Molecular Ecology (2016)

## Closely coupled evolutionary history of ecto- and endosymbionts from two distantly related animal phyla

JUDITH ZIMMERMANN,\*<sup>1</sup> CECILIA WENTRUP,\*†<sup>1</sup> MIRIAM SADOWSKI,\* ANNA BLAZEJAK,\* HARALD R. GRUBER-VODICKA,\* MANUEL KLEINER,\*‡ JÖRG A. OTT,§ BODIL CRONHOLM,¶ PIERRE DE WIT,\*\* CHRISTER ERSÉUS†† and NICOLE DUBILIER\*‡‡

\*Max Planck Institute for Marine Microbiology, Celsiusstrasse 1, D-28359 Bremen, Germany, †Department of Microbiology and Ecosystem Science, Division of Microbial Ecology, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria, ‡Department of Geoscience, University of Calgary, Calgary, 2500 University Drive, AB T2N 1N4, Canada, §Department of Limnology and Oceanography, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria, ¶Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, SE-104 05 Stockholm, Sweden, \*\*Department of Marine Sciences, Sven Lovén Centre for Marine Sciences Tjärnö, University of Gothenburg, Hättebäcksvägen 7, SE-452 96, Strömstad, Sweden, ††Department of Biological and Environmental Sciences, University of Gothenburg, Box 463, SE-405 30 Göteborg, Sweden, ‡‡Faculty of Biology/Chemistry, University of Bremen, Bibliothekstrasse 1, D-28359 Bremen, Germany

## Abstract

The level of integration between associated partners can range from ectosymbioses to extracellular and intracellular endosymbioses, and this range has been assumed to reflect a continuum from less intimate to evolutionarily highly stable associations. In this study, we examined the specificity and evolutionary history of marine symbioses in a group of closely related sulphur-oxidizing bacteria, called Candidatus Thiosymbion, that have established ecto- and endosymbioses with two distantly related animal phyla, Nematoda and Annelida. Intriguingly, in the ectosymbiotic associations of stilbonematine nematodes, we observed a high degree of congruence between symbiont and host phylogenies, based on their ribosomal RNA (rRNA) genes. In contrast, for the endosymbioses of gutless phallodriline annelids (oligochaetes), we found only a weak congruence between symbiont and host phylogenies, based on analyses of symbiont 16S rRNA genes and six host genetic markers. The much higher degree of congruence between nematodes and their ectosymbionts compared to those of annelids and their endosymbionts was confirmed by cophylogenetic analyses. These revealed 15 significant codivergence events between stilbonematine nematodes and their ectosymbionts, but only one event between gutless phallodrilines and their endosymbionts. Phylogenetic analyses of 16S rRNA gene sequences from 50 Cand. Thiosymbion species revealed seven well-supported clades that contained both stilbonematine ectosymbionts and phallodriline endosymbionts. This closely coupled evolutionary history of marine ecto- and endosymbionts suggests that switches between symbiotic lifestyles and between the two host phyla occurred multiple times during the evolution of the Cand. Thiosymbion clade, and highlights the remarkable flexibility of these symbiotic bacteria.

*Keywords*: cospeciation, gutless oligochaetes, nematodes, Phallodrilinae, Stilbonematinae, Thiosymbion

Received 27 October 2015; revision received 23 December 2015; accepted 19 January 2016

Correspondence: Judith Zimmermann, Fax: +49 421 2028 580; E-mail: jzimmern@mpi-bremen.deand Cecilia Wentrup; E-mail: wentrup@microbial-ecology.net <sup>1</sup>Both authors contributed equally.

## Introduction

The central role symbiosis plays in the ecology and evolution of eukaryotic organisms is now unquestioned

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

<sup>© 2016</sup> The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.

(Margulis & Fester 1991; McFall-Ngai *et al.* 2013). However, the evolutionary processes by which symbioses are established and maintained within host lineages are less clear for many symbioses. Several forms of symbiont integration are known. These range from ectosymbiotic associations, in which the symbionts are attached to the outside of the host, to endosymbiotic associations, in which the symbionts live extracellularly or intracellularly inside the host. It is often assumed that this range of lifestyles corresponds with their intimacy and evolutionary stability, that is the extent of metabolic and other biochemical interactions between host and symbiont, and the stability of the symbiotic association over time.

For example, intracellular endosymbioses with strict vertical transmission from the host to its offspring are the most intimate and evolutionarily stable associations (Bennett & Moran 2015). These symbioses often share long co-evolutionary histories between symbionts and hosts and are accompanied by severe symbiont genome degradation (Moya et al. 2008; Toft & Andersson 2010; McCutcheon & Moran 2012). These types of associations are well known from insects, but also from marine hosts such as Paracatenula flatworms, and are often reflected in congruent phylogenies of host and symbiont as a result of codiversification of the symbiotic partners over evolutionary time (Gruber-Vodicka et al. 2011; de Vienne et al. 2013). However, not all vertically transmitted endosymbioses are evolutionarily stable over extended periods of time. Symbiont loss and replacement can disrupt codiversification patterns, even in highly intimate intracellular symbioses (e.g. Stewart et al. 2008). Further evidence that intracellular symbioses are not always evolutionarily stable comes from those that are horizontally transmitted, that is symbionts are acquired from the environment. For example, in the highly intimate intracellular symbioses of the hydrothermal vent tube worm Riftia pachyptila and deep-sea bathymodiolin mussels, the endosymbionts are acquired from the surrounding seawater or cooccurring adults (Won et al. 2003; Nussbaumer et al. 2006; Fontanez & Cavanaugh 2014), and there is no evidence for codiversification (Vrijenhoek 2010). Ectosymbioses are often considered to be less intimate and evolutionarily stable than endosymbioses. Ectosymbiotic associations can be highly specific within an extant host species or genus (Goffredi 2010; Petersen et al. 2010; Flot et al. 2014), but almost nothing is known about their evolutionary stability over longer periods of time. Convincing evidence in ectosymbioses for evolutionary stability, that is strict codiversification, has to our knowledge only been described in the associations between termite gut flagellates and their ectosymbiotic bacteria (Desai et al. 2010).

Many studies have investigated the evolutionary stability of symbioses within a closely related group of either ecto- or endosymbiotic bacteria. Very few groups of closely related symbionts are known that contain both types of bacteria. One such example is a monophyletic clade of marine sulphur-oxidizing Gammaproteobacteria within the Chromatiaceae, here referred to as Candidatus Thiosymbion (H. R. Gruber-Vodicka, M. Kleiner, N. Leisch, C. Wentrup, J. Zimmermann, S. Bulgheresi, J. A. Ott, N. Dubilier, in preparation). Symbionts from this clade have adapted to different mutualistic lifestyles with three very distantly related animal groups, two from the phylum Nematoda and one from the Annelida (Ott et al. 2004; Dubilier et al. 2006, 2008). Cand. Thiosymbion species live (i) as ectosymbionts attached to the cuticle of stilbonematine nematodes (family Desmodoridae) (Ott et al. 2004), (ii) as extracellular endosymbionts in the space just below the cuticle and above the epidermal cells of gutless phallodriline annelids (oligochaetes, Clitellata, Naididae sensu Erséus et al. 2008) and (iii) as intracellular endosymbionts in the mouthless astomonematine nematode Astomonema jenneri (family Siphonolaimidae) (Ott et al. 1982). While not much is known about the diversity and distribution of astomonematine nematodes, numerous species of stilbonematines and gutless phallodrilines have been described from around the world, particularly from tropical and subtropical shallow-water habitats (Erséus 1992; Tchesunov 2013), where these hosts often co-occur (Dubilier et al. 2006).

In the ectosymbiotic associations between stilbonematine nematodes and *Cand.* Thiosymbion, different symbiont morphotypes, including rods, cocci and filaments, have been described, and most host species are associated with a single symbiont morphotype (Ott *et al.* 1991; Polz *et al.* 1992). The hosts are assumed to use their ectosymbionts as a food source by grazing on their symbiotic coat (Powell *et al.* 1979; Ott *et al.* 2004). Previous studies have shown that some *Cand.* Thiosymbion ectosymbionts are specific to a given host species, but nothing is known about their evolutionary stability over time (Polz *et al.* 1994; Bayer *et al.* 2009; Pende *et al.* 2014).

In the endosymbiotic associations of gutless phallodriline annelids with *Cand*. Thiosymbion, these bacteria have been found in all investigated host species so far and are called the primary symbionts. Each host species harbours a specific and phylogenetically distinct *Cand* Thiosymbion species, as well as up to five additional endosymbionts (which are not closely related to *Cand*. Thiosymbion) (Dubilier *et al.* 1999, 2001; Blazejak *et al*. 2005, 2006; Ruehland *et al.* 2008). Gutless phallodrilines have completely reduced digestive and excretory systems, and their endosymbionts provide them with nutrition and recycle their waste compounds (Woyke *et al.* 2006; Kleiner *et al.* 2012). It is assumed that the *Cand.* Thiosymbion bacteria are transmitted vertically from adult worms to the embryo (Krieger 2000; Dubilier *et al.* 2006), but nothing is known about how or whether these symbionts have diversified with their hosts.

The goal of this study was to better understand the evolutionary history of the Cand. Thiosymbion clade. To achieve this goal, we reconstructed the phylogeny of 19 species of stilbonematine nematodes and their Cand. Thiosymbion ectosymbionts and 22 species of gutless phallodriline annelids and their Cand. Thiosymbion endosymbionts, collected from locations around the world. We used a combined approach of multilocus phylogeny, codiversification analyses and statistics to examine the evolutionary stability of these associations, that is whether these ecto- and endosymbionts codiversified with their hosts. Second, we investigated the evolutionary processes that might explain the close phylogenetic relationships between the Cand. Thiosymbion bacteria associated with nematodes and annelids. Our question here was whether ecto- and endosymbiotic lifestyles evolved only once or multiple times within the Cand. Thiosymbion clade.

## Material and methods

#### Meiofauna collection and preparation

Stilbonematine nematodes and phallodriline annelids were collected from sites in the Atlantic and Pacific Oceans, the Mediterranean and the North Sea (Fig. 1), Table S1 and S2, Supporting information). Sediments containing the worms were collected using buckets or sediment cores, and the worms were extracted by decantation with a 64-µm mesh sieve for stilbonematines and a 250-µm mesh for phallodrilines. Live animals were sorted with a dissecting microscope and identified to the genus level (for stilbonematines) or species level (for phallodrilines) based on morphological characters. If a morphological identification of phallodrilines to the species level was not possible in the field, specimens were cut in two: the anterior part was fixed in Bouin's fluid, stained and mounted on a microscope slide as described in Erséus (1994) for a posteriori host species identification, while the posterior part was preserved in 95% ethanol for molecular analyses. All other specimens were washed in 0.2 µm filtered seawater (stilbonematines) or unfiltered seawater (phallodrilines) and immediately fixed in 70-95% ethanol for molecular host and symbiont identification. Samples were stored at 4 °C or -20 °C until further processing.

## DNA extraction, PCR and sequencing

For stilbonematine nematodes, two different DNA extraction methods were used. For nematodes from Sylt (North Sea, Germany) and Elba (Mediterranean, Italy), DNA was extracted and purified from single specimens as described previously (Schizas et al. 1997). Nematodes from Heron Island (Australian Great Barrier Reef) and the Caribbean (surrounding Carrie Bow Cay, Belize) were treated with three 25- to 40-s ultrasonication pulses (Bandelin Sonoplus HD 70, Berlin, Germany) and vortexed between pulses to mechanically remove the symbionts from the worms. DNA from worm and symbiont fractions was extracted separately and the latter pelleted by centrifugation, with the DNAeasy Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. For the host, we sequenced the 18S ribosomal RNA (rRNA) gene, because it has been shown to provide good phylogenetic resolution for nematodes, including the subfamily Stilbonematinae

> Fig. 1 Sampling sites of stilbonematine nematodes and gutless phallodriline annelids. Specimens were sampled from shallow-water habitats in the West Atlantic, Mediterranean Sea, North Sea, Australia, and one deeper site on the continental margin off Peru in the East Pacific. For details, see Tables S1 and S2 (Supporting information). The figure was generated using PanMap (Diepenbroek *et al.* 2002).





(e.g. Meldal et al. 2007; Van Megen et al. 2009; Ott et al. 2014). For the symbionts, we sequenced the partial 16S-23S rRNA operon, including the 16S rRNA gene, the internal transcribed spacer (ITS) between the 16S and 23S rRNA genes and the first 200 bp of the 23S rRNA, because the 16S rRNA gene alone did not provide sufficient resolution for relationships within some clades (e.g. Leptonemella, Catanema and Laxus) (Fig. S1, Supporting information). Genes were amplified by PCR with Phusion<sup>®</sup> DNA polymerase (Finnzymes, Finland) and the primers listed in Table S3 (Supporting information). Cycling conditions were as described in Ott et al. (2014). The purified host PCR products were sequenced bidirectionally with internal forward and reverse primers listed in Table S3 (Supporting information). For the analysis of symbionts in DNA extracts from whole worms collected on Sylt and Elba, partial 16S-23S rRNA operon clone libraries were constructed and screened for inserts with the correct size as described previously (Zimmermann et al. 2014). Up to 52 clones were partially sequenced with the universal primer 907RM (Table S1, Supporting information), and compared with sequences in GenBank using BLASTn (Altschul et al. 1997). The majority of the clones contained sequences with top BLASTn hits to previously sequenced stilbonematine ectosymbionts (>95% nucleotide identity), of which one to eight clones were selected per nematode specimen for bidirectional sequencing of the entire cloned fragment. Similar to previous studies showing that the ectosymbiont coat consisted of a specific gammaproteobacterial symbiont phylotype (Polz et al. 1994; Pende et al. 2014), the vast majority of the clones from each worm contained the same partial 16S-ITS ectosymbiont sequence. However, a few clones contained sequences from other marine bacteria that may have originated from the gut lumen. We therefore modified our sequencing approach for worms collected from Heron Island and Carrie Bow Cay and used DNA extracted from the ectosymbionts only (these were removed from the host cuticle through ultrasonification; see above). Partial 16S-23S rRNA PCR products of these ectosymbiont extracts were sequenced bidirectionally without cloning, with internal forward and reverse primers listed in Table S3 (Supporting information). Sanger sequencing of the PCR products from these extracts always yielded a single, very clear chromatogram, indicating that only a single ectosymbiont phylotype was dominant on these worms. For the stilbonematine nematode Laxus oneistus from Carrie Bow Cay, host 18S rRNA gene and symbiont 16S-23S rRNA operon sequences were extracted from a preliminary draft assembly of Illumina HiSeq short reads from a single worm metagenome. The optimization of the assembly including binning of host and symbiont genomes is

currently in progress. The raw reads and preliminary assemblies are available upon request.

For gutless phallodrilines, DNA was either extracted from whole or halved specimens using the DNAeasy Tissue Kit or the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. For the host, we sequenced a total of six nuclear and mitochondrial genetic markers because neither the nuclear 18S rRNA gene, nor the mitochondrial 16S rRNA and COI (cytochrome oxidase I) genes alone provide sufficient phylogenetic resolution (Nylander et al. 1999; Sjölin et al. 2005). We sequenced the mitochondrial (mt)12S rRNA, mt16S rRNA, mtCOI, 18S rRNA and 28S rRNA genes, as well as the ITS between the 18S and 28S rRNA genes. All PCR and sequencing reactions were carried out as described previously (White et al. 1990; Palumbi 1991; Folmer et al. 1994; Jamieson et al. 2002; Källersjö et al. 2005; Sjölin et al. 2005; Blazejak et al. 2006; Envall et al. 2006). Previous studies using clone library analyses of bacterial 16S rRNA gene PCR products showed that in all investigated individuals of a given gutless phallodriline species, the 16S rRNA gene sequences of the Candidatus Thiosymbion symbionts were identical (>99.95% nucleotide identity) (e.g. Blazejak et al. 2006; Ruehland et al. 2008). In this study, we therefore only amplified the 16S rRNA gene using a newly designed Cand. Thiosymbion-specific forward primer (G1\_55\_all) and a general bacterial reverse primer (GM4R, Muyzer et al. 1995) (Table S3, Supporting information). If available, we also amplified and sequenced host and symbiont genes from additional specimens of the same species. All host and symbiont PCR products were purified and sequenced bidirectionally with the primers listed in Table S3 (Supporting information).

Sequencing was performed either with an ABI PRISM 3100 or 377 genetic analyser (Applied Biosystems, Foster City, CA, USA), or outsourced to Macrogen Inc., South Korea. Sequences were assembled using SEQUENCHER v4.6 (GeneCodes Corporation, Ann Arbor, MI), manually curated and chimera-checked with the BELLEROPHON program (Huber *et al.* 2004). Host and symbiont genes were sequenced from multiple individuals per species, where available. For details, see Tables S1 and S2 (Supporting information).

## Phylogenetic analyses of host genes

For the nematode phylogeny, all generated 18S rRNA gene sequences and previously published sequences from other stilbonematine nematode species from which also the corresponding symbiont sequence was published, as well as four Draconematidae nematodes sequences as outgroup (*Prochaetosoma* sp. 2 and 3

(FJ182223 - 4), *Draconema japonicum* (FJ182217) and *Draconematidae* gen sp. (FJ182219)), which were downloaded from GenBank, were used for phylogenetic reconstruction. For the phallodriline annelid phylogeny, the generated mt12S, mt16S, 18S, 28S rRNA gene, ITS and mtCOI sequences of gutless phallodriline as well as five gut-bearing phallodriline annelid species as outgroup taxa were used for phylogenetic reconstruction (not every gene was obtained for every species, see Table S4, Supporting information).

Sequences for each gene were aligned separately using MAFFT v7 (Katoh & Standley 2013) with the Q-INS-I setting (Katoh & Toh 2008) that considers the predicted secondary structure of the RNA for the alignment for ribosomal genes and the L-INS-I setting (Katoh et al. 2005), that also considers gaps for phallodriline ITS and COI sequences. Alignments were manually adjusted and 5' and 3' end-trimmed using GENEIOUS v6.3 (Biomatters, New Zealand). For the phallodriline annelid phylogeny, the six separately aligned host markers were concatenated into one matrix and the matrix was partitioned according to genetic locus. Additionally, within the COI alignment, each codon position (1, 2 and 3) was treated as a separate unit to account for the fact that these evolve differently. The optimal substitution model for each alignment was assessed using the Akaike information criterion as implemented in **IMODELTEST 2.1.4** (Guindon & Gascuel 2003; Darriba et al. 2012), and the following nucleotide substitution models were chosen for phylogenetic analyses: GTR+I+G for the nematode 18S rRNA and the phallodriline ITS genes, GTR+G for the 12S and 16S rRNA genes and K80 + I+G for the 18S and 28S rRNA genes. To account for codon position in the COI gene, the alignment was partitioned into three parts, and the models assessed were GTR+G for the first, HKY+G for the second, and SYM+I+G for the third codon position.

Phylogenetic trees were reconstructed using Bayesian inference (MRBAYES v. 3.2.1) (Ronquist & Huelsenbeck 2003) maximum-likelihood-based and methods (RAxML) (Stamatakis et al. 2008). As specific nucleotide substitution models cannot be assigned separately to each gene partition using RAxML, the GTR+G+I substitution model was assigned to all partitions. For each single gene alignment as well as the concatenated data set, two separate MCMC analyses were run in MRBAYES with 4 Markov chains each (one cold and three hot) for 20 million generations and sampling every 1000 generations. For the concatenated data set, all loci were unlinked from each other for all parameters except topology. Prior to concatenation, we determined that no incongruences were statistically supported between the single-gene trees (Fig. S2, Supporting information). Convergence was evaluated by plotting the generations versus logL, and the resulting trees were summarized into a majority-rule consensus tree, after discarding the first 5000 sampled trees as the burn-in. Node stability was evaluated using posterior probabilities (pp, Bayesian inference) and bootstrap support (100 RAxML rapid bootstrap runs) with values above 0.80 considered significant.

## Phylogenetic analyses of symbiont genes

For the stilbonematine nematode ectosymbiont phylogeny, analyses were carried out with both the partial 16S-23S rRNA operon and the 16S rRNA gene sequences to test whether including the ITS and partial 23S rRNA sequences affects the ectosymbiont phylogeny (Fig. S1, Supporting information). In addition to the specimens examined in this study, we also included published stilbonematine 16S rRNA gene symbiont sequences in our phylogenetic reconstructions, if the corresponding host 18S rRNA gene sequence from the same individual was also available. The only exception was the ectosymbiont sequence from Catanema sp. ('Robbea sp. 2' in Figs 2 and 4), which was obtained from a pooled sample of 200 host individuals and may therefore not have originated from the same individual as the gene used for the host phylogeny. Sequences from pure cultures of free-living gammaproteobacterial sulphur-oxidizing bacteria (Allochromatium vinosum strain DSM180, Marichromatium purpuratum 984, Thiorhodococcus drewsii AZ1, Thiocapsa marina 5811, Thiorhodovibrio sp. 970) were downloaded from GenBank and used as an outgroup.

Both 16S rRNA and 23S rRNA genes (where available) were aligned separately using MAFFT Q-INS-I as described above. Because the ITS regions were difficult to align, ITS clusters were identified based on cluster analysis using CD-HIT (Li & Godzik 2006; Fu et al. 2012) applying a similarity threshold of 85%. Sequences that did not fall into a cluster or were single representatives of a host genus were excluded from the analysis. Five ITS clusters were identified in total (symbionts of 'undescribed genus A' species, Leptonemella species, Laxus Heron Island species, Stilbonema Heron Island species and Eubostrichus species) and aligned separately using MAFFT G-INS-I (Katoh et al. 2005). The optimal substitution model was assessed for all alignments separately as described above, which were GTR+I+G for the 16S and 23S rRNA genes, and HKY, HKY+I or GTR+G for the different ITS alignments. The 16S rRNA, five ITS and partial 23S rRNA gene alignments were concatenated and all loci unlinked from each other in all parameters except topology. All characters were weighted equally, and all gaps, including the missing ITS regions of some symbionts, were treated as missing data and did not contribute phylogenetic information to the trees.

For the phylogenetic reconstructions of the gutless phallodriline endosymbionts, we only used symbiont sequences produced in this study, as previously published symbiont sequences were not always sequenced from single individuals. We used the same outgroup as for the stilbonematine ectosymbiont phylogeny. The 16S rRNA data set was aligned, manually adjusted and trimmed as described for the host phylogenies above. The optimal substitution model suggested by JMODELTEST (Guindon & Gascuel 2003; Darriba *et al.* 2012) was the GTR+I+G.

For reconstruction of the combined Cand. Thiosymbion phylogeny based on the 16S rRNA gene, we included all sequences from this study as well as previously published 16S rRNA gene symbiont sequences from single nematodes (stilbonematines and Astomonema sp.) or gutless phallodriline annelids. Two previously published sequences amplified from coral samples off Panama were included in the tree, but their origin is not clear (GU118064, -120 Sunagawa et al. 2010) (Fig. 4) (see Discussion). To better reconstruct the ancestral state (described below), we added three of the most closely related 16S rRNA environmental sequences of bacteria from marine sediments (JF344100, -324, -607) to the analysis. The outgroup sequences were the same cultured sulphur oxidizers as described above. The optimal substitution model suggested by **JMODELTEST** (Guindon & Gascuel 2003; Darriba et al. 2012) was the GTR+I+G.

All symbiont phylogenetic trees were reconstructed with MRBAYES (Ronquist & Huelsenbeck 2003) and RAxML (Stamatakis *et al.* 2008) using the respective optimal substitution model but the same settings as described for the host trees. As specific nucleotide substitution models cannot be assigned separately to each gene partition using RAxML, the GTR+G+I substitution model was assigned to all partitions separately.

## Codivergence analysis of host and symbiont trees

To investigate codivergence between the ecto- or endosymbionts and their invertebrate hosts, we used the program JANE (version 4, http://www.cs.hmc.edu/ ~hadas/jane/; Conow *et al.* 2010). JANE is an eventbased software tool that reconciles tree topologies of hosts and symbionts by taking co-evolutionary events (codivergence, host switch, duplication, loss, failure to diverge) into account to find the best reconstruction by minimizing the global costs of these events. JANE automatically resolves polytomies so that the total cost of the resulting cophylogeny reconstruction is minimized. We used 1000 generations, a population size of 30 and the default cost settings of codivergence as the most likely event (cost = 0); duplication, loss and failure to diverge unlikely events (cost = 1); and host switches as least probable events (cost = 2) (0, 1, 1, 1, 2). When host switches were assigned equally likely events (cost = 1) as duplication, loss and failure to diverge (0, 1, 1, 1, 1,) or when loss was assigned the least probable event (0, 1, 2, 1, 1), results were similar to those obtained when using default parameters (data not shown). Global congruence was assessed by 1000 random tip mapping permutations with the same cost scheme used for the codivergence analyses. We used the consensus trees of the hosts and their respective symbionts to run the analyses (Figs 2-4 and S1, Supporting information). One of the supported codivergence events was predicted between two individuals of the same stilbonematine species (Robbea hypermnestra) and was therefore neglected in our results (Figs 2, S3 and S4, Supporting information). Furthermore, one of the supported codivergence events between host and symbionts of Catanema sp. and C. 'st. andrea' should be treated with caution because the ectosymbiont sequence of Catanema sp. was obtained from multiple individuals as explained above and may not have originated from the same individual used for the host phylogeny.

To investigate co-evolutionary events between all members of the *Cand*. Thiosymbion clade and both animal phyla, we 'artificially' connected the stilbonematine and phallodriline host trees as sister groups (Figs 2 and 3). Before the analysis, we removed all symbiont sequences from the *Cand*. Thiosymbion tree for which no corresponding host sequence was available, without changing the tree topology (Fig. 4). The analysis was run with the cost scheme, generation number and population size described above. Only events with support values higher than 80% were considered.

## Ancestral state reconstruction

Ancestral character state reconstructions were carried out with the combined ecto- and endosymbiont 16S rRNA phylogeny, using BAYESTRAITS V2.0 (Pagel & Meade 2014). Gutless phallodriline and Astomonema sp. sequences were assigned the character state 'endosymbiont', stilbonematine sequences the character state 'ectosymbiont'. The three closely related environmental sequences from marine sediments (JF344100, -324, -607) were assigned the character state 'uncertain'. We used 5 million post burn-in trees and ran the analyses for 50 million generations, 500 000 burn-in generations and three runs each. Statistical support for the Bayesian ancestral state reconstruction was determined using



**Fig. 2** Phylogeny of stilbonematine nematodes and their ectosymbiotic *Candidatus* Thiosymbion bacteria. Consensus tree of Bayesian inference (MRBAYES) and maximum-likelihood (RAxML) analyses of host 18S rRNA gene and partial 16S-23S rRNA operon sequences of their ectosymbiotic bacteria. For phylogeny based on the 16S rRNA gene only, see Fig. S1 (Supporting information). Coloured boxes show monophyletic groups of hosts and their symbionts, respectively, and coloured names the sampling locations. Sequences published prior to this study are discernible by their accession numbers. Accession numbers for sequences from this study are listed in Table S1 (Supporting information). Stars indicate 12 of the 15 statistically supported codivergence events and red vertical arrows statistically supported host switching events (>80%, JANE v.4, Conow *et al.* 2010) (for more details, see Fig. S3, Supporting information). Nodes that were statistically supported in both analyses (>80% bootstrap value and 0.80 posterior probability (pp) value) are highlighted in black, and nodes that were statistically supported only in Bayesian inference are highlighted in white. Provisional working names for undescribed species are in quotes. The scale bars represent average nucleotide substitutions per site. The host and symbiont trees were rooted with four *Draconematidae* and five free-living Chromatiaceae sequences, respectively.

posterior probabilities calculated with the 'MRCA' command in BAYESTRAITS. To test for significance of the reconstructed ancestral state, we used the log Bayes factor test on the harmonic means calculated by the 'fossil' command.

### Correlation analyses

Patristic distance matrices, estimated from Bayesian Inference phylogenies of stilbonematines, gutless phallodrilines and their respective symbionts were calculated with the program PATRISTIC (Fourment & Gibbs 2006). Jukes–Cantor distance matrices of single genes (16S rRNA gene of endo- and ectosymbionts and 18S rRNA gene of stilbonematines) and concatenated gene alignments (mt12S, mt16S, 18S, 28S rRNA, ITS and mtCOI of gutless phallodrilines) were calculated with the software package ARB (Ludwig *et al.* 2004). Gutless phallodriline hosts that had one of the marker sequences missing were excluded from the Jukes–Cantor genetic distance analyses. Geographical distance matrices were calculated based on sampling coordinates (Table S1 and S2, Supporting information) using the GEOGRAPHIC DIS-TANCE MATRIX GENERATOR v1.2.3 (Ersts 2014) while considering a sphere shape of the Earth. We used single-line linear distances between coordinates and did not take dispersal barriers such as continents into account. Linear distances were similar to the distances taking continents into account in all cases except for one (stilbonematine nematodes sampled in the Mediterranean and Sylt). We assumed that current geographical distances, whether they include dispersal barriers or not, have little meaning for the routes taken during diversification of stilbonematines and gutless phallodrilines millions of years ago. As the age of these symbioses is not known (there is no fossil record for these soft-bodied worms), we cannot estimate how continents were shaped during their distribution and the routes these worms took while diversifying. Mantel and partial Mantel tests to test the association of symbiont distances with host and geographical distances were performed with R (http://www.r-project.org/) using Pearson's r-statistic and 999 Monte Carlo permutations implemented in the R package 'vegan' (Oksanen et al. 2012).



Gutless phallodriline hosts

• statistically supported by pp and bootstrap \* statistically supported codivergence event O statistically supported by pp

Fig. 3 Phylogeny of gutless phallodriline annelids and their endosymbiotic *Candidatus* Thiosymbion bacteria. Consensus tree of Bayesian inference (MRBAYES) and maximum-likelihood (RAxML) analyses of six concatenated genetic markers for the host (mitochondrial (mt)12S, mt16S, 18S, 28S rRNA genes, internal transcribed spacer between 18S and 28S rRNA (ITS), mtCOI (cytochrome oxidase I)) and the 16S rRNA gene for their endosymbiotic bacteria. For trees based on single genes, see Fig. S2 (Supporting information). Coloured boxes show monophyletic groups of gutless phallodriline hosts and their endosymbiotic bacteria, respectively, and coloured names the sampling locations. Accession numbers for sequences from this study are listed in Table S4 (Supporting information). Stars indicate statistically supported codiversification events (>80%, JANE v.4, Conow *et al.* 2010). Nodes that were statistically supported in both analyses (>80% bootstrap value and 0.80 posterior probability (pp) value) are highlighted in black and nodes that were statistically supported only in Bayesian inference are highlighted in white. Provisional working names for undescribed species are in quotes. The scale bars represent average nucleotide substitutions per site. The host and symbiont trees were rooted with five gut-bearing Phallodriline and five free-living Chromatiaceae sequences, respectively.

## Results

## *Cophylogeny analyses of the stilbonematine nematode ectosymbiosis*

Host phylogeny: confirmation of the current host genus classification. To investigate the evolutionary history of stilbonematine nematode hosts and their *Candidatus* Thiosymbion ectosymbionts, we reconstructed host and symbiont phylogenies based on genetic markers from 19 nematode species collected around the world (Fig. 1, Table S1, Supporting information). We sequenced the 18S rRNA gene from individual stilbonematine worms and found that it provided sufficient phylogenetic resolution to distinguish between closely related species (Fig. 2). The 18S rRNA gene phylogeny was consistent with the current morphological classification of host genera (Fig. 2). All stilbonematine nematode sequences formed a highly supported monophyletic group clearly distinct from the nonsymbiotic desmodorid nematodes that we used as an outgroup (not shown), confirming previous studies on stilbonematines (Kampfer *et al.* 1998; Bayer *et al.* 2009; Ott *et al.* 2014). All stilbonematine species from the same genus formed highly supported clades independent of their sampling location (Fig. 2). In addition to the six previously described

**Bacterial endosymbionts** 

Fig. 4 *Candidatus* Thiosymbion phylogeny showing the evolutionary history of stilbonematine nematode and gutless phallodriline annelid symbionts. Consensus tree of Bayesian inference (MRBAYES) and maximum-likelihood (RAxML) analyses based on the 16S rRNA gene. Brown and yellow boxes highlight ectosymbiotic bacteria of stilbonematine nematodes and endosymbiotic bacteria of gutless phallodriline annelids, respectively. The red box highlights endosymbiotic bacteria of astomonematine nematodes. Coloured names show the sampling location of the symbionts. Sequences published prior to this study are discernible by their accession numbers. Accession numbers for sequences from this study are listed in Tables S1 and S4 (Supporting information). Nodes that were statistically supported in both analyses (>80% bootstrap value and 0.80 posterior probability (pp) value) are highlighted in black and nodes that were statistically supported only in Bayesian inference are highlighted in white. Stars indicate statistically supported codiversification events (>80%, JANE v.4, Conow *et al.* 2010, for more details, see Fig. S6, Supporting information). Red stars show well-supported clades that contain both ecto- and endosymbiotic bacteria. Provisional working names for undescribed species are in quotes. The scale bar represents average nucleotide substitutions.per site.

stilbonematine nematode genera, we discovered two new host genera that had 18S rRNA gene sequences that were distinct from all previously described genera. We preliminarily named these 'undescribed genus A' (two new species from Australia and Belize) and 'undescribed genus B' (one species from Belize) (Fig. 2).



<sup>© 2016</sup> The Authors. Molecular Ecology published by John Wiley & Sons Ltd.

While the relationships of species within the different genera were well resolved, the relationships between the different host genera were not well resolved (Fig. 2). The basal position of the *Eubostrichus* clade to all other nematode genera was the only node that was statistically supported in both maximum-likelihood and Bayesian inference analyses (Fig. 2).

Ectosymbiont phylogeny: high congruence with host phylogeny. To analyse the Cand. Thiosymbion ectosymbiont phylogeny, we sequenced the 16S rRNA gene, the 16S-23S internal transcribed spacer (ITS) and the partial 23S rRNA gene (partial 16S-23S rRNA operon) from the same individuals used for the host analyses in this study. Analyses of up to 52 clones from the same host individual revealed highly similar to identical Cand. Thiosymbion sequences (>99.93% nucleotide identity), indicating that each individual is associated with a dominant ectosymbiont type, as shown in earlier studies (Table S1, Supporting information, Polz et al. 1994; Bayer et al. 2009; Pende et al. 2014). Different individuals of a given stilbonematine species had symbionts with highly similar to identical 16S rRNA gene sequences (>99.7% nucleotide identity), but the ITS sequences sometimes differed considerably between individuals of the same species with as little as 96.4% nucleotide identity, and occasionally ITS regions contained long gaps or insertions. Symbionts were never shared among stilbonematine species suggesting highly species-specific associations. Nucleotide identities between 16S rRNA gene sequences ranged from 98.1% to 99.95% between symbionts of hosts from the same genus and from 93.9% to 98.3% between symbionts of hosts from different genera (Fig. 3).

Symbiont sequences from hosts that belonged to the same genus formed highly supported clades independent of their geographical location in both maximumlikelihood and Bayesian inference analyses based on partial 16S-23S rRNA operon sequences (Fig. 2). The only exceptions were symbionts of the genus Eubostrichus that were nested within the Laxus symbiont clade. Phylogenetic analyses of the 16S rRNA gene alone resulted in almost identical branching patterns, but with lower statistical support (Fig. S1, Supporting information). Within each host genus, symbiont and host branching patterns were always congruent (Fig. 2). The only exception was the genus Laxus, where host and symbiont phylogenies were not congruent (Fig. 2). In contrast to the congruence between host and symbiont phylogeny within each host genus, no congruence was observed between host genera and their corresponding symbiont clades. For example, Eubostrichus species belonged to a basal branch of the host tree and were not closely related to Laxus host species, while Eubostri*chus* symbionts were more derived and most closely related to the symbionts of *Laxus* hosts (Fig. 2). Mantel test results confirmed the highly significant correlation between symbiont and host genetic distances ( $R^2 = 0.4$ , P < 0.001 (partial 16S-23S rRNA), P = 0.001 (16S rRNA)), and the lack of correlation between symbiont genetic distances and their geography ( $R^2 = 0.11$ , P = 0.078) (Table 1).

# Cophylogeny analysis of the gutless phallodriline endosymbiosis

Host phylogeny: the current host genus classification needs revision. To investigate the evolutionary history of gutless phallodriline annelids and their Cand. Thiosymbion endosymbionts, we sequenced host and symbiont genetic markers from 22 host species (of which two are cryptic in the nominal Olavius imperfectus) collected around the world (Fig. 1, Table S2, Supporting information). Many of the gutless phallodriline symbioses cooccurred with the stilbonematines analysed in this study (Tables S1 and S2, Supporting information). Six host genetic markers encoding the mt12S rRNA, mt16S rRNA, mtCOI, 18S rRNA, 28S rRNA genes, and the ITS region between the 18S and 28S rRNA genes were analysed, because none of these alone could sufficiently resolve the host phylogeny (Fig. S2, Supporting information). A concatenated phylogenetic analysis of all six host genetic markers supported the monophyly of the gutless phallodrilines with their clear separation from Phallodrilinae species with a mouth and digestive tract (Fig. 3), as shown previously in phylogenetic analyses based on morphological and molecular characters (e.g. Erséus 1984, 1992; Ferraguti & Erséus 1999). Our analyses also confirmed the monophyly of the four Inanidrilus species we analysed in this study (Inanidrilus clade) (Nylander et al. 1999; Erséus et al. 2000, 2002). However, the 18 Olavius species investigated here did not form a monophyletic group (Fig. 3). This result contradicts previous morphological and molecular analyses that indicated the monophyly of this genus (Sjölin et al. 2005), but supports an earlier molecular study with only four Olavius species and a single mitochondrial gene (Nylander et al. 1999). We identified four highly supported Olavius clades (1, 2, 3, 4 in Fig. 3). Four Olavius species did not fall into any of these four clades, and two further Olavius species fell basal to all other gutless phallodrilines (Fig. 3). The paraphyly of the genus Olavius indicates that the taxonomy of gutless phallodrilines needs revision. Given the discrepancy between our molecular phylogeny and the current classification that names only two host genera, Inanidrilus and Olavius, in the gutless phallodrilines, we use the term 'host clades' instead of host genera in this paper.

Target group	Distance measure	Matrix A	Matrix B	$R^2$	Probability (P-value)
Gutless phallodriline endosymbioses	Patristic distances	Host multigene distances Geographical distance matrix*	Symbiont 16S rRNA gene distances Symbiont 16S rRNA gene distances	0.26 0.24	0.015 0.006
	Jukes-Cantor distances	Host multigene distances Geographical distance matrix*	Symbiont 16S rRNA gene distances Symbiont 16S rRNA gene distances	$-0.2 \\ 0.3$	0.95 0.001
Stilbonematine ectosymbioses	Patristic distances	Host 18S rRNA gene distances	Symbiont partial 16S-23S rRNA operon distances	0.42	<0.001
		Host 18S rRNA gene distances	Symbiont 16S rRNA gene distances	0.4	0.001
		Geographical distance matrix*	Symbiont partial 16S-23S rRNA operon distances	0.11	0.078
		Geographical distance matrix*	Symbiont 16S rRNA gene distances	0.08	0.136
	Jukes-Cantor distances	Host 18S rRNA gene distances Geographical distance matrix*	Symbiont 16S gene rRNA distances Symbiont 16S gene rRNA distances	0.22 0.07	0.068 0.187

Table 1 Mantel test of correlation between host and symbiont genetic distances and geographical distances, based on Pearson's product-moment correlation and 999 Monte Carlo permutations

\*Based on geographical coordinates.

Similar to stilbonematine nematodes, our analyses revealed little congruence between the phylogeny of gutless phallodrilines and their geographical location (Fig. 3). Many closely related gutless phallodriline species have a disjunct distribution, that is they are widely separated from each other geographically (Fig. 3). Only two clades, *Olavius* 1 and 4, consisted exclusively of species from the same geographical region, although *Olavius* 4 with only two species was most likely undersampled.

Endosymbiont phylogeny: weak congruence with host phylogeny but high congruence with geographical location. To analyse the phylogeny of the Cand. Thiosymbion endosymbionts of gutless phallodrilines, we sequenced the 16S rRNA gene from the same individuals used for host analyses (Fig. 3, Table S4, Supporting information). Symbiont 16S rRNA gene sequences from individuals of the same species were highly similar to identical (≥99.93% nucleotide identity), indicating that each species is associated with a specific Cand. Thiosymbion phylotype, a result also shown in earlier analyses (e.g. Blazejak et al. 2005; Ruehland et al. 2008). Cand. Thiosymbion phylotypes were never shared between any of the 22 host species, suggesting highly speciesspecific associations. Endosymbiont 16S rRNA gene sequence similarities ranged from 99.93% nucleotide identity (Olavius algarvensis and O. 'pianosa 1'), corresponding to a consistent single bp difference, to 94.7% nucleotide identity (O. imperfectus A and O. crassitunicatus) (for details on the investigated number of specimens per species, see Table S4, Supporting information). The only

exception was the 16S rRNA gene sequence of the *Olavius geniculatus* symbiont, which had a 65-bp insertion not present in the other symbionts.

The phylogeny of the endosymbiotic Cand. Thiosymbion showed little congruence with host phylogeny (Fig. 3). For example, the symbionts of hosts from three of the four highly supported Olavius clades did not cluster according to the phylogeny of their hosts but were rather interspersed throughout the tree. Furthermore, basal taxa in the host tree, such as O. 'st andrea 3' and Olavius loisae have symbionts with a more derived position in the symbiont tree, and vice versa, basal taxa in the symbiont tree, such as the O. algarvensis and O. 'pianosa 1' symbionts have hosts with a more derived position in the host tree (Fig. 3). Similarly, while all Inanidrilus species formed a highly supported clade, not all of their symbionts clustered together. Congruent phylogenies were only observed in closely related hosts from similar geographical regions, such as within the Olavius clade 1, and Inanidrilus species from the West Atlantic (Belize, Bermuda and the Bahamas) (Fig. 3).

In contrast to the weak congruence between symbiont and host phylogeny, both treeing methods showed strong support for the clustering of endosymbionts according to their geographical location (Fig. 3). For example, the symbionts of *Inanidrilus* species from the West Atlantic formed a well-supported clade and were not closely related to the Australian *I. manae* symbiont, while their *Inanidrilus* hosts were very closely related and formed a monophyletic clade (Fig. 3). Similarly, the symbiont of the Australian *Olavius* 'heron island A' clustered with symbionts from other Australian hosts in the *Olavius* 1 clade, but their hosts did not (Fig. 3). In fact, the four symbiont clusters with significant statistical support always contained symbionts from similar geographical locations. These analyses suggest that the phylogeny of the *Cand*. Thiosymbion endosymbionts of gutless phallodrilines was more strongly affected by geography than divergence of their hosts. This observation was also supported by the results of our Mantel test (Table 1) that showed strong support for a correlation between symbiont genetic distances and geographical distances ( $R^2 = 0.24$ , P = 0.006). However, the Mantel test also showed a significant correlation between symbiont and host genetic distances ( $R^2 = 0.26$ , P = 0.015; Table 1), although with less support than for geographical distances.

## *Phylogenetic reconstruction of the symbiont clade Candidatus Thiosymbion*

Phylogenetic analyses of the 16S rRNA genes from *Cand*. Thiosymbion bacteria sequenced in this study as well as those available in sequence databases revealed that these formed a monophyletic group that also included the endosymbionts of marine nematodes from the genus *Astomonema* (Fig. 4). The closest relatives of *Cand*. Thiosymbion are uncultured Gammaproteobacteria from marine sediments and cultivated, free-living phototrophic sulphur oxidizers of the family Chromatiaceae (Fig. 4).

The Cand. Thiosymbion clade consisted of 25 wellsupported clades (represented by black and white nodes in Fig. 4). Five of these clades contained only stilbonematine ectosymbionts while a sixth clade contained both stilbonematine ectosymbionts of Eubostrichus and sequences from coral reef sediments (Fig. 4) (Sunagawa et al. 2010). Four clades contained only phallodriline endosymbionts and another clade the published sequences of Astomonema nematode endosymbionts (Musat et al. 2007). Seven of the well-supported clades in the Cand. Thiosymbion phylogeny contained both ectoand endosymbionts (red stars in Fig. 4) and three of these consisted of ecto- and endosymbionts that originated from hosts collected in the same geographical region (Fig. 4). For example, the ectosymbiont of the nematode Laxus 'heron 2' from Australia formed a monophyletic group with the endosymbionts of Australian gutless phallodrilines (Fig. 4). Similarly, the ecto- and endosymbionts of Laxus oneistus and O. longissimus as well as those of Robbea hypermnestra and O. imperfectus fell into well-supported subclades, and all were sampled in the Caribbean region (i.e. Belize and Bahamas).

#### Ancestral state reconstruction

The ancestral state of the character association type (ectosymbiont or endosymbiont) was reconstructed based on the *Cand.* Thiosymbion consensus tree (Fig. 4). The analyses suggested that the ancestral state ectosymbiont had a slightly higher posterior probability (0.513) than endosymbiont (0.487). However, this result was not well supported by the log Bayes factor test (logBF = 0.3) where logBF > 6 is considered significant (Pagel & Meade 2014).

## Codivergence analyses

Codivergence analysis of stilbonematine nematodes and their Cand. Thiosymbion ectosymbionts with JANE based on the phylogeny of the partial 16S-23S rRNA operon resulted in a total of 15 codivergence events that were highly supported, as well as one host switch from Laxus oneistus to the ancestor of all Eubostrichus species (Table 2, Figs 2 and S3, Supporting information). Randomized trees showed an average overall cost of  $35.4 \pm 2.1$ , which was more than four times higher than the observed optimal cost, indicating highly significant global codivergence (P = 0.001) between stilbonematine nematodes and their ectosymbionts (Table 2). Codivergence analyses based on the phylogeny of the 16S rRNA ectosymbiont gene alone (without the ITS-23S sequences) revealed nearly identical results. The only differences to the analyses with the ITS-23S sequences were one less codivergence event (14 instead of 15), and one additional symbiont loss event, both caused by the different phylogenetic branching patterns within the Leptonemella symbiont clade (Table 2, Figs 2, S1 and S4, Supporting information). This indicates that the pronounced pattern of codivergence between stilbonematine hosts and their symbionts was not biased by our choice of phylogenetic markers.

For gutless phallodrilines and their *Cand.* Thiosymbion endosymbionts only one codivergence event between the Australian hosts *O. albidoides* and *O. prodigus* and their endosymbionts was highly supported (Table 2, Figs 3 and S5, Supporting information). Randomized trees showed an average overall cost of  $35 \pm 2.6$ , which was 1.4 times higher than the observed optimal cost indicating global congruence, that is that the phylogenies of these symbionts and their hosts are more similar than expected by chance alone (*P* = 0.002).

We next compared the phylogeny of all *Cand*. Thiosymbion ecto- and endosymbionts shown in Fig. 4 with a host tree that linked the stilbonematine nematodes and the gutless phallodriline annelids as sister groups (i.e. a host tree that combined the trees shown in Figs 2 and 3). JANE analyses resulted in three equally likely solutions, each of them revealing nine highly supported codivergence events between stilbonematines and their ectosymbionts. Additionally, up to 20 host switches were identified, of which 12 occurred between

	Statistically supported codivergence events/total	Statistically supported host switch events/total	Statistically supported loss events/total	Global cost (P-value) <sup>‡</sup>	Global cost randomized tip mappings
Phallodrilinae and their endosymbionts	1/10	0/12	0/1	25 (0.001)	35
Stilbonematinae and their ectosymbionts*	15/19	1/3	0/2	8 (0.001)	35
Stilbonematinae and their ectosymbionts <sup>†</sup>	14/18	1/4	1/5	13 (0.001)	35
Both hosts and their symbionts	Solution 1: 9/25 Solution 2: 9/26 Solution 3: 9/27	0/20 0/19 0/18	0/5 0/7 0/9	45 45 45	77 77 77

Table 2 Codivergence analyses with JANE version 4

For more details, see Fig. S3-S6 (Supporting information).

\*Based on the consensus partial 16S-23S rRNA operon tree of the ectosymbionts.

<sup>†</sup>Based on the consensus 16S rRNA gene tree of the ectosymbionts.

<sup>\*</sup>The overall sum of events multiplied by the cost we attributed to each event.

stilbonematine nematodes and gutless phallodrilines, but none of the 20 host switches was statistically supported (Table 2, Fig. S6, Supporting information). Randomized trees showed an average overall cost of 77.2  $\pm$  2.4, which was more than 1.7 times higher than the observed optimal cost, indicating highly significant global codivergence (P = 0.001) between *Cand*. Thiosymbion ecto- and endosymbionts and their hosts (Table 2).

## Discussion

# Long-term evolutionary stability of stilbonematine nematode ectosymbioses

Closely related stilbonematine nematode hosts were almost consistently associated with closely related Candidatus Thiosymbion ectosymbionts. This was visible in the clear congruence within all host and symbiont clades, with only the Laxus symbioses forming the only exception to this consistent pattern (Fig. 2). In contrast to the high congruence of host and symbiont phylogenies within each host genus, no congruence was observed between the different host genera and their corresponding symbionts. This could be due to the poor phylogenetic resolution of the trees at the nodes connecting host genera and symbiont clades, but could also indicate that symbiont switches between host genera occurred in the past. Mantel test results confirmed the highly significant relationship between symbiont and host patristic distances (Table 1), and JANE analyses confirmed the marked congruence between host and symbiont phylogeny with the identification of up to 15 codivergence events (Table 2, Figs S3 and S4, Supporting information).

The pronounced codivergence between stilbonematine nematodes and their ectosymbiotic bacteria is remarkable given that the symbionts are attached to the surface of their hosts and constantly exposed to the surrounding sediment. This exposure could cause detachment of the symbionts when the worms migrate through the sediment. Furthermore, as all marine nematodes, stilbonematines moult and shed their cuticle four times during their life cycle (Ott *et al.* 1995). This lifestyle would provide ample opportunities for free-living bacteria or symbionts from other hosts to invade and displace the host's symbionts, but our analyses revealed only one replacement event (symbionts of an ancestor of *Eubostrichus* were replaced by an ancestral *Laxus* symbiont (Fig. 2)).

The high genus and species specificity as well as evolutionary stability of the stilbonematine ectosymbiosis implies that highly specific mechanisms for recognition and maintenance have evolved to sustain these symbioses. These could include morphological adaptations such as a reduced body diameter of the nematode host to accommodate the thickness of the symbiont coat (Polz et al. 1992), possibly to avoid friction and thus detachment of symbionts during movement through the sediments. Longitudinal division of the symbionts with an unusual shift in the fission plane may prevent the symbionts from losing contact with their host's surface during growth, as suggested for some rod-shaped bacterial symbionts (Polz et al. 1992; Leisch et al. 2012). Furthermore, unique glandular sensory organs contribute to specificity by secreting species-specific lectins on the hosts' cuticle that recognize sugar residues on the symbiont's cell membrane and show symbiont specificity at the isoform level

(Bauer-Nebelsick *et al.* 1995; Nussbaumer *et al.* 2004; Bulgheresi *et al.* 2006, 2011).

Transmission of the symbionts from one generation to the next and during the four moults of the stilbonematines life cycle is also critical for maintaining the symbiosis. It has been assumed that the symbionts are acquired horizontally from a free-living population in the environment or co-occurring hosts based on (i) the absence of symbionts on juvenile worms that had not yet hatched from their egg (Bulgheresi and Ott unpublished in Bayer et al. 2009), and (ii) the amplification of partial 16S rRNA gene sequences with high similarity to symbiont sequences of stilbonematine nematodes from surface seawaters (Heindl et al. 2011) (but see below under 'Repeated host switches between animal phyla' for a more detailed discussion of environmental sequences closely related to stilbonematine ectosymbionts). Convincing evidence for codivergence in symbioses with horizontal acquisition is, however, rare (Bright & Bulgheresi 2010). It is therefore possible that vertical transmission between stilbonematine hosts and their symbionts or a mixed mode of vertical and occasional horizontal transmission occurs in these associa-(Ebert 2013). The ectosymbionts could be tions transferred onto the egg or juvenile during passage through the female genital opening, called the vulva. Intriguingly, the glandular sensory organs described above that secrete symbiont specific lectins are especially numerous and large in the vulva region of many stilbonematine nematodes (e.g. Hopper & Cefalu 1973; Ott et al. 1991, 1995). A similar mode of transmission, in which symbionts are transmitted 'pseudo-vertically' onto eggs or juveniles in the brood pouch, has also been suggested for the ectosymbionts of Niphargus amphipods found in freshwater caves (Dattagupta et al. 2009).

Evidence for codivergence in ectosymbiotic associations has so far only been shown for termite gut flagellates and their *Bacteroidales* ectosymbionts (Desai *et al.* 2010). Our study provides an example for a highly specific marine ectosymbiosis in which host and symbiont phylogenies showed a pronounced pattern of codivergence and highlights that ectosymbiotic associations can be highly stable over long evolutionary time periods.

## Weak congruence between gutless phallodriline hosts and endosymbiont phylogenies suggests repeated symbiont replacements

Our analyses of the phylogeny of gutless phallodrilines and their *Cand*. Thiosymbion endosymbionts showed little congruence between host and symbiont branching patterns (Fig. 3). This result was supported by our codivergence analyses with JANE that predicted only a single codivergence event between two closely related species (Table 2, Fig. 3). On the other hand, Mantel test analyses revealed a significant correlation between symbiont and host genetic distances, and JANE analyses showed significant support for global codivergence. These results indicate that the evolution of the annelid endosymbioses was influenced in part by the codiversification of the symbionts and their host. However, our phylogenetic analyses revealed that geography also had a strong influence on endosymbiont diversity, as shown by the many symbiont clades that contained symbionts from the same geographical region (Fig. 3). Mantel test results provided further support for a strong correlation between symbiont genetic distance and geographical distances (Table 1). Clearly, geographical location had a major impact on the evolutionary history of these symbioses.

The lack of consistent codivergence between gutless phallodrilines and their obligate Cand. Thiosymbion endosymbionts is surprising given that these symbionts are assumed to be transmitted vertically via a smear infection from the parent to offspring during egg deposition (Giere & Langheld 1987; Krieger 2000). Our data suggest recurring events of symbiont displacements, but how these occurred is not clear. One possible window of opportunity for displacement are the early developmental stages of the host. Gutless phallodrilines lay single eggs together with sperm into a cocoon that they deposit into the sediments they inhabit (Giere & Langheld 1987; Krieger 2000). The embryo develops within the cocoon before the juvenile worm eventually hatches. During this phase of exposure to the environment, symbionts from the environment or co-occurring hosts would have the opportunity to invade the cocoon, infect the developing embryo or juvenile and displace the vertically transmitted symbiont. For successful displacement, the newly acquired symbiont would have to have a strong selective advantage over the original symbiont to be able to be sustained in the population. For example, the newly acquired symbiont could be better adapted to the host's environment or enable it to invade new habitats or niches.

Occasional events of horizontal symbiont acquisition are well known from associations where the symbionts are predominantly transmitted vertically (reviewed by Bright & Bulgheresi 2010; Ebert 2013). For example, marine vesicomyid and solemyid clams that live in symbiosis with sulphur-oxidizing bacteria predominantly transmit their symbionts vertically, but molecular analyses indicate that horizontal acquisition of symbionts has happened in the past (Stewart *et al.* 2008, 2009; Decker *et al.* 2013). Other examples include bryozoan *Bugula* species and their *Cand.* Endobugula symbionts (Lim-Fong *et al.* 2008) and leaf-cutter ants that cultivate specific symbiotic fungi strains for food (Bot *et al.* 2001; Green *et al.* 2002).

In some hosts, such as insects, where vertically transmitted symbionts have undergone genome reduction, horizontal acquisition of a novel symbiont enables the host to escape dependency on a degenerated, inefficient symbiont (Lefèvre et al. 2004; Conord et al. 2008; Koga et al. 2013). So far, the genome of only one Cand. Thiosymbion has been sequenced from the gutless phallodriline Olavius algarvensis (Woyke et al. 2006). This symbiont does not appear to have a reduced genome although unusually high amounts of transposases indicate that it may be in transition to an obligate, hostassociated lifestyle (Kleiner et al. 2013). Additionally, we never found any evidence for the co-occurrence of more than one Cand. Thiosymbion species in any of the phallodriline hosts we examined. This suggests that cooccurrence of a primary vertically transmitted symbiont and a novel horizontally acquired Cand. Thiosymbion is rare in these hosts and that symbiont displacement occurs rapidly within a host population.

## Repeated host switches between animal phyla

Our phylogenetic reconstruction of 50 *Cand*. Thiosymbion ecto- and endosymbionts from three very distantly related host groups (Stilbonematinae, *Astomonema* and Phallodrilinae) from two animal phyla (Nematoda and Annelida), and from a wide range of environments around the world, showed that these are monophyletic and share a single common ancestor (Fig. 4). These results confirm earlier studies with lower taxonomic coverage that showed the remarkably close relationships between these ecto- and endosymbionts despite the large phylogenetic distances between their three host groups (e.g. Musat *et al.* 2007; Bulgheresi *et al.* 2011; Heindl *et al.* 2011; Pende *et al.* 2014). Seven highly supported clades contained both stilbonematine ectosymbionts and phallodriline endosymbionts (Fig. 4).

Two scenarios that are not mutually exclusive could explain the intermingled evolutionary history of the *Cand.* Thiosymbion ecto- and endosymbionts (we only discuss scenarios for the stilbonematine and phallodriline symbionts, because the *Astomonema* nematode symbiosis with only two published symbiont sequences is undersampled): (i) closely related free-living bacteria from the environment repeatedly displaced symbionts associated with stilbonematine and/or gutless phallodriline hosts, or (ii) the phallodriline endosymbionts evolved multiple times from stilbonematine ectosymbionts and/or vice versa. The first scenario is less likely for the stilbonematine ectosymbiosis because our phylogenetic and statistical analyses indicated that the evolution of these symbionts was largely driven by codivergence with their hosts. For the gutless phallodriline endosymbiosis, this scenario is possible given the lack of codivergence and the strong influence of geography on their evolution. However, to date there is little evidence for free-living bacteria from the Cand. Thiosymbion clade despite intensive sequencing of marine environments (including those in which stilbonematines and gutless phallodrilines occur) by others (e.g. Musat et al. 2006; Schöttner et al. 2011; Aravindraja et al. 2013) and ourselves (J. Wippler and N. Dubilier, unpublished results). 16S rRNA gene sequences that fall within the Cand. Thiosymbion clade have only been described in two studies. In the first study on shallowwater coral-associated bacteria (Sunagawa et al. 2010), two sequences were found (GU118064 and GU118120), which form a monophyletic clade with the ectosymbionts of stilbonematines from the genus Eubostrichus based on our analyses (Fig. 4). These sequences may have originated from symbionts of Eubostrichus individuals contained in the sampled coral fragments. The other sequences were found in surface waters close to sediments containing stilbonematines and gutless phallodrilines, but could only be amplified using the highly sensitive, and contamination-prone nested PCR approach (Heindl et al. 2011).

The second scenario for explaining the intermingled evolutionary history of seven clusters that contain Cand. Thiosymbion ecto- and endosymbionts is that these symbionts repeatedly switched between the two animal phyla, that is stilbonematine ectosymbionts displaced phallodriline endosymbionts and/or phallodriline endosymbionts displaced stilbonematine ectosymbionts. JANE analyses predicted 12 host switches between stilbonematine nematodes and gutless phallodrilines, although none of these were statistically supported (Fig. S6, Supporting information). This lack of statistical support for host switching can be explained by the low resolution of the Cand. Thiosymbion 16S rRNA gene phylogeny at critical nodes, which hampered a more precise reconstruction of their evolutionary history (Fig. 4). Assuming that host switching between the two animal phyla occurred, it appears more likely to have occurred from nematodes to annelids given that the evolution of the stilbonematine nematode ectosymbionts was much more closely linked to their host phylogeny than the evolution of the gutless phallodriline endosymbionts. Geography played a key role in the evolution of gutless phallodriline symbioses, and their common cooccurrence with stilbonematine nematodes would have provided numerous opportunities for uptake of the stilbonematine symbionts.

Host switching of *Cand*. Thiosymbion between the two animal phyla Nematoda and Annelida is intriguing as it predicts the ability of the bacterial symbiont to

shift between ecto- and endosymbiotic lifestyles. One explanation for this evolutionary flexibility is the similar function of the symbionts in the stilbonematine and phallodriline associations. All Cand. Thiosymbion bacteria are assumed to be autotrophic sulphur oxidizers that provide their hosts with nutrition. (It should be noted, however, that evidence for chemolithoautotrophy has only been shown in some symbionts (Schiemer et al. 1990; Polz et al. 1992; Hentschel et al. 1999; Blazejak et al. 2006; Woyke et al. 2006; Kleiner et al. 2012), and the contribution of the symbionts to their host's nutrition remains to be clarified.) A second explanation for the flexibility of these associations is that stilbonematines and phallodrilines provide similar benefits to their symbionts. Both hosts migrate between the upper oxidized and the lower reduced layers of the sediment and thus provide their symbionts with access to electron acceptors such as oxygen in the surface layers and electron donors such as reduced sulphur compounds in the lower layers (Giere et al. 1991; Ott et al. 1991). A third explanation is that major physiological adaptations might not have been needed for switches between an ectosymbiotic and an endosymbiotic lifestyle. Although the phallodriline associations are endosymbiotic, the bacteria are not intracellular, but extracellular and only separated from the sediment environment by a very thin cuticle that is highly permeable for charged and uncharged molecules of up to 70 kDa (Dubilier et al. 2006). Thus, in shifts from an ectosymbiotic lifestyle with nematodes to the endosymbiotic habitus with phallodrilines, these symbionts did not have to evolve the intricate adaptations that are needed to maintain and persist intracellularly within host cells.

## Conclusions and outlook

The Candidatus Thiosymbion association is, to our knowledge, the only known marine symbiosis in which host switches appear to have occurred repeatedly between animal phyla. There is one other group of closely related marine symbionts that associates with different animal phyla: sulphur-oxidizing bacteria that live as intracellular symbionts in deep-sea bathymodiolin mussels, in the tissues of a poecilosclerid sponge and as ectosymbionts on a marine terebellid polychaete (Nishijima et al. 2010; Petersen et al. 2012). However, the symbionts from these three host groups are not most closely related to each other, but rather to free-living sulphur-oxidizing bacteria in the environment, and there is no phylogenetic evidence for host switching between the mussels, sponge and polychaete (Petersen et al. 2012). In associations between terrestrial hosts and bacteria, host switching within an animal phylum is common, especially among arthropods (e.g.

Heath *et al.* 1999; Russell *et al.* 2003; Huigens *et al.* 2004; Weinert *et al.* 2009). However, clear evidence for host switching between animal phyla is rare. Intracellular *Wolbachia* bacteria that live in mutualistic and parasitic associations with nematodes and arthropods may have switched between these two host phyla (Casiraghi *et al.* 2005), but more recent studies indicate that this cannot be resolved with current phylogenetic methods (Bordenstein *et al.* 2009; Comandatore *et al.* 2013).

The intermingled evolutionary history within the Cand. Thiosymbion clade of ecto- and endosymbiotic bacteria therefore raises a number of questions: Are there conserved recognition mechanisms in Cand. Thiosymbion bacteria that allow them to engage in symbiotic associations with hosts from three animal groups and two animal phyla? Which factors favour codivergence, which lead to host switching? What are the morphological and physiological mechanisms that determine an ecto- or endosymbiotic lifestyle, and how have these evolved in switches between these two lifestyles? To answer these questions, comparative analyses of the symbionts' genomes, transcriptomes, proteomes and metabolomes would help identify the similarities and differences in their functional and metabolic interactions with phylogenetically distant hosts. In parallel, the collection and analysis of more hosts from a wider range of environments and locations would improve the reconstruction of the phylogenetic history of Cand. Thiosymbion and shed light on the factors that enabled their evolutionary flexibility and success.

## Acknowledgements

We thank the staff of the HYDRA Institute on Elba, Italy, the Wadden Sea Station Sylt of the Alfred Wegener Institute in Bremerhaven, Germany, the Smithsonian Institution and Carrie Bow Cay Marine Laboratory in Belize, the Heron Island Research Station in Australia and the Bermuda Institute of Ocean Sciences in St. George's for providing laboratory facilities and technical assistance. We gratefully acknowledge the collaboration with the Parco Nazionale Arcipelago Toscano for granting us access to the protected area of Pianosa, Italy (research permit no. 4440/2011) and the Bermuda Aquarium in Flatts Village for access to the sampling site and their facilities. Thanks to Lisa Levin from Scripps Institution of Oceanography at UC San Diego, California for O. crassitunicatus samples. Many thanks to Jillian M. Petersen, Lizbeth Sayavedra and Pier L. Buttigieg for valuable scientific discussions and Silke Wetzel and Sabine Kuehn for excellent technical assistance. This work is contribution 984 from the Carrie Bow Cay Laboratory, Caribbean Coral Reef Ecosystems Program, National Museum of Natural History, Washington, DC. It is also a contribution from work performed by CE and ND at the Lizard Island Research Station (Great Barrier Reef, Australia), and the former Caribbean Marine Research Centre on Lee Stocking Island (Bahamas); CE is indebted to Philippe Bouchet, and the Total

Foundation, for the invitation to participate in the LIFOU 2000 expedition to New Caledonia, to Fred E. Wells for hosting the International Marine Biological Workshop on The Marine Flora and Fauna of Rottnest Island, Western Australia, and to Anna Ansebo (University of Gothenburg) for invaluable assistance in the early phases of the molecular work on gutless phallodriline annelids. This work was supported by the Gordon and Betty Moore Foundation through Grant GBMF3811 to ND, the Max Planck Society and the Marie Sklodowska-Curie Initial Training Network 'Symbiomics'. CE was funded by the Swedish Research Council, and the Adlerbertska Foundation (Göteborg).

#### References

- Altschul S, Madden T, Schaffer A et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nuceic Acids Research, 25, 3389–3402.
- Aravindraja C, Viszwapriya D, Karutha Pandian S (2013) Ultradeep 16S rRNA sequencing analysis of geographically similar but diverse unexplored marine samples reveal varied bacterial community composition. *PLoS ONE*, **8**, e76724.
- Bauer-Nebelsick M, Blumer M, Urbancik W, Ott JA (1995) The glandular sensory organ of Desmodoridae (Nematoda)ultrastructure and phylogenetic implications. *Invertebrate Biology*, **114**, 211–219.
- Bayer C, Heindl NR, Rinke C *et al.* (2009) Molecular characterization of the symbionts associated with marine nematodes of the genus *Robbea*. *Environmental Microbiology Reports*, **1**, 136–144.
- Bennett GM, Moran NA (2015) Heritable symbiosis: the advantages and perils of an evolutionary rabbit hole. Proceedings of the National Academy of Sciences of the United States of America, 112, 10169–10176.
- Blazejak A, Erséus C, Amann R, Dubilier N (2005) Coexistence of bacterial sulfide oxidizers, sulfate reducers, and spirochetes in a gutless worm (Oligochaeta) from the Peru margin. Applied and Environmental Microbiology, 71, 1553–1561.
- Blazejak A, Kuever J, Erséus C, Amann R, Dubilier N (2006) Phylogeny of 16S rRNA, ribulose 1,5-bisphosphate carboxylase/ oxygenase, and adenosine 5'-phosphosulfate reductase genes from gamma- and alphaproteobacterial symbionts in gutless marine worms (Oligochaeta) from Bermuda and the Bahamas. *Applied and Environmental Microbiology*, **72**, 5527–5536.
- Bordenstein SR, Paraskevopoulos C, Hotopp JCD et al. (2009) Parasitism and mutualism in Wolbachia: what the phylogenomic trees can and cannot say. Molecular Biology and Evolution, 26, 231–241.
- Bot ANM, Rehner SA, Boomsma JJ (2001) Partial incompatibility between ants and symbiotic fungi in two sympatric species of *Acromyrmex* leaf-cutting ants. *Evolution*, **55**, 1980–1991.
- Bright M, Bulgheresi S (2010) A complex journey: transmission of microbial symbionts. *Nature Reviews Microbiology*, 8, 218– 230.
- Bulgheresi S, Schabussova I, Chen T *et al.* (2006) A new C-type lectin similar to the human immunoreceptor DC-SIGN mediates symbiont acquisition by a marine nematode. *Applied and Environmental Microbiology*, **72**, 2950–2956.
- Bulgheresi S, Gruber-Vodicka HR, Heindl NR *et al.* (2011) Sequence variability of the pattern recognition receptor

Mermaid mediates specificity of marine nematode symbioses. *ISME Journal*, **5**, 986–998.

- Casiraghi M, Bordenstein SR, Baldo L *et al.* (2005) Phylogeny of *Wolbachia pipientis* based on gltA, groEL and ftsZ gene sequences: clustering of arthropod and nematode symbionts in the F supergroup, and evidence for further diversity in the *Wolbachia* tree. *Microbiology*, **151**, 4015–4022.
- Comandatore F, Sassera D, Montagna M et al. (2013) Phylogenomics and analysis of shared genes suggest a single transition to mutualism in *Wolbachia* of nematodes. *Genome Biology* and Evolution, 5, 1668–1674.
- Conord C, Despres L, Vallier A *et al.* (2008) Long-term evolutionary stability of bacterial endosymbiosis in Curculionoidea: additional evidence of symbiont replacement in the Dryophthoridae family. *Molecular Biology and Evolution*, **25**, 859–868.
- Conow C, Fielder D, Ovadia Y, Libeskind-Hadas R (2010) Jane: a new tool for the cophylogeny reconstruction problem. *Algorithms for Molecular Biology*, **5**, 16.
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nature Methods, 9, 772.
- Dattagupta S, Schaperdoth I, Montanari A *et al.* (2009) A novel symbiosis between chemoautotrophic bacteria and a freshwater cave amphipod. *ISME Journal*, **3**, 935–943.
- Decker C, Olu K, Arnaud-Haond S, Duperron S (2013) Physical proximity may promote lateral acquisition of bacterial symbionts in vesicomyid clams. *PLoS ONE*, 8, e64830.
- Desai MS, Strassert JFH, Meuser K *et al.* (2010) Strict cospeciation of devescovinid flagellates and Bacteroidales ectosymbionts in the gut of dry-wood termites (Kalotermitidae). *Environmental Microbiology*, **12**, 2120–2132.
- Diepenbroek M, Grobe H, Reinke M *et al.* (2002) PANGAEA an information system for environmental sciences. *Computers* & *Geosciences*, **28**, 1201–1210.
- Dubilier N, Amann R, Erséus C et al. (1999) Phylogenetic diversity of bacterial endosymbionts in the gutless marine oligochete Olavius loisae (Annelida). Marine Ecology Progress Series, 178, 271–280.
- Dubilier N, Mulders C, Ferdelman T *et al.* (2001) Endosymbiotic sulphate-reducing and sulphide-oxidizing bacteria in an oligochaete worm. *Nature*, **411**, 298–302.
- Dubilier N, Blazejak A, Ruehland C (2006) Symbioses between bacteria and gutless marine oligochaetes. In: *Molecular Basis* of Symbiosis Progress in Molecular and Subcellular Biology (ed Overmann PDJ), pp. 251–275. Springer-Verlag, Berlin.
- Dubilier N, Bergin C, Lott C (2008) Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nature Reviews Microbiology*, 6, 725–740.
- Ebert D (2013) The epidemiology and evolution of symbionts with mixed-mode transmission. *Annual Review of Ecology*, *Evolution, and Systematics*, 44, 623–643.
- Envall I, Kallersjö M, Erséus C (2006) Molecular evidence for the non-monophyletic status of Naidinae (Annelida, Clitellata, Tubificidae). *Molecular Phylogenetics and Evolution*, 40, 570–584.
- Erséus C (1984) Taxonomy and phylogeny of the gutless Phallodrilinae (Oligochaeta, Tubificidae), with descriptions of one new genus and twenty-two new species\*. *Zoologica Scripta*, **13**, 239–272.
- Erséus C (1992) A generic revision of the Phallodrilinae (Oligochaeta, Tubificidae). *Zoologica Scripta*, **21**, 5–48.

<sup>© 2016</sup> The Authors. Molecular Ecology published by John Wiley & Sons Ltd.

- Erséus C (1994) The Oligochaeta. In: Taxonomic Atlas of the Santa Maria Basin and Western Santa Barbara Channel. Vol. 4 -The Annelida Part 1. Oligochaeta and Polychaeta: Phyllodocida (Phyllodocidae to Paralacydoniidae) (eds. Blake JA, Hilbig B), pp. 5–38. Santa Barbara Museum of Natural History, Santa Barbara, California.
- Erséus C, Prestegaard T, Källersjö M (2000) Phylogenetic analysis of Tubificidae (Annelida, Clitellata) based on 18S rDNA sequences. *Molecular Phylogenetics and Evolution*, **15**, 381–389.
- Erséus C, Källersjö M, Ekman M, Hovmöller R (2002) 18S rDNA phylogeny of the Tubificidae (Clitellata) and its constituent taxa: dismissal of the Naididae. *Molecular Phylogenetics and Evolution*, 22, 414–422.
- Erséus C, Wetzel MJ, Gustavsson L (2008) ICZN rules–a farewell to Tubificidae (Annelida, Clitellata). *Zootaxa*, **1744**, 66–68.
- Ersts PJ (2014) Geographic Distance Matrix Generator (version 1.2.3) American Museum of Natural History. Center for Biodiversity and Conservation. Available from http://biodiversityinformatics amnh org/open\_source/gdmg (Accessed on 16 June 2014).
- Ferraguti M, Erséus C (1999) Sperm types and their use for a phylogenetic analysis of aquatic clitellates. *Hydrobiologia*, 402, 225–237.
- Flot JF, Bauermeister J, Brad T *et al.* (2014) Niphargus-Thiothrix associations may be widespread in sulphidic groundwater ecosystems: evidence from southeastern Romania. *Molecular Ecology*, 6, 1405–1417.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biologz and Biotechnology*, 3, 294–299.
- Fontanez KM, Cavanaugh CM (2014) Evidence for horizontal transmission from multilocus phylogeny of deep-sea mussel (Mytilidae) symbionts. *Environmental Microbiology*, 16, 3608– 3621.
- Fourment M, Gibbs MJ (2006) PATRISTIC: a program for calculating patristic distances and graphically comparing the components of genetic change. BMC Evolutionary Biology, 6, 1.
- Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28, 3150–3152.
- Giere O, Langheld C (1987) Structural organisation, transfer and biological fate of endosymbiotic bacteria in gutless oligochaetes. *Marine Biology*, **93**, 641–650.
- Giere O, Conway NM, Gastrock G, Schmidt C (1991) "Regulation" of gutless annelid ecology by endosymbiotic bacteria. *Marine Ecology Progress Series. Oldendorf*, 68, 287–299.
- Goffredi SK (2010) Indigenous ectosymbiotic bacteria associated with diverse hydrothermal vent invertebrates. *Environmental Microbiology Reports*, 4, 479–488.
- Green AM, Mueller UG, Adams RMM (2002) Extensive exchange of fungal cultivars between sympatric species of fungus-growing ants. *Molecular Ecology*, **11**, 191–195.
- Gruber-Vodicka HR, Dirks U, Leisch N et al. (2011) Paracatenula, an ancient symbiosis between thiotrophic Alphaproteobacteria and catenulid flatworms. Proceedings of the National Academy of Sciences of the United States of America, 108, 12078–12083.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

- Heath BD, Butcher RDJ, Whitfield WGF, Hubbard SF (1999) Horizontal transfer of *Wolbachia* between phylogenetically distant insect species by a naturally occurring mechanism. *Current Biology*, **9**, 313–316.
- Heindl NR, Gruber-Vodicka HR, Bayer C et al. (2011) First detection of thiotrophic symbiont phylotypes in the pelagic marine environment. FEMS Microbiology Ecology, 77, 223–227.
- Hentschel U, Berger EC, Bright M, Felbeck H, Ott JA (1999) Metabolism of nitrogen and sulfur in ectosymbiotic bacteria of marine nematodes (Nematoda, Stilbonematinae). *Marine Ecology Progress Series*, **183**, 149–158.
- Hopper BE, Cefalu RC (1973) Free-living marine nematodes from Biscayne Bay, Florida V. Stilbonematinae: contributions to the taxonomy and morphology of the genus *Eubostrichus* Greeff and related genera. *Transactions of the American Microscopical Society*, **92**, 578–591.
- Huber T, Faulkner G, Hugenholtz P (2004) Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics*, **20**, 2317–2319.
- Huigens ME, de Almeida RP, Boons PAH, Luck RF, Stouthamer R (2004) Natural interspecific and intraspecific horizontal transfer of parthenogenesis--inducing Wolbachia in Trichogramma wasps. Proceedings of the Royal Society of London. Series B: Biological Sciences, 271, 509–515.
- Jamieson BG, Tillier S, Tillier A et al. (2002) Phylogeny of the Megascolecidae and Crassiclitellata (Annelida, Oligochaeta): combined versus partitioned analysis using nuclear (28S) and mitochondrial (12S, 16S) rDNA. Zoosystema-Paris, 24, 707–734.
- Källersjö M, Von Proschwitz T, Lundberg S, Eldenäs P, Erséus C (2005) Evaluation of ITS rDNA as a complement to mitochondrial gene sequences for phylogenetic studies in freshwater mussels: an example using Unionidae from northwestern European. Zoologica Scripta, 34, 415–424.
- Kampfer S, Sturmbauer C, Ott J (1998) Phylogenetic analysis of rDNA sequences from Adenophorean nematodes and implications for the Adenophorea-Secernentea controversy. *Invertebrate Biology*, **117**, 29–36.
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.
- Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, 9, 286–298.
- Katoh K, Kuma K, Toh H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, 33, 511–518.
- Kleiner M, Wentrup C, Lott C *et al.* (2012) Metaproteomics of a gutless marine worm and its symbiotic microbial community reveal unusual pathways for carbon and energy use. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, E1173–E1182.
- Kleiner M, Young JC, Shah M, VerBerkmoes NC, Dubilier N (2013) Metaproteomics reveals abundant transposase expression in mutualistic endosymbionts. *mBio*, 4, e00223–13.
- Koga R, Bennett GM, Cryan JR, Moran NA (2013) Evolutionary replacement of obligate symbionts in an ancient and diverse insect lineage. *Environmental Microbiology*, **15**, 2073– 2081.
- Krieger J (2000) Funktion und Übertragung endosymbiotischer Bakterien bei bakteriensymbiontischen, darmlosen marinen Oligochae-

ten. (Function and transmission of endosymbiotic bacteria in gutless marine oligochaetes.). PhD Thesis, University of Hamburg, Germany.

- Lefèvre C, Charles H, Vallier A *et al.* (2004) Endosymbiont phylogenesis in the Dryophthoridae weevils: evidence for bacterial replacement. *Molecular Biology and Evolution*, **21**, 965–973.
- Leisch N, Verheul J, Heindl NR *et al.* (2012) Growth in width and FtsZ ring longitudinal positioning in a gammaproteobacterial symbiont. *Current Biology*, **22**, R831–R832.
- Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, **22**, 1658–1659.
- Lim-Fong GE, Regali LA, Haygood MG (2008) Evolutionary relationships of "Candidatus Endobugula" bacterial symbionts and their Bugula bryozoan hosts. Applied and Environmental Microbiology, 74, 3605–3609.
- Ludwig W, Strunk O, Westram R et al. (2004) ARB: a software environment for sequence data. *Nucleic Acids Research*, **32**, 1363–1371.
- Margulis L, Fester R (1991) *Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis,* MIT Press, Cambridge, Massachusetts.
- McCutcheon JP, Moran NA (2012) Extreme genome reduction in symbiotic bacteria. *Nature Reviews Microbiology*, **10**, 13– 26.
- McFall-Ngai M, Hadfield MG, Bosch TCG *et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 3229–3236.
- Meldal BH, Debenham NJ, De Ley P *et al.* (2007) An improved molecular phylogeny of the Nematoda with special emphasis on marine taxa. *Molecular Phylogenetics and Evolution*, **42**, 622–636.
- Moya A, Peretó J, Gil R, Latorre A (2008) Learning how to live together: genomic insights into prokaryote–animal symbioses. *Nature Reviews Genetics*, **9**, 218–229.
- Musat N, Werner U, Knittel K *et al.* (2006) Microbial community structure of sandy intertidal sediments in the North Sea, Sylt-Rømø Basin, Wadden Sea. *Systematic and Applied Microbiology*, **29**, 333–348.
- Musat N, Giere O, Gieseke A *et al.* (2007) Molecular and morphological characterization of the association between bacterial endosymbionts and the marine nematode *Astomonema* sp. from the Bahamas. *Environmental Microbiology*, 9, 1345–1353.
- Muyzer G, Teske A, Wirsen C, Jannasch H (1995) Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Archives of Microbiology*, **164**, 165–172.
- Nishijima M, Lindsay DJ, Hata J *et al.* (2010) Association of thioautotrophic bacteria with deep-sea sponges. *Marine Biotechnology*, **12**, 253–260.
- Nussbaumer AD, Bright M, Baranyi C, Beisser CJ, Ott JA (2004) Attachment mechanism in a highly specific association between ectosymbiotic bacteria and marine nematodes. *Aquatic Microbial Ecology*, **34**, 239–246.
- Nussbaumer AD, Fisher CR, Bright M (2006) Horizontal endosymbiont transmission in hydrothermal vent tubeworms. *Nature*, **441**, 345–348.

- Nylander JAA, Erséus C, Källersjö M (1999) A test of monophyly of the gutless Phallodrilinae (Oligochaeta, Tubificidae) and the use of a 573-bp region of the mitochondrial cytochrome oxidase I gene in analysis of annelid phylogeny. *Zoologica Scripta*, **28**, 305–313.
- Oksanen J, Blanchet FG, Kindt R *et al.* (2012) Vegan: community ecology. Package. R package version 2.0–4. URL http:// CRAN.R-project.org/package=vegan.
- Ott J, Rieger G, Rieger R, Enderes F (1982) New mouthless interstitial worms from the sulfide system: symbiosis with prokaryotes. *P.S.Z.N. Marine Ecology*, **3**, 313–333.
- Ott JA, Novak R, Schiemer F *et al.* (1991) Tackling the sulfide gradient: a novel strategy involving marine nematodes and chemoautotrophic ectosymbionts. *P.S.Z.N. Marine Ecology*, **12**, 261–279.
- Ott JA, Bauer-Nebelsick M, Novotny V (1995) The genus *Laxus* Cobb, 1894 (Stilbonematinae: Nematoda): description of two new species with ectosymbiotic, chemoautotrophic bacteria. *Proceedings of the Biological Society Washington*, **108**, 508–527.
- Ott J, Bright M, Bulgheresi S (2004) Symbioses between marine nematodes and sulfur-oxidizing chemoautotrophic bacteria. *Symbiosis*, **36**, 103–126.
- Ott JA, Gruber-Vodicka HR, Leisch N, Zimmermann J (2014) Phylogenetic confirmation of the genus *Robbea* (Nematoda: Desmodoridae, Stilbonematinae) with the description of three new species. *Systematics and Biodiversity*, **12**, 434– 455.
- Pagel M, Meade A (2014) BayesTraits v.2.0 Computer program and documentation. Available at http://www.evolution.rdg.ac.uk/BayesTraits.html.
- Palumbi S (1991) Simple fool's guide to PCR. Zoology University of Hawaii, Honolulu. Available from: http://agris.fao.org/ agris-search/search.do?recordID=US201300300173
- Pende N, Leisch N, Gruber-Vodicka HR et al. (2014) Size-independent symmetric division in extraordinarily long cells. *Nature Communications*, 5, 1–10.
- Petersen JM, Ramette A, Lott C *et al.* (2010) Dual symbiosis of the vent shrimp Rimicaris exoculata with filamentous gamma-and epsilonproteobacteria at four Mid-Atlantic Ridge hydrothermal vent fields. *Environmental Microbiology*, 8, 2204–2218.
- Petersen JM, Wentrup C, Verna C, Knittel K, Dubilier N (2012) Origins and evolutionary flexibility of chemosynthetic symbionts from deep-Sea animals. *The Biological Bulletin*, 223, 123–137.
- Polz MF, Felbeck H, Novak R, Nebelsick M, Ott JA (1992) Chemoautotrophic, sulfur-oxidizing symbiotic bacteria on marine nematodes: morphological and biochemical characterization. *Microbial Ecology*, 24, 313–329.
- Polz MF, Distel DL, Zarda B *et al.* (1994) Phylogenetic analysis of a highly specific association between ectosymbiotic, sulfur-oxidizing bacteria and a marine nematode. *Applied and Environmental Microbiology*, **60**, 4461–4467.
- Powell EN, Crenshaw MA, Rieger RM (1979) Adaptations to sulfide in sulfide-system I. <sup>35</sup>S-sulfide accumulation and the presence of a sulfide detoxification system. *Journal o f Experimental Marine Biology and Ecology*, **37**, 57–76.
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19, 1572–1574.

© 2016 The Authors. Molecular Ecology published by John Wiley & Sons Ltd.

- Ruehland C, Blazejak A, Lott C *et al.* (2008) Multiple bacterial symbionts in two species of co-occurring gutless oligochaete worms from Mediterranean sea grass sediments. *Environmental Microbiology*, **10**, 3404–3416.
- Russell JA, Latorre A, Sabater-Muñoz B, Moya A, Moran NA (2003) Side-stepping secondary symbionts: widespread horizontal transfer across and beyond the Aphidoidea. *Molecular Ecology*, **12**, 1061–1075.
- Schiemer F, Novak R, Ott J (1990) Metabolic studies on thiobiotic free-living nematodes and their symbiotic microorganisms. *Marine Biology*, **106**, 129–137.
- Schizas NV, Street GT, Coull BC, Chandler GT, Quattro JM (1997) An efficient DNA extraction method for small metazoans. *Molecular Marine Biology and Biotechnology*, 6, 381– 383.
- Schöttner S, Pfitzner B, Grünke S et al. (2011) Drivers of bacterial diversity dynamics in permeable carbonate and silicate coral reef sands from the Red Sea. Environmental Microbiology, 13, 1815–1826.
- Sjölin E, Erséus C, Källersjö M (2005) Phylogeny of Tubificidae (Annelida, Clitellata) based on mitochondrial and nuclear sequence data. *Molecular Phylogenetics and Evolution*, **35**, 431– 441.
- Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Systematic Biology, 57, 758–771.
- Stewart FJ, Young CR, Cavanaugh CM (2008) Lateral symbiont acquisition in a maternally transmitted chemosynthetic clam endosymbiosis. *Molecular Biology and Evolution*, 25, 673–687.
- Stewart FJ, Baik AHY, Cavanaugh CM (2009) Genetic subdivision of chemosynthetic endosymbionts of *Solemya velum* along the southern New England coast. *Applied and Environmental Microbiology*, **75**, 6005–6007.
- Sunagawa S, Woodley CM, Medina M (2010) Threatened corals provide underexplored microbial habitats. *PLoS ONE*, **5**, e9554.
- Tchesunov AV (2013) Marine free-living nematodes of the subfamily Stilbonematinae (Nematoda, Desmodoridae): taxonomic review with descriptions of a few species from the Nha Trang Bay, Central Vietnam. *Meiofauna Marina*, **20**, 71– 94.
- Toft C, Andersson SGE (2010) Evolutionary microbial genomics: insights into bacterial host adaptation. *Nature Reviews Genetics*, 11, 465–475.
- Van Megen H, van den Elsen S, Holterman M et al. (2009) A phylogenetic tree of nematodes based on about 1200 fulllength small subunit ribosomal DNA sequences. Nematology, 116, 927–950.
- de Vienne DM, Refrégier G, López-Villavicencio M *et al.* (2013) Cospeciation vs host-shift speciation: methods for testing, evidence from natural associations and relation to coevolution. *New Phytologist*, **198**, 347–385.
- Vrijenhoek RC (2010) Genetics and evolution of deep-sea chemosynthetic bacteria and their invertebrate hosts. *In: Vent* and Seep Biota: Aspects from Microbes to Ecosystems (ed. Kiel S), pp. 15–49. Springer Netherlands, Dordrecht, the Netherlands.
- Weinert LA, Werren JH, Aebi A, Stone GN, Jiggins FM (2009) Evolution and diversity of *Rickettsia* bacteria. *BMC Biology*, 7, 6.

- White TJ, Bruns T, Lee S *et al.* (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protocols: A Guide for Methods and Applications*, **18**, 315–322.
- Won YJ, Hallam SJ, O'Mullan GD et al. (2003) Environmental acquisition of thiotrophic endosymbionts by deep-sea mussels of the genus Bathymodiolus. Applied and Environmental Microbiology, 69, 6785–6792.
- Woyke T, Teeling H, Ivanova NN *et al.* (2006) Symbiosis insights through metagenomic analysis of a microbial consortium. *Nature*, **443**, 950–955.
- Zimmermann J, Lott C, Weber M *et al.* (2014) Dual symbiosis with co-occurring sulfur-oxidizing symbionts in vestimentiferan tubeworms from a Mediterranean hydrothermal vent. *Environmental Microbiology*, **16**, 3638–3656.

J.Z. and C.W. developed the concept, conceived the manuscript, constructed large parts of the molecular data and performed the phylogenetic and statistical analyses; J.Z. wrote the manuscript; C.W. helped to write parts of the manuscript. M.S., A.B. and B.C. contributed to the construction of the molecular data; H.G.V helped to develop the concept and organized the sampling on Belize; M.K. organized the sampling of many gutless Phallodrilinae species; J.O. helped to develop the concept, contributed to the sampling and morphologically identified the stilbonematine nematodes; P.D.W. contributed to the phylogenetic and cophylogenetic analyses; C.E. helped to develop the concept, contributed with molecular data of the Phallodrilinae and their phylogenetic analyses, contributed to the sampling and morphologically identified the marine Phallodrilinae; N.D. helped to develop the concept, conceived and wrote some parts of the manuscript. All authors edited the manuscript.

### Data accessibility

All sequences were submitted to GenBank and are available under the following Accession nos.: Gutless phallodriline annelid genes: 18S rRNA: KP943792 -KP943817, 28S rRNA: KP943818 - KP943844, mtCOI: KP943845 - KP943866, ITS (between 18S rRNA and 28S rRNA genes): KP943867 - KP943884, mt12S rRNA: KP943885 - KP943908, mt16S rRNA: KP943909 -KP943931. Gutless phallodriline annelid endosymbiont 16S rRNA genes: KP943932 - KP943955 - KP943970 and KT826595. Stilbonematine nematode ectosymbiont partial 16S-23S rRNA operon: KP943971 - KP943988 and KT826595. Sequence alignments and tree files for all loci analysed in our study are available on Dryad Digital Repository (doi:10.5061/dryad.1tf8g).

## Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Stilbonematine nematode ectosymbiont phylogeny, based on the 16S rRNA gene.

Fig. S2 Phallodriline annelids phylogenies, based on different host genetic markers.

**Fig. S3** Cophylogeny of stilbonematine nematodes and *Candidatus* Thiosymbion ectosymbionts (based on the partial 16S-23S rRNA operon).

Fig. S4 Cophylogeny of stilbonematine nematodes and *Candidatus* Thiosymbion ectosymbionts (based on the 16S rRNA gene).

**Fig. S5** Cophylogeny of gutless phallodriline annelids and *Candidatus* Thiosymbion endosymbionts (based on the 16S rRNA gene).

Fig. S6 Cophylogeny of gutless phallodriline annelids and stilbonematine nematodes and their *Candidatus* Thiosymbion symbionts.

Table S1 Stilbonematine nematodes sampled for this study.

Table S2 Gutless phallodriline annelids sampled for this study.

Table S3 Primers used in this study.

**Table S4** Amplified genes and accession numbers for gutless phallodriline annelids and their endosymbionts.