Test of the correlation between body size and DNA content in *Pimelia* (Coleoptera: Tenebrionidae) from the Canary Islands

MIQUEL PALMER1,2*, EDUARD PETITPIERRE1,2 and JOAN PONS3

1IMEDEA (CSIC-UIB) Instituto Mediterráneo de Estudios Avanzados, Miquel Marquès, 21, E-07190 Esporles (Illes Balears), Spain; e-mail: ieampv@uib.es
2Department Biologia Ambiental. Lab. Genética, Campus Univ. Illes Balears, Cta Valldemossa km 7,5, E-07071 Palma de Mallorca, Spain
3The Natural History Museum, Department of Entomology, Cromwell Road, London SW7 5BD, UK

Key words. DNA content, comparative methods, phylogenetic inertia, satellite DNA, nucleotypic effects, microdensitometry, image analysis, *Pimelia* (Tenebrionidae, Coleoptera)

Abstract. Comparative analyses of interspecific data in evolutionary biology usually require specific methods to remove the effects of phylogenetic inertia. When phylogenetic inertia is not considered, the Canarian *Pimelia* species show a positive, and almost significant (Prob. = 0.066) correlation between nuclear genome size and body size. However, after controlling for phylogenetic inertia there was a negative and significant correlation (Prob. = 0.007 to 0.017, depending on the DNA fraction considered). Such a change in the relationship after controlling for phylogenetic inertia is rarely reported. Moreover, the relationship usually reported is positive and thought be a consequence of species having a similar number of cells at the same stage of development. The aim of the present study is to report a case of a negative correlation, but not to explain the causal mechanism involved in genome size variations or propose a formal hypothesis on the specific links between DNA content and body size. However, a common explanation of the change in the relationship, i.e., positive to negative, is suggested. Moreover, the data available on the highly repetitive, non-coding satellite DNA allows us to analyse the specific pattern exhibited by this fraction.

INTRODUCTION

Nucleotypic effects of genome size are defined as the effects exerted by the total amount of nuclear DNA on the phenotypic characters. They occur at both the cellular and whole-organism level (Bennett, 1971; Cavalier-Smith, 1985). The positive association between genome size and cell size, or between genome size and length of cell cycle is well documented in both plants or animals (Sparrow et al., 1965; Bennett, 1972; Bennett, 1973; Olino, 1983; Juan & Petitpierre, 1989a; Petitpierre et al., 1993; Petitpierre et al., 1998). A theoretical explanation of the positive correlation between genome size and cell cycle length was suggested by Vinogradov (1998): the non-coding DNA binds DNA-specific proteins and buffers their concentration. The benefit linked to such a passive-homoeostasis mechanism has also the inertial cost associated with retarding the nuclear machinery, which in turn results in longer cell cycles.

However, there is no clear relationship between genome size and developmental time at the organismic level. No such relationship seems to exist in either birds or mammals (John & Miklos, 1988; Monaghan & Metcalfe, 2000). But a positive correlation was proposed for other vertebrates (Sessions & Larson, 1987; Jockusch, 1997), and invertebrates (Ferrari & Rai, 1989; White & McLaren, 2000). The body mass-independent basal metabolic rate is negatively correlated with genome size (Vinogradov, 1995; 1997). Genome size is also known to have negative effects on the number of generations per year in some leaf beetles (Petitpierre & Juan, 1994). Two opposite relationships between body size and genome size have been described. Palmer & Petitpierre (1996) found a significant negative correlation between body size and nuclear DNA content in populations of a single species of darkling beetle (Coleoptera, Tenebrionidae). It was suggested that large genome sizes reduce the metabolic rate and hence increase the duration of the cell cycle, which results in small body size (the implicit assumption is an overall reduction in cell number). Conversely, a large number of studies report a positive correlation between genome and body size. This positive correlation is thought to be a consequence of the similar cell number in similar developmental stages in related species (White & McLaren, 2000). Hinegardner (1974) reported a positive correlation in molluscs, McLaren et al. (1988; 1989) in copepods (Crustacea) and Fox (1972) in dermestids (Coleoptera). Other studies recording positive correlations are those of Ferrari & Rai (1989), Finston et al. (1995) and Gregory et al. (2000).

The aim of the present paper is to report a case of a negative correlation between body size and genome size. We also explore the role that highly repetitive, non-coding satellite DNA plays in this relationship.

* Corresponding author.
Almost all (13 out of 14) of the Canarian species or subspecies of *Pimelia* F., 1775 were included in this analysis. Details of the geographical sources and number of specimens measured are given in Table 1. Most of the material is in the Museu de la Naturalista de les Illes Balears (MNIB, Palma de Mallorca, Spain). Two-dimensional coordinates of six landmarks (Rohlf & Marcus, 1993) on every specimen were determined using a Wild dissecting microscope, a video camera, and VIDAS21 image analyser software (Kontron Elektron, Munich). Landmarks were on the elytron (namely, left and right points of maximum width, left and right anterior-lateral angles, sutural apical angle, and basal middle point). The size of a specimen was estimated using their centroid size (i.e., the square root of the sum of the squared distances from each landmark to the centre of gravity; Rohlf & Marcus, 1993; Slice et al., 1996). Each specimen was measured three times and the three values of the centroid size obtained were then averaged. The centroid size for a species was determined by averaging the centroid sizes of all the specimens measured.

The reliability of centroid size as a body size measurement was tested using an independent database. Traditional (i.e., point to point distances) measurements of 12 body traits of ten of the species measured. The reliability of centroid size as a body size measurement was tested using an independent database. Traditional (i.e., point to point distances) measurements of 12 body traits of ten of the species measured.

**Measuring size**

Almost all (13 out of 14) of the Canarian species or subspecies of *Pimelia* F., 1775 were included in this analysis. Details of the geographical sources and number of specimens measured are given in Table 1. Most of the material is in the Museu de la Naturalista de les Illes Balears (MNIB, Palma de Mallorca, Spain). Two-dimensional coordinates of six landmarks (Rohlf & Marcus, 1993) on every specimen were determined using a Wild dissecting microscope, a video camera, and VIDAS21 image analyser software (Kontron Elektron, Munich). Landmarks were on the elytron (namely, left and right points of maximum width, left and right anterior-lateral angles, sutural apical angle, and basal middle point). The size of a specimen was estimated using their centroid size (i.e., the square root of the sum of the squared distances from each landmark to the centre of gravity; Rohlf & Marcus, 1993; Slice et al., 1996). Each specimen was measured three times and the three values of the centroid size obtained were then averaged. The centroid size for a species was determined by averaging the centroid sizes of all the specimens measured.

The reliability of centroid size as a body size measurement was tested using an independent database. Traditional (i.e., point to point distances) measurements of 12 body traits of ten of the species measured. The reliability of centroid size as a body size measurement was tested using an independent database. Traditional (i.e., point to point distances) measurements of 12 body traits of ten of the species measured.

**Measuring DNA content**

The relative nuclear DNA content was measured using microdensitometry of Feulgen-stained spermatids (Juan & Petitpierre, 1989b and Alvarez-Fuster et al., 1991). At least 50 (but less in a few cases; Table 1) round and homogeneously stained spermatids were measured per individual. So stained, the amount of DNA content of a spermatid (e.g., MacLeish & Sunderland, 1961; Rasch, 1985). Light transmission and spermatid area were measured using a microscopic (Zeiss Axioskop, 1600x) video-image. Images were processed by a VIDAS21, which measures area and light transmission (through grey intensity) through a spermatid. Spermatid area was determined by automatic edge selection using a threshold-based procedure. For reference, a number of spermatids of *Tribolium castaneum* Hbst. (mean genome size: 0.208 ± 0.002 pg, Juan & Petitpierre, 1989a) were treated and stained in the same way. This value was used to obtain absolute estimates of the DNA content (Alvarez-Fuster et al., 1991; Palmer, 1994; Palmer & Petitpierre, 1996). Satellite DNA is almost universally present in eukaryotic genomes. It consists of tandemly repeated sequence clusters, which can be several megabases in length (Elder & Turner, 1995). After isolating DNA from adults using standard phenol extraction and ethanol precipitation procedures (Sambrook et al., 1989), the presence of satellite DNA was determined by digestion of whole genomic DNA with restriction enzymes able to recognize and cut within the repeat unit. After electrophoresis of digested genomic DNA on an agarose gel, satellite DNA can be observed as a ladder of oligomer restriction fragments (i.e., the satellite DNA). The relative amount of satellite DNA in *Pimelia* species was determined from electrophoreses of Eco RI and Hind III digestions of genomic DNA on 1.5% agarose gels. The digitalization and densitometric measurement of the gels were performed with the aid of the Sun View program (Pharmacia). The satellite DNA percentage was estimated using the densitometric percentage of bands with respect to the whole DNA (Plohl & Ugarov, 1994).

**Data analysis**

The aim of the present paper is to test for relationships between body and genome size including species phylogeny within the statistical framework. Phylogenetic relationships of the species used in this study have been studied by Juan et al. (1995) and Pons et al. (2002). Both studies present the same phylogenetic tree, which we use in this study along with the branch lengths inferred from the data analysed by Pons et al. (2002). The phylogenetic hypothesis was obtained using neighbour-joining with maximum likelihood (Fig. 1). Phylogenetic distance between two species was estimated using the sum of all branches connecting the two species (i.e., additive distances).

We assessed the relationship between body and genome size using a newly developed method for comparative analysis: the phylogenetic eigenvalue regression (Diniz-Filho et al., 1998). Standard correlation analysis requires that observations (i.e., species) should be independent (Sokal & Rohlf, 1981). Some comparative methods circumvent the problem of statistical non-independence between species (i.e., phylogenetic relatedness) by using the information on tree topology and branch length to construct a set of sequential comparisons (or contrasts) between species and nodes (Felsenstein, 1985; Purvis & Rambaut, 1995). Alternatively, the phylogenetic eigenvalue regression (PVR; Diniz-Filho et al., 1998) used here derives an eigenvector matrix that summarises phylogeny. Such a matrix is obtained by principal coordinate analysis, on the double-centred pairwise phylogenetic distance matrix. The number of eigenvectors retained is determined by a broken-stick model (i.e., comparing the variance explained by the actual eigenvectors to that expected by chance; Legendre & Legendre, 1998). Effective removal of phylogenetic inertia was tested on the residuals after regressing the trait under study on the eigenvector matrix. The existence of a phylogenetic autocorrelation in the residuals was investigated using a phylogenetic correlogram. It involves 1) definition of some classes of phylogenetic distances and 2) calculation of a statistic (Moran's I coefficient, Gittleman & Kot, 1990) that measures the (within-class) covariance of the variable of interest. Three classes are defined (class marks: 0 to 0.06, 0.6 to 0.12 and more than 0.12). Positive Morans' I coefficient for a class indicates that the species belonging to species pairs falling
within this class tend to deviate from the mean body size (or DNA content) in the same direction.

The relationship between nuclear DNA content and body size was determined by partial multiple regression using CANOCO 4 (ter Braak & Smilauer, 1998). The model used allows us to assess the (partial) effects of an explanatory matrix (DNA content) on a response vector (i.e., body size), using the phylogenetic eigenvector matrix as matrix of covariables. Four variables related to DNA content were tested in turn: total DNA (1C DNA content, pg), satDNA (satellite DNA, pg), %satDNA (percentage of satDNA, squared arc-sin transformed), and non-satDNA (the fraction of total DNA remaining after subtracting non-satDNA; its biological significance will be discussed below). A variance decomposition between the explanatory variables and the matrix of covariables was performed as suggested by Legendre & Legendre (1998) and ter Braak & Smilauer (1998). The significance of the variance fractions reported by specific explanatory variables was determined by a Monte Carlo iteration procedure (ter Braak & Smilauer, 1998).

RESULTS

The 1C DNA values for all the specimens studied are given in Table 1. They show a rather wide range (from 0.228 pg in P. sparsa sparsa to 0.747 pg in P. laevigata costipennis). Interspecific body size differences are also noticeable (Table 1), ranging from a centroid size of 10.4 (P. sparsa albohumeralis) to 16.9 (P. laevigata validipes). Note that centroid size is given in mm and not mm$^2$. Moreover, it should be noted that the absolute value depends on the number of landmarks.

Positive relationships were found between DNA content and body size when a standard correlation analysis was used (i.e., when species were assumed to be statistically independent points). The best correlation was obtained when the two variables were ln-transformed ($r = 0.52$; $n = 13$; Prob. = 0.066; Fig. 2A). Note, however, that this nearly significant positive correlation is dependent on two outlying points.

Moreover, the traits under study (i.e., body size and DNA content) showed significant Moran’s I coefficients at the first distance class, and smaller values for successive classes. This indicates that closely related species (phylogenetic distance less than 0.06) tend to be similar (in terms of deviation from the mean), and that similarity decreases with increase in phylogenetic distance (Fig. 3).

A matrix of phylogenetic divergences between species was estimated from the pairwise additive distances on the phylogenetic tree (Fig. 1), and used as the input in a phylogenetic eigenvector regression (PVR, Diniz-Filho et al., 1998). Principal coordinate analysis of this matrix after double-centering extracted four eigenvectors with eigenvalues higher than those expected by chance. The eigenvectors summarize the phylogeny well, explaining 85.8% of the variability, and the correlation between the original matrix of phylogenetic divergence and the matrix of Euclidean distances derived from the four eigenvectors was high (Mantel test: $r = 0.95$; Prob. $< 0.002$).

The regression analysis of body size on these four eigenvectors was significant (body size: $r^2 = 0.72$, Prob.$ = 0.023$; DNA content: $r^2 = 0.94$, Prob.$ < 0.001$). The residuals of this multiple regression describe the variation in body size after removing the phylogenetic inertia. The correlogram of these residuals was not significant, demonstrating that phylogenetic inertia was successfully removed after PVR (Fig. 3).

The relationship between body size and the different fractions of the DNA content was determined by multiple regression, including as covariables the four eigenvectors describing the phylogeny. The best predictor of body size was non-satDNA content. The model including this variable and phylogeny explained 90.0% of the variance in body size. The negative relationship between body size and non-satDNA content after removing the phylogenetic inertia is seen when the corresponding residuals are plotted (Fig. 2B). Partial multivariate regression indicate that this negative relationship is significant (Prob.$ = 0.007$; Table 2). Variance decomposition between non-satDNA content and the matrix of the four eigenvectors describing the phylogeny is detailed in Table 2. The significance of

\begin{table}
\centering
\begin{tabular}{lcccccc}
\hline
Species & Centroid size & SD & n & PCA & DNA & SD & n & %satDNA \\
\hline
P. canariensis Brullé 1838 (CAN) & 16.068 & 0.968 & 6 & 0.113 & 0.486 & 0.037 & 100 & 38.1 \\
P. estevezi Oromi 1990 (EST) & 16.438 & 0.835 & 9 & - & 0.348 & 0.027 & 59 & 31.3 \\
P. fernandez-lopezi Machado1979 (FER) & 13.939 & 0.831 & 5 & - & 0.414 & 0.039 & 50 & 27.0 \\
P. granulicollis Wollaston 1864 (GRA) & 16.134 & 1.312 & 17 & 0.885 & 0.324 & 0.029 & 53 & 30.5 \\
P. laevigata costipennis Wollaston 1864 (LCO) & 14.582 & 0.922 & 20 & -0.178 & 0.747 & 0.065 & 30 & 36.3 \\
P. laevigata laevigata Brullé 1838 (LLA) & 14.882 & 0.790 & 5 & -0.472 & 0.646 & 0.036 & 55 & 35.5 \\
P. lataria Brullé 1838 (LUT) & 15.282 & 1.314 & 8 & 0.520 & 0.589 & 0.043 & 50 & 36.0 \\
P. laevigata validipes Wollaston 1864 (LVA) & 16.851 & 0.964 & 16 & 1.099 & 0.639 & 0.058 & 59 & 35.9 \\
P. radula ascendenso Wollaston 1864 (RAS) & 15.570 & 0.404 & 5 & 0.657 & 0.438 & 0.037 & 50 & 34.2 \\
P. radula granulata Wollaston 1864 (RGR) & 15.373 & 0.563 & 12 & -0.258 & 0.521 & 0.038 & 50 & 32.1 \\
P. sparsa albokuneralis Lindberg 1947 (SAL) & 10.427 & 0.571 & 8 & - & 0.246 & 0.027 & 6 & 28.9 \\
P. sparsa servimargo Wollaston 1864 (SSE) & 11.463 & 0.861 & 15 & -2.443 & 0.278 & 0.023 & 30 & 29.3 \\
P. sparsa sparsa Brullé 1838 (SSE) & 14.614 & 0.756 & 15 & 0.077 & 0.228 & 0.016 & 60 & 26.3 \\
\hline
\end{tabular}
\caption{Morphometric and cytogenetic data for the 13 Pimelia species. Morphometric data are centroid size (mm), standard deviation of centroid size (SD), sample size (n), and first factor scores extracted from Principal Component Analysis of 12 body traits (PCA). Cytogenetic data are 1C DNA contents (DNA, in pg), standard deviation of 1C DNA contents (SD), sample size (n, spermatids measured), and percentage of satDNA (%satDNA).}
\end{table}
Fig. 2. Apparent effects of genome size on body size in 13 species of Canarian *Pimelia*: A – the relationship of the non-transformed data; B to D – the relationship of body size to each of three explanatory variables after removal of the effects of phylogenetic inertia (B: total DNA, C: non-satDNA, D: satDNA).

The variance fractions explained by the other putative explanatory variables are also indicated in Table 2. Partial regression of size on total DNA indicate the same significant negative relationship (Prob. = 0.017; residuals are shown in Fig. 2C). Conversely, partial regression of size on satDNA was not significant, either using absolute values (Prob. = 0.146; residuals are shown in Fig. 2D) or percentages (Prob. = 0.70).

**DISCUSSION**

Many cases of a positive correlation between genome size and body size have been reported in invertebrate animals. Some of these have been obtained by comparing different species without considering their phylogenetic background, but others have used some type of statistical correction for non-independence of samples (e.g., McLaren et al., 1989). Therefore, the existence of a positive correlation appears well supported.

DNA content and body size of *Pimelia* from the Canary Islands also appear (nearly but not significantly) positively correlated when species phylogeny is not considered. However, an analysis of the data when phylogenetic inertia is taken into consideration does not support the existence of a positive correlation. Instead, body size and DNA content (total DNA and non-satDNA) are significantly, negatively, correlated after adjusting for phylogeny. In contrast, the relationship between satDNA and body size is clearly non-significant. This is an unexpected result because satDNA accounts for a large percentage

<table>
<thead>
<tr>
<th>Source</th>
<th>%variance</th>
<th>F-value</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-satDNA</td>
<td>18.3</td>
<td>13.6</td>
<td>0.007</td>
</tr>
<tr>
<td>Phylogenetic inertia</td>
<td>71.8</td>
<td>53.5</td>
<td>0.002</td>
</tr>
<tr>
<td>Residual</td>
<td>9.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total DNA</td>
<td>15.8</td>
<td>9.3</td>
<td>0.017</td>
</tr>
<tr>
<td>SatDNA</td>
<td>7.9</td>
<td>2.8</td>
<td>0.146</td>
</tr>
</tbody>
</table>
Fig. 3. Phylogenetic correlogram of interspecific variation in body size in 13 species of the Canarian Pimelia, using Moran’s I coefficients at three distance classes of pairwise phylogenetic distance (open points: untransformed data) and residuals (PVR residuals; solid points). High (significantly different from zero) Moran’s I value at class 1 denotes that closely related species tend to display a similar body size (or similar DNA content). These trends are removed using the residuals of PVR.

(e.g., 26 to 38% in the species studied) of the total DNA, with noticeable interspecific variability (Pons, 1999).

Another case of negative correlation between genome size and body size was previously reported in the species Phylan semicostatus Mulsant and Rey (Coleoptera, Tenebrionidae). Palmer & Petitpierre (1996) took advantage of the island distribution of this flightless species to control for phylogeny in testing correlates of genome size. In this particular case, a rising sea level after the last ice-age partitioned a single population into eight insular populations, each residing on a different island (Cuerda, 1975). These populations thus provide nearly independent samples for a test of the body size-genome size correlation.

The evolution of body size is affected, constrained and modulated by a number of environmental factors (e.g., Damuth & MacFadden, 1990). Moreover, species living on islands (such as the Canarian Pimelia) can experience important changes in body size (Heaney, 1978), linked, for example, to resource availability (e.g., Brown and Lomolino, 1998). It appears that the genus Pimelia colonised the Canary Islands from North Africa over a span of 8 My, in a series of stepping-stone colonisations (Juan et al., 1995; Juan et al., 2000). Therefore, every colonisation by Pimelia has probably produced specific changes in body size depending on the environmental characteristics of the newly colonized island. Consequently, the amount of DNA could be not only a target for selection, but also a tool (among others) for adjusting body size to a specific environment.

As stated above, we do not propose a formal hypothesis on the specific link between DNA content and body size, but delineate a possible common explanation for different patterns of correlation between genome size and body size. An explanation for the positive correlation that exists, for example, in some copepods, is straightforward because these organisms have similar numbers of cells and developmental times. Therefore, body size and genome size are necessarily positively correlated (McLaren et al., 1989). However, one can easily imagine an (obviously oversimplified) example where genome size, cell size, and cell cycle duration are all linearly and positively related. In that case, when 1) genome size increases and 2) development time remains the same, then cell number decreases, and there is no change in body size (i.e., genome size and body size would become uncorrelated). And by way of an example, a negative correlation would emerge in the same scenario from an exponential relationship between development time and cell cycle length. Therefore, the specific negative correlation found in the case of Pimelia will be explained only after exploring the complex web of relationships between genome size, cell size, cell cycle length, development time and body size, and more importantly, their relative scaling.

Finally, the possibility that the observed correlation is non-adaptive should also be considered. For example, an increased genome size may simply be due to increased transposable element activity. This increased activity may be slightly deleterious and lead to smaller body size (Heartl, 2000).

ACKNOWLEDGEMENTS. We thank José Alexandre Felizola Diniz-Filho, Tonyo Alcover, Enric Descals, Carlos Juan, Jesus Gómez-Zurita, Stuard Longhorn and Pere Oromi, and two anonymous referees for their suggestions. This research was supported financially by the scientific project BOS2000-0822, from the Ministry of Science and Technology (Spain).

REFERENCES


HEANEY L.R. 1978: Insular area and body size of insular mammals: Evidence from the tri-coloured squirrel (Callosciurus) and the dwarfing effect of insular mam­mals. Evolution 32: 29–44.


Received April 8, 2002; revised August 14, 2002; accepted October 10, 2002