Worms by number

C. J. Glasby¹*, S. P. Glasby² and F. Pleijel³

¹Museum and Art Gallery of the Northern Territory, GPO Box 4646, Darwin, NT 0801, Australia
²Department of Mathematics, Central Washington University, Ellensburg, WA 98926, USA
³Tjärnö Marine Biological Laboratory, Department of Marine Ecology, Göteborg University, 452 96 Strömstad, Sweden

This paper investigates alternation patterns in length, shape and orientation of dorsal cirri (fleshy segmental appendages) of phyllodocids, a large group of polychaete worms (Annelida). We document the alternation patterns in several families of Phyllodocida (Syllidae, Hesionidae, Sigalionidae, Polynoidae, Aphroditidae and Acoetidae) and identify the simple mathematical rule bases that describe the progression of these sequences. Two fundamentally different binary alternation patterns were found on the first four segments: 1011 for nereidiform families and 1010 for aphroditiform families. The alternation pattern in all aphroditiform families matches a simple one-dimensional cellular automaton and that for Syllidae (nereidiform) matches the Fibonacci string sequence. Hesionidae (nereidiform) showed the greatest variation in alternation patterns, but all corresponded to various known substitution rules. Comparison of binary patterns of the first 22 segments using a distance measure supports the current ideas on phylogeny within Phyllodocida. These results suggest that gene(s) involved in postlarval segmental growth employ a switching sequence that corresponds to simple mathematical substitution rules.

Keywords: Fibonacci numbers; binary sequence; symmetry; substitution; growth rule; Annelida

1. INTRODUCTION

The number of applications of mathematics to biology is large and increasing. The study of biological patterns and symmetry mathematics, in particular, number theory, has a great application. Living organisms display a wide range of symmetries: reflectional, rotational, spiral and fractal. In the plant kingdom, symmetries and numerical patterns abound. For example, the spiral arrangement of petals in a flower, the bracts in a pine cone, the seeds on strawberries and the leaves of artichokes. The basic spiral pattern in these examples is the same and, not surprisingly, also the same as the arrangement of microscopic primordia in the tips of shoots of plants; both can be described using Fibonacci numbers (Adler et al. 1997). The configuration of each spiral is consecutive Fibonacci numbers, with the first two numbers, 1, 1, and subsequent numbers generated by the recurrence relation, \( F_n = F_{n-1} + F_{n-2} \), for \( n \geq 3 \). Its values are

\[
\begin{align*}
F_1 & = 1, \\
F_2 & = 1, \\
F_3 & = 2, \\
F_4 & = 3, \\
F_5 & = 5, \\
F_6 & = 8, \\
F_7 & = 13, \\
F_8 & = 21, \\
F_9 & = 34, \\
F_{10} & = 55, \\
\ldots
\end{align*}
\]

Compared with plants, animals have fewer examples of striking numerical patterns. However, one group, Mol-lusca, is worth mentioning because their numerical patterns involve both shape and colour—the complex markings on the shells of many marine molluscs, in particular cone shells, can be modelled accurately by cellular automata (geometric arrays of cells, each in one of a finite number of states; Herman & Liu 1973; Wolfram 2002), while other mathematical models can emulate the coiled form of gastropod shells (Raup 1962). Annelids, in particular the largely marine Polychaeta, also show a wide range of interesting patterns including the spiral shells of spirorbines, possible fractal patterns in the vividly coloured tentacular crowns of Sabellidae, and the complex reflection symmetry of the jaw apparatus of eunicimorphs; however, none appear to have been previously investigated from a mathematical point of view.

One group of Polychaeta, the Phyllodocida, shows a remarkable alternation pattern in the shape, length and orientation of fleshy segmental appendages called dorsal cirri (DC, e.g. Gidholm 1966), which can be characteristic of a species, genus or family. This study investigates the alternation patterns in the DC and seeks simple rules for describing them. Also, we investigate whether the mathematical similarity between sequences of different Phyllodocida groups agrees with our current knowledge of phylogeny in the group.

2. POLYCHAETE MORPHOLOGY AND DC ALTERNATION PATTERNS

Polychaetes, like other annelids, comprise a head (pros-tomium), a tail (pygidium) and in between a series of similar segments arranged in a linear sequence along the anteroposterior axis. Each segment carries a pair of lateral bristle-bearing parapodia (hence the origin of the name from the greek, polys, for many and chaities, for hair). Typically, the parapodium comprises a dorsal component, the notopodium (bearing bristles called notochaetae) and a ventral component, the neuropodium (bearing bristles called neurochaetae), although one or both elements and their chaetae may be reduced or absent. In Phyllodocida, DC are usually slender, tapering fleshy structures that arise from the upper edge of the notopodium (or on the
body wall above the notopodium) but they may be modified, becoming flattened and fleshy (e.g. Phyllodocidae) or scale like (e.g. Polynoidae). Once formed the type of DC does not change with growth, apart from the elongation of anterior ones associated with cephalization. This last assumption has been observed and confirmed in the Hesionidae (F.P.).

The DC alternation pattern differs among Phyllodocida: in the nereidiforms—Chrysopetalidae, Hesionidae and Syllidae—the DC alternation pattern involves regular changes in position (elevated or horizontal) and relative length along the body (figures 1 and 2). By contrast, Nereididae and Pilargidae have non-alternating DC, although the DC of some nereids may become progressively larger and more laterally positioned on the parapodium but this is not an alternation pattern. In the aphroditiforms—Acoetidae, Aphroditidae, Polynoidae and Sigalionidae—DC alternate with elytrae (scales; figure 2). Elytrae are probably homologous to the elevated or long type DC of Nereidiformia because like the latter they are more dorsally (and medially) positioned than ‘normal’ DC. As Rouse & Pleijel (2001, p. 73) point out though, the nature of elytra is not straightforward: non-elytra-bearing segments in many taxa have dorsal tubercles in addition to DC, so it may be that the dorsal tubercles in fact are the homologues of nereidiform DC. Although the issue warrants further study, it probably has no bearing on the present study since the dorsal tubercles occur on the same segment as the DC in the taxa studied.

Phylogenetic relationships within Phyllodocida are reasonably well known, which provide us with a framework for comparing DC alternation patterns. Within Nereidiformia, evidence from morphology suggests that Chrysopetalidae is the sister group of Hesionidae, with Nereididae, Pilargidae and Syllidae successively more distantly related (figure 3; Glasby 1993; Pleijel & Dahlgren 1998). Recent phylogenetic studies within Aphroditiformia based on both morphological and molecular evidence have suggested that some family level taxa should be subsumed within others—Polynoidae (and Eulepethidae) are the sister groups to Acoetidae, with Sigalionidae (including Pholoidae and Pisionidae) sister to these and Aphroditidae more distant (Struck et al. 2005; Wiklund et al. 2005). If DC alternation patterns are heritable traits then we would expect to find similarities at the family level that are congruent with our current knowledge of phylogeny.

3. OBSERVATIONS AND ANALYSIS
Observations of DC alternation patterns were based on analysis of living and dead specimens and scanning electron and light microscope micrographs. Alternation patterns were scored for seven families—Acoetidae, Aphroditidae, Chrysopetalidae, Hesionidae (see also Pleijel 1998; Ruta & Pleijel 2006), Polynoidae, Sigalionidae and Syllidae—from segments 1 to 39 as long (L), short (S) or elytra (E). Elytrae were assumed to be homologous with long DC (i.e. E = L). For some groups, characterization of the DC on segments 1–4 in adults was problematic. In Hesionidae, the DC of normal locomotory parapodia become elongated during the cephalization process and transformed into sensory DC; in this case we used larval information, when available, to clarify their original form. In Chrysopetalidae, the small size of adults (often only a few mm length), reduced size of anterior DC in some groups (e.g. Dysponetus) and their retractile nature in others (e.g. Bhawania), meant that the form of the DC on segments 1–4 could not be determined accurately in the preserved specimens available (observations on living specimens are required).
In order to analyse the binary sequence of these patterns, we assigned L the value 0 and S the value 1 (in one instance, we reversed the roles of 0 and 1: the symbols used are not important but the pattern of symbols is). Binary sequences were scored for the first 22 segments, as after that we found that alternation patterns for several groups break down (see also Fauchald 2000). Next, we searched the on-line encyclopaedia of integer sequences (OEIS; Sloane 2003), a database with over 100 000 sequence entries and citations to the refereed literature. Sequence hits were then filtered on the basis of (i) length of match, (ii) mathematical simplicity and (iii) biological plausibility, such that long-match simple sequences that are biologically plausible were favoured.

4. RESULTS

The primary data on alternation sequences for all seven families is archived online (see electronic supplementary material). Alternation sequences for Chrysopetalidae are incomplete (see above) and for Sigalionidae (segment 3) the form of the DC was variable but most commonly represented by a short DC (assumed to be the plesiomorphic state from here on). The DC alternation pattern between species within most families is relatively constant. Exceptions are the Hesionidae, for which there are at least four different patterns and the Syllidae, which all have the same pattern up to segment 28, but thereafter members of the subfamily Autolytinae vary (Gidholm 1966; Nygren 2004). In terms of common patterns, the first thing to note is that the first four terms of all nereidiforms is SLSS (Z1011) and that of the aphroditiforms is SESE (Z1010; table 1). This fits well with our predictions based on current knowledge of the phylogeny of Phyllodocida, viz. the nereidiform families are more closely related to each other than any are to aphroditiforms.

Sequence analysis indicated that the first 11 terms of the sequence for the Syllidae (Sy) match the first 11 terms of the Fibonacci string sequence (A036299). The binary sequences of Acoetidae (Ac), Aphroditidae (Ap), Polynoidae (Po) and Sigalionidae (Si) are the same, but different from Syllidae—the first 21 terms of sequences correspond to an OEIS subsequence of A076404: the parity of the perfect powers. The perfect powers are 1, 4, 8, 9, 16, 25, ..., and their parities are 1, 0, 0, 1, 0, 1, ... (the parity of an even number is 0 and an odd number is 1). The chance that the first 21 terms match is $2^{-21}$, or approximately 1 in 2 000 000! This sequence appears to be generated by only a global rule, and hence should be discarded as being biologically implausible. However, if 0 and 1 are reversed, then the first 21 terms are generated by the sequence A071028 corresponding to a local rule, a particularly simple one-dimensional cellular automaton called “Rule 50” by Wolfram (2002, ch. 3). See §5 for a discussion of local rules.

Hesionidae show the greatest variability in DC alternation pattern. The most similar two species, Psamathe fusca and Sirsoe methanicola, have the same pattern up to segment 31, after which it deviates slightly.
In both species, the alternation sequence from segment 5 (the first non-tentacular DC) to segment 21 match exactly the 17 terms of the sequence $t_n$ described by the substitution rule $t_1 = L$, and $L \rightarrow LSS$, $S \rightarrow LS$. If $L \rightarrow 1$ and $S \rightarrow 0$ then this is sequence A004641, *viz.*

<table>
<thead>
<tr>
<th>taxon</th>
<th>Am</th>
<th>He</th>
<th>Ne</th>
<th>Ps/Sm</th>
<th>Sy</th>
<th>Ac/Ap/Po/Si</th>
</tr>
</thead>
<tbody>
<tr>
<td>Am</td>
<td>0</td>
<td>$\leq 2^{-21}$</td>
<td>$2^{-20}$</td>
<td>$2^{-20}$</td>
<td>$2^{-8}$</td>
<td>$2^{-5}$</td>
</tr>
<tr>
<td>He</td>
<td>0</td>
<td>$2^{-20}$</td>
<td>$2^{-20}$</td>
<td>$2^{-8}$</td>
<td>$2^{-8}$</td>
<td>$2^{-5}$</td>
</tr>
<tr>
<td>Ne</td>
<td>0</td>
<td>$2^{-23}$</td>
<td>$2^{-8}$</td>
<td>$2^{-8}$</td>
<td>$2^{-8}$</td>
<td>$2^{-5}$</td>
</tr>
<tr>
<td>Ps/Sm</td>
<td>0</td>
<td>$2^{-8}$</td>
<td>$2^{-8}$</td>
<td>$2^{-8}$</td>
<td>$2^{-5}$</td>
<td>$2^{-5}$</td>
</tr>
<tr>
<td>Sy</td>
<td>0</td>
<td>$2^{-5}$</td>
<td>$2^{-5}$</td>
<td>$2^{-5}$</td>
<td>$2^{-5}$</td>
<td>$2^{-5}$</td>
</tr>
<tr>
<td>Ac/Ap/Po/Si</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Two other types of DC alternation pattern may be recognized in the family. Members of *Heteropodarke* and *Gyptis* have the same pattern as *Amphidurus fuscescens*, and members of *Syllidia* have the same pattern as *Nereimyra punctata*. The alternation pattern in *P. fusca* and *S. methanicola* is the same as that occurs in members of *Micropodarke*. These groupings based on DC alternation pattern show high congruence with hesionid phylogeny as we know it today (Ruta *et al.* 2007), with groupings such as *Micropodarke*, *Psamathe*, *Sylsoe*, *Syllidia* and *Nereimyra* corresponding to one clade and having a common pattern until segment 18, and the closely related *Micropodarke* and *Psamathe*, and *Syllidia* and *Nereimyra*, respectively, having identical patterns.

Another measure of the similarity between the DC alternation sequences between different polychaetes is by introducing a distance function on the sequences. Suppose that two sequences compared from left to right (i.e. head to tail) differ first in the $n$th segment. If $n$ is large we say that the alternation sequences are close, and if $n$ is small they are far apart (table 2). We find that the DC sequences for the hesionid taxa (*Amphidurus*, *Hesione*, *Nereimyra*, *Psamathe* and *Sirsoe*) are close to those of *Syllidia*, and both of these are less close to *Acoetidae* (*Ac*), *Aphroditidae* (*Ap*), *Polynoidae* (*Po*) and *Sigalionidae* (*Si*). The distance between two binary strings is defined as $2^{-n}$ if the first instance where the strings differ is in the $n$th position; the distance between identical strings is $2^{-\infty} = 0$.)

<table>
<thead>
<tr>
<th>$t_1$</th>
<th>$t_2$</th>
<th>$t_3$</th>
<th>$t_4$</th>
<th>$t_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSS</td>
<td>LSSLSLS</td>
<td>LSLSLS</td>
<td>LSLSLS...</td>
<td>LSLS LSLS</td>
</tr>
</tbody>
</table>

The Fibonacci string sequence has three noteworthy properties. First, it yields a constant morphology of the anterior-most segments, as the left end of the sequence is always the same. Second, the string $s_n$ at the $n$th growth stage is a concatenation of the previous two growth stages: specifically $s_n = s_{n-1} + s_{n-2}$. For example, $s_1 = SL$, $s_2 = SLSS$ and $s_4 = SLSSLS = (SLSSLS)(SLSS) = s_5s_3$. This second observation is a key to why the anterior end morphology is constant. Third, the substrings LL and SSS never occur.

When using the OEIS, care must be taken to select only sequences that correspond to potentially plausible biological rules. Such rules tend to be ‘local’ rather than ‘global’. Local rules may be generated by substitution systems, cellular automata, Turing machines, or tag systems (Wolfram 2002). An example of a local rule is the substitution system ‘L → S’ and ‘S → SL’ defining the Fibonacci string sequence. It uses only local information: substrings of length 1. By contrast, the global rule $s_n = s_{n-1} + s_{n-2}$ for the same Fibonacci string

---

**Table 2.** Similarity matrix comparing binary sequences on segments 1–22 (table 1). (Rows and columns ordered to show that sequences for *Amphidurus fuscescens* (Am), *Hesione splendida* (He), *Nereimyra punctata* (Ne), *Psamathe fusca* (Ps), *Sirsoe methanicola* (Sm; all Hesionidae) are close to *Syllidae* (Sy), and both of these are less close to *Acoetidae* (Ac), *Aphroditidae* (Ap), *Polynoidae* (Po) and *Sigalionidae* (Si). The distance between two binary strings is defined as $2^{-n}$ if the first instance where the strings differ is in the $n$th position; the distance between identical strings is $2^{-\infty} = 0$.)

5. DISCUSSION

**a) Polychaete growth and growth rules**

Similar to other annelids, polychaetes grow by the addition of new segments (each equipped with parapodia-bearing cirri of various lengths and orientations) from the posterior end of the body called the posterior growth zone (Anderson 1973). The number of adult segments varies considerably; it may be fixed or segments may be added throughout life. Once a segment is formed, the basic alternation pattern in length and orientation—short/lateral or long/elevated—of the DC of each segment does not change during growth, apart from the elongation of the ones on anterior segments related to cephalization in the early stages of some species (figure 2). Unfortunately, nothing is known about the mechanism controlling segmentation.

There may be many growth rules that account for a given string of short/lateral (S) and long/elevated (L) DC. For example, simple addition of new segments at the growth zone requires the polychaete to repeatedly defer to its genetic code to know, for example, whether segment 20 should have short or long DC. Hence, the polychaete must in some sense be able to count to 20. Thus, this rule, although simple in description, is not simple in implementation. By contrast, under a local-rule-based Fibonacci string sequence, new segments are produced using the substitution rules $L \rightarrow S$ and $S \rightarrow SL$. If the most recently added segment has long DC, the next one will have short DC; if the last one has short DC, the next two will have short and long DC, respectively. If the first segment is an S-segment, then a five-segment polychaete SLSSL, for example, grows to become an eight-segment polychaete (SL)(SL)(SL)S or SLSSLSSL. If $s_n$ denotes the $n$th growth stage, then

$$s_1 = SL, s_2 = SLSS, s_3 = SLSSL, s_4 = SLSSLSS, s_5 = SLSSLSSSL,$$

The global rule $s_n = s_{n-1} + s_{n-2}$ for the same Fibonacci string

---

2074  C. J. Glasby et al.  *Worms by number*
sequence requires that substrings of arbitrary length be remembered. For this reason rules such as A071028, called Rule 50 (Wolfram 2002: ch. 3), are more biologically plausible than rule A076404.

The Fibonacci growth rule is certainly simple: being determined by an initial segment S and two substitution rules \( L \rightarrow S \) and \( S \rightarrow SL \). A polychaete seeking to grow by these rules is not required to count or to constantly defer to its genetic code for information on DC form, although it does need to know its head from its tail. The Fibonacci string sequence is related to the sequence of Fibonacci numbers which have a well-known connection with plant growth and morphology. A simple growth rule such as this would facilitate regeneration of the correct DC sequence of posterior segments lost, for example, by damage from predation. However, as it does not work in the opposite direction, we predict that under this rule worms that lose their head (e.g. through predation)—and have the ability to regenerate anteriorly—are incapable of replicating the original DC alternation pattern. Among the families investigated here, only certain Syllidae are known to be able to regenerate anteriorly (Bely 2006).

(b) Evolution of alternating DC pattern in Phyllodocida

The phylogenetic origin of alternating DC in Phyllodocida is not entirely clear. The most parsimonious scenario involving three evolutionary steps is that the common ancestor of Aphroditiformia and Nereidiformia had the 1011 alternating pattern of the latter, it was transformed to the 1010 elytrae pattern of scale worms, and the 1011 alternation pattern was lost in Nereididae and Pilargidae (figure 3a). A loss hypothesis has also been suggested for certain rare cases of elytrae absence in, for example, Psioine and Pamilia (Struck et al. 2005; Wiklund et al. 2005). A less probable scenario in terms of the number of evolutionary steps (four) is that even-length cirri (1111) is ancestral, and alternating DC was acquired independently in scale worms via a two-step transformation (1111 → 1110 → 1010), and one step in each of Syllidae (1111 → 1011) and Chrysopetalidae + Hesionidae (1111 → 1011; figure 3b). In both scenarios, selective pressures favouring alternating DC are most likely the result of greater exposure to water currents because DC have a chemoreceptive function (Boilly-Marer 1972; Storch & Schlotzer-Schrehardt 1988; Heffneran 1990), and elytra in particular also have a respiratory function (Wiklund et al. 2005). This fits with what we know about the biology of these groups, viz. groups with alternating DC (and elytrae) tend to be epifaunal compared with taxa with even-length cirri (Pilargidae and Nereididae) and the elytrae-less Psioine, which both have a higher proportion of infanlal species.

Simple rules encoded in the genetic code of Phyllodocida provide an explanation of how a phyllodocid polychaete can 'know' the position in time and space of each segment. These same rules may also be involved in other examples of segmental awareness in Phyllodocida, including (i) fixed segment number—a few species of scale worms (figure 2b) and Hesionidae have a relatively limited, and fixed, compliment of segments as adults, i.e. once a certain number of segments is reached the growth zone in front of the pygidium becomes inactivated, and (ii) fixed position of segment transformation—upon reaching sexual maturity many nereidids and some syllids transform their whole body into a pelagic sexual form (an ‘epitoke’). The parapodia of middle and posterior segments become much more lobate and flattened compared with parapodia of the relatively unmodified anterior segments. The segment where this change occurs is more or less constant for each sex of each species.

A mathematical rule base underpinning segmental awareness may explain another ability of phyllodocids—to divide their true segments into two or three smaller units (called merosegments by Minelli (2003, p. 202); but note that this is not unique to Phyllodocida). The regular banding pattern in some syllids, for example, appears not to correspond to true segmental boundaries but rather to these merosegments. A connection between the ability of certain annelids to produce merosegments and the regular alternation pattern of elytrae in scale worms has also been suggested (Minelli 2003, p. 203), but this connection seems tenuous, because the pattern of elytrae (= DC alternation pattern; table 2) does not match the rather simple merosegmentation pattern, either two or three merosegments per true segment. Possibly, merosegmentation may be responsible for the binary alternation pattern in the DC of posterior segments, but it does not appear to be a good general explanation.

Segment identity and control of post-embryonic segmentation is regulated by homeotic (=Hox) genes, which in turn control the activity of possibly hundreds of other target genes (Irvine & Martindale 2001; Shimizu & Nakamoto 2001; Martínez & Amemiya 2002). To date, the Hox genes of very few polychaete species have been studied, but two studies have included species of Phyllodocida. The Hox clusters of the nereidids Neanthes virens and Platynereis dumerilii have 11 and 9 genes, respectively, with each being activated at different stages of early development. The patterns of expression were similar in the homologous genes of both species leading the investigators to conclude that the genes were involved in the anterior–posterior regionalization of the larval body (Andreeva et al. 2001; Kulakova et al. 2007). Despite this promising finding, the involvement of Hox genes in the post-larval segmentation and regionalization of the body is still not clear (Dorresteijn 2005). If Hox genes are shown to be active during post-larval segmental growth period, as seems likely, then it would be reasonable to suggest that Hox gene regulation may involve a switching sequence that corresponds to a substitution rule such as the Fibonacci string sequence or a cellular automaton.

REFERENCES


