



# Degenerate Tetraploidy Was Established Before Bdelloid Rotifer Families Diverged

## Citation

Hur, Jae H., Karine Van Doninck, Morgan L. Mandigo, and Matthew Meselson. 2009. Degenerate tetraploidy was established before bdelloid rotifer families diverged. *Molecular Biology and Evolution* 26 (2):375-383.

## Published Version

<http://dx.doi.org/10.1093/molbev/msn260>

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:3322470>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# Degenerate Tetraploidy Was Established Before Bdelloid Rotifer Families Diverged

Jae H. Hur,\* Karine Van Doninck,\*† Morgan L. Mandigo,\*‡ and Matthew Meselson\*

\*Department of Molecular and Cellular Biology, Harvard University; †Department of Biology, University of Namur (FUNDP), URBO, Namur, Belgium; and ‡Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD

Rotifers of Class Bdelloidea are abundant freshwater invertebrates known for their remarkable ability to survive desiccation and their lack of males and meiosis. Sequencing and annotation of approximately 50-kb regions containing the four *hsp82* heat shock genes of the bdelloid *Philodina roseola*, each located on a separate chromosome, have suggested that its genome is that of a degenerate tetraploid. In order to determine whether a similar structure exists in a bdelloid distantly related to *P. roseola* and if degenerate tetraploidy was established before the two species separated, we sequenced regions containing the *hsp82* genes of a bdelloid belonging to a different family, *Adineta vaga*, and the histone gene clusters of *P. roseola* and *A. vaga*. Our findings are entirely consistent with degenerate tetraploidy and show that it was established before the two bdelloid families diverged and therefore probably before the bdelloid radiation.

## Introduction

Rotifers of Class Bdelloidea are common freshwater invertebrates thought to reproduce solely by asexual parthenogenesis (Hsu 1956a, 1956b; Normark et al. 2003) and are also unusual in being able to survive and reproduce after desiccation at any life stage (Ricci 1987) and after exposure to extremely high doses of ionizing radiation (Gladyshev and Meselson 2008). The class comprises some 460 described species (Segers 2007) classified among four families (Adinetidae, Habrotrochidae, Philodinidae, and Philodinidae). The high synonymous nucleotide divergence ( $K_s$ ) between orthologous genes in bdelloids of different families and the presence of bdelloid remains in 35–40-My-old amber indicate that bdelloids arose tens of millions of years ago (Waggoner and Poinar 1993; Mark Welch and Meselson 2000).

Sequencing of four 45–70-kb regions, each containing one of the four copies of the *hsp82* heat shock gene in the bdelloid *Philodina roseola* (family Philodinidae), has indicated that *P. roseola* is a degenerate tetraploid (Mark Welch et al. 2008). The four regions constitute a quartet comprised of two colinear gene-rich pairs with only a minority of genes common to both pairs, in the same order and orientation and separated by gene-rich segments present in only one or the other pair. The average nucleotide difference between aligned members of a colinear pair is 4% and the non-weighted average  $K_s$  between genes within a pair is 0.06 (range 0–0.15). There are tracts of sequence identity or near identity up to several kilobases in length between members of a pair, as could result from recent conversion. There are no conspicuous tracts of identity between members of different colinear pairs, and they can be aligned only within the genes they share, for which the nonweighted average  $K_s$  is 0.93 (range 0.67–1.29).

Consistent with the above evidence that *P. roseola* is a degenerate tetraploid, its karyotype, ignoring dot chromosomes, consists of 11 equal-size chromosomes and one

chromosome of about twice their length that appears to be an isochromosome (Mark Welch et al. 2004). Taking its two arms as a colinear pair, the *P. roseola* genome may therefore consist of three quartets, each made up of two colinear pairs. The karyotype of the distantly related bdelloid *Adineta vaga* (family Adinetidae) consists of 12 equal-size chromosomes, suggesting that it too may be a degenerate tetraploid (Mark Welch and Meselson 1998).

Here, we ask if degenerate tetraploidy is a general characteristic of bdelloid rotifers and, if so, whether it was established before the separation of bdelloid families or independently in different bdelloid families or taxa of lower rank. To address this question, we sequenced 38–67-kb regions containing the *hsp82* genes of *A. vaga* and 25–39-kb regions containing the histone gene clusters of *A. vaga* and *P. roseola*. Our results are entirely consistent with degenerate tetraploidy. Moreover, the order and orientation of genes in sequenced regions are largely conserved between species, indicating that degenerate tetraploidy was established before the two families separated. Our findings are consistent with the occurrence of a whole genome duplication followed by numerous segmental deletions before the separation of the two bdelloid families but after the separation of Bdelloidea from rotifers of their sister Class, the Monogononta (Melone et al. 1998; Mark Welch 2000), which are facultatively sexual diploids (Robotti 1975; King and Snell 1977; Mark Welch and Meselson 2000).

## Materials and Methods

### Library Construction and Screening

*Adineta vaga* was cultured and DNA extracted from purified eggs as previously described except CsCl density-gradient purification was replaced by phenol:chloroform extraction (Mark Welch et al. 2004; Gladyshev and Meselson 2008). The fosmid library was constructed with an Epicentre CopyControl Fosmid Library Production Kit (Epicentre, Madison, WI), giving a primary fosmid library in *Escherichia coli*. The library was arranged onto GeneScreen Plus positively charged nylon membranes (PerkinElmer, Waltham, MA) using a Genetix Q-Bot (Genetix, New Milton, Hampshire, United Kingdom). <sup>32</sup>P-labeled probes were made by nick translation of

Key words: *hsp82*, histone gene cluster, *Adineta vaga*, *Philodina roseola*.

E-mail: msm@wjh.harvard.edu.

*Mol. Biol. Evol.* 26(2):375–383. 2009

doi:10.1093/molbev/msn260

Advance Access publication November 7, 2008

732-bp polymerase chain reaction (PCR) amplicons (primers 5' GTGAATTGATTTCAAATGCATC 3' and 5' CATCTTTTTTCTTGTGCATCATC 3') of the previously cloned 927-bp region of *A. vaga* *hsp82* (Mark Welch and Meselson 2000). Fosmids that hybridized with the probe were screened by PCR with the above primers, and the resulting amplicons were purified and sequenced. ClustalW (Higgins et al. 1994; Larkin et al. 2007) alignments and Blast searches (Altschul et al. 1990) were used to determine which copy of *hsp82* was present in each fosmid. As a further screen, *hsp82* regions were sequenced directly from fosmids.

### Fosmid Sequencing and Contig Construction

Fosmid DNA was prepared with an Eppendorf Perceptprep Plasmid Midi kit (Eppendorf, Westbury, NY), and approximately 10  $\mu$ g of each fosmid was sheared to 3–4-kb fragments (Genemachines Hydroshear, Genomic Solutions, Ann Arbor, MI), subcloned into pCR 4-TOPO vectors (TOPO TA Cloning Kit for Sequencing, Invitrogen, Carlsbad, CA), and sequenced as described (Mark Welch et al. 2008). Approximately 800 bp at both ends of approximately 300 plasmid subclones from each fosmid were sequenced (ca., 9-fold coverage of a 50-kb-fosmid), and remaining gaps were closed by direct sequencing from fosmid templates. Plasmid sequences and any additional gap-spanning sequences from each fosmid were assembled into contigs using STRAW (Josephine Bay Paul Center Computational Facility, Marine Biological Laboratories, Woods Hole, MA), a script that combines PHRED, PHRAP, and CROSS\_MATCH (Ewing and Green 1998; Ewing et al. 1998) to call bases from sequence chromatograms, align overlapping sequences, and organize them into contigs. There was complete identity within the approximately 8-kb overlap of the two sequenced fosmids containing *hsp82-1*, and these were combined into one 67-kb-contig designated Avhsp-1. All other contigs presented in this study were single fosmids.

### Southern Blot

The 732-bp regions of *hsp82-1* and *hsp82-3* used to screen the library were also used to probe a Southern blot of *A. vaga* genomic DNA (gDNA) (Sambrook et al. 1989). Probes were separately labeled and hybridization to a lane containing equal amounts of separate NsiI digests of three fosmids containing *hsp82-1*, *hsp82-3*, and *hsp82-4* showed them to have equal specific activity (fig. 1A, lane 2). The intensities of the bands in each lane were read with a Storm 860 Imager (GE Healthcare, Piscataway, NJ), measured with IMAGEJ (Abramoff et al. 2004), and corrected for background.

### Fluorescent In Situ Hybridization (FISH)

Published protocols (Mark Welch et al. 2004) were followed using red (Alexa 568-dUTP [deoxyuridine triphosphate]) or green (Alexa 488-dUTP) fluorophores (Invitrogen).

### Annotation

Putative genes were identified by BlastX searches of the GenBank nonredundant database (Altschul et al. 1990; Benson et al. 2005); by de novo prediction (GeneMark.hmm ES-3.0, self-trained on *Caenorhabditis elegans*) (Lomsadze et al. 2005); and by TblastX searches for conserved regions between contigs within and between species. Candidate genes were then screened against the Pfam (Bateman et al. 2004) and NCBI Conserved Domain (Marchler-Bauer et al. 2005) databases for homology to known protein domains. Intron–exon boundaries were identified by comparison with homologous amino acid sequences using GENEWISE2 (mode—local, organism—worm, no intronic bias, splice site-GT/AG only, model-flat, and algorithm—GeneWise 623) (Birney et al. 2004) and, in a few cases, were identified manually as regions with GT/AG boundaries containing stop codons and joining open reading frames. The results for *P. roseola* *hsp82* contigs differed slightly from the previous annotation in the number of putative genes and the assignment of gene boundaries (Mark Welch et al. 2008).

### Divergence Calculations

Contigs were aligned with ClustalW, and total divergence was measured with DNASP 4.0 (Rozas et al. 2003). Genes were aligned based on translation with REVTRANS 1.4 (Wernersson and Pedersen 2003), and synonymous ( $K_s$ ) and nonsynonymous ( $K_a$ ) divergence was measured with DIVERGE (GCG Wisconsin Package, Accelrys, San Diego, CA), which corrects for multiple hits and transversion–transition bias (Li et al. 1985; Li 1993; Pamilo and Bianchi 1993). Where the ratio of transversions to transitions was  $\geq 1$ , adjustments were made only for multiple hits (Jukes and Cantor 1969) using DNASP 4.0. No corrections were made in the two cases in which raw divergence was  $\geq 0.75$ .

## Results

### *hsp82* Regions in *A. vaga*

We constructed a primary genomic fosmid library of *A. vaga* and probed it with a mixture of approximately 700-bp segments from the same region of each of the three different previously described copies of *hsp82* that had been found in this species (Mark Welch and Meselson 2000). Of the approximately 110,000 fosmids screened, 102 exhibited hybridization signals. These were screened by PCR and direct sequencing with *hsp82* primers, yielding 67 that contained an *hsp82* probe sequence. The *hsp82* genes represented in these fosmids were of three kinds, designated *hsp82-1* (35 fosmids), *hsp82-3* (17 fosmids), and *hsp82-4* (15 fosmids), with sequences identical to the three copies found previously and designated copies 1, 2, and 3, respectively (Mark Welch and Meselson 2000). No other *hsp82* sequences were found. Two fosmids containing *hsp82-1* and one each containing *hsp82-3* and *hsp82-4* were fully sequenced, giving contigs Avhsp-1 (EU637017), Avhsp-3 (EU637018), and Avhsp-4 (EU831279), respectively.

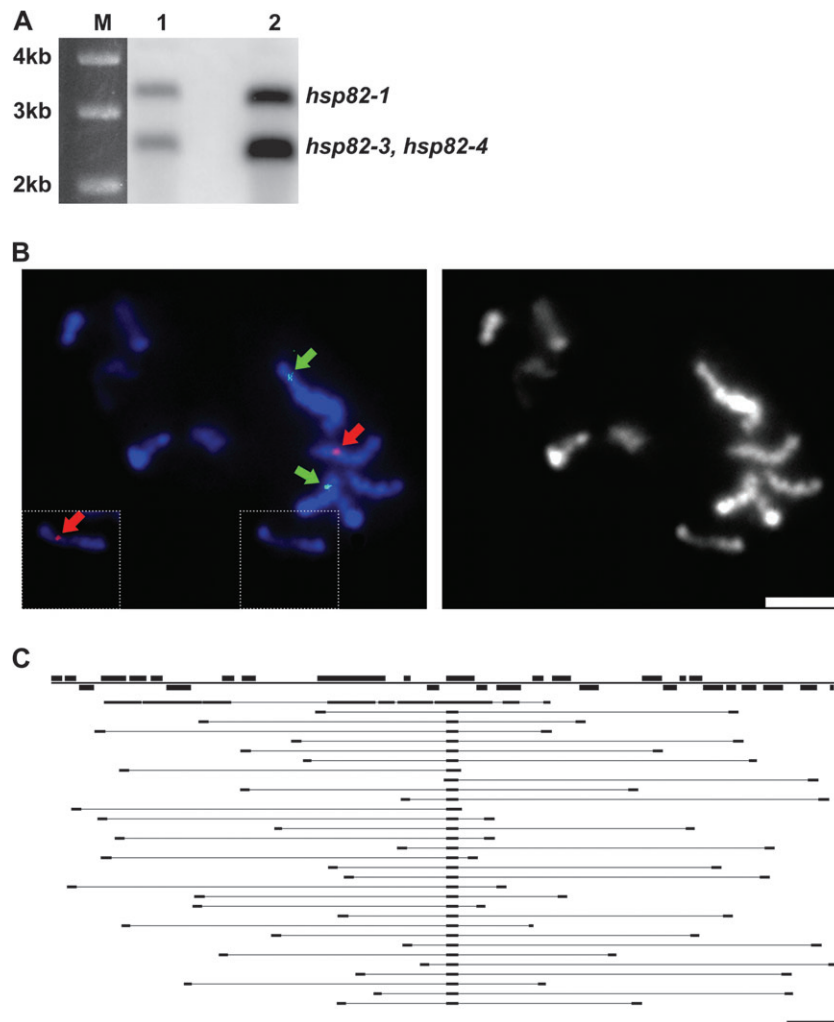


FIG. 1.—*Adineta vaga* contains four *hsp82*-containing regions, two of which are identical, on four separate chromosomes. (A) Southern blots of *Nsi*I-digested *A. vaga* gDNA hybridized to a mixture of *hsp82-1* and *hsp82-3* probes of equal specific activity. The *hsp82-3* probe diverges by 1.6% from the corresponding region of *hsp82-4* and hybridizes equally with *hsp82-3* and *hsp82-4* (data not shown). The 3.5-kb fragment derives from *hsp82-1*, and the 2.6-kb fragment derives from *hsp82-3* and *hsp82-4*. Lane M, size marker. Lane 1, 7- $\mu$ g *Nsi*I-digested gDNA; the ratio of corrected intensity of the 2.5-kb band to that of the 3.5-kb band is 0.92. Lane 2, 5 ng each of digested fosmids containing *hsp82-1*, *hsp82-3*, and *hsp82-4* mixed with 7  $\mu$ g of *Nsi*I-digested gDNA; the ratio of background-corrected intensity of the 2.5-kb band to that of the 3.5-kb band is 0.52. (B) Left, *A. vaga* chromosomes hybridized to a mixture of an *hsp82-1* containing fosmid labeled with green fluorophore and *hsp82-3* and *hsp82-4* containing fosmids labeled with a red fluorophore. Arrows point to hybridization signals. One of the two red signals is photographed in a different focal plane (inset). Right, the same squash stained with 4',6-diamidino-2-phenylindole. Scale bar, 5  $\mu$ m. (C) Alignment of the termini of 32 fosmids containing *hsp82-1* on Avhsp-1. Each line below Avhsp-1 represents a fosmid that contains *hsp82-1*. Thickened lines represent the sequenced regions at the termini and in the *hsp82* gene of each fosmid insert and more extensive regions sequenced in one of the fosmids, all of which were found to be identical to the corresponding sequence in Avhsp-1. Scale bar, 5 kb.

The finding of approximately twice as many fosmids containing *hsp82-1* as those containing *hsp82-3* or *hsp82-4* suggested that there are two copies of the *hsp82-1* gene in the *A. vaga* genome that are identical within the sequenced region. This possibility was tested by Southern blotting and by FISH.

When probed with a mixture of the  $^{32}$ P-labeled *hsp82-1* and *hsp82-3* PCR products used to screen the fosmid library, an *Nsi*I digest of *A. vaga* gDNA revealed two bands of equal intensity, corresponding to the expected 3.5-kb fragment from *hsp82-1* and the expected 2.6-kb fragment from both *hsp82-3* and *hsp82-4* (fig. 1A, lane 1). That the specific activities and hybridization efficiencies of the two probes were the same was verified by the observation

that the restriction fragment from *hsp82-3* and *hsp82-4* was labeled with twice the intensity as that from *hsp82-1* when an equal mixture of *Nsi*I digests of the three fosmids, each containing a different copy of *hsp82*, was simultaneously hybridized on the same membrane with both probes (fig. 1A, lane 2). The observation that the 2.6- and 3.5-kb bands are of equal intensity therefore shows that the *A. vaga* genome contains twice as many copies of *hsp82-1* as of *hsp82-3* or *hsp82-4*.

The chromosomal distribution of the various *hsp82*-containing regions was visualized by hybridizing squash preparations of *A. vaga* embryos with one of the *hsp82-1* containing fosmids labeled with a green fluorophore together with a fosmid containing *hsp82-3* and a fosmid

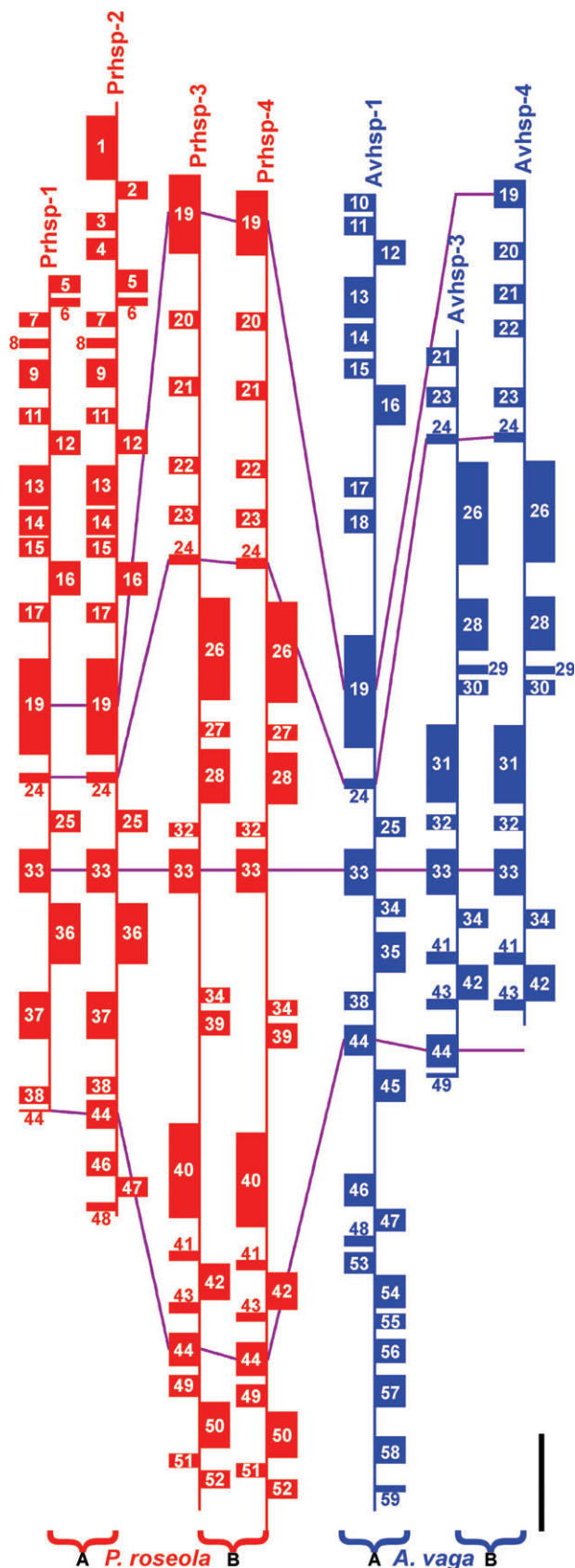


FIG. 2.—Distribution of genes in *hsp82* regions of *Adineta vaga* and *Philodina roseola*. Contigs are shown as lines with putative genes shown as boxes numbered as in table 1. Boxes on opposite sides of a line

containing *hsp82-4*, both labeled with a red fluorophore. As shown in figure 1B, there is a green signal on each of two chromosomes and a red signal on each of two others, consistent with the presence of two copies of Avhsp-1 and one each of Avhsp-3 and Avhsp-4, all on separate chromosomes.

Seeking to determine how far the region of identity extends beyond the *hsp82* genes in the two regions containing *hsp82-1*, we sequenced the terminal approximately 500 bp of the inserts in each of 32 fosmid containing the *hsp82-1* sequence from the library screen. These 64 end sequences fall randomly along Avhsp-1 and should be equally likely to derive from one or the other of the two regions (fig. 1C). In no case was any difference found between a fosmid insert end sequence and the corresponding sequence in Avhsp-1. It may be concluded that the two regions containing *hsp82-1* have sequences that are identical or nearly identical over a region of at least 67 kb containing the gene.

Contigs Avhsp-3 and Avhsp-4 differ by 3% in the approximately 32-kb region of their overlap (gene 23 through gene 43) and contain the same genes in the same order and orientation, with nonweighted average  $K_s = 0.04$  (range 0–0.15) (fig. 2, table 1). The only exception involves a region of Avhsp-4 containing four tandem copies of a putative 7 transmembrane receptor (genes 20–23). An approximately 2.5-kb segment of Avhsp-4 containing one of these copies is not present in Avhsp-3. In contrast to the colinearity of Avhsp-3 and Avhsp-4 and the apparent identity of the two regions containing *hsp82-1*, only 5 of the 19 different putative genes present in Avhsp-3 or Avhsp-4 are also present in Avhsp-1. These five genes are in the same order and orientation, separated by regions up to 12 kb in length that are present in only one or the other of the two colinear pairs. The nonweighted average  $K_s$  between genes in different colinear pairs is 0.63 (range 0.32–1.16). Thus, the organization of the *hsp82* regions of *A. vaga*, like the *hsp82* regions of *P. roseola*, is that of a degenerate tetraploid.

#### Histone Gene Cluster Regions in *A. vaga* and *P. roseola*

Further evidence for bdelloid degenerate tetraploidy is seen in the regions surrounding the clusters of canonical H4, H3, and H2B histone genes in *A. vaga* and *P. roseola*. The genomes of both bdelloid species contain four such clusters comprising two colinear pairs with nonweighted average  $K_s = 0.004$  (range 0–0.01) between corresponding histone genes within a colinear pair in *A. vaga* and  $K_s = 0.04$  (range 0.02–0.06) in *P. roseola*. The nonweighted average  $K_s$  between corresponding histone genes in different colinear pairs is 0.38 (range 0.24–0.62) in *A. vaga* and 0.51 (range 0.43–0.59) in *P. roseola* (table 2). In order to investigate larger regions containing these histone gene clusters,

← represent genes transcribed in opposite directions. Genes common to all colinear pairs are connected by purple lines. Colinear pairs of one species that are similar to those in the other are designated by letters A and B. *Philodina roseola*, red; *A. vaga*, blue. Scale bar, 5 kb.

**Table 1**  
**Putative Genes Identified in *hsp82* Regions**

| No. | Gene Designation                 | Avhsp-1 | Prhsp-1/2 | Avhsp-3/4 | Prhsp-3/4 | Avhsp-1,<br>Avhsp-3/4 | Prhsp-1/2,<br>Prhsp-3/4 |
|-----|----------------------------------|---------|-----------|-----------|-----------|-----------------------|-------------------------|
| 1   | A/G cyclase                      |         | ✓         |           |           |                       |                         |
| 2   | Hypothetical 1                   |         | ✓         |           |           |                       |                         |
| 3   | Ubiquitin assoc.                 |         | ✓         |           |           |                       |                         |
| 4   | Cbl-interacting prot. Sts-1-like |         | ✓         |           |           |                       |                         |
| 5   | TAR                              |         | 0.00      |           |           |                       |                         |
| 6   | Zinc finger prot.                |         | 0.00      |           |           |                       |                         |
| 7   | Hypothetical 2                   |         | 0.00      |           |           |                       |                         |
| 8   | Minor capsid                     |         | 0.00      |           |           |                       |                         |
| 9   | Transcriptional regulator        |         | 0.03      |           |           |                       |                         |
| 10  | Glycoprotein                     | ✓       |           |           |           |                       |                         |
| 11  | Hypothetical 3                   | ✓       | 0.00      |           |           |                       |                         |
| 12  | Hypothetical 4                   | ✓       | 0.00      |           |           |                       |                         |
| 13  | DUF288                           | ✓       | 0.09      |           |           |                       |                         |
| 14  | Hypothetical 5                   | ✓       | 0.15      |           |           |                       |                         |
| 15  | Methionine sulfoxide reductase   | ✓       | 0.13      |           |           |                       |                         |
| 16  | Strabismus                       | ✓       | 0.05      |           |           |                       |                         |
| 17  | s/t/y kinase                     | ✓       | 0.11      |           |           |                       |                         |
| 18  | Integral membrane prot.*         | ✓       |           |           |           |                       |                         |
| 19  | pmca                             | ✓       | 0.09      | ✓         | 0.09      | 0.79                  | 1.11                    |
| 20  | 7tmr1                            |         |           | ✓         | 0.11      |                       |                         |
| 21  | 7tmr2                            |         |           | 0.03      | 0.14      |                       |                         |
| 22  | 7tmr3                            |         |           | †         | 0.10      |                       |                         |
| 23  | 7tmr4                            |         |           | 0.15      | 0.08      |                       |                         |
| 24  | Myosin light chain alkali        | ✓       | 0.02      | 0.00      | 0.02      | 0.39                  | 0.68                    |
| 25  | Hypothetical 6                   | ✓       | 0.06      |           |           |                       |                         |
| 26  | Sodefirin precursor repeats A    |         |           | 0.04      | 0.04      |                       |                         |
| 27  | Sodefirin precursor repeats B*   |         |           |           | 0.00      |                       |                         |
| 28  | Sodefirin precursor repeats C    |         |           | 0.04      | 0.08      |                       |                         |
| 29  | Metal binding prot.*             |         |           | 0.03      |           |                       |                         |
| 30  | tRNA methyltransferase*          |         |           | 0.01      |           |                       |                         |
| 31  | Laminin-G*                       |         |           | 0.03      |           |                       |                         |
| 32  | TPR-containing prot. A           |         |           | 0.02      | 0.10      |                       |                         |
| 33  | hsp82                            | ✓       | 0.01      | 0.04      | 0.10      | 0.32                  | 0.70                    |
| 34  | Hypothetical 7                   | ✓       |           | 0.06      | 0.09      | 1.05                  |                         |
| 35  | DUF288rc*                        | ✓       |           |           |           |                       |                         |
| 36  | LRR containing prot.*            |         | 0.07      |           |           |                       |                         |
| 37  | TPR containing prot. B*          |         | 0.10      |           |           |                       |                         |
| 38  | Hypothetical 8                   | ✓       | 0.04      |           |           |                       |                         |
| 39  | Hypothetical 9*                  |         |           |           | 0.06      |                       |                         |
| 40  | 7tmrA*                           |         |           |           | 0.08      |                       |                         |
| 41  | Hypothetical 10                  |         |           | 0.03      | 0.07      |                       |                         |
| 42  | Zinc carboxypeptidase            |         |           | 0.05      | 0.08      |                       |                         |
| 43  | Hypothetical 11                  |         |           | 0.01      | 0.04      |                       |                         |
| 44  | saicars                          | ✓       | 0.12      | ✓         | 0.08      | 0.76                  | [1.22]                  |
| 45  | IBR containing*                  | ✓       |           |           |           |                       |                         |
| 46  | Hypothetical 12                  | ✓       | ✓         |           |           |                       |                         |
| 47  | Hypothetical 13                  | ✓       | ✓         |           |           |                       |                         |
| 48  | Hypothetical 14                  | ✓       | ✓         |           |           |                       |                         |
| 49  | Hypothetical 15                  |         |           | ✓         | 0.06      |                       |                         |
| 50  | Lung 7tmr                        |         |           |           | 0.09      |                       |                         |
| 51  | Hypothetical 16                  |         |           |           | 0.01      |                       |                         |
| 52  | Phosphoglycerate mutase          |         |           |           | 0.10      |                       |                         |
| 53  | NHL repeats                      | ✓       |           |           |           |                       |                         |
| 54  | Ubiquitin ligase                 | ✓       |           |           |           |                       |                         |
| 55  | snf7                             | ✓       |           |           |           |                       |                         |
| 56  | C2-containing prot.              | ✓       |           |           |           |                       |                         |
| 57  | PKD channel                      | ✓       |           |           |           |                       |                         |
| 58  | Synaptotagmin                    | ✓       |           |           |           |                       |                         |
| 59  | Hypothetical 17                  | ✓       |           |           |           |                       |                         |

NOTE.—Average  $K_2$  values corrected for multiple hits and transition/transversion bias. ✓—sequenced in only one contig; †—present in; \*—present in one colinear pair of one species; [ ]—only corrected for multiple hits using Jukes–Cantor corrections.

fosmids covering each cluster of each species were fully sequenced and annotated (Avhis-1, EU652315; Avhis-2, EU850438; Avhis-3, EU652316; Avhis-4, EU850439; Prhis-1, EU850440; Prhis-2, EU652317; Prhis-3,

EU652318; and Prhis-4, EU850441), revealing a pattern of degenerate tetraploidy like that of the *hsp82* regions of these species (fig. 3, table 2). For example, between the most distantly separated genes common to both colinear

**Table 2**  
**Putative Genes Identified in Histone Gene Cluster Regions**

| No. | Gene Designation                              | Avhis-1/2 | Prhis-1/2 | Avhis-3/4 | Prhis-3/4 | Avhis-1/2, Avhis-3/4 | Prhis-1/2, Prhis-3/4 |
|-----|---|-----------|-----------|-----------|-----------|----------------------|----------------------|
| 1   | Hypothetical 18                               | ✓         |           |           |           |                      |                      |
| 2   | Hypothetical 19                               | ✓         |           |           |           |                      |                      |
| 3   | Guanosine triphosphate binding                |           | ✓         |           |           |                      |                      |
| 4   | Coatomer                                      | ✓         | ✓         |           |           |                      |                      |
| 5   | G protein coupled receptor                    | ✓         | ✓         |           |           |                      |                      |
| 6   | Beta-transducin repeat                        | ✓         | ✓         |           |           |                      |                      |
| 7   | Speckle-type Pox virus and zinc finger domain | 0.03      | ✓         |           |           |                      |                      |
| 8   | Adenosine kinase                              | 0.02      | ✓         |           |           |                      |                      |
| 9   | Prefoldin subunit                             | 0.01      | ✓         |           |           |                      |                      |
| 10  | Dihydrogenpyrimidine dehydrogenase            |           |           | 0.01      | ✓         |                      |                      |
| 11  | Translation init. factor eIF-4E               | 0.03      | ✓         | 0.02      | ✓         | 0.85                 | 1.07                 |
| 12  | Rho GTPase activating prot.                   | 0.03      | 0.13      | 0.03      | ✓         | 1.18                 | 1.45                 |
| 13  | 7tm prot. 66                                  |           |           | 0.03      | 0.10      |                      |                      |
| 14  | Histone H4                                    | 0.00      | 0.03      | 0.00      | 0.03      | 0.24                 | 0.48                 |
| 15  | Histone H3                                    | 0.01      | 0.04      | 0.00      | 0.03      | 0.29                 | 0.58                 |
| 16  | Histone H2B                                   | 0.00      | 0.02      | 0.01      | 0.06      | 0.61                 | 0.48                 |
| 17  | Histone H2Av                                  | 0.02      | 0.03      |           |           |                      |                      |
| 18  | Osteopetrosis assoc. tm prot.                 |           |           | 0.02      | 0.17      |                      |                      |
| 19  | Cleft lip and palate assoc. tm prot.          |           | 0.10      | 0.03      | 0.09      |                      | 0.75                 |
| 20  | Mito. ribosomal prot. s22                     | 0.03      | 0.06      |           |           |                      |                      |
| 21  | Ribosomal prot. s15a                          | ✓         | 0.02      | 0.02      | 0.03      | 0.54                 | 0.93                 |
| 22  | Hypothetical 20                               | ✓         | 0.06      | 0.02      | 0.07      | 1.37                 | [1.76]               |
| 23  | START domain-containing prot.                 | ✓         | 0.07      | ✓         | 0.06      | 0.59                 | 1.05                 |
| 24  | Arginine vasopressin receptor*                | ✓         |           |           |           |                      |                      |
| 25  | Filament prot.                                | ✓         | 0.06      |           | 0.08      |                      | 1.55                 |
| 26  | Nucleolar complex assoc. prot.                | ✓         | ✓         |           |           |                      |                      |
| 27  | Hypothetical 21                               |           | ✓         |           | 0.08      |                      | 1.19                 |
| 28  | Putative G protein coupled receptor           |           | ✓         |           | 0.18      |                      | 1.31                 |
| 29  | Hypothetical 22                               |           | ✓         |           | 0.10      |                      | 1.21                 |

NOTE.—Corrections for multiple hits and symbols as described in table 1.

pairs in *P. roseola* (genes 11 and 25), there are eight genes present in both pairs, in the same order and orientation, and four genes in only one or the other pair. As in the *hsp82* regions of *A. vaga* and *P. roseola*, colinear regions containing the histone gene clusters are closely similar, differing by 2% in *A. vaga* and 6% in *P. roseola* and contigs belonging to different colinear pairs can be aligned only within the genes they share. The nonweighted average  $K_s$  for genes in members of a colinear pair is 0.02 (range 0–0.03) in *A. vaga* and 0.07 (range 0.02–0.18) in *P. roseola*, whereas the nonweighted average  $K_s$  for genes in different colinear pairs is 0.68 (range 0.24–1.46) in *A. vaga* and 1.02 (range 0.43–1.87) in *P. roseola* (tables 1 and 2).

#### Introns and Transposable Elements

The *hsp82* and histone gene cluster contigs in both species are gene rich with exons constituting approximately 50% of the sequenced 800 kb. The lack of transposable genetic elements and the abundance of small introns in these regions are consistent with findings reported for *P. roseola* *hsp82* regions and other gene-rich regions of bdelloid genomes (Arkhipova and Meselson 2000, 2005; Mark Welch and Meselson 2001; Mark Welch et al. 2008; Gladyshev, Meselson, and Arkhipova 2008). There are introns even in the canonical histone genes contrary to the general lack of introns in histone genes of other species.

#### Discussion

As has been found in the *hsp82* region of *P. roseola*, we find that this region in the distantly related bdelloid species *A. vaga* and the histone gene cluster regions of *A. vaga* and *P. roseola* each comprise two colinear pairs with only a minority of genes common to both pairs, in the same order and orientation. Thus, our findings are entirely consistent with degenerate tetraploidy in both species.

Within regions in which both colinear pairs can be compared in both *A. vaga* and *P. roseola* (*hsp82* region genes 19–44 and histone gene cluster region genes 11–23), we find that 30 of 39 genes are present in both species (figs. 2 and 3, tables 1 and 2). Further, each colinear pair in one species resembles a colinear pair in the other species, as seen in the sequence of genes and the pattern of presence and absence of gene-containing segments. Moreover, there is no case in which a gene present in only one colinear pair in a given species is found only in the dissimilar pair in the other species. This interspecies conservation almost certainly means that degenerate tetraploidy was established before the separation of *A. vaga* and *P. roseola*. Although it remains to be seen if the same pattern is characteristic of the two other bdelloid families, it appears likely that degenerate tetraploidy was established before the bdelloid radiation and that, consistent with other evidence, Class Bdelloidea is monophyletic (Melone et al. 1998; Mark Welch 2000).

In sexually reproducing species, the colinearity of homologous chromosomes is maintained by segregation and

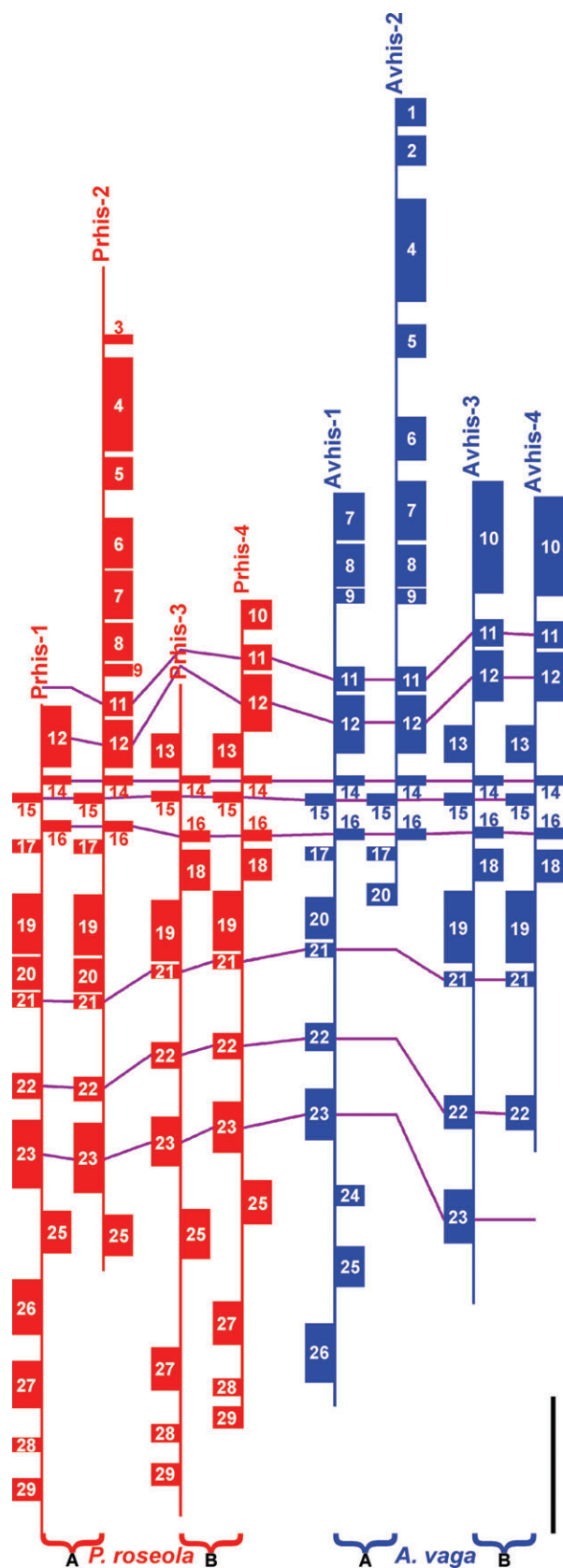


FIG. 3.—Distribution of genes in histone gene cluster regions of *Adineta vaga* and *Philodina roseola*. Contigs are shown as lines with putative genes shown as boxes numbered as in table 2. Other designations as in figure 1. Scale bar, 5 kb.

haplotype drift. The presence of colinear chromosome pairs in bdelloids, however, may have an altogether different explanation, being required in G1 nuclei for faithful template-dependent repair of DNA double strand breaks (DSBs) as may be particularly frequent as a consequence of the anhydrobiotic bdelloid life style (Gladyshev and Meselson 2008). As in the *hsp82* regions of *P. roseola* (Mark Welch et al. 2008), there are numerous tracts of sequence identity between members of a colinear pair, possibly resulting from gene conversion associated with DSB repair. The longest such tract entirely within sequenced regions is a 1.3 kb tract between Avhsp-3 and Avhsp-4. Further, the entire 67 kb covered by Avhsp-1, which is present as two identical or nearly identical regions, may be an unusually long conversion tract, as could result from break-induced replication, or, if the identity extends to one or both ends of the colinear chromosomes involved, may be the result of mitotic crossing-over in G2 or chromosome loss and reduplication, respectively. The possibility is not excluded, however, that some of the regions of identity we see are the result of some form of homologous DNA exchange between closely related individuals.

In contrast to the tracts of identity or near identity seen within colinear pairs, the longest stretch of identity between contigs belonging to different colinear pairs is only 99 bp, found within the *hsp82*-coding region of Avhsp-1 and Avhsp-4. Considering the possibility that conversion tracts are associated with DSB repair, it is interesting that even ignoring the two apparently identical *hsp82* regions in *A. vaga*, the average  $K_s$  as well as the total divergence within colinear pairs in *A. vaga* is consistently lower than the corresponding values for *P. roseola* (tables 1 and 2), possibly reflecting a higher frequency of DSB repair in the more ephemeral aquatic habitats characteristic of *A. vaga* (Ricci 1998).

The apparent lack of conversion between different colinear pairs also suggests that, all else being equal and assuming no homologous exchange between individuals,  $K_s$  between corresponding genes in different colinear pairs in one species would be the same as that in the other species. Instead, the nonweighted average  $K_s$  in the 12 genes present in both colinear pairs in both species (tables 1 and 2) is 0.96 (range 0.48–1.76) in *P. roseola* and 0.66 (range 0.24–1.37) in *A. vaga* and is greater in the former species in 11 of these 12 genes (Spearman's rank order correlation;  $P = 0.001$ ). The observed difference, if it is a general characteristic of the two species, could reflect conversion or mitotic crossing-over between colinear pairs early in the ancestry of *A. vaga* or a lower average mutation rate in *A. vaga*, as might result if it has spent more time than *P. roseola* in a desiccated state (Ricci 1998).

There are two genes present in both colinear pairs of one species that, although surrounded by genes that are present in both colinear pairs, are present in only one pair of the other. Gene 34 in the *hsp82* region is present in all *A. vaga* *hsp82* contigs but absent in Prhsp-1 and Prhsp-2 (fig. 2). Gene 19 in the histone gene cluster regions is present in all *P. roseola* histone gene cluster contigs but missing in Avhis-1 and Avhis-2 (fig. 3). The presence of these genes in only one colinear pair in only one species suggests that although most deletions occurred before the two species separated, a few have occurred subsequently.



There are 12 putative genes, indicated by asterisks in tables 1 and 2, which were found in only one species and in only one colinear pair. These genes, if initially present in the sequenced regions, may have been lost from all but one colinear pair in one species or may have been added as a result of duplication, translocation, or, perhaps, horizontal transfer (Gladyshev, Meselson, and Arkhipova 2008). In the only case we tested, probing the *A. vaga* genomic library with gene 32 from the *A. vaga hsp82* region failed to reveal any additional copies of the gene, indicating that the copy in Avhsp-3 and Avhsp-4 is not a result of duplication with retention of the parent copy elsewhere. That duplications do occur, however, is seen in Avhsp-1, where two copies of the same gene (13 and 35) are present in convergent orientation, separated by 31 kb. The substantial  $K_s$  value between the copies ( $K_s = 0.35$ ) indicates that the duplication occurred long ago, whereas the even greater  $K_s$  value between them and the corresponding genes in *P. roseola* ( $K_s = 1.22$ ) suggests that it may have occurred after the two bdelloid species separated. The lower nonsynonymous divergence ( $K_a$ ) value relative to the  $K_s$  value between the two copies of the gene in *A. vaga* ( $K_a/K_s = 0.56$ ) and also between these and the copies in *P. roseola* ( $K_a/K_s = 0.21$ – $0.30$ ) suggests these duplicated genes have been under purifying selection in both species.

Our results may be explained most simply as the consequence of a whole genome duplication in a progenitor of modern bdelloids, followed by segmental deletions from one or the other colinear pair, giving rise to a degenerate tetraploid in which most genes are present in only one or the other of the two colinear pairs. The Adineta–Philodina separation then occurred after most such deletions had taken place. A different scenario in which extensive segmental deletion and divergence between homologous chromosomes in an ancient asexual diploid ancestor was followed by whole genome duplication is less likely because, until the duplication, bdelloid progenitors would lack homologous templates for DSB repair in G1 detrimental condition that would be especially severe if bdelloids had already adapted to anhydrobiosis and the necessity of frequent DSB repair.

### Supplementary Material

Supplementary table S1 (Divergence values for *hsp82* region genes within and between species.  $K_s$  values corrected for multiple hits and transition/transversion bias: [ ]—corrected only for multiple hits, { }—uncorrected) and table S2 (Divergence values for histone gene cluster region genes within and between species. Corrected for multiple hits as indicated in table S1) are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

### Acknowledgments

We thank David and Jessica Mark Welch for helpful discussions and technical assistance at every stage of the experiments. Other members of the Josephine Bay Paul Center for Comparative Molecular Biology and Evolution

at the Marine Biological Laboratory, Woods Hole, MA provided assistance in fosmid sequencing and contig construction. At Harvard, Beth Weiner and others in the Kleckner laboratory provided access to their fluorescent microscope and instruction in its use, Christian Daly, George Kenty, and Peter Wang provided assistance in library construction, and members of the Maniatis laboratory and Kevin McMullen provided assistance in library screening. Financial support was provided by the US National Science Foundation (M.M.), the Belgian–American Educational Foundation (K.V.D.), and the Fulbright Program (K.V.D.).

### Literature Cited

- Abramoff MD, Magelhaes PJ, Ram SJ. 2004. Image processing with ImageJ. *Biophoton International*. 11:36–42.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol*. 215:403–410.
- Arkhipova I, Meselson M. 2000. Transposable elements in sexual and ancient asexual taxa. *Proc Natl Acad Sci USA*. 97:14473–14477.
- Arkhipova I, Meselson M. 2005. Diverse DNA transposons in rotifers of the class Bdelloidea. *Proc Natl Acad Sci USA*. 102:11781–11786.
- Bateman A, Coin L, Durbin R, Finn RD, et al. (13 co-authors). 2004. The Pfam protein families database. *Nucleic Acids Res*. 32:D138–D141.
- Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. 2005. Genbank. *Nucleic Acids Res*. 33:D34–D38.
- Birney E, Clamp M, Durbin R. 2004. GeneWise and Genomewise. *Genome Res*. 14:988–995.
- Ewing B, Green P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res*. 8:186–194.
- Ewing B, Hillier L, Wendl M, Green P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res*. 8:175–185.
- Gladyshev E, Meselson M. 2008. Extreme resistance of bdelloid rotifers to ionizing radiation. *Proc Natl Acad Sci USA*. 105:5139–5144.
- Gladyshev E, Meselson M, Arkhipova I. 2008. Massive horizontal gene transfer in bdelloid rotifers. *Science*. 320:1210–1213.
- Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 22:4673–4680.
- Hsu WS. 1956a. Oogenesis in the bdelloidea rotifer *Philodina roseola* (Ehrenberg). *La Cellule*. 5:283–296.
- Hsu WS. 1956b. Oogenesis in *Habrotrocha tridens* (Milne). *Biol Bull*. 111:364–374.
- Jukes TH, Cantor CR. 1969. Evolution of protein molecules. Munro HN, editor. In: *Mammalian protein metabolism vol III*. New York: Academic Press. p. 21–132.
- King CE, Snell TW. 1977. Sexual recombination in rotifers. *Heredity*. 39:357–360.
- Larkin MA, Blackshields G, Brown NP, et al. (13 co-authors). 2007. ClustalW and ClustalX version 2. *Bioinformatics*. 23:2947–2948.
- Li WH. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J Mol Evol*. 36:96–99.
- Li WH, Wu CL, Luo CC. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol Biol Evol*. 2:150–174.

- Lomsadze A, Ter-Hovhannisyanyan V, Chernoff Y, Borodovsky M. 2005. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.* 33:6494–6506.
- Marchler-Bauer A, Anderson JB, Cherukuri PF, et al. (24 co-authors). 2005. CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res.* 33:D192–D196.
- Mark Welch DB. 2000. Evidence from a protein-coding gene that acanthocephalans are rotifers. *Invert Biol.* 119:17–26.
- Mark Welch DB, Mark Welch JL, Meselson M. 2008. Evidence for degenerate tetraploidy in bdelloid rotifers. *Proc Natl Acad Sci USA.* 105:5145–5149.
- Mark Welch DB, Meselson M. 2000. Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science.* 288:1211–1215.
- Mark Welch DB, Meselson M. 2001. A survey of introns in three genes of rotifers. *Hydrobiologia.* 446/447:333–336.
- Mark Welch JL, Mark Welch DB, Meselson M. 2004. Cytogenetic evidence for asexual evolution of bdelloid rotifers. *Proc Natl Acad Sci USA.* 101:1618–1621.
- Mark Welch JL, Meselson M. 1998. Karyotypes of bdelloid rotifers from three families. *Hydrobiologia.* 387/388:403–407.
- Melone G, Ricci C, Segers H, Wallace RL. 1998. Phylogenetic relationships of phylum rotifera with emphasis on the families of bdelloidea. *Hydrobiologia.* 387/388:101–107.
- Normark BB, Judson OP, Moran NA. 2003. Genomic signatures of ancient asexual lineages. *Biol J Linn Soc Lond.* 79:69–84.
- Pamilo P, Bianchi NO. 1993. Evolution of the *Zfx* and *Zfy* genes: rates and interdependence between the genes. *Mol Biol Evol.* 10:271–281.
- Ricci C. 1987. Ecology of bdelloids: how to be successful. *Hydrobiologia.* 147:117–127.
- Ricci C. 1998. Anhydrobiotic capabilities of bdelloid rotifers. *Hydrobiologia.* 387/388:321–326.
- Robotti C. 1975. Chromosome complement and male haploidy of *Asplanchna priodonta* Gosse 1850 (Rotatoria). *Experientia.* 31:1270–1272.
- Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics.* 19:2496–2497.
- Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. New York: Cold Spring Harbor Laboratory Press.
- Segers H. 2007. Annotated checklist of the rotifers (phylum rotifera), with notes on nomenclature, taxonomy and distribution. *Zootaxa.* 1564:1–104.
- Waggoner BM, Poinar GO Jr. 1993. Fossil habrotrichid rotifers in Dominican amber. *Experientia.* 48:408–410.
- Wernersson R, Pedersen AG. 2003. RevTrans—constructing alignments of coding DNA from aligned amino acid sequences. *Nucleic Acids Res.* 31:3537–3539.

Martin Embley, Associate Editor

Accepted November 3, 2008