New phylogenomic and comparative analyses provide corroborating evidence that Myxozoa is Cnidaria

Jin-Mei Feng, Jin-Yong Zhang, Ya-Lin Yang, Bin Yao, Zhi-Gang Zhou, Wei Miao

A R T I C L E   I N F O

Article history:
Received 2 April 2014
Revised 5 August 2014
Accepted 14 August 2014
Available online 1 September 2014

Keywords:
Myxozoa
Thelohanellus kitauei
Phylogenomic analysis
Comparative analysis
Cnidaria

A B S T R A C T

Myxozoa, a diverse group of morphologically simplified endoparasites, are well known fish parasites causing substantial economic losses in aquaculture. Despite active research, the phylogenetic position of Myxozoa remains ambiguous. After obtaining the genome and transcriptome data of the myxozoan Thelohanellus kitauei, we examined the phylogenetic position of Myxozoa from three different perspectives. First, phylogenomic analyses with the newly sequenced genomic data strongly supported the monophyly of Myxozoa and that Myxozoa is sister to Medusozoa within Cnidaria. Second, we detected two homologs to cnidarian-specific minicollagens in the T. kitauei genome with molecular characteristics similar to cnidarian-specific minicollagens, suggesting that the minicollagen homologs in T. kitauei may have functions similar to those in Cnidaria and that Myxozoa is Cnidaria. Additionally, phylogenetic analyses revealed that the minicollagens in myxozoans and medusozoans have a common ancestor. Third, we detected 11 of the 19 proto-mesodermal genes in the T. kitauei genome, which were also present in the cnidian Hydra magnipapillata, indicating Myxozoa is within Cnidaria. Thus, our results robustly support Myxozoa as a derived cnidarian taxon with an affinity to Medusozoa, helping to understand the diversity of the morphology, development and life cycle of Cnidaria and its evolution.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Myxozoans (phylum Myxozoa) are a diverse group of microscopic obligate endoparasites with characteristic multicellular spores, distinct polar capsules and an extrudable polar filament used in the invasion of hosts. They are composed of two subgroups, the Malacosporea and the Myxosporea, which include more than 2000 species (Canning and Okamura, 2004; Kent et al., 2001; Lom and Dykova, 2006). Several malacosporean species have been discovered to date, including the muscular, vermiform-like para-Lom and Dykova, 2006). Several malacosporean species have been described for only one malacosporean, Tetracapsuloides bryosalmonae (Grabner and El-Matbouli, 2008; Morris and Adams, 2006; Tops et al., 2004), and less than 40 myxosporeans (Lom and Dykova, 2006). The myxosporean Thelohanellus kitauei, a pathological parasite that infects carp (of the family Cyprinidae), is one of the most important cultured freshwater fish species in the world, and causes intestinal giant cystic disease (IGCD) (Kitaue, 1980). This disease has been recognized as one of the most serious myxospordiosest causes economic losses in aquaculture (Seo et al., 2012).

The phylogenetic position of Myxozoa has been controversial for decades, and resolving the phylogenetic placement of Myxozoa has become a recent hot topic in evolutionary biology. Myxozoans were originally assigned to protozoans, but years later, a series of studies revealed that myxozoans possess the features of multicellular spores, cellular junctions, separation of generative and somatic cells, and the differentiation of somatic cells (Canning and Okamura, 2004; Siddall et al., 1995). These features, along with the metazoan-like 18S ribosomal DNA sequences in myxozoans and Okamura, 2004; Siddall et al., 1995). These features, along with the metazoan-like 18S ribosomal DNA sequences in myxozoans
(Smothers et al., 1994), resulted in the general acceptance of Myxozoa within Metazoa. However, the relative placement of Myxozoa within Metazoa remains uncertain; it has been placed either at the base of Bilateria or within Cnidaria, depending on the characteristics analyzed and the phylogenetic analysis employed (Evans et al., 2010; Jimenez-Guri et al., 2007; Nesnidal et al., 2013). An ultrastructure study revealed the triplastic nature of the malacosporean B. plumatellae by showing that this species possesses four sets of longitudinal muscle and indicated that it is closely related to the worm-like bilaterians (Okamura et al., 2002). Additionally, the phylogenetic analyses based on the rDNA sequences robustly supported Myxozoa as a sister group to Bilateria (Siddall et al., 1995; Zrzavy and Hyspa, 2003).

However, several lines of evidence have also suggested that Myxozoa is Cnidaria. Morphologically, all myxozoans possess polar capsules, which are extraordinarily similar to the cnidarian nematocysts in structure and function (Weill, 1938). Molecula- rly, a homolog to minicollagens, which are cnidarian-specific nemato- cyst proteins, was detected in the malacosporean T. bryosalmonae (Holland et al., 2010), further indicating a close relationship between myxozoan polar capsules and cnidarian nematocysts. Additionally, an ultrastructure study revealed tetra-radially arranged musculature in B. plumatellae consistent with that in medusozoans, indicating an affinity between Myxozoa and Medusozoa within Cnidaria (Gruhl and Okamura, 2012). Furthermore, two phylogenomic analyses have also posited Myxozoa as a highly derived Cnidaria, likely within Medusozoa (Jimenez-Guri et al., 2007; Nesnidal et al., 2013). However, in a study using only EST data from malacosporean B. plumatellae, a bilaterian origin for Myxozoa could not be rejected with the topology test (Jimenez-Guri et al., 2007); in another study using EST data from B. plumatellae and genomic data from myxosporean Myxobolus cerebralis, Myxozoa was recovered into Cnidaria strongly supported by a maximum likelihood (ML) analysis but not as strongly supported by a Bayesian inference analysis with the CAT model (Nesnidal et al., 2013). Therefore, the relative phylogenetic placement of Myxozoa within Metazoa remains unresolved in certain aspects and is compelling for further study.

Recently, phylogenomic analyses have generally improved the robustness of molecular phylogenetic reconstructions and resolved many debated phylogenies (Chiari et al., 2012; Hampel et al., 2009; Liang et al., 2013; Philippe et al., 2009, 2005; Philippe and Telford, 2006; Ryan et al., 2013; Struck and Fisse, 2008). However, adequate genomic data are still lacking for Myxozoa. Herein, we obtained the newly sequenced genome of the malacosporean T. kitauei (belonging to the Thelohanellus genus of the myxosporean subgroup) with 150.7 mega-base pairs and containing 16,638 predicted protein-coding genes. These data were used to revisit the phylogenetic placement of Myxozoa in a phylogenetic analysis that combined the genomic data of various other taxa with that of T. kitauei. Additionally, comparative analyses of the biological characteristics among T. kitauei and others were also conducted at the level of the genome.

2. Materials and methods

2.1. T. kitauei genome and transcriptome data

The genome and transcriptome of T. kitauei were sequenced and annotated in our other work (submitted manuscript). The whole genome and transcriptome sequences of T. kitauei have been deposited in the National Center for Biotechnology Information (NCBI) under the project number PRJNA193083 (http://www.ncbi.nlm.nih.gov/bioproject/193083).

2.2. Data assembly

Using each of the 128 protein-encoding genes from 57 taxa (kindly provided by Nesnidal et al. (2013)) as queries, we conducted a BLAST (blastp) search against the genomic sequence data of T. kitauei and extracted the corresponding orthologs of each gene. Only hits with an E value less than 10^{-10} were retained, and the hit lists were filtered to discard the confounding paralogs. The hits were retained only if the score of the second best hit of each gene in a given taxon was less than half of the score of the best hit. For confirmation, the filtered hits were reciprocally BLAST searched against the GenBank non-redundant (nr) protein database. These confirmed hits were considered orthologs in the T. kitauei genome. Because the genomic data from the ctenophore Mnemiopsis leidyi had recently become available (Ryan et al., 2013), we used the same criteria as those used for T. kitauei to yield orthologs from the M. leidyi genome. Then, we combined the corresponding orthologous sequences in T. kitauei with each of the 128 protein-encoding genes from 57 taxa to construct a data set. Additionally, we replaced the EST-derived orthologous sequences with genome-based sequences of M. leidyi for the data set constructed above. To minimize the potential effects imposed by missing data, we assembled a trimmed data set by retaining only those sequences that were present in all three representative myxozoans and reevaluated the phylogenetic placement of Myxozoa.

2.3. Sequence alignment and phylogenetic analyses

Each gene set was aligned using MUSCLE version 3.6 software with the default parameters except for changing the output format from ClustalW to FASTA (Edgar, 2004), and the ambiguously aligned regions were detected and trimmed using Gblocks version 0.91b (Castresana, 2000) (-b2 = 0.65, -b3 = 10, -b4 = 5, -b5 = a). The 128 individual alignments of 58 taxa were concatenated into a supermatrix by SCAFOs version 4.42 software using the data set assembling panel (Roure et al., 2007).

The ML analysis was conducted using RAxML version 7.2.6 (Stamatakis, 2006) with an LG amino acid substitution matrix, frequencies empirically estimated (+F), and a Γ model of site heterogeneity with four categories (LG + F + Γ4). Bootstrap support for the ML analysis was evaluated with 100 replicates. Bayesian inference analysis was performed with PhyloBayes 3.3 (Lartillot et al., 2009) under a mixture model of CAT + POI + Γ4 with two independent Markov Chain Monte Carlo (MCMC) runs. The detailed parameters used to run PhyloBayes were as follow: a discrete gamma distribution of rate variation with four rate categories (Γ4), the relative exchange rates modeled by the Poisson process (POI), and the amino acid profiles estimated using the CAT model (CAT). The bpcmp program within PhyloBayes was used to determine whether the two independent MCMC runs reached convergence. The bpcmp program compares the discrepancy of bipartition frequencies between the two runs and outputs a consensus tree (PhyloBayes 3.3 manual). According to the PhyloBayes manual, a “good run” and adequate convergence of the two runs occurs when the largest discrepancy (maxdiff) in the bipartition frequencies between the two runs is less than 0.1. In this study, the maxdiff equaled 0.0370.

The alternative hypothesis that Myxozoa is a sister group to Bilateria (Evans et al., 2010; Kim et al., 1999) was previously proposed. To test this hypothesis, we manually created constraint tree topologies with reference to an alternative hypothesis using Mesquite version 2.75 software (Maddison and Maddison, 2011) and then performed the RAxML analysis with each constraint using the -g option and LG + F + Γ4 model. The optimal ML tree was
considered the constrained ML tree. The site likelihood values for both the constrained and unconstrained ML trees were computed using RAxML, and the obtained likelihood values were analyzed using the approximately unbiased (AU) and Shimodaira–Hasegawa (SH) tests with CONSEL v.0.1j software (Shimodaira and Hasegawa, 2001).

2.4. Identification and characterization of minicollagen genes

All known minicollagen proteins were collected from the NCBI protein database and a previous study (Holland et al., 2010). The proteins were aligned and a profile hidden Markov model (HMM) was built with hmmbuild to search against the predicted proteome of _T. kitauei_ with hmmsearch using the HMMER 3.0 software package (Finn et al., 2011). In addition, the _T. kitauei_ genome sequences were searched using the tblastn algorithm with sequences of all known minicollagen proteins as queries. The E-value cutoff was set at 0.001. The hits were assessed in the online Pfam database (http://pfam.sanger.ac.uk/) (Punta et al., 2012), and only those hits containing the characteristic domain structures of the known minicollagen proteins were considered candidate homologs in _T. kitauei_.

The putative full-length sequences of candidate _T. kitauei_ minicollagen homologs were obtained by comparing the _T. kitauei_ genome sequences with the transcriptome data. The sequences were then aligned with the known minicollagen protein sequences using MEGA version 5.2 software (Tamura et al., 2011). The putative signal peptides were predicted using SignalP version 3.0 software (Bendtsen et al., 2004). The three-dimensional models of the N- and C-terminal cysteine rich domains (CRDs) of the _T. kitauei_ minicollagen homologs were constructed following the homolog modeling procedures described in a previous study (Holland et al., 2010). The candidate genes possessing the sequence characteristics of known minicollagen proteins were considered minicollagen homologs in the _T. kitauei_ genome. To evaluate the relationship between myxozoan minicollagen homologs and the known cnidian-specific minicollagens, phylogenetic analyses were conducted using the neighbor-joining method with MEGA version 5.2 (Tamura et al., 2011), the maximum-likelihood analysis with PhyML version 3.0 (Guindon and Gascuel, 2003), and Bayesian inference with MrBayes version 3.2 software (Ronquist and Huelsenbeck, 2003).

2.5. Identification and comparison of proto-mesodermal genes

Protein sequences in _Homo sapiens_ from the proto-mesodermal genes were collected from the NCBI database and used as queries to reciprocally search with blastp and tblastn against the _T. kitauei_ predicted proteome and genome, respectively. The E-value cutoff was set at 0.001. The domain architecture of candidate protein sequences was assessed in the online Pfam database (http://pfam.sanger.ac.uk/) (Punta et al., 2012), and only those containing the characteristic domain structures of the human proteins were considered corresponding homologs in the _T. kitauei_ genome. Homologs to human mesodermal genes in representatives of other metazoa and Choanoflagellata have been previously reported (Ryan et al., 2013); those were collected and the components of their proto-mesodermal genes compared with those of _T. kitauei_.

3. Results

3.1. Phylogenetic position of Myxozoa inferred from new phylogenomic analyses

In our newly sequenced genomic data of myxosporean _T. kitauei_, 86 orthologs of the 128 protein-coding genes selected by Philippe et al. (Philippe et al., 2009) were detected, and they were incorporated into the supermatrix comprising the 128 proteins from 57 taxa used by Nesnidal et al. (2013). Moreover, orthologs of all 128 protein-coding genes found in the recently released genomic data from the ctenophore _M. leidyi_ (Ryan et al., 2013) were observed in the present study. However, only 118 orthologs of the 128 proteins were detected in the EST-derived data (Nesnidal et al., 2013). Thus, in the present study, the 118 EST-derived orthologs were replaced by the 128 genome-based orthologous sequences of ctenophore _M. leidyi_ in a data set comprising 31,647 ambiguously aligned amino acid sites of 128 proteins for 58 taxa (including the three myxozoan representatives that were most current). Our phylogenomic analyses provided strong support for the phylogenetic position of monophyletic Myxozoa as a sister group to Medusozoa within Cnidaria based on the concatenated data set (Fig. 1). In the ML tree calculated with RAxML under the LG + F + F4 model, myxozoans were recovered into a clade (_Buddenbrockia, _Theolhanellus, _Myxobolus_) with maximal bootstrap support (BS = 100%), and the monophyletic Myxozoa formed a sister group to Medusozoa within Cnidaria with a strong support value (BS = 98%) (Fig. 1). In contrast to the robust support of the monophyletic Myxozoa within Medusozoa within Cnidaria, Placozoa was reconstructed as a derived member of Porifera with extremely weak bootstrap support (Fig. 1). In addition, the AU likelihood-based tests significantly rejected the proposed alternative hypotheses in which Myxozoa was a sister group to Bilateria (Table 1). In concordance with the results of the ML tree, the Bayesian inference analysis with the CAT mixture model yielded the same tree topologies with high support values, depicting monophyletic Myxozoa as a sister group to Medusozoa within Cnidaria (Supplementary data S1). In addition, maximal Bayesian posterior probabilities (PP = 1.0) were obtained for the monophyletic Myxozoa and a sister group relationship between Myxozoa and Medusozoa within Cnidaria (Supplementary data S1).

To minimize any potential effects imposed by missing data, we pruned the missing data and retained 7837 unambiguously aligned amino acid alignment positions representing only those sequences that were present in all three representative myxozoans. The results of the phylogenetic analyses of the trimmed data set also revealed that the monophyletic Myxozoa was a sister group to Medusozoa within Cnidaria using both ML tree and Bayesian analyses (BS = 79%; PP = 0.99) (Supplementary data S2), consistent with the results from the concatenated data set with multiple missing data.

Collectively, the phylogenomic analyses robustly supported the monophyly of Myxozoa and Myxozoa as a cnidian taxon sister group to Medusozoa.

3.2. Identification of two minicollagens in _T. kitauei_

Two putative homologs to cnidian-specific minicollagens were detected in the myxosporean _T. kitauei_ genome (designated _Tk_Ncol-1_ and _Tk_Ncol-2_), and their corresponding ESTs were observed in the transcriptome data for _T. kitauei_. Through comparing the transcribed minicollagen sequences with _T. kitauei_ genomic sequences, the full-length sequence of the _Tk_Ncol-1_ gene and the nearly full-length sequence of the _Tk_Ncol-2_ gene were obtained. The _Tk_Ncol-1_ gene consisted of a 481-base pair (bp) genomic DNA sequence containing an intron 22 bp in size and could be translated into an open reading frame (ORF) with 152 amino acid residues (Supplementary data S3a). The obtained _Tk_Ncol-2_ gene contained a 426-bp genomic DNA fragment possessing a 22-bp intron without a stop codon because the fragmented gene was immediately followed by a long fragment (1,880 bp in length) of an unknown sequence at its 3’ terminal. Even without a stop codon, the partial _Tk_Ncol-2_ could be translated into 134 amino
acid residues possessing all the domain structures of minicollagens (Supplementary data S3b). In addition, we found that the intron–exon boundaries in both the Tk_Ncol-1 and Tk_Ncol-2 genes conformed to the known GT/AG donor/acceptor site rule (Senapathy et al., 1990).

Both Tk_Ncol-1 and Tk_Ncol-2 possess diagnostic domain structures of the cnidarian-specific group 1 minicollagens (Fig. 2), including a central collagen-like domain of 14 repeated Gly-X-Y units flanked by two polyproline sequences of 22 residues (N-terminal) and three residues (C-terminal) in Tk_Ncol-1, and by seven residues (N-terminal) and three residues (C-terminal) in Tk_Ncol-2. The N- and C-terminals of the polyproline sequences contained CRDs with six cysteine residues arranged as CxxxCxxxCxxxCxxxCC in both Tk_Ncol-1 and Tk_Ncol-2. Additionally, signal peptide cleavage sites were present in sequences of both Tk_Ncol-1 and Tk_Ncol-2.

Fig. 1. Maximum likelihood tree topology from analysis of a phylogenomic matrix comprising 31,647 unambiguously aligned amino acid residues for 58 taxa inferred using RAxML software under an LG + F + Γ4 model. Bootstrap values from 100 replicates are given on the nodes. Scale bar, substitutions per position.
Table 1
Comparisons of alternative phylogenetic hypotheses regarding the phylogenetic position of Myxozoa.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Phylogenetic hypothesis</th>
<th>Likelihood</th>
<th>A/Likelihood</th>
<th>AU</th>
<th>SH</th>
</tr>
</thead>
<tbody>
<tr>
<td>T0</td>
<td>ML tree (Fig. 1)</td>
<td>−576.211</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>T1</td>
<td>Myxozoa sister group to Bilateria</td>
<td>−576.402</td>
<td>191</td>
<td>4e−11</td>
<td>3e–04</td>
</tr>
</tbody>
</table>

Tree topologies can be rejected when the p-values of the approximately unbiased (AU) and Shimodaira–Hasegawa (SH) tests are lower than 0.05.

Fig. 2. Multiple sequence alignment of minicollagens. Gaps are indicated by dashes. The conserved KR endopeptidase cleavage site, the cysteine arrangement in the N- and C-terminal CRDs, and the non-helical alanines are highlighted. Glycine residues in the collagen-like domain are bolded. Abbreviations: Tk, *Thelohanellus kitauei*; Tb, *Tetracapsuloides bryosalmonae*; Hm, *Hydra magnipapillata*; Ch, *Clytia hemisperica*; He, *Hydractinia eichinata*; Mk, *Malo kingi*; Ho, *Hydra oligactis*; Pc, *Podocoryne carnea*; Ms, *Metridium senile*; Nv, *Nematostella vectensis*.
the conserved endopeptidase cleavage sites KR and KK, respectively (Supplementary data S3). In Tk_Ncol-1, exon 1 contained the signal peptide, exon II was the mature minicollagen peptide, and the intron interrupted the propeptide (Supplementary data S3a), whereas in Tk_Ncol-2, the intron interrupted the signal peptide (Supplementary data S3b). Both Tk_Ncol-1 and Tk_Ncol-2 were transcribed, and the scores for their Fragments Per Kilobase of exon model per Million mapped fragments (FPKM) were 64.25 and 9.88, respectively, indicating that both genes were highly expressed in T. kitauei. The CRDs of both Tk_Ncol-1 and Tk_Ncol-2 possessed the same canonical cysteine sequence pattern as that in other known minicollagens, although the side chains were variable. The three-dimensional structures of the CRDs showed that the positions of the cysteines and their side chains in both Tk_Ncol-1 and Tk_Ncol-2 were similar to those in Hydra minicollagen (Hm-Ncol-1) (Supplementary data S4), indicating that the disulphide arrangements in the T. kitauei homologs were the same as those in the cnidarian minicollagen. Taken together, these characteristics of Tk_Ncol-1 and Tk_Ncol-2 indicated that they were functional homologs to cnidarian minicollagens.

Phylogenetic analyses recovered the two minicollagen homologs in T. kitauei together with the known cnidarian-specific group 1 minicollagens into three groups (Supplementary data S5); Group A consisted of the two identified minicollagen homologs in T. kitauei and all known medusozoan minicollagens harboring a single collagen-like domain with 14 Gly-X-Y repeats. Group B was composed of the malacosporean homolog Tb_Ncol-1 and all medusozoan minicollagens with a double collagen-like domain having 14 plus 13 Gly-X-Y repeats. Group C comprised all anthozoan minicollagens possessing a single or double collagen-like domain. Surprisingly, the myxosporean homologs (Tk_Ncol-1 and Tk_Ncol-2) and the malacosporean homolog Tb_Ncol-1 did not cluster into the same group, although both myxosporean and malacosporean homologs were close to medusozoan minicollagens.

3.3. Identification of proto-mesodermal genes in T. kitauei and comparative analyses

Two prevalent hypotheses regarding the phylogenetic position of Myxozoa within Metazoa have been put forward (Evans et al., 2010). One hypothesis placed Myxozoa within Cnidaria, suggesting a diploblastic nature of Myxozoa. The alternative hypothesis posited Myxozoa as a sister taxon to Bilateria, suggesting a triploblastic nature of Myxozoa. Therefore, the diploblastic versus triploblastic nature of Myxozoa is under debate. Generally, the triploblastic bilaterians possess three germ layers (endoderm, ectoderm and mesoderm), and the diploblastic non-bilaterians have only two germ layers (endoderm and ectoderm). Thus, the mesoderm is exclusively present in Bilateria and evolved in the bilaterian lineage. Although Ctenophora and Cnidaria are commonly recognized as diploblastic non-bilaterians, some of the proto-mesodermal genes were previously observed in them (Martindale et al., 2004; Ryan et al., 2013). With a myxozoan genome in hand, we investigated the proto-mesodermal genes in the myxozoan T. kitauei and explored the phylogenetic position of Myxozoa by comparing the phylogenetic distribution of proto-mesodermal genes among T. kitauei and representative organisms from the five metazoan lineages (Ctenophora, Porifera, Placozoa, Cnidaria and Bilateria) and Chaoanoflagellata. Eleven (Brachury, Snail, gli, glis, Timm, Nkx2.6, nk2.1/NKX21, Bagpipe/NKX32, Csc, Forkhead/HNF3 FoxA/group 1, Ms2, GATA and LIM/CSR3P) out of the 19 human proto-mesodermal genes were identified in the T. kitauei genome (Table 2). The phylogenetic distributions of the proto-mesodermal genes in representatives of other metazoans and Chaoanoflagellata were summarized in a previous study (Ryan et al., 2013). Here, we compared the phylogenetic distributions of proto-mesodermal genes in T. kitauei with representatives of other metazoans and Chaoanoflagellata. We found that only four (Eomesoderm/TBR2, Troponin C, Troponin T and Troponin T) out of the 19 proto-mesodermal genes were exclusively present in Bilateria; six (Brachury, Snail, Gli, Me2, GATA and LIM) out of the 19 proto-mesodermal genes were widely spread in Metazoa, except that Snail was absent in Porifera and both Gli and LIM were absent in Placozoa; four (Twist, Myf5, Ladybird and Pax3) out of the 19 proto-mesodermal genes were present only in Bilateria and in one representative cnidian Nemastostella vectensis; the remaining five genes (Tinman/Nkx2.6, nk2.1/NKX21, Bagpipe/ NKX32, Csc and Forkhead/HNF3 FoxA/group 1) were present in Bilateria and many non-bilaterian lineages, but were completely absent in Ctenophora and Chaoanoflagellata. Thus, the number of proto-mesodermal genes in T. kitauei was greater than that observed in Ctenophora, Ctenophora, Porifera, and Placozoa, but less than that found in Bilateria and the representative cnidian N. vectensis. Importantly, all components of the proto-mesodermal genes detected in T. kitauei were present in the representative cnidian H. magnipapillata (Table 2).

In addition to investigating the phylogenetic distribution, we also conducted a preliminary phylogenetic analysis of the 19 proto-mesodermal genes (data not shown). Notably, in the majority of the gene trees containing homologous sequences of representatives of both Cnidaria and Myxozoa, we found that the anthozoan cnidarians (represented by N. vectensis and Acropora digitifera) were not grouped with the medusozoan cnidian H. magnipapillata, and that H. magnipapillata was grouped with the myxozoan T. kitauei.

4. Discussion

4.1. Myxozoa is sister to Medusozoa within Cnidaria according to a new Phylogenomic analysis

Previous phylogenetic studies on metazoan phylogeny have failed to obtain a clear consensus on the phylogenetic position of myxozoans, with the various studies often yielding ambiguous and sometimes conflicting results. This discord may likely have occurred because the myxozoan genes have undergone a rapid evolution (Brinkmann and Philippe, 2008), resulting in an intrinsically difficult phylogenetic problem. Previous phylogenetic studies addressing this evolutionary issue used either small data sets based on rDNA genes (Evans et al., 2010; Kim et al., 1999; Smothers et al., 1994) or multiprotein data sets with only one or two representatives of Myxozoa (Jimenez-Guri et al., 2007; Nesnidal et al., 2013). Phylogenomic analyses that used data from large portions of or from entire genomes obtained in various taxa to infer phylogenetic relationships have resolved many previously debated phylogenies (Chiari et al., 2012; Hampl et al., 2009; Liang et al., 2013; Philippe et al., 2011; Pick et al., 2010; Ryan et al., 2013) and have shown promise for resolving the phylogenetic position of Myxozoa. However, the phylogenomic analysis of Myxozoa was impeded by a lack of adequate representative genomic data. The phylogenetic position of Myxozoa was explored by phylogenomic analysis after obtaining the first genomic data from the myxosporean M. cerebralis (Nesnidal et al., 2013). In the present study, after obtaining genomic data from another myxosporean T. kitauei (an important pathogen), we revisited the phylogenetic position of Myxozoa by reconstructing phylogenetic trees using genomes from the largest number of myxozoans to date.

Myxozoa was consistently recovered as a monophyle with maximum support value in all our trees. This result is consistent with several previous studies investigating rDNA sequences (Evans
Table 2

Identification of the proto-mesodermal genes in T. kitauei and comparative analyses of their distributions in Choanoflagellata and metazoans.

<table>
<thead>
<tr>
<th>Bila</th>
<th>Cnidaria</th>
<th>Myxozoa</th>
<th>Plac</th>
<th>Pori</th>
<th>Cten</th>
<th>Choá</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hs</td>
<td>Nv</td>
<td>Hm</td>
<td>Tk</td>
<td>Ta</td>
<td>Aq</td>
<td>Ml</td>
</tr>
<tr>
<td>Twot</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Myogenin/Myf5</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Brachyury</td>
<td>√</td>
<td>√</td>
<td>evm.model.scaffold01439.5</td>
<td>√</td>
<td>√(−)</td>
<td>√</td>
</tr>
<tr>
<td>Snail</td>
<td>√</td>
<td>√</td>
<td>evm.model.scaffold04751.3</td>
<td>√</td>
<td>–</td>
<td>√(−)</td>
</tr>
<tr>
<td>Eomesodermin/TBR2</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>gl/gl</td>
<td>√</td>
<td>√</td>
<td>evm.model.scaffold01581.11</td>
<td>√</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Unnan/Nkx2.6</td>
<td>√</td>
<td>√</td>
<td>evm.model.scaffold04913.8</td>
<td>√</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>nkn2.1/NKX21</td>
<td>√</td>
<td>√</td>
<td>evm.model.scaffold04913.8</td>
<td>√</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Bagpipe/NKX22</td>
<td>√</td>
<td>√</td>
<td>evm.model.scaffold04913.8</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ladybird</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Pax3</td>
<td>√</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Gsc</td>
<td>√</td>
<td>√</td>
<td>evm.model.scaffold04913.87</td>
<td>√</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Forkhead/HNF1FoxA/group 1</td>
<td>√</td>
<td>√(−)</td>
<td>evm.model.scaffold04470.9</td>
<td>√</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Mel2</td>
<td>√</td>
<td>√</td>
<td>evm.model.scaffold03170.40</td>
<td>√</td>
<td>–</td>
<td>√(−)</td>
</tr>
<tr>
<td>GATA</td>
<td>√</td>
<td>√</td>
<td>evm.model.scaffold02630.54</td>
<td>√</td>
<td>√(−)</td>
<td>–</td>
</tr>
<tr>
<td>LIM/CSR2</td>
<td>√</td>
<td>√</td>
<td>evm.model.scaffold03330.3</td>
<td>–</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Troponin T</td>
<td>√</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Troponin I</td>
<td>–</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Troponin C</td>
<td>√</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Gene identifiers are given in T. kitauei. √: presence; √(−): not present in the representative species but present in other species in a given group; –: absence. Abbreviations: HS, Homo sapiens; Nv, Nematomastella vectensis; Hm, Hydra magnipapillata; Tk, Thelohanellus kitauei; Ta, Trichoplax adhaerens; Aq, Amphimedon queenslandica; Ml, Mnemiopsis leidyi; Mb, Monosiga brevicollis; Bila, Bilateria; Plac, Placozoa; Pori, Porifera; Cten, Ctenophora; Choá, Choanoflagellata.

4.2. Evidence that Myxozoa is Cnidaria using cnidarian-specific minicollagen

Minicollagens are cnidarian-specific genes encoding key nematocyst proteins (Kurz et al., 1991). It has long been known that minicollagens are present exclusively in cnidarians. Polar capsules are intracellular organelles found in all myxozoans and are similar to the cnidarian nematocysts in structure and function (Weill, 1938). In the present work, two actively transcribed homologs to cnidarian-specific minicollagens were observed in the T. kitauei genomic data, and their sequence characteristics and gene organization patterns were similar to those of cnidarian-specific minicollagens, indicating that homologs to cnidarian-specific minicollagens are present in myxosporean T. kitauei. The CRDs of the cnidarian minicollagens are thought to play pivotal roles in forming the stable capsule wall of nematocysts by covalently cross-linking with similar CRDs through disulphide bonds to form a stable network. Herein, both Tk_Ncol-1 and Tk_Ncol-2 possessed canonical cysteine sequence patterns in their N- and C-terminal CRDs. In addition, the three-dimensional structures of these CRDs showed that the positions of the cysteines and their side chains were similar to those of Hydra minicollagen Hm-Ncol-1, suggesting that the disulphide bond arrangements in T. kitauei minicollagens are the same as those in Hm-Ncol-1. Thus, the minicollagen homologs in T. kitauei would be involved in forming structures similar to those of the cnidarian nematocysts. Based on these results together with the evidence that the minicollagen homologue Tb_Ncol-1 is present in malacosporean T. bryosalmonae (Holland et al., 2010), we speculate that homologs to the cnidarian-specific minicollagen are present in all myxozoans and that they may be key constituents of the myxozoan polar capsule in a manner similar to the minicollagens being key proteins in cnidarian nematocysts. The testing of this hypothesis awaits further experimental evidence. Using available genomic data, we investigated the existence of homologs of cnidarian-specific minicollagens in...
many other eukaryotes; however, no homologs were found. Therefore, minicollagens currently appear restricted to cnidarians and myxozoa, offering further corroborating evidence that Myxozoa is Cnidaria. Furthermore, our phylogenetic analyses of minicollagens revealed that although minicollagen homologs in myxosporean and malacosporean have diverged, both show a close relationship to medusozoan minicollagens, suggesting that Myxozoa is close to Medusozoa within Cnidaria. This conclusion is consistent with a previous study that examined the ultrastructure of B. plumatellae (Gruhl and Okamura, 2012).

Tk_Ncol-1 and Tk_Ncol-2 recovered as a clade that was clustered with the group A minicollagens comprising medusozoan minicollagens, which possessed a single collagen-like domain with 14 Gly-X-Y repeats. These results were consistent with the sequence characteristics of both Tk_Ncol-1 and Tk_Ncol-2, which possessed single 14 Gly-X-Y repeats. However, the malacosporean minicollagen homolog Tb_Ncol-1 clustered with the group B minicollagens, which encompassed medusozoan minicollagens harboring a double collagen-like domain with 14 plus 13 Gly-X-Y repeats. These results suggested that both group A and group B minicollagens emerged in the common ancestors of Medusozoa and Myxozoa. Then, after the split of Malacosporea and Myxosporea, one of the two groups was retained in Myxosporea and one in Malacosporea. However, our phylogenetic analyses of minicollagens were restricted by the paucity of available minicollagen sequences, and these were dominated by hydrozoan sequences. Additionally, a complete set of minicollagen homologs in the malacosporean T. bryosalmonae has not yet been revealed because the whole genome sequence is unavailable. Therefore, greater resolution of the relationships among medusozoan minicollagens and minicollagen homologs in the two subgroups of Myxozoa will require increased representation of these genes from the various cnidarians and increased sampling from the whole genome data for myxozoans.

Collectively, our results identifying cnidarian-specific nematocyst minicollagen homologs in the T. kitauei genome and our comparative analyses between minicollagen homologs in Myxozoa and cnidian minicollagens further corroborate Myxozoa is Cnidaria and a close relative of Medusozoa.

4.3. Evidence that Myxozoa has an affinity with Cnidaria using comparative genomic analyses of proto-mesodermal genes

Generally, bilaterians are triploblastic animals with three germ layers (endoderm, ectoderm and mesoderm), and cnidarians are diploblastic animals with two germ layers (endoderm and ectoderm). Thus, resolving the controversy over the diploblastic versus triploblastic nature of Myxozoa would provide insight to explore the phylogenetic position of Myxozoa within Metazoa. The mesoderm is the embryonic germ layer developing between the endoderm and the ectoderm, and it is considered a pivotal innovation in Bilateria. However, previous studies have revealed that some of the proto-mesodermal genes are observed in non-bilaterian animals (Ryan et al., 2013) and that the mesoderm evolved from the endoderm (Martindale et al., 2004). In the present study, we found that the majority of proto-mesodermal genes were detected in the T. kitauei genome. Comparative analyses revealed that the number of proto-mesodermal genes in T. kitauei was larger than that in representatives of Chaoanellagellata, Ctenophora, Porifera and Placozoa but smaller than that in Bilateria and the cnidian N. vectensis. However, all of the components of the proto-mesodermal genes that were detected in T. kitauei were present in H. magnipapillata, indicating that Myxozoa has an affinity to Hydra within Cnidaria and supporting the hypothesis that Myxozoa is cnidian. Four genes (Twist, Myf5, Ladybird and Pax3) were present in N. vectensis but not in H. magnipapillata and T. kitauei, indicating a secondary loss of proto-mesodermal genes in H. magnipapillata and T. kitauei relative to N. vectensis. Previous studies revealed that Twist is a positive regulator of mesodermal development in Drosophila (Leptin, 1991), Pax3 and Myf5 are involved in regulatory cascade functions at the onset of myogenesis (Sato et al., 2010), and Ladybird participates in the determination of heart lineages and is required to specify the identities of subpopulations of heart cells (Jagla et al., 1997). In addition to epitheliomuscular cells, there are greater degrees of muscle cell type specializations in N. vectensis than in H. magnipapillata (Chapman et al., 2010), and we postulated that the four genes present in N. vectensis but not in H. magnipapillata might be involved in the muscle cell type specialization in N. vectensis. This postulation awaits further experimental evidence.

5. Conclusions

In conclusion, both phylogenomic analyses using new genomic data from T. kitauei and comparative analyses of minicollagen genes and the proto-mesodermal genes robustly supported Myxozoa as a highly derived cnidarian taxon and sister to Medusozoa. These results will help us to understand the diversity of Cnidaria and to interpret the evolution of the distinct morphological, developmental, and molecular features of myxozoans within cnidarians. Furthermore, they may offer insights for controlling parasitosis in fish by comparative genomic analyses among myxozoans and the closely related free-living cnidarians.

Acknowledgments

The authors are indebted to Prof. Bernhard Hausdorf (Zoologisches Museum der Universitaet Hamburg) for providing the seed material. We are grateful to Prof. Peng Zhang (Sun Yat-Sen University) for his advice on conducting this work. This work was supported by grants from the National Natural Science Foundation of China (Grant 31372168), the National Science and Technology Support Program Project of China (Grants 2013BAD10B01-2 and 2012BAD25B0203) and the Special Fund for Basic Research of the Chinese Academy of Agricultural Sciences (Grant 2013ZL043).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.mpev.2014.08.016.

References

J.-M. Feng et al. / Molecular Phylogenetics and Evolution 81 (2014) 10–18


