ISSN 0026-2617, Microbiology, 2013, Vol. 82, No. 4, pp. 415–422. © Pleiades Publishing, Ltd., 2013.
Original Russian Text © D.G. Naumoff, 2013, published in Mikrobiologiya, 2013, Vol. 82, No. 4, pp. 408–416.

EXPERIMENTAL ARTICLES

# Multiple Lateral Transfers and Duplications of Genes as Sources of Diversity of α-L-Rhamnosidases in *Clostridium methylpentosum* DSM5476

**D. G. Naumoff**[1]

*Winogradsky Institute of Microbiology, Russian Academy of Sciences, pr. 60-letiya Oktyabrya 7, k. 2, Moscow, 117312 Russia*

Received October 15, 2012

**Abstract**—α-L-Rhamnosidases are an important group of glycoside hydrolases represented in many organisms from various prokaryotic phyla. Based on the homology of catalytic domains, all these proteins are assigned to the GH78 and GH106 families of glycoside hydrolases. However, most prokaryotic genomes contain no genes encoding proteins from these two families. We found that the unique genome of *Clostridium methylpentosum* DSM5476 contains 83 genes of proteins from these families and undertook investigation of their phylogeny. The absence of homologous genes in most of strains of the genus *Clostridium* suggests an important ecological role of these genes, in *C. methylpentosum* in particular. Phylogenetic analysis revealed multiple lateral transfers and duplications of the corresponding genes.

*Keywords: Clostridium methylpentosum*, glycoside hydrolase, α-L-rhamnosidase, GH78 family, GH106 family, protein evolution, protein phylogenetic tree, lateral transfer, gene duplication, paralogue

**DOI:** 10.1134/S0026261713040085

α-L-Rhamnosidases (EC 3.2.1.40) are a widespread and industrially important group of glycoside hydrolases; they are responsible for cleavage of terminal α-L-rhamnose residues from the nonreducing end of carbohydrates and their derivatives [1–5]. Based on homology of the catalytic domains, virtually all enzymes of this group are assigned to the GH78 and GH106 families of glycoside hydrolases in the Carbohydrate-Active enZymes (CAZy) classification; no other enzymatic activities have been reported for proteins of these two families [6]. However, enzymes of different families somewhat differ in terms of their substrate specificities [7, 8]. According to the CAZy database, the GH78 and GH106 families contain 483 and 86 proteins, respectively. Most of the sequenced prokaryotic| genomes do not encode proteins of these families, while some genomes contain several of their paralogues [6]. Our recent comparative study of α-L-rhamnosidases demonstrated that the lists of the GH78 and GH106 families at the CAZy site are incomplete. The total number of proteins containing catalytic domains of the two families in the GenPept database (section of non-redundant protein sequences) equaled 1981 according to the results of blastp screening performed in the end of August 2012 [9].

In 1989, strain ATCC43829 = DSM5476, an isolate from human intestine, was described as a new species, *Clostridium methylpentosum*, characterized by an unusual ability to ferment only pentoses and methylpentoses, including L-rhamnose [10]. The latter ability indicates possible presence of α-L-rhamnosidase activity in this strain. Recently, the genome sequence of this organism, represented by 146 contigs, was deposited in GenBank (ACEC00000000.1). Our analysis of this genome revealed 83 genes encoding proteins that contain domains of the GH78 and GH106 families of glycoside hydrolases [11]. Meanwhile, the majority of sequenced genomes of bacteria of the genus *Clostridium* do not encode proteins of the GH78 and GH106 families [6]. Thus, the evolutionary origin of the hypothetical α-L-rhamnosidase genes in *C. methylpentosum* DSM5476 remained unclear: did they emerge as a result of a limited number of lateral transfers followed by multiple duplications, or were the lateral transfers numerous? In the present work, we attempted to answer this question.

## RESULTS

### *GH106 Family of Glycoside Hydrolases*

Analysis of the amino acid sequences of the 46 proteins of *C. methylpentosum* that we previously found to contain the GH106 family domain [11] showed that three of them are not full-size proteins. Two of the three proteins (GenPept, EEG29402.1 and EEG29403.1) turned out to be fragments of a single protein whose gene is interrupted by a stop codon. The third protein (EEG30591.1) can be elongated through

[1] Corresponding author; e-mail: daniil_naumoff@yahoo.com

choice of an alternative start codon. We edited the above proteins and used the two resulting full-size variants in further analyses.

Analysis of one of the fragments (GenBank, ACEC01000062.1) of the *C. methylpentosum* genome sequence allowed us to detect one more candidate gene that encodes a protein containing the domain of the GH106 glycoside hydrolase family. Pairwise comparison of the deduced amino acid sequences revealed the CLOSTMETH_00114 (GenPept, EEG32242.1) gene as its closest evolutionary relative (42% amino acid identity). Detection of multiple stop codons in the relevant potential reading frame, as well as a large deletion, allowed us to conclude that this is a pseudogene. However, the corresponding region of the genome was annotated as containing two short genes: CLOSTMETH_01918 (EEG30477.1) and CLOSTMETH_01919 (EEG30478.1). This pseudogene was excluded from further analysis.

We used 343 proteins containing the GH106 family domain for the phylogenetic analysis; 45 of them were *C. methylpentosum* proteins. The trees were constructed using the Neighbor-Joining (NJ; Fig. 1) and Maximum Parsimony (MP; this tree is not shown) algorithms. The NJ method is based on repeated pairwise joining of most similar sequences, and the MP algorithm implies minimization of the number of evolutionary changes. Due to the essential differences in these algorithms, similarity of the generated topologies suggests that they most probably reflect the real evolutionary events. In both trees, proteins of *C. methylpentosum* belonged to three well-separated clusters. One of the clusters (cluster 1 in Fig. 1; 100 and 99% bootstrap support in NJ and MP trees respectively) contained 38 *C. methylpentosum* proteins, several proteins of *Ascomycota* and *Actinobacteria*, and a protein of a *Lachnospiraceae bacterium* (GenPept, EGN36798.1). The second cluster (cluster 2 in Fig. 1; 100% bootstrap support in both trees) contained six *C. methylpentosum* proteins and proteins of bacteria from the *Bacteroidetes* phylum. The third cluster (cluster 3 in Fig. 1; 81 and 96% bootstrap support in NJ and MP trees respectively) contained only one protein of *C. methylpentosum* (EEG28461.1) together with proteins of the *Bacteroidetes* and *Proteobacteria*.

### *GH78 Family of Glycoside Hydrolases*

Pairwise comparison of the amino acid sequences of the 37 proteins of *C. methylpentosum* that we previously [11] found to contain the GH78 family domain revealed that 28 of them may be united into two groups of closely related proteins; each of these groups emerged as a result of an ancestor gene duplication. One of the groups contained 22 *C. methylpentosum* proteins. Phylogenetic analysis of these 22 proteins, together with 161 close homologues, showed that all of these *C. methylpentosum* proteins belonged to a cluster that comprised 49 proteins with a bootstrap support of 94% in the NJ tree (Fig. 2; the cluster is outlined with a dashed line) and 89% support in the MP tree (not shown). Within this cluster, there are stable subclusters (at least 80% support in NJ and MP trees) formed by proteins of representatives of the phyla *Actinobacteria* (subcluster number 4 in Fig. 2) and *Bacteroidetes* (subclusters 1, 2, and 3 in Fig. 2). In addition, the cluster contained two proteins from other bacteria of the *Firmicutes* phylum: *Lachnospiraceae bacterium* (GenPept, EGN42892.1) and *Paenibacillus* sp. (ACX65306.1).

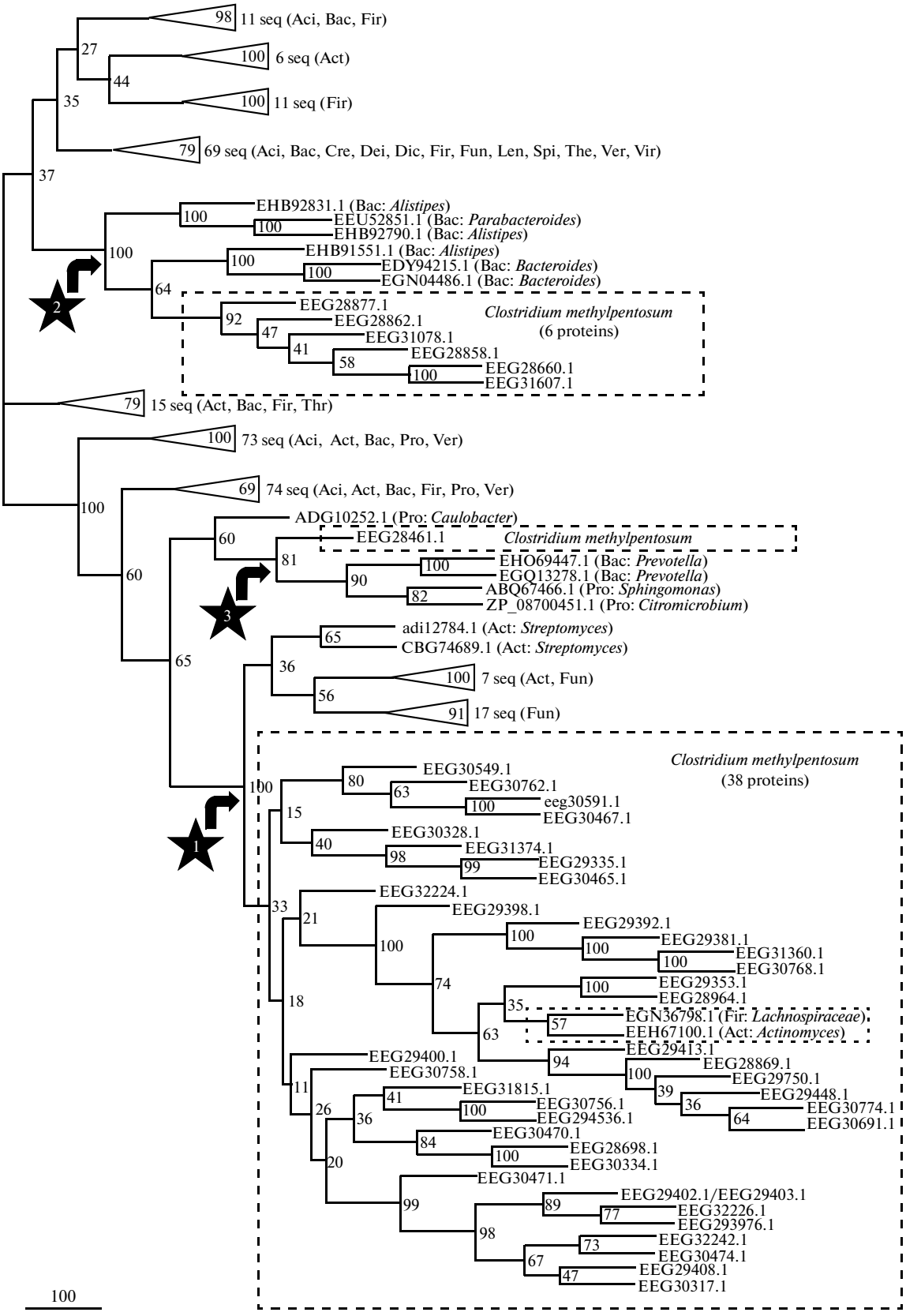
The six *C. methylpentosum* proteins of the second group, together with 57 their close homologues, were used to construct another phylogenetic tree. All these proteins of *C. methylpentosum*, together with 33 proteins from other organisms (mostly of the *Bacteroidetes* phylum) formed a stable cluster with 100% bootstrap support in NJ (not shown) and MP trees (Fig. 3).

The list of the rest nine proteins of *C. methylpentosum* containing the GH78 family domain, together with the information on their closest homologues, is presented in the table. Two of these proteins (EEG28703.1 and EEG28728.1), with 47% identical amino acid sequences, apparently emerged as a result of a recent ancestor gene duplication.

## DISCUSSION

Glycoside hydrolases are well-known as a diverse and widespread group of enzymes represented in almost all living organisms [6, 12]. Genes coding for glycoside hydrolases and their homologues make up approximately 1% of all sequenced genes. However,

**Fig. 1.** Scheme of the phylogenetic tree of GH106 family glycoside hydrolases constructed using the neighbor-joining method and comprising 343 proteins. Statistical significance of the nodes was evaluated by bootstrap analysis; the number of supporting pseudoreplicas (out of 100) is indicated at each node. All proteins are labeled with identifiers of the GenPept database; also (with the exception of *C. methylpentosum* proteins), host organism affiliation (genus and a higher taxon, typically, phylum) is indicated. Lower case letters in GenPept identifiers (for example, eeg30591.1) mark proteins whose amino acid sequences underwent editing in the course of multiple alignment. Protein clusters containing proteins of *Clostridium methylpentosum* are labeled and outlined with dashed rectangles. Within these clusters, subclusters of the proteins of other organisms are marked with thinner dashed lines. In the triangles are stable clusters containing no *C. methylpentosum* proteins. Inside a triangle, bootstrap support of the cluster is reported, and on the right of the triangles, the number of proteins in the cluster and the host taxonomy. Numbers in stars with arrows indicate clusters of proteins discussed in the text. The following taxon abbreviations are used: Aci, *Acidobacteria*, Act, *Actinobacteria*, Bac, *Bacteroidetes*, Chl, *Chloroflexi*, Cre, *Crenarchaeota*, Dei, *Deinococci*, Dic, *Dictyoglomi*, Fir, *Firmicutes*, Fun, *Fungi*, Len, *Lentisphaerae*, Pla, *Planctomycetes*, Pro, *Proteobacteria*, Spi, *Spirochaetes*, The, *Thermotogae*, Thr, *Thermobaculum terrenum*, Ver, *Verrucomicrobia*, and Vir, *Viridiplantae*.


11 seq (Aci, Bac, Fir)
6 seq (Act)
11 seq (Fir)
69 seq (Aci, Bac, Cre, Dei, Dic, Fir, Fun, Len, Spi, The, Ver, Vir)
EHB92831.1 (Bac: Alistipes)
EEU52851.1 (Bac: Parabacteroides)
EHB92790.1 (Bac: Alistipes)
EHB91551.1 (Bac: Alistipes)
EDY94215.1 (Bac: Bacteroides)
EGN04486.1 (Bac: Bacteroides)
EEG28877.1
EEG28862.1
EEG31078.1
EEG28858.1
EEG28660.1
EEG31607.1
Clostridium methylpentosum
(6 proteins)
15 seq (Act, Bac, Fir, Thr)
73 seq (Aci, Act, Bac, Pro, Ver)
74 seq (Aci, Act, Bac, Fir, Pro, Ver)
ADG10252.1 (Pro: Caulobacter)
EEG28461.1
Clostridium methylpentosum
EHO69447.1 (Bac: Prevotella)
EGQ13278.1 (Bac: Prevotella)
ABQ67466.1 (Pro: Sphingomonas)
ZP_08700451.1 (Pro: Citromicrobium)
adi12784.1 (Act: Streptomyces)
CBG74689.1 (Act: Streptomyces)
7 seq (Act, Fun)
17 seq (Fun)
Clostridium methylpentosum
(38 proteins)
EEG30549.1
EEG30762.1
eeg30591.1
EEG30467.1
EEG30328.1
EEG31374.1
EEG29335.1
EEG30465.1
EEG32224.1
EEG29398.1
EEG29392.1
EEG29381.1
EEG31360.1
EEG30768.1
EEG29353.1
EEG28964.1
EGN36798.1 (Fir: Lachnospiraceae)
EEH67100.1 (Act: Actinomyces)
EEG29413.1
EEG28869.1
EEG29750.1
EEG29448.1
EEG30774.1
EEG30691.1
EEG29400.1
EEG30758.1
EEG31815.1
EEG30756.1
EEG294536.1
EEG30470.1
EEG28698.1
EEG30334.1
EEG30471.1
EEG29402.1/EEG29403.1
EEG32226.1
EEG293976.1
EEG32242.1
EEG30474.1
EEG29408.1
EEG30317.1
100

relative and absolute abundance of these genes and the set of enzymatic activities in the proteins they encode vary in wide ranges, even in closely related microorganisms. Thus, according to the CAZy database [6], 146 genes of hypothetical glycoside hydrolases (or 2.91% of all protein-coding genes) were detected in *Flavobacterium johnsoniae* UW101 (GenBank, CP000685.1), whereas only 9 genes (0.37%) occur in *Flavobacterium psychrophilum* JIP02/86 (AM398681.1), 8 genes (1.00%), in *Mycoplasma fermentans* JER (CP001995.1), and no genes, in *Mycoplasma bovis* Hubei-1 (CP002513.1). Such a distribution pattern can be explained by the important role that gene duplication, loss, and lateral transfer play in the evolution of glycoside hydrolase genes in practically all taxonomic groups of living organisms, and is primarily determined by the ecological niches they occupy.

Today, genome sequences of several thousand prokaryotic species are available. This raises a possibility of investigation of the abundance and representation of any gene in organisms of various taxa. *Actinobacteria, Firmicutes*, and *Proteobacteria* are evident leaders among the prokaryotic phyla in both the number of validly described species [13] and the number of sequenced genomes [14]. Underrepresentation of a gene in one of these phyla (as well as its overrepresentation in some other phylum) indicates possible ecological significance of the relevant protein. In this context, proteins exhibiting α-L-rhamnosidase activity are of interest. According to the CAZy database [6], proteins of the GH78 family are overrepresented in the *Bacteroidetes* (108 proteins) and underrepresented in *Proteobacteria* (29) as compared to *Actinobacteria* (98) and *Firmicutes* (127). At the same time, proteins of the GH106 family are overrepresented in *Acidobacteria* (10) and *Bacteroidetes* (57) as compared to *Actinobacteria* (10) and *Proteobacteria* (5), and are absent from *Firmicutes* [6]. These statistics indicate selective accumulation of α-L-rhamnosidase genes in organisms occupying ecological niches rich in relevant substrates, in particular, the gastrointestinal tract of animals.

Another indication of probable ecological importance of a protein may be the presence in a particular taxonomic group both of organisms harboring a large number of its homologues (paralogues) and of organisms harboring a few homologues or lacking this protein at all. α-L-Rhamnosidases of bacteria of the genus *Clostridium* are just the case. Among the 39 genomes of bacteria of this genus represented in the CAZy database [6], only the genome of *C. saccharolyticum* WM1 encodes two proteins of the GH78 family (GenPept, ADL03771.1 and ADL04034.1) (and no proteins of the GH106 family). The other 38 genomes do not encode proteins of these two families. The genome of *C. methylpentosum* DSM5476, which is not represented in the CAZy database [6], encodes 37 and 45 hypothetical α-L-rhamnosidases belonging to GH78 and GH106 families, respectively, and additionally contains a pseudogene of a protein of the GH106 family (GenPept, EEG30477.1 and EEG30478.1). Pairwise comparison of amino acid sequences (table) and phylogenetic analysis (Figs. 1–3) that we conducted indicate multiple lateral transfers of genes coding for hypothetical α-L-rhamnosidases into the genome of *C. methylpentosum*.

Most likely, there have been three lateral transfers of the GH106 family genes. One of these transfers must have been from a *Bacteroidetes* bacterium (see cluster 2 in Fig. 1), whereas the origins of the other two transferred genes are more difficult to speculate on. It may be supposed that the second transfer occurred from a bacterium of the phylum *Proteobacteria* or *Bacteroidetes* (see cluster 3 in Fig. 1), and the third one, from some *Actinomycetes* or fungi (see cluster 1 in Fig. 1). The lateral transfers were followed by multiple duplications of genes, which have led to the observed diversity of the hypothetical GH106 family α-L-rhamnosidases in *C. methylpentosum*. Presence of stable protein groups within relevant clusters (Fig. 1) suggests that these duplications occurred not immediately after the lateral transfer, but over a long period of time. After most duplications had happened, there was lateral transfer of one of the formed genes from *C. methylpentosum* to the genomes of other bacteria (*Actinomyces urogenitalis* DSM 15434, GenPept, EEH67100.1; and *Lachnospiraceae* bacterium 1_4_56FAA, EGN36798.1). However, one may not exclude alternative hypotheses of evolutionary events. Thus, the largest cluster, comprising 38 proteins of *C. methylpentosum*, has a very low bootstrap support (33% in both NJ and MP trees). This allows an assumption to be made that these proteins emerged as a result of two or more lateral transfers of relatively closely related genes, rather than as a result of a single gene transfer. Alternatively, a sole lateral transfer could have happened much earlier, resulting in the appearance of these 38 proteins together with one more *C. methylpentosum* protein, EEG28461.1. Their common cluster has a much higher bootstrap support (65 and 64% in NJ and MP trees, respectively). However, such course of events implies multiple lateral transfers from *C. methylpentosum* to other organisms, including fungi (Fig. 1), which does not seem likely.

The number of lateral transfers to the *C. methylpentosum* genome of genes encoding GH78 family proteins was significantly higher, and only a few of them were followed by duplications. Most of these *C. methylpentosum* proteins (22 of 37) belong to a single stable cluster in the phylogenetic tree (Fig. 2), suggesting that the encoding genes emerged through multiple duplications of a single ancestor gene acquired by *C. methylpentosum* via lateral transfer form an unknown organism. These duplications also occurred over an evolutionarily long period of time. Some of the resulting genes (at least four) were then transferred, directly or indirectly, to genomes of a number of other

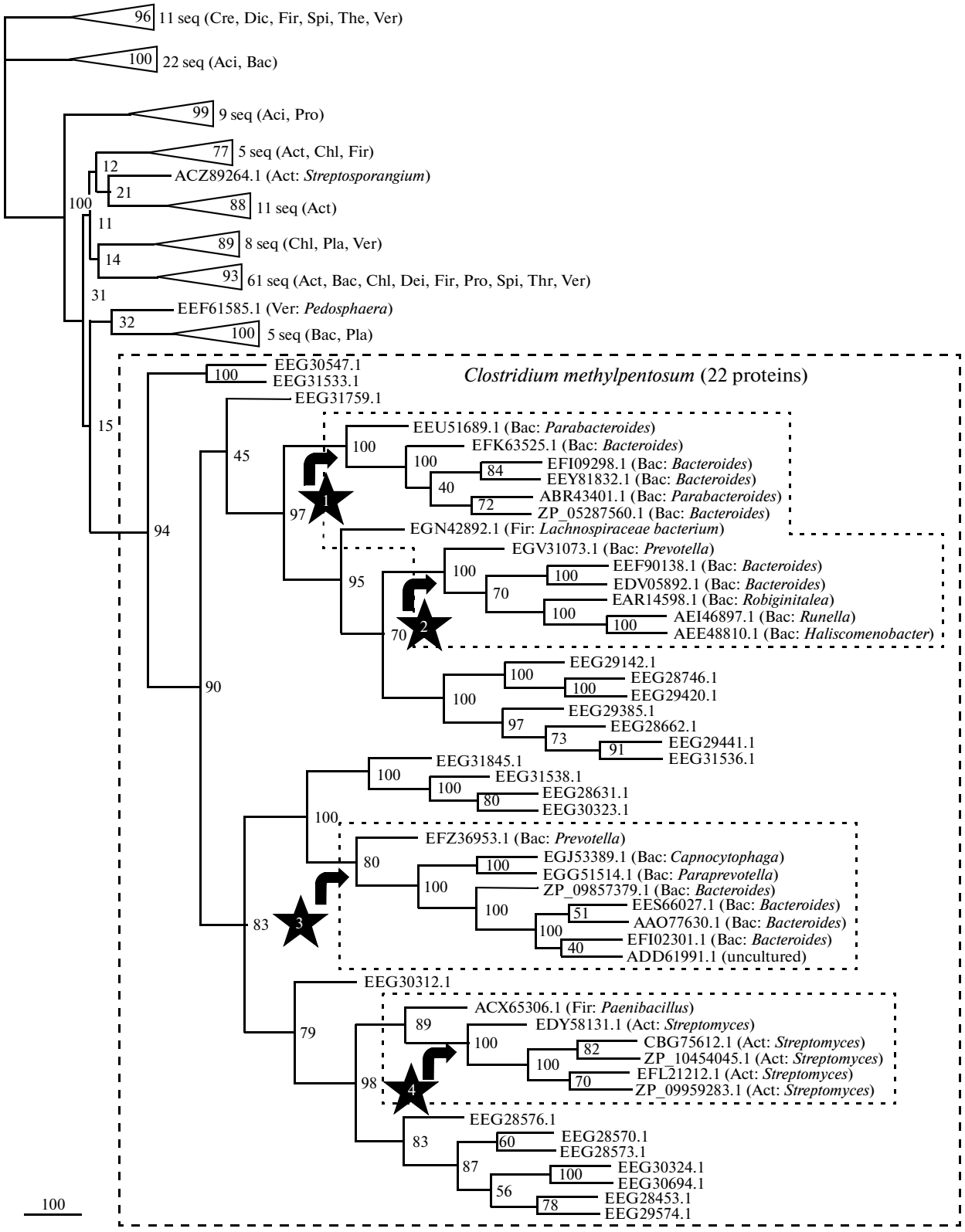



**Fig. 2.** Scheme of the phylogenetic tree of GH78 family glycoside hydrolases constructed using the neighbor-joining method and comprising 183 proteins. See legend to Fig. 1 for designations.

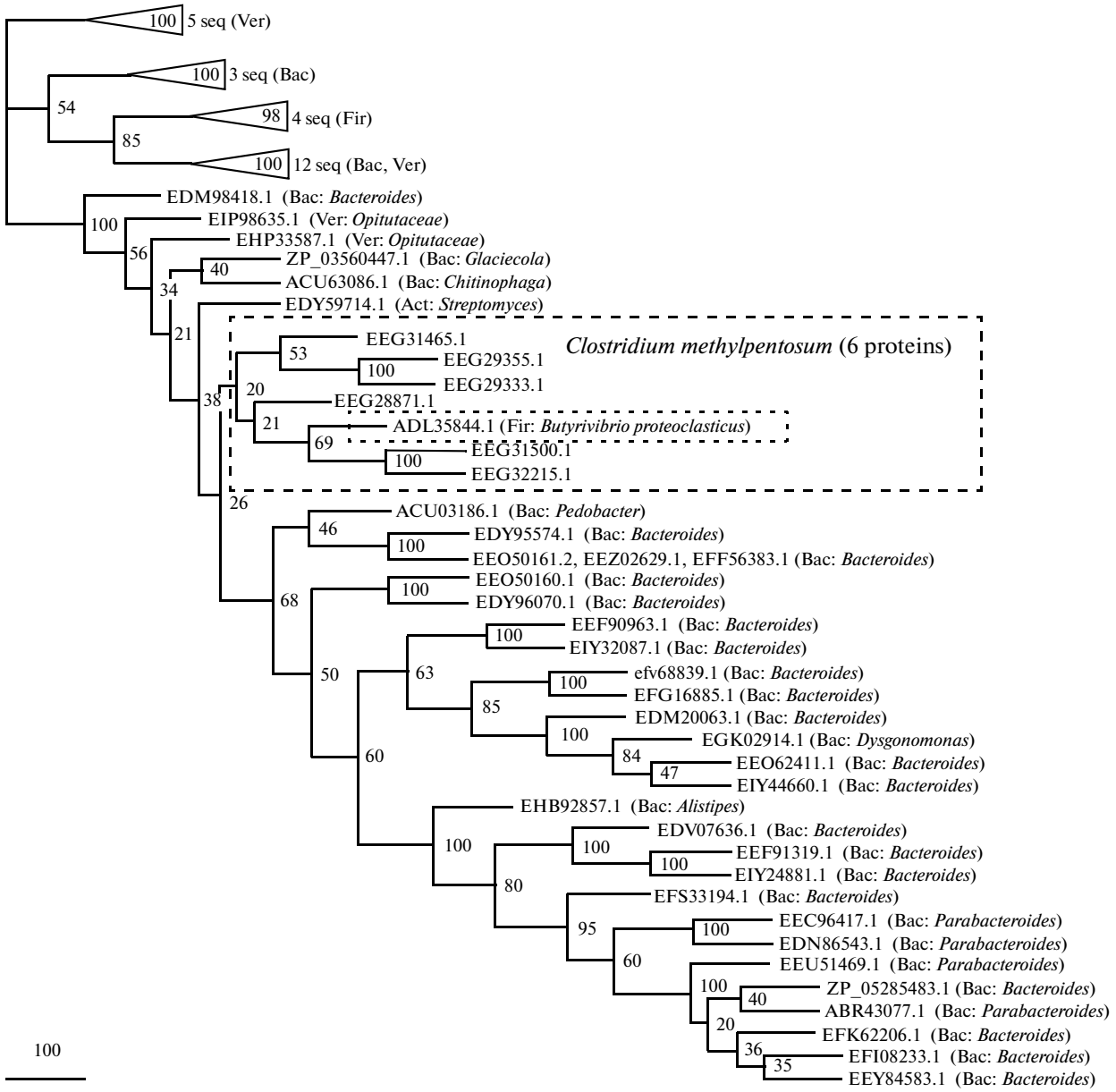



**Fig. 3.** Scheme of the phylogenetic tree of GH78 family glycoside hydrolases constructed using the maximum parsimony method and comprising 63 proteins. See legend to Fig. 1 for designations.

bacteria representing the *Actinobacteria, Bacteroidetes*, and *Firmicutes* phyla (Fig. 2). Another six proteins of *C. methylpentosum* (Fig. 3) may also originate from a single lateral transfer from a bacterium of the phylum *Bacteroidetes*. In this case as well, after a series of duplications, there was lateral transfer of one of the resulting genes to another organism (*Butyrivibrio proteoclasticus* B316, GenPept, ADL35844.1). However, the cluster of these seven proteins has a very low bootstrap support in the MP tree (20%), and it is split into two independent clusters (with 47 and 96% support) in the NJ tree (data not shown). One may not exclude that the genes coding for these six *C. methylpentosum* proteins appeared as a result of two or more lateral transfers of rather closely related genes.

Genes of two more proteins (EEG28703.1 and EEG28728.1) were also formed through duplication of a gene acquired via lateral transfer. The source of this transfer was probably a bacterium of the phylum *Actinobacteria* or *Firmicutes* (table). Apparently, the genes of the remaining seven proteins were brought by independent lateral transfers (primarily from bacteria of the phyla *Bacteroidetes* and *Firmicutes*) and were not subjected to further duplications. Among these seven proteins only one (EEG31690.1) has its closest homologue (as judged from pairwise comparison of amino acid sequences) in another bacterium of the genus

Nine proteins of *Clostridium methylpentosum* and their closest homologues

| Protein (GenPept) | Gene | Size (aa) | Fragment (domain) | Closest homologue | Closest homologue host organism | Among the 20 best blastp hits |
|---|---|---|---|---|---|---|
| EEG28617.1 | CLOSTMETH_03801 | 1704 | 877–1384 | EHR63658.1 | *Saccharomonospora cyanea* | *Actinobacteria*—7, *Firmicutes*—6 |
| EEG28703.1 | CLOSTMETH_03602 | 1516 | 554–1067 | EEG28728.1 | *Clostridium methylpentosum* | *Actinobacteria*—6, *Firmicutes*—5 |
| EEG28728.1 | CLOSTMETH_03627 | 1756 | 530–1035 | EEG28703.1 | *Clostridium methylpentosum* | *Firmicutes* —8, *Actinobacteria*—6 |
| EEG29427.1 | CLOSTMETH_02944 | 1079 | 343–811 | EHB50645.1 | *Paenibacillus lactis* | *Firmicutes*— 8, *Crenarchaeota*—8 |
| EEG29432.1 | CLOSTMETH_02949 | 1466 | 265–735 | ACT93515.1 | *Dyadobacter fermentans* | *Bacteroidetes*—17 |
| EEG30703.1 | CLOSTMETH_01630 | 1890 | 540–1019 | EHQ30823.1 | *Mucilaginibacter paludis* | *Bacteroidetes*—18 |
| EEG31496.1 | CLOSTMETH_00806 | 1244 | 654–1146 | ADY50907.1 | *Pedobacter saltans* | *Bacteroidetes*—13, *Firmicutes*—5 |
| EEG31602.1 | CLOSTMETH_00748 | 1832 | 276–822 | EGG34219.1 | *Paenibacillus* sp. | *Firmicutes* —9, *Bacteroidetes*—5 |
| EEG31690.1 | CLOSTMETH_00725 | 1804 | 557–1038 | EDO60159.1 | *Clostridium leptum* | *Firmicutes*—11, *Bacteroidetes*—6 |

Note: For each protein, GenPept identifier is indicated in the first column, the gene name, in the second column, data on the length of amino acid sequence, in the third column, and the coordinates of the GH78 glycoside hydrolase family domain, in the fourth column. This domain was used to screen the GenPept database using the blastp software. The evolutionarily closest homologue (according to the blastp data) and its host organism are indicated in the fifth and the sixth columns, respectively. The seventh column presents the prokaryotic phyla top-represented among the 20 best blastp hits of the particular *C. methylpentosum* protein (the query protein is included).

*Clostridium*, *C. leptum* DSM753 (EDO60159.1). This fact vividly demonstrates that the majority of lateral transfers discussed occurred between organisms that are not closely related.

Thus, this study allowed us to demonstrate that the diversity of hypothetical α-L-rhamnosidase genes in *C. methylpentosum* DSM5476 emerged as a result of multiple lateral transfers followed by multiple duplications of the acquired genes. The evolutionary significance of this unusual accumulation of a large number of functionally similar genes may be explained by adaptation to life in human gastrointestinal tract, where substrates containing α-L-rhamnose residues are easily available. The consequence of such an adaptation is the narrow specialization of the organism, which is capable of fermentation of only four monosaccharides (pentoses and methylpentoses): D-arabinose, L-lyxose, L-rhamnose, and L-fucose [10].

It should be noted that in the course of this work, we revealed one more microorganism containing multiple copies of genes encoding hypothetical α-L-rhamnosidases. This is a gammaproteobacterium *Glaciecola* sp. HTCC2999. Its genome (GenBank, ABST00000000.1) encodes at least seven and five proteins containing the GH78 and GH106 domains, respectively (data not shown). At the same time, according to the CAZy database [6], the genomes of *Glaciecola agarilytica* 4H-3-7+YE-5 (CP002526.1), *G. nitratireducens* FR1064 (CP003060.1), and *G. psychrophila* 170 (CP003837.1) do not encode proteins of the two families. Thus, in this case as well, evolutionarily recent multiple lateral transfers of the relevant genes occurred, some of them followed by duplications.

Another important result of the present study is the finding that proteins of the GH106 family have a significantly wider phylogenetic distribution than it has previously been acknowledged [6]. In particular, we were the first to detect the encoding genes in genomes of bacteria of the following phyla: *Deinococcus-Thermus, Dictyoglomi, Firmicutes, Lentisphaerae, Spiro-*

*chaetes*, and *Thermotogae*, as well as in the genomes of some archaea, algae, and fungi (Fig. 1).

As of June 5, 2013, there are 515 and 90 proteins of GH78 and GH106 families in the CAZy database. In this database, there are data on 43 genomes of bacteria of the genus *Clostridium*. None of these genomes encodes GH106 family proteins and only three of them encode GH78 family proteins: *C. saccharolyticum* WM1 (GenPept, ADL03771.1 and ADL04034.1), *C. saccharoperbutylacetonicum* N1-4(HMT) (AGF55542.1, AGF56667.1, AGF56675.1, and AGF56983.1), and *C. stercorarium* subsp. *stercorarium* DSM 8532 (AGC68061.1). None of the seven relevant proteins has closest homologues (as judged from pairwise comparison of amino acid sequences by blastp) among proteins of *C. methylpentosum* DSM5476, which indicates their acquisition via independent lateral transfers. Our screening of the GenPept database performed on February 28, 2013, revealed 1654 and 619 proteins containing GH78 and GH106 family domains, respectively.

## ACKNOWLEDGMENTS

The author is grateful to E.S. Naumova (State Research Center GosNIIgenetika, Russia) for critical remarks on the manuscript.

## REFERENCES

1. Rzaeva, O.M. and Varbanets, L.D., α-L-Rhamnosidase of microorganisms, *Mikrobiol. Z.*, 2006, vol. 68, no. 1, pp. 69–84. (Article in Ukrainian).
2. Manzanares, P., Vallés, S., Ramón, D., and Orejas, M., α-L-Rhamnosidases: old and new insights, *Industrial Enzymes: Structure, Function and Applications. Section A*, Polaina, J. and MacCabe, A.P., Eds., Berlin–Heidelberg: Springer, 2007, ch. 8, pp. 117–140.
3. Yadav, V., Yadav, P.K., Yadav, S., and Yadav, K.D.S., α-L-Rhamnosidase: a review, *Process Biochem.*, 2010, vol. 45, no. 8, pp. 1226–1235.
4. Ribeiro, M.H., Naringinases: occurrence, characteristics, and applications, *Appl. Microbiol. Biotechnol.*, 2011, vol. 90, no. 6, pp. 1883–1895.
5. Puri, M., Updates on naringinase: structural and biotechnological aspects, *Appl. Microbiol. Biotechnol.*, 2012, vol. 93, no. 1, pp. 49–60.
6. Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B., The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics, *Nucleic Acids Res.*, 2009, vol. 37, database issue, pp. D233–D238. http://www.cazy.org/
7. US Patent, no. 5468625, 1995.
8. Miake, F., Satho, T., Takesue, H., Yanagida, F., Kashige, N., and Watanabe, K., Purification and characterization of intracellular α-L-rhamnosidase from *Pseudomonas paucimobilis* FP2001, *Arch. Microbiol.*, 2000, vol. 173, no. 1, pp. 65–70.
9. Naumoff, D.G. and Dedysh, S.N., Lateral gene transfer between the *Bacteroidetes* and *Acidobacteria*: the case of α-L-rhamnosidases, *FEBS Lett.*, 2012, vol. 586, no. 21, pp. 3843–3851.
10. Himelbloom, B.H. and Canale-Parola, E., *Clostridium methylpentosum* sp. nov.: a ring-shaped intestinal bacterium that ferments only methylpentoses and pentoses, *Arch. Microbiol.*, 1989, vol. 151, no. 4, pp. 287–293.
11. Naumoff, D.G., Evolution of the α-L-rhamnosidases: history of the lateral gene transfers and the gene duplications, *Abstracts, The Eighth International Conference on Bioinformatics of Genome Regulation and Structure/Systems Biology (BGRS/SB'12)*, June 25–29, 2012, Novosibirsk, Russia, p. 217.
12. Naumoff, D.G., Hierarchical classification of glycoside hydrolases, *Biochemistry* (Moscow), 2011, vol. 76, no. 6, pp. 622–635.
13. Yarza, P., Ludwig, W., Euzéby, J., Amann, R., Schleifer, K.-H., Glöckner, F.O., and Rosselló-Móra, R., Update of the all-species living tree project based on 16S and 23S rRNA sequence analyses, *Syst. Appl. Microbiol.*, 2010, vol. 33, no. 6, pp. 291–299.
14. Pagani, I., Liolios, K., Jansson, J., Chen, I.-M.A., Smirnova, T., Nosrat, B., Markowitz, V.M., and Kyrpides, N.C., The Genomes OnLine Database (GOLD) v. 4: status of genomic and metagenomic projects and their associated metadata, *Nucleic Acids Res.*, 2012, vol. 40, database issue, pp. D571–D579.

*Translated by N. Kuznetsova*