Genome Size Diversity and Patterns within the Annelida

by

Alison Christine Forde

A Thesis
presented to
The University of Guelph

In partial fulfillment of requirements
for the degree of
Master of Science
in
Environmental Biology

Guelph, Ontario, Canada

ABSTRACT

GENOME SIZE DIVERSITY AND PATTERNS WITHIN THE ANNELIDA

Alison Christine Forde
University of Guelph, 2013

Advisors:
Dr. T. Ryan Gregory
Dr. Jonathan Newman

This thesis concerns genomic variation within the Annelida, for which genome size studies are few and provide data for only a handful of groups. Genome size estimates were generated using Feulgen image analysis densitometry for 35 species of leeches and 61 polychaete species. Relationships were explored utilizing collection location and supplementary biological data from external sources. A novel, inverse correlation between genome size and maximum adult body size was found across all leeches. Leeches that provide parental care had significantly larger genome sizes than leeches that do not. Additionally, specimens identified as *Nephelopsis obscura* exhibited geographic genome size variation. Within the Polychaeta, Polar region polychaete genomes were significantly larger than those of Atlantic and Pacific polychaetes. These studies represent the first exploration of leech genome sizes, and provide base evidence for numerous future studies to examine relationships between genome size and life history traits across and within different annelid groups.

ACKNOWLEDGEMENTS

I have been extraordinarily fortunate to have a strong support system during my undergraduate and graduate studies at the University of Guelph. A very sincere thank you goes to my advisor Dr. T. Ryan Gregory for his trust, leadership and guidance over the past few years, and for taking an interest in the lesser-loved leeches. I would also like to express my gratitude to my co-advisor, Dr. Jonathan Newman, for his assistance throughout this process.

My parents, with their love, support and understanding during my various struggles and successes, were a huge part in making my academic career possible; for this I cannot thank them enough.

Reaching back to the early fall semester in 2008, I must also thank Dr. Doug Larson for his advice given during his Ecological Methods course, in a particular lecture titled "So you want to go to grad school?" After class, I asked how I might go about finding a lab to volunteer with, to see if research was the right post-undergrad path for me. Doug told me to read up on the various labs in the department and talk to the professor(s) doing research that interested me the most. This led me to approach Ryan, my then-Integrative Vertebrate Zoology professor, about volunteering in his lab. Shortly afterwards, Ryan had me under the wing of Jill Smith, who had collected mammal samples for genome sizing as part of her Masters research on bats. I am grateful for these connections and the academic pathway they started me on.

Over the past four years, fellow members of the Gregory Lab proved to be invaluable to my research as both an undergraduate and graduate student. Without their support in the lab, field and in meetings, I would not have reached where I am today. I would especially like to thank Nick Jeffery and Diana Petlock for accompanying me and assisting me on my April 2012

collecting trip to Auburn University in Alabama. Many thanks also to Patrick Moldowan for collecting specimens from the Algonquin area over the summer of 2011.

A great number of family members, friends and colleagues also assisted me in obtaining specimens. Thanks goes to: Allison Brown, Christina Davies, Dr. Ken Halanych, Peter Hovingh, Dr. Donald Reish, Mary Reynolds, Dr. Dan Shain, and Margery Thomas. As well, enormous thanks to my grandma, Audrey Woytiuk, and my uncle, Peter Woytiuk for their help and generosity in collecting leeches from Saint John, New Brunswick.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

An Overview of Annelid Diversity and Genome Size

Annelid Diversity

The phylum Annelida, the segmented worms, represents one of the most ecologically diverse phyla of animals on the planet. While their soft bodies made fossilization rare, these invertebrates were present in the seas as far back as the Palaeozoic (Rouse 2002). At present there are more than 17,000 described species which, apart from arid or aerial environments, can be found worldwide in a wide range of habitats, from hydrothermal vents to glacial ice.

All annelids share the same basic characteristic of a segmented, bilaterally symmetrical body plan; each segment is typically divided by septa but united by the digestive, vascular, muscular and nervous systems (Rouse 2002). Based on differences such as the presence of chaetae or antennae, the phylum Annelida is divided into two classes: Polychaeta and Clitellata, with the clitellates further divided into subclasses Oligochaeta and Hirudinea.

Physiologically and phylogenetically, annelids are well studied; the literature is rich with accounts of their various methods of reproduction and development (Sella 2006; Kuo 2009). Among species, adult forms can also differ greatly in morphology; in length they can be less than a millimetre (e.g., *Neotenotrocha* sp.) to upwards of three metres (e.g., *Eunice aphroditois, Megascolides australis*). As annelids radiated and adapted to new and diverse habitats, a number of different reproductive strategies to evolved among the groups, including asexual, gonochoric, and hermaphroditic methods (Rouse 2002; Sella 2006). A diversity of diets, also linked tightly with habitat, is also found within the annelids, including filter-feeding, selective deposit-feeding, geophagy, macrophagy, and sanguivory. Recent initiatives in deep sea exploration and genetic research have allowed us to move past the tip of the iceberg in discovering more about this fascinating phylum (e.g. Pradillon and Gaill 2009; Lang and Shain 2009).

Applications of Annelids in Science

A wide variety of annelids have been used for decades -- some even centuries -- in science and medicine (Lang and Shain 2009; Wells et al. 1993). The best known example of this is the European medicinal leech, *Hirudo medicinalis*, which has served as a popular alternative to blood-letting instruments for hundreds of years, and is now used to treat numerous illnesses and injuries from osteoarthritis to nerve compression (Adams 1988; Pilcher 2004; Heckmann et al. 2005). Despite the long-term use of the European medicinal leech, it was only recently discovered through the use of microsatellite analyses and DNA barcoding that some leeches widely marketed as *H. medicinalis* were in fact *H. verbena* (Siddall et al. 2007).

Modern science has also used annelids in practical applications outside the medical field. Polychaetes, a predominantly marine subclass of annelids, have more recently been used in concert with other fauna in human impact assessments (or biomonitoring) of marine environments (Díaz-Castañeda and Reish 2009). Earthworms (class Clitellata, subclass Oligochaeta) are similarly used to test and monitor contaminated soils (Römbke and Egeler 2009). In freshwater environments, the widely-distributed oligochaete *Lumbriculus variegatus* is frequently used to test for bioaccumulation of contaminants in benthic and epibenthic communities (Phipps et al. 1992).

Annelid Evolution

Despite the widespread use of annelids in many branches of science, there are still significant gaps in our understanding of their biology and evolution. The lack of whole-body preservation within the fossil record has left much to speculation and debate in determining phylogeny and evolutionary changes in morphological characteristics (Rouse 2002). To date

there is still much controversy in annelid phylogenetic relationships, in part because of the numerous differences between molecular and morphological analyses used to determine those relationships (Bartolomaeus et al. 2005; Bleidorn 2007). Phylogenetic analyses based on morphological data are often confounded by attempts at identifying character loss, while analyses based on molecular data (e.g., 18S rRNA) often do not find evidence to support the same relationships among annelids that morphological data suggest (McHugh 2000; Bleidorn 2007).

Previous research has already identified habitat as being a key factor in annelid diversification by comparing groups and separating specific habitat adaptations from plesiomorphies (primitive characters) of ancestral (stem) species (Purshke 1999). A select number of studies have also explored this relationship with habitat by measuring the genome sizes of certain groups in relation to physiological and life history characteristics (Sella et al. 1993; Soldi et al. 1994; Gambi et al. 1997; Gregory and Hebert 2002). While genome size data alone do not provide enough evidence to resolve phylogenetic relationships, when accompanied by biological parameters and life history data they can provide insight into other aspects of annelid evolution.

A large-scale, international effort to definitively assemble the annelid tree of life was initiated in 2011, funded by the US National Science Foundation and led by Dr. Kenneth Halanych (Principle Investigator) of Auburn University. This collaboration utilises a multi-tiered approach to fill in knowledge gaps of annelid diversity and evolutionary history in order to piece together a new phylogeny, utilizing approximately 3000 species. Genome size data will be useful to the project to increase knowledge of both annelid diversity and evolution.

Genome Size

Genome size is the total amount of DNA found within a single copy of one genome; it may also be referred to as C-value, which is a quantitative estimate of nuclear haploid DNA that is consistent within species (Vendrely and Vendrely 1948; Hinegardner 1976). Measurements of genome size are given either in picograms (pg) or number of nucleotide base pairs (bp), where 1 pg is equal to 978 Mbp (Doležel et al. 2003). C-value estimations are important for deciding on large-scale genomic sequencing targets, as the costs associated with sequencing are directly related to the quantity of DNA in a genome (Gregory 2005).

In animals, C-values range greatly in size from 0.02 pg (*Pratylenchus coffeae,* a plant-parasitic nematode) to 132.83 pg (*Protopterus aethiopicus*, the marbled lungfish) (Gregory 2012). Given this wide range of C-values, genome size is entirely unrelated to organismal complexity (Thomas 1971); this is because most eukaryotic DNA is non-coding, so a large genome does not imply a large number of genes. This large amount of non-coding DNA is central to a number of new studies that research the evolution of the genome.

Genome size has been shown to positively correlate positively with cell size in a variety of animal, plant, and protozoan taxa (see Gregory 2005). Genome size also correlates negatively with cell division rate in various taxa (Gregory 2005).

Additionally, in angiosperms, genome size is known to decrease with increasing latitude at high latitudes (Bennett et al. 1982). This is attributed to a selection against larger genome sizes given the short growing seasons and low temperatures at higher latitudes (Bennett et al. 1982).

The correlations between cell size and cell division rate mean that an organism's morphology, physiology, and/or development may be influenced by genome size. Most of the

studies on these relationships have focused on vertebrates and plants, while the much more diverse invertebrates have been less well examined. Some studies have shown positive correlations between body size and genome size and negative relationships with developmental rate in select invertebrate groups (see Gregory 2005). However, this is an area in need of much more investigation.

Genome Sizing in Annelids

Despite genome size estimation being one of the most basic forms of genomic measurement, very little work has been done on the annelids in this area. Polychaete species dominate the existing annelid genome size dataset (Connor et al. 1972; Sella et al. 1993; Soldi et al. 1994; Gambi et al. 1997). Oligochaete genome sizes have been explored in only one published study (Gregory and Hebert 2002), and leeches have no published genome size data whatsoever. Of the 17,000 known annelids, data for only 140 species (or approximately 0.8% of all known species) have been recorded in the Animal Genome Size Database (Gregory 2012). It is important to collect data representing a wide range of taxonomies with different life history traits to better understand the evolutionary consequences of genome size variation across species.

Despite a general lack of data, there is some preliminary indication that genome size may relate to developmental and environmental parameters in annelids. A few decades following the pioneering annelid genome size study by Connor et al. (1972), Sella et al. (1993) measured C-values for 9 benthic polychaete species of the genus *Ophryotrocha* and analysed the data for genome size relationships with body size and developmental rate. While no significant relationships were found, the species exhibited a rather low C-value average (0.4 pg), which led

Sella et al. (1993) to suggest that this trait could be unique to species occupying interstitial habitats.

On the other hand, Gregory and Hebert (2002) found no correlation between genome size and life history traits in earthworms. They did suggest, however, that a broader survey of genome sizes for annelids with varying life history traits may yield correlations that are, as of yet, unseen.

Objectives

To date, little has been done to investigate genome size diversity in annelids. Most studies focus on simply characterizing new species or physiological traits of medical interest. This limits both the ability to understand annelid evolution as well as the evolution of the genome itself.

The primary objective of the present study is to provide large scale investigations of genome size diversity in two subclasses of the annelid class Clitellata: Hirudinea (Chapter 2) and Polychaeta (Chapter 3). These genome size data can then be used to investigate possible patterns relating to development, body size, and geographical distribution, based on what has been observed in other animal groups.

In Chapter 2, genome size relationships with geographical location and life history traits such as body size, parental care, and diet are examined in leeches. Leeches were selected as a study group for their widespread range, unique life history traits, and entire absence within the literature in regards to genome size studies.

Chapter 3 explores the genome size diversity of polychaetes, with attention paid to taxonomic groups and geographical patterns. With previous groundwork laid out by Connor et al. (1972), Sella et al. (1993), Soldi et al. (1994), and Gambi et al. (1997), the goal of this study is to build on past findings and expand on current knowledge of polychaete genome sizes and evolution.

Questions and Predictions for Chapter 2

1. *What is the genome size range across the subclass Hirudinea?*

Based on existing data for other annelids, it is not obvious what the range of genome sizes will be in leeches. If they are similar to polychaetes, then their genome sizes may be quite variable. However, there are some developmental and other considerations that may lead to an expectation of more constrained genome sizes among leeches (see below).

2. *Is there a significant difference in genome size between the two leech orders, Arhynchobdellida and Rhynchobdellida?*

If the varied physiological and behavioural differences associated with parental care cause differential developmental rates, then the Arhynchobdellida and Rhynchobdellida will differ in genome size. The parental care provided by the Rhynchobdellida may allow for larger genome sizes than the Arhynchobdellida, given the reduced pressure for rapid development.

3. *Do leech C-values correlate with body size, and if so, does body size pose any kind of constraint to genome size?*

If leech body size is determined primarily by cell volume rather than cell number, then species with large body sizes will be expected to possess large genomes.

4. *Do leeches with differing diets have significantly different genome sizes?*

If there are opposing developmental pressures associated with the behavioural and physiological differences between macrophagous and sanguivorous leeches, then genome size will differ between the two groups.

5. *Is there a relationship between genome size and geographic region?*

If freshwater leeches in colder climates with shorter growing seasons face increased pressure to develop quickly, then leeches from higher latitudes will have smaller genome sizes.

Questions and Predictions for Chapter 3

1. *Do polychaete genome sizes still fall within the same range from previous studies?*

If genome size within the Polychaeta is relatively unconstrained, then polychaete C-values are expected to fall within 0.06 – 7.20 pg.

2. *Is there evidence of genome duplication within any polychaete groups?*

If the average congeneric species ratio is 2 or more, then polyploidy and cryptopolyploidy is likely responsible for increasing genome sizes within polychaetes.

3. *Is there a geographic difference in genome size for polychaetes?*

If latitude and its associated environmental factors affect genome size in polychaetes, then there will be a difference among groups from Polar and Temperate marine regions.

4. *Does genome size differ between polychaete subclasses Aciculata, Canalipalpata and Scolecida?*

If the three polychaete subclasses differ in genome size, then they likely have greatly different rates of development.

CHAPTER TWO

Hirudinid Genome Size Diversity and Biological Correlates

INTRODUCTION

There are currently approximately 700 described species within the subclass Hirudinea (the leeches), of which 43 are known to be found in Canada (Klemm 1985). This annelid subclass can be further subdivided into two orders: the Rhynchobdellida, which possess an exsertile proboscis, and the Arhynchobdellida, which lack this structure. They are found across a wide range of habitats distributed across six continents (Antarctica excluded), with species thriving in freshwater, marine and terrestrial environments (Sket and Trontelj 2008).

The leeches possess a wide range of life history traits. The maximum adult body size for this group varies greatly, from less than two centimetres in several glossiphoniid species to nearly half a metre in *Haementeria ghilianii* (Klemm 1985; Smithsonian Institution 2012). While sanguivory is thought to represent the ancestral diet, the majority of leech species are not in fact blood-feeders, but will feed on other invertebrates either by consuming them whole, or sometimes feeding on their body fluids or other tissues (Borda and Siddall 2004). Lastly, parental care can be found within the Rhynchobdellida, which brood their eggs (on the ventrum or stationary on a substrate) until they hatch and bring their offspring to their first meal (Sawyer 1971). The Arhynchobdellida, on the other hand, maintain the ancestral characteristic of laying their eggs within a protective cocoon and then cementing it on a hard surface prior to abandonment (Siddall and Burreson 1996).

Among some invertebrates, body size and genome size are positively correlated, while negative correlations have been observed between genome size and developmental rate in various groups of animals and plants (see Gregory 2005). Very little genome size data exist for annelids; while a handful of studies have looked at a few species of polychaetes (Sella et al.

1993; Soldi et al. 1994; Gambi et al. 1997) and oligochaetes (Gregory and Hebert 2002), the leeches have been left entirely unexplored. To date there have been no published genome size data for the Hirudinidea; some unpublished data exist in the Animal Genome Size Database but lack any analyses to explore potential significance. The purpose of this study is twofold: 1) to serve as an exploration of genome size diversity in leeches, expanding the Animal Genome Size Database and 2) to act as a preliminary examination of relationships between genome size and life history traits among leeches.

MATERIALS AND METHODS

Collection

Approximately 100 leeches were collected by hand from tundra ponds and lakes in Churchill, Manitoba in July 2009. Additional field collection of leeches took place from May – September 2011 in ponds and lakes within several Ontario locations (Guelph, Algonquin Provincial Park, and the Bruce Peninsula region) as well as Saint John, NB. Collecting continued in Guelph between May and July 2012. One leech was also collected near Auburn, Alabama in April 2012. During July of 2011, specimens were collected and identified from several locations in Utah and Nevada, U.S.A. by Peter Hovingh. A total of 454 specimens were collected from all locations in this manner.

Additional data for the species measured were collected by direct observation, from Klemm (1985), and elsewhere in the literature (Sawyer 1971; Siddall and Burreson 1995; Siddall and Burreson 1996; Rousset et al. 2008; Sket and Trontelj 2008). Such data included: collection site, body size (length-wise, relaxed state), environment (marine/freshwater/terrestrial), diet (sanguivorous or macrophagous), and reproductive method (brooding + type, non-brooding).

Identification

Specimens collected from Churchill were identified through DNA barcoding in November 2009 (see Hebert et al. 2003; Ratnasingham and Hebert 2007). Specimens were separated into eight species groups with the use of a phylogenetic tree builder on the Barcode of Life Database project management site (barcodinglife.org); three groups were identified down to species level and all were identified to at least family. The freshwater annelid guide by Klemm (1985) was employed to confirm or ascertain the morphological identification of leech specimens collected in 2009, 2011 and 2012 to species, genus, family, or order. Using the aforementioned methods, thirty-three specimen groups were identified to species, three to genus, three to family, and one to order.

All leech species names were verified as being the current accepted form utilizing the Integrated Taxonomic Information System (http://www.itis.gov). The following species name changes were made to be consistent with the ITIS: *Erpobdella obscura* to *Nephelopsis obscura*; *Helobdella elongata* to *Gloiobdella elongata*; *Batracobdella picta* to *Desserobdella picta*; and *Theromyzon biannulatum* to *Theromyzon bifarium*.

Preparation and Measurement

Live specimens were either placed in cryotubes and frozen at -80ºC (for later preparation) or prepared onto slides in the following way: each specimen was placed in a 5% ethanol solution for anaesthesia before bisection of the coelomic cavity for collection of haemocoelomic fluid on a slide. Slides were then air-dried and subsequently stained and prepared for Feulgen Image Analysis Densitometry (FIAD) following the methods outlined in Hardie et al. (2002). Leech specimens which were frozen were briefly thawed and prepared for FIAD in a similar fashion as

outlined above, minus the anesthetisation step. Feulgen Image Analysis Densitometry was chosen above other genome sizing methods (i.e. Flow Cytometry) for its previous known success rate in measuring annelid genome sizes, relative specimen processing speed, measurement repeatability, and other logistical considerations such as the ability to work with frozen materials and prepare samples in the field.

Measurements of the stained slides were performed using a Leica DM2500 microscope; images of the stained nuclei were captured using a Retiga EXi high resolution digital camera and analysed using the Bioquant Life Science image analysis system. This system allowed for the integrated optical density (IOD) of multiple nuclei to be measured on a given viewing field. To obtain an accurate IOD estimate, at least 50 nuclei were measured per viable sample. Leech nuclei were confirmed as diploid through comparison of nuclei with sperm from the same sample. *Gallus gallus domesticus* erythrocytes (1C = 1.25pg) were used as a standard for conversion of the unknown mean IOD values into C-values due to the relative similarity in DNA compaction levels (see Hardie et al. 2002).

Data Analysis

Genome size and life history data were organized according to species (or family/next known taxonomic level, where species data were unavailable). Analyses comparing genome size and body size were performed using Pearson's correlations at the species, genus and family level for all leeches, taking into account the non-independence of phylogenetically related species by using hierarchical taxonomic correlations (Gregory 2000) using nested averages at the species, genus and family levels. Both body size data and genome size data were compared at the order level, with and between groups, using two-sample equal variance t-tests. For analysis of genome size and geographic range, collection sites were divided into three regions: Western USA

(Nevada, Utah), Arctic (Churchill), and Eastern Canada (sites in Ontario and New Brunswick), with site-specific species genome size data used (Table 2.2). Genome size and region were compared at the species level using one-way ANOVA and Tukey *post hoc* tests. Unpublished genome size data for five leech species not represented in the specimen collection were also used from the Animal Genome Size Database (Gregory 2012) for addition to the analyses for parental care, body size and diet.

RESULTS

Leech Genome Size Overview

New genome size estimates for 35 species of leeches measured ranged from 0.24 pg (*Haemopis marmorata*) to 1.23 pg (*Placobdella ornata*) (Table 3.1), which is approximately a five-fold difference, and have an average of 0.62 pg. Sixteen species in the Arhynchobdellida had a range of 0.24 pg (*Haemopis marmorata*) to 0.84 pg (*Mooreobdella tetragon*), a 3.5-fold difference, while nineteen species in the Rhynchobdellida ranged from 0.28 pg (Unknown #7) to 1.23 pg (*Placobdella ornata*), which is a 4-fold difference.

Five of the 35 species analysed overlapped with unpublished estimates logged in the Animal Genome Size Database (Gregory 2012); apart from *P. ornata*, which was double the previous estimate, new estimates averaged 5% higher than previously reported data.

Parental Care

Leeches that brood their offspring, the Rhynchobdellida, had genome sizes significantly larger than the non-brooding Arhynchobdellida (t-test, $p < 0.001$) (Table 2.2; Figure 2.1). Within

the Rhynchobdellida, leech species which brood their offspring on a substrate had significantly larger genome sizes than those which brood on their ventrum (t-test, $p < 0.05$) (Table 2.2).

Maximum Adult Body Size

Across all leech species there was a significant negative correlation between genome size and maximum adult body size ($r^2 = 0.2178$, $p < 0.004$) (Table 2.2; Figure 2.2). This relationship maintained significance with the removal of the largest body size outlier, *H. ghilianii* ($r^2 = 0.2671$, $p < 0.002$). Significance was also maintained at the genus level ($r^2 = 0.242$, $p < 0.04$), but not at the family level ($r^2 = 0.138$, $p = 0.4684$). Within the Arhynchobdellida and the Rhynchobdellida, there were no correlations between genome size and maximum adult body size ($r^2 = 0.1058$, $p = 0.19$; $r^2 = 0.0406$, $p = 0.39$). Maximum adult body size did not differ significantly between the Arhynchobdellida and the Rhynchobdellida (t-test, $p = 0.25$).

Diet

There was no significant difference found between macrophagous leeches and sanguivorous leeches (t-test, $p = 0.19$) (Table 2.2).

Geography

There was no significant difference in genome size found between leeches collected in Western USA (mean = 0.61 ± 0.11 pg; n = 7), the Arctic (mean = 0.63 ± 0.07 pg; n = 7), or Eastern Canada (mean = 0.68 ± 0.05 pg; n = 30) (ANOVA, $p = 0.78$) (Table 2.2).

A one-way ANOVA test yielded no significant difference between *G. complanata* collected in Western USA (mean = 1.06 ± 0.09 pg; n = 3), the Arctic (mean = 1.01 ± 0.09 pg; n = 13) or Eastern Canada (mean = 0.92 ± 0.05 pg; n = 10) (i = 0.64), nor for *H. stagnalis* found in

these regions (Western USA: mean = 0.55; n = 1) (Arctic: mean = 0.73; n = 1) (Eastern Canada: mean = 0.78 ± 0.06 pg; n = 9) ($p = 0.52$). There was a statistically significant difference between *N. obscura* groups collected in the three major geographic regions (F (2,40) = 5.989, $p < 0.005$) (Table 3.3). A Tukey *post-hoc* test revealed that genome sizes were significantly larger in *N. obscura* found in Western USA (0.62 ± 0.07 pg, q = 4.108, $p < 0.01$) and Eastern Canada (0.55 ± 0.05 pg, q = 4.472, $p < 0.01$) compared to Arctic *N. obscura* (0.47 ± 0.15 pg). *N. obscura* collected in Western USA also had a significantly larger genome size than those from Eastern Canada (q = 1.917, $p < 0.01$).

DISCUSSION

Parental Care

The full genome size range of the leeches studied spanned a five-fold difference from 0.24 pg to 1.23 pg, with the Rhynchobdellida presenting a wider range and significantly higher genome sizes than the more restricted Arhynchobdellida. This difference may be related to the unique brooding behaviour found only within the glossiphoniid family of the Rhynchobdellida (Sawyer 1971). Brooding behaviour within these leeches is believed to be a recently acquired trait, while cementing and abandoning a hardened cocoon of eggs is generally accepted to be a pleisomorphy (Sawyer 1971; Siddall and Burreson 1996). The stable environment provided by brooding for offspring to develop may have eased the pressure for rapid development, which would have allowed for an increase in genome size.

There are traits other than parental care that divide the Arhynchobdellida and the Rhynchobdellida, however, such as whether or not the body is dorso-ventrally flattened. Dorso-ventral flattening is a characteristic of the family Glossiphoniidae, and could be a physiological

prerequisite that allowed for brooding behaviour to evolve. As the Rhynchobdellida represented in the data all belong to Glossiphoniidae (apart from, possibly, Unknown #6), a strong argument can be made for brooding behaviour to be a major contributing factor to the observed difference. Additionally, a marked, significant difference in genome size can be found within the glossiphoniids depending on their particular method of brooding. As Sawyer (1971) describes, three types of parental care can be found within the glossiphoniids: brooding eggs in a permanent internal pouch, brooding eggs on a substrate, or brooding eggs attached to the ventrum. The latter two types were represented in the leeches from the present study, where it was found that substrate-brooders (of the subfamilies Glossiphoniinae and Theromyzinae) had significantly larger genome sizes than ventrum-brooders (subfamily Haementeriinae). As it was previously mentioned, parental care evolved more recently, so it would be rather unsurprising to observe an increase in genome size between the two groups expressing different behaviours. However, this is not the case; within the glossiphoniids there is a subsequent decrease in genome size between substrate-brooding species and ventrum-brooding species, when substrate-brooding is seen as an intermediate step between non-brooding behaviour and brooding eggs on the ventrum (Sawyer 1971). As suggested with brooding overall, perhaps the substrate-brooding behaviour provides a more stable environment (i.e. hiding under a rock or within a crevice), while ventrum-brooding species are more motile and therefore could face a higher predation rate, which would increase the rate of development, decreasing genome size. This unique relationship within Glossiphoniidae should be further explored, and include "marsupial" species with an internal brood pouch such as *Marsupiobdella africana* and *Maiabdella batracophila* (Sawyer 1986a).

Maximum Adult Body Size

A significant inverse correlation between genome size and maximum adult body size was found across all leech species and genera. This is the opposite pattern of what has previously been observed among some invertebrates when comparing genome size with body size; the relationship between these two variables has been reported as significantly positive in turbellarian flatworms and copepod crustaceans (Gregory et al. 2000), aphids (Finston et al. 1995), flies (Ferrari and Rai 1989), and molluscs (Hinegardner 1974). This means that it is unlikely that body size is restricted by a small genome size (or vice-versa) and also that body size is not related to cell volume, as cell volume increases as genome size increases (Cavalier-Smith 1985). It can also be noted that genome size is not constrained by segment number, as the majority of leeches have a fixed number of segments with 32 (Weisblat et al. 1988). What could cause this negative correlation between body size and genome size? Flemming et al. (2000) found that the hypodermis of rhabditid nematodes consists of a single syncytium, which exhibits a high level of endopolyploidy, and that the syncytium drives body size rather than number of cells overall, as with mammals and many invertebrates. The hypodermis of leeches has been examined and was not observed to contain polyploid cells, and therefore unlikely to determine body size in the same way as nematodes (Rolleston and Jackson 1888). It is then possible that growth rate drives body size, if both segment number and growth time are relatively constant, as fast growth rate has been observed in species with small genome sizes (Wyngaard et al. 2005).

Unlike genome size, body size did not differ significantly between the two leech orders. It is possible this is an artifact of sample size, and perhaps given a larger number of species, the Rhynchobdellida would exhibit smaller body sizes (given the significantly larger genomes); the beginnings of this trend is evident when the Arhynchobdellida and the Rhynchobdellida are differentiated on the body size-genome size regression (Figure 2.3).

While it is not significant compared to the Rhynchobdellida, the Arhynchobdellida exhibit a large diversity of maximum adult body sizes. Members of the Hirudiniformes (which include Haemadipsidae, Haemopidae, and Hirudinidae from the present study) exhibit some of the largest body sizes within the Arhynchobdellida, and also have a limited C-value range (0.24 – 0.32 pg). Zulhisyam et al. (2011) demonstrated that light intensity and temperature play a significant role in the development from egg to juvenile in *Hirudinea* sp. It is not yet known how or if these conditions affect the adult growth of leeches. For instance, as in Zulhisyam et al. (2011), if zero light intensity and 25 - 28°C water temperature allows for optimal juvenile growth and survival, then surviving adults from the same treatment would be larger than counterparts from other treatments. Given the varied habitat and latitude range of the Hirudiniformes, a study based on a hypothesis such as this may reveal a gradient.

Future studies should explore environmental effects on leech lifetime development among contrasting groups, such as the Hirudiniformes and Erpobdelliformes, using both laboratory and field studies. This would also provide a basis for the developmental rate within the Arhynchobdellida for any future comparisons with the Rhynchobdellida in regards to genome size and developmental rate.

Geography

Genome size does not differ between leech species found in the categorized geographical ranges of Western USA, the Arctic, and Eastern Canada. This suggests that latitude has little or no effect on patterns of genome size diversity in leeches on a large scale. A significant difference was observed on a small scale, however, among *N. obscura* found in each geographic range (Table 3.3). These differences are potentially caused by two types of intraspecific variation.

"Orthodox" intraspecific variation at the species level is due to chromosome polymorphisms and/or spontaneous aberrations, or due to non-recognized taxonomic heterogeneity (cryptic species), while "unorthodox" variation requires *ad hoc* assumptions for explanation (i.e. genome plasticity during the course of development or reproduction) (Greilhuber 1981). It is likely this is orthodox intraspecific variation, with the variation possibly indicating that *N. obscura* is in fact a cryptic species complex that has diversified under the unique selection pressures each range applies, and/or simply due to the abiotic reproductive barriers between the populations. On the other hand, given that the genome sizes of *N. obscura* decrease with higher latitudes, it is possible that the lifestyle of *N. obscura* requires faster development in more northern regions, resulting in smaller C-values. In either case, it would be beneficial to utilize additional genetic analyses, such as DNA barcoding, to resolve the relationships between the *N. obscura* groups found in distinct geographic locations (cf. Bely and Weisblat 2006). DNA barcoding has previously proven effective in identifying cryptic species within the annelid class Polychaeta; one study found 34 morphologically identified species as representing 88 provisional barcode-supported species (Carr et al. 2011).

Genome Size and Chromosome Number

Variation in genome size was observed among some of the more closely related leech groups. The species-rich genera *Mooreobdella*, *Helobdella*, and *Placobdella* exhibit variation that appears to be somewhat stepwise or discontinuous. Several studies have previously quantified chromosome numbers for a handful of leech species (see Chichocka and Bielecki 2008), but not enough species data yet exist to provide a comparison between karyotype and genome size to determine whether the observed variation is due to polyploidy or

cryptopolyploidy (where genome contents multiply without affecting the number of chromosomes).

Notes on Methodology

Sanguivorous leech specimens engorged with blood did not provide good data unless extra care was taken to rupture only the body wall and not the gut. However, in several cases where the gut was ruptured and it was known what species the leech had been feeding on (it was directly removed from its prey), the genome size provided by FIAD was comparable to that of the prey. For example, the gut contents of two *Desserobdella picta* specimens removed from *Rana sylvatica* yielded genome sizes of 5.32 pg, and the reported estimates for *R. sylvatica* range from 5.49 – 6.50 pg (Gregory 2012). This indicates it may be possible to acquire accurate genetic information from the gut contents of sanguivorous leeches, which is congruous with the findings within other blood-feeding invertebrates (Gariepy et al. 2012).

Gut contents of macrophagous leeches appeared on a stained slide as debris. This typically entirely renders a sample useless, again demonstrating the importance of taking care to avoid cutting into the gut while preparing slides. If specimens have been kept frozen rather than preserved in ethanol, it may be possible to re-sample from the specimen to produce new slides for staining and measurement, or to use alternative genome size estimation methods such as Flow Cytometry.

Future Directions

One of the main goals of the present study was to serve as a foundation and starting point for future studies. This has been accomplished given the number of new questions and hypotheses arising from the data. As mentioned in the preceding sections, there are several new

directions that should be taken to build upon the results of the present study. DNA barcoding is a necessary step to be taken in order to identify 6 potentially new species, resolve phylogenetic relationships among all species, and add to the Barcode of Life database.

CONCLUSIONS

This study presented the first comprehensive analysis of genome size estimates of species from the clitellate subclass Hirudinea. New genome size estimates generated represent approximately 60% of the 43 leech species found in Canada, and nearly 6% of all leeches. Thirty of the new genome size estimates belong to species yet to be recorded in the Animal Genome Size Database.

Several statistically significant findings were made using the new leech genome size estimates, unpublished leech genome size data from the Animal Genome Size Database, and life history data. Leech species that provide parental care to their offspring have larger genome sizes than those which do not, although this generally segregates along taxonomic lines and could be related to additional characteristics of the taxa in question. However, the mode of parental care within leech groups may also be linked to differences in genome size. A novel inverse relationship between genome size and body size was also discovered, which raises interesting questions about the relationships between genome size, cell size, cell division, cell number, and body size in these animals. Lastly, the genome sizes of erpobdellid *Nephelopsis obscura* were found to differ between geographic locations, with C-values decreasing at higher latitudes, suggesting this one species may actually be several species within a cryptic species complex.

Table 2.1. New mean haploid genome size estimates (GS, in pg) for hirudinean species, including the standard error (SE) for estimates obtained from more than one individual (N = number of specimens per species). Genome size estimates were measured using Feulgen Image Analysis Densitometry with at least 50 nuclei measured per specimen, compared to a blood standard from *Gallus gallus domesticus* (GS = 1.25 pg).
[a] denotes values from the Animal Genome Size Database

Order	Suborder	Family	Subfamily	Species	GS (pg)	SE	N
Arhynchobdellida					**0.46**	**0.04**	**18**
Arhynchobdellida	Erpobdelliformes	Erpobdellidae	Erpobdellinae	*Erpobdella lahontana*	0.54		1
Arhynchobdellida	Erpobdelliformes	Erpobdellidae	Erpobdellinae	*Erpobdella punctata punctata*	0.53	0.01	36
Arhynchobdellida	Erpobdelliformes	Erpobdellidae	Erpobdellinae	*Mooreobdella bucera*	0.51	0.01	3
Arhynchobdellida	Erpobdelliformes	Erpobdellidae	Erpobdellinae	*Mooreobdella fervida*	0.52		1
Arhynchobdellida	Erpobdelliformes	Erpobdellidae	Erpobdellinae	*Mooreobdella melanostoma*	0.56	0.03	2
Arhynchobdellida	Erpobdelliformes	Erpobdellidae	Erpobdellinae	*Mooreobdella microstoma*	0.46	0.08	3
Arhynchobdellida	Erpobdelliformes	Erpobdellidae	Erpobdellinae	*Mooreobdella tetragon*	0.84	0.17	2
Arhynchobdellida	Erpobdelliformes	Erpobdellidae	Erpobdellinae	*Nephelopsis obscura*	0.53	0.02	43
Arhynchobdellida	Erpobdelliformes	Erpobdellidae	Erpobdellinae	*Nephelopsis sp. (Unknown #1)*	0.45	0.03	40
Arhynchobdellida	Erpobdelliformes	Erpobdellidae		*Dina dubia*	0.68	0.01	63
Arhynchobdellida	Erpobdelliformes	Erpobdellidae		*Dina parva*	0.28		1
Arhynchobdellida	Erpobdelliformes	Erpobdellidae		*Unknown #2*	0.69		1
Arhynchobdellida	Hirudiniformes	Haemadipsidae	Haemadipsinae	*Philaemon pungens*	0.27[a]		
Arhynchobdellida	Hirudiniformes	Haemopidae	Haemopinae	*Haemopis grandis*	0.26	0	2
Arhynchobdellida	Hirudiniformes	Haemopidae	Haemopinae	*Haemopis lateromaculata*	0.32	0.06	5
Arhynchobdellida	Hirudiniformes	Haemopidae	Haemopinae	*Haemopis marmorata*	0.24	0.01	2
Arhynchobdellida	Hirudiniformes	Hirudinidae	Hirudinariinae	*Hirudo medicinalis*	0.23[a]		
Arhynchobdellida	Hirudiniformes	Hirudinidae	Macrobdellinae	*Macrobdella decora*	0.32	0.01	12
Rhynchobdellida					**0.75**	**0.05**	**22**
Rhynchobdellida		Glossiphoniidae	Glossiphoniinae	*Desserobdella picta*	0.73	0.03	8
Rhynchobdellida		Glossiphoniidae	Glossiphoniinae	*Glossiphonia complanata*	0.98	0.05	26
Rhynchobdellida		Glossiphoniidae	Glossiphoniinae	*Placobdella hollensis*	1.03	0.53	7
Rhynchobdellida		Glossiphoniidae	Glossiphoniinae	*Placobdella montifera*	0.64		1
Rhynchobdellida		Glossiphoniidae	Glossiphoniinae	*Placobdella multilineata*	0.77	0.18	4
Rhynchobdellida		Glossiphoniidae	Glossiphoniinae	*Placobdella ornata*	1.23	0.18	12
Rhynchobdellida		Glossiphoniidae	Glossiphoniinae	*Placobdella papillifera*	0.97	0.17	11
Rhynchobdellida		Glossiphoniidae	Glossiphoniinae	*Placobdella parasitica*	0.69	0.06	10
Rhynchobdellida		Glossiphoniidae	Haementeriinae	*Alboglossiphonia australiensis*	0.46[a]		
Rhynchobdellida		Glossiphoniidae	Haementeriinae	*Gloiobdella elongata*	0.89	0.01	9
Rhynchobdellida		Glossiphoniidae	Haementeriinae	*Haementeria ghilianii*	0.52[a]		
Rhynchobdellida		Glossiphoniidae	Haementeriinae	*Helobdella europaea*	0.35[a]		
Rhynchobdellida		Glossiphoniidae	Haementeriinae	*Helobdella papillata*	1.04		1
Rhynchobdellida		Glossiphoniidae	Haementeriinae	*Helobdella robusta*	0.37		1
Rhynchobdellida		Glossiphoniidae	Haementeriinae	*Helobdella sp. (Unknown #3)*	0.91		1

Rhynchobdellida	Glossiphoniidae	Haementeriinae	*Helobdella stagnalis*	0.76	0.05	10
Rhynchobdellida	Glossiphoniidae	Haementeriinae	*Helobdella triserialis*	0.83		1
Rhynchobdellida	Glossiphoniidae	Theromyzinae	*Theromyzon bifarium*	0.85	0.09	2
Rhynchobdellida	Glossiphoniidae	Theromyzinae	*Theromyzon rude*	0.92		1
Rhynchobdellida	Glossiphoniidae		*Unknown #4*	0.64		1
Rhynchobdellida	Glossiphoniidae		*Unknown #5*	0.28		1
Rhynchobdellida			*Unknown #6*	0.61	0.07	10

Table 2.2. Leech life history and collection data used for analyses. Body size refers to the maximum adult body size. Habitat type is indicated as freshwater (FW), terrestrial (TR) and amphibious (AM). Parental care includes brooding (B), brooding on substrate (BS), brooding on ventrum (BV) and non-brooding (NB). Diet type is either macrophagous (M) or sanguivorous (S). Region of collection is indicated as Western USA (WU), Eastern Canada (EC) or Arctic (AR).

Order	Family	Genus	Body Size (mm)	Habitat	Parental Care	Diet	Region
Arhynchobdellida	Erpobdellidae	*Unknown #2*	-	FW	NB	M	EC
Arhynchobdellida	Erpobdellidae	*Dina dubia*	60	FW	NB	M	EC, WU
Arhynchobdellida	Erpobdellidae	*Dina parva*	30	FW	NB	M	EC
Arhynchobdellida	Erpobdellidae	*Erpobdella lahontana*	30	FW	NB	M	WU
Arhynchobdellida	Erpobdellidae	*Erpobdella punctata punctata*	100	FW	NB	M	AR, EC
Arhynchobdellida	Erpobdellidae	*Mooreobdella bucera*	30	FW	NB	M	EC
Arhynchobdellida	Erpobdellidae	*Mooreobdella fervida*	50	FW	NB	M	EC
Arhynchobdellida	Erpobdellidae	*Mooreobdella melanostoma*	55	FW	NB	M	EC
Arhynchobdellida	Erpobdellidae	*Mooreobdella microstoma*	50	FW	NB	M	EC
Arhynchobdellida	Erpobdellidae	*Mooreobdella tetragon*	40	FW	NB	M	EC
Arhynchobdellida	Erpobdellidae	*Nephelopsis obscura*	100	FW	NB	M	AR, EC, WU
Arhynchobdellida	Erpobdellidae	*Nephelopsis sp. (Unknown #1)*	150	FW	NB	M	AR
Arhynchobdellida	Haemadipsidae	*Philaemon pungens*	20	TR	NB	S	-
Arhynchobdellida	Haemopidae	*Haemopis grandis*	225	FW	NB	M	EC
Arhynchobdellida	Haemopidae	*Haemopis lateromaculata*	60	FW	NB	M	EC, WU
Arhynchobdellida	Haemopidae	*Haemopis marmorata*	100	FW	NB	M	EC
Arhynchobdellida	Hirudinidae	*Hirudo medicinalis*	200	AM	NB	S	-
Arhynchobdellida	Hirudinidae	*Macrobdella decora*	150	FW	NB	S	EC
Rhynchobdellida	Glossiphoniidae	*Alboglossiphonia australiensis*	-	FW	BV	M	-
Rhynchobdellida	Glossiphoniidae	*Desserobdella picta*	25	FW	BS	S	EC
Rhynchobdellida	Glossiphoniidae	*Gloiobdella elongata*	25	FW	BV	M	EC
Rhynchobdellida	Glossiphoniidae	*Glossiphonia complanata*	25	FW	BS	M	AR, EC, WU
Rhynchobdellida	Glossiphoniidae	*Haementeria ghilianii*	457	FW	BV	S	-
Rhynchobdellida	Glossiphoniidae	*Helobdella europaea*	25	FW	BV	M	-
Rhynchobdellida	Glossiphoniidae	*Helobdella papillata*	14	FW	BV	M	EC
Rhynchobdellida	Glossiphoniidae	*Helobdella robusta*	30	FW	BV	M	WU
Rhynchobdellida	Glossiphoniidae	*Helobdella sp. (Unknown #3)*	-	FW	BV	M	EC
Rhynchobdellida	Glossiphoniidae	*Helobdella stagnalis*	14	FW	BV	M	AR, EC, WU
Rhynchobdellida	Glossiphoniidae	*Helobdella triserlialis*	29	FW	BV	M	EC
Rhynchobdellida	Glossiphoniidae	*Placobdella hollensis*	30	FW	BS	S	EC
Rhynchobdellida	Glossiphoniidae	*Placobdella montifera*	16	FW	BS	S	EC
Rhynchobdellida	Glossiphoniidae	*Placobdella multilineata*	50	FW	BS	S	EC
Rhynchobdellida	Glossiphoniidae	*Placobdella ornata*	40	FW	BS	S	EC
Rhynchobdellida	Glossiphoniidae	*Placobdella papillifera*	45	FW	BS	S	EC
Rhynchobdellida	Glossiphoniidae	*Placobdella parasitica*	65	FW	BS	S	EC
Rhynchobdellida	Glossiphoniidae	*Theromyzon bifarium*	26	FW	BS	S	EC
Rhynchobdellida	Glossiphoniidae	*Theromyzon rude*	30	FW	BS	S	EC
Rhynchobdellida	Glossiphoniidae	*Unknown #4*	-	FW	B	-	EC
Rhynchobdellida	Glossiphoniidae	*Unknown #5*	3	FW	B	-	AR
Rhynchobdellida		*Unknown #6*	3	FW	BV	-	AR

Table 2.3. Genome size variation among three distinct geographical ranges for *Nephelopsis obscura*.

	N	Genome Size (pg)	SD
Arctic	12	0.47	0.15
Eastern Canada	25	0.55	0.05
Western USA	6	0.62	0.07

Figure 2.1. Significant genome size differences between the Arhynchobdellida and the Rhynchobdellida.

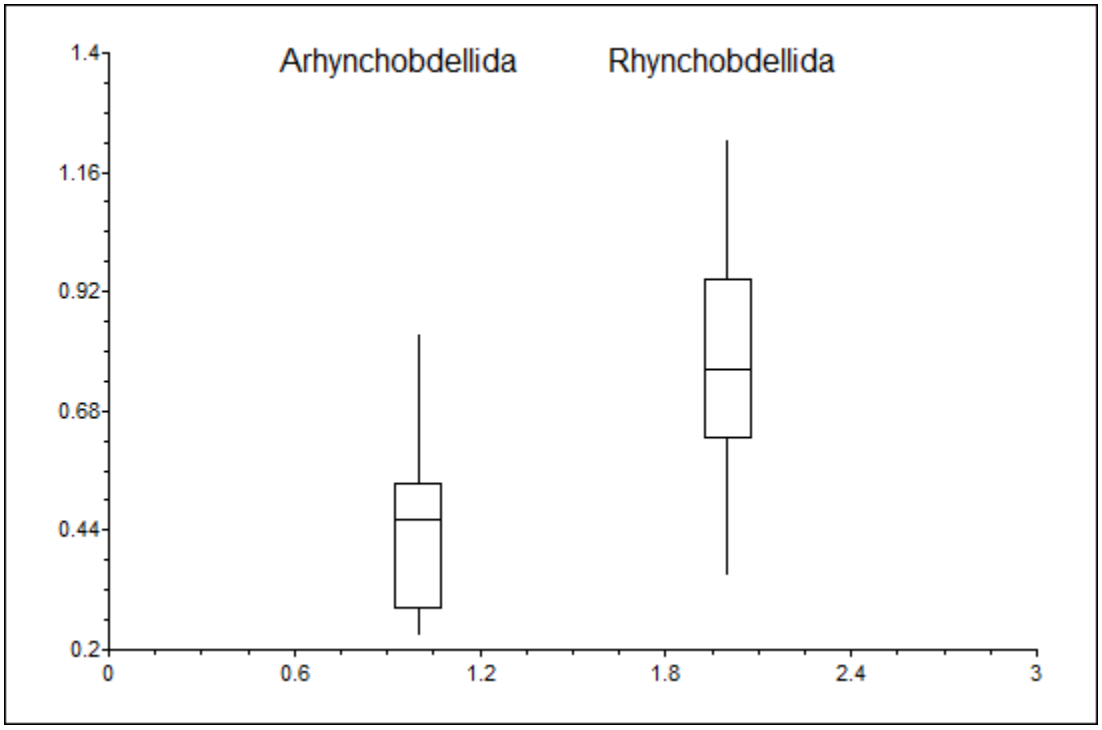


Figure 2.2. Relationship between leech genome size and body size across species. Body size data were not available for all species.

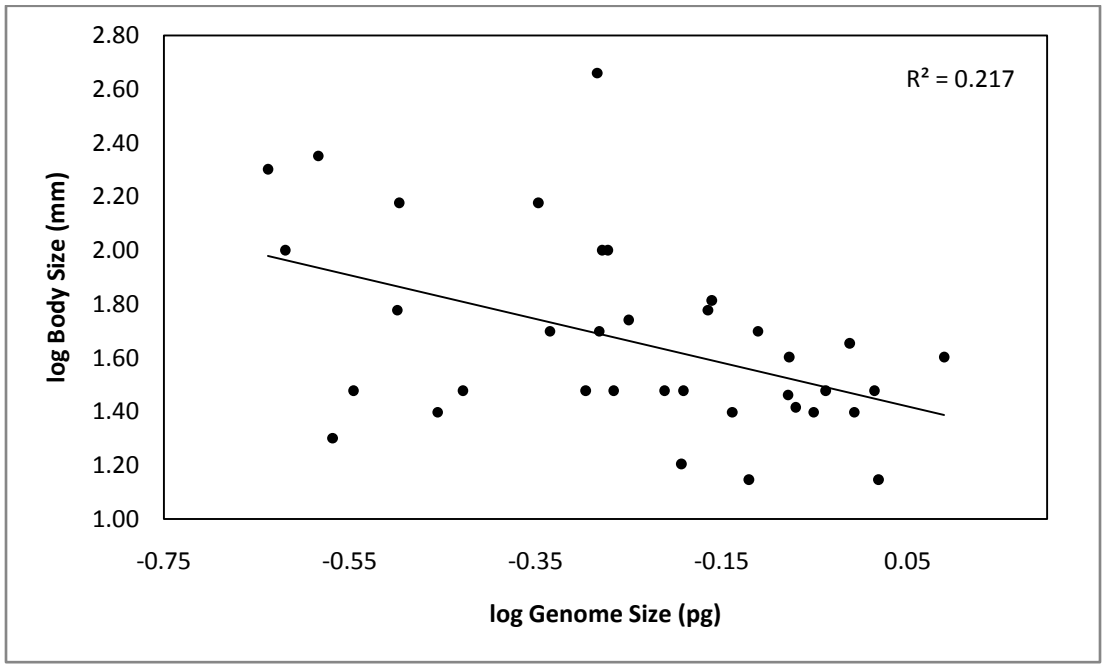
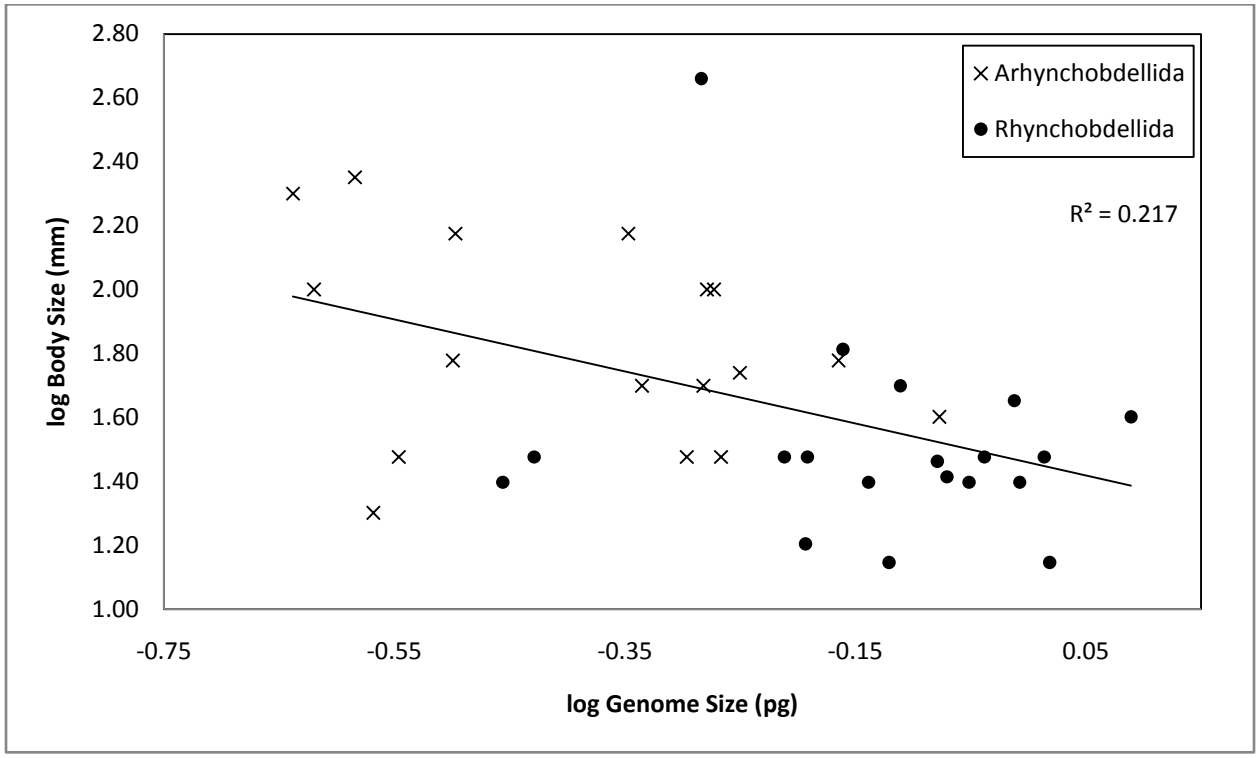


Figure 2.3. Relationship between genome size and body size in the Arhynchobdellida and the Rhynchobdellida. Body size data were not available for all species.



CHAPTER THREE

Expanding on Marine Polychaete Genome Sizes

INTRODUCTION

The Polychaeta is a diverse class of annelids that is distinguished by the paired parapodia found on each body segment, each of which have chitinous bristles called chaetae. They are predominantly marine (and, unlike oligochaetes or leeches, are strictly aquatic), with habitats ranging from intertidal mud- and sand-flats to the deepest ocean depths.

Research on this group has primarily focused on taxonomy, morphology, ecology and larval biology. Many different types of polychaetes have also been researched for their various unique properties, such as neurotoxin secretions and regeneration, which could someday be applied to research in the medical field (Bon et al. 1985; Pfeifer et al. 2012). Despite their widespread use within many branches of science, there are still significant gaps in understanding polychaete evolution. Recent evolutionary work on polychaetes has progressed from the evo-devo movement, comparing developmental stages of various species to better understand characteristics such as segmentation (Hill and Savage 2009). However, as with other annelids, the lack of whole-body preservation within the fossil record has left much to speculation and debate in determining phylogeny and evolutionary changes in morphological characteristics (Rouse 2002). Phylogenetic relationships among polychaetes and across the broader annelid phylum have been a point of contention as morphological and molecular analyses do not always reach the same conclusions (Bartolomaeus et al. 2005; Bleidorn 2007). For these reasons and many others, a large-scale comprehensive research project called WormNet II: Assembling the Annelid Tree of Life has been launched at Auburn University, AL, and other international institutions with funding from the National Science Foundation.

In one of the first annelid genome size studies since the 1970s, Sella et al. (1993) measured C-values for 9 benthic polychaete species of the genus *Ophryotrocha* and compared them with body size and developmental rate. While no significant relationship was found, the low C-values (0.4pg) for these polychaetes led Sella et al. (1993) to suggest that this characteristic could be unique to species occupying interstitial habitats.

Following Sella et al. (1993), Gambi et al. (1997) demonstrated that certain types of environment may play a role in shaping genome sizes in polychaetes. Through the measurement of C-values for 43 polychaete genera and species, this study found that interstitial taxa have a restricted range of C-values compared to macrobenthic polychaete species. They suggested that the variable interstitial environment placed selective pressure on polychaetes inhabiting this environment (*r*-selection, e.g. Satchell 1980) which resulted in a reduction in body size, increased rate of development and subsequent reduction in genome size. Their analysis also suggested that, in polychaetes, genome size has no clear phyletic significance at the order level.

The present study aims to augment the current polychaete dataset available in the Animal Genome Size Database. This additional data will be useful in furthering research currently underway as part of WormNet II.

MATERIALS AND METHODS

Collection

Whole or partial polychaete specimens were provided either live or frozen (-80ºC) from Don Reish at California State University, Long Beach (Long Beach, CA, USA) or obtained by sub-sampling the frozen tissue collection in the lab of Ken Halanych at Auburn University

(Auburn, AL, USA). Approximately 155 specimens were acquired from these collaborators, accompanied by collection location and taxonomic identification data.

All polychaete species names included in the present study were verified as being the current accepted form utilizing the Integrated Taxonomic Information System (http://www.itis.gov) and the World Register of Marine Species (http://www.marinespecies.org).

Preparation and Measurement

All polychaete specimens were frozen at -80ºC prior to slide preparation. The "Freeze-Squash" method was employed in preparing frozen samples. First, a small tissue sample from approximately the head region was placed in a drop of 40% acetic acid on a slide and torn apart using pins until no significant tissue aggregations remained. Then a cover slip was placed on top of the sample, held in place by three clothes-pegs, and placed on dry ice for approximately 3 minutes or until the slide sample had frozen through. Following this, the cover slip was removed using a razor blade to pry it off the slide. Finally, the slide was fixed in reagent alcohol for several minutes before drying on a tray. Feulgen Image Analysis Densitometry was chosen above other genome sizing methods (i.e. Flow Cytometry) for its previous known success rate in measuring annelid genome sizes, relative specimen processing speed, measurement repeatability, and other logistical considerations such as the ability to work with frozen materials and prepare samples in the field.

Measurements of the stained slides were performed using a Leica DM2500 microscope; images of the stained nuclei were captured using a Retiga EXi high resolution digital camera and analysed using the Bioquant Life Science image analysis system. This system allowed for the integrated optical density (IOD) of multiple nuclei to be measured on a given viewing field. To

obtain an accurate IOD estimate, no less than 50 nuclei were measured per viable sample. *Gallus gallus domesticus* erythrocytes (1C = 1.25 pg) and *Oncorhynchus mykiss* (1C = 2.5 pg) were used as standards for conversion of the unknown mean IOD values into C-values (see Hardie et al. 2002).

Data Analysis

Genome size data were organized according to species (or genera, family, or next known taxonomic level, where species data were unavailable). For data analysis, collection sites were divided into two geographic marine coastal regions: Polar (Antarctica and Norway) and Temperate (British Columbia, Washington, California, Maine, North Carolina, South Carolina, and Alabama). Genome size and geographic region were compared the species level using a two-sample equal variance t-test. Relationships between genome size and phylogeny were analysed using hierarchical taxonomic correlations (Gregory 2000) in which nested averages were calculated at the species, genus, family, and subclass level apart from Polychaeta incertae sedis (a subclass for taxa of uncertain position in the Polychaeta class phylogeny). Subclasses comprised of nested family averages were compared using one-way ANOVA and Tukey-Kramer MSD.

The average congeneric species ratio (CSR) was calculated for the 22 genera within the dataset containing multiple species. In a given genus, the largest C-value was divided by the second largest, then the second largest C-value divided by the third largest, and so on, until the smallest value had been used as the divisor (see below for sample formula). The results of these divisions were then averaged together to calculate the CSR for that genus. The CSR values for all genera were averaged to obtain the average CSR.

$$\textbf{CSR}: \frac{\frac{\text{C-value 3}}{\text{C-value 2}} + \frac{\text{C-value 2}}{\text{C-value 1}}}{2}$$

Additional species data from the Animal Genome Size Database were included for hierarchical taxonomic correlations and congeneric species ratios.

RESULTS

Polychaete Genome Size Overview

New genome size estimates for 61 polychaete species within 29 families ranged from 0.41 pg (*Glycera tridactyla*) to 10.29 pg (Unknown #6), a 25-fold difference, and an average of 1.65 pg (Table 3.1). Seven species had previous records in the Animal Genome Size Database; new estimates of these species were approximately 19% smaller than previous numbers. The new genome size estimates increase the polychaete data in the Animal Genome Size Database from 97 species records to 151. The number of families represented in the Database increased by six, to 40.

Congeneric Species Ratio

The average congeneric species ratio was 2.01. There was no significant relationship between the average ratio of congeners and the average genome size of genera for 22 genera ($r^2 = 0.0415$, $p = 0.35$) (Figure 3.1).

Geography

A two-sample equal variance t-test found a significant difference in C-values between collection regions ($p < 0.001$). Polar region polychaetes had significantly larger genomes than Temperate polychaetes (Table 3.2).

Subclasses

A one-way ANOVA did not find a significant difference in genome sizes between subclasses Aciculata (mean = 2.26 ± 0.39 pg; n = 14), Canalipalpata (mean = 1.66 ± 0.24 pg; n = 13), and Scolecida (mean = 1.89 ± 0.70 pg; n = 7) ($F(2,31) = 0.666$, $p = 0.521$).

DISCUSSION

Polychaete Genome Size Overview

New genome size estimates indicate the genome size range within polychaetes may be much larger than previously thought; three unidentified maldanid worms from the Antarctic region demonstrated C-value estimates over 3 pg more than *Nephtys incisa*, which last held the record for largest polychaete genome size (Table 3.1). The range, from 0.06 pg (*Dinophilus gyrociliatus*) to 10.29 pg (Unknown #6) is greater than a 170-fold difference, and is much less constrained than other annelid groups (Gregory 2012; Chapter 2). This variability was generally distributed across the entire class; that is, particularly large or small genomes were not restricted to specific lower-level taxa (Table 3.1; see also Gambi et al. 1997). For example, the Maldanidae have a C-value range from 0.46 pg to 10.29 pg, and the Spionidae have a range from 0.50 pg to 7.23 pg.

Genome Size Variation

The large range in nuclear DNA contents among polychaetes raises questions regarding the mechanisms that have generated this much diversity. In many groups, transposable elements are thought to be responsible for much of the variability in genome size among species, but it is also worth considering the role of large-scale duplications in polychaetes, as polyploidy is known

in other annelids (Gregory and Mable 2005). Some patterns of discontinuous genome size variation have been noted in other invertebrates, though this may not be associated with differences in chromosome number – for this reason, such patterns have been dubbed "cryptopolyploidy" (e.g., Gregory et al. 2000). There is some evidence of discontinuous variation within the current polychaete dataset. For example, the genome size of *Alitta succinea* is approximately twice the size of *A. virens* (Table 3.1). The Orbiniid genus *Leitoscoloplos* also exhibits discontinuous variation, with the genome size estimates increasing by multiples of approximately 0.6 pg in all three species. The average ratio for congeneric species in the present study was 2.01, suggesting that genome duplication is a potential source of variability within genera. This value is largely driven by the genus Tharyx, however, as it exhibits an 8.5-fold variation among the two species with genome size estimates. The frequency distribution of average congeneric species ratio values for genera is heavily skewed left, indicating that only speculation can be made at this point whether whole genome duplication is a significant factor in other polychaete genera and species (Figure 3.2). However, to determine whether the observed quantum variation in Alitta and Leitoscoloplos is explained by polyploidy or cryptopolyploidy, additional data from karyotypic analyses as well as sequence comparisons from small and large genomes is required, as has been done in some plants and amphibians (Blanc et al. 2000; Beçak and Kobashi 2004).

Body Size and Genome Size

Previous polychaete genome size analyses indicated that interstitial species had a restricted genome size range of less than 1 pg, while macrobenthic species had more variable, larger genome sizes (Soldi et al. 1994; Gambi et al. 1997). New estimates present reason to believe otherwise, however, as species with C-values less than 1 pg such as *Glycera tridactyla*,

Leitoscoloplos fragilis and *Alitta virens* are all macrobenthic. Conversely, the tiny (< 0.5mm) interstitial, cold-temperate *Paramphinome jeffreysii*, has a larger genome size of 2.12 pg (Schuckel et al. 2010). In other words, there is evidence to suggest interstitial species are not limited to genome sizes less than 1 pg, and perhaps body size does not present any sort of constraint on genome size, or vice-versa.

Geography

The world-wide range of polychaetes throughout the oceans leaves open the possibility of correlations between specific geographic range and genome size. A comparison between genome size estimates and two latitudinally contrasting regions from which the specimens were collected yielded a significant result between Polar polychaetes and those found in the Temperate coastal regions of North America (Table 3.1). Polychaetes collected from the polar waters surrounding Norway and Antarctica had significantly larger genome size estimates than temperate polychaetes; in fact, the C-value range for Polar polychaetes was double that of the others. Amphipods are also known to exhibit significantly larger genome sizes in polar regions than their lower-latitude counterparts (Rees et al. 2007). Depth of species habitation is unknown whether to affect genome size without complete collection data and appropriate depth profiles giving salinity, temperature, light intensity, oxygen, and other dissolved elements at the site of collection. Without further taxonomic identification, and morphological and life history characteristics of the four unidentified polar species, the cause for the increase in genome size is yet to be determined. Given the larger genome sizes in the polar polychaetes, we would expect to see a relatively decreased rate of development and larger body size among these species. Further research, including a larger sample size and accompanying morphological, habitat and life history data, would benefit the pursuit of explaining this polar trend.

Extremophiles

A number of polychaete species and families have long been known to habituate extreme environments, such as hydrothermal vents or cold seeps (Grassle 1985). Many species within Alvinellidae, such as *Alvinella pompejana*, *A. caudata*, *Paralvinella palmiformis*, and *P. grasslei* are among the most thermotolerant metazoans known (Chevaldonné et al. 1992; Cary et al. 1998). The extremophile species represented in the present study do not present a large enough dataset for formal analyses, however preliminary observations do not suggest this highly adapted lifestyle is associated with genome size. The range of C-values for the four aforementioned species is limited to 0.96 – 1.23 pg, and as the Alvinellidae are endemic to hydrothermic vents it cannot be said whether this is a family trait or adaptation for thermotolerance. Other non-extremophile families, such as Siboglinidae, also present restricted ranges of genome sizes around 1 pg, so thermotolerance cannot necessarily be attributed to this narrow range. However, *Lamellibrachia luymesi*, a cold seep-habituating tubeworm with a remarkable lifespan of more than 250 years, also has a relatively small genome size of 1.37 pg. With this, it is possible to suggest that extremophile polychaetes are restricted to genome sizes smaller than 2 pg; however the reason for such a small genome size is unclear, as very few species are able to colonize or predate in such a habitat, so there would be less pressure for fast development.

CONCLUSIONS

Much is still unknown about the life history for countless polychaete species, information which would contribute enormously to understanding the widely ranging C-values within this group. New evidence disputes the previous claim that segregated meio- and macrobenthic species by genome size, instead suggesting that the harsh interstitial environment does not

impose as great a restriction on genome size due to fast developmental requirements and small body size. Polychaetes found in polar regions are more likely to have larger C-values than those found in the waters along the temperate coasts of North America. Extremophile polychaetes may be restricted to genome sizes less than 2 pg. Future polychaete studies should explore potential genome size relationships with egg size, reproductive method, habitat type and depth, longevity, and sessile versus motile lifestyles.

Table 3.1. Haploid genome size estimates for polychaete species, including collection location data for new estimates. Genome size estimates were measured using Feulgen Image Analysis Densitometry with at least 50 nuclei measured per specimen, compared to a blood standard from *Gallus gallus domesticus* (1C = 1.25 pg) or *Oncorhynchus mykiss* (1C = 2.5 pg).
Bolded C-values denote data from the Animal Genome Size Database.

Subclass	Family	Genus	GS (pg)	Collection Region
Aciculata	Amphinomidae	*Linopherus ambigua*	**2.40**	-
Aciculata	Amphinomidae	*Paramphinome jeffreysii*	2.12	Polar
Aciculata	Aphroditidae	*Aphrodita aculeata*	**0.62**	-
Aciculata	Aphroditidae	*Laetmonice producta*	1.01	Polar
Aciculata	Aphroditidae	*Laetmonice sp.* (*Unknown #1*)	1.51	Polar
Aciculata	Chrysopetalidae	*Paleanotus debile*	**1.09**	-
Aciculata	Dorvilleidae	*Ophrytrocha costlowi*	**0.40**	-
Aciculata	Dorvilleidae	*Ophrytrocha diadema*	**0.39**	-
Aciculata	Dorvilleidae	*Ophrytrocha gracilis*	**0.36**	-
Aciculata	Dorvilleidae	*Ophrytrocha hartmanni*	**1.04**	-
Aciculata	Dorvilleidae	*Ophrytrocha labronica*	**0.39**	-
Aciculata	Dorvilleidae	*Ophrytrocha labronica pacifica*	**0.36**	-
Aciculata	Dorvilleidae	*Ophrytrocha macrovifera*	**0.72**	-
Aciculata	Dorvilleidae	*Ophrytrocha notoglandulata*	**0.32**	-
Aciculata	Dorvilleidae	*Ophrytrocha puerilis*	**0.40**	-
Aciculata	Dorvilleidae	*Ophrytrocha robusta*	**0.34**	-
Aciculata	Dorvilleidae	*Protodorvillea kefersteini*	**0.32**	-
Aciculata	Dorvilleidae	*Schistomeringos neglecta*	**0.39**	-
Aciculata	Eunicidae	*Eunice norvegica*	4.08	Polar
Aciculata	Glyceridae	*Glycera americana*	2.97	Temperate
Aciculata	Glyceridae	*Glycera tridactyla*	0.41	Temperate
Aciculata	Glyceridae	*Glycera lapidum*	**1.33**	-
Aciculata	Goniadidae	*Goniada brunnea*	4.18	Temperate
Aciculata	Hesionidae	*Hesiolyra bergi*	**1.31**	-
Aciculata	Hesionidae	*Hesiospina sp.*	**0.48**	-
Aciculata	Hesionidae	*Kefersteinia sp.*	**0.20**	-
Aciculata	Hesionidae	*Ophiodromus flexuosus*	**0.31**	-
Aciculata	Hesionidae	*Ophiodromus obscurus*	**1.60**	-
Aciculata	Hesionidae	*Ophiodromus pugettensis*	2.04	Temperate
Aciculata	Lumbrineridae	*Ninoe nigripes*	5.33	Temperate
Aciculata	Lumbrineridae	*Scoletoma tenuis*	**2.40**	-
Aciculata	Nephtyidae	*Nephtys brachycephala*	5.10	Temperate
Aciculata	Nephtyidae	*Nephtys incisa*	**7.20**	-
Aciculata	Nephtyidae	*Nephtys sp.*	**2.20**	-
Aciculata	Nereididae	*Alitta succinea*	1.53	Temperate
Aciculata	Nereididae	*Alitta virens*	0.67	Temperate
Aciculata	Nereididae	*Laeonereis culveri*	**0.80**	-
Aciculata	Nereididae	*Neanthes arenaceodentata*	2.02	Temperate
Aciculata	Nereididae	*Neanthes caudata*	**2.51**	-
Aciculata	Nereididae	*Nereis neoneanthes*	0.67	Temperate
Aciculata	Nereididae	*Platynereis dumerilii*	**0.89**	-
Aciculata	Nereididae	*Platynereis megalops*	1.76	Temperate

Aciculata	Onuphidae	*Americonuphis magna*	**1.20**	-
Aciculata	Onuphidae	*Diopatra cuprea cuprea*	**2.00**	-
Aciculata	Onuphidae	*Nothria conchylega*	0.92	Temperate
Aciculata	Onuphidae	*Onuphis elegans*	2.14	Temperate
Aciculata	Onuphidae	*Onuphis eremita oculata*	**1.70**	-
Aciculata	Onuphidae	*Onuphis quadricuspis*	1.28	Temperate
Aciculata	Onuphidae	*Onuphis sp.*	**1.70**	-
Aciculata	Phyllodocidae	*Nereiphylla paretti*	**2.70**	-
Aciculata	Polynoidae	*Branchinotogluma trifurcus*	**1.83**	-
Aciculata	Polynoidae	*Branchinotogluma tunnicliffae*	1.62	Temperate
Aciculata	Polynoidae	*Branchipolynoe pettiboneae*	**2.56**	-
Aciculata	Polynoidae	*Branchipolynoe seepensis*	**3.10**	-
Aciculata	Polynoidae	*Branchipolynoe sp.*	**1.00**	-
Aciculata	Polynoidae	*Gattyana cirrhosa*	1.39	Temperate
Aciculata	Polynoidae	*Halosydna brevisetosa*	1.93	Temperate
Aciculata	Polynoidae	*Lepidonotopodium jouinae*	**4.16**	-
Aciculata	Polynoidae	*Lepidonotus squamatus*	**1.50**	-
Aciculata	Polynoidae	*Lepidonotus sublevis*	**2.20**	-
Aciculata	Polynoidae	*Polyeunoa laevis*	2.01	Polar
Aciculata	Polynoidae	*Unknown #2*	2.24	Polar
Aciculata	Syllidae	*Erinaceusyllis erinaceus*	**0.73**	-
Aciculata	Syllidae	*Exogone dispar*	1.70	Temperate
Aciculata	Syllidae	*Odontosyllis fulgurans*	**0.50**	-
Aciculata	Syllidae	*Parapionosyllis elegans*	**0.11**	-
Aciculata	Syllidae	*Plakosyllis brevipes*	**0.26**	-
Aciculata	Syllidae	*Prosphaerosyllis campoyi*	**0.48**	-
Aciculata	Syllidae	*Salvatoria limbata*	**0.39**	-
Aciculata	Syllidae	*Sphaerosyllis pirifera*	**0.11**	-
Aciculata	Syllidae	*Syllis prolifera*	**0.40**	-
Canalipalpata	Acrocirridae	*Macrochaeta clavicornis*	**0.58**	-
Canalipalpata	Alvinellidae	*Alvinella caudata*	**0.97**	-
Canalipalpata	Alvinellidae	*Alvinella pompejana*	**0.96**	-
Canalipalpata	Alvinellidae	*Paralvinella grasslei*	**1.23**	-
Canalipalpata	Alvinellidae	*Paralvinella palmiformis*	1.01	Temperate
Canalipalpata	Ampharetidae	*Melinna macualta*	2.96	Temperate
Canalipalpata	Chaetopteridae	*Chaetopterus variopedatus*	1.00	Temperate
Canalipalpata	Chaetopteridae	*Mesochaetopterus taylori*	1.31	Temperate
Canalipalpata	Chaetopteridae	*Phyllochaetopterus prolifica*	2.81	Temperate
Canalipalpata	Chaetopteridae	*Spiochaetopterus costarum oculatus*	1.06	Temperate
Canalipalpata	Cirratulidae	*Cirratulus grandis*	**0.70**	-
Canalipalpata	Cirratulidae	*Cirratulus spectabilis*	0.82	Temperate
Canalipalpata	Cirratulidae	*Cirriformia filigera*	**1.00**	-
Canalipalpata	Cirratulidae	*Cirriformia luxuriosa*	**3.40**	-

Canalipalpata	Cirratulidae	*Tharyx acutus*	0.85	Temperate
Canalipalpata	Cirratulidae	*Tharyx multibranchiis*	**0.10**	-
Canalipalpata	Pectinariidae	*Pectinaria gouldii*	1.28	Temperate
Canalipalpata	Sabellariidae	*Neosabellaria cementarium*	1.56	Temperate
Canalipalpata	Sabellidae	*Amphiglena mediterranea*	**0.35**	-
Canalipalpata	Sabellidae	*Branchiomma luctuosum*	**1.07**	-
Canalipalpata	Sabellidae	*Branchiomma nigromaculatum*	**1.30**	-
Canalipalpata	Sabellidae	*Chone aurantiaca*	2.39	Temperate
Canalipalpata	Sabellidae	*Dodecaceria fewkesi*	1.73	Temperate
Canalipalpata	Sabellidae	*Myxicola infundibulum*	**3.10**	-
Canalipalpata	Sabellidae	*Sabella spallanzanii*	**0.58**	-
Canalipalpata	Sabellidae	*Schizobranchia insignis*	1.69	Temperate
Canalipalpata	Serpulidae	*Cricigera zygophora*	1.53	Temperate
Canalipalpata	Serpulidae	*Hydroides elegans*	**0.75**	-
Canalipalpata	Serpulidae	*Pomatoceros lamarcki*	**1.23**	-
Canalipalpata	Serpulidae	*Serpula vermicularis*	1.06	Temperate
Canalipalpata	Siboglinidae	*Lamellibrachia luymesi*	1.37	Temperate
Canalipalpata	Siboglinidae	*Oasisia alvinae*	**0.97**	-
Canalipalpata	Siboglinidae	*Riftia pachyptila*	**0.79**	-
Canalipalpata	Siboglinidae	*Seepiophila jonesi*	1.47	Temperate
Canalipalpata	Siboglinidae	*Tevnia jerichonana*	**1.06**	-
Canalipalpata	Spionidae	*Laonice weddellia*	7.23	Polar
Canalipalpata	Spionidae	*Prionospio malmgreni*	**0.50**	-
Canalipalpata	Spionidae	*Scolecolepides viridis*	3.39	Temperate
Canalipalpata	Sternaspidae	*Sternaspis scutata*	2.40	Temperate
Canalipalpata	Terebellidae	*Amphitrite ornata*	1.39	Temperate
Canalipalpata	Terebellidae	*Thelepus crispus*	1.22	Temperate
Canalipalpata	Terebellidae	*Unknown #3*	0.65	Polar
Polychaeta incertae sedis	Dinophilidae	*Dinophilus gyrociliatus*	**0.06**	-
Polychaeta incertae sedis	Nerillidae	*Mesonerilla intermedia*	**0.33**	-
Polychaeta incertae sedis	Polygordiidae	*Polygordius appendiculatus*	**0.68**	-
Polychaeta incertae sedis	Protodrilidae	*Protodrilus sp.*	**0.24**	-
Polychaeta incertae sedis	Saccocirridae	*Saccocirrus papillocerus*	**0.48**	-
Scolecida	Arenicolidae	*Abarenicola claparedi*	0.98	Temperate
Scolecida	Arenicolidae	*Abarenicola pacifica*	1.32	Temperate
Scolecida	Arenicolidae	*Arenicola cristata*	**0.90**	-
Scolecida	Capitellidae	*Capitella capitata*	**0.24**	-
Scolecida	Capitellidae	*Heteromastus filiformis*	0.96	Temperate
Scolecida	Capitellidae	*Notomastus latericeus*	**1.20**	-
Scolecida	Capitellidae	*Notomastus tenuis*	1.01	Temperate
Scolecida	Maldanidae	*Axiothella rubrocincta*	2.76	Temperate
Scolecida	Maldanidae	*Clymenella mucosa*	**2.70**	-
Scolecida	Maldanidae	*Clymenella torquata*	2.01	Temperate

Scolecida	Maldanidae	*Euclymene collaris*	**0.46**	-
Scolecida	Maldanidae	*Unknown #4*	9.29	Polar
Scolecida	Maldanidae	*Unknown #5*	9.80	Polar
Scolecida	Maldanidae	*Unknown #6*	10.29	Polar
Scolecida	Opheliidae	*Polyophthalmus pictus*	**0.17**	-
Scolecida	Orbiniidae	*Leitoscoloplos fragilis*	0.62	Temperate
Scolecida	Orbiniidae	*Leitoscoloplos pugettensis*	1.18	Temperate
Scolecida	Orbiniidae	*Leitoscoloplos robustus*	2.26	Temperate
Scolecida	Orbiniidae	*Orbinia bioreti*	**0.53**	-
Scolecida	Orbiniidae	*Scoloplos rubra*	**3.10**	-
Scolecida	Paraonidae	*Aricidea cerruti*	**0.62**	-
Scolecida	Paraonidae	*Aricidea fragilis*	**4.60**	-
Scolecida	Paraonidae	*Aricidea quadrilobata*	1.28	Temperate
Scolecida	Scalibregmatidae	*Scalibregma inflatum*	1.57	Temperate

Figure 3.1. The relationship between average congeneric species ratio and average genome size per genus for 22 polychaete genera.

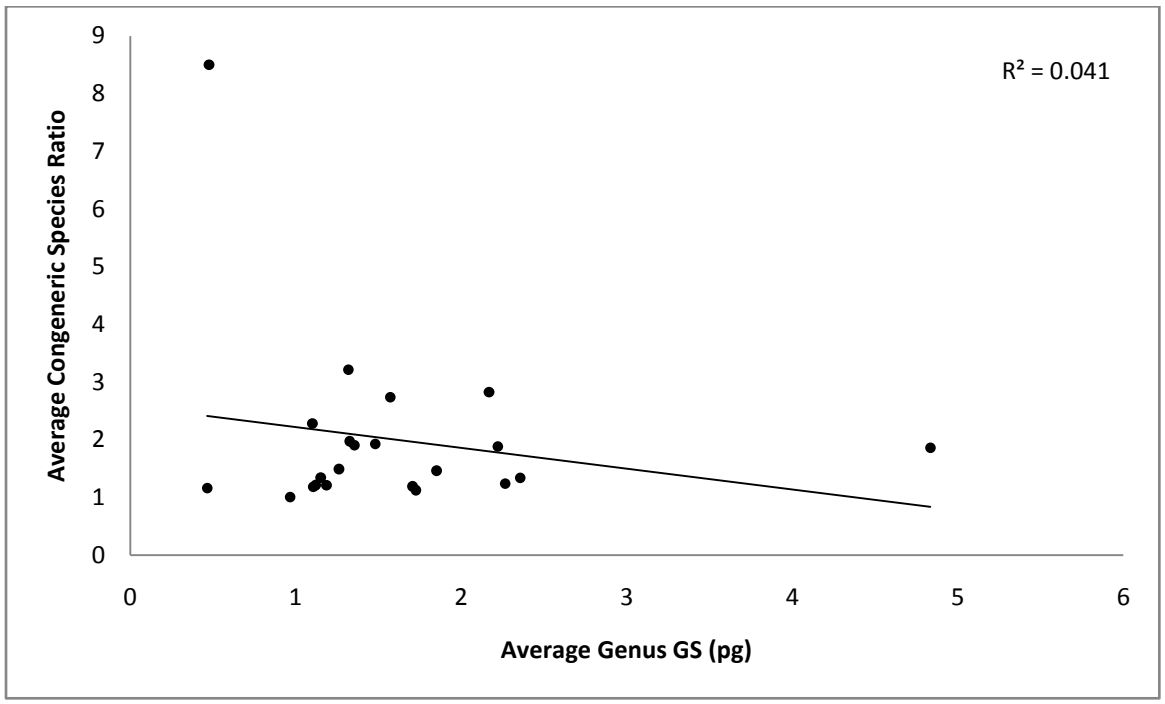
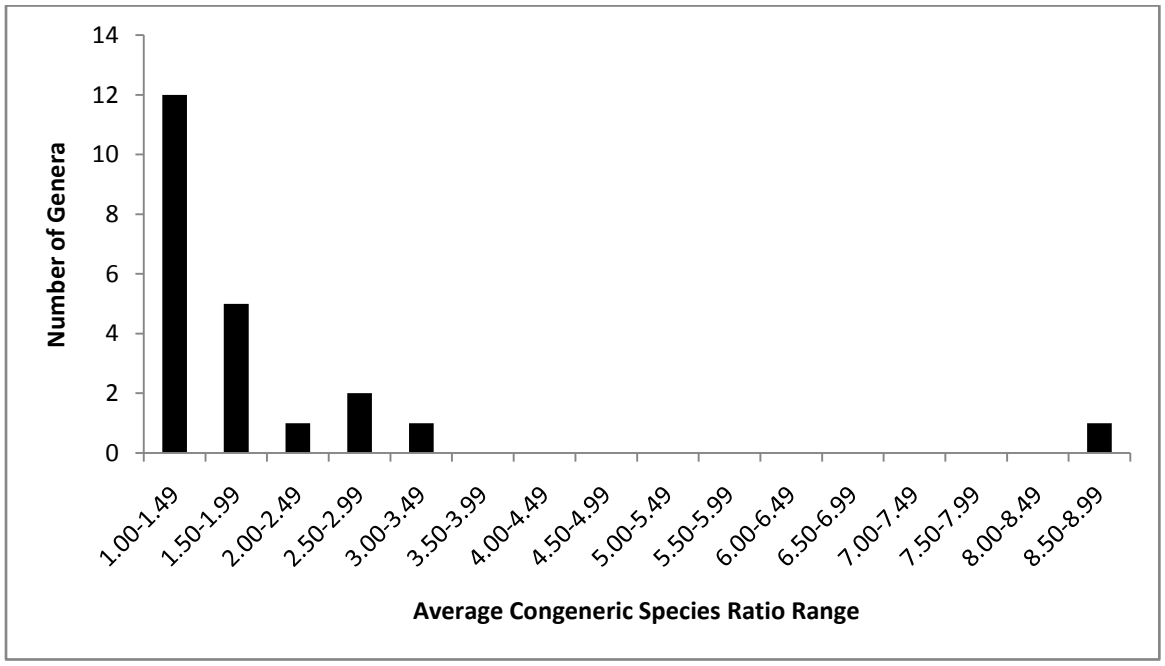


Figure 3.2. Frequency distribution of average congeneric species ratio for 22 polychaete genera.



CHAPTER FOUR

Discussion and Conclusions

Annelid Genome Sizes

The combined number of new genome size estimates presented in chapters 2 and 3 represents more than a 60% increase in annelid species represented in the Animal Genome Size Database (AGSD) (Gregory 2012). With the increasing popularity and capability to sequence entire genomes, it is important to build a library of annelid genome sizes to be available for planning future studies. The longevity, regenerative capabilities and thermal or toxic tolerance of many annelids make them prime candidates for sequencing to understand the genetic framework behind these characteristics (e.g. Nyberg et al. 2012).

Of all the annelids, the leeches present the most constrained genome size range, varying only five-fold from 0.24 – 1.23 pg with an average of 0.62 pg. While the polychaetes presented a much larger range of genome sizes (0.06 – 10.28 pg), the majority of C-values occur within a range below 2.00 pg, similar to the leeches (Figure 4.1). Oligochaetes, the sister clade to leeches within Clitellata, range more widely from 0.43 – 7.64 pg; within the oligochaetes, however, the lumbricid earthworms share a similar constraint to leeches with a genome size range of 0.43 – 1.24 pg, while naidid worm genome sizes range from 0.78 – 7.64 pg (Gregory and Hebert 2002; Gregory 2012). The disparity unlikely lies in habitat differences, as nadid worms are aquatic, like leeches. This matter, and many other questions regarding the subject cannot yet be addressed as still only 28 oligochaete species have genome size estimates (Gregory 2012). It is evident, however, that a threshold may exist around the 2.00 pg level that fewer species exceed than remain below. Future genomic studies should seek to expand the number of annelid species, particularly the clitellates, represented in the AGSD to explore emerging trends.

The ongoing assemblage of a new annelid phylogeny through the efforts involved in WormNet II will make it easier to infer larger scale phylogenetic relationships with genome size in future studies, and therefore better understand annelid genome evolution.

Questions and Predictions Revisited for Chapter 2

1. *What is the genome size range across the subclass Hirudinea?*

Leech genome size estimates fell well within the previously estimated annelid range; in fact they were quite constrained, with only a five-fold difference between 0.24 – 1.23 pg.

2. *Is there a significant difference in genome size between the two leech orders, Arhynchobdellida and Rhynchobdellida?*

The Rhynchobdellida exhibited significantly larger C-values than Arhynchobdellida; parental care within the Rhynchobdellida may have eased the pressure for rapid development, thereby allowing for an increase in genome size. The cause of an observed decrease in genome size between substrate-brooding species and ventrum-brooding species within the Rhynchobdellida requires further investigation.

3. *Do leech C-values correlate with body size, and if so, does body size pose any kind of constraint to genome size?*

An inverse relationship between genome size and body size indicates that within leeches, body size is not associated with cell volume. Growth rate may instead be responsible for body size within this group, if both segment number and growth time are constant.

4. *Do leeches with differing diets have significantly different genome sizes?*

There was no difference observed between leeches with macrophagous and sanguivorous diets.

5. Is there a relationship between genome size and geographic region?

Only one leech species, *Nephelopsis obscura*, exhibited a latitudinal gradient in genome size; however this could reflect either developmental pressures at higher latitudes, or the possibility that *N. obscura* is a cryptic species complex.

Questions and Predictions Revisited for Chapter 3

1. Do polychaete genome sizes still fall within the same range from previous studies?

Polychaete genome sizes were significantly larger than data reported in previous studies. The C-value range now extends from 0.06 – 10.29 pg, an approximately 170-fold difference.

2. Is there evidence of genome duplication within any polychaete groups?

The average congeneric species ratio for polychaetes was 2.01, and discontinuous variation observed in two genera, indicating that there is evidence of polyploidy and/or cryptopolyploidy within the polychaetes.

3. Is there a geographic difference in genome size for polychaetes?

Polychaetes found in Polar regions had significantly larger genomes than temperate-region polychaetes.

4. Does genome size differ between polychaete subclasses Aciculata, Canalipalpata and Scolecida?

Genome size did not differ between Aciculata, Canalipalpata and Scolecida.

Notes on Methodology

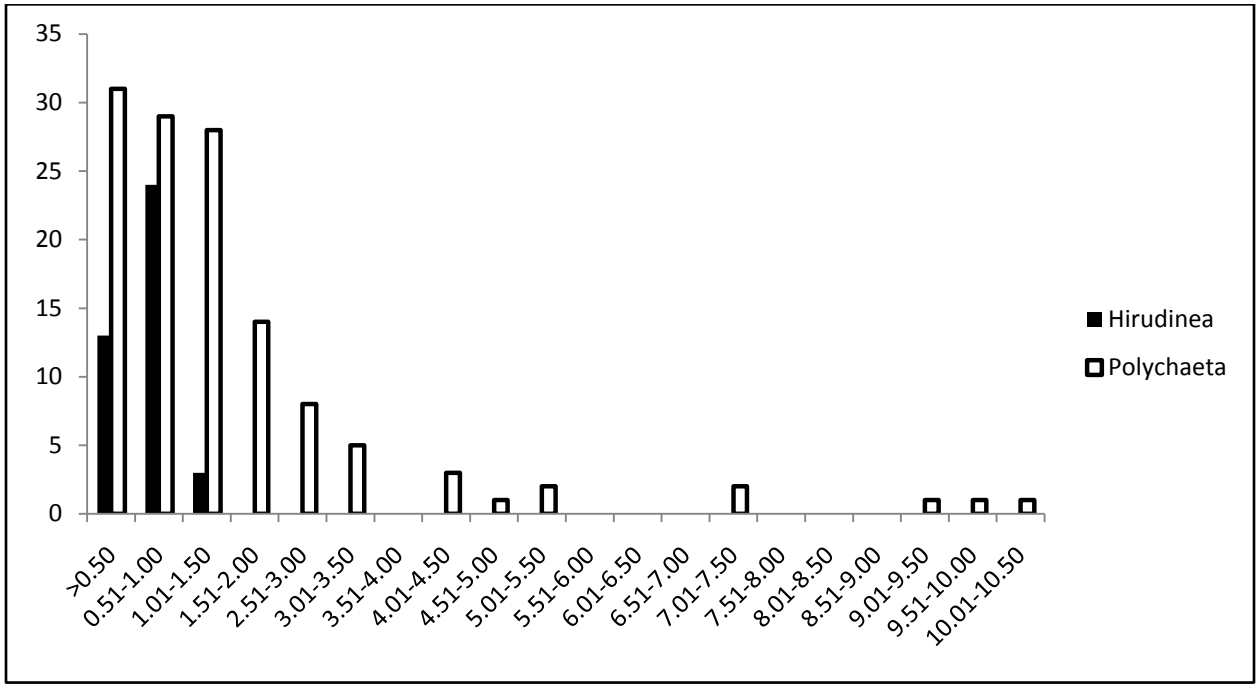
Approximately 80% of leech slides prepared for FIAD yielded useful genome size data, with 50 or more stained nuclei and a coefficient of variation of ≤ 5%. Compared to fresh specimens, slides prepared from frozen leech specimens were slightly less successful at producing data due to the frequency of lysed nuclei, which are too diffuse to measure.

It is important to note the success rate of acquiring data from frozen polychaete tissue, as it is often impractical to prepare slides prior to preservation at the site of collection in remote marine areas. Specimens preserved in ethanol are unusable due to the dehydrated state of the tissue. All polychaete slides in the present study were prepared from frozen material and had a 66% success rate at producing useful data. As tissue from the head region was used (when possible), the two most common issues faced were debris obscuring stained nuclei and nuclei aggregated too close together for measurement. Debris typically consisted of chaetae/bristles and sediment which may have been stuck to the outside of the specimen. Tissue from areas other than the head (or possibly posterior end) proved to be too thin, covered in too much chaetae and often carried with it sediment and other debris from gut contents.

Concluding Remarks

This was the first study to quantify genome sizes within leeches and seek relationships with life history traits and geographic distribution. It also raises new questions about previous conceptions of the genome size relationship with body size in both polychaetes and invertebrates as a whole. The results lay the groundwork for future annelid studies involving evolutionary genomics, construction and analyses of phylogenetic trees (through the use of DNA barcoding) and sequencing.

Figure 4.1. Frequency distribution of genome size observed between the Hirudinea (0.24 – 1.23 pg) and the Polychaeta (0.06 – 10.28 pg).



CHAPTER FIVE

Literature Cited

Adams, S.L. 1988. The Medicinal Leech: A Page from the Annals of Internal Medicine. *Annals of Internal Medicine* **109**: 399-405.

Bartolomaeus, T., Purschke, G., and Hausen, H. 2005. Polychaete Phylogeny Based on Morphological Data – A Comparison of Current Attempts. *Hydrobiologia* **535/536**: 341-356.

Beçak, M.L., and Kobashi, L.S. 2004. Evolution by polyploidy and gene regulation in Anura. *Genetics and Molecular Research* **3**: 195-212.

Bely, A.E., and Weisblat, D.A. 2006. Lessons from leeches: a call for DNA barcoding in the lab. *Evolution and Development* **8**: 419-501.

Bennett, M.D., Smith, J.B., and Smith, R.I.L. 1982. DNA amounts of angiosperms from the Antarctic and South Georgia. *Environmental and Experimental Botany* **22**: 307-318.

Blanc, G., Barakat, A., Guyot, R., Cooke, R., and Delseny, M. 2000. Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell* **12**: 1093-1101.

Bleidorn, C. 2007. The Role of Character Loss in Phylogenetic Reconstruction as Exemplified for the Annelida. *Journal of Zoological Systematics and Evolutionary Research* **45**: 299-307.

Bon, C., Saliou, B., Thieffry, M., and Manaranche, R. 1985. Partial purification of α-glycerotoxin, a presynaptic neurotoxin from the venom glands of the polychaete annelid *Glycera convoluta. Neurochemistry International* **7**: 63-75.

Borda, E., and Siddall, M.E. 2004. Review of the evolution of life history strategies and phylogeny of the Hirudinea (Annelida: Oligochaeta). *Lauterbornia* **52**: 5-25.

Carr, C.M., Hardy, S.M., Brown, T.M., Macdonald, T.A., and Hebert, P.D.N. 2011. A tri-oceanic perspective: DNA barcoding reveals geographic structure and cryptic diversity in Canadian polychaetes. *PLoS ONE* **6**: e22232.

Cary, C.S., Shank, T., and Stein, J.L. 1998. Worms bask in extreme temperatures. *Nature* **391**: 545-546.

Cavalier-Smith, T. 1985. Cell volume and the evolution of eukaryotic genome size. *In* The Evolution of Genome Size (ed. T. Cavalier-Smith). Wiley, Chichester, UK. pp. 104-184.

Chevaldonné, P., Desbruyères, D., and Childress, J.J. 1992. Some like it hot... and some even hotter. *Nature* **359**: 593-594.

Cichocka, J., and Bielecki, A. 2008. Biological diversity of leeches (Clitellata: Hirudinida) based on characteristics of the karyotype. *Wiadomości Parazytologiczne* **54**: 309-314.

Conner, W.G., Hinegardner, R., and Bachmann, K. 1972. Nuclear DNA Amounts in Polychaete Annelids. *Experientia* **28**: 1502-1504.

Crow, K.D., and Wagner, G.P. 2006. What is the role of genome duplication in the evolution of complexity and diversity? *Molecular Biology and Evolution* **23**: 887-892.

Desser, S.S. 2001. The blood parasites of anurans from Costa Rica with reflections on the taxonomy of their trypanosomes. *Journal of Parisitology* **87**: 152-160.

Díaz-Castañeda, V., and Reish, D.J. 2009. "Polychaetes in Environmental Studies". *In* Annelids in Modern Biology (ed. D.H. Shain). John Wiley and Sons, Ltd., Hoboken, NJ. pp. 217-219.

Doležel, J., Bartoš, J., Voglmayr, H., and Greilhuber, J. 2003. Letter to the Editor: Nuclear DNA Content and Genome Size of Trout and Human. *Cytometry Part A* **51A**: 127-128.

Ferrari, J.A., and Rai, K.S. 1989. Phenotypic correlates of genome size variation in *Aedes albopictus*. *Evolution* **43**: 895-899.

Finston, T.L., Hebert, P.D.N., and Foottit, R. 1995. Genome size variation in aphids. *Insect Biochemistry and Molecular Biology* **25**: 189-196.

Gambi, M.C., Ramella, L., Sella, G., Protto, P., and Aldieri, E. 1997. Variation in Genome Size in Benthic Polychaetes: Systematic and Ecological Relationships. *Journal of the Marine Biological Association of the United Kingdom* **77**: 1045-1057.

Gariepy, T.D., Lindsay, R., Ogden, N., and Gregory, T. R. 2012. Identifying the last supper: utility of the DNA barcode library for bloodmeal identification in ticks (Ixodida: Ixodidae). *Molecular Ecology Resources* **12**: 646-652.

Ghiselin, M.T. 1969. The Evolution of Hermaphroditism Among Animals. *Quarterly Review of Biology* **44**: 189-208.

Govedich, F.R., Moser, W.E., Davies, R.W. 2010. "Annelida (Clitellata): Oligochaeta, Branchiobdellida, Hirudinida, and Acanthobdellida". *In* Ecology and Classification of North American Freshwater Invertebrates (eds. J.H. Thorp and A.P. Covich). Elsevier, London, UK. pp. 429.

Govedich, F.R., Negm-Eldin, M.M., and Bain, B.A. 2004. Prey preference and feeding behaviour of two Australian glossiphoniid leeches, and the avoidance response of their snail prey. *Victorian Naturalist* **121**: 51-55.

Grassle, J.F. 1985. “The Ecology of Deep-Sea Hydrothermal Vent Communities”. *In* Advances in Marine Biology Volume 23 (ed. J. H. S. Baxter). Academic Press Inc., London, UK. pp. 326-328.

Gregory, T.R. 2000. Nucleotypic effects without nuclei: genome size and erythrocyte size in mammals. *Genome* **43**: 895-901.

Gregory, T.R. 2005. “Genome Size Evolution in Animals”. *In* The Evolution of the Genome (ed. T.R. Gregory). Elsevier Academic Press, San Diego, CA. pp. 41-60.

Gregory, T.R. 2012. Animal Genome Size Database. http://www.genomesize.com.

Gregory, T.R., and Hebert, P.D.N. 2002. Genome Size Estimates for Some Oligochaete Annelids. *Canadian Journal of Zoology* **80**: 1485–1489.

Gregory, T.R., Hebert, P.D.N., and Kolasa, J. 2000. Evolutionary implications of the relationship between genome size and body size in flatworms and copepods. *Heredity* **84**: 201-208.

Gregory, T.R., and Mable, B.K. 2005. “Polyploidy in Animals”. *In* The Evolution of the Genome (ed. T.R. Gregory). Elsevier Academic Press, San Diego, CA. pp 427-517.

Greilhuber, J. 1981. Intraspecific variation in genome size: a critical reassessment. *Annals of Botany* **82**: 27-35.

Hackmann, J.G., Dütsch, M., Neundörfer, B., Dütsch, F., and Hartung, U. 2005. Leech therapy in the treatment of median nerve compression due to forearm haematoma. *Journal of Neural Neurosurgery and Psychiatry* **76**: 1465-1468.

Hardie, D.C., Gregory, T.R., and Hebert, P.D.N. 2002. From Pixels to Picograms: A Beginners' Guide to Genome Quantification by Feulgen Image Analysis Densitometry. *Journal of Histochemistry and Cytochemistry* **50**: 735–749.

Hebert, P.D.N., Cywinska, A., Ball, S.L., and deWaard, J.R. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B* **270**: 313-321.

Hickman, Jr., C.P., Roberts, L.S., Larson, A., I'Anson, H., and Eisenhour, D.J. 2006. "Segmented Worms". *In* Integrated Principles of Zoology. McGraw-Hill, New York, NY. pp. 373-374.

Hill, S.D., and Savage, R.M. 2009. "Evolution, Development and Ecology of *Capitella* sp. I: A Waxing Model for Polychaete Studies". *In* Annelids in Modern Biology (ed. D.H. Shain). John Wiley and Sons, Ltd., Hoboken, NJ. pp. 88-111.

Hinegardner, R. 1974. Cellular DNA content of the Mollusca. *Comparative Biochemistry and Physiology* **47A**: 447-460.

Hinegardner, R. 1976. "Evolution of genome size". *In* Molecular Evolution (ed. F. Ayala). Sinauer Association Inc., Sunderland, MA. pp. 179-199.

Hovingh, P., and Klemm, D.J. 2000. *Erpobdella lahontana* (Annelida: Hirudinea: Arhynchobdellida: Erpobdellidae), a new species of freshwater leech from North America. *Proceedings of the Biological Society of Washington* **3**: 155-161.

Klemm, D.J. 1985. "Freshwater Leeches (Annelida: Hirudinea)". *In* A Guide to the Freshwater Annelida (Polychaeta, Naidid and Tubificid Oligochaeta, and Hirudinea) of North America (ed. D.J. Klemm). Kendall/Hunt Publishing Company, Dubuque, IA. Pp. 70-165.

Kuo, D.-H. 2009. "Annelids in Evolutionary Developmental Biology". *In* Annelids in Modern Biology (ed. D.H. Shain). John Wiley and Sons, Ltd., Hoboken, NJ. pp. 65-66.

Lai, Y.-T., Chang, C.-H., and Chen, J.-H. 2009. Two new species of *Helobdella* Blanchard a896 (Hirudinida: Rhynchobdellida: Glossiphoniidae) from Taiwan, with a checklist of hirudinea fauna of the island. *Zootaxa* **2068**: 27-46.

Lang, S.A., and Shain, D.H. 2009. "Stem Cell Genesis and Differentiation in Leech". *In* Annelids in Modern Biology (ed. D.H. Shain). John Wiley and Sons, Ltd., Hoboken, NJ. pp. 116-117.

McHugh, D. 2000. Molecular phylogeny of the Annelida. *Canadian Journal of Zoology* **78**: 1873-1884.

Nyberg, K.G., Conte, M.A., Kostyun, J.L., Forde, A. and Bely, A.E. 2012. Transcriptome characterization via 454 pyrosequencing of the annelid *Pristina leidyi*, an emerging model for studying the evolution of regeneration. *BMC Genomics* **13**: 287.

Oceguera-Figueroa, A., León-Règagnon, V., and Siddall, M.E. 2010. DNA barcoding reveals Mexican diversity within the freshwater leech genus *Helobdella* (Annelida: Glossiphoniidae). *Mitochondrial DNA* **21**: 24-29.

Pfeifer, K., Dorresteijn, A.W., and Fröbius, A.C. 2012. Activation of Hox genes during caudal regeneration of the polychaete annelid *Platynereis dumerilii*. *Development Genes and Evolution* **222**: 165-179.

Phipps, G.L., Ankley, G.T., Benoit, D.A., and Mattson, V.R. 1992. Use of the aquatic oligochaete *Lumbriculus variegates* for assessing the toxicity and bioaccumulation of sediment-associated contaminants. *Environmental Toxicology and Chemistry* **12**: 269-279.

Pilcher, H. 2004. Stuck on you. *Nature* **432**: 10-11.

Pradillon, F., and Gaill, F. 2009. "Hydrothermal Vent Annelids". *In* Annelids in Modern Biology (ed. D.H. Shain). John Wiley and Sons, Ltd., Hoboken, NJ. pp. 287-297.

Purshke, G. 1999. Terrestrial Polychaetes – Models for the Evolution of Clitellata (Annelida)? *Hydrobiologia* **406**: 87-99.

Ratnasingham, S., and Hebert, P.D.N. 2007. BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes* **7**: 355-364.

Rees, D.J., Duresne, F., Glémet, H., and Belzile, C. 2007. Amphipod genome sizes: first estimates for Arctic species reveal genomic giants. *Genome* **50**: 151-158.

Rolleston, G., and Jackson, W.H. 1888. "Medicinal Leech (*Hirudo medicinalis*)" *In* Forms of Animal Life: A manual of comparative anatomy with descriptions of selected types. Clarendon Press, Oxford, UK. pp. 212-216.

Römbke, J., and Egeler, P. 2009. "Oligochaete worms for Ecotoxicological Assessment of Soils and Sediments". *In* Annelids in Modern Biology (ed. D.H. Shain). John Wiley and Sons, Ltd., Hoboken, NJ. pp. 237

Rouse, G.W. 2002. "Annelida (Segmented Worms)". *In* Encyclopedia of Life Sciences. John Wiley and Sons, Ltd., Hoboken, NJ. pp. 1-8.

Satchell, J.E. 1980. "'*r*' Worms and '*K*' Worms: A Basis for Classifying Lumbricid Earthworm Strategies". *In* Soil Biology as Related to Land Use Practices. (ed. D.L. Dindal). U.S. Environmental Protection Agency, Washington, D.C. pp. 848–864.

Sawyer, R.T. 1971. The phylogenetic development of brooding behaviour in the Hirudinea. *Hydrobiologia* **37**: 197-204.

Schückel, U., Ehrich, S., and Kröncke, I. 2010. Temporal variability of three different macrofauna communities in the northern North Sea. *Esuarine, Coastal and Shelf Science* **89**: 1-11.

Sella, G., Redi, C.A., Ramella, L., Soldi, R., and Premoli, M.C. 1993. Genome size and karyotype length in interstitial polychaete species of the genus *Ophryotrocha* (Dorvilleidae). *Genome* **36**: 652-657.

Sella, G. 2006. "Sexual Strategies and Mating Systems". *In* Reproductive Biology and Phylogeny of Annelida (eds. G.W. Rouse and F. Pleijel). Science Publishers, Enfield, NH. pp. 77-89.

Shankland, M., Bissen, S.T., and Weisblat, D.A. 1992. Description of the Californian leech *Helobdella robusta* sp. nov., and comparison with *Helobdella triserialis* on the basis of morphology, embryology, and experimental breeding. *Canadian Journal of Zoology* **70**: 1259-1263.

Siddall, M.E. 2002. Phylogeny of the leech family Erpobdellidae (Hirudinea : Oligochaeta). *Invertebrate Systematics* **16**: 1-6.

Siddall, M.E., and Burreson, E.M. 1995. Phylogeny of the Euhirudinea: Independent evolution of blood feeding by leeches? *Canadian Journal of Zoology* **73**: 1048-1064.

Siddall, M.E., and Burreson, E.M. 1996. Leeches (Oligochaeta?:Euhirudinea), their phylogeny and the evolution of life-history strategies. *Hydrobiologia* **334**: 277-285.

Siddall, M.E., Trontelj, P., Utevsky, S.Y., Nkamany, M., and Macdonald III, K.S. 2007. Diverse molecular data demonstrate that commercially available medicinal leeches are not *Hirudo medicinalis*. *Proceedings of the Royal Society B* **274**: 1481-1487.

Sket, B., and Trontelj, P. 2008. Global diversity of leeches (Hirudinea) in freshwater. *Hydrobiologia* **595**: 129-137.

Smithsonian Institution 2012. "*Haementeria ghilianii* de Filippi, 1849 (Giant Amazon Leech) "Grandma Moses"". National Museum of Natural History, Department of Invertebrate Zoology. http://invertebrates.si.edu/Features/stories/haementeria.html.

Soldi, R., Ramella, L., Gambi, M.C., Sordino, P., and Sella, G. 1994. "Genome Size in Polychaetes: Relationships with Body Length and Life Habitat". *In* Actes de la 4éme Conference Internationale des Polychètes, Angers, France, 27 Juillet – 1 Aout 1992 (ed. J.-C. Dauvin et al.). Editions du Muséum National d'Histoire Naturelle, Paris. pp. 129-135.

Tasmania Parks and Wildlife Service 2011. Leeches. http://www.parks.tas.gov.au/index.aspx?base=17016.

Vendrely, R., and Vendrely, C. 1948. La Teneur du Noyau Cellulaire en Acide Désoxyribonucléique à Travers les Organes, les Individus et les Espèces Animales: Techniques et Premiers Résultats. *Experientia* **4**: 434–436.

Weisblat, D.A., Price, D.J., and Wedeen, C.J. 1988. Segmentation in leech development. *Development* **104**: 161-168.

Wells, M.D., Manktelow, R.T., Boyd, J.B., and Bowen, V.B. 1993. The Medical Leech: An Old Treatment Revisited. *Microsurgery* **14**: 183-186.

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* **352**: 123-137.

Zulhisyam, A.K., Ismail, A.A., and Omar, I.C. 2011. Optimization of growth conditions of *Hirudinea* sp. *Australian Journal of Basic Applied Sciences* **5**: 268-275.