



# Origins and Evolution of the $\alpha$ -L-Fucosidases: From Bacteria to Metazoans

Jia You<sup>1</sup>, Shujin Lin<sup>2</sup> and Tao Jiang<sup>3\*</sup>

<sup>1</sup> Department of Hepatology, The Liver Center, The First Affiliated Hospital of Fujian Medical University, Fuzhou, China, <sup>2</sup> College of Biological Science and Engineering, Fuzhou University, Fuzhou, China, <sup>3</sup> Department of Urology, The First Affiliated Hospital of Fujian Medical University, Fuzhou, China

## OPEN ACCESS

### Edited by:

Baolei Jia,  
Chung-Ang University, South Korea

### Reviewed by:

Maximiliano Juri Ayub,  
National University of San Luis,  
Argentina

Cong Chen,  
Chinese Academy of Medical  
Sciences, Peking Union Medical  
College, China

Xianlin Han,  
Peking Union Medical College  
Hospital (CAMS), China

### \*Correspondence:

Tao Jiang  
jiangtao2607@fjmu.edu.cn

### Specialty section:

This article was submitted to  
Evolutionary and Genomic  
Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 16 April 2019

**Accepted:** 16 July 2019

**Published:** 27 August 2019

### Citation:

You J, Lin S and Jiang T (2019)  
Origins and Evolution of the  
 $\alpha$ -L-Fucosidases: From Bacteria  
to Metazoans.  
Front. Microbiol. 10:1756.  
doi: 10.3389/fmicb.2019.01756

$\alpha$ -L-fucosidases (EC 3.2.1.51, FUC), belonging to the glycoside hydrolase family 29 (GH29), play important roles in several biological processes and are markers used for detecting hepatocellular carcinoma. In this study, a protein sequence similarity network (SSN) was generated and a subsequent evolutionary analysis was performed to understand the enzymes comprehensively. The SSN indicated that the proteins in the FUC family are mainly present in bacteria, fungi, metazoans, plants, as well as in archaea, but less abundantly. The sequences in bacteria were found to be more diverse than those in other taxonomic groups. The SSN and a phylogenetic tree both supported that the proteins in the FUC family can be classified into 3 subfamilies. FUCs in each subfamily are under the pressure of negative selection. The enzymes from metazoans, fungi, and plants separated into the three subfamilies and shared high similarity with the bacterial homologs. The multiple sequence alignment results indicated that the amino acid residues for binding  $\alpha$ -L-fucosidase and catalysis are highly conserved in the 3 subfamilies; however, the evolutionary patterns were different, based on the coevolution analysis in the subfamily of metazoans and bacteria. Finally, gene duplication plays an important role for  $\alpha$ -L-fucosidase evolution, not only in metazoans, but also in bacteria and fungi.

**Keywords:**  $\alpha$ -L-fucosidase, evolution, sequence similarity network, bacteria, metazoa

## INTRODUCTION

Fucose is a hexose deoxy sugar with a relatively low abundance in the biosphere. Fucoses are usually attached to oligosaccharides, oligolipids, and other glycoconjugates by  $\alpha$ -1,3 linkages to glucose (Glc),  $\alpha$ -1,2 linkages to galactose (Gal), or 1,3/4/6 linkages to N-acetylglucosamine (GlcNAc) (Becker and Lowe, 2003).  $\alpha$ -L-fucosidase (FUC) (EC 3.2.1.51) is a group of glycoside hydrolase (EC 3.2.1) that specifically catalyzes the reaction to remove the Non-reducing terminal L-fucose; recently, an FUC from the bacterium *Elizabethkingia meningoseptica* was found to have catalytic activity against the substrates with core  $\alpha$ -1,3-fucosylation (Moreti et al., 2013; Li et al., 2018). FUC can be classified into two different glycoside hydrolase families: FUCs employing a retaining mechanism fall under the glycoside hydrolase family 29 (GH29), while those employing an inverting mechanism fall under the glycoside hydrolase family 95 (GH95) (Katayama et al., 2004). FUCs in the GH29 family are widespread in bacteria, fungi, molluscs, ascidians, and mammals

(Intra et al., 2007). GH29 FUCs are well-studied because of the biological importance of L-fucose and the fucosylated conjugates in several critical biological processes such as the immune response, early embryogenesis and development, signal transduction, adhesion of pathogens, apoptosis, and extravasation of leukocytes (Intra et al., 2007). For example, fucose and FUC play important roles in gamete interactions in bull, ascidians, and molluscs (Matsumoto et al., 2002). FUC and/or L-Fucose may also have function in sperm-egg interaction in amphibians and mammals (Venditti et al., 2007).

In humans, FUC has great value for the diagnosis of hepatocellular carcinoma (HCC) as a serum marker (Gan et al., 2014; Wang et al., 2014). FUCs in humans are encoded by two genes: *hFUCA1*, which encodes the tissue enzyme, and *hFUCA2*, which encodes the plasma  $\alpha$ -L-fucosidase. A deficiency of *hFUCA1* causes fucosidosis, a disease characterized by the excretion of Non-degraded oligosaccharides via urine and excessive deposition of oligosaccharides, mucopolysaccharides, and glycolipids in tissues (Willems et al., 1999). Fucosidosis patients have symptoms of neurodegeneration with progressive motor and mental deterioration. There are several studies that indicate a link between *hFUCA1* and tumorigenesis. For example, p53 can target *hFUCA1* and regulate the growth and survival of cancer cells (Ezawa et al., 2016). *hFUCA2* is a secreted enzyme that is fundamental for the adhesion of *Helicobacter pylori*, particularly the duodenal ulcer- and gastric cancer-specific strains (Liu T.W. et al., 2009), suggesting that *hFUCA2* is a potential target for therapeutic intervention and clinical diagnosis of the diseases related to *H. pylori*.

FUCs in the GH29 family have been recognized widely in bacteria, ascidians, human, and other mammals. In the current research, sequence similarity networks (SSNs), multiple sequence alignments (MSAs), and other evolutionary analyses were carried out to obtain insights into the diversity and evolution of the FUC family across biospheres. We found that the FUC sequences were immensely more diverse than had been estimated previously, and that the proteins can be separated into 3 subfamilies.

## MATERIALS AND METHODS

### Putative FUCs Sequences Collection

For a detailed analysis of FUCs across biospheres, the complete protein sequences of the FUCs that have been experimentally characterized were acquired from literature. These protein sequences were used as queries and putative FUCs were collected from UniProt. The *E*-value threshold is  $10^{-2}$  during BLASTP. The Pfam database was further employed to identify the presence of the FUC domain (Marchler-Bauer et al., 2015). The protein sequences are listed in the **Supplementary Sheet**.

### MSAs and SSNs Analysis

MSAs of the putative FUC sequences were performed using the Multiple Alignment using Fast Fourier Transform (MAFFT) (version 7.0) program (Kato et al., 2017). The SSNs of FUC sequences were created by the EFI-EST (Gerlt et al., 2015). Cytoscape was used to show the SSNs Shannon et al. (2003).

The nodes in the networks indicate putative FUCs. An edge implies that the two putative FUCs linked by that edge show high sequence similarity Gerlt et al. (2015).

## Phylogenetic Analysis and Coevolving Protein Residues

MEGA7 was used to construct the unrooted phylogenetic trees using the maximum likelihood approach with 1000 bootstrap replicates (Kumar et al., 2016). The MISTIC web server was employed to calculate the coevolution in FUCs by the mutual information (MI) method in the MSAs (Simonetti et al., 2013). The three-dimensional structure of *hFUCA1* was modeled using the I-TASSER server<sup>1</sup> and visualized by PyMOL<sup>2</sup>.

### $K_a/K_s$ Ratio Calculation

The Non-synonymous to synonymous substitution rate is indicated by  $K_a/K_s$ . PAL2NAL and was used to calculate the values (Suyama et al., 2006).

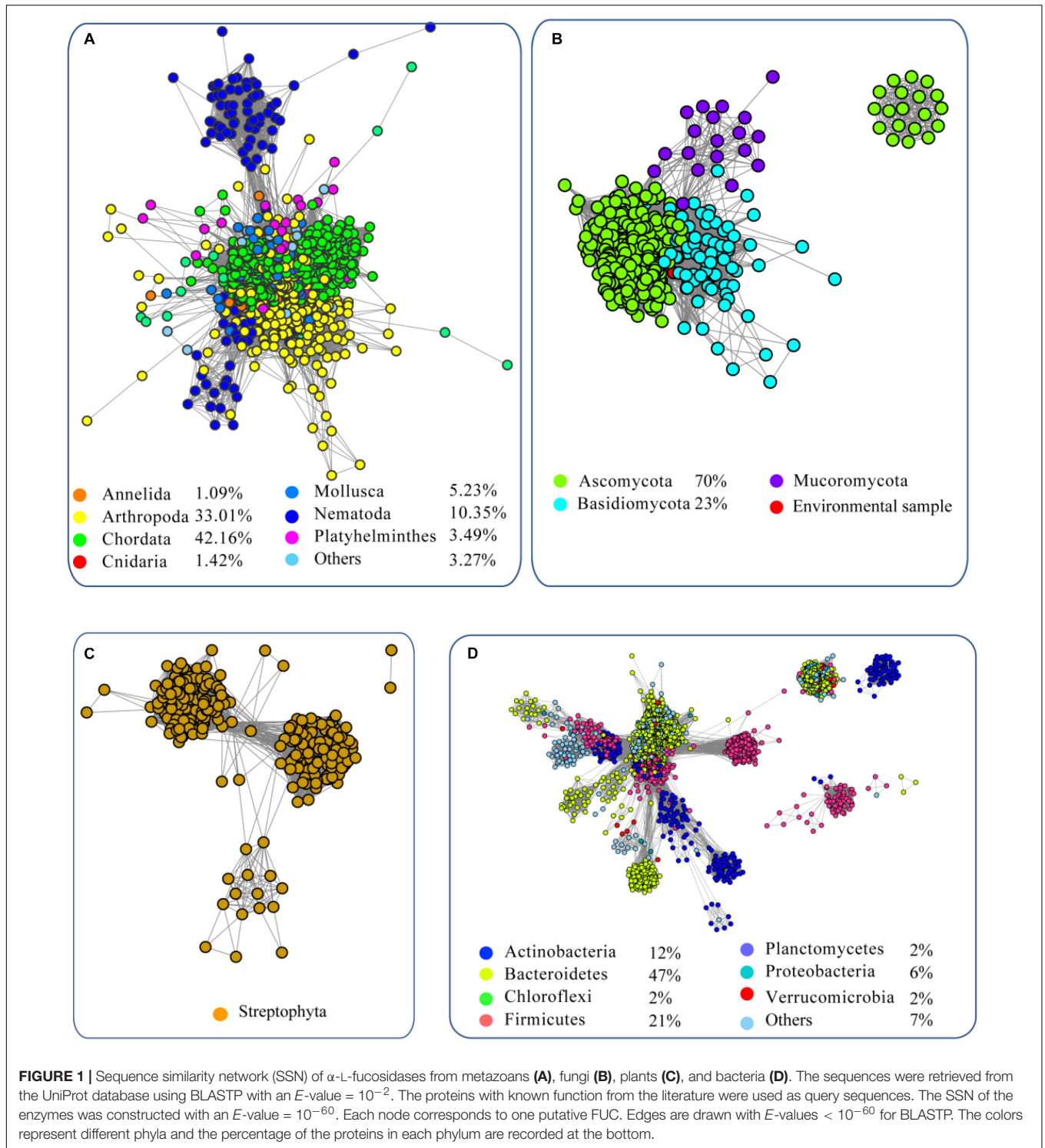
## RESULTS

### Distribution of FUCs in Biospheres

To analyze the evolution of FUCs, we searched the UniProt database using the FUCs from humans (UniProt ID: P04066), *Arabidopsis* (Q8GW72), *Dictyostelium discoideum* (P10901), and *Bacteroides thetaiotaomicron* (Q9WYE2) as query sequences, which have been experimentally characterized. In total, 6208 sequences were collected (**Supplementary Sheet**) and FUC homologs were found to be widespread in metazoans, plant, fungi, and bacteria; in addition, a few homologs were also found in archaea. To provide a more comprehensive view of these sequences, the SSNs for the sequences from the domain Bacteria, kingdom Metazoa, clade Viridiplantae, and kingdom Fungi were constructed using a protein sequence similarity cut-off of 60% (**Figure 1**). The network was further illustrated by classifying the included taxa. In Metazoa, members of the FUC family were mainly found in Chordata, Arthropoda, Nematoda, Mollusca, and Platyhelminthes (**Figure 1A**). All of the homologs in Metazoa are clustered together, indicating that these enzymes are highly conserved. In fungi, FUCs can be found in the three common phyla: Ascomycota, Basidiomycota, and Mucoromycota. The proteins from Basidiomycota separated into two clusters at a protein sequence similarity cut-off is 60% (**Figure 1B**). In Viridiplantae, FUCs only occurred in Streptophyta, but not in the green algae such as Spirotaenia and Chlorophyta. The homologs in Streptophyta can be grouped into one cluster, suggesting that the sequences are also relatively conserved (**Figure 1C**). In Bacteria, homologs of FUCs mainly occurred in Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria. The proteins in bacteria are relatively diverse because they separated into four clusters when the SSN was constructed using the same cut-off value (60% identity) (**Figure 1D**). In the taxonomic groups of Bacteria, Metazoa, Viridiplantae,

<sup>1</sup><http://zhanglab.ccmb.med.umich.edu/I-TASSER/>

<sup>2</sup><https://pymol.org/>

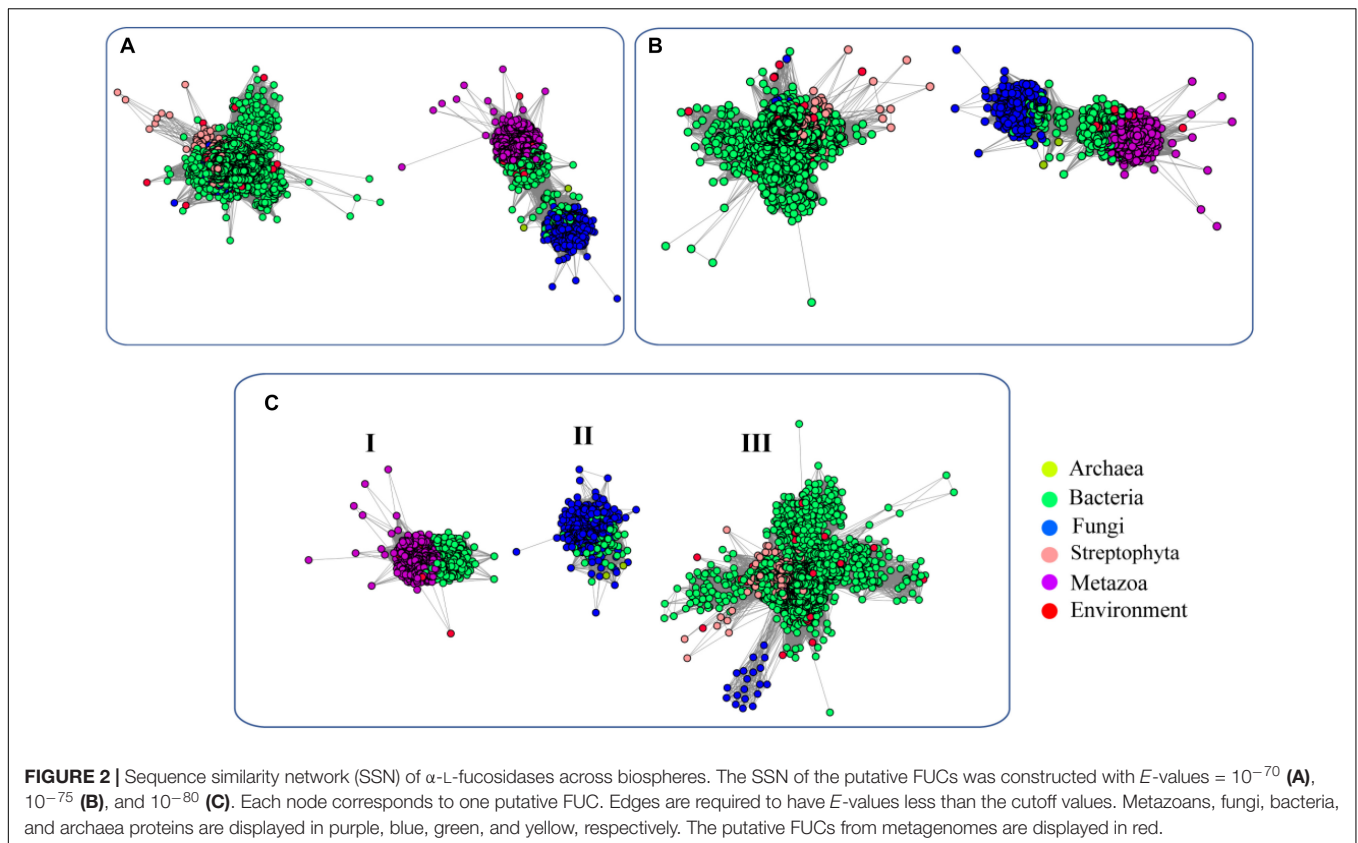


and Fungi, FUCs were relatively low in Fungi. FUCs were encoded in *Aspergillus niger*, a model filamentous ascomycete fungus that is widely used in biotechnology. However, the gene cannot be found in *Neurospora crassa*, *Saccharomyces cerevisiae*, and other model fungi, implying that a loss of the gene *FUC* occurred widely in fungi. In conclusion, FUCs

widely occurred in the biosphere and gene loss happened widely in fungi.

### Evolutionary Relationship of FUCs

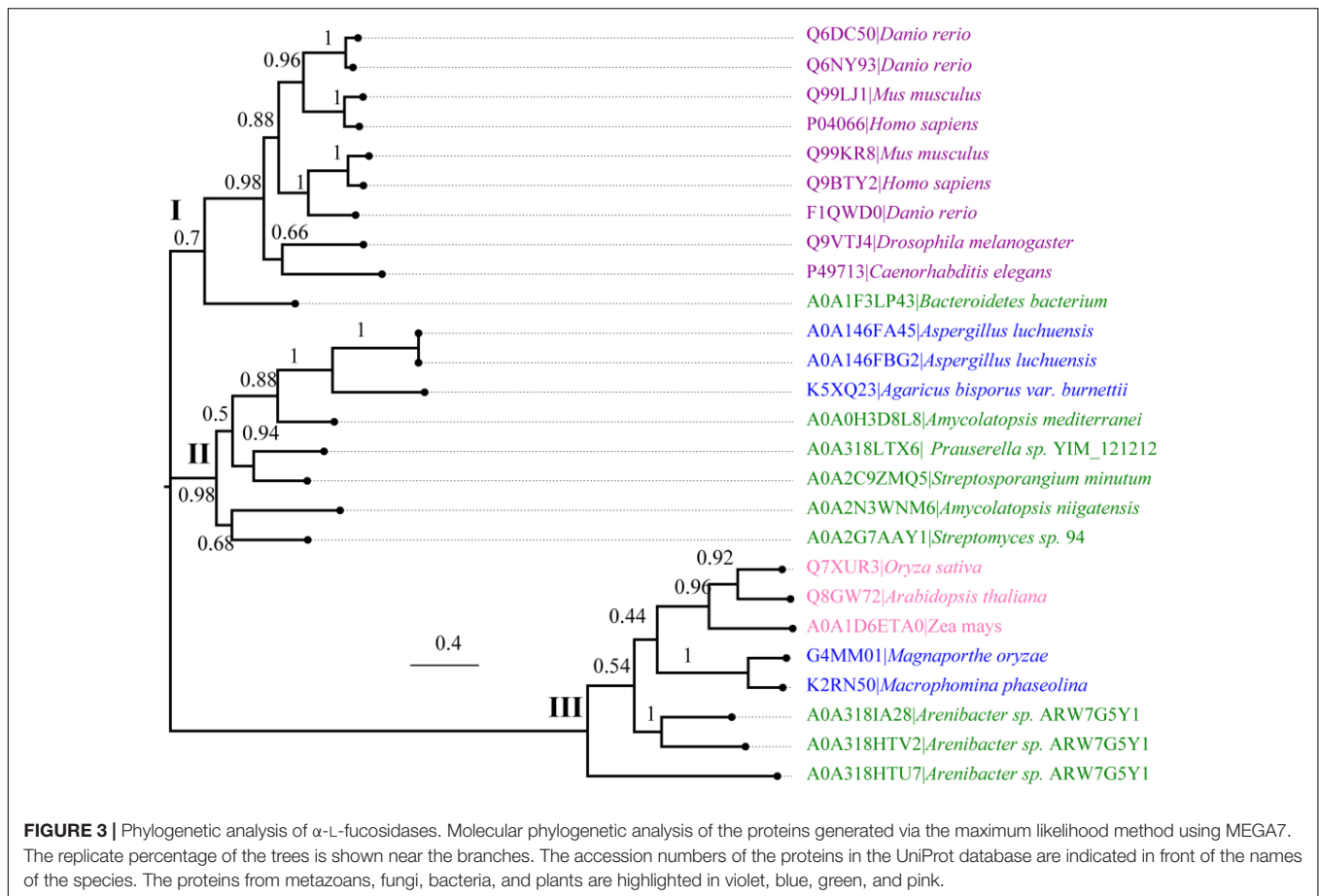
To elucidate the evolutionary relationship of FUCs among the species, we performed an SSN analysis of the identified proteins.



The results showed that the proteins from fungi and metazoans and some homologs in bacteria show similarity and fall into the same cluster at an  $E$ -value of  $10^{-70}$  (protein sequence identity > 30%), but the proteins from plants and the bacterial homologs separated and formed another cluster (Figure 2A). If the  $E$ -value =  $10^{-75}$  (protein sequence identity > 35%), the classification of these proteins shows a similar pattern as that if the  $E$ -value =  $10^{-70}$  (Figure 2B). When the  $E$ -value cutoff =  $10^{-80}$  (protein sequence identity > 40%), two clusters from metazoans and fungi separate (Figure 2C). In case of the two clusters, one cluster contains the proteins from bacteria and metazoans, and the proteins from the second cluster were from bacteria and fungi. The clusters containing the proteins from plant and bacteria cannot be separated at the cutoff value, suggesting that these enzymes are relatively highly conserved. In summary, the network analysis suggested that FUCs can be classified into 3 subfamilies, and FUCs from metazoans, fungi, and plants may share the common ancestors with the corresponding bacterial homologs.

We further performed an evolutionary analysis using the representative proteins in each subfamily assigned based on SSNs to provide more detailed evolutionary relationships among the subfamilies (Figure 3 and Supplementary Figure S1). Proteins from subfamily I included the enzymes from humans, *Mus musculus*, *Danio rerio*, *Drosophila melanogaster*, and *Caenorhabditis elegans*. In human and *M. musculus*, there are two copies of each gene. In *D. rerio*, one copy of the gene was further

duplicated, resulting in three copies. However, there is only one copy of FUC in *D. melanogaster* and in *C. elegans*. These proteins are clustered together with the enzyme from *B. thetaiotaomicron*, which has a known crystal structure (Wright et al., 2013). The proteins in subfamily II can be found in both fungi and bacteria. Generally, there is one copy of the gene in this subfamily, but the gene can also be duplicated in fungi, such as that in *Aspergillus luchuensis*. Most of the proteins in subfamily III are from bacteria and plants, while a few are from fungi. Similar to subfamily II, the gene in this subfamily can either be one or two copies because of gene duplication. In the genomes of *Oryza sativa* and *A. thaliana*, two model plant species, FUC from the GH29 family exists as a single copy. In this subfamily, the genes in some bacteria were duplicated, such as that in *Arenibacter* sp. A high bootstrap value supports the separation of these subfamilies in the phylogenetic tree. To further analyze the enzymes in the three subfamilies, the proteins from metazoans, fungi, and plants, as well as the corresponding enzymes in bacteria, were selected and aligned (Figure 4 and Supplementary Table S1). The enzymes from different subfamilies share conserved residues for catalysis and substrate binding (Figure 4). The enzymes in the same subfamily also shared relatively high identity in terms of amino acid sequences even though they are from different taxonomic groups (Supplementary Table S1). Based on SSN, phylogenetic tree, and sequence alignment, we suggested that the GH29 FUC family can be separated into 3 subfamilies, and that gene duplication occurred during evolution.



## Selection Pressure on FUCs

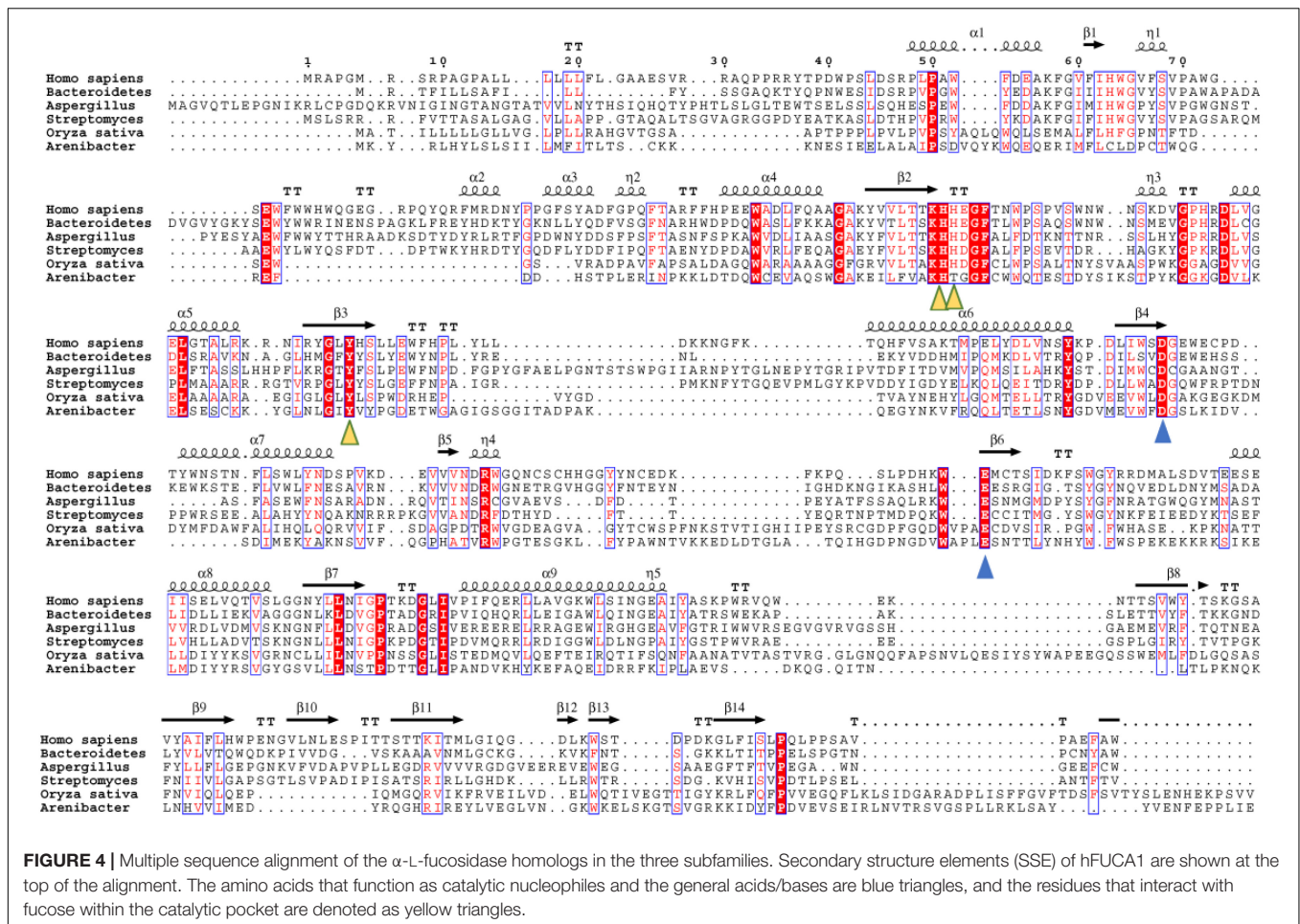
We measured the ratio of  $dN/dS$  of FUCs in the 3 groups with pairwise combinations of genes using PAL2NAL to analyze the selection pressure on the genes (Table 1 and Supplementary Figure S2). The results showed that the  $dN/dS$  ratios for the representative genes in the 3 subfamilies are  $< 1$ , indicating that the FUCs are under the pressure of negative selection. However, the ratios in the 3 subfamilies were slightly different: the ratios for  $dN/dS$  for the genes in metazoans are 0.16, while the values for the genes in other subfamilies are  $< 0.08$ . These results indicated that the FUCs in metazoans are subject to relaxed negative selection, while other subfamilies are under stringent negative selection.

## Conservation and Evolution Analysis of the Amino Acid Residues in FUCs of Bacteria and Metazoans

To elucidate the similarity of FUCs between bacteria and metazoans, the sequences from the two groups were analyzed by MSAs. The sequences of hFUCA1 (UniProt ID: P04066) and FUC from *B. thetaiotaomicron* (PDB ID: 6hzy; named bFUC) were employed to display the results of MSA (Figure 5). To better illustrate the conservation of the amino acids, the 3D structure of hFUCA1 was constructed by the I-TASSER server (Figure 5). The MSA results showed that the residues in the

N-termini are highly conserved compared to those in C-termini in both bacteria and metazoans. In the N-termini, the residues of hFUCA1 that bind fucose within the catalytic pocket include H63, H138, H139, and Y182, which is consistent with that found in a previous study (Liu S.W. et al., 2009). The corresponding amino acids in bFUC are H66, H135, H136, and Y178. The sequence WxDx that contains the catalytic nucleophile aspartate (D230 in hFUCA1 and D229 in bFUC) was also highly conserved. The D230N mutant lost catalytic activity significantly (Liu S.W. et al., 2009). Mapping the top conserved amino acids residues onto the structures indicated that the conserved amino acids that bind fucose and catalyze the reaction form a pocket in the N-terminus of the enzyme. Other conserved amino acid residues, including tryptophan, tyrosine, and proline, were located on the surface of the enzyme.

We further investigated the coevolution of the amino acids in FUCs using MI based on MSA (Figure 5A). Two residues are likely to be coevolving if they have a high MI score, which means that one amino acid mutation is related to another particular mutation in order to keep the function (Petit et al., 2014). The network of MI for the FUC members in bacteria indicated that the top 10% of the values were spread in the N-termini of the enzymes, while the values in the C-termini were relatively low (Figure 4A). However, the high values of the proteins in metazoans were in the C-termini. The analysis suggested that the



**TABLE 1 |** Tabular representation of the dN/dS calculations of  $\alpha$ -L-fucosidases in the three subfamilies.

Subfamily	UniProt ID	Number of Synonymous site (S)	Number of Non-synonymous site (N)	Synonymous substitution rate (dS)	Non-synonymous substitution rate (dN)	dN/dS
I	Q9BTY2	320.6	1035.4	0.6313	0.1015	0.1608
II	A0A146FBG2	410.6	1311.4	9.3385	0.5992	0.0642
V	Q7XUR3	402.6	1103.4	4.6373	0.3412	0.0736

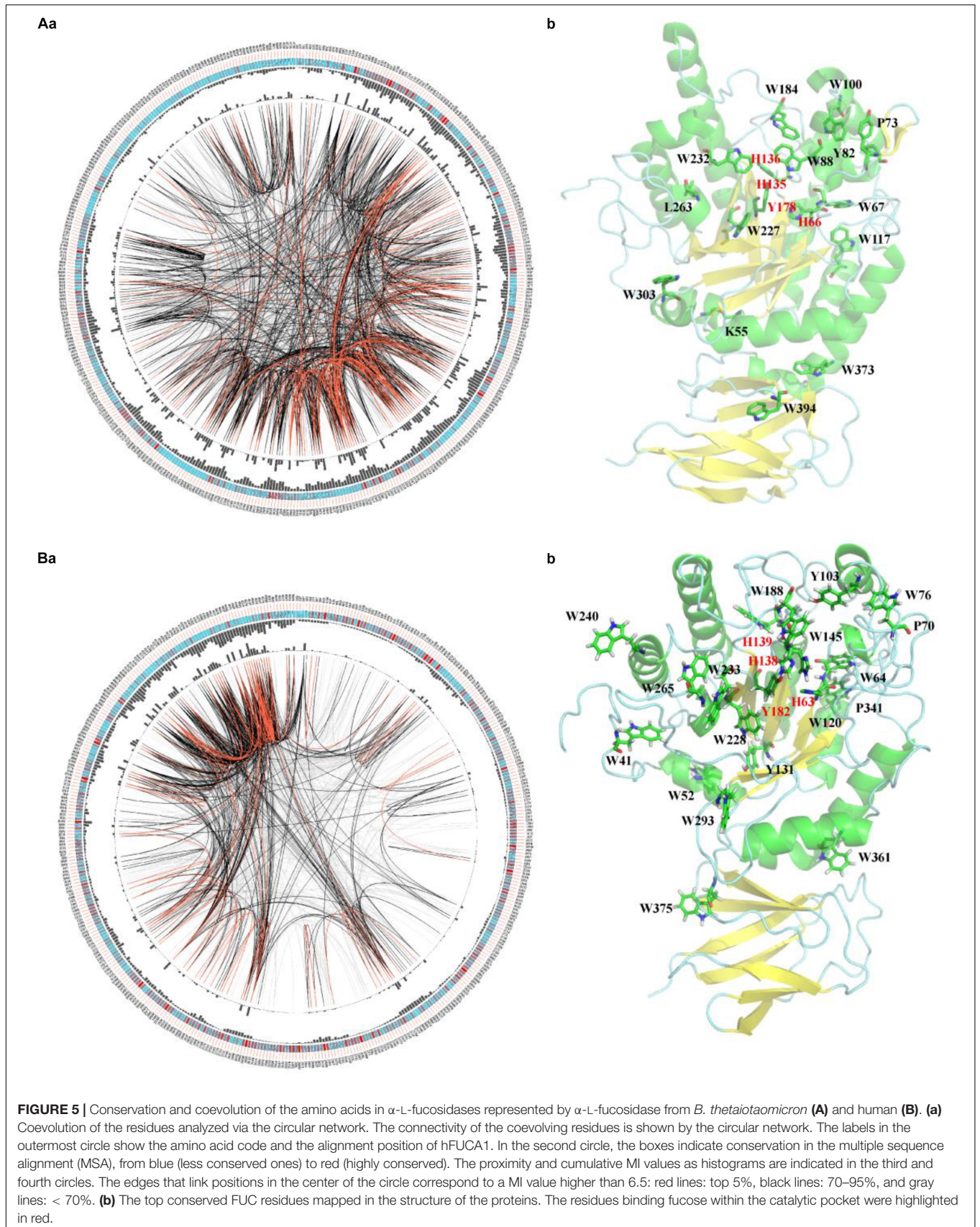
evolutionary patterns of FUCs in bacteria and metazoans were different even though the conserved residues were highly similar.

### DISCUSSION

FUCs, which are responsible for fucosylated glycoconjugate processing, are involved in cystic fibrosis, inflammation, lysosomal storage disease, cancer development, and in the interactions between gametes in both invertebrates and vertebrates. In this study, a comprehensive *in silico* study of FUCs was performed, which displayed that the homologs are distributed among bacteria, fungi, plants, and metazoans. Based on SSN and phylogenetic tree analyses, the proteins in the FUC family can be classified into 3 subfamilies and

all of the subfamilies are under relaxed negative selection during evolution.

A previous study showed that FUCs were mainly from bacteria, fungi, and metazoans (Intra et al., 2007). In this study, we showed that FUCs also occurred in plants. In *Arabidopsis*, three enzymes (AtFuc95A, AtFUC1, and AtFXG1) were identified to have  $\alpha$ -L-fucosidase activity (Frankova and Fry, 2013). Only AtFUC1 belongs to the GH29 family among the three enzymes (De La Torre et al., 2002), and we only considered the homologs of AtFUC1 in this study. The homologs from plants shared high sequence identity and showed high similarity to the bacterial homologues. Another previous study also showed that two copies of FUC genes are present in all vertebrates. Our research showed that one copy of the gene can be further duplicated, as the three copies of the gene encode the enzyme in



*D. rerio*. In bacteria and fungi, the gene encoding FUC was also duplicated as two copies of the gene were found in many species.

The MSA with structure analysis showed that the conserved amino acids in FUCs can form a pocket in the N-termini for binding and catalyzing. The C-termini composed of  $\beta$ -sheets were less conserved. Several hydrophobic and large amino acids such as Trp and Tyr located near the surface of the enzyme were also conserved. The hydrophobicity of these amino acids may affect the solubility and function of the enzyme (Podorieszsch and Huttunen-Hennelly, 2010), which in turn may also affect crystal growth important for structure determination. This may explain why the structures of hFUCA1 and hFUCA2 have not been solved until now. For the conserved amino acids, E289 may be responsible for general acid/base catalysis (Liu S.W. et al., 2009); however, according to the alignment, this amino acid is outside of the active site (Shaikh et al., 2013). Our study also showed that E289 is far from the active site and is less conserved. Cysteine was thought to be part of the active site, but cysteine was less conserved based on our analysis. These analyses suggested that the catalytic function of the amino acid residues in FUC is still unclear.

The FUC family can be divided into 3 subfamilies. Among them, the proteins from subfamily I were from bacteria and metazoans, the proteins from subfamily II were from bacteria and fungi, and the enzymes from subfamily III were from both bacteria and plants. This suggested that FUC appeared before bacteria and eukaryotes diverged. Furthermore, the results together with selection pressure analysis also suggested that FUCs were also conserved during evolution. Finally, we proposed that

gene duplication is a common event in metazoans, as well as in bacteria and fungi during the evolution of FUC, as 2–3 copies of the gene can be found in the species.

## DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

## AUTHOR CONTRIBUTIONS

JY and TJ designed the experiments, performed all analysis, and wrote the manuscript. SL revised the manuscript and the data.

## FUNDING

This study was supported by the Natural Science Foundation of Fujian Province (No. 2016J01529).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2019.01756/full#supplementary-material>

## REFERENCES

- Becker, D. J., and Lowe, J. B. (2003). Fucose: biosynthesis and biological function in mammals. *Glycobiology* 13, 41R–53R. doi: 10.1093/glycob/cwg054
- De La Torre, F., Sampedro, J., Zarra, I., and Revilla, G. (2002). AtFXG1, an *Arabidopsis* gene encoding alpha-L-fucosidase active against fucosylated xyloglucan oligosaccharides. *Plant Physiol.* 128, 247–255. doi: 10.1104/pp.128.1.247
- Ezawa, I., Sawai, Y., Kawase, T., Okabe, A., Tsutsumi, S., Ichikawa, H., et al. (2016). Novel p53 target gene FUCA1 encodes a fucosidase and regulates growth and survival of cancer cells. *Cancer Sci.* 107, 734–745. doi: 10.1111/cas.12933
- Frankova, L., and Fry, S. C. (2013). Biochemistry and physiological roles of enzymes that ‘cut and paste’ plant cell-wall polysaccharides. *J. Exp. Bot.* 64, 3519–3550. doi: 10.1093/jxb/ert201
- Gan, Y., Liang, Q., and Song, X. (2014). Diagnostic value of alpha-L-fucosidase for hepatocellular carcinoma: a meta-analysis. *Tumour Biol.* 35, 3953–3960. doi: 10.1007/s13277-013-1563-8
- Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D. R., et al. (2015). Enzyme function initiative-enzyme similarity tool (EFI-EST): a web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta* 1854, 1019–1037. doi: 10.1016/j.bbapap.2015.04.015
- Intra, J., Perotti, M.-E., Pavesi, G., and Horner, D. (2007). Comparative and phylogenetic analysis of  $\alpha$ -L-fucosidase genes. *Gene* 392, 34–46. doi: 10.1016/j.gene.2006.11.002
- Katayama, T., Sakuma, A., Kimura, T., Makimura, Y., Hiratake, J., Sakata, K., et al. (2004). Molecular cloning and characterization of *Bifidobacterium bifidum* 1,2- $\alpha$ -L-fucosidase (AfcA), a novel inverting glycosidase (Glycoside Hydrolase Family 95). *J. Bacteriol.* 186, 4885–4893. doi: 10.1128/jb.186.15.4885-4893.2004
- Kato, K., Rozewicki, J., and Yamada, K. D. (2017). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* doi: 10.1093/bib/bbx108 [Epub ahead of print].
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Li, T., Li, M., Hou, L., Guo, Y., Wang, L., Sun, G., et al. (2018). Identification and characterization of a core fucosidase from the bacterium *Elizabethkingia meningoseptica*. *J. Biol. Chem.* 293, 1243–1258. doi: 10.1074/jbc.M117.804252
- Liu, S. W., Chen, C. S., Chang, S. S., Mong, K. K., Lin, C. H., Chang, C. W., et al. (2009). Identification of essential residues of human alpha-L-fucosidase and tests of its mechanism. *Biochemistry* 48, 110–120. doi: 10.1021/bi801529t
- Liu, T. W., Ho, C. W., Huang, H. H., Chang, S. M., Papat, S. D., Wang, Y. T., et al. (2009). Role for  $\alpha$ -L-fucosidase in the control of *Helicobacter pylori*-infected gastric cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14581–14586. doi: 10.1073/pnas.0903286106
- Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., et al. (2015). CDD: NCBI’s conserved domain database. *Nucleic Acids Res.* 43, D222–D226. doi: 10.1093/nar/gku1221
- Matsumoto, M., Hirata, J., Hirohashi, N., and Hoshi, M. (2002). Sperm-egg binding mediated by sperm  $\alpha$ -L-fucosidase in the ascidian, *Halocynthia roretzi*. *Zool. Sci.* 19, 43–48. doi: 10.2108/zsj.19.43
- Moreti, R., Perrella, N. N., and Lopes, A. R. (2013). Carbohydrate digestion in ticks and a digestive  $\alpha$ -L-fucosidase. *J. Insect. Physiol.* 59, 1069–1075. doi: 10.1016/j.jinsphys.2013.08.008
- Petit, D., Teppa, E., Mir, A.-M., Vicogne, D., Thisse, C., Thisse, B., et al. (2014). Integrative view of  $\alpha$ 2,3-sialyltransferases (ST3Gal) molecular and functional evolution in deuterostomes: significance of lineage specific losses. *Mol. Biol. Evol.* 32, 906–927. doi: 10.1093/molbev/msu395
- Podorieszsch, A. P., and Huttunen-Hennelly, H. E. (2010). The effects of tryptophan and hydrophobicity on the structure and bioactivity of novel indolicidin derivatives with promising pharmaceutical potential. *Org. Biomol. Chem.* 8, 1679–1687. doi: 10.1039/b921248e



- Shaikh, F. A., Lammerts Van Bueren, A., Davies, G. J., and Withers, S. G. (2013). Identifying the catalytic acid/base in GH29  $\alpha$ -L-fucosidase subfamilies. *Biochemistry* 52, 5857–5864. doi: 10.1021/bi400183q
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Simonetti, F. L., Teppa, E., Chernomoretz, A., Nielsen, M., and Marino Buslje, C. (2013). MISTIC: mutual information server to infer coevolution. *Nucleic Acids Res.* 41, W8–W14. doi: 10.1093/nar/gkt427
- Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34, W609–W612. doi: 10.1093/nar/gkl315
- Venditti, J. J., Donigan, K. A., and Bean, B. S. (2007). Crypticity and functional distribution of the membrane associated  $\alpha$ -L-fucosidase of human sperm. *Mol. Reprod. Dev.* 74, 758–766. doi: 10.1002/mrd.20666
- Wang, K., Guo, W., Li, N., Shi, J., Zhang, C., Lau, W. Y., et al. (2014). Alpha-1-fucosidase as a prognostic indicator for hepatocellular carcinoma following hepatectomy: a large-scale, long-term study. *Br. J. Cancer* 110, 1811–1819. doi: 10.1038/bjc.2014.102
- Willems, P. J., Seo, H. C., Coucke, P., Tonlorenzi, R., and O'Brien, J. S. (1999). Spectrum of mutations in fucosidosis. *Eur. J. Hum. Genet.* 7, 60–67. doi: 10.1038/sj.ejhg.5200272
- Wright, D. W., Moreno-Vargas, A. J., Carmona, A. T., Robina, I., and Davies, G. J. (2013). Three dimensional structure of a bacterial  $\alpha$ -L-fucosidase with a 5-membered iminocyclitol inhibitor. *Bioorg. Med. Chem.* 21, 4751–4754. doi: 10.1016/j.bmc.2013.05.056

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 You, Lin and Jiang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.