Identification of two MMP genes from sea cucumber *Apostichopus japonicus*

Ting Miao

1. Introduction

Sea cucumbers (*Echinodermata, Holothuroidea*) are capable of regenerating damaged organs and body parts, such as the intestine, respiratory tree, gonads and the body wall. They can be induced to eviscerate intestines and other organs through the cloaca. Within a few weeks, the eviscerated organs are completely regenerated (Yang, Hamel, & Mercier, 2015). During animal regeneration, precursor cells, like pluripotent stem cells, migrate from the blastema to the wounded area (e.g., torn mesenteric edges) and proliferate and differentiate to replace wounded tissue (Suarezcastillo, Medinaortiz, Roiglopez, & Garciaarraras, 2004). Garcia-Arraras *et al.* (2011) reported that at early stages of sea cucumber intestine regeneration, cells for regeneration initially come from dedifferentiating mesothelium (Garcia-Arraras *et al*., 2011). Remodeling of the extracellular matrix (ECM) regulated by matrix metalloproteinases (MMPs) is important for organogenesis and tissue regeneration in terms of regulating cell migration (Brown & Badylak, 2014). In our previous studies, we found two MMP genes whose expression level were significantly upregulated in regenerative intestines of sea cucumber *Apostichopus japonicus* and cloned full-length cDNA of the two genes. However, it is hard to identify MMPs genes in sea cucumber due to limited information of MMPs sequences from Echinoderm and other lower level species. In the present study, the two target genes, *A. japonicus* 1 and *A. japonicus* 2, which were named as MMP-2 and MMP-16, were identified by interpreting and comparing the phylogenetic trees from maximum parsimony, distance and maximum likelihood analysis.

2. Methods

2.1 BLAST

The cloned cDNA sequences of *A. japonicus* 1 (GenBank Accession No. KX372219) and *A. japonicus* 2 (GenBank Accession No. KX372220) were analyzed using the respective National Center for Biotechnology Information (NCBI) BLAST programs (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.2 Multiple sequence alignments

28 amino acid sequences of MMPs from invertebrate and vertebrate species were obtained from NCBI and were aligned with amino acid sequences of *A. japonicus* 1 and *A. japonicus* 2 by MAFFT. The auto strategy was used for the alignment and the output was saved to both fasta and phylip formats for further analysis.

2.3 Maximum parsimony

The aligned file in 2.2 was converted into Nexus format and executed in PAUP. Heuristic search was used in maximum parsimony search. Parsimony bootstrap analysis was conducted with TBR branch swapping and 200 bootstrap replicates using

random stepwise addition sequence. The consensus tree was built according to majority rule.

2.4 Distance analysis

Distance analysis by using neighbor-joining (N-J) algorithm was done in PHYLIP. Two hundred replicates bootstrapping were set before creating the protein distance matrix. Jones-Taylor-Thornton (JTT) model was used in the protein distance algorithm with 200 data sets. 200 N-J trees were built according to each distance matrix and a single majority rule consensus tree was built in the end.

2.5 Likelihood analysis

Likelihood analysis was done in RAxML. The maximum likelihood trees from PROTCAT model with 100 bootstrap replicates were used to build the consensus tree.

3. Results

According to the BLAST result, all the sequences chosen shared at least 40% identity with target genes. To be more specific, *A. japonicus* 1 shared 49% identity with MMP-2 from *Strongylocentrotus purpuratus* (XP_780356.3) and 46% identity with two MMP-14 isoforms from *Acanthaster planci* (XP_022102156.1 and XP_022102157.1). *A. japonicus* 2 shared 54% identity with MMP-16 from *Strongylocentrotus purpuratus* (XP_781575.3) and 54% identity with MMP-14 from *Acanthaster planci* (XP_022110634.1). My target sequences together with other 28 MMPs amino acid sequences from other species were aligned in MAFFT using auto strategy (data not shown).

The best score of the trees found in PAUP 200 bootstrap parsimony heuristic search was 5358. In the consensus tree from parsimony analysis (**Fig. 1**), *S.purpuratus* MMP-2 was set as the outgroup. *A. japonicus* 1 was not clustered with other sequences, while *A. japonicus* 2 was on the same branch with *S. purpuratus* MMP-16, *A. planci* MMP-19, *S. purpuratus* MMP-18 and *S. purpuratus* MMP-14.

A rooted tree was built in distance analysis by using N-J algorithm (**Fig. 2**). *A. japonicus* 1 and *S.purpuratus* MMP-2 were set as outgroup. *A. japonicus* 2 was clustered with *A. planci* MMP-19, *S. purpuratus* MMP-16, *S. purpuratus* MMP-18 and *S. purpuratus* MMP-14. The similar result was shown in maximum likelihood tree (**Fig. 3**).

4. Discussion

All of the sequences in the present study share more than 40% identity and have representative protein domains for MMP family proteins like HX, ZnMc_MMP and peptidase_M10 conserved domain, so those sequences are homologous. The trees from distance and likelihood analysis are pretty similar. According to the trees, *A. japonicus* 1 has a relatively far distant of relationship with other MMPs included in this study and has the most recent common ancestor with *S.purpuratus* MMP-2. *A. japonicus* 2 clusters with MMPs from other echinodermata animals. A strong conclusion based on these analyses is hard to reach, however, it is likely that *A.*

japonicus 1 could be MMP-2 according to both the multiple sequence alignment and the trees. *A. japonicus* 2 shares the same identity with MMP-16 from *S.purpuratus* and MMP-14 from *A. planci* according to the alignment. *S.purpuratus* MMP-16 is grouped with *A. japonicus* 2, while *A. planci MMP-14* is in a relatively far distance from *A. japonicus* 2 according to the trees. Thus, *A. japonicus* 2 could be MMP-16 based on analyses in the present study.

However, almost all of the echinoderm MMPs sequences uploaded to NCBI database were from genome sequences and most of them were not resequenced. Moreover, the annotations of these genes may be not correct since they don't share a high identity with corresponding MMP sequences in the database according to alignments (data not shown). Thus, although the conclusion of the present study is in consistent with the original identification of the two genes in NCBI, the proof is not solid enough. Further studies on reannotating echinoderm sequences is required for a more accurate identification of those genes.

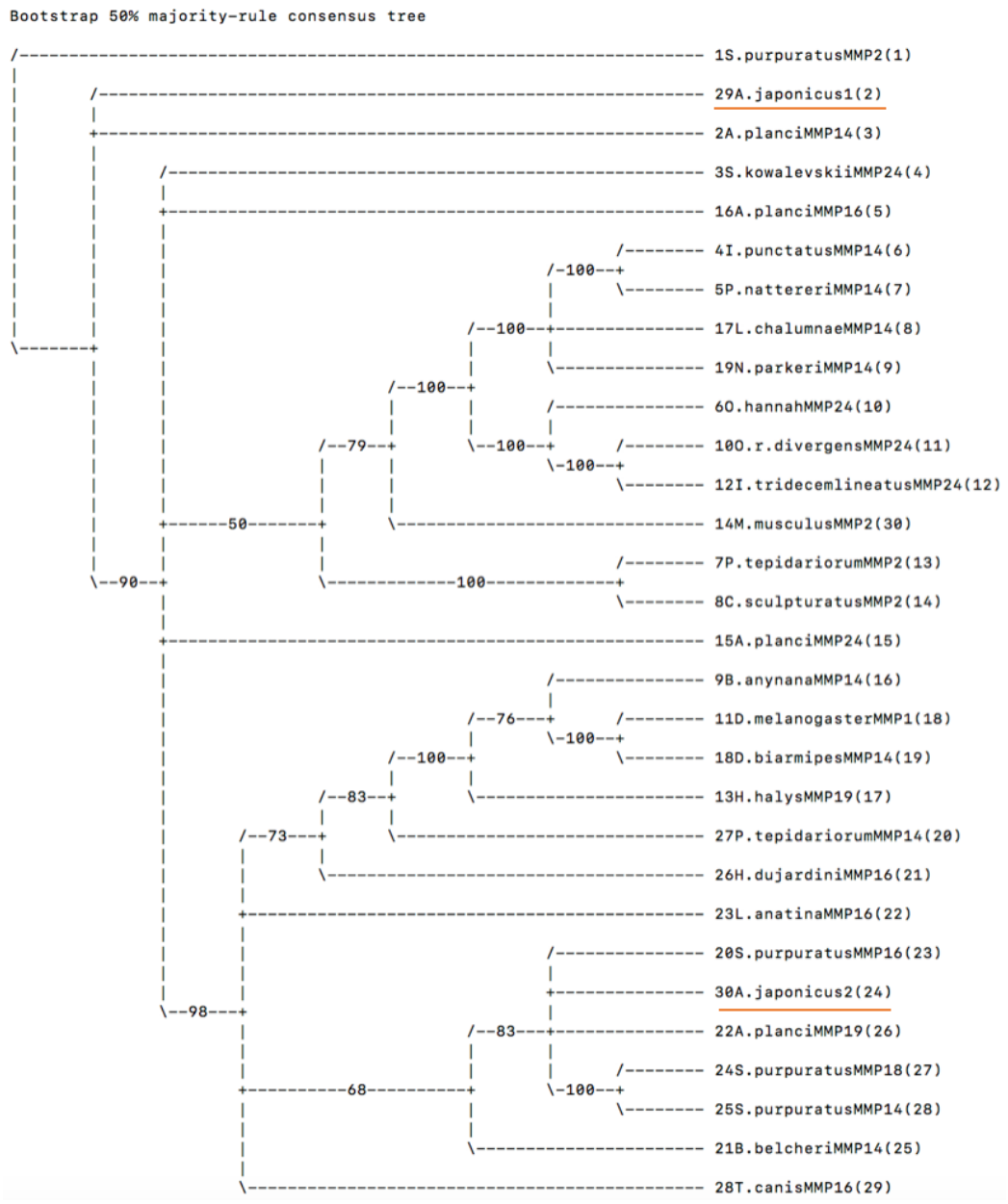

Fig.1Bootstrap 50% majority-rule consensus tree of parsimony heuristic search

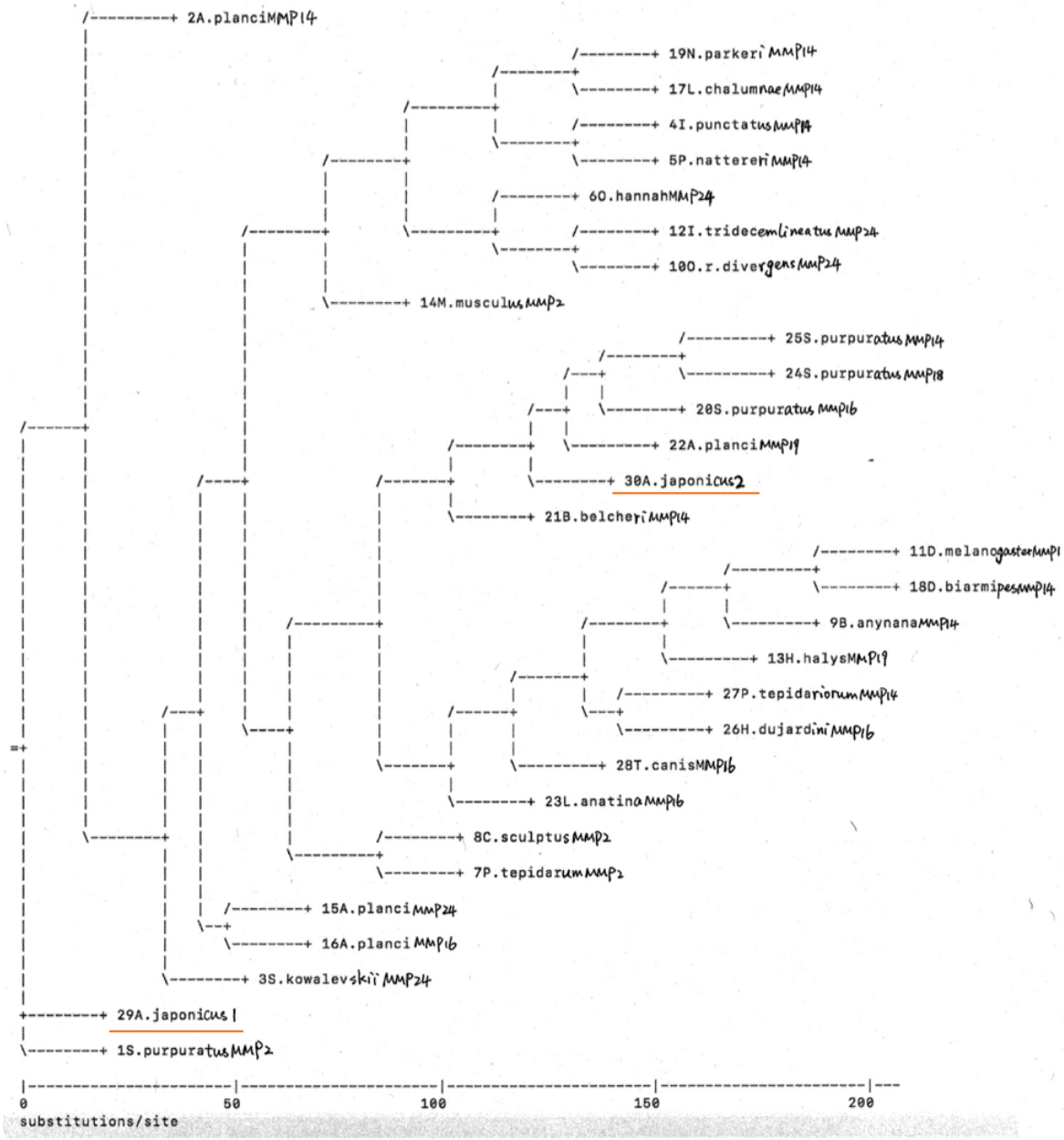

Fig.2 Neighbor-joining consensus tree

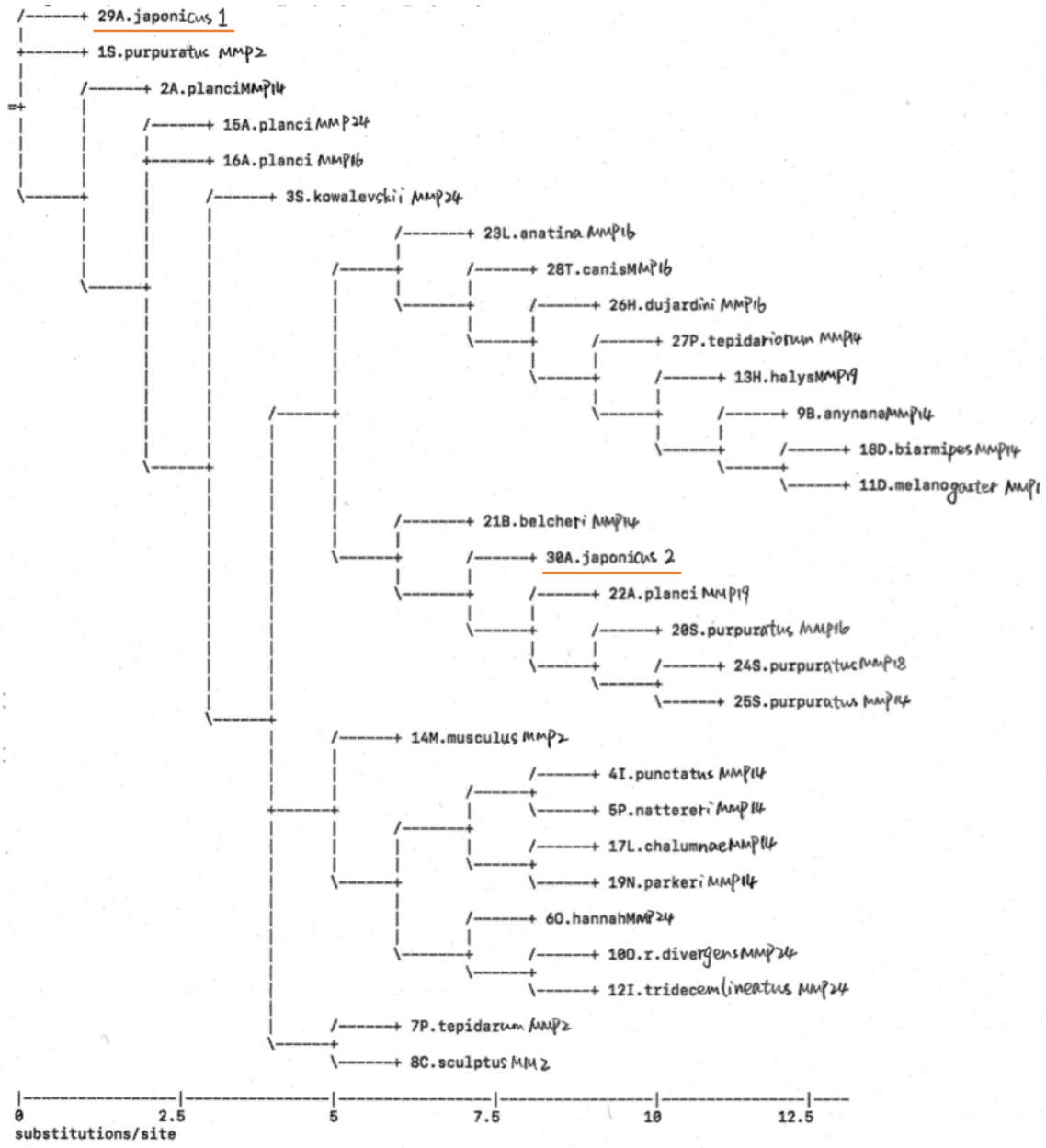

Fig.3 Maximum likelihood consensus tree

References:

Yang, H., Hamel, J.-F., & Mercier, A. (2015). Chapter 11 Aestivation and regeneration. In: *The sea cucumber Apostichopus japonicus: history, biology and aquaculture* (Vol. 39): Academic Press.

Suarezcastillo, E., Medinaortiz, W., Roiglopez, J., & Garciaarraras, J. (2004). Ependymin, a gene involved in regeneration and neuroplasticity in vertebrates, is overexpressed during regeneration in the echinoderm Holothuria glaberrima. *Gene, 334*, 133-143.

Garcia-Arraras, J. E., Valentin-Tirado, G., Flores, J. E., Rosa, R. J., Rivera-Cruz, A., San Miguel-Ruiz, J. E., & Tossas, K. (2011). Cell dedifferentiation and epithelial to mesenchymal transitions during intestinal regeneration in H. glaberrima. *BMC Dev Biol, 11*, 61. doi:10.1186/1471-213X-11-61

Brown, B. N., & Badylak, S. F. (2014). Extracellular matrix as an inductive scaffold for functional tissue reconstruction. *Translational Research, 163*(4), 268-285.