

**ALTERNATE CHAPTER – NOT INCLUDED IN FINAL DRAFT**  
(T. Ryan Gregory, unpublished)

**- Chapter XXX -**

**Genome sizes of miscellaneous invertebrates**

**- Chapter XXX -**  
**Genome sizes of miscellaneous invertebrates<sup>1</sup>**

**Abstract**

Only a few groups of invertebrates other than arthropods have been studied in terms of genome size. The majority of non-arthropod C-value estimates have been from polychaete annelids, echinoderms, molluscs, and turbellarian flatworms, and a great many phyla have yet to be evaluated. This chapter provides a summary of the genome sizes currently available for the remaining invertebrate taxa and includes 55 new measurements, most from groups that have not been studied previously.

---

<sup>1</sup>

This chapter consists largely of material not yet published, but contains information from Gregory et al. (2000) and Gregory and Hebert (2002).

## **Introduction**

Thus far, the discussion of invertebrate genome sizes has focused exclusively on three large groups of arthropods, the crustaceans (Chapter 8), insects (Chapter 9), and spiders (Chapter 10). Various other groups of invertebrates have been the subject of genome size study in the past, although in most cases the coverage of invertebrate phyla remains slight. This chapter provides a summary of the patterns observed in the less commonly studied invertebrates, based on available published data as well as new estimates for some previously neglected groups.

### **Summary of the dataset(s)**

#### *Previously published data*

To date, nearly 420 genome sizes have been reported for invertebrates other than crustaceans, insects, and spiders. Most (> 80%) of these have been from polychaete annelids, echinoderms, molluscs, and turbellarian flatworms. Invertebrate groups besides these have generally been very poorly studied (Table 11.1). Previously published data used in this chapter are grouped by taxa in a series of appendices, as outlined below.

Taxonomic updates for this chapter relied heavily on the *Integrated Taxonomic Information System* database (<http://www.itis.usda.gov/>).

#### *New genome size estimates*

As with the insects (Chapter 9), the present survey represents an attempt to address some of the major gaps in the current invertebrate genome size dataset. To this end, new estimates are given for 55 different species, providing at least preliminary coverage to several phyla that had not yet been examined (Table 11.1; Appendix 11.1). Most of the sample preparation and measurement details for these groups were outlined

in Chapter 7, and only a few additional comments are provided below as necessary. Information regarding the cell types measured, standards used, and collection information are provided in the appendices listed in the following sections. Many of the specimens used in this analysis were provided by colleagues whose research programs require information on genome size of one or a few species. In many cases, these measurements were solicited by the researchers in question, and at other times material was provided at my request. In both cases, the contributions of many interesting and previously unstudied species is gratefully acknowledged.

### **Patterns of variation**

Many of the invertebrate taxa treated in this chapter are represented in Fig. 1.3, which provides a basic overview of the variation among the best-known groups of organisms. Simple citations of absolute ranges cannot provide an adequate picture of genome size variation, however, and as such the following sections provide some more detailed insights into the patterns found in the major groups of invertebrates not yet discussed. By necessity, these discussions are largely descriptive in nature, although an element of statistical rigour has been infused wherever possible. The major groups are treated here in order of decreasing amounts of available data.

#### *Molluscs*

The Mollusca represent the most speciose aquatic animal phylum, and the second largest group of animals overall. Molluscs also make up one of the best studied invertebrate groups with regard to absolute number of C-value estimates, behind the Crustacea and Insecta (but note that all three are very poorly covered in relative terms). Thus far, about 170 species of mollusc genome sizes have been assessed, including about

80 species each of bivalves and gastropods (Appendix 11.2). Cephalopods have been poorly studied, and are currently represented by only five species. Seven species of chiton have also been analyzed to date. Overall, molluscan C-values range nearly 14-fold – with the entire range found among gastropods – from about 0.4pg in the owl limpet *Lottia gigantea* to 5.9pg in the Antarctic whelk *Neobuccinum eatoni*. The average C-value for all molluscs is  $1.8\text{pg} \pm 0.07$  (but again, this is based almost entirely on gastropods and bivalves).

The large genome of *N. eatoni* has been attributed to unspecified selective pressures operating in the harsh environment of Antarctic continental waters (Libertini et al. 1993), but relatively high genome sizes also appear to characterize whelks, mud snails, and cone shells (Order Neogastropoda) in general, and similar values are found in some cephalopods (Appendix 11.2). Other prosobranch gastropods, including abalones and keyhole limpets (Order Archaeogastropoda), applesnails (Order Architaenioglossa), periwinkles (Order Neotaenioglossa), and limpets (Order Patellogastropoda), typically display much smaller genomes (means of 1.6pg, 1.1pg, 1.2pg, and 0.61pg, respectively). The opisthobranchs display a similar mean genome size of about 1.7pg. Notably, a small sample of unidentified snails inhabiting sub-oceanic thermal vents (an environment at least as “harsh” as the Antarctic) did not appear to differ greatly from related coastal species (Dixon et al. 2001).

Among the pulmonate gastropods, it is apparent that terrestrial snails and slugs (Order Stylommatophora) possess larger genomes than their freshwater relatives (Order Basommatophora), as has been pointed out by Vinogradov (2000). In fact, the terrestrial pulmonates studied so far display an average genome size roughly twice that of

freshwater forms<sup>2</sup> (2.7 vs. 1.4pg; Appendix 11.2). Here, Vinogradov (2000) notes the similarity with the enlargement of genome sizes in lungfishes and amphibians by arguing that “this independent phylogenetic branch shows a parallelism to the vertebrates, indicating that genome enlargement in the land pioneers is not incidental”. However, this similarity is only superficial, and probably not produced by similar causes, as shown by several observations. First, terrestrial gastropod genomes are only slightly (2x) larger than those of freshwater species, whereas lungfish and urodele C-values are, in some cases, two orders of magnitude larger than those of the presumed comparison group, the teleost fishes (Chapters 4 and 5). Second, lungfishes and urodeles have the largest genomes of any vertebrates (and indeed, of any animals), but terrestrial pulmonate genomes are only half the size of those of some marine molluscs. Third, the increase in genome in amphibians and lungfishes is known to be a secondary feature, almost certainly related to a simplification of development (Chapter 4), and not to any aspects of a terrestrial lifestyle *per se*. Nevertheless, the pattern of increase genome size among terrestrial pulmonates is an interesting one worthy of more detailed study, even if the resemblance to the vertebrate case is of no functional relevance.

---

2

No molluscan genome models currently exist, and in light of their smaller genomes it seems reasonable that the first one chosen should be a freshwater pulmonate snail. To this end, a growing consortium of authors led by Coen Adema and Eric Loker (University of New Mexico) has been lobbying to make the freshwater snail *Biomphalaria glabrata* the “fruit fly” of molluscan genomics. At the request of these colleagues, I performed genome size estimates on haemocytes taken from 14 individuals of three different strains (M-line, 13-16R1, and BS90), and found the C-value to be  $0.95\text{pg} \pm 0.01$ . This relatively small genome size should prove tractable for a genome sequencing project, and there are many good reasons for choosing this particular species, including its primacy as a model of molluscan developmental biology and its role as a vector for parasites such as *Schistosoma mansoni* (more details can be found at <http://biology.unm.edu/biomphalaria-genome/>). For what it’s worth, I hereby endorse the effort to have *B. glabrata* provide the first sequenced mollusc genome.

Among bivalves, the largest genomes appear to reside in the palaeoheterodonts (mean of 3.1pg) and protobranchs (3.5 pg), although neither group has been the subject of intensive study. The best studied bivalves, the heterodonts (clams and cockles) and pteriomorphs (mussels, oysters, and scallops) possess genomes about half as large as this (means of 1.6pg and 1.5pg, respectively). Aside from this difference, the variation in bivalves is not substantial (Appendix 11.2).

Following from the first (but very limited) survey of invertebrate genome sizes, Mirsky and Ris (1951) suggested that “among the molluscs the more primitive members have the lowest DNA content”. This was based on the notion that “the squid is a far more highly developed animal than are the limpet, snail and chiton; and the squid has far more DNA per sperm than is found in the lower molluscs”. Based on the sizeable dataset now available for molluscs, it is apparent that this pattern no longer holds. Recall that the largest genome is found in a whelk, not a squid. And while it cannot be denied that cephalopods are the most complex of the molluscs, it must be noted that the arrow squid (*Loligo plei*) measured by Hinegardner (1974a) and the bobtail squid (*Euprymna scolopes*) measured in the present study both have C-values smaller than the members of some supposedly less “highly developed” groups. Still, Hinegardner (1974a) echoed his commentary on crustacean genome size evolution by invoking the following definitions with regards to molluscs: “A generalized mollusc is one that has many of the features characteristic of the group to which it belongs. A specialized one will have few of these characteristics very well developed and often completely lack others.” As with the crustaceans, Hinegardner (1974a) outlined a tendency for higher DNA contents to be found in generalized groups. Some recent authors have supported this assertion for the

case of bivalves (Rodríguez-Juíz et al. 1996) while other have refuted it (González-Tizón et al. 2000a; Thiriot-Quievreux 2002). In molluscs overall, the larger-than-average genomes of cephalopods and terrestrial gastropods would seem to deflate any such trend among major groups.

Hinegardner (1974a) also claimed that there is a positive correlation between genome size and chromosome number in molluscs at large, but this appears to be the product of comparisons between a few related groups, and not species within groups. Thus, more recent studies of particular groups of molluscs have found no such relationship between C-value and chromosome number (Iyema et al. 1994; Vitturi et al. 1995; Libertini et al. 1996; Rodríguez-Juíz et al. 1996). It is nevertheless the case that polyploidy has played an appreciable role in the evolution of molluscs. For example, Hinegardner (1974a) suggests that the neogastropods were formed by polyploidization, and polyploidy has also been identified in various bivalves (e.g., González-Tizón et al. 2000b). In the dwarf surfclam, *Mulinia lateralis*, the induction of triploidy results in larger eggs and adult body sizes as compared to diploids (Guo and Allen 1994), suggesting that variation in DNA content in molluscs could be linked to body size. In terms of basal genome size, Hinegardner (1974a) suggested such a link between genome size and body size, but only in a general way and only within families.

Among populations of the pond snail *Viviparus contectus*, Vinogradov (1998b) found minor but significant variation in genome size which was correlated positively with shell length and negatively with shell size-corrected aperture size. Both of these parameters of ecological significance, and may be related to food availability and other such environmental features. Whether genome size is selected for on this basis in this

species is not clear, but this seems an interesting possibility, and one which could potentially be extended to comparisons across species as well. Intraspecific variation was also reported for bivalve species by Rodríguez-Juíz et al. 1996), although this never reached a CV of 8% within species. Once again, a minimum level of variation of 10% is suggested as a reasonable threshold for attributing biological significance to reports of intraspecific variation in genome size (Chapter 5).

### *Annelids*

C-value estimates are currently available for approximately 80 species of polychaetes, covering some 16 orders and 32 families (Conner et al. 1972; Sella et al. 1993; Soldi et al. 1994; Gambi et al. 1997). These range 120-fold from about 0.06 to 7.2pg, with a mean of  $1.4\text{pg} \pm 0.14$  (Fi. 11.2; Appendix 11.3). Within most genera, polychaete C-values appear relatively constant, but in some groups there is evidence of cryptopolyploidy with genomes varying in a quantum fashion (Sella et al. 1993).

As pointed out by Gambi et al. (1997), variation in polychaete genome sizes is not distributed evenly among taxa. Species inhabiting interstitial environments display small C-values (about 0.06 to 1.1pg), while the genome sizes of macrobenthic species are larger and more variable (0.4 to 7.2pg; Soldi et al. 1994; Gambi et al. 1997; Appendix 11.3). This difference has generally been attributed to adaptation to the harsh and variable interstitial environment, which favours small body size, rapid development, and other r-selected traits (Sella et al. 1993; Soldi et al. 1994; Gambi et al. 1997). Macrobenthic polychaetes, on the other hand, are largely free of these constraints on features such as cell size and division rates which can be directly influenced by DNA content (Appendix 11.3). It bears mentioning, however, that the Pompeii worm (*Alvinella pompejana*)

measured in the present study is one of the most heat-tolerant animals known and lives in the extremely harsh environment of sub-oceanic thermal vents, but does not display a particularly large or small genome (~0.8pg; Appendix 11.4). Other vent-dwelling invertebrates similarly show no striking increase or decrease in genome size despite inhabiting these challenging environments (Dixon et al. 2001).

Oligochaetes have so far received only modest attention in terms of genome size variation. Relative nuclear DNA contents of roughly 40 species of potworms (Order Enchytraeida, Family Enchytraeidae) were assessed by Christensen (1966), but unfortunately could not be converted to absolute genome sizes because no internal reference standard was included. Chromosomal studies have revealed that polyploidy is widespread in several oligochaete families (e.g., Muldal 1952; Christensen 1966, 1980; Casellato 1987; Coates 1995), but absolute nuclear DNA content measurements have not been part of this work. And while the complete mitochondrial genome of *Lumbricus terrestris* has been sequenced (Boore and Browb 1995), nuclear genome sizes have so far been reported for only two oligochaetes, the earthworms *Eisenia fetida* at 0.7pg and *Octodrilus complanatus* at 0.86pg (Vitturi et al. 2000).

Earthworms not only play an essential role in nutrient cycling and soil turnover, they impose a slow but substantial influence on the landscape itself. Economically, they are important participants in many large-scale composting programs, and in the multi-million-dollar fishing bait industry (e.g., Tomlin 1983). As Charles Darwin (1881) put it, “Worms have played a more important part in the history of the world than most persons would at first suppose”. Unfortunately, this importance has obviously not translated into interest among genome biologists. In similar fashion, only one genome

size has been estimated for leeches (Class Hirudinea). This discrepancy is dealt with in at least a preliminary way in the present study, which includes genome size estimates for 15 species of earthworms and six leeches (Appendix 11.4). Estimates for 12 species of freshwater oligochaetes performed previously in the Hebert lab are also discussed in this context. These newer oligochaete data form the basis of most of the present discussion, as detailed reviews of polychaete genome size variation are available elsewhere (Soldi et al, 1994; Gambi et al. 1997).

Most of the earthworms analyzed in the present study were collected from wild populations in southern Ontario from various habitats with different soil types (Appendix 11.4). Species were collected from under rocks, logs, and leaf litter, by digging and manual soil sorting, and by chemical extraction using formalin (Schwert 1990). Three additional species were acquired from commercial compost colonies<sup>3</sup>. Species were identified according to Reynolds (1977), Sims and Gerard (1985), and Schwert (1990). Coelomocytes of the various species were compared against the same cell type from *Eisenia fetida*, which has a genome size of 0.7pg, as determined by flow cytometric measurements of testes (Vitturi et al. 2000) and Feulgen image analysis of testes and sperm (present study). Mid-ranges of body size data for the various earthworm species were taken from Reynolds (1977) and Sims and Gerard (1985), and included values for length, diameter, and number of segments. Average body volume was calculated for each species ( $\text{Volume} = \text{B} \times \text{Length} \times [\frac{1}{2} \text{Diameter}]^2$ ), as were average lengths and volumes of individual segments (Table 11.2). The freshwater oligochaetes were collected from a few

---

3

Courtesy of Vermitech Systems Ltd., Toronto, ON.

sites in Ontario (primarily from Peche Island, Pike Creek, and Lake St. Clair), and were measured by scanning densitometry of Feulgen-stained whole-body squashes. As body size data were not readily available for these taxa, a measurement of the length of a single segment (the 8<sup>th</sup>) was chosen for comparison among species.

C-values for the 12 species of freshwater oligochaetes ranged nearly 10-fold, from 0.8 to 7.6pg (Appendix 11.3). While the current sample of aquatic oligochaetes is small, some interesting patterns are apparent. Notably, within the genus *Limnodrilus* there is a clear division between species with ~0.8pg and those with genomes almost exactly four times larger at ~3.2pg. It is tempting to ascribe this discontinuity to polyploidy, which is known to be common in the Tubificidae (Christensen 1980). However, *L. hoffmeisteri* (0.87pg) and *L. udekemianus* (3.17pg) are reported to have approximately equal chromosome numbers ( $2n = 100$ ; Christensen 1980), which suggests that this difference may instead be due to cryptopolyploidy.

Similar patterns of discontinuous genome size variation, not explainable in terms of polyploidy, have been reported for polychaete annelids (Sella et al. 1993) and several other invertebrate groups (see below and Chapters 8 and 9). It is also of interest that some of the largest genomes among oligochaetes are found in the Naididae (Appendix 11.3), even though polyploidy is rare in this group (Christensen 1980). Thus, it appears that shifts in genome size in freshwater annelids – even large and discrete ones – likely occur independently of changes in the karyotype. These also appear to occur without any obvious effects on body size, as assessed by the admittedly crude measure of 8<sup>th</sup> segment length ( $p > 0.88$ ; Gregory and Hebert 2002).

All six leech genome sizes measured here were relatively small, ranging from

about 0.3 to 0.6pg. The smallest was found in the Australian terrestrial leech *Philaemon pungens*, but this was not substantially larger than those of several freshwater species (*Helobdella* spp.). The largest C-value was found in the ribbon leech, *Nephelopsis obscura*, which is commonly used for fishing bait and which is not a blood-feeder. A great deal more work will be required before any meaningful conclusions can be drawn regarding leech genome size variation, but it is interesting that species from two different orders and with very different lifestyles all appear to have genomes of a similar size (Appendix 11.4).

Genome sizes of the 15 species of earthworms were also relatively constrained, ranging only 3-fold, from about 0.4 to 1.2pg (Appendix 11.4). Variation within genera was limited, but when present appeared bimodal. Based on this sample of lumbricids, it is clear that differences in nuclear DNA content among these species also occur independently of variation in chromosome numbers ( $r = 0.1$ ,  $p > 0.74$ ), and are therefore not likely due to polyploidy. Polychaete annelids similarly show no significant relationship between genome size and chromosome number (Gambi et al. 1997). A cross-class comparison also reveals that whereas some polychaete annelids have genome sizes of 0.4pg with  $2n = 6$  chromosomes (Gambi et al. 1997), other polychaetes and some earthworms possess similar genome sizes despite having chromosome numbers several times higher (Gambi et al. 1997; Appendix 11.3; Table 11.2). Nevertheless, polyploidy is known to have played an important role in the general evolution of the Lumbricidae (Christensen 1980; Casellato 1987). *Perionyx excavatus* (Family Megascolecidae), the only non-lumbricid earthworm here analyzed, has a relatively small C-value similar to that of the other worms, although chromosome numbers are unknown for this species

(Appendix 11.4).

A weak but positive association between C-value and body length has been found in polychaetes (Soldi et al. 1994). However, no significant relationships were found between genome size and body length, body volume, or segment size in the lumbricids (all  $p = 0.13$  to  $0.3$ ). Moreover, some of the smallest earthworms studied here had the largest C-values (Table 11.2). It is also relevant in this regard that although two distinct size morphs of *E. rosea* were found (one large and one small), there was no significant difference in their genome sizes ( $t$ -test,  $p > 0.45$ ).

The environments inhabited by most earthworms are quite stable, and many species take months or even years to mature (Sims and Gerard 1985). It is therefore unlikely that developmental constraints play a significant role in shaping their genome sizes, in contrast to the case of polychaetes. It is worth noting that polychaetes generally possess a more sophisticated capability for regeneration than do oligochaetes, which again suggests that cell division requirements probably do not exert as great an influence on earthworm C-values.

Earthworm species can be classified along the r-K continuum according to their lifestyles (Satchell 1980; Sims and Gerard 1985; Table 11.2). In general, r-selected species inhabit less stable surface environments and have high fecundity and mortality, shorter incubation and maturation times, high metabolic rate, small body size, high mobility, and more developed pigmentation (Satchell 1980). K-selected species, by contrast, typically inhabit stable subterranean environments and display the opposite trends as those for r-selected species (Satchell 1980). In the present survey, there was no obvious association between genome size and r-K classification ( $t$ -test,  $p > 0.74$ , with

intermediate r/K worms included in both categories for comparison).

There is, however, a tentative indication that genome size may be related to reproductive system in earthworms. Specifically, the obligately and facultatively parthenogenetic species studied here have larger genomes on average than strictly sexual species (*t*-test,  $p < 0.03$ ). However, this significant difference was partly dependent on the characterization of *Dendrobaena* species, which have the largest genomes in the current survey, as facultatively parthenogenetic. Indeed, polyploid parthenogenetic forms are common in *D. rubida* and especially *D. octaedra* (Reynolds 1977; Sims and Gerard 1985), but in the present study spermatozoa were obtained from *D. rubida*, suggesting that this was a sexual diploid population. Nevertheless, a link has long been recognized between polyploidy and parthenogenesis in earthworms (Muldal 1952; Christensen 1980; Casellato 1987), and relatively large haploid genome sizes may similarly be more common among parthenogenetic species. In any case, the present results indicate that more detailed surveys of oligochaete species with different ecological and reproductive attributes may reveal some interesting patterns.

### *Flatworms*

To date, genome sizes have been estimated for 56 species in the Phylum Platyhelminthes, nearly all of them free-living members of the Class Turbellaria (Appendix 11.5). Overall, these vary nearly 350-fold, from a mere 0.06pg in *Stenostomum brevipharyngium* to 20.5pg in *Otomesostoma auditivum*. This latter value is one of the largest yet found for an invertebrate. Relatively large genomes are found in the Orders Allocoela, Rhabdocoela, and Tricladida (means 8.0, 3.6, and 2.1, respectively), whereas the Orders Catenulida, Macrostomida, and Proseriata display much

smaller and more tightly constrained C-values (0.5, 0.4, and 1.3, respectively). The simplest group of turbellarians, the Order Acoela, is currently represented only by the marine flatworm *Neochildia fusca*, which was estimated in the present study to have a C-value of about 1.9pg.

Figure 11.1 shows the correlation between genome size and body size reported for 31 species of turbellarian flatworms by Gregory et al. (2000). Given the strength of this relationship, it seems likely that body size constraints play a role in shaping genome size variation among free-living flatworms. Development time may also be related to genome size in flatworms (Göltenboth 1973; Hebert and Beaton 1990), although this has not yet been studied in sufficient detail. Interestingly, *Mesostoma ehrenbergii* apparently displays a generation time of only 16 days, despite its genome size of nearly 15pg; as Hebert and Beaton (1990) put it, “further study of the factors permitting *Mesostoma*’s rapid rate of development is merited”. Turbellarians are also important from a cytogenetic point of view because they exhibit few, large chromosomes as well as variation in B chromosome frequencies, ploidy levels (both within and among individuals), and reproductive modes (Martens et al. 1989; Hebert and Beaton 1990; Beukeboom et al. 1998). These parameters do raise issues regarding the accuracy of genome size estimates, however, and may explain some of the discrepancies observed among studies (J. Baguña, personal communication). Whether they explain the apparent (but by no means conclusively demonstrated) patterns of quantum shifts among related flatworms (Gregory et al. 2000) will remain unknown in the absence of further investigation.

Prior to the present study, no other groups of Platyhelminthes had been measured

for genome size. Four species of flukes (Class Trematoda) have been analyzed here, and these range in C-value from about 0.9 to 1.3pg. The three species in the genus *Trichobilharzia* (Family Schistosomatidae, Order Digenea) are parasites of birds with a snail intermediate host, but are also implicated in the appearance of cercarial dermatitis (“swimmers’ itch”) in humans. All three have similar genome sizes (~1.1 to 1.3pg), whereas *Diplostomum pseudospathaceum* (Family Diplostomidae, Order Strigeidida), which has both a fish and snail host and which is also a parasite of birds, has a smaller C-value of about 0.9pg (Appendix 11.1). In light of the complex life cycle of these trematodes, it may not be surprising that they have relatively small genomes, although whether this is a general feature of the class cannot be determined without a great deal more data. No tapeworm (Class Cestoda) genome size data are currently available, but their relatively simple life cycle might be taken to predict large maximum C-values in the class.

### *Echinoderms*

C-value estimates for echinoderms currently total 44 species, with 39 of these presented by Hinegardner (1974b). Suffice it to say, much work remains to be done in this group, and that so far the necessary effort has not been forthcoming. Overall, the echinoderm genome sizes currently available vary from 0.5pg in the sea star *Dermasterias imbricata* to 4.4pg in the sea cucumber *Thyonella gemmata*, about an 8-fold range (Appendix 11.6). The mean for the phylum is  $1.4\text{pg} \pm 0.13$ . Of the major groups studied so far, the asteroidean stelleroids (that is to say, sea stars) display a mean genome size of about 0.7pg, whereas the ophiurideans (serpent stars) have much larger genomes averaging 2.6pg. Echinoids (sea urchins, sand dollars) have a mean genome

size of about 1.0pg, and the Holothuroidea (sea cucumbers) average about 2.0pg.

Crinoids (feather stars and sea lilies) have not yet been studied in terms of genome size.

The same correlation between specialization and small genome size is attributed to the echinoderms by Hinegardner (1974b), although it is obviously difficult to assess with the limited dataset currently available. It does bear mentioning that the sea cucumbers themselves cover almost the entire range observed in the phylum (0.8 to 4.4pg), which may partially negate any such pattern.

### *Nematodes*

Estimates for nematode C-values total 21 species so far, all but one of which are from the Class Secernentea. These range from 0.05pg in root-knot nematodes (*Meloidogyne* spp.) to about 2pg in the horse roundworm (*Parascaris univalens*). The other is the Class Adenophorea, represented by the trichinosis roundworm, *Trichinella spiralis*, at about 0.3pg (Appendix 11.7). The mean for all nematodes is about  $0.25\text{pg} \pm 0.09$ , although without the single large (germline) C-value of *P. univalens*, the only one higher than 0.9pg, the average drops to  $0.16\text{pg} \pm 0.02$ . Thus, nematodes possess some of the smallest genomes among animals. It should not be surprising, therefore, that the first metazoan genome sequenced in its entirety belonged to a member of this phylum (*Caenorhabditis elegans*).

Although integral to both comparative genomics and developmental biology, nematodes have not been considered especially interesting from the perspective of genome size evolution. This neglect seems unwarranted, however, since there are several intriguing aspects of nematode biology that make them useful in the study of the C-value enigma, even with their apparent lack of large-scale variation in genome size. Most

famously, the parasitic ascarid nematodes undergo chromatin diminution, and indeed the process was first discovered in a member of this group 115 years ago (see Chapters 1 and 8 for discussion). Notably, the large germline genome of *P. univalens* and the much smaller one of *Ascaris lumbricoides* are both reduced to about 0.25pg in the somatic line, a remarkable 85% reduction in the former case (Moritz and Roth 1976).

Like copepods, but in an even more strict sense, nematodes exhibit determinate growth with each individual being comprised of a set number of cells. In this regard, it is tempting to assume that variation in DNA content would necessarily be linked to such features as developmental rate and body size (as in the copepods), and therefore to explain the prevalence of chromatin diminution in parasitic nematodes (see Chapter 8 for more on copepods). However, it appears that endopolyploidy, and not shifts in basal genome size, may drive body size variation in nematodes. Endopolyploid cells had been identified more than 15 years ago in the intestinal and hypodermal tissues of *C. elegans* by Hedgecock and White (1985), but the relevance of this process to morphological variation was not recognized until recently.

Unlike in vertebrates, body size variation among nematodes is not produced (almost) exclusively by differences in cell numbers. In nematodes, increases in body size consist in large part of cell growth during early adulthood, in the absence of cell division. The organ primarily responsible for shaping adult body size in nematodes is the hypodermis, which covers the worm and excretes the exoskeletal cuticle, and which consists of a single large syncytium (a mass of fused cells containing many nuclei). In their study of 12 free-living terrestrial nematodes from three families, Flemming et al. (2000) examined the possible relationships between body size, number of syncytial nuclei

(the functional equivalent of “cell number”), haploid genome size, and level of endopolyploidy in hypodermal nuclei. They found that nuclear number was a poor predictor of body size, suggesting that other factors must influence the size of the hypodermal syncytium and therefore adult size. Genome size was found to exert at best only a very weak influence on overall body size. The level of endopolyploidy, on the other hand, appears to be an important determinant of body size in these nematodes. Moreover, certain dwarfism mutations in *C. elegans* affecting the hypodermis were shown to exert their effects at least in part by causing a deficiency of endoreduplication in hypodermal nuclei (Flemming et al. 2000). Taken together, these results provide a convincing demonstration of the importance of DNA content modulation in the evolution of nematodes.

### *Tardigrades*

Tardigrades, commonly known as “water bears”, are aquatic microinvertebrates that possess a superficial resemblance to arthropods (and also to ursid mammals!). These tiny animals (0.1-1.2mm) are of interest in the study of the C-value enigma because they display determinate growth with cell number constancy (usually ~40,000 cells), variation in reproductive mode and ploidy level, and high levels of endopolyploidy in certain tissues (Bertolani et al. 1994; Garagna et al. 1996). In terms of haploid genome sizes, a total of 17 tardigrade species has so far been examined. Some tardigrade C-values are among the smallest found in animals, and range about 10-fold overall, from 0.08 to 0.82pg and with a mean of  $0.38\text{pg} \pm 0.05$  (Appendix 11.8). Variation in tardigrades appears to be found largely among individual species, and does not segregate along family or higher taxonomic lines (Garagna et al. 1996). Little is known about the reasons

for variation among species, and the only distinctive C-value correlation yet reported (or indeed, examined) is with sperm head morphology, such that species with short, round-headed spermatozoa possess higher DNA contents than those with more specialized long-headed sperm cells (Garagna et al. 1996). Thus, the tardigrades may also be superficially similar to insects in this regard as well (see Chapter 9).

### *Gastrotrichs*

Members of the Phylum Gastrotricha are small (0.5-4mm), colourless, free-living aquatic worms most closely related to nematodes. Gastrotrichs share many features with tardigrades in terms of genome size variation, including a similar number of species studied to date ( $n = 15$ ), and a small range from 0.08 to 0.63pg (Balsamo and Manicardi 1995; Appendix 11.8). The mean C-value for the available gastrotrichs is  $0.23\text{pg} \pm 0.04$ , similar to the nematodes. The small genome sizes of these animals are of potential significance to some of the topics discussed in these volumes, given their small body sizes and extremely rapid development (i.e., sometimes as little as two days to sexual maturity). Again like the tardigrades, the variation in gastrotrich genome sizes does not appear segregate according to class/order, although the current data are insufficient to determine whether there are clear distinctions among families.

### *Myriapods*

Prior to the present study, no data were available on the genome sizes of centipedes or millipedes. As part of the present survey, 14 species of myriapods (3 centipedes, 11 millipedes) were assayed for variation in C-value. In total, the genome sizes in this group ranged from 0.28 to 2.14pg, with a mean of  $0.78\text{pg} \pm 0.16$ . Small and relatively large C-values were found in both centipedes (Class Chilopoda, 0.39 to 2.14pg)

and millipedes (Class Diplopoda, 0.28 to 1.4pg).

The two most well known centipedes, the common house centipede<sup>4</sup> (*Scutigera coleoptrata*) and the rock centipede (*Lithobius forficatus*) had the largest genomes in the sample (2.03pg and 2.14pg, respectively). With the exception of one unidentified millipede at 1.4pg, all other myriapods analyzed here had C-values less than 0.9pg. Like spiders and ametabolous insects, myriapods undergo no metamorphosis during their development, such that juveniles exhibit the morphology of a miniaturized adult. As with these other groups, the maximum known genome size for myriapods exceeds the 2pg threshold for metamorphosing insects (Chapter 9).

### *Cnidarians*

Prior to the present study, only three species of the Phylum Cnidaria (formerly Coelenterata), one hydra and two jellyfishes, had previously been examined for genome size (Appendix 11.8). C-values were estimated for two additional species in this study, including a sea anemone and a coral (Appendix 11.8). Although this covers the better known groups of cnidarians (and all three classes), such a small dataset provides very limited insights into the evolution of genome sizes in this diverse group. The two jellyfishes and the anemone studied so far all have relatively small genomes (0.2-0.7pg), whereas the hydra and coral display larger C-values (1.1-1.9pg). However, neither the generality nor the significance (or lack thereof) of this apparent difference can be discussed without the acquisition of additional data from each of these groups. Because of the differences in life history among the different classes, the Cnidaria remains a very

---

4

Better known to basement-dwelling undergraduates as “What the hell is *that?!?*”.

good candidate group for more detailed future study.

### *Rotifers*

Rotifers are tiny (most 200 to 500 $\mu$ m) filter-feeding animals found in aquatic (primarily freshwater) environments and moist soil. Very little is known of genome size variation in the Phylum Rotifera, with only four species studied thus far. The genomes of these species range from about 0.35 to 1.1pg (Appendix 11.8), although the actual degree of variation in the phylum cannot be assessed with such a small sample. Polyploidy is known in rotifers, and is associated with body size variation in at least one species (*Euchlanis dilatata*; Walsh and Zhang 1992). Endopolyploid nuclei have also been found in rotifers, with the yolk gland of *Asplanchna sieboldi* ranging from 36- to 300-ploid (Jones and Gilbert 1977).

Bdelloid rotifers are well known in evolutionary biology because of their status as ancient asexuals. These animals reproduce by obligate parthenogenesis, and yet appear to have survived for vast lengths of time (thereby indicating that they have not read the relevant theoretical literature). In terms of genome size evolution, the asexual nature of these rotifers is of interest because it allows a test of neutralist approaches to the C-value enigma. For example, proponents of the DNA loss hypothesis have sometimes suggested that a lack of sexual reproduction will curb the recombinational spread of transposable elements, and therefore promote the loss of DNA by deletion bias as the elements cease to be replenished<sup>5</sup>. Despite the current paucity of data, it does seem apparent that this

---

5

This is based on personal communications and manuscripts I have reviewed; no published statements to this effect are available for citation at this stage. See Chapter 6 for a more detailed treatment of the DNA loss hypothesis.

prediction fails in this phylum. Specifically, the three asexual bdelloid rotifers all have genomes *larger* (0.5-1.1pg) than the one known C-value from a sexually-reproducing monogonont rotifer (0.35pg).

### *Sponges*

The first two sponge genome size estimates were provided by Mirsky and Ris (1951). Since then, only two more species have been examined (Imsiecke et al. 1995). Genome size estimates vary about 30-fold in the Phylum Porifera, but it seems possible that this is due to measurement artifacts. That is, both species analyzed by Mirsky and Ris (1951) had reported C-values of 0.06pg, whereas the two species assessed by Imsiecke et al. (1995) had genomes around 1.6-1.8pg (Appendix 11.8). However, these species were all derived from different orders, so it is possible that such variation is legitimate. Given the “primitive” status, extremely simple morphological organization, and very long evolutionary lifespan of sponges, this is clearly an issue worth examining.

### *Miscellaneous invertebrates*

The genome sizes of animals from several of the lesser known phyla have been estimated over the past several decades, although these have been few and far between (see Appendix 11.8 for the short but complete list). Some of them are of particular interest because of their unique biological characteristics or important phylogenetic positions, and further study of these and other groups is highly recommended.

One of the most interesting examples is the only extant member of the Phylum Placozoa, *Trichoplax adhaerens*. This animal displays the simplest organization yet observed in a metazoan, consisting of only four cell types arranged in three layers. It possesses no definitive shape, no organs, and resembles a sort of multicellular ciliated

amoeba. At a minuscule 0.04pg, it also has the smallest C-value yet reported for a metazoan. Interestingly, despite this extreme morphological simplicity, endopolyploidy is nevertheless known in this species, with the two epithelial layers being diploid and the middle cell layer being tetraploid (Ruthmann and Wenderoth 1975; Ruthmann 1977). Other small marine phyla for which at least one genome size is known include the Phylum Priapulida (phallus worms), represented by *Halicryptus spinulosus* at 0.56pg, and the Phylum Nemertea (ribbon worms), with *Cerebratulus lacteus* at 1.4pg.

A single representative of the Phylum Onychophora (velvet worms) has also been analyzed, which is of interest in the context of ideas presented in these volumes. Velvet worms represent an ancient group, close relatives of which were present in the Cambrian more than 500 million years ago. These animals are sometimes considered a group of arthropods, and are generally believed to at least be related to the Arthropoda (and the Tardigrada), perhaps remaining similar to their common ancestor. It is of interest in this context that these terrestrial relatives of arthropods have relatively simple development (and in some cases are viviparous), and appear to possess large genomes: the single species studied so far, *Epiperipatus biolleyi*, has a C-value of about 4.4pg.

Another poorly characterized group for which limited genome size data are available, this time a relative of the annelids, is the Phylum Pogonophora (beard worms). The taxonomic placement of this group is unclear; it is sometimes referred to as the Phylum Vestimentifera, and is considered by some to be a part of the Phylum Annelida. Only one species has been studied from this phylum, the Atlantic hydrothermal vent beard worm *Riftia pachyptila*, which has a genome size of 0.64pg. This is consistent with the C-values of many annelids (Appendix 11.3), and does not differ greatly from those of

vent-dwelling annelids like the Pompeii worm (0.8pg) and other polychaetes (1.0-1.5pg).

Three additional phyla not previously investigated were dealt with in a preliminary way in the present study (see Appendix 11.1). These include two representatives of the Phylum Chaetognatha (arrow worms), *Flaccisagitta enflata* at 0.7pg and *Spadella cephaloptera* at about 1pg. Two species from the Phylum Ctenophora (comb jellies) from two different orders were also assayed, and these were found to differ by an order of magnitude from 0.3pg in *Mnemiopsis leidy* to 3.2pg in *Haeckelia rubra*. Finally, a single species of the Phylum Sipunculida (peanut worms), *Themiste langeniiformis*, was found to have a C-value of 1.3pg.

### **Genome size evolution in invertebrates: a summary**

In Volume One, it was established that while there are some universal correlates of genome size in animals (e.g., cell size), the implications of these may vary greatly among taxa. Thus, metabolism and not development is important in mammals and birds, whereas the reverse is true of amphibians (and perhaps fishes) (Chapters 3, 4, and 5). Similarly, the implications of genome size variation in different groups of invertebrates are diverse and related to the particular biology of the animals in question.

In some crustaceans, DNA content modulation by polyploidy and quantum shifts in C-value have probably been of substantial import, particularly in the copepods where variation in genome size is linked to body size and developmental rate. Some crustaceans display complex and developmentally important mechanisms of DNA content modulation, including endopolyploidy and chromatin diminution (Chapter 8). Endopolyploidy is also important among insects, and in this case genome size appears unrelated to body size but may be associated with developmental complexity. The

spiders display a relatively large range in DNA content, although the exact implications of this remain unknown. The two-spotted spider mite, by contrast, has one of the smallest arthropod genomes yet reported, and its relation to development seems clear.

Only a few groups of non-arthropod vertebrates have been studied in any kind of detail from the perspective of genome size, but some interesting patterns are nonetheless apparent. In gastropod molluscs, there may be an association between terrestriality and large genome size, and in nematodes the modulation of DNA content by endopolyploidization is a prime determinant of body size. In flatworms, there is a strong positive association between basal genome size and body size. Among polychaete annelids, there are associations between C-value and ecological lifestyle, while in earthworms genome size seems more closely related to reproductive mode. Again, while the implications of genome size variation may differ widely among groups, the general importance of such cannot be denied.

The current lack of information on many invertebrate phyla can perhaps be explained (but not excused) on the basis of unfamiliarity and relative difficulty of acquiring specimens. However, since many of these groups display biological characteristics of potentially great interest, it is apparent that the additional effort required to study their genome sizes is warranted. A great many phyla remain totally unexamined, and those that have been are often represented by only one or a few species. Interesting patterns of genome size variation in invertebrates are beginning to illuminate some of the puzzles that make up the C-value enigma, but in many cases these remain no more than glimpses. Nevertheless, they indicate that the time has come to shed light on groups formerly eclipsed by biology's vertebrate bias.

**Table 11.1.** Summary of the various invertebrate groups examined for genome size to date, excluding those arthropod groups that were dealt with in previous chapters: crustaceans (Chapter 8), insects (Chapter 9), and spiders (Chapter 10). The number of species for which C-values have been estimated is given for all previous studies combined and for the present study.

<b>Taxon</b>	<b>Number of species studied previously</b>	<b>Number of species in the present study</b>
<b>Phylum Annelida</b>	<b>85</b>	<b>26</b>
Class Hirudinea	1	6
Class Oligochaeta	2	15 <sup>1</sup>
Class Polychaeta	82	5
<b>Phylum Arthropoda</b>		
<b>Subphylum Chelicerata</b>	<b>1</b>	<b>0</b>
Class Merostomata	1	0
<b>Subphylum Myriapoda</b>	<b>0</b>	<b>14</b>
Class Chilopoda	0	3
Class Diplopoda	0	11
<b>Phylum Chaetognatha</b>	<b>0</b>	<b>2</b>
<b>Phylum Cnidaria</b>	<b>3</b>	<b>2</b>
Class Anthozoa	0	2
Class Hydrozoa	1	0
Class Scyphozoa	2	0
<b>Phylum Ctenophora</b>	<b>0</b>	<b>2</b>
<b>Phylum Echinodermata</b>	<b>44</b>	<b>0</b>
<b>Subphylum Asterozoa</b>		
Class Stellerioidea	15	0
<b>Subphylum Echinozoa</b>		
Class Echinoidea	18	0
Class Holothuroidea	11	0

<b>Taxon</b>	<b>Number of species studied previously</b>	<b>Number of species in the present study</b>
<b>Phylum Gastrotricha</b>	<b>15</b>	<b>0</b>
<b>Phylum Mollusca</b>	<b>167</b>	<b>3</b>
Class Bivalvia	80	0
Class Cephalopoda	4	1
Class Gastropoda	76	2
Class Polyplacophora	7	0
<b>Phylum Nematoda</b>	<b>21</b>	<b>0</b>
Class Adenophorea	1	0
Class Secernentea	20	0
<b>Phylum Nemertea</b>	<b>1</b>	<b>0</b>
<b>Phylum Onychophora</b>	<b>1</b>	<b>0</b>
<b>Phylum Placozoa</b>	<b>1</b>	<b>0</b>
<b>Phylum Platyhelminthes</b>	<b>51<sup>2</sup></b>	<b>5</b>
Class Turbellaria	51	1
Class Trematoda	0	4
<b>Phylum Pogonophora<sup>3</sup></b>	<b>1</b>	<b>0</b>
<b>Phylum Porifera</b>	<b>4</b>	<b>0</b>
<b>Phylum Priapulida</b>	<b>1</b>	<b>0</b>
<b>Phylum Rotifera</b>	<b>4</b>	<b>0</b>
<b>Phylum Sipuncula</b>	<b>0</b>	<b>1</b>
<b>Phylum Tardigrada</b>	<b>17</b>	<b>0</b>
Class Eutardigrada	16	0
Class Heterotardigrada	1	0
<b>TOTAL</b>	<b>417</b>	<b>55</b>

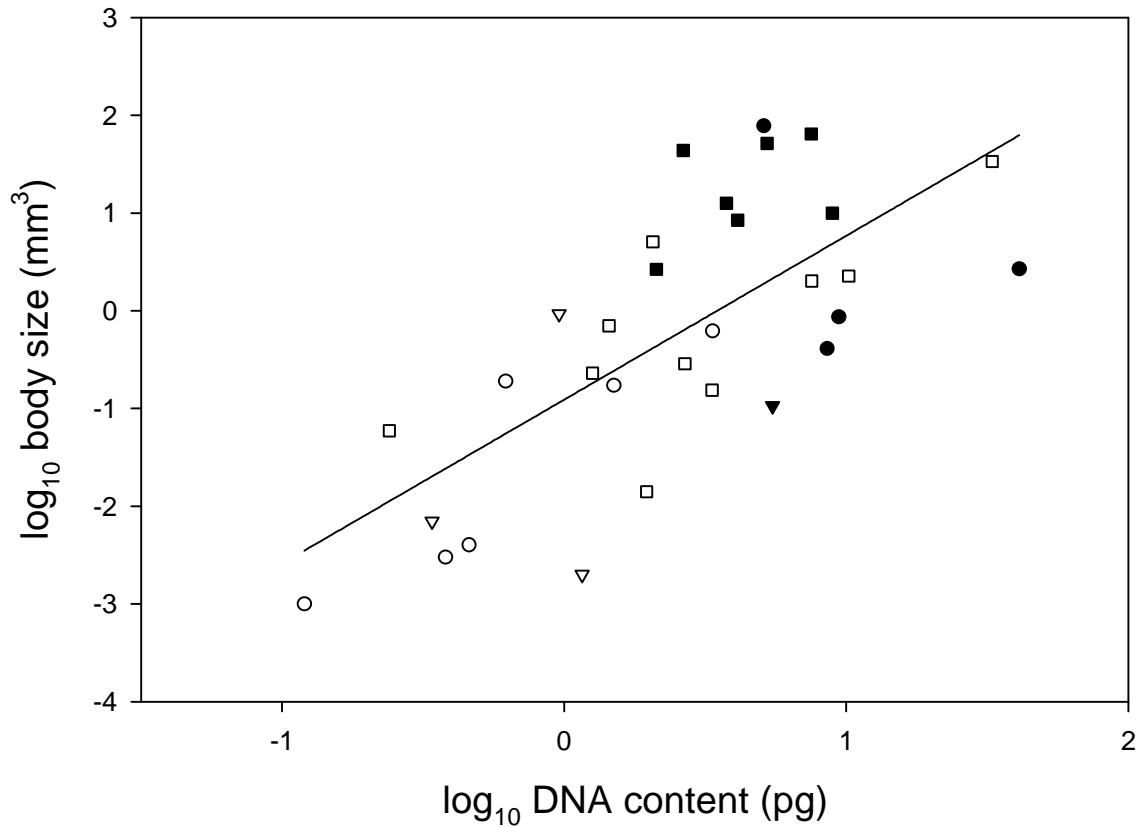
**Table 11.1 notes:**

- 1) Does not include 12 species of freshwater oligochaetes previously measured in the Hebert lab, and included in Gregory and Hebert (2002).
- 2) Includes 38 species of turbellarian flatworms previously measured in the Hebert lab, and published in Gregory et al. (2000).
- 3) The taxonomy of this group is not clear. This is sometimes referred to as the Phylum Vestimentifera, and is sometimes considered a part of the Phylum Annelida.

**Table 11.2.** Summary of C-value estimates, chromosome numbers, body sizes, and ecological and reproductive attributes for 15 species of earthworms (Phylum Annelida, Class Oligochaeta, Order Haplotaxida). Karyotype (2n) data were taken from Christensen (1980). An asterisk indicates species for which polyploid forms or conflicting 2n values have been reported. N = number of individuals examined in the present study. Body size data (L = length in mm, D = diameter in mm, Seg = segment number) follow Reynolds (1977) and Sims and Gerard (1985). Designations of worm species as either “r-selected” or “K-selected” are based on the discussions of Satchell (1980) and Sims and Gerard (1985). Reproductive systems (RL: S = sexual, P = parthenogenetic) follow Reynolds (1977) and Sims and Gerard (1985).

Species	C-value (pg)	2n	L	D	Seg	r/K	RL
<b>Family Lumbricidae</b>							
<i>Allolobophora chloritica</i>	1.02	32	30-70	3-5	80-138	K	S
<i>Aporrectodea caliginosa</i> ( <i>Allolobophora caliginosa</i> , <i>Aporrectodea turgida</i> )	0.65	36*	60-85	3.5-5	130-168	K	S
<i>Aporrectodea tuberculata</i>	0.67		90-150	4-8	146-196		S
<i>Dendrobaena octaedra</i>	1.2	108	17-60	3-5	60-100	r	S/P
<i>Dendrobaena rubida</i> ( <i>Dendrodrilus rubidus</i> )	1.24	34*	20-90	2-5	50-120	r	S/P
<i>Dendrobaena veneta</i> ( <i>Eisenia veneta</i> )	0.46	36	50-155	4-8	86-255		S
<i>Eisenia fetida</i> ( <i>Eisenia foetida</i> )	0.7	22	35-170 (usually < 70)	3-5	80-110	r	S
<i>Eisenia rosea</i> ( <i>Aporrectodea rosea</i> )	1.1	54*	25-85	3-5	120-150	K	P
<i>Eiseniella tetraedra</i>	0.9	54*	30-60	2-4	60-90	r	P
<i>Lumbricus castaneus</i>	0.43	36	30-50 (usually < 35)	3-5	70-100	r	S
<i>Lumbricus rubellus</i>	0.43	36	50-150 (usually > 60)	4-6	70-120	r/K	S

<b>Species</b>	<b>C-value (pg)</b>	<b>2n</b>	<b>L</b>	<b>D</b>	<b>Seg</b>	<b>r/K</b>	<b>RL</b>
<i>Lumbricus terrestris</i>	0.6	36	90-300	6-10	120-160	r/K	S
<i>Octolasion cyaneum</i>	0.65	190	65-180	7-8	140-158	K	P
<i>Octolasion tyrtaeum</i> ( <i>Octolasion lacteum</i> )	0.83	36- 38*	25-130	3-6	75-150	K	P
<b>Family</b>							
<b>Megascolecidae</b>							
<i>Perionyx excavatus</i>	0.45		30-180	3-7	123-178		



**Figure 11.1.** The relationship between genome size and body size in turbellarian flatworms. Different symbols represent different orders ( $\tilde{Z}$  = Alloecoela,  $\bullet$  = Catenulida,  $—$  = Kalythorhynchia,  $L$  = Macrostomida,  $\text{€}$  = Tricladida,  $\sim$  = Typhloplanida). The relationship is highly significant ( $r^2 = 0.50$ ,  $p < 0.0001$ ,  $n = 31$ ) in flatworms at large, and appears to hold within orders as well. From Gregory et al. (2000).