A review of molecular markers used for Annelid phylogenetics

Kenneth M. Halanych¹ and Alexis M. Janosik

Auburn University, Department of Biological Sciences, Auburn, AL 36830, USA

Synopsis Annelida, one of the most successful animal phyla, exhibits an amazing variety of morphological forms. Disparity between some of the forms is so great that until molecular tools were used, some annelid lineages (for example, echiurids and pogonophorans) were not commonly recognized as belonging to the group. Although it is easy to assign annelids to a given family, understanding the deeper relationships within the group has been difficult. The main working hypothesis for annelid phylogeny is based on morphological cladistic analysis. However, the recent work using molecular tools has caused a revision of our view of annelid evolutionary history. For example, Scolecida and Palpata appear not to be natural groups, and the phylogenetic positions of some aberrant taxa (for example, Siboglinidae, Poeobiidae, Pisidae) have been determined. Herein, we discuss some of the main molecular markers that have been used to elucidate annelid phylogeny and the contribution that such work is making to our understanding. A table highlighting the molecular literature and the genes used is included.

Annelida is one of the most successful major animal lineages. Segmented worms, with over 16,500 recognized species (Brusca and Brusca 2003), are one of the dominant fauna in most marine habitats and have successfully radiated into fresh water and terrestrial environments. Despite their success and ecological importance, the evolutionary history of the group is still an enigma in numerous respects.

The purpose of this contribution is to elucidate issues and present tools that will help us unravel the evolutionary history of annelids. Herein, we assess the outlook for reconstructing annelid history. We will not present a thorough review of the recent phylogenetic literature for two reasons. First, McHugh (2000, 2001, 2005) has relatively recently summarized the literature on annelid systematics. Since those reviews considerable work has been undertaken, but the emphasis has been within specific annelid lineages and the understanding of relationships among major lineages has changed to a more limited degree. Second, at the time of writing this manuscript several laboratories were on the verge of making major contributions to the understanding of annelid evolutionary history, making such a review premature.

Understanding the evolutionary history of any group can be decomposed into several distinct questions. Before evolution of organismal features can be inferred, we need to have a framework for comparison. Phylogeny, the actual relationships between lineages, is that framework. There is only one phylogeny or true set of relationships. In other words, barring the presence of parallel universes, there is only one evolutionary history of any given organism. Our attempt to determine that phylogeny takes the form of phylogenetic hypotheses, often referred to as topologies, trees, and so on. Once a well-supported estimation of phylogeny, or a reasonable number of alternatives is obtained, we can begin to explore how lineages have changed over time. In any given study we may be interested in changes in morphology, behavior, genes, physiology, reproduction, or some other aspect of organismal biology. To understand any one of these, we may have to try and assess the environment, organismal interactions, and other possible selective forces that acted over the time period in question. Obviously, assessing these factors can be very difficult and, in some cases, speculative at best. In the case of annelid evolution, our ability to address evolutionary history is still at the first step of reconstructing a phylogenetic framework. Limited knowledge of annelid phylogeny hinders our ability to fully understand their evolution.

There has been a long standing recognition (see Fauchald 1977; Fauchald and Rouse 1997) that polychaete annelids could, for the most part, be relatively easily placed into 70–80 recognized families, but the relationships between those groups are not well understood. Historically, there has been little cross talk between researchers working on polychaetes and...
those working on clitellates (that is, oligochaetes and leeches). However, with the growth of phylogenetic theory and interest in the past 15 years, research on annelid phylogeny has been proceeding.

**Inclusiveness of annelids**

Issues that are generally perceived to be the large problems for understanding annelid phylogeny deal with events early in annelid evolution. The reviews of McHugh (2001 and reiterated in 2005) indicated that important unanswered questions included: “what are the relationships between polychaete annelids?, what is the sister group to Clitellata?, what extant group is most basal on the annelid tree?, and what group is sister to the Annelida?” Unfortunately, many of these remain unanswered.

Among annelid workers there has really only been one working hypothesis for the interrelationships of major annelid lineages. Rouse and Fauchald’s (1997) morphological cladistic analysis of Annelida has served as a useful tool by providing an explicit and testable phylogenetic hypothesis. This analysis has been subsequently updated (Rouse 1999; Rouse and Pleijel 2001) as scoring of characters was refined and new taxa were added (Fig. 1). In particular, placement of some groups in the original analyses was problematic because the difference between primary absence of a feature and secondary loss was not recognized (for example, Westheide and others 1999; Purschke and others 2000; Hessling 2002). In other words, a character that has never been present and a character that has been present and then lost in a descendant lineage will be scored the same way (that is, as absent) in the data matrix. Scoring of several morphological characters, for example presence of segmentation, presence of parapodia, state of the coelom, and presence of nuchal organs, are subject to this problem. Recognition of primary absence versus secondary loss requires knowledge about evolutionary history of the group in question. Making assumptions of phylogeny to determine character state evolution can lead to circular reasoning if those characters are then used to resolve phylogeny. Furthermore in the case of morphology, so little is understood about developmental mechanisms and functional constraints of morphology over evolutionary time, it is usually impossible to a priori objectively distinguish primary absence from secondary loss. In practice, primary absence versus secondary loss is less of a problem with DNA sequence data because of the more limited range of possible character states and patterns character state change are better understood.

In the case of the original Rouse and Fauchald trees, the net effect of not assuming secondary loss was to place derived taxa (Clitellata, Echiura, and Siboglinidae) basal. Despite this issue, the Rouse and Fauchald tree has been a major tool for advancing our knowledge of annelid relationships because it has served as a testable working hypothesis. They formalized novel terms like Scolecida, Canalipalpata, and Aciculata, as well as several terms that were already in use (highlighted in Fig. 1).

Other “global” hypotheses of annelid relationships have been put forward based mainly on 18S nuclear ribosomal data (for example, Colgan and others 2001; Rota and others 2001; Struck Hessling, and Purschke 2002; Bleidorn and others 2003a; Hall and others 2004). Unfortunately, when using 18S data alone, relationships deep within annelids are not well supported. In particular, recognized families that are well supported by morphology often come out as polyphyletic. Thus, topologies from such studies have been ignored as working hypotheses of annelid evolution. This lack of support and short internal branch lengths deep in the tree has lead to the suggestion that early annelid evolution was characterized by a rapid radiation of the major lineages (McHugh 2000, 2001). Addition of more sequence data or taxa may help to provide resolution.

Intra-annelid relationships have typically employed a limited number of genes with the most heavily used being 18S, followed by CO1 and 28S. Table 1 summarizes molecular phylogenetic studies focusing on deep annelid issues. A few interesting trends are readily apparent in the table. Most obviously, the bulk of work on relationships within annelids is very recent. Additionally, based on 2005, the amount of effort appears to be increasing. However, the choice of markers is still limited. Below, we outline the utility of these different genes and how they might be applicable to future efforts within Annelida.

**18S**—This nuclear ribosomal gene is also commonly referred to as the nuclear small ribosomal subunit (SSU) and is ~1800–2000 nucleotides in length. This gene is part of a tandem repeated element in the nuclear genome that also includes the 28S, 5.8S, and the internal transcribed spacer regions (ITS) (reviewed in Hillis and Dixon 1991). There are hundreds of copies of this repeat in the genome that are typically homogenized by concerted evolution. 18S data have been used to address intra-annelid relationships mainly for historical reasons in that the early article of Field and others (1988) began to build a database that was easy to add onto. Additionally, conserved regions throughout animals has allowed for the development of universal primers for amplification via polymerase chain reaction (PCR), and variation in nucleotide sequence in different gene regions facilitates obtaining information at several different phylogenetic levels.
Fig. 1  The current understanding of annelid phylogeny based on Rouse and Pleijel (2001) which is a modification of Rouse and Fauchald (1997). Interested parties should also see the more extensive “metatree” in Rouse and Pleijel (2001). This phylogenetic hypothesis is based largely on non-molecular data using a parsimony approach.
<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Annelid origins</th>
<th>Interfamily</th>
<th>18S</th>
<th>28S</th>
<th>CO1</th>
<th>16S</th>
<th>12S</th>
<th>mtDNA genome</th>
<th>EF-1α</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eernisse</td>
<td>1997</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim and others</td>
<td>1996</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martin</td>
<td>2001</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moon and others</td>
<td>1996</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peterson and Eernisse</td>
<td>2001</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winnepenninckx and others</td>
<td>1995</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winnepenninckx and others</td>
<td>1998</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boore and Staton</td>
<td>2002</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Borda and Siddall</td>
<td>2004</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martin and others</td>
<td>2000</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Struck and others</td>
<td>2006</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bleidorn and others</td>
<td>2005</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rousset and others</td>
<td>2004</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apakupakul and others</td>
<td>1999</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erseus and others</td>
<td>2000</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Halanych and others</td>
<td>1998</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Struck, Purschke, and Halanych</td>
<td>2005</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiklund and others</td>
<td>2005</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bleidorn</td>
<td>2005</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bleidorn and others</td>
<td>2003a</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bleidorn and others</td>
<td>2003b</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burnett and others</td>
<td>2005</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erseus</td>
<td>2005</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erseus and others</td>
<td>2002</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erseus and Källersjö</td>
<td>2003</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erseus and Källersjö</td>
<td>2004</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gelder and Sidall</td>
<td>2001</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Halanych and others</td>
<td>2001</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hall and others</td>
<td>2004</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jördens and others</td>
<td>2004</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nygren and Sundberg</td>
<td>2003</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rota and others</td>
<td>2001</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Siddall and others</td>
<td>2001</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Struck, Hessling, and Purschke</td>
<td>2002</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Struck, Westheide, and Purschke</td>
<td>2002</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worsaae and others</td>
<td>2005</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christensen and Theisen</td>
<td>1998</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown and others</td>
<td>1999</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H3, U2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colgan and others</td>
<td>2001</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rousset and others</td>
<td>2003</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spolin and others</td>
<td>2005</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dahlgren and others</td>
<td>2000</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nylander and others</td>
<td>1999</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Siddall and Burreson</td>
<td>1998</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boore and Brown</td>
<td>2000</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Articles using molecular tools to focus on deep phylogenetic issues of annelids.
In the case of annelids, the 18S has both pros and cons (Hillis and Dixon 1991; Abouheif and others 1998; Halanych 2004). Issues with variation of nucleotide substitution rates across lineages are well known and, to some degree, can be factored into analyses. Although it has not been commonly done in annelids, it has been successfully considered in other groups (for example, Strepsiptera flies—Huelsenbeck 1998; Mollusks—Passamaneck and others 2004). As mentioned, 18S has failed to elucidate deep annelid relationships. However, it has been helpful within recognized families (for example, Nygren and Sundberg 2003; Bleidorn 2005), with placing some recognized families within others (for example, Erseus and others 2002; Burnette and others 2005), and with occasionally identifying sister taxa (Rousset and others 2004). It has also been particularly useful within some specific clades such as Clittelata (for example, Erseus and others 2000; Martin and others 2000; Siddall and others 2001; Erseus and Källersjö 2003, 2004; Borda and Siddall 2004). The 18S will remain an important tool for annelid systematics and is the most explored gene within the taxon.

28S—The nuclear large ribosomal subunit (LSU), or 28S, is physically linked to the 18S in the tandem repeat and is typically 2800–3000 nucleotides in length. We are still learning about the utility of this gene for annelid phylogeny. Several studies (for example, Nygren and Sundberg 2003; Bleidorn 2005), with placing some recognized families within others (for example, Erseus and others 2002; Burnette and others 2005), and with occasionally identifying sister taxa (Rousset and others 2004). It has also been particularly useful within some specific clades such as Clittelata (for example, Erseus and others 2000; Martin and others 2000; Siddall and others 2001; Erseus and Källersjö 2003, 2004; Borda and Siddall 2004). The 18S will remain an important tool for annelid systematics and is the most explored gene within the taxon.

28S—The nuclear large ribosomal subunit (LSU), or 28S, is physically linked to the 18S in the tandem repeat and is typically 2800–3000 nucleotides in length. We are still learning about the utility of this gene for annelid phylogeny. Several studies (for example, Nygren and Sundberg 2003; Bleidorn 2005), with placing some recognized families within others (for example, Erseus and others 2002; Burnette and others 2005), and with occasionally identifying sister taxa (Rousset and others 2004). It has also been particularly useful within some specific clades such as Clittelata (for example, Erseus and others 2000; Martin and others 2000; Siddall and others 2001; Erseus and Källersjö 2003, 2004; Borda and Siddall 2004). The 18S will remain an important tool for annelid systematics and is the most explored gene within the taxon.

CO1—The mitochondrial cytochrome c oxidase subunit 1 gene, part of the electron transport chain, is the most commonly sequenced mitochondrial gene for non-vertebrate animals. As with 18S and 28S, this is largely due to availability of primers that amplify an ~710 nucleotide fragment. However, the Folmer et al. primers (Folmer and others 1994) do not work well on several groups of annelids (and some mollusks), and new annelid specific primers should be developed. In contrast to the ribosomal genes, protein coding genes usually capture phylogenetic signal at two levels. Nucleotide substitutions accumulate most quickly in the third, or wobble, position of the codon. Typically, the third codon position is most useful for intraspecific (for example, Kojima and others 2002, 2003) to intrageneric analyses (for example, Black and others 1997). In contrast, first and second (the most conservative) position changes usually cause amino acid substitutions and thus accumulate slower presumably due to selective constraints. Variation in the first two positions can be examined at either the nucleotide or amino acid level and is helpful with elucidating intergeneric to interfamilial relationships. One potential problem of coding genes is that situations can be encountered where third positions have experienced multiple substitutions obliterating phylogenetic signal (that is, saturation) before amino acid changes have occurred. The result is a region of the evolutionary history that is not resolvable with the marker in question (Halanych and Robinson 1999).

16S—This marker is the mitochondrial version of the large ribosomal subunit. In most analyses to date, only a short 450–500 nucleotide fragment of this gene has been examined corresponding primers designed by
Palumbi’s group (1991). This region is typically useful for intraspecific and intragenic level relationships (for example, Dahlgren and others 2001; Halanych and others 2001; Jolly and others, 2006; Schulze 2006) and has limited utility at higher levels (Struck and others 2006). However, the utility of a larger region of, or the complete, 16S gene is unknown. With the increase of known mitochondrial genomes available for annelids, it should be possible to explore the utility of the 16S and design novel primers that span a longer region.

**mtDNA genome**—In most animals, the mitochondrial genome is ~15,000 bp and holds phylogenetic information that can be examined as gene rearrangement data, amino acid data, or nucleotide data. As of April 2006, only four complete annelid genomes were available in GenBank, with an additional three taxa for which a considerable portion was known. However, there is active work in this area (Valles and Boore 2006). The nearly complete genomes are the result of difficulties with amplifying the control region (also called the D-loop or unknown region) of mtDNA genomes (Boore and Brown 2000; Jennings and Halanych 2005). All the reported annelid mtDNA genomes show a remarkable degree of conservation in gene order suggesting that analysis of concatenated coding and ribosomal genes may be more promising. It is noteworthy that sipunculans share the same gene order as many annelids (Boore and Staton 2002; Jennings and Halanych 2005; Bleidorn and others 2006).

**EF-1α**—Elongation factor-1α is nuclear gene involved in part of the cell’s protein synthesis machinery. This marker has been used for deep level questions such as annelid origins or inclusiveness. To date Kojima and co-workers (Kojima and others 1993; Kojima 1998), and McHugh (McHugh 1997) are the only researchers involved mainly on siboglinid relationships and origins. This marker will be useful to some degree within annelids and it can be exploited at both the amino acid or nucleotide levels. Expect more data to be forthcoming from EF-1α.

**Other genes**—The other genes that have been used in deeper level annelid analysis include the mitochondrial 12S (or mitochondrial small ribosomal subunit) and cytB (cytochrome oxidase subunit B) genes, and the nuclear H3 (Histone subunit 3) and U2 snRNA genes. The utility of the mitochondrial genes for such issues is not well known, whereas these nuclear genes are of limited use because they are too conserved or too short (see Brown and others 1999). Clearly, additional markers need to be developed.

Although advances in our understanding of annelid relationships has been slow in coming, the application of genomic tools, more labs addressing key issues, and additional taxon sampling are certain to yield considerable insight over the next several years. In fact, the recent molecular work is already reshaping our thoughts on annelid evolution. Figure 2 shows a version of the morphological tree (Rouse and Fauchald 1997; Rouse and Pleijel 2001) that has been modified according to recent findings with molecular tools (Table 1). At first glance, there are clearly some taxa observed in both versions of the tree. For example, Aciculata, Phyllodocida, several Phyllodocida subclades, and several Canalipalpata subclades appear to be consistent with both morphology and molecules. Note that many of these clades have not been rigorously tested, but at present there is no reason to doubt them.

There are, however, some notable differences between the trees. Molecular work consistently fails to recover “Scolecida” as monophyletic and they appear to be a disparate group of unrelated taxa. Likewise, “Palpata” is not recovered. Given these results, we must consider that the possession of palps is an ancestral character for annelids that has been subsequently lost in several taxa. Although myzostomids, a bizarre group of parasites on echinoderms, have been included in the Rouse and Pleijel tree, no molecular evidence to date supports their inclusion in annelids (Eeckhaut and others 2000; Passamaneck and Halanych 2006).

As mentioned above, there is a growing body of data that suggest annelids and sipunculans are closely related (Boore and Staton 2002; Jennings and Halanych 2005; Passamaneck and Halanych 2006; Fig. 2 node A). This evidence is also consistent with a recent morphological study of neural and muscle formation in the sipunculan *Plaiceolus strombus* (Wanninger and others 2005). Another major advance in our understanding is the placement of echinurids next to capitellids (Bleidorn and others 2003b; Hall and others 2004; Fig. 2 node B) showing that closely related groups can have very different segmentation patterns. Arenicolid is sister to Maldanids (Bleidorn and others 2005; Fig. 2 node C). Combined molecular and morphological evidence also suggests that Siboglinids may be close to Oweniids (Rousset and others 2004; Fig. 2 node D) and close to Sabellids (see Rouse and Fauchald 1997). Although there is abundant evidence that Clitellates are derived polychaetes, their placement is still uncertain.

We also have evidence for the placement of some recognized families or problem taxa within other groups. For example, *Questa* is within Orbinidae (Bleidorn and others 2003a, b; Fig. 2 node E), and Ctenodrilidae within Cirratulidae (Bleidorn and others 2003b; Fig. 2 node F). Also Poeobiidae, and probably
An updated view of annelid phylogeny that incorporates molecular findings of Table 1 into the general backbone of the Rouse and Pleijel (2001) tree (Fig. 1). Although no explicit analysis was conducted, nodes which lacked support were collapsed (for example, Scolecida and Palpata). Other groupings of interest are indicated by the letters and referred to in the text.

Fig. 2 An updated view of annelid phylogeny that incorporates molecular findings of Table 1 into the general backbone of the Rouse and Pleijel (2001) tree (Fig. 1). Although no explicit analysis was conducted, nodes which lacked support were collapsed (for example, Scolecida and Palpata). Other groupings of interest are indicated by the letters and referred to in the text.
Flota, are within Flabelligeriidae (Burnette and others 2005; Fig. 2 node G). Phylogenetic hypotheses are also emerging for some of the better-known worm groups. In particular, two recent articles (Struck, Purschke, and Halanych 2005; Wiklund and others 2005; Fig. 2 node H) have shown that sigalionids, pisionids, and pholoids are nested within Aphroditoida, while placing Aphroditidae as the most basal group. Molecular data has also challenged our views of Eunicida evolution (Fig. 2 node I). The monophyly of Dorveillidae has been into question (Struck, Westheide, and Purschke 2002; Struck, Halanych, and Purschke 2005; Struck and others 2006) and Lumbrinerids are more basal than traditionally believed (Struck and others 2006).

Future directions
To date, most of the ambiguous attempts (for example, Brown and others 1999; Hall and others 2004) to resolve annelid interfamilial relationships have been unable to resolve the base of the annelid tree and have low support for deep branches. This situation suggests that annelids may be the result of a rapid radiation event (McHugh 2001), for example, the Cambrian explosion. However, even if this is the case, the next several years will hold tremendous advances for our understanding of annelid phylogeny. As genetic tools continue to become more powerful and cost efficient, our knowledge will improve.

Based on how annelid phylogenetic research has progressed, it will be more realistic to expect relationships within several recognized clades (for example, Eunicida, Sabellida, Phyllodocida, and so on) to be reasonably worked out before the deepest branches are elucidated. In addition to increased cooperation within the community, there are several issues that should be addressed by all working on deeper annelid relationships:

1. Develop a common suite of genes to be used by annelid researchers. This does not mean that researchers should be limited to these genes. Initial possible choices include the 18S, the 2.1 kb 28S fragment, and CO1.

2. Develop additional markers. The number of genes used in phylogenetic studies across most animals, other than arthropods and vertebrates, is limited. With the advent of whole nuclear genome sequencing and EST projects in annelids, this situation should be remedied in the near future.

3. Demand that morphological and molecular vouchers be deposited to an appropriate museum. Even between taxonomic specialists there can be disagreement on taxonomic issues and identification. Enforcing vouchers can be achieved through the review and editorial process.

4. As a community, push other groups of researchers that use genomic tools (for example, developmental biologists, physiologists, and genetists) to embrace annelid models. As more researchers become interested in annelids, the knowledge base and the resources will increase to the advantage of all.

Acknowledgments
We are grateful for support in various forms from the Society of Integrative and Comparative Biology (divisions support from DDEB, DSEB, and DIZ) and Friday Harbor Laboratories, University of Washington. Constructive comments by anonymous reviewers were most helpful. This work was supported by the WormNet grant (EAR-0120646) from the National Science Foundation to K.M.H. Contribution #13 to the AU Marine Biology Program.

Conflict of interest: None declared.

References


Vestimentifera (tube worms) and Annelida revealed by amino acid sequence of elongation factor - 1a. J Mol Evol 37:66–70.


