIAD and FCM in picograms (pg) along with their maximum body sizes and the maximum depth they are found at. For all FCM estimates n=1 with 3-5 pseudoreplicates.

Species	Mean FIAD Genome size (pg) ± SE	n=	Mean FCM GS (pg) ± SE	Max. Body size ^a (mm)	Max. Depth ^b (m)
<i>Acanthogammarus brevispinus</i>	5.82	1		36	70
<i>Acanthogammarus godlewskii</i>	8.32±0.25	2	8.52	50	180
<i>Acanthogammarus victorii</i>	5.43	1		67	90
<i>Baikalogammarus pullis</i> ^c	2.15	1		5	10
<i>Brachiuropus grewingkii</i>	7.94	1		90	1300
<i>Brandtia latissima</i>	6.13	1		18	30
<i>Carinurus bicarinatus</i>	7.24	1	8.15	30	200
<i>Cornugammarus maximus</i>	5.94	1		67	121
<i>Crypturopus inflatus</i>	4.93	1		16	150
<i>Dorogostaiskia parasitica</i>	6.30	1		11.5	60
<i>Echiuropus macronychus</i>	3.33	1		12	50
<i>Eucarinogammarus wagii</i>	5.39	1		57	400
<i>Eulimnogammarus cyaneus</i>	3.82±0.07	4		15	5
<i>Eulimnogammarus czerskii</i>	5.69	1		30	25
<i>Eulimnogammarus marituji</i>	3.77	1		20	10
<i>Eulimnogammarus melanochlorus</i>	4.22	1		32.5	10
<i>Eulimnogammarus similis</i>	4.72	1		15	26
<i>Eulimnogammarus verrucosus</i>	6.10±0.29	2	7.13	36	10
<i>Eulimnogammarus violaceus</i>	5.65	1		30	30
<i>Eulimnogammarus vittatus</i>	3.95±0.06	2		18	30
<i>Garjajewia cabanisi</i>	10.57±0.16	2		80	1250

<i>Gmelinoides fasciatus</i>	3.68	1		15	10
<i>Hyalellopsis carinata</i>	5.85	1		10	40
<i>Linevichella vortex</i>	4.71	1		7	10
<i>Macrohectopus branickii</i>	4.20	1		30	250 ^d
<i>Micruropus parvulus</i>	3.66	1		3	10
<i>Micruropus wohli</i>	3.94±0.11	2		9	4
<i>Odontogammarus calcaratus</i>	4.10	1		30	200
<i>Palicarinus puzylli</i>	5.64	1		50	200
<i>Pallasea brandtii</i>	7.60	1	8.58	30	60
<i>Pallasea cancellus</i>	4.56	1		70	52
<i>Pallasea cancelloides</i>	5.21	1		25	178
<i>Pallasea kessleri</i>	6.26	1		33	61
<i>Parapallasea borowskii</i>	7.44	1	7.84	55	600
<i>Poekilogammarus pictoides</i>	3.18	1		30	30
<i>Sluginella kietlinskii</i>	16.63	1		81	630

Notes: ^a from Bazikalova (1945) or measured body lengths of largest specimens found

^b from Bazikalova (1945)

^cOnly one slide was made from a single individual due to its small size

^dPlanktonic species; maximal depth unknown

Figure 1. Log-transformed phylogenetic independent contrasts of body size versus genome size for 32 species of amphipods from Lake Baikal, showing a significant, positive relationship ($r^2=0.29$, $p=0.001$).

Figure 2. Log-transformed phylogenetic independent contrasts of maximum depth and genome size for 32 species of amphipods from Lake Baikal, showing a significant, positive relationship ($r^2=0.38$, $p=0.0001$).

Figure 3. A positive relationship ($r^2=0.44$, $p=0.018$) exists between average genome size and broad speciation rate, measured as $\text{rate}=\log_2(N)/T$, where N is the number of species within the clade, and T is the time since the clade's origin.

Supplementary Tables and Figures

Table 1. Clades defined from the constructed phylogeny (Figure S1) for which we had genome size data. The number of species currently defined per clade, as well as the average genome size and branch length for the clade are listed.

Clade	Nsp broad	Average Genome Size (pg)	Branch Length (ds)
1) <i>Baikalogammarus</i>	1	2.15	0.0507
2) <i>Micruropus et al</i>	49	4.19	0.0357
3) <i>Hyallelopsis</i>	13	5.85	0.01448
4) <i>Echiuropus</i>	3	3.33	0.0183
5) <i>Eucarinogammarus et al</i>	6	5.25	0.0129
6) <i>Eulimnogammarus (Phylolimnogammarus)</i>	13	4.37	0.00675
7) <i>Eulimnogammarus (s. str)</i>	10	5.69	0.004
8) <i>Acanthogammarus et al</i>	24	6.78	0.00898
9) <i>Pallasea, (Spinacanthus, Brandtia)</i>	32	5.61	0.0161
10) <i>Parapallasea</i>	4	6.54	0.0046
11) <i>Garjajewia</i>	4	10.57	0.0074
12) <i>Sluginella (Corophiomorphus)</i>	10	16.63	0.0035

Figure S1. An example flow cytometry histogram for the amphipod *Eulimnogammarus verrucosus* (B) relative to the fluorescent signal of chicken blood *Gallus domesticus* (A) with a reference genome size of 1.25pg. The channel ‘FL3’ is the fluorescence channel for propidium iodide which was used to stain the nuclei.

Figure S2. A transcriptome-based phylogeny of Lake Baikal amphipods used to conduct phylogenetic independent contrasts in the present study. Provided by Naumenko et al. (in prep.).