

Testing the Hypothesis That the Nylonase NylB Protein Arose De Novo via a Frameshift Mutation

Salvdor Cordova, John Sanford

Submitted date: 30/07/2019 • Posted date: 31/07/2019

Licence: CC BY-NC 4.0

Citation information: Cordova, Salvdor; Sanford, John (2019): Testing the Hypothesis That the Nylonase NylB Protein Arose De Novo via a Frameshift Mutation. ChemRxiv. Preprint.

Bioinformatic tables indicating nylonases from the families of betalactamases, amidases, transpeptidases. Identification of likely catalytic residues of NylB nylonase based on bioinformatic alignments.

File list (2)

nylonase_manuscript_preprints_v2.pdf (246.02 KiB)

[view on ChemRxiv](#) • [download file](#)

nylonase_supplemental_preprints_v1.pdf (4.55 MiB)

[view on ChemRxiv](#) • [download file](#)

Title Page

Testing the Hypothesis that the Nylonase NylB Protein Arose *de novo* via a Frameshift Mutation

Authors: Salvador Cordova¹ and John Sanford²

¹FMS Foundation, 75 Lafayette Center, Lafayette Avenue, Canandaigua, NY 14424, USA;

²FMS Foundation, 75 Lafayette Center, Lafayette Avenue, Canandaigua, NY 14424, USA, corresponding author.

Salvador Cordova email – acordov4@alumni.jh.edu

John Sanford email – jsanford@fmsfound.org

Corresponding author details:

Dr. John C. Sanford

Cornell University - Associate Professor (Retired).

President FMS Foundation

Phone: 1-585-233-4046

Abstract

Background

In 1984, Susumu Ohno hypothesized that the nylon-degrading enzyme NylB arose *de novo* via a frameshift mutation within a hypothetical precursor protein (PR.C). However, Ohno never tested his hypothesis or provided supporting biological evidence. For decades, Ohno's famous frameshift hypothesis has been uncritically accepted as the correct explanation for the origin of NylB. In this paper we have surveyed the literature relevant to Ohno's NylB claims as well as the various alternative models that have been published regarding the origin of NylB. We have employed a new and rigorous method to test Ohno's hypothesis.

Results

We searched multiple protein databases to determine the distribution of NylB and any possible homologs. We then determined the distribution of other known nylonases and their possible homologs. We also determined the distribution of Ohno's hypothetical PR.C. protein and any possible homologs. Lastly, we determined what protein families the various nylonases belong to.

We found that the NylB protein is widely occurring, has thousands of homologs, and is found in diverse organisms and diverse habitats. Likewise, we found that the other known nylonases are also widely occurring, have thousands of homologs, and are found in diverse organisms and diverse habitats. However, the hypothetical PR.C protein does not show up in any of the same databases, and there is no evidence of any homologs. Conserved domain searches showed that NylB is a member of the beta lactamase protein family – a highly conserved family of enzymes. Likewise, the other known nylonases belong to well-characterized enzyme families.

Conclusions

Our results very effectively falsify Ohno's frameshift hypothesis, regarding the origin of the NylB nylonase. Like NylB, none of the other nylonases that we examined were substantively novel, all had thousands of homologs, and they were found in diverse organisms and diverse habitats. These findings not only falsify Ohno's frameshift hypothesis, they also falsify the long-held assumption that all nylonases evolved after the invention of nylon in 1935.

Keywords

Nylonase, NylB, NylA, NylC, Beta lactamase, 6-aminohexanoate hydrolase, Frameshift

Background

Nylon is a synthetic polymer that was invented in 1935. From the mid-1950's onward, a variety of enzymes have been discovered that can degrade nylon-6 oligomers into their monomer components. [1,2] Such enzymes have been colloquially referred to as “nylonases”. It has been widely assumed that all nylonases have evolved since 1935. However, there are now many reasons to doubt this assumption.

Below we review various hypotheses that have been developed to understand the origin of nylonase enzymes – in particular the origin of NylB. The three primary hypotheses for the origin of NylB are: 1) Okada *et al.*'s post-1935 gene duplication and mutation hypothesis; [3] 2) Ohno's post-1935 frame shift hypothesis; [4] and 3) Yomo *et al.*'s pre-1935 NylB homologs hypothesis. [5]

In 1957, Ebata and Morita discovered the first enzyme that could breakdown nylon. [6] They found that Trypsin, a widely conserved enzyme in mammals, was shown to degrade nylon-6 oligomers. This capability obviously existed in Trypsin prior to the invention of nylon in 1935. Trypsin is a protease, and nylon has some protein-like molecular features (Figure 1). Therefore, it should not be surprising that Trypsin might degrade nylon. It is important to note that many enzymes that existed long before the invention of nylon might still manifest “nylonase” activity. This does not necessarily imply a newly evolved enzyme function. Hence application of the term “nylonase” can be ambiguous. We will use the term *nylonase* to refer to all enzymes with measureable or predicted nylonase activity.

Figure 1

Structures of 6-aminohexanoic acid, lysine, and amide bonds in peptides and nylon-6 linear oligomers. *Top*: structures for 6-aminohexanoic acid and lysine are shown. *Middle*: a generic oligopeptide is shown with amide bonds highlighted in blue with green boxes. *Bottom*: a generic linear nylon-6 oligomer is shown with amide bonds highlighted in blue with green boxes. For both oligomer structures, the individual monomers are shown in black (first), red (second), and light blue (third). The squiggly line and dots at the right indicate additional monomers.

In 1966, Fukumura first discovered that a bacterium (*Corynebacterium aurantiacum B-2*) could metabolize nylon [7] and he isolated two of the enzymes involved. [8]

From the mid 1970's to early 1980's Kinoshita, Okada and others published a series of papers on isolation of two nylonase enzymes (eventually named NylA and NylB) from *Achromobacter gutatus KI72* (renamed *Flavobacteria KI72*). [1,3] The corresponding genes were on the plasmid pOAD in KI72. [9] A paralog of NylB named NylB' was also discovered, which had substantially lower nylonase capability than NylB. [2] In 1993 one more nylonase, called NylC, was discovered on the same plasmid in the same bacterium. [10] The natural ability of the KI72 strain to metabolize nylon is apparently due to the coordinated action of this set of four linked complementary nylonase genes. [2] Yet Kinoshita claimed that all these genes were “newly evolved” since 1935. [11]

Okada *et al.* was the first to present a specific hypothesis regarding the origin of NylB. In 1983, Okada argued NylB was a paralog that arose via a gene duplication event from a linked gene coding NylB'. He assumed this must have occurred sometime after 1935. [3] His model required that the duplicate gene, NylB, would need to acquire 47 residue substitutions via point mutations in just a few decades. Although the paralogous nature of NylB and NylB' suggests a gene duplication event, there was no direct evidence that it happened post 1935, and he gave no reason why NylB' might not have arisen from NylB instead of the reverse.

In his 1984 paper, Susumu Ohno offered a second major hypothesis for the origin of NylB. Ohno criticized Okada's 1983 hypothesis because it required too many point mutations to effect so many amino acid substitutions in so little time. Ohno said, "so extensive an amino acid divergence is not expected to occur in so short a time span."

Ohno took Okada's published sequence known as RS-IIA (which encoded NylB) and constructed a hypothetical sequence he called PR.C by simply deleting a single nucleotide from the RS-IIA sequence and relabeling it as PR.C. Ohno then claimed PR.C was the ancestral sequence of NylB. He claimed that shortly after 1935, a single nucleotide insertion in the gene encoding his hypothetical PR.C protein yielded the present-day RS-IIA sequence that now encodes NylB. Ohno criticized Okada's hypothesis as being unrealistic because it required so many point mutations, yet Ohno's hypothesis required an essentially random amino acid sequence arising from a frameshift to instantly form a stable, functional, and specific enzyme. Ohno had no direct evidence that the hypothetical PR.C protein even existed, and his frameshift mutation was purely hypothetical. Yet Ohno put forward his hypothesis so forcefully that readers accepted his model as if it were history, and his paper continues to be cited as if the hypothetical frameshift mutation was actually an observed fact. [12,13,14,15,16,17]

Just 5 years after Ohno published his frameshift hypothesis, Kanagawa *et al.* discovered another NylB enzyme in another bacterium, *Pseudomonas* NK87, which also had the ability to degrade nylon-6. [18] This effectively falsified Ohno's claim that NylB was unique. This new NylB gene sequence was highly divergent, having only 53% DNA similarity, [19] and had only 35% protein sequence similarity compared to Kinoshita's NylB in KI72 (the one Ohno claimed was truly unique). Kanagawa designated this newly discovered NylB as p-NylB and re-named the previously discovered NylB and NylB' proteins as f-NylB and f-NylB', respectively.

In 1991, Kato *et al.* attempted to explore Okada's hypothesis by experimentally mutating the 47 amino acids in NylB' that were divergent from NylB. [20] They discovered that only two of the 47 amino acids were required to enhance nylonase activity in NylB' up to the level of NylB. The two linked genes coding for NylB and NylB' were substantially divergent (making duplication and divergence in just a few decades very unlikely), yet they were still sufficiently homologous to rule out a single frame shift for the origin of two proteins simultaneously. [4]

In 1992, in response to Kanagawa's discovery of p-NylB, Yomo *et al.* co-authored a paper with Urabe and Okada, to put forward a third competing hypothesis regarding the origin of NylB. Yomo *et al.* argued that Kinoshita's f-NylB and Kanagawa's p-NylB homologs descended from a common ancestor that existed about 140 million years ago. [5] Yomo *et al.* wrote: "The distance between P-nylB and F-nylB (or F-NylB') is much larger than between F-nylB and F-NylB'. The time divergence of F-nylB and P-nylB is estimated to be at least 1.4×10^8 years... Therefore, most of the amino acid substitutions from the ancestor of the *nylB* gene family to its descendants of today might have occurred before the beginning of nylon manufacture."

In 1995, experiments by Prijambada, Negoro, Yomo, and Urabe, showed that strains of the bacteria *Pseudomonas aeruginosa* PAO1 which initially lacked activity toward nylon-6 linear and cyclic dimers could be selectively evolved into a strain that could digest these dimers. [21] The evolved descendant from the ancestral PAO1 that had nylon digesting capability was designated PAO5502. However, Prijambada *et al.* point out, "a molecular basis for the emergence of nylon oligomer metabolism in PAO5502 is still unknown."

In 2007, Sudhakar demonstrated that strains of *Bacillus cereus* found in the Indian Ocean could digest nylon-6. [22] To confirm this, we searched protein data bases for evidence of NylB in *Bacillus cereus*.

We will show that multiple lines of evidence falsify Ohno's hypothesis, but are consistent with Yomo's model. Ohno had three primary claims: a) he claimed the NylB protein never existed until sometime after 1935; b) he claimed NylB arose as a *de novo* protein as the result of a frameshift mutation in a precursor protein; c) he claimed he knew the exact sequence of his hypothetical precursor protein.

Since the sequences of the NylB protein and Ohno's hypothetical protein are both known, Ohno's hypothesis is now readily testable using protein databases.

If Ohno's hypotheses were correct, then a protein database search should reveal evidence that Ohno's hypothetical precursor protein actually existed, had a history, and so should have many protein homologs. At the same time there should be clear evidence that the NylB protein really is a unique protein, with no history, and no protein homologs.

Conversely, if Ohno's hypotheses were wrong, then a protein database search should reveal evidence that the hypothetical precursor protein never existed, had no history, and should have few if any homologs. At the same time there should be evidence that the NylB protein is *not* unique, has a history, and has numerous homologs.

Results

We used UniProt to determine the number of organisms with various explicit nylonase names, along with Ohno's PR.C sequence. Searches for potential remote homologs were also made using psi-BLAST and SPARCLE. The SPARCLE numbers were not adjusted for redundancies, but were provided to give an idea of the degree of representation of the homologs in the databases which SPARCLE surveys. The results are shown in Table 1.

Table 1
Search results for homologs of NylB, NylB', NylA, NylC, 6-aminohexanoate hydrolase, and Ohno's hypothetical PR.C.

Protein Name	Organisms with Protein in UniProt	psi-BLAST Potential Homologs	NIH SPARCLE Potential Homologs
NylB	201	20,000+	68,000+
NylB'	134	20,000+	68,000+
NylA	193	20,000+	53,000+
NylC	7	6,936	44
6-aminohexanoate Hydrolase	2115	20,000+	68,000+
Ohno's Hypothetical PR.C	0	0	0

From the lists of predicted nylonases generated through UNIPROT, data was gathered in the Conserved Domain Database (CDD) and then tallied to see the most common domain family for each nylonase (Table 2).

Table 2
Homology of various nylonases to known enzyme families such as beta lactamases, amidases, and peptidases.

Protein	Most Representative Conserved Domain Family	Conserved Domain Accession Numbers
----------------	--	---

NylB	beta lactamase	COG1680, pfam00144
NylB'	beta lactamase	COG1680, pfam00144
NylA	amidase	COG0154, c118951, pfam01425, PRK07869, PRK06061, pfam07501, PRK06529
NylC	peptidase	cd00123, c100603, cd02252

The data used to construct Table 1 and 2 can be found in the in Supplementary Tables S1, S2, S3, S4, and S5.

Several proteins labeled as NylB were also labeled as beta lactamases. Several proteins labeled as NylA were also labeled as amidases. The listing of these proteins with their accession numbers can be found in the Supplementary Tables S6 and S7.

CDD analysis (as provided by GenBank), scored the similarity of the NylB that Ohno studied and COG1680 beta lactamase at 130 bits, which implies the probability that a random amino acid polymer would achieve that level of similarity to the archetypal COG1680 beta lactamase is one chance in 2^{130} .

A representative sample of these proteins (mostly those reported with experimentally verified nylonase activity) were also aligned using the MUSCLE alignment algorithm to show some of the conserved features of the NylB homologs, particularly the Serine-X-X-Lysine motif. This Serine-X-X-Lysine motif has been confirmed by X-ray crystallography of NylB in *Arthrobacter KI72*. [23] The first 8 of 10 proteins in the alignment (figure 2) were proteins from organisms that had experimental evidence of nylonase NylB activity, and the last two were provided for comparison as they are remote homologs with only predicted NylBs (as of this writing).

We searched the databases for a NylB homolog in *Pseudomonas aeruginosa* PAO1 and found a sequence (accession AAG07735.1) that had 100% identity (96% coverage) to a predicted NylB in another strain of *Pseudomonas aeruginosa* (accession CKI08964.1). AAG07735.1 was used as one of the proteins featured in the MUSCLE alignment.

Figure 2

MEGA 6.0 MUSCLE Alignment of a critical segment of selected NylB and NylB' homologs showing conserved motifs, particularly the Serine-X-X-Lysine motif confirmed by X-ray crystallography. Using the GenBank Accession numbers for each row: (1) NylB *Flavobacteria KI72* CAA24927.1, (2) NylB' *Flavobacteria KI72* CAA26616.1, (3) NylB *Agromyces KYR5* BAE97621.1, (4) NylB' *Agromyces KYR5* BAE97630.1, (5) NylB *Bacillus cereus* SME41718.1, (6) NylB' *Bacillus cereus* SME45458.1 (7) NylB *Pseudomonas aeruginosa* PAO1 AAG07735.1, (8) NylB *Pseudomonas NK87* BAA01524.1, (9) NylB' *Paenibacillus sp.* AD87 OAX47860.1, (10) Beta-Lactamase/NylB *Escherichia coli* SCQ13749.1.

The complete MUSCLE alignment is available in Supplementary Table S8.

For completeness we did a distance computation using MUSCLE aligned sequences and found that the NylB and NylB' paralogs in it had 11% sequence divergence in KI72 and similarly the NylB and NylB'

paralogs in *Bacillus cereus* had 12% sequence divergence, yet the distance between the NylB in K172 and NylB in *Bacillus cereus* was around 75%. The distance matrix can be also found in Supplementary Table S8.

Discussion

Ohno's famous 1984 paper claimed to show how a frameshift mutation might have given rise to a totally novel protein such as a nylonase enzyme. Ever since that time it has been widely accepted that this was the correct explanation for the origin of the NylB enzyme. [12,13,14,15,16,17] It has been widely assumed that this happened in an extremely short timeframe, soon after the invention of nylon in 1935. Some people, by extension, have assumed that Ohno's frameshift claim might help explain other nylonases such as NylA and NylC. Most broadly, Ohno's frameshift paper is considered by many to be the best proof of the rapid evolution of a *de novo* gene/protein.

Many readers have not realized that Ohno's 1984 claims were not supported by any type of evidence – his model was entirely speculative. Ohno presented his assertions very forcefully, as if they were facts. It seems that many readers of that paper got the impression that Ohno actually had observational evidence for the existence of his specified precursor protein and his specified frameshift mutation.

Experiments by Kato *et al.* in 1991 and Prijambada *et al.* in 1995 failed to confirm that NylB evolved via frameshift mutation, and in fact they argue against Ohno's hypothesis. In the case of Kato's experiment, it suggests NylB nylonase evolution is feasible by as little as two amino acid changes in an ancestral homolog rather than a frameshift mutation affecting 400+ amino acids (as in Ohno's hypothesis).

In the case of Prijambada *et al.*'s experiment where nylon digesting ability was evolved via directed evolution in the lab, the presence of a NylB homolog in strains of *Pseudomonas aeruginosa* PAO1 suggests the lab-based directed evolution of nylon digestion from PAO1 involved point mutations of a pre-existing NylB homolog in PAO1, not a frameshift mutation.

Careful reading shows that Ohno's proposed precursor protein and his proposed frameshift mutation were only inferred. Therefore, at that time Ohno did not even have a testable hypothesis. In addition to the experimental evidence that could reasonably be deemed sufficient to refute Ohno's hypothesis, now, in the age of bioinformatics, we can do what Ohno could not do – we can further test his model using bioinformatic tools.

If Ohno's hypotheses were correct, then a protein database search should reveal evidence for the existence of Ohno's hypothetical precursor protein, which should have a history and should have protein homologs. On the flip side, there should be evidence that the NylB protein is a unique protein, with no history and no protein homologs.

Conversely, if Ohno's hypothesis were wrong, then a protein database search should reveal evidence that the hypothetical precursor protein never existed, has no history, and has few if any homologs. At the same time there should be evidence that the NylB protein is *not* unique, and so has a history and numerous homologs.

An objection might be raised that the psi-BLAST and SPARCLE results were artificially inflated by redundancies and spurious hits for NylB because of the search parameters we used. This objection would not be valid. First, these parameters were the defaults set by the NIH, and second, the point of the comparison was to show that even under relaxed and generous parameters, no remote potential homologs of PR.C could be detected. Because NylB is in the family of beta lactamases, and NylA is in the family of amidases, therefore NylB and NylA are both clearly members of very well-known protein families (independent of the psi-BLAST and SPARCLE results).

Our results are very clear. Numerous protein databases show that there is no evidence that Ohno's hypothetical precursor protein ever existed. We found zero instances of the protein, and zero protein homologs.

This by itself is strong evidence against Ohno's hypothesis. But the most conclusive proof that Ohno's hypothesis is false is that the NylB gene is not at all unique – it is found in many organisms, in many habitats, and has a vast number of homologs.

Our search results indicate that homologs of NylB and various other 6-aminohexanoate hydrolases are very abundant. Some organisms with these homologous proteins have been experimentally shown to have the ability to digest nylon, [22,24] but most were not enzymatically tested. While sequence-based gene predictions cannot prove that all such NylB homologs can necessarily degrade nylon, such predictions point to a family of proteins that have very significant homology. All of the genes with the NylB designation (from our UNIPROT-developed list which also had available CDD pages) had beta lactamase domains (Supplementary Table S1). Beta lactamases are considered one of the most ancient proteins. [25] The divergence within the NylB class of enzymes was often very substantial. This precludes the possibility that all such enzymes arose from an isolated frame shift mutation that arose sometime after 1935. It should be obvious that a single frameshift mutation, in just a few decades, could not possibly have proliferated via horizontal gene across a very large number of unrelated organisms found all around the world.

Ohno's hypothesis was based upon Kinoshita's NylB protein sequence. We computed the sequence similarity of NylB to the architecture of the beta lactamase protein family. Based on the bits score assigned to this particular NylB gene by CDD (accession COG1680), we found these sequences were strikingly similar. According to CDD, the probability that this similarity to a COG1680 beta lactamase would arise by chance is 2^{-130} . Given the degree of non-random similarity of NylBs to beta lactamase domains (Supplementary Table S1), there is no doubt of the homology of NylB with beta lactamases, and this is illustrated by the fact that several entries in GenBank list the same protein as a NylB homolog and beta lactamase simultaneously (Supplementary Table S6).

Ohno's hypothesis was premised upon numerous assumptions that we now know are incorrect, and so his hypothesis is falsified on several levels:

1. The widely held assumption that all nylonase enzymes evolved since 1935 was incorrect.
2. Therefore, Ohno's assumption that the NylB protein was a new and unique protein was incorrect.
3. Ohno assumed a hypothetical but specific precursor protein that now appears to have never existed, and thus Ohno's hypothetical frame-shift mutation appears to have never happened.
4. Ohno claimed that a random string of amino acids could reasonably be expected to give rise to a specific, functional, beneficial, and stable enzyme. Having all these things happen by chance is so incredibly unlikely that it is hard to imagine. This is especially clear in light of the fact that CDD database indicates that the probability of NylB would be so similar to beta lactamase by chance would be essentially impossible (2^{-130}).
5. Lastly, it was established 8 years after Ohno published his paper that NylB was not operating independently, but was clearly part of a catabolic chain, functioning in coordination with three other nylonases on the same plasmid (NylA, NylB', and NylC). Indeed, NylB was shown to be co-regulated with NylC, sharing the very same promoter.[10] Thus, Ohno had no access to this data at the time he published his hypothesis.

Ironically, Ohno pointed out that the level of divergence of paralogous pair of NylB and NylB' in K172 suggests that this paralogous pair must have existed prior to 1935. [4] A similar level of divergence exists in the paralogous NylB and NylB' proteins in *Bacillus cereus* even though the NylB in *Bacillus cereus* is around 75% divergent from the NylB in K172. These considerations themselves cast serious doubt on any post-1935 gene duplication hypothesis. Furthermore, it appears that in the case of *Pseudomonas NK87* (with a functioning nylonase NylB), that having a paralog is unnecessary for the

evolution of NylB nylonase activity. Importantly, since Ohno's hypothesis only applies to K172, it cannot account for the presence of the NylB paralogs in *Bacillus cereus* that are over 75% divergent from their counterparts in K172, nor the NylB orthologs in the *Pseudomonas* strains.

Taken collectively, our findings very clearly refute Ohno's frame-shift hypothesis. However, our findings are consistent with Yomo *et al.*'s hypothesis that the NylB gene and its homologs have been around for a long time.

In the beginning of Ohno's 1984 paper, he describes why he developed his nylonase hypothesis. He had no particular interest in nylonases – he wanted to show that frameshift mutations might be a powerful way to evolve novel and beneficial proteins instantly. The falsification of Ohno's nylonase hypothesis does not in itself prove that beneficial novel proteins might not possibly arise via frame shift mutations. However, the spectacular failure of Ohno's nylonase model puts his broader thesis in very serious doubt.

Overlapping reading frames are known to exist in biology, [26,27] and Okamura [28] has speculated that several such human genes may have originated via frame shift mutations. However, it is worth noting that Okamura believed that Ohno's claims were correct when he proposed this. With the falsification of Ohno's famous hypothesis, it becomes very reasonable to question the evolutionary importance of novel and useful proteins arising instantly by frameshift mutations.

We extended our search to look for homologs of other nylonases such as NylB', NylA, and NylC (all of which were assumed to have evolved since 1935). While Kinoshita did not detect physiological amidase activity for NylA, [1,9] our analysis clearly shows that NylA has amidase homology. Similarly, we found that NylC was homologous to a rare peptidase. We found several proteins had dual classifications such as beta lactamase and 6-aminohexanoate hydrolase (NylB), or amidase and 6-aminohexanoate cyclic hydrolase (NylA). In addition to experiments with proteases like Trypsin,[6]experiments have shown that even triacylglycerol lipases can act as nylonases.[29] Thus it appears that the term "nylonase" could be applied to members of the protease, beta lactamase, amidase, peptidase, and lipase enzyme families. This is in broad agreement with some of Yasuhira *et al.* and Negoro's findings that NylB and NylB' are in the beta lactamase family, NylA is in the amidase family, and some nylonases have similarity to lipases. [13,23,30] In every case the proteins were found in various organisms and in various natural habitats - along with a great many homologs. We conclude that all of these nylonases and their close homologs existed prior to 1935, although in some cases there may have been adaptive modifications after 1935. It appears that these various naturally occurring enzymes that happen to be able to degrade nylon have historically acted upon alternative nylon-like substrates.

Conclusions

The focus of this research has been to test Ohno's claim that sometime after 1935 the "nylonase" NylB arose *de novo* via a frameshift mutation in a precursor gene/protein. Ohno's hypothesis has been historically impactful - being considered a powerful proof that new genes and enzymes can instantly arise *de novo*. While Ohno's model was speculative and was never actually tested, it has been uncritically accepted within the scientific community for several decades.

In the last three decades, there have been some authors who have questioned Ohno's hypothesis and have proposed alternative explanations of how NylB might have arisen. While that work did suggest Ohno's hypothesis might be wrong, these authors did not rigorously falsify Ohno's hypothesis, and so the frameshift model has continued to be cited as a fact – even up to the time of this writing.

Thanks to protein databases that were not available until quite recently, we have been able to unambiguously falsify Ohno's hypothesis. We have shown that Ohno's hypothesis can be falsified on multiple levels.

More broadly, we have examined the widely-held assumption that there were no enzymes having nylonase activity prior to the invention of nylon in 1935. Ohno shared this assumption with most of the scientists of his day. However, the primary “nylonases” that have been studied (NylA, NylB, NylB', and NylC), were all found on the same plasmid, functioning in coordination, suggesting that none of these genes/proteins could have arisen *de novo* in the very recent past. Our database searches show that all of these enzymes are widely distributed in the biosphere and have thousands of homologs. We also show that these enzymes belong to well characterized enzyme families that are ancient. It is clear that numerous enzymes existed prior to the invention of nylon, which had previously been acting on other substrates, but also happened to have “nylonase-like” activity. In the future the term *nylonase* might be used with more caution.

Methods

The GenBank accession numbers for the database searches for NylB were established by typing in the sequence from Ohno's original 1984 paper into the NCBI BLAST tool. The sequences in Ohno's paper were derived from Okada's 1983 paper that reported the sequences of several nylonases in *Flavobacteria* K172. Subsequent papers discovered more nylonases on the same plasmid, and all of the K172 nylonases were designated NylA, NylB, NylB', and NylC.[3,10] These sequences were used to define the searches for nylonase homologs under those names, specifically in terms of GenBank accession numbers BAA05090.1 for NylA, CAA24927.1 for NylB, CAA26616.1 for NylB', and BAA01528.1 for NylC.

The gene coding for the NylB protein was contained in a segment of DNA Okada *et al.* called RS-IIA. [3] It is worth mentioning that it appears Ohno mislabeled Okada's RS-IIA as R-IIA in his paper. [3,4] Also, it appears Ohno made either a typo in transcription or failed to clearly account for the creation of a premature stop codon in construction of his PR.C from the RS-IIA sequence. Okada's paper and GenBank indicate that the end of Ohno's PR.C (derived from RS-IIA) should be “GCGGCGTGA,” not “GCGGCTGA” as was the case in Ohno's paper. Given that Okada's paper was the source of the actual sequence data, with Ohno's work deriving from that paper, the error must be Ohno's and not Okada's.

The UNIPROT searches were easily conducted by simply going to the uniprot.org website and typing search terms such as “NylB”, “NylB'”, “NylA”, “NylC”, and “6-aminohexanoate hydrolase.” Lists of proteins for each of these nylonases were created by using a simple Java program to filter out duplicate experimental entries. Afterward, manual review of the filtered lists was also conducted to remove spurious search results. An example of such a spurious result was the result induced by organisms like *Nylandria* that happened to have the “Nyla” string in them. These lists were used as the foundation for tallying CDD data found in the Supplementary Tables.

Gathering of the CDD results for each protein can be illustrated by the example of NylB in K172. The gene accession number for the *nylB* gene in K172 was determined to be X00046.1, and for the NylB protein CAA24927.1. We then went to Protein Database at the URL <https://www.ncbi.nlm.nih.gov/protein> and entered CAA24927.1 as the search term. This brought up a page that provided a hyperlink titled “identify conserved domains” which links to a Conserved Domain Database (CDD) page. The CDD page showed similarity of this NylB homolog to a beta lactamase domain with the CDD ID of COG1680. Also, on this page is the probability bit score of similarity of this NylB homolog to the COG1680 beta lactamase domain.

The CDD page also provided a hyperlink titled “domain architecture ID 10004149”, which links to the SPARCLE viewer that gives the statistics for the variety of entries in the databases that conform to the NylB architecture (ID 10004149) associated with Ohno/Okada's NylB homolog. Because of the dynamic nature of databases constantly being updated, the exact numbers provided by the SPARCLE viewer change from day to day. A snapshot of the statistics reported by SPARCLE were then recorded in the results section, but these numbers should not be expected to be the same over time given the ever-expanding size of the databases which SPARCLE is surveying.

Similar procedures were used to tally the conserved domains and SPARCLE numbers for all the other proteins in the lists generated by UNIPROT for NylB, NylB', NylA, and NylC. The CDD results for NylB, NylB', NylA, and NylC were listed in Supplementary Tables S1, S2, S3, S4. No CDD analysis was conducted for the general 6-aminohexanoate hydrolases since they are a collection of several enzymes that include NylB, NylB', NylA, and NylC plus many others.

The psi-BLAST searches were accomplished through the NCBI protein BLAST website. The psi-BLAST algorithm option was enabled by checking the appropriate box in the program selection section. Under the algorithm parameters section, Max Target sequences was reset from the default of 500 to 20,000. All other algorithm parameters were at the default settings which were: Automatically adjust parameters for short input sequences, Expect threshold = 10, word size = 3, Max matches in a query range = 0, Matrix = BLOSUM62, Gap Cost = Existence 11 Extension 1, Compositional adjustments = Conditional compositional score matrix adjustment, no filters or masks, psi-blast threshold = .005, pseudocount = 0. Only 1 psi-BLAST iteration was run for each sequence.

The homologs chosen to search for NylB, NylB', NylA, NylC were from Flavobacteria KI72 since these were the first named homologs in Ohno's paper as well as subsequent papers studying nylonases in Flavobacteria KI72. Since NylA, NylB, NylC were all 6-aminohexanoate hydrolases, the union of all the entries with homologs to these three nylonases would be at least as large as the largest set, namely that for NylB, thus the figure of 20,000+ for 6-aminohexanoate hydrolases was entered in Table 1.

To verify the possible conflict in assigning functions to protein predictions, a manual search through several thousand potential NylB and NylA homologs created by psi-BLAST was painstakingly reviewed and a listing of some of the notable conflicts were listed in Supplementary Tables S6 and S7.

As shown in Supplementary Table S9, we constructed a set of 10 protein sequences in FASTA format that was composed of a mix of NylB homologs from organisms that had been mentioned in literature as having credible evidence of nylonase activity and for comparison included in the mix one NylB having a weak e-value compared to NylB from KI72 as well as a NylB homolog from a well-known organism, *E. coli*. In the case of *Pseudomonas aeruginosa* PAO1, although the sequence (accession AAG07735.1) is not listed as NylB, it is 100% identical (96% coverage) to a sequence in another strain of *Pseudomonas aeruginosa* in GenBank that goes by the name of NylB (accession CKI08964.1). For *E. Coli*, the gene was called *nylB*, even though the predicted expressed protein is called a beta lactamase (accession SCQ13749.1). Data in Supplementary Table S9 was the basis of the MUSCLE alignment.

The sequences in Table S9 were then put in MEGA 6.0 to generate MUSCLE alignments. We confirmed by inspection that the Serine-X-X-Lysine motif that appears in the MUSCLE alignment (Figure 2) agreed with Negoro's X-ray crystallography of NylB. [20,23] The amino acid sequences in PAO1, which corresponded to the sequence in KI72 in Kato *et al.* experiment, were determined by inspection of the MUSCLE alignment. The distance matrix was generated also by MEGA with the parameters stated in Supplementary Table S8.

List of Abbreviations

BLAST – Basic Local Alignment Search Tool
CDD – Conserved Domain Database
MEGA – Molecular Evolutionary Genetics Analysis
MUSCLE – Multiple Sequence Comparison by Log-Expectation
NCBI – National Center for Biotechnology Information
NIH – National Institutes of Health
SPARCLE – Subfamily Protein Architecture Labeling Engine
UniProt – Universal Protein Resource

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files which were derived from public databases of UNIPROT, GenBank, SPARCLE, CDD as described in the materials and methods section.

Competing interests

The authors declare they have no competing interests.

Funding

Not applicable.

Authors' contributions

SC did the bioinformatics searches and compiled results. SC and JS analyzed the results and wrote the main manuscript. Both authors reviewed the manuscript.

Acknowledgements

Not applicable.

Authors' information

JS is a retired research professor, Cornell University and President of FMS Foundation, Inc. SC is a research associate for FMS Foundation, Inc.

References

1. Kinoshita S, Negoro S, Muramatsu M, Bisaria VS, Sawada S, Okada H. 6-Aminohexanoic Acid Cyclic Dimer Hydrolase A New Cyclic Amide Hydrolase Produced by *Acromobacter guttatus* KI72. *Eur. J. Biochem.* 1977;80:489-495.
2. Negoro S. Biodegradation of nylon oligomers. *Appl Microbiol Biotechnol* 2000 October;54:461-466.
3. Okada H, Negoro S, Kimura H, Nakamura S. Evolutionary adaptation of plasmid-encoded enzymes for degrading nylon oligomers. *Nature.* 1983 November 10;306.
4. Ohno S. Birth of a unique enzyme from an alternative reading frame of the preexisted, internally repetitious coding sequence. *Proc. Natl. Acad. Sci. USA, Evolution.* 1984 April;81:2421-2425.
5. Yomo T, Urabe I, Okada H. No stop codons in the antisense strands of the genes for nylon oligomer degradation. *Proc. Natl. Acad. Sci. USA, Evolution.* 1999 May;89:3780-3784.
6. Ebata M, Morita K. Hydrolysis of ϵ -aminocaproyl compounds by trypsin. *The Journal of Biochemistry.* 1959;46(4).
7. Fukumura T. Hydrolysis of cyclic and linear oligomers of 6-aminocaproic acid by bacterial cell extract. *The Journal of Biochemistry.* 1966;59(6).

8. Fukumura, T. Two bacterial enzymes hydrolyzing oligomers of 6-aminocaproic acid. *The Journal of Biochemistry*. 1966;59(6).
9. Negoro S, Taniguchi T, Kanaoka M, Kimura H, Okada H. Plasmid-determined enzymatic degradation of nylon oligomers. *Journal of Bacteriology*. 1983 July;155(1): 22-31.
10. Negoro S, Kakudo S, Urabe I, Okada H. A new nylon oligomer degradation gene (nylC) on plasmid pOAD2 from a *Flavobacterium* sp. *Journal of Bacteriology*. 1992 Dec;174(24):7948-7953.
11. Kinoshita S, Terada T, Taniguchi T, Takene Y, Masuda S, Matsunaga N, Okada H. Purification and Characterization of 6-aminohexanoic-acid-oligomer hydrolase of flavobacterium sp. Ki72. *Eur J. Biochem*. 1981;116:547-551.
12. Sabath N, Wagner A, Karlin D. Evolution of viral proteins originated *de novo* by overprinting. *Mol Biol Evol*. 2012 Dec;29(12):3767-3780.
13. Ruiz-Orera J, Hernandez-Rodriguez J, Chiva C, Sabido E, Kondova I, Bontrop R, *et al*. Origins of *de novo* genes in human and chimpanzee. *PLoS Genet*. 2015 Dec 31; 11(12):e1005721. doi: 10.1371/journal.pgen.1005721
14. Dias R, Austin M, Kolaczkowski O, Kolaczkowski B. Convergence of Domain Architecture, Structure, and Ligand Affinity in Animal and Plant RNA-Binding Proteins. *Mol Biol Evol*. 2017 Jun;34(6):1429-1444.
15. Fellner L, Simon S, Scherling C, Witting M, Schober S, Polte C, *et al*. Evidence for the recent origin of a bacterial protein-coding, overlapping orphan gene by evolutionary overprinting. *BMC Evol Biol*. 2015 Dec 18;15: 283. doi: 10.1186/s12862-015-0558-z
16. Neme R, Tautz D. Phylogenetic patterns of emergence of new genes support a model of frequent *de novo* evolution. *BMC Genomics*. 2013; 14:117.
17. Guogin Y, Stoltzfus A. Population diversity of ORFan Genes in *Escherichia coli*. *Genome Biol Evol*. 2012; 4(11):1176-1187.
18. Kanagawa K, Negoro S, Takada N, Okada H. Plasmid dependence of *Pseudomonas* sp. Strain NK87 enzymes that degrade 6-aminohexanoate-cyclic dimer. *Journal of Bacteriology*. 1989 June;171(6):3181-3186.
19. Kanagawa K, Mayumi O, Negoro S, Urabe I, Okada H. Characterization of the 6-aminohexanoate-dimer hydrolase from *Pseudomonas* sp NK87. *Journal of General Microbiology*. 1993;139:787-795.
20. Kato K, Fujiyama K, Hatanaka HS, Prijambada ID, Negoro S, Urabe I, Ikada H. Amino acid alterations essential for increasing the catalytic activity of the nylon-oligomer-degradation enzyme of *Flavobacterium* sp. *Eur J. Biochem*. 1991 Aug 15;200(1):165-169.
21. Prijambada ID, Negoro S, Yomo T, Urabe I. Emergence of nylon oligomer degradation enzymes in *Pseudomonas aeruginosa* PAO through experimental evolution. *Applied and Environmental Microbiology*. 1995 May;61(5):2020-2022.
22. Sudhakar M, Priyadarshini C, Doble M, Murthy PS, Venkatesan R. Marine bacteria mediated degradation of nylon 66 and 6. *International Biodeterioration & Biodegradation*. 2007;60(3):144-151.

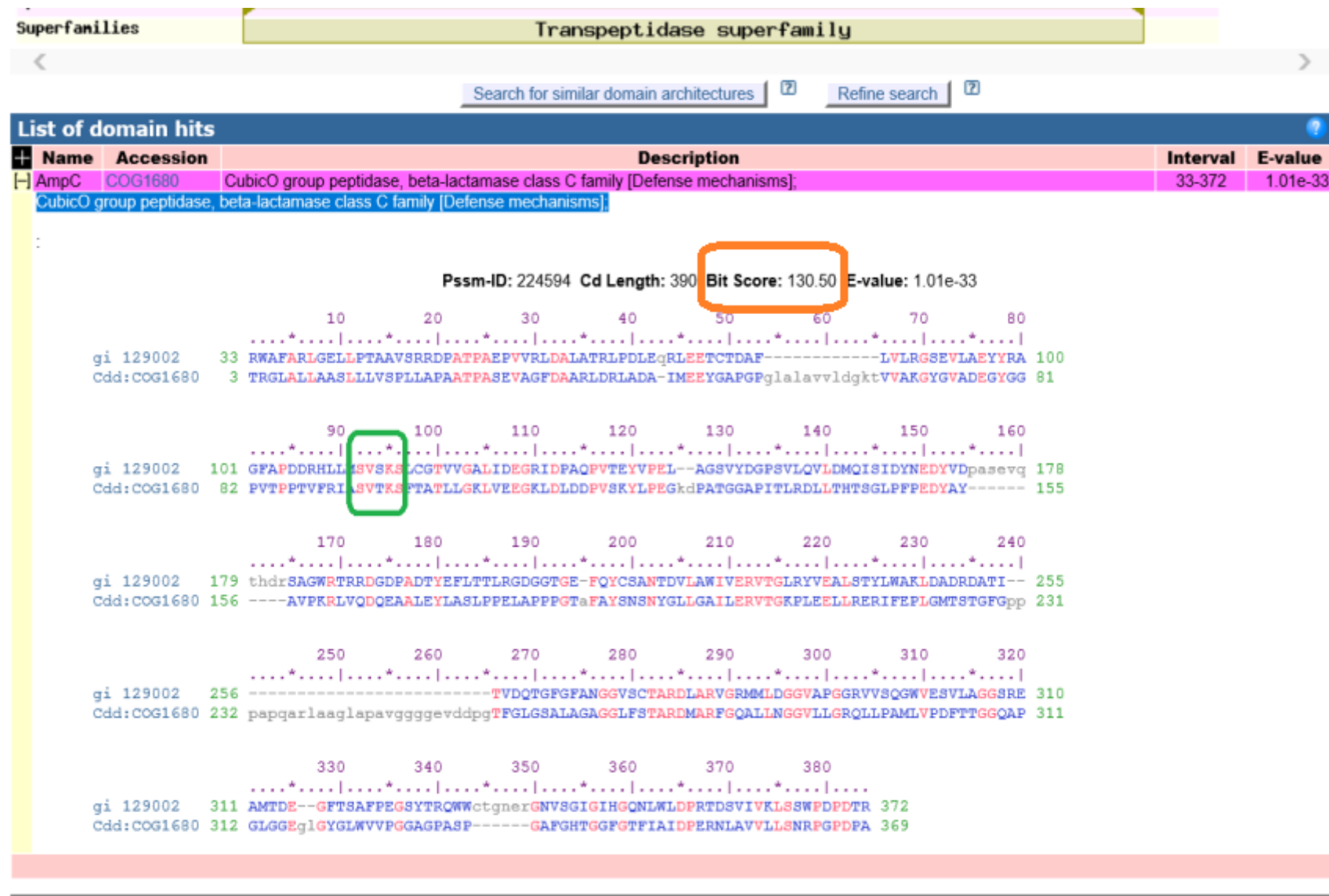
23. Negoro S, Ohki T, Shibata N, Mizuno N, Wakitani Y, Tsurukame J, *et al.* X-ray crystallographic analysis of 6-aminoohexanoate-dimer hydrolase. *The Journal of Biological Chemistry*. 2005 Nov 25;280(47):39644-39652.
24. Adeyemi SH, Ogunjobi AA, Fagade OE. Hydrolytic degradation of nylon-6 by *Pseudomonas aeruginosa* HE858284 isolated from solid waste dumpsites in Lagos State, Nigeria. *Journal of Natural Sciences Research*. 2015;5(2).
25. Hall BG, Barlow M. Evolution of the serine beta-lactamases; past, present and future. *Proc Natl. Acad. Sci. USA*. 1992 May, 89:3780-3784.
26. Rancurel C, Khosravi M, Dunker AK, Romero PR, Karlin D. Overlapping genes produce proteins with unusual sequence properties and offer insight into de novo protein creation. *J Virol*. 2009 Oct;83(20):10719-36. Doi: 10.1128/JVI.00595-09.
27. Sherr CJ. Divorcing ARF and p53: an unsettled case. *Nat Rev Cancer*. 2006 Sep;6(9):663-73.
28. Okamura K, Feuk L, Marques-Bonet T, Navarro A, Scherer SW. Frequent appearance of novel protein-coding sequences by frameshift translation. *Genomics*. 2006 December;88(6):600-697.
29. Kiumarsi A, Parvinzadeh M. Enzymatic hydrolysis of nylon 6 fiber using lipolytic enzyme. *The Journal of Biochemistry*. 1996;59(6).
30. Yasuhira K, Shibata N, Mongami G, Uedo Y, Atsumi Y, Kawashima Y, Hibino A, *et al.* X-ray Crystallographic analysis of the 6-aminoohexanoate cyclic dimer hydrolase. *Journal of Biological Chemistry*. 2010 Jan 8;285(2):1239-1248.

nylonase_manuscript_preprints_v2.pdf (246.02 KiB)

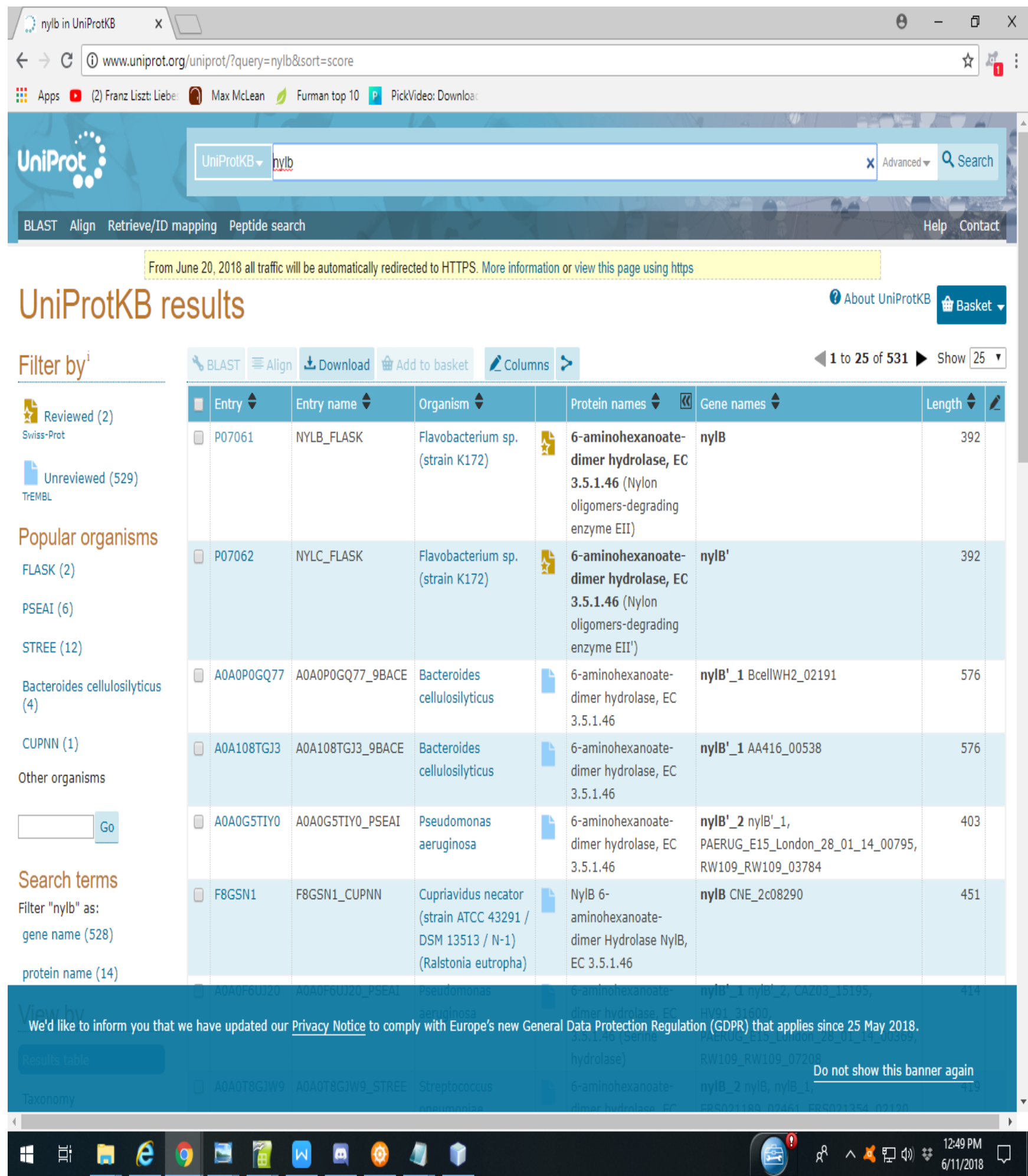
[view on ChemRxiv](#) • [download file](#)

A0A165T3A5	KZL05366.1	1 pfam00144	85.25	Pseudovibrio axinellae
G8PHS5	AEV36518.1	1 COG1680	133.20	Pseudovibrio sp. (strain FO-BEG1)
A0A1U6GQ15	SLJ85613.1	1 COG1680	134.35	Psychrobacter sp. DAB_AL43B
D8NJ41	CBJ43063.1	1 COG1680	180.58	Ralstonia solanacearum CFBP2957
D8NBY7	CBJ37975.1	1 COG1680	186.74	Ralstonia solanacearum CMR15
D8NTL3	CBJ51172.2	1 COG1680	178.65	Ralstonia solanacearum PS107
F6G7H0	AEG70906.1	1 COG1680	130.12	Ralstonia solanacearum (strain Po82)
G3A899	CCA86522.1	1 COG1680	180.96	Ralstonia solanacearum (strain Po82)
A0A1Y6GHY9	SMQ89731.1	1 COG1680	150.15	Ralstonia solanacearum (strain Po82)
W6S0Z5	CDM60151.1	1 COG1680	126.65	Ralstonia solanacearum (strain Po82)
A0A1W6KTY5	ARM90893.1	1 COG1680	150.53	Ralstonia solanacearum (strain Po82)
A0A1H8SL49	SEI12907.1	1 COG1680	131.66	Ralstonia solanacearum (strain Po82)
D5ANB4	ADE86404.1	1 COG1680	135.89	Ralstonia solanacearum (strain Po82)
X0Q0I2	GAF44337.1	1 COG1680	126.65	Ralstonia solanacearum (strain Po82)
A0A173SRD8	CUM92347.1	1 pfam00144	121.45	Ralstonia solanacearum (strain Po82)
A0A1X6ZH4	SLN50150.1	1 COG1680	131.66	Ralstonia solanacearum (strain Po82)
Q161I2	ABG33361.1	1 COG1680	146.29	Ralstonia solanacearum (strain Po82)
A0A1X6ZR95	SLN59152.1	1 pfam00144	106.05	Ralstonia solanacearum (strain Po82)
A0A0P1E3V5	CUH43141.1	1 COG1680	156.69	Ralstonia solanacearum (strain Po82)
A0A0P1I8K0	CUJ97681.1	1 COG1680	139.36	Ralstonia solanacearum (strain Po82)
A0A174R2F8	CUP79693.1	1 COG1680	147.06	Ralstonia solanacearum (strain Po82)
A0A0U6MUP1	CVA82934.1	1 COG1680	155.15	Ralstonia solanacearum (strain Po82)
A0A0P1FES5	CUH52959.1	1 COG1680	146.29	Ralstonia solanacearum (strain Po82)
A0A0M9EKA2	KPA23465.1	1 COG1680	132.04	Ralstonia solanacearum (strain Po82)
A0A173KVC3	ANI78048.1	1 COG1680	124.72	Ralstonia solanacearum (strain Po82)
A0A245ZH8	OWK29334.1	1 pfam00144	131.85	Ralstonia solanacearum (strain Po82)
A0A245ZS72	OWK32584.1	1 COG1680	121.64	Ralstonia solanacearum (strain Po82)
A0A220W924	ASK89323.1	1 COG1680	117.02	Ralstonia solanacearum (strain Po82)
A0A1V5WIK4	removed	0 n/a		Ralstonia solanacearum (strain Po82)
A0A0T8Y0L9	CKG36687.1	1 COG1680	150.92	Ralstonia solanacearum (strain Po82)
A0A117E7E7	GAQ53093.1	1 pfam00144	138.02	Ralstonia solanacearum (strain Po82)
A0A1D2IE66	ODA72800.1	1 pfam00144	89.10	Ralstonia solanacearum (strain Po82)
A0A1K2FPU2	SFY49811.1	1 COG1680	100.84	Ralstonia solanacearum (strain Po82)
A0A1B9EIG0	OCC07415.1	1 COG1680	124.72	Ralstonia solanacearum (strain Po82)
A0A1D3URL4	SCQ22896.1	1 COG1680	142.44	Ralstonia solanacearum (strain Po82)
A0A0P1FIN3	CUH67776.1	1 COG1680	135.12	Ralstonia solanacearum (strain Po82)
A0A0P1GCY7	CUH73549.1	1 COG1680	124.72	Ralstonia solanacearum (strain Po82)
I3TJC5	AFK52863.1	1 COG1680	156.31	Ralstonia solanacearum (strain Po82)
A0A1Y5S3B8	SLN31670.1	1 COG1680	145.14	Ralstonia solanacearum (strain Po82)
A0A1C6JX26	SCJ86692.1	1 COG1680	180.58	Ralstonia solanacearum (strain Po82)
A0A1C6G057	SCJ38643.1	1 COG1680	147.06	Ralstonia solanacearum (strain Po82)
A0A1C5PZB2	SCH42505.1	1 COG1680	122.80	Ralstonia solanacearum (strain Po82)
A0A1C5SG29	SCH72758.1	1 COG1680	115.48	Ralstonia solanacearum (strain Po82)
A0A1C6HH82	SCJ56877.1	1 COG1680	133.97	Ralstonia solanacearum (strain Po82)
A0A1C5YBZ6	SCI44810.1	1 COG1680	147.06	Ralstonia solanacearum (strain Po82)
A0A0H2LTX8	KLN53679.1	1 COG1680	118.17	Ralstonia solanacearum (strain Po82)
A0A1V6JU22	removed	0 n/a		Ralstonia solanacearum (strain Po82)
A0A1V6JCM4	removed	0 n/a		Ralstonia solanacearum (strain Po82)
U3AN41	GAD74717.1	1 COG1680	145.91	Ralstonia solanacearum (strain Po82)
A0A1C3JFB5	SBT13765.1	1 pfam00144	112.59	Ralstonia solanacearum (strain Po82)
A0A1B1NP12	ANS85294.1	1 COG1680	132.04	Ralstonia solanacearum (strain Po82)
A0A1Y2SKJ5	OTA18532.1	1 COG1680	85.05	Ralstonia solanacearum (strain Po82)

SCREEN SAMPLE OF CDD PAGE:



SCREEN SAMPLE OF UNIPROT PAGE FOR NYLB AND NYLB':



Testing the Hypothesis that the Nylonase NylB Protein
Arose *de novo* via a Frameshift Mutation

Salvador Cordova and John Sanford

Supplementary Table S2

List of organisms with NylB' from UNIPROT, and conserved domains of NylB' homologs with sample screen captures from NIH CDD and UNIPROT websites at the bottom

NOTES:

COG1680 = CubicO group peptidase, **beta-lactamase** class C family [Defense mechanisms];

cd10802 = Thermus thermophiles TTHB029 and similar proteins; This subfamily is represented by an YdjC-family protein TTHB029 from Thermus thermophilus HB8; it is similar to Escherichia coli YdjC, a hypothetical protein encoded by the celG gene. TTHB029 functions as a homodimer. Each of monomer consists of (beta/alpha)-barrel fold. The molecular function of TTHB029 is unclear.

cl21491 = Penicillin binding protein transpeptidase domain; for GenBank A0A011QU85 alignment was detected with superfamily member COG1680

pfam00144 = **Beta-lactamase**; This family appears to be distantly related to pfam00905 and PF00768 ...

pfam03724 = META domain; Small domain family found in proteins of unknown function. Some are secreted and implicated in motility in bacteria. Also occurs in Leishmania spp. as an essential gene. Over-expression in L.amazonensis increases virulence. A pair of cysteine residues show correlated conservation, suggesting that they form a disulphide bond.

n/a = not applicable because the data was embargoed pending publication

UNIPROT ID	Genbank ID	Conserved Domain Data Available	Primary Conserved Domain ACCESSION ID	Primary Conserved Domain Bit score	Secondary Conserved Domain ACCESSION ID	Secondary Conserved Domain Bit score	ORGANISM	Beta Lactamase Domain Bit Score
		no. avail. cdds	121					Average
		count nylb's	134					134.83
A0A0M7J5X6	CUI87454.1	1	COG1680	126.26			Achromobacter sp.	126.26
A0A1G5L686	SCZ08385.1	1	COG1680	184.04			Acinetobacter baumannii	184.04
A0A1R3U120	SCX34286.1	1	COG1680	132.43			Agrobacterium sp. DSM 25559	132.43
A0A1B3NSJ3	AOG08337.1	1	COG1680	156.69			Agrobacterium sp. RAC06	156.69
Q14T91	BAE97630.1	1	COG1680	130.12			Agromyces sp. KY5R	130.12
A0A0D6I854	CKH88456.1	1	COG1680	113.55			Alcaligenes xylosoxydans xylosoxydans (Achromo	113.55
A0A1B2ADK5	ANY20242.1	1	COG1680	142.44			Altererythrobacter dongtanensis	142.44
A0A0N0GL59	KPC49626.1	1	COG1680	158.23			Amantichitium ursilacus	158.23

A0A136WJN9	KXL54559.1	1 COG1680	131.66	Anaerotignum neopropionicum	131.66
A0A222E5F8	ASP21360.1	1 COG1680	137.43	Antarctobacter heliothermus	137.43
A0A1Y5T7Z2	SLN55856.1	1 COG1680	99.3	Aquimixticola soesokkakensis	99.3
A0A1Y6AVV0	PEV28669.1	1 COG1680	111.63	Bacillus cereus	111.63
A0A1V5JGZ1	removed	0 n/a		bacterium ADurb.BinA028	
A0A0A1PJ91	CEJ15450.1	1 COG1680	112.78	bacterium YEK0313	112.78
A0A180FM12	OAV73346.1	1 COG1680	138.59	Bacteroidales bacterium Barb6	138.59
A0A180FCF0	OAV69997.1	1 COG1680	136.28	Bacteroidales bacterium Barb6XT	136.28

secondary domain rather than primary domain had the beta lactamase

A0A0P0GQ77	ALJ59432.1	1 cd10802	265.32	COG1680	121.64	Bacteroides cellulosilyticus	121.64
A0A174JUS8	CUP00900.1	1 COG1680	157.46			Bacteroides xylanisolvens	157.46
A0A1V6CSN1	removed	0 n/a				Bacteroidetes bacterium ADurb.Bin123	
A0A1V6BCZ2	removed	0 n/a				Bacteroidetes bacterium ADurb.Bin145	
A0A1V5P201	removed	0 n/a				Bacteroidetes bacterium ADurb.Bin397	
A0A173XH31	CUN50989.1	1 COG1680	132.04			Bifidobacterium pseudocatenulatum	132.04
A0A221AGH6	ASL47615.1	1 COG1680	124.34			Burkholderia sp. AD24	124.34
A0A1V5GNM7	removed	0 n/a				candidate division BRC1 bacterium ADurb.BinA364	
A0A011QU85	EXI92700.1	1 cl21491	55.39			Candidatus Accumulibacter sp. BA-94	55.39
A0A084Y3V0	KFB69394.1	1 COG1680	79.65			Candidatus Accumulibacter sp. SK-01	79.65
A0A1Q9NG96	OLS21127.1	1 COG1680	170.95			Candidatus Thorarchaeota archaeon AB_25	170.95
M7NIZ7	EMR01730.1	1 COG1680	150.92			Cesiribacter andamanensis AMV16	150.92
A0A1X6W741	SHE13305.1	1 COG1680	157.85			Chlamydia abortus (Chlamydophila abortus)	157.85
A0A1V5QH62	removed	0 n/a				Chloroflexi bacterium ADurb.Bin344	
A0A1V5PDI7	removed	0 n/a				Chloroflexi bacterium ADurb.Bin360	
A0A1A7KJK9	OBW41946.1	1 pfam00144	121.45			Chryseobacterium sp. MOF25P	121.45
A0A1Y3FFF3	OUE17692.1	1 COG1680	96.99			Clavibacter michiganensis subsp. michiganensis	96.99
A0A143Y291	CVI73461.1	1 pfam00144	128.39			Clostridiales bacterium CHKCI001	128.39
A0A1R4UEF4	SJT07095.1	1 COG1680	145.14			Clostridioides difficile (Peptoclostridium difficile)	145.14
A0A174SE58	CUP94651.1	1 COG1680	140.13			[Clostridium] clostridioforme	140.13
A0A168MLP5	OAA84870.1	1 COG1680	103.15			Clostridium ljungdahlii	103.15
A0A1V4IZP8	OPJ65244.1	1 COG1680	120.1			Clostridium oryzae	120.1
A0A1S8TMH4	OOM78792.1	1 COG1680	97.76			Clostridium puniceum	97.76
A0A0U5PQ27	CUX75402.1	1 COG1680	139.75			Clostridium sp. C105KSO14	139.75
A0A0F0CJV5	KJJ74223.1	1 COG1680	141.29			Clostridium sp. FS41	141.29
A0A132HH88	AVA35740.1	1 COG1680	106.62			Cupriavidus metallidurans	106.62
A0A239WW24	SNV37894.1	1 pfam00144	141.48			Cutibacterium granulosum	141.48
A0A1V5J955	removed	0 n/a				Deltaproteobacteria bacterium ADurb.BinA014	
A0A174A4R5	CUN83297.1	1 COG1680	157.46			Dorea longicatena	157.46
A0A1E7WFA0	OEZ97024.1	1 COG1680	140.13			Duganella sp. HH101	140.13
A0A1E7V5C4	OEZ55850.1	1 COG1680	139.36			Duganella sp. HH105	139.36
A0A0D5BQQ6	AJW63928.1	1 COG1680	134.74			Elizabethkingia miricola (Chryseobacterium miricol	134.74

A0A157WDE4	SAJ28520.1	1 COG1680	162.47	Enterobacter cloacae	162.47
A0A1V6AIT8	removed	0 n/a		Firmicutes bacterium ADurb.Bin146	
Q59710	BAA05089.1	1 COG1680	134.35	Flavobacterium sp. KI723T1	134.35
P07062	CAA26616.1	1 COG1680	135.12	Flavobacterium sp. (strain K172)	135.12
A0A174KSI3	CUP12139.1	1 COG1680	125.49	Fusicatenibacter saccharivorans	125.49
W2UE25	ETN92390.1	1 COG1680	161.32	Gammaproteobacteria bacterium MOLA455	161.32
A0A128EWW5	CZF78501.1	1 COG1680	130.89	Grimontia celer	130.89
A0A174B3K2	CUN94340.1	1 COG1680	154.77	Hungatella hathewayi	154.77
A0A0D1EBH5	KIT15094.1	1 COG1680	114.71	Jannaschia aquimarina	114.71
A0A0M6YFF1	CTQ48680.1	1 COG1680	144.37	Jannaschia donghaensis	144.37
A0A0M6XQ32	CTQ32271.1	1 COG1680	149.76	Jannaschia rubra	149.76
A0A1E7W4C8	OEZ90584.1	1 COG1680	130.5	Janthinobacterium sp. HH106	130.5
A0A0M2WN34	KKO63150.1	1 COG1680	129.73	Janthinobacterium sp. KBS0711	129.73
A0A181XAP6	SBL54634.1	1 COG1680	152.84	Klebsiella oxytoca	152.84
A0A0M6Y0X3	CTQ43334.1	1 COG1680	143.6	Labrenzia aggregata	143.6
A0A174LCT8	CUP21913.1	1 COG1680	118.56	Lachnospira pectinoschiza	118.56
A0A1D8B005	AOS46469.1	1 COG1680	122.8	Lacunisphaera limnophila	122.8
A0A0W0UST3	KTD10888.1	1 COG1680	129.73	Legionella gratiana	129.73
A0A1V5UW73	removed	0 COG1680		Lentisphaerae bacterium ADurb.Bin242	
A0A0P0M7I9	ALK87917.1	1 COG1680	129.73	Limnohabitans sp. 63ED37-2	129.73
A0A143PQU0	AMY10736.1	1 COG1680	77.54	Luteitalea pratensis	77.54
A0A1C3JR26	SBT17678.1	1 COG1680	186.35	Marinomonas gallaica	186.35
A0A161SK65	KZE91550.1	1 COG1680	127.42	Microbacterium sp. TNHR37B	127.42
A0A1C9W5R1	AOS96494.1	1 COG1680	133.97	Microbulbifer aggregans	133.97
A0A1J5Q3Y3	OIQ78417.1	1 COG1680	116.25	mine drainage metagenome	116.25
A0A1N4H1G0	SII01201.1	1 COG1680	121.64	Mycobacterium abscessus subsp. abscessus	121.64
D5PH26	EFG74633.1	1 COG1680	137.82	Mycobacterium parascrofulaceum ATCC BAA-614	137.82
A0A200GVL6	OUZ04257.1	1 COG1680	136.28	Mycobacterium paratuberculosis	136.28
A0A0H5D2L7	CRL10978.1	1 COG1680	141.67	Nautella italica	141.67
A0A0U1NN44	CRK75919.1	1 COG1680	147.06	Nereida ignava	147.06
A0A1A9GPB2	ANH39275.1	1 COG1680	125.88	Nocardioides dokdonensis FR1436	125.88
A0A1Y5T0T5	SLN49476.1	1 COG1680	121.64	Oceanibacterium hippocampi	121.64
A0A238KNR1	SMX44435.1	1 COG1680	135.12	Octadecabacter ascidiaceicola	135.12
A0A0K0Y7L4	AKS46964.1	1 COG1680	136.28	Octadecabacter temperatus	136.28
A0A0B0I7C6	KHF36079.1	1 COG1680	109.31	Paenibacillus sp. P1XP2	109.31
A0A1Y5SE37	SLN37902.1	1 COG1680	136.28	Pelagicola litorisediminis	136.28
A0A238L5C6	SMX50178.1	0 query exception		Pelagimonas varians	
A0A0P1INL0	CUJ87550.1	1 COG1680	148.99	Phaeobacter sp. CECT 7735	148.99
A0A0G5TIY0	CRP67525.1	1 COG1680	135.89	Pseudomonas aeruginosa	135.89
A0A0D0TTV4	KIR24245.1	1 COG1680	128.96	Pseudomonas fluorescens	128.96
A0A1B2F7B1	ANY88065.1	1 COG1680	139.36	Pseudomonas putida (Arthrobacter siderocapsulat	139.36
A0A1C3GPS1	SAM32366.1	1 COG1680	133.97	Pseudomonas sp. 1 R 17	133.97
A0A1B5FEP0	CRM97274.1	1 COG1680	134.35	Pseudomonas sp. 22 E 5	134.35
A0A1B5DAM1	CRM30808.1	1 COG1680	137.43	Pseudomonas sp. 24 E 1	137.43
A0A1B5DJF5	CRM42528.1	1 pfam00144	109.9	Pseudomonas sp. 44 R 15	109.9
A0A1B5EAS6	CRM75146.1	1 pfam00144	105.28	Pseudomonas sp. 58 R 3	105.28
A0A0S4HRC6	CRL48425.1	1 COG1680	147.06	Pseudomonas sp. URMO17WK12:111	147.06

A0A099SNY9	KUR43944.1	1 COG1680	140.52	Pseudomonas syringae pv. tomato	140.52
A0A1Y5SNP5	SLN44728.1	1 COG1680	133.58	Pseudooctadecabacter jejudonensis	133.58
A0A1Y5RVI3	SLN23768.1	1 COG1680	158.62	Pseudoruegeria aquimaris	158.62
A0A165ZGJ7	KZL19875.1	1 COG1680	135.51	Pseudovibrio axinellae	135.51
A0A165V164	KZL13197.1	1 COG1680	136.28	Pseudovibrio sp. Ad26	136.28
A0A165Q4T3	KZK98727.1	1 COG1680	136.66	Pseudovibrio sp. W74	136.66
A0A161XKM1	KZL25998.1	1 COG1680	134.74	Pseudovibrio sp. WM33	134.74
A0A1X7BXA6	SMC14336.1	1 COG1680	145.91	Roseovarius aestuarii	145.91
A0A1X6ZMI7	SLN53765.1	1 COG1680	148.22	Roseovarius gaetbuli	148.22
A0A1X6YYU5	SLN33443.1	1 COG1680	140.9	Roseovarius halotolerans	140.9
A0A1V0RJX9	ARE82093.1	1 COG1680	137.05	Roseovarius mucosus	137.05
A0A1V1UE16	GAW33969.1	1 COG1680	150.92	Roseovarius sp. A-2	150.92
A0A0L6CVV1	KNX41805.1	1 COG1680	154.77	Roseovarius tolerans	154.77
A0A238KXG2	SMX47524.1	1 COG1680	157.46	Ruegeria arenilitoris	157.46
A0A1X7A821	SLN72802.1	1 COG1680	166.32	Ruegeria meonggei	166.32
A0A0P1IH18	CUK12632.1	1 COG1680	147.45	Ruegeria sp. CECT 5091	147.45
A0A0S2I4P7	ALO17296.1	1 COG1680	156.31	Salinivirga cyanobacteriivorans	156.31
A0A0U6YWJ4	CVF13851.1	1 COG1680	152.84	Serratia marcescens	152.84
A0A0M9EEL0	KPA20195.1	1 COG1680	128.96	Shimia sp. SK013	128.96
A0A220W7L8	ASK88821.1	1 COG1680	142.83	Sphingorhabdus flavimaris	142.83
A0A1V5HU64	removed	0 n/a		Spirochaetes bacterium ADurb.BinA120	
A0A0T8JK43	CJG33358.1	1 COG1680	112.01	Streptococcus pneumoniae	112.01
A0A221K285	ASM73079.1	1 COG1680	145.14	Sulfitobacter pseudonitzschiae	145.14
A0A109ZY28	AMH39403.1	1 COG1680	165.55	Tardiphaga robiniae	165.55
A0A0P1FIH6	CUH67767.1	1 COG1680	149.76	Thalassobius gelatinovorans (Ruegeria gelatinovorans)	149.76
A0A0P1FSF4	CUH71927.1	1 COG1680	133.97	Thalassobius mediterraneus	133.97
A0A1Y5RU04	SLN25029.1	1 COG1680	156.69	Tropicibacter litoreus R37	156.69
A0A0N7M0Q4	CUH81113.1	1 COG1680	117.79	Tropicibacter multivorans	117.79
A0A238JFC7	SMX29115.1	1 COG1680	130.12	Tropicibacter phthalicus	130.12
A0A1C6HKF7	SCJ57940.1	1 COG1680	154.77	uncultured Bacteroides sp.	154.77
A0A1C5VKN5	SCI11242.1	1 COG1680	148.22	uncultured Blautia sp.	148.22
A0A1C5WZV3	SCI28406.1	1 pfam00144	119.14	uncultured Clostridium sp.	119.14
A0A1C5Q3J1	SCH43822.1	1 COG1680	154.77	uncultured Dorea sp.	154.77
A0A1C6IY35	SCJ74355.1	1 COG1680	123.95	uncultured Lachnospira sp.	123.95
A0A1C5PT61	SCH40294.1	1 COG1680	133.97	uncultured Ruminococcus sp.	133.97
A0A0S3PSD0	BAT58843.1	1 COG1680	159.01	Variibacter gotjawalensis	159.01
A0A1V6GRF0	removed	0 COG1680		Verrucomicrobia bacterium ADurb.Bin070	
A0A1W6VLV6	ARP40995.1	1 COG1680	103.54	Vibrio alginolyticus	103.54
A0A1C3IG13	PMF19269.1	1 COG1680	102	Vibrio splendidus	102

SAMPLE NYLB' Conserved Domains Database Page:

The screenshot displays the NCBI Conserved Domains Database interface. It features a protein sequence with three conserved domains highlighted: SH3 (green), SH2 (blue), and SH1 (red). The NCBI logo is visible on the left side of the interface.

Conserved domains on [gi|1326315889|gb|PMF19269.1|]

View

6-aminohexanoate hydrolase [Vibrio splendidus]

Protein Classification

class C beta-lactamase-related serine hydrolase (domain architecture ID 10004149)
 class C beta-lactamase-related serine hydrolase similar to Flavobacterium 6-aminohexanoate-dimer hydrolase, which degrades linear oligomers of 6-aminohexanoate by exo-type cleavage, removing residues sequentially from the N-terminus of nylon-6 oligomers

Graphical summary

Zoom to residue level



List of domain hits

Name	Accession	Description	Interval	E-value
AmpC	COG1680	CubicO group peptidase, beta-lactamase class C family [Defense mechanisms]; CubicO group peptidase, beta-lactamase class C family [Defense mechanisms];	63-414	6.98e-24

Pssm-ID: 224594 Cd Length: 390 Bit Score: 102.00 E-value: 6.98e-24

```

      10      20      30      40      50      60      70      80
gi 1326315889 63 LSRFYPTAQIERDGPVYQFPYQIDQNIIGE-ISA TVHGESKTLNEHLDsypvdaFLVVKSGEVVFERYNT-MGKTDKHNMF 140
Cdd:COG1680  19 LLAPAATPASEVAGFDAARLDRDLADAIMEeYGA PGPLALAWLDGK-----TVVAKGYGVADEGYGgpVTPPTVFRIA 92

      90     100     110     120     130     140     150     160
gi 1326315889 141 SNSKITVGIE LAKLVNEGQVNLSDPISKYIPELKGSDWD--TVSVSDTANMATGLNATEhDEPEADsrinPEQPWFKQAV 218
Cdd:COG1680  93 SVTKSFTATLLGKLVEEGKLDLDDPVSKYLPFGKDPATGgaPITLRDLLTHTSGLPFPE-DYAYAV---PKRLVQDQEA 167

     170     180     190     200     210     220     230     240
gi 1326315889 219 SigvfeggnqspLEVLAEMKRRAPG--GKTFEYNSINTFV LARLIENVRNLP MNEIVSRDMWQKMGANNDAYTV----- 291
Cdd:COG1680  168 A-----LEYLASLPPELAPppGTAFAYSNSNYGLLGAILERVTKPLLELLRERIFEPLGMTSTGFpPapq 235

     250     260     270     280     290     300     310     320
gi 1326315889 292 -----VSPQGGYPLMFFS-----MNTTIEDMTKFGMLLtpsatkLGNGGVSKEVI---TLIQDSGNP-- 345
Cdd:COG1680  236 arlaaglapavggggeVDDPGTFGLGSALagaggLFSTARDMARFGQAL-----LNGGVLLGRQLlpaMLVPDFTTgq 309

     330     340     350     360     370     380     390
.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....
    
```

SCREEN SAMPLE OF UNIPROT PAGE FOR NYLB AND NYLB':



UniProtKB hylb

Advanced Search

BLAST Align Retrieve/ID mapping Peptide search

Help Contact

From June 20, 2018 all traffic will be automatically redirected to HTTPS. More information or view this page using https

UniProtKB results

About UniProtKB Basket

Filter byⁱ

Reviewed (2)
Swiss-Prot

Unreviewed (529)
TrEMBL

Popular organisms

FLASK (2)

PSEAI (6)

STREE (12)

Bacteroides cellulosilyticus (4)

CUPNN (1)

Other organisms

Go

Search terms

Filter "nylb" as:
gene name (528)
protein name (14)

BLAST Align Download Add to basket Columns 1 to 25 of 531 Show 25

Entry	Entry name	Organism	Protein names	Gene names	Length
P07061	NYLB_FLASK	Flavobacterium sp. (strain K172)	6-aminohexanoate-dimer hydrolase, EC 3.5.1.46 (Nylon oligomers-degrading enzyme EII)	nylb	392
P07062	NYLC_FLASK	Flavobacterium sp. (strain K172)	6-aminohexanoate-dimer hydrolase, EC 3.5.1.46 (Nylon oligomers-degrading enzyme EII')	nylB'	392
A0A0P0GQ77	A0A0P0GQ77_9BACE	Bacteroides cellulosilyticus	6-aminohexanoate-dimer hydrolase, EC 3.5.1.46	nylB'_1 BcellWH2_02191	576
A0A108TGJ3	A0A108TGJ3_9BACE	Bacteroides cellulosilyticus	6-aminohexanoate-dimer hydrolase, EC 3.5.1.46	nylB'_1 AA416_00538	576
A0A0G5TIY0	A0A0G5TIY0_PSEAI	Pseudomonas aeruginosa	6-aminohexanoate-dimer hydrolase, EC 3.5.1.46	nylB'_2 nylB'_1, PAERUG_E15_London_28_01_14_00795, RW109_RW109_03784	403
F8GSN1	F8GSN1_CUPNN	Cupriavidus necator (strain ATCC 43291 / DSM 13513 / N-1) (Ralstonia eutropha)	NylB 6-aminohexanoate-dimer Hydrolase NylB, EC 3.5.1.46	nylB CNE_2c08290	451
A0A0F0UJ20	A0A0F0UJ20_PSEAI	Pseudomonas aeruginosa	6-aminohexanoate-dimer hydrolase, EC 3.5.1.46 (Serine hydrolase)	nylB'_1 nylB'_2, CA203_15195, HV91_31690, PAERUG_E15_London_28_01_14_00795, RW109_RW109_07208	414
A0A0T8GJW9	A0A0T8GJW9_STREE	Streptococcus pneumoniae	6-aminohexanoate-dimer hydrolase, EC 3.5.1.46	nylB_2 nylB, nylB_1, PNC021140_02441, PNC021140_03190	419

View by

We'd like to inform you that we have updated our [Privacy Notice](#) to comply with Europe's new General Data Protection Regulation (GDPR) that applies since 25 May 2018.

Results table

Taxonomy

Do not show this banner again

Testing the Hypothesis that the Nylonase NylB Protein
Arose *de novo* via a Frameshift Mutation

Salvador Cordova and John Sanford

Supplementary Table S3

List of organisms with NylA from uniprot, and conserved domains of NylAs
with sample screen captures from NIH CDD and UNIPROT websites at the bottom

NOTES:

COG0154 = Asp-tRNAAsn/Glu-tRNA^{Gln} amidotransferase A subunit or related amidase
[Translation, ribosomal structure and biogenesis];

cl18951 = Amidase;

pfam01425 = Amidase;

PRK07869 = amidase; Provisional

PRK06061 = amidase; Provisional

pfam07501 = G5 domain; This domain is found in a wide range of extracellular proteins. It is found ...

PRK06529 = amidase; Provisional

n/a = not applicable because the data was embargoed pending publication

no. avail. cdds 182
count NylAs 193

UNIPROT ID	Genbank ID	Conserved Domain Data Available	Primary Conserved Domain ACCESSION ID	Primary Conserved Domain Bit score	Secondary Conserved Domain ACCESSION ID	Secondary Conserved Domain Bit score	ORGANISM
D4XAB8	EFF76383.1	1	COG0154	257.22			Achromobacter piechaudii ATCC 43553
A0A0M7DUQ9	CUI57268.1	1	COG0154	242.2			Achromobacter sp.
A0A066PM46	KDM66920.1	1	COG0154	327.33			Acidiphilium sp. JA12-A1
A0A0D8HH63	KJF16396.1	1	COG0154	303.06			Acidithrix ferrooxidans
A0A239RVT3	SNU14988.1	1	cl18951	361.83			Acinetobacter johnsonii
A0A150I2M3	KXZ73999.1	1	cl18951	364.14			Acinetobacter venetianus
A0A2P4URV2	POM27744.1	1	COG0154	330.41			Actinomadura sp. 5-2
D4YI68	EFG49104.1	1	cl18951	544.03			Aerococcus viridans ATCC 11563 = CCUG 4311
A0A1R3U232	SCX35430.1	1	COG0154	230.26			Agrobacterium sp. DSM 25559
A0A0D6INF8	CKI16847.1	1	COG0154	252.6			Alcaligenes xylosoxydans xylosoxydans (Achromobacter xylosoxidans)

A0A1B2ABP4	ANY19580.1	1 COG0154	266.08
A0A178X0F6	OAP23195.1	1 cl18951	219.86
A0A0D6QPK1	GAO05363.1	1 COG0154	299.98
A0A0U5BA19	BAU27772.1	1 cl18951	278.41
A0A1D7QVD8	AOM82949.1	1 cl18951	377.24
A0A1Q9BAH0	OLP65740.1	1 cl18951	501.66
L8AQW9	BAM53880.1	1 pfam01425	389.35
A0A090IS94	CEE00941.1	1 cl18951	529.78
A0A1V5J9J8	removed	0 n/a	
A0A2H6H4J4	GBE21173.1	1 cl18951	317.7
A0A2H5ZNF0	GBD27233.1	1 COG0154	344.28
A0A2H6AWV7	GBD42414.1	1 COG0154	309.99
A0A2H6AZD7	GBD43293.1	1 cl18951	314.23
A0A2H6B6G8	GBD45731.1	1 COG0154	350.06
A0A0A1PD21	CEJ11088.1	1 cl18951	307.3
H6RLD2	CCG02458.1	1 COG0154	311.15
A0A157S5C9	SAI65612.1	1 COG0154	276.48
A0A0H4WRC0	AKQ57708.1	1 COG0154	278.02
A0A140KUK9	CZQ35579.1	1 cl18951	98.13
A0A157SG71	SAI69344.1	1 COG0154	281.88
A0A0H2WG49	AAU48320.1	1 cl18951	321.16
A2S6W5	ABN01555.1	1 cl18951	321.16
A0A1S0SB87	CFL38648.1	1 cl18951	320.78
Q3JY82	ABA50489.1	1 cl18951	321.55
A0A221AB94	ASL45854.1	1 COG0154	257.22
V9TR00	AHC73309.1	1 COG0154	313.46
K8EDK6	CCO09833.2	1 cl18951	411.91
A0A0E9FL18	CRH82220.1	1 cl18951	136.49
F8GPZ8	AEI79351.1	1 cl18951	265.69
A0A239WBU5	SNV31669.1	1 COG0154	271.09
A0A1V6KGF2	removed	0 n/a	
A0A1V6IQ42	removed	0 n/a	
A0A1V6G4I0	removed	0 n/a	
A0A1V6AWT2	removed	0 n/a	
A0A1V5HE73	removed	0 n/a	
A0A239V818	SNV18335.1	1 cl18951	368.06
F5SJP9	EGK08499.1	1 cl18951	489.72
C0X1J8	EEl13129.1	1 cl18951	527.47
A0A0D8FVF4	KJE77106.1	1 cl18951	314.84
P13398	BAA05090.1	1 COG0154	288.42
A0A222TFI4	ASR01485.1	1 PRK07869	683.64
C8NG30	EEW37516.1	1 cl18951	691.18
A0A059NVE7	CDQ23279.1	1 COG0154	316.93
A0A2C9D0T4	SON53778.1	0 query exception	
A0A1E7WAM3	OEZ93467.1	1 COG0154	292.28
A0A0M2WHD7	KKO62249.1	1 COG0154	283.8
A0A1S6IQB3	AQS53746.1	1 cl18951	591.41

Altererythrobacter dongtanensis
Amycolatopsis sp. M39
Anaeromyxobacter sp. PSR-1
Aneurinibacillus soli
Bacillus beveridgei
Bacillus pumilus (Bacillus mesentericus)
Bacillus subtilis BEST7613
Bacillus thermoamylovorans
bacterium ADurb.BinA028
bacterium BMS3Bbin01
bacterium HR30
bacterium HR39
bacterium HR40
bacterium HR41
bacterium YEK0313
Blastococcus saxobsidens (strain DD2)
Bordetella ansorpii
Bordetella hinzii
Bordetella pertussis
Bordetella trematum
Burkholderia mallei (strain ATCC 23344)
Burkholderia mallei (strain NCTC 10229)
Burkholderia pseudomallei (Pseudomonas pseudomallei)
Burkholderia pseudomallei (strain 1710b)
Burkholderia sp. AD24
Candidatus Endolissoclinum faulkneri L5
Carnobacterium maltaromaticum LMA28
Chlamydia trachomatis
Cupriavidus necator (strain ATCC 43291 / DSM 13513 / N-1) (Ralstonia eutropha)
Cutibacterium granulorum
Deltaproteobacteria bacterium ADurb.Bin002
Deltaproteobacteria bacterium ADurb.Bin022
Deltaproteobacteria bacterium ADurb.Bin072
Deltaproteobacteria bacterium ADurb.Bin151
Deltaproteobacteria bacterium ADurb.BinA179
Dermatophilus congolensis
Desmospora sp. 8437
Enterococcus faecalis TX0104
Ferrimicrobium acidiphilum DSM 19497
Flavobacterium sp. (strain K172)
Gordonia rubripertincta (Rhodococcus corallinus)
Granulicatella adiacens ATCC 49175
Halobacillus karajensis
Hartmannibacter diazotrophicus
Janthinobacterium sp. HH106
Janthinobacterium sp. KBS0711
Jeotgalibaca dankookensis

A0A181C565	SAY46675.1	1 COG0154	311.92
A0A0M6Z8M8	CTQ64276.1	1 cl18951	297.67
C2D288	EEI71107.1	1 cl18951	494.72
C0WQT6	EEI19652.1	1 cl18951	494.72
J9W3Z9	AFS01208.1	1 cl18951	482.78
A0A2N9KZL5	SPE16009.1	1 cl18951	517.45
C0WXY8	EEI22090.1	1 cl18951	513.6
C0XGF7	EEI25501.1	1 cl18951	498.19
C8PAP3	EEW52532.1	1 cl18951	410.75
C2FBN4	EEI68757.1	1 cl18951	553.27
E7FSC6	EFZ34169.1	1 cl18951	546.73
A0A1E7X8H4	OFA09389.1	1 cl18951	489.72
C2ESU0	EEJ41066.1	1 cl18951	521.3
A0A173M330	BAV01662.1	1 cl18951	415.37
A0A078KW70	CDZ75999.1	1 COG0154	298.05
A0A128ZD65	CZI36732.1	1 COG0154	244.51
A0A110AWJ2	BAU43134.1	1 COG0154	341.97
Q72PS5	AAS70961.1	1 cl18951	398.04
A0A2P2D2K7	GBF38899.1	1 cl18951	368.38
A0A0N7J8F9	ALK88080.1	1 COG0154	291.51
A0A0Y7J7Z3	CWW90658.1	1 cl18951	527.85
A0A238LBI3	SMY06326.1	1 COG0154	299.59
A0A238K804	SMX38617.1	1 COG0154	298.44
E4PFN8	ADP96787.1	1 PRK07869	677.47
H8WBP3	CCG93690.1	1 COG0154	368.16
A0A1W6K3W3	ARM82123.1	1 cl18951	353.36
A0A1C3JP51	A0A1C3JP51	1 COG0154	316.93
A0A1A8T5V6	SBS27780.1	1 COG0154	296.9
A0A068DFK2	AID33410.1	1 COG0154	308.45
A0A090FKT7	CDX44609.1	1 COG0154	321.55
A0A090E4Q5	CDX21843.1	1 COG0154	318.47
A0A090ENA9	CDX32438.1	1 COG0154	318.47
A0A1E3H3I1	ODN70879.1	1 COG0154	179.03
A0A0F0LDS8	KJL31283.1	1 cl18951	255.29
A0A2P8AY42	PSK65391.1	1 COG0154	318.08
A0A1J5RNJ1	OIQ97289.1	1 cl18951	295.96
A0A0U0ZTE6	CPV65311.1	1 COG0154	334.65
A0A1N0DAQ6	SHV30349.1	1 COG0154	330.41
A0A1N6AFE1	SIN32756.1	1 COG0154	332.72
A0A1T7W6V6	SKM64002.1	1 COG0154	332.72
A0A0J6W9F6	KMO78503.1	1 PRK07869	789.57
A0A0J6W7L7	KMO77627.1	1 PRK06061	769.25
A0A1V3XQ88	OOK80926.1	1 cl18951	46.61
X7Y501	EUA01906.1	1 cl18951	46.61
A0A0J6Y9F2	KMO69576.1	1 COG0154	318.85
D5P7P7	EFG77891.1	1 PRK06061	789.28

Komagataeibacter rhaeticus
Labrenzia alba
Lactobacillus brevis subsp. *gravesensis* ATCC 27305
Lactobacillus buchneri ATCC 11577
Lactobacillus buchneri CD034
Lactobacillus fermentum
Lactobacillus fermentum ATCC 14931
Lactobacillus hilgardii DSM 20176 = ATCC 8290
Lactobacillus iners DSM 13335
Lactobacillus paracasei subsp. *paracasei* ATCC 25302
Lactobacillus ruminis ATCC 25644
Lactobacillus sunkii
Lactobacillus vaginalis DSM 5837 = ATCC 49540
Lactococcus garvieae
Legionella massiliensis
Legionella pneumophila
Leptolyngbya sp. O-77
Leptospira interrogans serogroup *Icterohaemorrhagiae*
serovar *copenhageni* (strain *Fiocruz L1-130*)
Leptospira sp. E8
Limnohabitans sp. 63ED37-2
Listeria monocytogenes
Loktanella marina
Maliponia aquimaris
Marinobacter adhaerens (strain DSM 23420 / HP15)
Marinobacter hydrocarbonoclasticus ATCC 49840
Marinobacter salarius
Marinomonas gallaica
Marinomonas spartinae
Mesorhizobium huakuii 7653R
Mesorhizobium plurifarium
Mesorhizobium sp. ORS3359
Mesorhizobium sp. SOD10
Methylobrevia pamukkalensis
Microbacterium azadirachtae
Micromonospora sp. MH33
mine drainage metagenome
Mycobacterium abscessus
Mycobacterium abscessus subsp. *abscessus*
Mycobacterium abscessus subsp. *bolletii*
Mycobacterium abscessus subsp. *massiliense*
Mycobacterium chlorophenolicum
Mycobacterium chubuense
Mycobacterium kansasii
Mycobacterium kansasii 824
Mycobacterium obuense
Mycobacterium parascrofulaceum ATCC BAA-614

A0A202FU52	OUZ03240.1	1 COG0154	327.33
A0A0D6IBQ3	CKH95081.1	1 COG0154	251.44
L8F948	ELQ87854.1	1 PRK06061	786.2
A0A240AD36	SNV81209.1	1 cl18951	321.16
A0A0T9S9J6	CNJ51382.1	1 cl18951	79.26
X8FKH9	EUA93672.1	1 cl18951	102.27
X8BHR1	EUA43006.1	0 no entry	
Q1CYM7	ABF91800.1	1 COG0154	346.97
F9EZJ5	EGQ75050.1	1 COG0154	299.98
G4CK47	EGY51774.1	1 COG0154	302.29
A0A2P6L7Y9	PRD34697.1	1 cl18951	357.98
W9V8H0	EXJ12342.1	1 cl18951	115.47
A0A231GW69	OXR40869.1	1 COG0154	284.19
A0A0H5NGA5	CRY74523.1	1 PRK07869	682.1
A0A1A9GJH0	ANH38404.1	1 COG0154	263.77
A0A1Y5TGA2	SLN63409.1	1 COG0154	326.94
A0A0A1MHT8	CEI82653.1	1 cl18951	464.29
W9AFE3	CDO04193.1	1 cl18951	470.46
R4YTJ3	CCK77043.1	1 COG0154	366.23
A0A0E3WGY6	CQR54353.1	1 cl18951	490.49
A0A2M9NE97	PJN62497.1	1 cl18951	514.75
A0A239SAD1	SNU82172.1	1 COG0154	275.33
A0A0U1KJ51	CQR46008.1	1 cl18951	479.7
A0A1X7AFR7	SMA36991.1	1 COG0154	364.69
A0A238KTK6	SMX46135.1	0 query exception	
A0A0P1GXR8	CUH87486.1	1 cl18951	114.7
A0A073CF91	KEI66786.1	1 COG0154	341.97
K8WC88	EKT58174.1	1 COG0154	299.98
A0A024H881	CCQ47959.1	1 COG0154	273.02
A0A127MTZ1	AMO76736.1	1 COG0154	291.89
A0A1Q9QYC4	OLS60161.1	1 cl18951	287.65
A0A1C3GU74	SAM33941.1	1 cl18951	319.62
A0A1B5ENX2	CRM88048.1	1 cl18951	308.45
A0A2H2YM22	CRM80544.1	1 cl18951	308.45
A0A1B5DRR8	CRM50573.1	1 cl18951	318.85
P13397	AAA25908.1	1 COG0154	285.73
A0A105T8Z6	KVV04240.1	1 cl18951	318.08
A0A099SXA7	KUR41590.1	1 COG0154	305.76
A0A1Y2NAK7	OSY43938.1	1 cl18951	612.76
A0A165TLS8	KZL06191.1	1 cl18951	211.99
A0A1R4EGM1	SJM37569.1	1 COG0154	298.05
A0A1R4GUP9	SJM71966.1	1 COG0154	285.34
F5SNM6	EGK14717.1	1 COG0154	285.34
A0A1U6GJN2	SLJ83708.1	1 COG0154	284.19
A0A166H147	KZX19744.1	1 cl18951	130.49
A0A1W6KTY7	ARM90891.1	1 cl18951	281.88
A0A1H8N0M8	SEI09855.1	1 cl18951	165.54

Mycobacterium paratuberculosis
Mycobacterium smegmatis
Mycobacterium smegmatis (strain MKD8)
Mycobacterium terrae
Mycobacterium tuberculosis
Mycobacterium ulcerans str. Harvey
Mycobacterium xenopi 3993
Myxococcus xanthus (strain DK 1622)
Neisseria macacae ATCC 33926
Neisseria shayeganii 871
Nephila clavipes (Golden silk orbweaver)
Nitriicola nitratireducens
Nocardia cerradoensis
Nocardia farcinica
Nocardioides dokdonensis FR1436
Oceanibacterium hippocampi
Oceanobacillus oncorhynchi
Oceanobacillus picturae
Oleispira antarctica RB-8
Paenibacillus riograndensis SBR5
Paenibacillus sp. GM1FR
Pandoraea sputorum
Paraliobacillus sp. PM-2
Parendozoicomonas haliclona
Pelagimonas varians
Phaeobacter sp. CECT 5382
Planktothrix agardhii NIVA-CYA 126/8
Providencia burhodogranariae DSM 19968
Pseudarthrobacter siccitolerans
Pseudomonas citronellolis
Pseudomonas putida (*Arthrobacter siderocapsulatus*)
Pseudomonas sp. 1 R 17
Pseudomonas sp. 22 E 5
Pseudomonas sp. 31 E 6
Pseudomonas sp. 58 R 3
Pseudomonas sp. (strain NK87)
Pseudomonas sp. TAD18
Pseudomonas syringae pv. tomato
Pseudonocardia autotrophica (*Amycolata autotrophica*) (*Nocardia autotrophica*)
Pseudovibrio sp. Ad26
Psychrobacter pasteurii
Psychrobacter piechaudii
Psychrobacter sp. 1501(2011)
Psychrobacter sp. DAB_AL43B
Rathayibacter tanacetii
Rhizobium sp. CIAT894
Rhizobium tibeticum

A0A1F2Q4G2	OFV77896.1	1 cl18951	291.12		Rhodococcus erythropolis (Arthrobacter picolinophilus)
A0A165LS60	AMY53540.1	1 COG0154	325.4		Rhodococcus fascians D188
A0A0D8I4R7	KJF25039.1	1 PRK07869	738.72		Rhodococcus sp. AD45
A0A2I9DAU8	GBF16438.1	1 PRK07869	749.12		Rhodococcus sp. Br-6
A0A143Q7E3	AMY18317.1	1 cl18951	315.99		Rhodococcus sp. PBTS 1
A0A143QHT7	AMY22499.1	1 COG0154	326.56		Rhodococcus sp. PBTS 2
A0A1Y5RXJ7	SLN27523.1	1 COG0154	207.92		Roseisalinus antarcticus
A0A1Q2YXL0	GAV35946.1	1 COG0154	271.86		Roseomonas sp. TAS13
A0A0P1E894	CUH44666.1	1 COG0154	209.84		Ruegeria atlantica
Q5LQG3	AAV95777.1	1 COG0154	280.72		Ruegeria pomeroyi (strain ATCC 700808 / DSM 15171 / DSS-3) (Silicibacter pomeroyi)
A0A0P1I6U8	CUJ93919.1	1 COG0154	199.06		Ruegeria sp. CECT 5091
A0A0X2PD81	CUW25003.1	1 COG0154	265.31		Serratia
A0A2H1LLX6	SMZ58976.1	1 COG0154	265.31		Serratia grimesii
A0A021XH97	EYR84628.1	1 COG0154	301.91		Shinella sp. DD12
F7XC84	AEH81375.1	1 COG0154	278.41		Sinorhizobium meliloti (strain SM11)
A0A1V5I0C5	removed	0 n/a			Spirochaetes bacterium ADurb.BinA120
A0A239WPY8	SNV36266.1	1 cl18951	339.87	pfam07501	51.38 Streptococcus acidominimus
A0A0G6P5S0	CRV30531.1	1 PRK06529	900.34		Streptococcus equi subsp. equi
A0A239ST89	SNU88701.1	1 cl18951	790.56		Streptococcus merionis
Q8DTV5	AAN58903.1	1 PRK06529	909.2		Streptococcus mutans serotype c (strain ATCC 700610 / UA159)
A0A0Y1XIJ5	CVX74856.1	1 cl18951	112.22		Streptococcus pneumoniae
F8LIY3	CCB93303.1	1 cl18951	813.67		Streptococcus salivarius (strain CCHSS3)
A0A116M201	CYV23554.1	1 cl18951	786.71		Streptococcus suis
A0A2N9B277	SOR77440.1	1 COG0154	290.35		Streptomyces chartreusis NRRL 3882
A0A2K8PAY7	ATZ23280.1	1 COG0154	240.66		Streptomyces lavendulae subsp. lavendulae
A0A1Y2NKS5	OSY48102.1	1 COG0154	282.26		Streptomyces platensis
A0A0U5HFC4	CUW32201.1	1 COG0154	325.4		Streptomyces reticuli
A0A124C331	GAQ60006.1	1 COG0154	338.11		Streptomyces scabiei
A0A2P7ZYE4	PSK53232.1	1 cl18951	253.37		Streptomyces sp. 111WW2
A0A1D2IJJ3	ODA74583.1	1 cl18951	620.08		Streptomyces sp. AVP053U2
A0A1K2FZV5	SFY52671.1	1 COG0154	279.56		Streptomyces sp. F-1
A0A2K8REQ6	AUA15301.1	1 COG0154	264.54		Streptomyces sp. M56
A0A2P9FEF6	SPF06182.1	1 COG0154	339.27		Streptomyces sp. MA5143a
A0A068N7F4	AIE76027.1	1 pfam01425	377.02		Synechocystis sp. (strain PCC 6714) (Aphanocapsa sp. (strain PCC 6714))
A0A1V4XGE4	removed	0 n/a			Syntrophaceae bacterium PtaB.Bin095
A0A238D0E3	SBP86699.1	1 COG0154	340.81		Thiomonas delicata (Thiomonas cuprina)
A0A238DI14	SCC91137.1	1 COG0154	345.82		Thiomonas sp. X19
I3TLI8	AFK53626.1	1 cl18951	304.21		Tistrella mobilis (strain KA081020-065)
A0A0H2LVZ3	KLN54384.1	1 COG0154	262.62		Variovorax paradoxus
A0A1C3JAB8	SBT12094.1	1 COG0154	268.01		Vibrio celticus
A0A1C3II24	SBS61074.1	1 COG0154	266.85		Vibrio splendidus
A0A024QGQ4	CDQ41684.1	1 cl18951	473.15		Virgibacillus massiliensis
A0A2K9J0P4	AUJ25274.1	1 cl18951	460.83		Virgibacillus pantothenticus

SAMPLE SCREEN SHOT OF CONSERVED DOMAINS DATABASE SCREEN:

Conserved domains on [gi|1318755899|gb|AUJ25274.1|]

View Concise Results

6-aminohexanoate-cyclic-dimer hydrolase [Virgibacillus pantothenicus]

Graphical summary Zoom to residue level [show extra options >](#)

Query seq. Superfamilies

Search for similar domain architectures [?](#) Refine search [?](#)

List of domain hits

Name	Accession	Description	Interval	E-value
Amidase super family	cl18951	Amidase;	10-491	3.23e-159

The actual alignment was detected with superfamily member PRK06529:

Pssm-ID: 327489 Cd Length: 482 Bit Score: 460.83 E-value: 3.23e-159

```

      10      20      30      40      50      60      70      80
gi 1318755899 10 DAVDLADLVKKGETT EEELLQASLEQLAQMPTLHAVISERKEKALQEWKDITY-NRPF
Cdd:PRK06529  5 DATAMAQAVQQGVPTLELVTAIYKAKKLNPTLNAIVSERYE EALEEAKQRDF
      90     100     110     120     130     140     150     160
gi 1318755899 89 GSKLLKQYapVAKQDSHFVRAFRDAGFQFIHNTPEFGLKKNITEPETDGPTRNPWNVAYSP
Cdd:PRK06529  85 GSRLFKNY--QATKTDLYVKRLEDLGF IILGRSNTPEFGFKNISDSSLHGPVNL
      170     180     190     200     210     220     230     240
gi 1318755899 169 AGASDGGGSIIRIPASFTGLFGLKPTRGRTPVGPGAGRQWGHGAIDFVLSRSVRD
Cdd:PRK06529  163 AAASDGGGSIIRIPASFNGLI GLKPSRGRIPVGPGRGWSQASVHFALTKSVRDR
      250     260     270     280     290     300     310     320
gi 1318755899 249 IETMKQPPKEELRIAYSVTSPVGTSPVNEAKEAVMKTQWLERLGYIEE-NDPP
Cdd:PRK06529  241 -ESLQSLQRPLKIAFYQRSPDGPVSLDAALKQAVTFLREQGHEVVEIEEFPLDMTE
      330     340     350     360     370     380     390     400
gi 1318755899 328 EKMIGRKITSDDVELETWLLHCAGQSVSAAAFSSSIASWDVAAAAMADFHKS
Cdd:PRK06529  320 EDALGRPMTKDDMETMTWAIYQSGQDIPAKRYSQVLQKWDTYSATMASFHETYD
      410     420     430     440     450     460     470     480
gi 1318755899 408 wITKMENAKANEKQEIYDMFLPSLTYPTQLANLTGQPAMSLPLHVTSENMP
Cdd:PRK06529  400 -LAQAEIFSSEEQQNLVETMFEKSLAITPYTALANLTGQPAISLPTYETKEGLPMG
      488     491
gi 1318755899 488 LWQG 491
Cdd:PRK06529  479 LLKI 482
    
```


SAMPLE SCREEN SHOT OF UNIPROT NYLA SEARCH:

UniProtKB

BLAST Align Retrieve/ID mapping Peptide search Help Contact

From June 20, 2018 all traffic will be automatically redirected to HTTPS. [More information](#) or [view this page using https](#)

UniProtKB results

[About UniProtKB](#) [Basket](#)

Filter byⁱ

◀ 76 to 100 of 443 ▶ Show 25 ▾

Entry	Entry name	Organism	Protein names	Gene names	Length
<input type="checkbox"/> A0A116LMD4	A0A116LMD4_STRSU	Streptococcus suis	Amidase	nylA ERS132421_01386	482
<input type="checkbox"/> A0A0Z8E6P7	A0A0Z8E6P7_STRSU	Streptococcus suis	Amidase	nylA ERS132370_01358	482
<input type="checkbox"/> A0A0Z8L219	A0A0Z8L219_STRSU	Streptococcus suis	Amidase	nylA ERS132440_01857	482
<input type="checkbox"/> A0A0Z8F9P1	A0A0Z8F9P1_STRSU	Streptococcus suis	Amidase	nylA ERS132502_00463	482
<input type="checkbox"/> A0A0Z8SP32	A0A0Z8SP32_STRSU	Streptococcus suis	Amidase	nylA ERS132536_01456	482
<input type="checkbox"/> A0A0Z8S8C1	A0A0Z8S8C1_STRSU	Streptococcus suis	Amidase	nylA ERS132551_00941	482
<input type="checkbox"/> A0A0Z8CTP8	A0A0Z8CTP8_STRSU	Streptococcus suis	Amidase	nylA ERS132374_01083	482
<input type="checkbox"/> A0A0Z8GQ06	A0A0Z8GQ06_STRSU	Streptococcus suis	Amidase	nylA ERS132393_01944	482
<input type="checkbox"/> A0A123T3Y1	A0A123T3Y1_STRSU	Streptococcus suis	Amidase	nylA ERS132392_02100	482
<input type="checkbox"/> A0A116LQC9	A0A116LQC9_STRSU	Streptococcus suis	Amidase	nylA ERS132410_01796	482
<input type="checkbox"/> A0A0Z8DJA8	A0A0Z8DJA8_STRSU	Streptococcus suis	Amidase	nylA ERS132399_00060	482
<input type="checkbox"/> A0A0Z8QS84	A0A0Z8QS84_STRSU	Streptococcus suis	Amidase	nylA ERS132444_00393	482
<input type="checkbox"/> A0A123T9P9	A0A123T9P9_STRSU	Streptococcus suis	Amidase	nylA ERS132444_00393	482

Search terms

Filter "nyla" as:

gene name (374)

organism (65)

protein name (7)

strain (1)

We'd like to inform you that we have updated our [Privacy Notice](#) to comply with Europe's new General Data Protection Regulation (GDPR) that applies since 25 May 2018. [Do not show this banner again](#)

Testing the Hypothesis that the Nylonase NylB Protein
Arose de novo via a Frameshift Mutation

Salvador Cordova and John Sanford

Supplementary Table S4

List of organisms with NylC from uniprot, and conserved domains of NylCs
with sample screen captures from NIH CDD and UNIPROT websites at the bottom

NOTES:

cd00123 = DmpA/OAT superfamily; composed of L-aminopeptidase D-amidase/D-esterase (DmpA), ornithine acetyltransferase (OAT) and similar proteins. DmpA is an aminopeptidase that releases N-terminal D and L amino acids from peptide substrates. This group represents one of the rare aminopeptidases that are not metalloenzymes. DmpA shows similarity in catalytic mechanism to N-terminal nucleophile (Ntn) hydrolases, which are enzymes that catalyze the cleavage of amide bonds through the nucleophilic attack of the side chain of an N-terminal serine, threonine, or cysteine.

OAT catalyzes the first and fifth steps in arginine biosynthesis, coupling acetylation of glutamate with deacetylation of N-acetylornithine, which allows recycling of the acetyl group in the arginine biosynthetic pathway. The superfamily also contains an enzyme, endo-type 6-aminohexanoate-oligomer hydrolase, that have been shown to be involved in nylon degradation. Proteins in this superfamily undergo autocatalytic cleavage of an inactive precursor at the site immediately upstream to the catalytic nucleophile.

cl00603 = DmpA/OAT superfamily; composed of L-aminopeptidase D-amidase/D-esterase (DmpA), ornithine acetyltransferase (OAT) and similar proteins. DmpA is an aminopeptidase that releases N-terminal D and L amino acids from peptide substrates. This group represents one of the rare aminopeptidases that are not metalloenzymes. DmpA shows similarity in catalytic mechanism to N-terminal nucleophile (Ntn) hydrolases, which are enzymes that catalyze the cleavage of amide bonds through the nucleophilic attack of the side chain of an N-terminal serine, threonine, or cysteine. OAT catalyzes the first and fifth steps in arginine biosynthesis, coupling acetylation of glutamate with deacetylation of N-acetylornithine, which allows recycling of the acetyl group in the arginine biosynthetic pathway. The superfamily also contains an enzyme, endo-type 6-aminohexanoate-oligomer hydrolase, that have been shown to be involved in nylon degradation. Proteins in this superfamily undergo autocatalytic cleavage of an inactive precursor at the site immediately upstream to the catalytic nucleophile.

cd02252 = nylC-like family; composed of proteins with similarity to Flavobacterium endo-type 6-aminohexanoate-oligomer hydrolase (EIII), the product of the nylon oligomer degradation gene, nylC. EIII is an amide hydrolase that catalyzes the degradation of highly-polymerized 6-aminohexanoate oligomers. Together with other nylon degradation enzymes, such as 6-aminohexanoate cyclic dimer hydrolase (EI) and 6-aminohexanoate dimer hydrolase (EII), EIII plays a role in the detoxification and biological removal of the synthetic by-products of nylon manufacture. EIII shows sequence similarity to L-aminopeptidase D-amidase/D-esterase (DmpA), an aminopeptidase that releases N-terminal D and L amino acids from peptide substrates. Like DmpA, EIII undergoes autocatalytic cleavage in front of a nucleophile to form a heterodimer. DmpA shows similarity in catalytic mechanism to N-terminal nucleophile (Ntn) hydrolases, which are enzymes that catalyze the cleavage of amide bonds through the nucleophilic attack of the side chain of an N-terminal serine, threonine, or cysteine.

no. avail. cdds 7
 count NylCs 7

UNIPROT ID	Genbank ID	Conserved Domain Data Available	Primary Conserved Domain ACCESSION ID	Primary Conserved Domain Bit score	Secondary Conserved Domain ACCESSION ID	Secondary Conserved Domain Bit score	ORGANISM
Q79F77	BAA01528.1	1	cd00123	313.18			Flavobacterium sp.
Q57326	BAA05088.1	1	cd00123	313.18			Flavobacterium sp. KI723T1
Q1EPR5	BAE95769.2	1	cd00123	305.47			Agromyces sp. KY5R
Q1EPR4	BAE95770.1	1	cd00123	303.16			Kocuria sp. KY2
U2XM45	GAD34180.1	1	cl00603	196.68			Microbacterium sp. TS-1
C0R138	ACN83826.1	1	cd02252	306.46			Brachyspira hyodysenteriae (strain ATCC 49526 / WA1)
G0EJE2	AEM21141.1	1	cd02252	304.15			Brachyspira intermedia (strain ATCC 51140 / PWS/A) (Serpulina intermedia)

SAMPLE CDD WEBSITE SCREEN SHOT:

Protein Classification
 DmpA_OAT domain-containing protein (domain architecture ID 10062641)
 DmpA_OAT domain-containing protein

Graphical summary Zoom to residue level [show extra options >](#)

List of domain hits

Name	Accession	Description	Interval	E-value
DmpA_OAT	cd00123	DmpA/OAT superfamily; composed of L-aminopeptidase D-amidase/D-esterase (DmpA), ornithine ...	55-354	1.13e-101

DmpA/OAT superfamily; composed of L-aminopeptidase D-amidase/D-esterase (DmpA), ornithine acetyltransferase (OAT) and similar proteins. DmpA is an aminopeptidase that releases N-terminal D and L amino acids from peptide substrates. This group represents one of the rare aminopeptidases that are not metalloenzymes. DmpA shows similarity in catalytic mechanism to N-terminal nucleophile (Ntn) hydrolases, which are enzymes that catalyze the cleavage of amide bonds through the nucleophilic attack of the side chain of an N-terminal serine, threonine, or cysteine. OAT catalyzes the first and fifth steps in arginine biosynthesis, coupling acetylation of glutamate with deacetylation of N-acetylornithine, which allows recycling of the acetyl group in the arginine biosynthetic pathway. The superfamily also contains an enzyme, endo-type 6-aminohexanoate-oligomer hydrolase, that have been shown to be involved in nylon degradation. Proteins in this superfamily undergo autocatalytic cleavage of an inactive precursor at the site immediately upstream to the catalytic nucleophile.

Pssm-ID: 238070 Cd Length: 286 Bit Score: 305.47 E-value: 1.13e-101

```

      10      20      30      40      50      60      70      80
gi 112181151 55 FPGVSIQAAYEE---GPTGATVIHIPAGARTAVDARGGAV-----GLSGGYDFNHAICLAGGASYGLEAGAGVSGALL 125
Cdd:cd00123  1 PRGVVGTAPVGEaddGRDGFVIASTAPATVSVVFRGRFagplcrEAVAGGQFRHGVVVLARNEGEENAREVREAVAR 80

      90     100     110     120     130     140     150     160
gi 112181151 126 ERLEYRTGFAEQLVSSAVIYDF SARSTAVYPDKALGRAALEFAVPGFPGGAGmsasAGKVDNDRTETITGQGA AFR 205
  
```



```

Cdd:cd00123 81 ARGLPRTGFAEEGE*LIASYDIGRQYTP*ESIRAHRLRTALWPAGEGGFDRGRAS-----AGAARAI*TTDTGPGEARR 154
      170      180      190      200      210      220      230      240
      .....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|
gi 112181151 206 RLGDVRIILAVVVPNPVGVIMDRAGTVVRGnydaqtgvrhrpvyfdyQEAFAEQVPPVT EAGNTTISAIVTNVRMSPVELNQ 285
Cdd:cd00123 155 SVGGATIVAIKNG*LEIVDRAGTVVRG-----QEAFAEQVPPVTPD*ATLITFFATDARLDP AELDR 218
      250      260      270      280      290      300
      .....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|
gi 112181151 286 FAKQVHSSMHRGIQPfHTDMDGDTLFAVTTDEIDLPTTPGSSRGRLSVNATALGAIASEVMWDAVLEAG 354
Cdd:cd00123 219 LARV*DRTFNVRVSIID-TDSTGDTAVAFATGLAGLPTTPGSSRGRLEVDAGEFEAAHTAALAAVKDAA 286

```

References:

Morabler Bauer A et al. (2017). "CD0/SNAPCLE: functional classification of proteins via subfamily domain architectures." *Nucleic Acids Res* 45(D)2002

SAMPLE SCREEN CAPTURE FROM UNIPROT:

The screenshot shows the UniProtKB search results page. At the top, there is a search bar with the query 'nylc' and a search button. Below the search bar, there are navigation links for 'BLAST', 'Align', 'Retrieve/ID mapping', and 'Peptide search'. A yellow banner indicates that from June 20, 2018, all traffic will be automatically redirected to HTTPS. The main heading is 'UniProtKB results'. On the left side, there are filters for 'Reviewed (1) Swiss-Prot' and 'Unreviewed (10) TrEMBL'. Below the filters, there are 'Popular organisms' listed: FLASP (1), FLASK (1), Kocuria sp. KY2 (1), Agromyces sp. KY5R (2), and Flavobacterium sp. KI723T1 (2). The main content area shows a table of search results with columns for 'Entry' and 'Organism'. The table contains 11 entries, all from the genus Flavobacterium. The first entry is P07062, Flavobacterium sp. (strain K172). The second entry is Q79F77, Flavobacterium sp. The third entry is Q57326, Flavobacterium sp. KI723T1. The fourth entry is Q1EPR5, Agromyces sp. KY5R. The fifth entry is Q1EPR4, Kocuria sp. KY2. The sixth entry is U2XM45, Microbacterium sp. TS-1. The seventh entry is Q14T92, Agromyces sp. KY5R. The table also includes a 'Columns' button and a 'Show 250' dropdown menu.

Other organisms

Search terms

Filter "nyc" as:

gene name (6)

protein name (3)

<input type="checkbox"/>	Q59710	Flavobacterium sp. KI723T1
<input type="checkbox"/>	C0R138	Brachyspira hyodysenteriae (strain ATCC 49526 / WA1)
<input type="checkbox"/>	G0EJE2	Brachyspira intermedia (strain ATCC 51140 / PWS/A) (Serpulina intermedia)
<input type="checkbox"/>	A0A146TFP9	Fundulus heteroclitus (Killifish) (Mummichog)

1 to 11 of 11 Show 250 ▾

View by

We'd like to inform you that we have updated our [Privacy Notice](#) to comply with Europe's new General Data Protection Regulation (GDPR) that applies since 25 May 2018.

Results table

[Do not show this banner again](#)

Taxonomy

Testing the Hypothesis that the Nylonase NylB Protein
Arose *de novo* via a Frameshift Mutation

Salvador Cordova and John Sanford

Supplementary Table S5
List of organisms with a 6-aminohexanoate hydrolase from UNIPROT
with sample screen capture from UNIPROT website at the bottom

UNIPROT ID	ORGANISM
1 A0A1F2PJC9	Acetobacterium wieringae
2 A9NHQ9	Acholeplasma laidlawii (strain PG-8A)
3 A0A1Z3GY09	Achromobacter denitrificans (Alcaligenes denitrificans)
4 A0A225MGN	Achromobacter marplatensis
5 A0A0L6IER7	Achromobacter piechaudii
6 D4XAB8	Achromobacter piechaudii ATCC 43553
7 A0A1D8IES6	Achromobacter ruhlandii
8 A0A0M7D486	Achromobacter sp.
9 A0A162N388	Achromobacter sp. AR476-2
10 A0A246Q7U7	Achromobacter sp. HZ28
11 A0A0U1QAU	Achromobacter sp. LC458
12 A0A0Q8K4P9	Achromobacter sp. Root170
13 A0A0Q8BA96	Achromobacter sp. Root565
14 A0A0Q8X5T1	Achromobacter sp. Root83
15 A0A095X3S6	Achromobacter sp. RTa
16 A0A0L0QTC4	Achromobacter spanius
17 A0A066PM46	Acidiphilium sp. JA12-A1
18 F7S2L8	Acidiphilium sp. PM
19 A0A0D8HH63	Acidithrix ferrooxidans
20 A0A165JNH4	Acidovorax sp. GW101-3H11
21 K0HY31	Acidovorax sp. KKS102
22 A0A235EL10	Acidovorax sp. KNDSW-TSA6
23 A0A0S9F8F4	Acidovorax sp. Leaf76
24 A0A0Q4RIF5	Acidovorax sp. Leaf78
25 H0BTY1	Acidovorax sp. NO-1
26 A0A0Q8TAG	Acidovorax sp. Root217
27 A0A0Q8T7V0	Acidovorax sp. Root219
28 A0A0Q9CUG	Acidovorax sp. Root275
29 A0A0Q6ZDN	Acidovorax sp. Root402
30 A0A0Q8LML1	Acidovorax sp. Root70
31 A0A1E3ZU93	Acidovorax sp. SCN 65-28
32 A0A0Q0SWC	Acidovorax sp. SD340
33 A0A1G5L686	Acinetobacter baumannii

34 A0A239RVT3 *Acinetobacter johnsonii*
35 S7WVF4 *Acinetobacter junii* MTCC 11364
36 A0A150HSM2 *Acinetobacter venetianus*
37 A0A1V5IP05 *Actinobacteria bacterium ADurb.BinA094*
38 W7JEG2 *Actinokineospora spheciospongiae*
39 A0A1Q9TSC2 *Actinomadura* sp. CNU-125
40 U5W414 *Actinoplanes friuliensis* DSM 7358
41 G8SAD2 *Actinoplanes* sp. (strain ATCC 31044 / CBS 674.73 / SE50/110)
42 A0A0M3DZD2 *Afipia massiliensis*
43 A0A1M2ZX57 *Afipia* sp. 62-7
44 A0A1H0JGI6 *Afipia* sp. GAS231
45 A0A0Q6Z796 *Afipia* sp. Root123D2
46 A0A1Z9VA19 *Afipia* sp. TMED4
47 R9PTA2 *Agarivorans albus* MKT 106
48 K2QYS9 *Agrobacterium albertimagini* AOL15
49 A0A0D8KR75 *Agrobacterium arsenijevicei*
50 A0A1S7RI05 *Agrobacterium fabrum* str. J-07
51 A0A1S7MVW *Agrobacterium genomosp. 1* str. TT111
52 A0A1S7MNB7 *Agrobacterium genomosp. 2* str. CFBP 5494
53 B9JH68 *Agrobacterium radiobacter* (strain K84 / ATCC BAA-868)
54 A0A071I2G0 *Agrobacterium rhizogenes*
55 A0A1R3U120 *Agrobacterium* sp. DSM 25559
56 A0A1B3NSJ3 *Agrobacterium* sp. RAC06
57 A0A1E3Y6R8 *Agrobacterium* sp. SCN 61-19
58 F0LFF7 *Agrobacterium* sp. (strain H13-3) (*Rhizobium lupini* (strain H13-3))
59 H0HHN7 *Agrobacterium tumefaciens* 5A
60 A0A202F2X2 *Agrobacterium vitis* (*Rhizobium vitis*)
61 Q14TA0 *Agromyces* sp. KY5R
62 A0A164DTT4 *Agromyces* sp. NDB4Y10
63 A0A0M9GN66 *Ahrensia marina*
64 A0A1H2T3N9 *Albimonas donghaensis*
65 A0A1I3JY37 *Albimonas pacifica*
66 A0A0A2NEC5 *Alcaligenes faecalis*
67 A0A1B6AXE2 *Alcaligenes faecalis* subsp. *faecalis* NBRC 13111
68 Q5YFS2 *Alcaligenes xylosoxydans xylosoxydans* (*Achromobacter xylosoxidans*)
69 A0A0P1J495 *Alcanivorax dieselolei*
70 A0A0B4XK86 *Alcanivorax pacificus* W11-5
71 A0A154C0V0 *Alcanivorax* sp. KX64203
72 A0A1V0FQE3 *Alcanivorax xenomutans*
73 A0A090VEN9 *Algibacter lectus*
74 B7DSV3 *Alicyclobacillus acidocaldarius* LAA1
75 A0A0P9CFM2 *Alicyclobacillus ferrooxydans*
76 A0A2A2G9D6 *Aliifodinibius* sp. WN023
77 A0A0P7KJ31 *Aliiroseovarius crassostreae*
78 A0A1B9PJ06 *Aliivibrio* sp. 1S175

79 A0A259NTY8 Alishewanella sp. 34-51-39
80 A0A011SHZ7 Alkalibacterium sp. AK22
81 A0A225Q6Z4 Alkalimicrobium pacificum
82 A0A1S8CRT6 Alkanindiges sp. H1
83 A0A0N1LKH9 alpha proteobacterium AAP38
84 A0A061QAK6 alpha proteobacterium Q-1
85 A0A1V1PU51 alpha proteobacterium U9-1i
86 A0A2A5BK32 Alphaproteobacteria bacterium
87 A0A258H4W7 Alphaproteobacteria bacterium 32-64-14
88 A0A1Q3JWA7 Alphaproteobacteria bacterium 64-6
89 A0A1Q3IWY5 Alphaproteobacteria bacterium 65-7
90 A0A1V6FWC7 Alphaproteobacteria bacterium ADurb.Bin100
91 A0A1V5FSP3 Alphaproteobacteria bacterium ADurb.BinA280
92 A0A1L9QFD7 Alphaproteobacteria bacterium AO1-B
93 A0A257HQ80 Alphaproteobacteria bacterium PA1
94 A0A257J3M2 Alphaproteobacteria bacterium PA2
95 A0A257JNK2 Alphaproteobacteria bacterium PA3
96 A0A1F3AL54 Alphaproteobacteria bacterium RIFCSPHIGHO2_12_FULL_63_12
97 A0A1Z9D0B1 Alphaproteobacteria bacterium TMED110
98 A0A1Z9WK57 Alphaproteobacteria bacterium TMED89
99 A0A0F7KSK3 Altererythrobacter atlanticus
100 A0A1B2A9T7 Altererythrobacter dongtanensis
101 A0A0M4M6L1 Altererythrobacter epoxidivorans
102 A0A1C7D5F0 Altererythrobacter namhicola
103 A0A0T2QIV8 Altererythrobacter sp. Root672
104 A0A1Z9BMI0 Alteromonadaceae bacterium TMED101
105 A0A0N0XJM8 Amantichitinum ursilacus
106 A0A142M3Y6 Aminobacter aminovorans (Chelatobacter heintzii)
107 A0A0Q6LJV1 Aminobacter sp. Root100
108 A0A178X0M9 Amycolatopsis sp. M39
109 A0A229RJG4 Amycolatopsis thailandensis
110 A0A1B7V2W7 Anabaena sp. LE011-02
111 A0A1B7WVL8 Anabaena sp. MDT14b
112 A0A0D6QPK7 Anaeromyxobacter sp. PSR-1
113 A7HFI3 Anaeromyxobacter sp. (strain Fw109-5)
114 A0A136WJN9 Anaerotignum neopropionicum
115 A0A1S1V4U0 Andreesenia angusta
116 A0A0U5BA19 Aneurinibacillus soli
117 A0A222E5F8 Antarctobacter heliothermus
118 A0A011UTK5 Aquamicrobium defluvii
119 A0A023BRB5 Aquimarina atlantica
120 A0A1Y5T7Z2 Aquimixticola soesokkakensis
121 A0A224VWD9 Aquitalea magnusonii
122 A0A1Q8A510 archaeon 13_1_20CM_2_54_9
123 A0A1Q3HRQ9 Archangium sp. Cb G35

124 A0A1B6A3B9 *Arenibacter* sp. C-21
125 N1UZN7 *Arthrobacter crystallopoietes* BAB-32
126 A0A0S7CR24 *Arthrobacter* sp. Hiyo1
127 A0A0K2QWB *Arthrobacter* sp. Hiyo4
128 A0A0K2RNW *Arthrobacter* sp. Hiyo8
129 A0A0D4BXM: *Arthrobacter* sp. IHBB 11108
130 J7LU22 *Arthrobacter* sp. Rue61a
131 A0A1E8FMB1 *Arthrobacter* sp. SW1
132 A2QLY1 *Aspergillus niger* (strain CBS 513.88 / FGSC A1513)
133 V4MYJ7 *Asticcacaulis* sp. AC402
134 V4PU99 *Asticcacaulis* sp. AC460
135 A0A0Q6FA40 *Aurantimonas* sp. Leaf443
136 A0A0B1Q326 *Aureimonas altamirensis*
137 A0A0Q5HAK: *Aureimonas* sp. Leaf324
138 A0A0Q6CPC: *Aureimonas* sp. Leaf427
139 A0A0Q6CP7: *Aureimonas* sp. Leaf454
140 A0A0Q6DW7 *Aureimonas* sp. Leaf460
141 A0A175R826 *Aureimonas ureilytica*
142 A0A087SMR: *Auxenochlorella protothecoides* (Green microalga) (*Chlorella protothecoides*)
143 A0A235H8Q6 *Azospirillum brasilense*
144 D3P1A2 *Azospirillum* sp. (strain B510)
145 A0A0F2KK19 *Azospirillum thiophilum*
146 A0A2C3GRV: *Bacillus anthracis*
147 A0A1J6VSB4 *Bacillus aquimaris*
148 A0A1D7QVD: *Bacillus beveridgei*
149 A0A023P794 *Bacillus bombysepticus* str. Wang
150 A0A2C2HETC *Bacillus cereus*
151 C2NZG8 *Bacillus cereus* 172560W
152 C2TIZ3 *Bacillus cereus* 95/8201
153 B5ULQ5 *Bacillus cereus* AH1134
154 C2YSH7 *Bacillus cereus* AH1271
155 C2Z8T6 *Bacillus cereus* AH1272
156 C2XV04 *Bacillus cereus* AH603
157 C2PWE8 *Bacillus cereus* AH621
158 C2N382 *Bacillus cereus* ATCC 10876
159 C2QVJ8 *Bacillus cereus* ATCC 4342
160 R8HVQ8 *Bacillus cereus* BAG1O-1
161 S3IIV2 *Bacillus cereus* BAG1O-3
162 R8DWH5 *Bacillus cereus* BAG1X1-1
163 R8FKZ9 *Bacillus cereus* BAG1X2-1
164 R8FW19 *Bacillus cereus* BAG1X2-2
165 R8GEZ8 *Bacillus cereus* BAG1X2-3
166 R8K6C8 *Bacillus cereus* BAG2O-1
167 C2T1Y1 *Bacillus cereus* BDRD-Cer4
168 C2SL82 *Bacillus cereus* BDRD-ST196

169 C2S4A4	Bacillus cereus BDRD-ST26
170 C2NIK6	Bacillus cereus BGSC 6E1
171 C2XCG9	Bacillus cereus F65185
172 R8D9F4	Bacillus cereus HuA2-9
173 R8D3V5	Bacillus cereus HuA3-9
174 R8M143	Bacillus cereus HuB13-1
175 R8RYL3	Bacillus cereus HuB4-4
176 R8Q0D3	Bacillus cereus ISP2954
177 R8IPR3	Bacillus cereus K-5975c
178 C2MLL8	Bacillus cereus m1293
179 C2RAP0	Bacillus cereus m1550
180 R8K9D9	Bacillus cereus MC118
181 C2PFW9	Bacillus cereus MM3
182 B3Z4G2	Bacillus cereus NVH0597-99
183 C2QCQ2	Bacillus cereus R309803
184 C2UES0	Bacillus cereus Rock1-15
185 C2UWA5	Bacillus cereus Rock3-28
186 C2VCU2	Bacillus cereus Rock3-29
187 C2VUP1	Bacillus cereus Rock3-42
188 C2W8N2	Bacillus cereus Rock3-44
189 C2X0T7	Bacillus cereus Rock4-18
190 C2WPT9	Bacillus cereus Rock4-2
191 R8CAB4	Bacillus cereus str. Schrouff
192 A0A158RQH7	Bacillus cereus (strain 03BB102)
193 B7HXV3	Bacillus cereus (strain AH187)
194 B7JFB1	Bacillus cereus (strain AH820)
195 Q734U4	Bacillus cereus (strain ATCC 10987 / NRS 248)
196 Q81CK3	Bacillus cereus (strain ATCC 14579 / DSM 31 / JCM 2152 / NBRC 15305 / NCIMB 9373 / NRRL B-3711)
197 B7HC46	Bacillus cereus (strain B4264)
198 B7IPG8	Bacillus cereus (strain G9842)
199 B9IRP0	Bacillus cereus (strain Q1)
200 Q63AJ3	Bacillus cereus (strain ZK / E33L)
201 R8Y226	Bacillus cereus TIAC219
202 D8H0A0	Bacillus cereus var. anthracis (strain CI)
203 R8HXD7	Bacillus cereus VD021
204 R8QEP8	Bacillus cereus VD118
205 R8LMS7	Bacillus cereus VD131
206 R8E0R6	Bacillus cereus VD133
207 R8P6H2	Bacillus cereus VD136
208 J7WYF9	Bacillus cereus VD142
209 R8MZI1	Bacillus cereus VD146
210 R8TXJ3	Bacillus cereus VD184
211 R8H2U2	Bacillus cereus VD196
212 R8QBI6	Bacillus cereus VDM006
213 R8ET56	Bacillus cereus VDM019

214 R8TE56 Bacillus cereus VDM021
215 R8P5Z7 Bacillus cereus VDM053
216 A0A0C5CBF9 Bacillus coagulans
217 A0A1D3QXB5 Bacillus cytotoxicus
218 A7GQG2 Bacillus cytotoxicus (strain DSM 22905 / CIP 110041 / 391-98 / NVH 391-98)
219 A0A073JYG4 Bacillus gaemokensis
220 A0A0T6BV65 Bacillus glycinifermentans
221 A0A2C6EZGC Bacillus halotolerans
222 W4QEL0 Bacillus hemicellulosilyticus JCM 9152
223 A0A1V4FWV1 Bacillus intestinalis
224 A0A1Q8GU15 Bacillus licheniformis
225 A0A073JVJ5 Bacillus manliponensis
226 A0A1C3T9V5 Bacillus mycoides
227 W4EZC5 Bacillus mycoides FSL H7-687
228 C3AUD0 Bacillus mycoides Rock1-4
229 A0A1L8ZGH6 Bacillus obstructivus
230 A0A2C3KNT7 Bacillus pseudomycoides
231 C3BM98 Bacillus pseudomycoides DSM 12442
232 A0A1Q9BAH1 Bacillus pumilus (Bacillus mesentericus)
233 A0A0Q3VNL2 Bacillus solani
234 A0A1K0AIB5 Bacillus sp. 0711P9-1
235 A0A2A8MF21 Bacillus sp. AFS001701
236 A0A2A7X098 Bacillus sp. AFS002410
237 A0A2A8V769 Bacillus sp. AFS015896
238 A0A2A8S7B5 Bacillus sp. AFS018417
239 A0A2C3JV36 Bacillus sp. AFS029533
240 A0A2C3H8D5 Bacillus sp. AFS029637
241 A0A2C2LA07 Bacillus sp. AFS033286
242 A0A2C0YT17 Bacillus sp. AFS041924
243 A0A2C3KYA3 Bacillus sp. AFS051223
244 A0A2B9AS36 Bacillus sp. AFS053548
245 A0A2B8UE39 Bacillus sp. AFS054943
246 A0A2B8TV50 Bacillus sp. AFS055030
247 A0A2B4FNX2 Bacillus sp. AFS059628
248 A0A2B4M774 Bacillus sp. AFS075034
249 A0A2B4N4N4 Bacillus sp. AFS075960
250 A0A2B1GIX3 Bacillus sp. AFS077874
251 A0A2A9S957 Bacillus sp. AFS088145
252 A0A2A7E576 Bacillus sp. AFS094611
253 A0A2A7I6R0 Bacillus sp. AFS096315
254 A0A2A7H3Z5 Bacillus sp. AFS098217
255 A0A1V9W725 Bacillus sp. CDB3
256 A0A0N0E9MC Bacillus sp. CHD6a
257 A0A1J9ZJK2 Bacillus sp. EB422
258 A0A0M4GRK Bacillus sp. FJAT-18017

259 A0A0M0W3R Bacillus sp. FJAT-18019
260 A0A0M0WY2 Bacillus sp. FJAT-21945
261 A0A1B8W0Y Bacillus sp. FJAT-26390
262 A0A1B9API9 Bacillus sp. FJAT-27225
263 A0A1B8WPD Bacillus sp. FJAT-27264
264 A0A0W7Y5B Bacillus sp. G3(2015)
265 N1LTX0 Bacillus sp. GeD10
266 A0A154AY22 Bacillus sp. GZT
267 W7ZHI9 Bacillus sp. JCM 19045
268 W7ZIP8 Bacillus sp. JCM 19046
269 W7ZMG2 Bacillus sp. JCM 19047
270 A0A135WX0C Bacillus sp. JH7
271 A0A2D0AU11 Bacillus sp. K2I17
272 A0A229M5W! Bacillus sp. KbaB1
273 A0A229MG3E Bacillus sp. KbaL1
274 A0A0J6NW31 Bacillus sp. LK2
275 A0A0J5GSX0 Bacillus sp. LL01
276 A0A226R6S0 Bacillus sp. M13(2017)
277 A0A246US95 Bacillus sp. MB353a
278 A0A1Q9KUVz Bacillus sp. MB366
279 A0A1J9Z069 Bacillus sp. MN5
280 A0A1J9SQX6 Bacillus sp. N35-10-2
281 A0A1J9T6B4 Bacillus sp. N35-10-4
282 A0A1J9VQB7 Bacillus sp. NH24A2
283 Q2B3E8 Bacillus sp. NRRL B-14911
284 A0A0Q9GJS1 Bacillus sp. Root11
285 A0A0Q9GLHf Bacillus sp. Root131
286 A0A1G1UPW Bacillus sp. RZ2MS9
287 A6CHJ8 Bacillus sp. SG-1
288 A0A1K0A8L5 Bacillus sp. TD41
289 A0A1K0BSB8 Bacillus sp. TD42
290 A0A0F5R JL6 Bacillus sp. UMTAT18
291 A0A150L8N9 Bacillus sporothermodurans
292 A0A0D6Z7E6 Bacillus subterraneus
293 A0A136GAMf Bacillus subtilis
294 L8AQW9 Bacillus subtilis BEST7613
295 A0A0D0G9L6 Bacillus thermoamylovorans
296 A0A2B7CJL1 Bacillus thuringiensis
297 A0A1C9BVDf Bacillus thuringiensis Bt18247
298 J3UJR5 Bacillus thuringiensis HD-771
299 K0FV12 Bacillus thuringiensis MC28
300 C3G580 Bacillus thuringiensis serovar andalousiensis BGSC 4AW1
301 A0A243KDT7 Bacillus thuringiensis serovar argentinensis
302 A0A242Z7T1 Bacillus thuringiensis serovar balearica
303 A0A242WUYf Bacillus thuringiensis serovar cameroun

304 A0A243NER8 Bacillus thuringiensis serovar canadensis
305 A0A242YHW1 Bacillus thuringiensis serovar chanpasis
306 A0A193CUP1 Bacillus thuringiensis serovar coreanensis
307 A0A243B XK5 Bacillus thuringiensis serovar graciosensis
308 A0A242XVI3 Bacillus thuringiensis serovar guiyangensis
309 C3H3I2 Bacillus thuringiensis serovar huazhongensis BGSC 4BD1
310 A0A243KKJ1 Bacillus thuringiensis serovar iberica
311 Q3EMT6 Bacillus thuringiensis serovar israelensis ATCC 35646
312 A0A243A015 Bacillus thuringiensis serovar kim
313 A0A242ZE72 Bacillus thuringiensis serovar londrina
314 A0A242X1Z4 Bacillus thuringiensis serovar malayensis
315 A0A242WC76 Bacillus thuringiensis serovar mexicanensis
316 A0A243A8I4 Bacillus thuringiensis serovar navarrensis
317 A0A242Y6U4 Bacillus thuringiensis serovar novosibirsk
318 A0A243FPY2 Bacillus thuringiensis serovar ostriniae
319 C3E5Q2 Bacillus thuringiensis serovar pakistani str. T13001
320 A0A243AXF7 Bacillus thuringiensis serovar pingluonsis
321 A0A243JFK4 Bacillus thuringiensis serovar pirenaica
322 A0A243BKT3 Bacillus thuringiensis serovar poloniensis
323 C3GJS5 Bacillus thuringiensis serovar pondicheriensis BGSC 4BA1
324 A0A243MNF8 Bacillus thuringiensis serovar sinensis
325 C3DKN7 Bacillus thuringiensis serovar sotto str. T04001
326 A0A243DHS4 Bacillus thuringiensis serovar subtoxicus
327 A0A243HVV7 Bacillus thuringiensis serovar thailandensis
328 M1QNZ9 Bacillus thuringiensis serovar thuringiensis str. IS5056
329 C3C4S0 Bacillus thuringiensis serovar tochigiensis BGSC 4Y1
330 A0A243E9U6 Bacillus thuringiensis serovar toumanoffi
331 A0A243D1M8 Bacillus thuringiensis serovar vazensis
332 A0A243C4A6 Bacillus thuringiensis serovar yosoo
333 A0A2C9Z9R1 Bacillus thuringiensis serovar yunnanensis
334 A0A2C9ZDH6 Bacillus thuringiensis serovar zhaodongensis
335 A0REW9 Bacillus thuringiensis (strain Al Hakam)
336 A0A243I197 Bacillus thuringiensis subsp. aizawai
337 A0A243LK93 Bacillus thuringiensis subsp. dakota
338 A0A243DV89 Bacillus thuringiensis subsp. darmstadiensis
339 A0A243GJ97 Bacillus thuringiensis subsp. finitimus
340 F0PJA2 Bacillus thuringiensis subsp. finitimus (strain YBT-020)
341 A0A243K649 Bacillus thuringiensis subsp. higo
342 A0A0K0SAE3 Bacillus thuringiensis subsp. indiana
343 A0A160LB34 Bacillus thuringiensis subsp. israelensis
344 A0A243MJ97 Bacillus thuringiensis subsp. jegathesan
345 A0A243IJB1 Bacillus thuringiensis subsp. konkukian
346 Q6HHY0 Bacillus thuringiensis subsp. konkukian (strain 97-27)
347 A0A243EUG6 Bacillus thuringiensis subsp. kumamotoensis
348 A0A2C9YWK Bacillus thuringiensis subsp. kyushuensis

349 A0A243P328 *Bacillus thuringiensis* subsp. medellin
350 A0A0G4D072 *Bacillus thuringiensis* subsp. tolworthi
351 A0A0F6J5L2 *Bacillus thuringiensis* T01-328
352 A0A2B4TBY1 *Bacillus toyonensis*
353 A9VLT1 *Bacillus weihenstephanensis* (strain KBAB4)
354 A0A1A9PS17 *Bacillus wiedmannii*
355 A0A222FPD5 *Bacterioplanes sanyensis*
356 A0A1V5K7C0 bacterium ADurb.Bin478
357 A0A1V5JGZ1 bacterium ADurb.BinA028
358 A0A0A1P7J7 bacterium YEK0313
359 A0A180EZ91 *Bacteroidales* bacterium Barb4
360 A0A180FM12 *Bacteroidales* bacterium Barb6
361 A0A180FCF0 *Bacteroidales* bacterium Barb6XT
362 A0A0P0GQ77 *Bacteroides cellulosilyticus*
363 A0A120A2Y7 *Bacteroides stercoris*
364 A0A174H7V0 *Bacteroides uniformis*
365 A0A174JUS8 *Bacteroides xylanisolvens*
366 A0A1V6JEF7 *Bacteroidetes* bacterium ADurb.Bin008
367 A0A1V6IDS0 *Bacteroidetes* bacterium ADurb.Bin035
368 A0A1V6HGC bacterium ADurb.Bin041
369 A0A1V6CTM4 *Bacteroidetes* bacterium ADurb.Bin123
370 A0A1V6BDD1 *Bacteroidetes* bacterium ADurb.Bin141
371 A0A1V6BCZ2 *Bacteroidetes* bacterium ADurb.Bin145
372 A0A1V5P201 *Bacteroidetes* bacterium ADurb.Bin397
373 A0A1V5JE97 *Bacteroidetes* bacterium ADurb.BinA012
374 A0A1V5H839 *Bacteroidetes* bacterium ADurb.BinA174
375 A0A1B6Y5M2 *Balneola* sp. EhC07
376 A0A0N1B6M2 beta proteobacterium AAP99
377 A0A1V5QXL4 Betaproteobacteria bacterium ADurb.Bin341
378 D2Q8T9 *Bifidobacterium dentium* (strain ATCC 27534 / DSM 20436 / JCM 1195 / Bd1)
379 A0A173XH31 *Bifidobacterium pseudocatenulatum*
380 A0A087CCG7 *Bifidobacterium pullorum*
381 A0A087D044 *Bifidobacterium saeculare* DSM 6531 = LMG 14934
382 H6RLD2 *Blastococcus saxobsidens* (strain DD2)
383 A3ZNI7 *Blastopirellula marina* DSM 3645
384 A0A173YDV8 *Blautia obeum*
385 A0A173WUL5 *Blautia wexlerae*
386 A0A1U9VNG6 blood disease bacterium A2-HR MARDI
387 G2ZV67 blood disease bacterium R229
388 A0A146AU06 *Bordetella ansorpii*
389 A0A0H4WRC *Bordetella hinzii*
390 A0A140KUK9 *Bordetella pertussis*
391 A9IF86 *Bordetella petrii* (strain ATCC BAA-461 / DSM 12804 / CCUG 43448)
392 A0A1E3Z1Y9 *Bordetella* sp. SCN 67-23
393 A0A157SG71 *Bordetella trematum*

394 A0A257KPA7 Bradyrhizobiaceae bacterium PARB1
395 F7QJB4 Bradyrhizobiaceae bacterium SG-6C
396 A0A1I7JKF5 Bradyrhizobium arachidis
397 A0A1X3G9T0 Bradyrhizobium canariense
398 A0A2A6N274 Bradyrhizobium diazoefficiens
399 A0A0D1NEM! Bradyrhizobium elkanii
400 A0A1H4P7Y1 Bradyrhizobium erythrophlei
401 A0A1B1UH73 Bradyrhizobium icense
402 A0A0M9BGF! Bradyrhizobium japonicum
403 A0A023XFH2 Bradyrhizobium japonicum SEMIA 5079
404 A0A0R3M534 Bradyrhizobium jicamae
405 A0A1M7AIC6 Bradyrhizobium lablabi
406 A0A151G036 Bradyrhizobium liaoningense
407 K5C0A7 Bradyrhizobium lupini HPC(L)
408 A0A0R3DJI8 Bradyrhizobium manausense
409 A0A270RGP4 Bradyrhizobium ottawaense
410 A0A0R3BRG! Bradyrhizobium pachyrhizi
411 A0A1B9Z1Y9 Bradyrhizobium paxllaeri
412 A0A0R3MQN Bradyrhizobium retamae
413 A0A265PJ62 Bradyrhizobium sp. 2(2017)
414 A0A265QTQ1 Bradyrhizobium sp. 3(2017)
415 A0A1Q5R6Y1 Bradyrhizobium sp. AS23.2
416 A0A150U3Z1 Bradyrhizobium sp. AT1
417 A0A176YT67 Bradyrhizobium sp. BR 10245
418 A0A109K0E8 Bradyrhizobium sp. BR 10303
419 A0A1V5EMA! Bradyrhizobium sp. BR10280
420 A0A2A6NSU! Bradyrhizobium sp. C9
421 A0A0X8CI62 Bradyrhizobium sp. CCGE-LA001
422 A0A1I3SZ38 Bradyrhizobium sp. cf659
423 U1HQ00 Bradyrhizobium sp. DFCI-1
424 A0A0S6UGC! Bradyrhizobium sp. DOA9
425 A0A1I3NHY1 Bradyrhizobium sp. Gha
426 A0A1I5UD35 Bradyrhizobium sp. Ghvi
427 A0A0Q6AS12 Bradyrhizobium sp. Leaf396
428 A0A1B9Z0V9 Bradyrhizobium sp. LMTR 3
429 A0A0D7PXW! Bradyrhizobium sp. LTSP849
430 A0A0D7PLA1 Bradyrhizobium sp. LTSP885
431 A0A0D7Q971 Bradyrhizobium sp. LTSPM299
432 A0A1Q5RAV! Bradyrhizobium sp. NAS80.1
433 A0A1Q5RKK! Bradyrhizobium sp. NAS96.2
434 A0A1H8I081 Bradyrhizobium sp. OK095
435 A0A1Y6KQM! Bradyrhizobium sp. ORS 285
436 A0A1G7AFE4 Bradyrhizobium sp. R5
437 A0A1G7WSM Bradyrhizobium sp. Rc2d
438 A0A1I4YQE0 Bradyrhizobium sp. Rc3b

439 I0G9L1 Bradyrhizobium sp. S23321
440 H0TB50 Bradyrhizobium sp. STM 3809
441 H0TIR1 Bradyrhizobium sp. STM 3843
442 A5E828 Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)
443 A4YJA9 Bradyrhizobium sp. (strain ORS 278)
444 H0SSJ3 Bradyrhizobium sp. (strain ORS 375)
445 A0A1C2ESQ5 Bradyrhizobium sp. UASWS1016
446 A0A1R1QGU1 Bradyrhizobium sp. UFLA 03-321
447 A0A2A2VDA7 Bradyrhizobium sp. UFLA03-84
448 A0A2A6PE34 Bradyrhizobium sp. Y36
449 A0A0R3KPP5 Bradyrhizobium valentinum
450 A0A0R3CAP5 Bradyrhizobium yuanmingense
451 A0A0J6BF55 Brevibacillus brevis (Bacillus brevis)
452 A0A0H0SQH1 Brevibacillus formosus
453 A0A0F6XZR3 Brevibacillus laterosporus (Bacillus laterosporus)
454 A0A075QW51 Brevibacillus laterosporus LMG 15441
455 A0A0K9YP49 Brevibacillus reuszeri
456 A0A177XKZ1 Brevibacillus sp. SKDU10
457 A0A0B9A9M5 Brevibacterium linens
458 A0A1S6ER20 Brevundimonas sp. LM2
459 A0A258HLW4 Brevundimonas subvibrioides
460 D7H3I5 Brucella abortus bv. 5 str. B3196
461 C4IUP8 Brucella abortus str. 2308 A
462 Q2YIN0 Brucella abortus (strain 2308)
463 A0A292DRS8 Brucella canis
464 A9MCQ6 Brucella canis (strain ATCC 23365 / NCTC 10854)
465 C0GAV5 Brucella ceti str. Cudo
466 A0A1Z1ZS86 Brucella melitensis
467 Q8YD58 Brucella melitensis biotype 1 (strain 16M / ATCC 23456 / NCTC 10094)
468 C0RMB9 Brucella melitensis biotype 2 (strain ATCC 23457)
469 C7LJ97 Brucella microti (strain CCM 4915)
470 C9V4N3 Brucella neotomae 5K33
471 A0A0E1WUX Brucella pinnipedialis M292/94/1
472 A0A1P8S6Z0 Brucella sp. 09RB8910
473 A0A1M4LF47 Brucella sp. 10RB9215
474 A0A1G4F5J7 Brucella sp. 141012304
475 D1D0I3 Brucella sp. 83/13
476 E2PLG0 Brucella sp. BO2
477 D6LPC0 Brucella sp. NVSL 07-0026
478 D0P7U4 Brucella suis bv. 5 str. 513
479 A0A0U5MSK1 Brucella vulpis
480 A0A2A7TX17 Burkholderia ambifaria
481 Q0B2Z2 Burkholderia ambifaria (strain ATCC BAA-244 / AMMD) (Burkholderia cepacia (strain AMMD))
482 A0A118PBH1 Burkholderia anthina
483 A0A1V2WW2 Burkholderia cenocepacia

484 B1K5X2 Burkholderia cenocepacia (strain MC0-3)
485 A0A0U3TXS6 Burkholderia cepacia JBK9
486 A0A1Z3YPG9 Burkholderia cepacia (Pseudomonas cepacia)
487 A0A1E3FJI4 Burkholderia contaminans
488 A0A1B4BNP4 Burkholderia diffusa
489 A0A104JVG7 Burkholderia gladioli (Pseudomonas marginata) (Phytomonas marginata)
490 A0A0B6S3A2 Burkholderia glumae PG1
491 A0A0J5YLS6 Burkholderia lata (strain ATCC 17760 / DSM 23089 / LMG 22485 / NCIMB 9086 / R18194 / 383)
492 A0A1B4NNG6 Burkholderia latens
493 C4B1X1 Burkholderia mallei GB8 horse 4
494 A0A153L7U9 Burkholderia mallei (Pseudomonas mallei)
495 A0A0H2WG4 Burkholderia mallei (strain ATCC 23344)
496 A2S6W5 Burkholderia mallei (strain NCTC 10229)
497 A0A1B4HPS8 Burkholderia metallica
498 A0A228DPT2 Burkholderia multivorans
499 A0A0H3KPQ6 Burkholderia multivorans (strain ATCC 17616 / 249)
500 A0A1B4AQ54 Burkholderia oklahomensis C6786
501 A0A1B4AEU8 Burkholderia oklahomensis EO147
502 A0A0P0L247 Burkholderia plantarii
503 A0A1X4K397 Burkholderia pseudomallei (Pseudomonas pseudomallei)
504 A0A0H3HQ04 Burkholderia pseudomallei (strain 1026b)
505 Q3JY82 Burkholderia pseudomallei (strain 1710b)
506 A0A132E8W6 Burkholderia pseudomultivorans
507 A0A1X1PCN6 Burkholderia puraquae
508 A0A104MRE1 Burkholderia pyrrocinia (Pseudomonas pyrrocinia)
509 A0A118FQP0 Burkholderia seminalis
510 A0A103E805 Burkholderia singularis
511 A0A1E4YV47 Burkholderia sp. A2
512 A0A1X0MLR7 Burkholderia sp. A27
513 A0A103B5M3 Burkholderia sp. ABCPW 11
514 A0A132CC19 Burkholderia sp. ABCPW 14
515 A0A221AGH6 Burkholderia sp. AD24
516 A0A228HD74 Burkholderia sp. AU15512
517 A0A228I1E1 Burkholderia sp. AU16741
518 A0A228I2B6 Burkholderia sp. AU17325
519 A0A228KBI4 Burkholderia sp. AU27893
520 A0A228PZY0 Burkholderia sp. AU6039
521 A0A118E0C2 Burkholderia sp. BDU6
522 A0A118E585 Burkholderia sp. BDU8
523 A0A211YXS8 Burkholderia sp. Bk
524 A0A1B4R1K5 Burkholderia sp. Bp7605
525 A0A1J9PLS5 Burkholderia sp. DNA89
526 A0A124QG08 Burkholderia sp. FL-7-2-10-S1-D7
527 A0A1N6KWV Burkholderia sp. GAS332
528 A0A0M5LLG7 Burkholderia sp. HB1

529 A0A228RJT8 Burkholderia sp. HI2714
530 A0A228QM82 Burkholderia sp. HI2761
531 A0A2A4C586 Burkholderia sp. IDO3
532 A0A132CUT8 Burkholderia sp. LA-2-3-30-S1-D2
533 A0A0Q5NY68 Burkholderia sp. Leaf177
534 A0A0J6M412 Burkholderia sp. LK4
535 A0A084DN29 Burkholderia sp. MSh2
536 A0A1Y9SLT7 Burkholderia sp. MSMB0266
537 A0A1B4M7W Burkholderia sp. MSMB0856
538 A0A104LGU7 Burkholderia sp. MSMB1552
539 A0A118LL25 Burkholderia sp. MSMB1835
540 A0A1B4S4G4 Burkholderia sp. MSMB617WGS
541 A0A1B4DLK8 Burkholderia sp. NRF60-BP8
542 A0A160FG61 Burkholderia sp. OLGA172
543 A0A0X8L436 Burkholderia sp. PAMC 26561
544 A0A102ET31 Burkholderia sp. RF2-non_BP3
545 A0A1Q8J3Z7 Burkholderia sp. SRS-W-2-2016
546 A0A0N0IFV3 Burkholderia sp. ST111
547 A0A103EHI2 Burkholderia sp. TSV86
548 A0A0D8DX57 Burkholderia sp. USM B20
549 A0A1H5A9M8 Burkholderia sp. WP9
550 A0A1D7ZDR2 Burkholderia stabilis
551 A0A106NXD4 Burkholderia stagnalis
552 A0A105VIA1 Burkholderia territorii
553 A0A1B4K3N0 Burkholderia thailandensis
554 Q2STU4 Burkholderia thailandensis (strain ATCC 700388 / DSM 13276 / CIP 106301 / E264)
555 A0A108CIY8 Burkholderia ubonensis
556 A0A118EHR4 Burkholderia vietnamiensis
557 A0A0F0ESZ6 Burkholderiaceae bacterium 16
558 A0A0F0DFE5 Burkholderiaceae bacterium 26
559 A0A1Q3SCP3 Burkholderiales bacterium 64-34
560 A0A1Q3Y230 Burkholderiales bacterium 67-32
561 A0A257CSA2 Burkholderiales bacterium PBB3
562 A0A257LT33 Burkholderiales bacterium PBB5
563 A0A1F4ES80 Burkholderiales bacterium RIFCSPHIGHO2_01_FULL_64_960
564 A0A1F4IVC3 Burkholderiales bacterium RIFCSPHIGHO2_12_FULL_65_48
565 A0A1F4JN14 Burkholderiales bacterium RIFCSLOWO2_02_FULL_57_36
566 A0A158CKL0 Caballeronia hypogeia
567 A0A0J1FT51 Caballeronia mineralivorans PML1(12)
568 A0A242M8D8 Caballeronia sordidicola
569 A0A1V5GNM candidate division BRC1 bacterium ADurb.BinA364
570 A0A011PTN3 Candidatus Accumulibacter sp. BA-92
571 A0A011QU85 Candidatus Accumulibacter sp. BA-94
572 A0A084Y3V0 Candidatus Accumulibacter sp. SK-01
573 A0A080MKD1 Candidatus Accumulibacter sp. SK-02

574 A0A1V6B2J0 Candidatus Aminicenantes bacterium ADurb.Bin147
575 V9TR00 Candidatus Endolissoclinum faulkneri L5
576 A0A2A5EE89 Candidatus Hydrogenedentes bacterium
577 A0A1V6EEM6 Candidatus Hydrogenedentes bacterium ADurb.Bin101
578 X5MEY0 Candidatus Phaeomarinobacter ectocarp
579 A0A1Z8UQW Candidatus Puniceispirillum sp. TMED52
580 A0A0J9E550 Candidatus Rhodobacter lobularis
581 A0A1Q9P628 Candidatus Thorarchaeota archaeon AB_25
582 A0A0B7IGW5 Capnocytophaga canimorsus
583 F9YPG4 Capnocytophaga canimorsus (strain 5)
584 A0A0B7INU9 Capnocytophaga canis
585 A0A0B7H2X8 Capnocytophaga cynodegmi
586 A0A0R2JAN1 Carnobacterium maltaromaticum (Carnobacterium piscicola)
587 A0A0R2IS35 Carnobacterium maltaromaticum DSM 20342
588 K8EDK6 Carnobacterium maltaromaticum LMA28
589 A0A0U3MEA6 Carnobacterium sp. CP1
590 A0A0M2NMF7 Catabacter hongkongensis
591 A0A174HD35 Catenibacterium mitsuokai
592 Q9A8N0 Caulobacter crescentus (strain ATCC 19089 / CB15)
593 A0A0H3C844 Caulobacter crescentus (strain NA1000 / CB15N)
594 A0A0P0NZJ7 Caulobacter henricii
595 A0A257X495 Caulobacter sp. 12-67-6
596 A0A258IZM8 Caulobacter sp. 32-67-35
597 A0A258KDE6 Caulobacter sp. 35-67-4
598 A0A259ID67 Caulobacter sp. 39-67-4
599 A0A0Q7TB71 Caulobacter sp. Root1455
600 A0A0Q7WNG Caulobacter sp. Root1472
601 A0A0Q6TSV5 Caulobacter sp. Root342
602 A0A0Q6UP28 Caulobacter sp. Root343
603 A0A0Q8FWR Caulobacter sp. Root655
604 A0A0Q8GDX1 Caulobacter sp. Root656
605 A0A290MJ25 Caulobacter vibrioides (Caulobacter crescentus)
606 A0A1B1AF90 Caulobacteraceae bacterium OTSz_A_272
607 A0A258FJV6 Caulobacterales bacterium 32-69-10
608 A0A1Q3JML5 Caulobacterales bacterium 68-7
609 K1LUT6 Cecembia lonarensis LW9
610 A0A291GCNC Celeribacter ethanolicus
611 A0A1Z9I936 Cellvibrionales bacterium TMED157
612 A0A1Z8MX29 Cellvibrionales bacterium TMED21
613 A0A1Z8RZJ5 Cellvibrionales bacterium TMED47
614 A0A1Z8TWE7 Cellvibrionales bacterium TMED48
615 A0A1Z8Z7N2 Cellvibrionales bacterium TMED79
616 M7NIZ7 Cesiribacter andamanensis AMV16
617 A0A0K2DK09 Chelatococcus sp. CO-6
618 C7PKB0 Chitinophaga pinensis (strain ATCC 43595 / DSM 2588 / NCIB 11800 / UQM 2034)

619 A0A1X6W3T6 Chlamydia abortus (Chlamydophila abortus)
620 A0A0E9FL18 Chlamydia trachomatis
621 A0A136K2T4 Chlorobi bacterium OLB4
622 A0A136PFT0 Chlorobi bacterium OLB7
623 A0A1V5WGQ Chloroflexi bacterium ADurb.Bin222
624 A0A1V5RV19 Chloroflexi bacterium ADurb.Bin325
625 A0A1V5QH62 Chloroflexi bacterium ADurb.Bin344
626 A0A1V5PDI7 Chloroflexi bacterium ADurb.Bin360
627 A0A1Q8TA39 Chromohalobacter japonicus
628 A0A083WRM Chryseobacterium antarcticum
629 A0A1A7KL28 Chryseobacterium sp. MOF25P
630 A0A077KPP6 Chryseobacterium sp. StRB126
631 A0A239XLG5 Chryseobacterium taklimakanense
632 D0DE53 Citreicella sp. SE45
633 A0A1Y3FFF3 Clavibacter michiganensis subsp. michiganensis
634 A0A143Y291 Clostridiales bacterium CHKC1001
635 A0A1R4UEF4 Clostridioides difficile (Peptoclostridium difficile)
636 A0A0C1U0F9 Clostridium argentinense CDC 2741
637 A0A174SE58 [Clostridium] clostridioforme
638 A0A174Z2S0 Clostridium innocuum
639 A0A168MLP5 Clostridium ljungdahlii
640 A0A2A7ME47 Clostridium neonatale
641 A0A1V4IZP8 Clostridium oryzae
642 A0A1D9N0N6 Clostridium pasteurianum
643 A0A1S8TH57 Clostridium puniceum
644 A0A1A6B1J1 Clostridium ragsdalei P11
645 A0A0U5PQ27 Clostridium sp. C105KSO14
646 A0A0U5HF17 Clostridium sp. C105KSO15
647 A0A0F0CE38 Clostridium sp. FS41
648 A0A1F2CYH6 Clostridium sp. HMSC19B10
649 W6N4Q2 Clostridium tyrobutyricum DIVETGP
650 A0A0J8QLY7 Coccidioides immitis RMSCC 3703
651 A0A0J6IKJ4 Coccidioides posadasii RMSCC 3488
652 E9D435 Coccidioides posadasii (strain RMSCC 757 / Silveira) (Valley fever fungus)
653 G0ACC8 Collimonas fungivorans (strain Ter331)
654 A0A257EI85 Comamonadaceae bacterium PBBC1
655 A0A257G061 Comamonadaceae bacterium PBBC2
656 A0A0M2DUU1 Comamonas sp. E6
657 A0A2A7UQ41 Comamonas terrigena
658 A0A096HCR5 Comamonas testosteroni (Pseudomonas testosteroni)
659 B7WS76 Comamonas testosteroni (strain DSM 14576 / KF-1)
660 A0A076PZE3 Comamonas testosteroni TK102
661 A0A176TST0 Comamonas thiooxydans
662 W7E892 Commensalibacter sp. MX01
663 A0A0N9NR37 Confluentimicrobium sp. EMB200-NS6

664 A0A174DWD: Coprococcus comes
665 H8MKX6 Corallococcus coralloides (strain ATCC 25202 / DSM 2259 / NBRC 100086 / M2) (Myxococcus coralloides)
666 A0A1R3FXV9 Corchorus capsularis (Jute)
667 A0A0B6TVK5 Corynebacterium marinum DSM 44953
668 A0A1Q7LPF6 Crenarchaeota archaeon 13_1_40CM_3_53_5
669 A0A1B1BHDC Cryobacterium arcticum
670 H1SAR3 Cupriavidus basilensis OR16
671 A0A0M4KIS7 Cupriavidus gilardii CR3
672 A0A132HH88 Cupriavidus metallidurans
673 Q1LCI4 Cupriavidus metallidurans (strain ATCC 43123 / DSM 2839 / NBRC 102507 / CH34) (Ralstonia metallidurans)
674 A0A142JSV0 Cupriavidus nantongensis
675 F8GPZ8 Cupriavidus necator (strain ATCC 43291 / DSM 13513 / N-1) (Ralstonia eutropha)
676 Q46SW2 Cupriavidus necator (strain JMP 134 / LMG 1197) (Ralstonia eutropha (strain JMP 134))
677 L2EK81 Cupriavidus sp. HMR-1
678 V2J8V3 Cupriavidus sp. HPC(L)
679 A0A1U9UZA5 Cupriavidus sp. NH9
680 A0A069I7N4 Cupriavidus sp. SK-3
681 A0A022FUI7 Cupriavidus sp. SK-4
682 A0A1D9H6U3 Cupriavidus sp. USMAA2-4
683 B3RAT4 Cupriavidus taiwanensis (strain DSM 17343 / BCRC 17206 / CIP 107171 / LMG 19424 / R1) (Ralstonia taiwanensis (strain LMG 19424))
684 A0A1Y0N366 Curvibacter sp. AEP1-3
685 A0A2C6LKM4 Cutibacterium acnes (Propionibacterium acnes)
686 A0A2B7I057 Cutibacterium acnes subsp. acnes
687 A0A2B7HL54 Cutibacterium acnes subsp. defendens
688 A0A1B3Q1Y9 Cutibacterium avidum
689 G4CUD1 Cutibacterium avidum ATCC 25577
690 A0A239WW2 Cutibacterium granulosum
691 U1GEY5 Cutibacterium granulosum DSM 20700
692 A0A159Z148 Defluviimonas alba
693 A0A059IY1 Defluviimonas sp. 20V17
694 A0A1U7NW2 Deinococcus marmoris
695 Q9RXS1 Deinococcus radiodurans (strain ATCC 13939 / DSM 20539 / JCM 16871 / LMG 4051 / NBRC 15346 / NCIMB 9279 / R1 / VKM B-1422)
696 A0A291JE33 Delftia acidovorans (Pseudomonas acidovorans) (Comamonas acidovorans)
697 A9BN79 Delftia acidovorans (strain DSM 14801 / SPH-1)
698 A0A081J655 Delftia sp. 670
699 A0A1C7L5Y4 Delftia sp. JD2
700 A0A210W8E3 Delftia sp. K82
701 A0A031IT02 Delftia sp. RIT313
702 F6AL57 Delftia sp. (strain Cs1-4)
703 A0A1D8J8Y0 Delftia tsuruhatensis
704 A0A1V6KGF2 Deltaproteobacteria bacterium ADurb.Bin002
705 A0A1V6IQ42 Deltaproteobacteria bacterium ADurb.Bin022
706 A0A1V6G442 Deltaproteobacteria bacterium ADurb.Bin072
707 A0A1V6AWT: Deltaproteobacteria bacterium ADurb.Bin151
708 A0A1V5J955 Deltaproteobacteria bacterium ADurb.BinA014

709 A0A1V5HE73 Deltaproteobacteria bacterium ADurb.BinA179
710 A0A0S7YV12 Deltaproteobacteria bacterium SG8_13
711 A0A239V818 Dermatophilus congolensis
712 A0A1K1LEH2 Desulfovibrio piger
713 A0A0F5FSK2 Devosia geojensis
714 A0A0F5LW9C Devosia limi DSM 17137
715 A0A0F5L6I5 Devosia soli
716 A0A1M3KS5C Devosia sp. 66-22
717 A0A0Q4P836 Devosia sp. Leaf64
718 A0A178HXF3 Devosia sp. S37
719 A0A177LBJ1 Dietzia cinnamea
720 E6J575 Dietzia cinnamea P4
721 A0A199ATT5 Dietzia sp. 111N12-1
722 A0A1F1K8X7 Dietzia sp. HMSC21D01
723 A0A167H4E9 Dokdonella koreensis DS-123
724 A0A174A4R5 Dorea longicatena
725 A0A1W4UIN8 Drosophila ficusphila (Fruit fly)
726 A0A1E7WFA1 Duganella sp. HH101
727 A0A1E7V5C4 Duganella sp. HH105
728 A0A0Q5D7A2 Duganella sp. Leaf126
729 A0A0Q4P4G1 Duganella sp. Leaf61
730 A0A1E3AQH5 Eisenbergiella tayi
731 A0A0D5BQQ1 Elizabethkingia miricola (Chryseobacterium miricola)
732 A0A1X7AI28 Endozoicomonas sp. S-B4-1U
733 W8IBY1 Ensifer adhaerens OV14
734 A0A0L8BM63 Ensifer adhaerens (Sinorhizobium morelense)
735 A0A178XU98 Ensifer glycinis
736 A0A1C0SKE3 Ensifer sp. LC14
737 A0A1C0SQ09 Ensifer sp. LC54
738 A0A1S1T112 Ensifer sp. LCM 4579
739 A0A0Q7QGS1 Ensifer sp. Root142
740 A0A0Q9E8V7 Ensifer sp. Root278
741 A0A1E3V7D0 Ensifer sp. YIC4027
742 A0A0J0HGF8 Enterobacter cloacae
743 G8LI19 Enterobacter cloacae EcWSU1
744 A0A181D6J5 Enterobacter kobei
745 R3L1W8 Enterococcus faecalis ATCC 6055
746 A0A0M2AE38 Enterococcus faecalis EnGen0302
747 A0A0M2AW7 Enterococcus faecalis EnGen0354
748 R3JGQ0 Enterococcus faecalis EnGen0359
749 A0A0M2CLC1 Enterococcus faecalis EnGen0426
750 V7ZSW4 Enterococcus faecalis PF3
751 Q836S5 Enterococcus faecalis (strain ATCC 700802 / V583)
752 A0A246MX18 Enterococcus faecalis (Streptococcus faecalis)
753 C0X1J8 Enterococcus faecalis TX0104

754 A0A1L8TTG8 Enterococcus gallinarum
755 A0A219BFR0 Enterococcus hirae 67-03-C5
756 J0NUD5 Enterococcus sp. C1
757 T0UCQ9 Enterococcus sp. HSIEG1
758 A0A135I9A2 Enterovibrio corali
759 A0A1C3EPQ5 Enterovibrio pacificus
760 A0A165RC74 Erythrobacter sp. HI0063
761 A0A074M2K7 Erythrobacter sp. JL475
762 A0A0P6X3L4 Erythrobacter sp. SG61-1L
763 A0A1D3U0X1 Escherichia coli
764 A0A174ZK95 [Eubacterium] eligens
765 A0A173ZD16 Faecalicatena contorta
766 A0A0D8FW25 Ferrimicrobium acidiphilum DSM 19497
767 I0K217 Fibrella aestuarina BUZ 2
768 I2GH59 Fibrisoma limi BUZ 3
769 A0A160ILQ3 Fictibacillus phosphorivorans
770 A0A1V6AIT8 Firmicutes bacterium ADurb.Bin146
771 A0A1V5Y736 Firmicutes bacterium ADurb.Bin182
772 A0A1V5U0S1 Firmicutes bacterium ADurb.Bin248
773 A0A1V5TRM2 Firmicutes bacterium ADurb.Bin262
774 R7N860 Firmicutes bacterium CAG:95
775 A3J172 Flavobacteria bacterium BAL38
776 A0A1J0LM14 Flavobacteriaceae bacterium UJ101
777 A0A0D0EX75 Flavobacterium hibernum
778 A0A085ZZ56 Flavobacterium hydatis (Cytophaga aquatilis)
779 A0A226I074 Flavobacterium oncorhynchi
780 A0A226IS47 Flavobacterium plurextorum
781 Q79F77 Flavobacterium sp.
782 A0A1Q3ZT62 Flavobacterium sp. 38-13
783 A0A244D1Q8 Flavobacterium sp. AJR
784 Q57326 Flavobacterium sp. KI723T1
785 A0A0C1DR67 Flavobacterium sp. KMS
786 A0A0Q5V6GC Flavobacterium sp. Leaf359
787 P13398 Flavobacterium sp. (strain K172)
788 A0A1V2AUBC [Flexibacter] sp. ATCC 35208
789 A0A0Q9PIG2 Frateuria sp. Soil773
790 A0A0M9EWU Fusarium langsethiae
791 N4TTL8 Fusarium oxysporum f. sp. cubense (strain race 1) (Panama disease fungus)
792 N1RYA1 Fusarium oxysporum f. sp. cubense (strain race 4) (Panama disease fungus)
793 A0A174KSI3 Fusicatenibacter saccharivorans
794 A0A174XLK7 Fusicatenibacter sp. 2789STDY5834925
795 Q1YU97 gamma proteobacterium HTCC2207
796 W2UH15 Gammaproteobacteria bacterium MOLA455
797 A0A1G0K8X0 Gammaproteobacteria bacterium RIFCSPLOWO2_12_FULL_52_10
798 A0A0S8A877 Gammaproteobacteria bacterium SG8_11

799 A0A0S8DTT8 Gammaproteobacteria bacterium SG8_47
800 A0A1Z9CP29 Gammaproteobacteria bacterium TMED112
801 A0A1Z9NL67 Gammaproteobacteria bacterium TMED219
802 A0A1Z9PFW(Gammaproteobacteria bacterium TMED225
803 A0A1Z9PCI7 Gammaproteobacteria bacterium TMED226
804 A0A1Z9R0K5 Gammaproteobacteria bacterium TMED242
805 A0A1Z9TMI5 Gammaproteobacteria bacterium TMED278
806 A0A1Z8Q2H1 Gammaproteobacteria bacterium TMED36
807 A0A061NHNC Geomicrobium sp. JCM 19037
808 A0A061NYJ9 Geomicrobium sp. JCM 19038
809 A0A061PHA8 Geomicrobium sp. JCM 19039
810 A0A061NOV9 Geomicrobium sp. JCM 19055
811 F3S710 Gluconacetobacter sp. SXCC-1
812 G7GNR0 Gordonia amarae NBRC 15530
813 M3VJE8 Gordonia paraffinivorans NBRC 108238
814 A0A222TFI4 Gordonia rubripertincta (Rhodococcus corallinus)
815 C8NIE3 Granulicatella adiacens ATCC 49175
816 A0A128EWV! Grimontia celer
817 A0A212AFU6 Haematobacter genomosp. 1
818 A0A086Y029 Haematobacter massiliensis
819 A0A212AN47 Haematobacter missouriensis
820 M5E0U2 Halanaerobium saccharolyticum subsp. saccharolyticum DSM 6643
821 A0A2A5WR2 Halieaceae bacterium MED-G27
822 A0A059NVE7 Halobacillus karajensis
823 A0A126V1M6 Halocynthiibacter arcticus
824 A0A1J0VHN5 Halomonas aestuarii
825 A0A265DV80 Halomonas boliviensis LC1
826 A0A246S039 Halomonas campaniensis
827 A0A243T4U5 Halomonas desiderata SP1
828 A0A291JTQ4 Halomonas hydrothermalis
829 A0A098RAP9 Halomonas salina
830 A0A1P8R8J2 Halomonas sp. 1513
831 A0A101D421 Halomonas sp. 54_146
832 T2LAH3 Halomonas sp. A3H3
833 A0A1A9R1GC Halomonas sp. ALS9
834 A0A1A0FKY5 Halomonas sp. G11
835 A0A2A3XJE5 Halomonas sp. JB37
836 A0A0C3DNU(Halomonas sp. KHS3
837 A0A2A4HNA3 Halomonas sp. MBT G8648
838 A0A220RGN2 Halomonas sp. N3-2A
839 A0A1J8PCU6 Halomonas sp. QHL1
840 A0A1U9RGR Halomonas sp. 'Soap Lake #7'
841 A0A2A2EQK1 Halomonas sp. WRN001
842 A0A172YJR4 Halotalea alkalilenta
843 A0A2A2F7D0 Halovibrio sp. YL5-2

844 A0A2C9D0T4 Hartmannibacter diazotrophicus
845 A2RQ43 Herbaspirillum seropedicae
846 D8J0D7 Herbaspirillum seropedicae (strain SmR1)
847 A4G8V7 Herminiimonas arsenicoxydans
848 A9B2R3 Herpetosiphon aurantiacus (strain ATCC 23779 / DSM 785)
849 A0A0P6YIJ7 Herpetosiphon geysericola
850 C6XP16 Hirschia baltica (strain ATCC 49814 / DSM 5838 / IFAM 1418)
851 A0A1C1YZ89 Hoeflea olei
852 F6EJ06 Hoyosella subflava (strain DSM 45089 / JCM 17490 / NBRC 109087 / DQS3-9A1) (Amycolicoccus subflavus)
853 K1TDC0 human gut metagenome
854 A0A174B3K2 Hungatella hathewayi
855 A0A085WHH1 Hyalangiium minutum
856 A0A0W8FM01 hydrocarbon metagenome
857 A0A1L1PDA4 Hydrogenophaga intermedia
858 A0A1Q3QEU1 Hydrogenophaga sp. 70-12
859 A0A1C9V9S3 Hydrogenophaga sp. PBC
860 Q31IT9 Hydrogenovibrio crunogenus (strain XCL-2) (Thiomicrospira crunogena)
861 A0A1L7SBN4 hydrothermal vent metagenome
862 A0A1E4BXY7 Hyphomicrobium sp. SCN 65-11
863 A0A0F2QIR6 Hyphomonadaceae bacterium BRH_c29
864 A0A069E6Q0 Hyphomonas adhaerens MHS-3
865 A0A059FPN2 Hyphomonas hirschiana VP5
866 A0A059FF25 Hyphomonas jannaschiana VP2
867 A0A059FE75 Hyphomonas johnsonii MHS-2
868 Q0C2K3 Hyphomonas neptunium (strain ATCC 15444)
869 A0A059G865 Hyphomonas oceanitis SCH89
870 A0A062VLC9 Hyphomonas polymorpha PS728
871 A0A259LK04 Hyphomonas sp. 34-62-18
872 A0A0F2RDK4 Hyphomonas sp. BRH_c22
873 A0A1Z9X064 Hyphomonas sp. TMED31
874 A0A0K8NTA2 Ideonella sakaiensis (strain 201-F6)
875 A0A211ZN00 Inquilinus limosus
876 A0A0A0DB64 Inquilinus limosus MP06
877 A0A161IJZ3 Isoptericola dokdonensis DS-3
878 A0A0D1EBH5 Jannaschia aquimarina
879 A0A0M6YFF1 Jannaschia donghaensis
880 A0A0M6XQ32 Jannaschia rubra
881 A0A031GX16 Janthinobacterium lividum
882 A0A1I9XVN2 Janthinobacterium sp. 1_2014MBL_MicDiv
883 A0A0S4NRP2 Janthinobacterium sp. CG23_2
884 A0A1E7WAM Janthinobacterium sp. HH106
885 A0A0M2WHD Janthinobacterium sp. KBS0711
886 A0A1S6F9H1 Janthinobacterium sp. LM6
887 A6T266 Janthinobacterium sp. (strain Marseille) (Minibacterium massiliensis)
888 A0A290WYZ5 Janthinobacterium svalbardensis

889 A0A1S6IQB3 Jeotgalibaca dankookensis
890 A0A0M0KXQ Jeotgalibacillus marinus
891 A0A1H5TFQ Jhaorihella thermophila
892 F9Y4Y6 Ketogulonicigenium vulgare (strain WSH-001)
893 K6VJY8 Kineosphaera limosa NBRC 100340
894 A0A1E7N9G1 Kitasatospora aureofaciens (Streptomyces aureofaciens)
895 A0A145MIT9 Klebsiella aerogenes (Enterobacter aerogenes)
896 A0A181XAP6 Klebsiella oxytoca
897 A0A1W1KFI7 Klebsiella pneumoniae
898 A0A0E1CG0 Klebsiella pneumoniae 30660/NJST258_1
899 W8UHE6 Klebsiella pneumoniae 30684/NJST258_2
900 Q1EPR4 Kocuria sp. KY2
901 A0A0M0EIV8 Komagataeibacter europaeus (Gluconacetobacter europaeus)
902 A0A181C565 Komagataeibacter rhaeticus
903 A0A070A9B3 Komagataeibacter rhaeticus AF1
904 W5Y864 Komagataeibacter xylinus E25
905 D6U8H7 Ktedonobacter racemifer DSM 44963
906 A0A240UKM Kushneria marisflavi
907 W7SGA0 Kutzneria sp. 744
908 A0A0M6Y0X3 Labrenzia aggregata
909 A0A0M6Z4Q2 Labrenzia alba
910 A0A174LCT8 Lachnospira pectinoschiza
911 A0A0R1LS40 Lactobacillus acidifarinae DSM 19394
912 A0A0R2KCN Lactobacillus acidipiscis
913 M5ADU8 Lactobacillus brevis KB290
914 C2D288 Lactobacillus brevis subsp. gravesensis ATCC 27305
915 C0WQT6 Lactobacillus buchneri ATCC 11577
916 J9W3Z9 Lactobacillus buchneri CD034
917 K6QIN7 Lactobacillus casei 21/1
918 K6QA34 Lactobacillus casei 32G
919 A0A0E2BPL3 Lactobacillus casei A2-362
920 S4ZS02 Lactobacillus casei LOCK919
921 K6QYG7 Lactobacillus casei M36
922 A0A125UAK7 Lactobacillus casei (strain BD-II)
923 K6S6R9 Lactobacillus casei UCD174
924 K6RGF8 Lactobacillus casei UW4
925 K0NAA2 Lactobacillus casei W56
926 A0A0R1SRI6 Lactobacillus diolivorans DSM 14421
927 A0A0R1W69 Lactobacillus farraginis DSM 18382 = JCM 14108
928 A0A158SLK4 Lactobacillus fermentum
929 C0WXY8 Lactobacillus fermentum ATCC 14931
930 D8II88 Lactobacillus fermentum (strain CECT 5716)
931 A0A0R1P1Q Lactobacillus frumenti DSM 13145
932 C0XGF7 Lactobacillus hilgardii DSM 20176 = ATCC 8290
933 C8PAP3 Lactobacillus iners DSM 13335

934 A0A0R1P766 *Lactobacillus mucosae* DSM 13345
935 A0A0R1KMJC *Lactobacillus nodensis* DSM 19682 = JCM 14932 = NBRC 107160
936 A0A0R1RGC *Lactobacillus oligofermentans* DSM 15707 = LMG 22743
937 A0A081BKN6 *Lactobacillus oryzae* JCM 18671
938 S4PQ60 *Lactobacillus otakiensis* DSM 19908 = JCM 15040
939 A0A1X1FFMC *Lactobacillus parabuchneri*
940 A0A1V0Q8X5 *Lactobacillus paracasei*
941 C2FBN4 *Lactobacillus paracasei* subsp. *paracasei* ATCC 25302
942 S2TJL4 *Lactobacillus paracasei* subsp. *paracasei* Lpp49
943 K8QAN6 *Lactobacillus rhamnosus* LRHMDP2
944 R6S9Q8 *Lactobacillus ruminis* CAG:367
945 F7R1I8 *Lactobacillus ruminis* SPM0211
946 A0A0R1R0G7 *Lactobacillus spicheri* DSM 15429
947 A0A1E7X8H4 *Lactobacillus sunkii*
948 A0A0R1L4P7 *Lactobacillus sunkii* DSM 19904
949 A0A0R1IY18 *Lactobacillus tuceti* DSM 20183
950 C2ESU0 *Lactobacillus vaginalis* DSM 5837 = ATCC 49540
951 A0A1Y6JZ04 *Lactobacillus zymae*
952 A0A0R1MN18 *Lactobacillus zymae* DSM 19395
953 A0A173M330 *Lactococcus garvieae*
954 K2PPG6 *Lactococcus garvieae* DCC43
955 G8P4E4 *Lactococcus lactis* subsp. *cremoris* A76
956 A0A161TZC1 *Lactococcus lactis* subsp. *cremoris* (*Streptococcus cremoris*)
957 U6ETZ5 *Lactococcus lactis* subsp. *lactis* Dephy 1
958 A0A0B8QXB1 *Lactococcus lactis* subsp. *lactis* (*Streptococcus lactis*)
959 A0A1C3SYZ1 *Lactococcus piscium*
960 A0A0D6DZS4 *Lactococcus piscium* MKFS47
961 I7JH02 *Lactococcus raffinolactis* 4877
962 A0A139MG44 *Lactococcus* sp. DD01
963 A0A1D8B005 *Lacunisphaera limnophila*
964 A3XGY6 *Leeuwenhoekella blandensis* (strain CECT 7118 / CCUG 51940 / MED217) (*Flavobacterium* sp. (strain MED217))
965 A0A0W0UST *Legionella gratiana*
966 A0A078KW7C *Legionella massiliensis*
967 A0A128ZD65 *Legionella pneumophila*
968 G8UUA1 *Legionella pneumophila* subsp. *pneumophila* ATCC 43290
969 A0A1S8GMP1 *Legionella pneumophila* subsp. *pneumophila* str. *Mississauga*
970 Q5ZRM8 *Legionella pneumophila* subsp. *pneumophila* (strain Philadelphia 1 / ATCC 33152 / DSM 7513)
971 S7X834 *Leifsonia rubra* CMS 76R
972 A0A1B8RWV *Leisingera* sp. JC1
973 A0A1H3KUN *Lentibacter algarum*
974 A0A1V5UW7 *Lentisphaerae bacterium* ADurb.Bin242
975 U9VSG3 *Leptolyngbya* sp. Heron Island J
976 A0A110AWJ2 *Leptolyngbya* sp. O-77
977 A0A166V5S7 *Leptolyngbya valderiana* BDU 20041
978 A0A1D7UZC1 *Leptospira alstonii*

979 Q72PS5 *Leptospira interrogans* serogroup Icterohaemorrhagiae serovar copenhageni (strain Fiocruz L1-130)
980 A0A1X8WRS *Leptospira interrogans* serovar Canicola
981 A0A098MSF5 *Leptospira interrogans* serovar Lai
982 A0A150FF05 *Leptospira* sp. ZV016
983 A0A147ENK4 *Leucobacter chromiirensistens*
984 A0A0D0ILT8 *Leucobacter komagatae*
985 A0A0M2LXR5 *Leucobacter* sp. Ag1
986 A0A061LXD1 *Leucobacter* sp. UCD-THU
987 A6GNC6 *Limnobacter* sp. MED105
988 A0A0L8AWU1 *Limnohabitans planktonicus* II-D5
989 A0A0P0M7I9 *Limnohabitans* sp. 63ED37-2
990 A0A0Y7J7Z3 *Listeria monocytogenes*
991 A0A1M4V7X3 *Loktanella atrilutea*
992 A0A238LBI3 *Loktanella marina*
993 A0A0Q1BAI4 *Loktanella* sp. 1ANDIMAR09
994 A0A0Q2UI66 *Loktanella* sp. 3ANDIMAR09
995 A0A0Q3F5K0 *Loktanella* sp. 5RATIMAR09
996 A0A0F4RI65 *Loktanella* sp. S4079
997 A0A0F3L255 *Luteibacter yejuensis*
998 A0A290XED3 *Luteimonas* sp. 100111
999 A0A0H1AJ03 *Luteimonas* sp. FCS-9
1000 A0A1B3W664 *Luteimonas* sp. JM171
1001 A0A143PQUC *Luteitalea pratensis*
1002 V4RKZ2 *Lutibaculum baratangense* AMV1
1003 A0A2A4D532 *Lysinibacillus fusiformis*
1004 A0A0A3I369 *Lysinibacillus manganicus* DSM 26584
1005 A0A0A3HV67 *Lysinibacillus sinduriensis* BLB-1 = JCM 15800
1006 A0A0M9WVL *Lysinibacillus* sp. FJAT-14222
1007 A0A0M0X3A7 *Lysinibacillus* sp. FJAT-14745
1008 A0A0P7L9V9 *Lysinibacillus* sp. ZYM-1
1009 A0A0K9FFE7 *Lysinibacillus xylanilyticus*
1010 A0A120AHI2 *Lysobacter capsici* AZ78
1011 A0A0A0EL72 *Lysobacter concretionis* Ko07 = DSM 16239
1012 A0A0A0M5Y1 *Lysobacter defluvii* IMMIB APB-9 = DSM 18482
1013 W6KA50 *Magnetospira* sp. QH-2
1014 A0A238KTS8 *Maliponia aquimaris*
1015 A0A0B3S1J9 *Mameliella alba*
1016 A0Z923 marine gamma proteobacterium HTCC2080
1017 A0A225NPB5 *Marinibacterium profundimaris*
1018 A0A1R4JC33 *Marinilactibacillus psychrotolerans* 42ea
1019 A0A1E3C726 *Marinobacter adhaerens*
1020 E4PFN8 *Marinobacter adhaerens* (strain HP15)
1021 A6F4G5 *Marinobacter algicola* DG893
1022 A0A137SG28 *Marinobacter excellens* LAMA 842
1023 H8WBP3 *Marinobacter hydrocarbonoclasticus* ATCC 49840

1024 A0A1M2UY0E *Marinobacter hydrocarbonoclasticus* (*Pseudomonas nautica*)
1025 A1U5T4 *Marinobacter hydrocarbonoclasticus* (strain ATCC 700491 / DSM 11845 / VT8)
1026 G6YQE9 *Marinobacter manganoxydans* Mnl7-9
1027 A0A072MYZ0 *Marinobacter nitratireducens*
1028 A0A1W6K3W *Marinobacter salarius*
1029 A0A1D9GK4E *Marinobacter salinus*
1030 U7P5N7 *Marinobacter* sp. C1S70
1031 A0A0K1UDP8 *Marinobacter* sp. CP1
1032 U7HQ79 *Marinobacter* sp. EN3
1033 U7NU24 *Marinobacter* sp. EVN1
1034 A0A142FMH5 *Marinobacter* sp. LQ44
1035 A0A139DBK7 *Marinobacter* sp. T13-3
1036 A0A1E7PWU *Marinobacter* sp. X15-166B
1037 A0A1C3JR26 *Marinomonas gallaica*
1038 A6VTQ7 *Marinomonas* sp. (strain MWYL1)
1039 A0A1A8T0Y1 *Marinomonas spartinae*
1040 A3V9I8 *Maritimibacter alkaliphilus* HTCC2654
1041 A0A1X4P2B1 *Marivita cryptomonadis*
1042 A0A1X4NFB4 *Marivita geojedonensis*
1043 A0A098UEM2 *Massilia* sp. JS1662
1044 A0A085FAJ7 *Massilia* sp. LC238
1045 I6ZPY4 *Melioribacter roseus* (strain JCM 17771 / P3M-2)
1046 H0I240 *Mesorhizobium alhagi* CCNWXJ12-2
1047 G6YEM4 *Mesorhizobium amorphae* CCNWGS0123
1048 A0A143NGF9 *Mesorhizobium ciceri*
1049 E8TFY8 *Mesorhizobium ciceri* biovar *biserrulae* (strain HAMBI 2942 / LMG 23838 / WSM1271)
1050 A0A068DC64 *Mesorhizobium huakuii* 7653R
1051 A0A271L571 *Mesorhizobium mediterraneum*
1052 M5F152 *Mesorhizobium metallidurans* STM 2683
1053 F7YAB6 *Mesorhizobium opportunistum* (strain LMG 24607 / HAMBI 3007 / WSM2075)
1054 A0A090GTY8 *Mesorhizobium plurifarum*
1055 A0A1R3V877 *Mesorhizobium prunedense*
1056 A0A1Q3M0U8 *Mesorhizobium* sp. 61-13
1057 A0A1Q4A7E5 *Mesorhizobium* sp. 65-26
1058 A0A1A5T406 *Mesorhizobium* sp. AA22
1059 A0A1L3SYH6 *Mesorhizobium* sp. B7
1060 A0A2A6FLH7 *Mesorhizobium* sp. BSA136
1061 X6KLJ0 *Mesorhizobium* sp. L103C105A0
1062 X6K7M6 *Mesorhizobium* sp. L103C119B0
1063 V7HT44 *Mesorhizobium* sp. L103C120A0
1064 X6JTW4 *Mesorhizobium* sp. L103C131B0
1065 X6J3L4 *Mesorhizobium* sp. L103C565B0
1066 X6J8R5 *Mesorhizobium* sp. L2C054A000
1067 X6I7N2 *Mesorhizobium* sp. L2C066B000
1068 X6HJR0 *Mesorhizobium* sp. L2C067A000

1069 X6H8J8 Mesorhizobium sp. L2C084A000
1070 X6G5S1 Mesorhizobium sp. L2C085B000
1071 V7H9P6 Mesorhizobium sp. L2C089B000
1072 X6FZ61 Mesorhizobium sp. L48C026A00
1073 A0A0H1AES7 Mesorhizobium sp. LC103
1074 A0A1S1SG27 Mesorhizobium sp. LCM 4576
1075 X6FHW4 Mesorhizobium sp. LNHC209A00
1076 X6EHB4 Mesorhizobium sp. LNHC220B00
1077 X6DRB8 Mesorhizobium sp. LNHC221B00
1078 X6EXE1 Mesorhizobium sp. LNHC229A00
1079 X6CV37 Mesorhizobium sp. LNHC232B00
1080 X6CTS0 Mesorhizobium sp. LNHC252B00
1081 X6C982 Mesorhizobium sp. LNJC372A00
1082 X6AVM4 Mesorhizobium sp. LNJC386A00
1083 A0A0E2NY95 Mesorhizobium sp. LNJC391B00
1084 V7G350 Mesorhizobium sp. LNJC394B00
1085 X5ZN91 Mesorhizobium sp. LNJC395A00
1086 X5YTQ4 Mesorhizobium sp. LNJC399B00
1087 X5YED4 Mesorhizobium sp. LNJC405B00
1088 X5XHV3 Mesorhizobium sp. LSHC412B00
1089 A0A0E2NJ11 Mesorhizobium sp. LSHC414A00
1090 V7FFU5 Mesorhizobium sp. LSHC420B00
1091 X5VQD4 Mesorhizobium sp. LSHC422A00
1092 X5UK31 Mesorhizobium sp. LSHC426A00
1093 V7F871 Mesorhizobium sp. LSJC264A00
1094 X5RZX9 Mesorhizobium sp. LSJC265A00
1095 X5R0C6 Mesorhizobium sp. LSJC269B00
1096 X5PEK1 Mesorhizobium sp. LSJC277A00
1097 X5QIX6 Mesorhizobium sp. LSJC280B00
1098 X5QKZ5 Mesorhizobium sp. LSJC285A00
1099 A0A090DDQ€ Mesorhizobium sp. ORS3324
1100 A0A090ETB7 Mesorhizobium sp. ORS3359
1101 A0A1S1TFB8 Mesorhizobium sp. ORS3428
1102 A0A0Q6P2V0 Mesorhizobium sp. Root102
1103 A0A0Q8AKC! Mesorhizobium sp. Root157
1104 A0A0Q8K469 Mesorhizobium sp. Root172
1105 A0A0Q7W1A; Mesorhizobium sp. Root552
1106 A0A0Q7WPM Mesorhizobium sp. Root554
1107 A0A1D2SS43 Mesorhizobium sp. SCN 65-20
1108 A0A1E2SRV1 Mesorhizobium sp. SEMIA 3007
1109 A0A090ENA9 Mesorhizobium sp. SOD10
1110 M5FMD4 Mesorhizobium sp. STM 4661
1111 A0A1C2DE92 Mesorhizobium sp. UASWS1009
1112 A0A1W6X12C Mesorhizobium sp. WSM1497
1113 A0A2A3DTH¿ Mesorhizobium sp. WSM3859

1114 A0A2A3DEK6 Mesorhizobium sp. WSM3860
1115 A0A2A3CR40 Mesorhizobium sp. WSM3864
1116 A0A2A3B450 Mesorhizobium sp. WSM3866
1117 A0A1A5S3L8 Mesorhizobium sp. WSM3873
1118 A0A2A3CDF7 Mesorhizobium sp. WSM3876
1119 A0A2A3BZI7 Mesorhizobium sp. WSM4308
1120 A0A2A3AXT4 Mesorhizobium sp. WSM4313
1121 A0A271KQH1 Mesorhizobium sp. WYCCWR 10019
1122 A0A271LNK6 Mesorhizobium temperatum
1123 A0A0E3RYX1 Methanosarcina mazei C16
1124 A0A0E3LW73 Methanosarcina mazei LYC
1125 A0A0F8FYB0 Methanosarcina mazei (Methanosarcina frisia)
1126 A0A0E3RF04 Methanosarcina mazei S-6
1127 A0A0E3RBS4 Methanosarcina mazei SarPi
1128 Q8PSC2 Methanosarcina mazei (strain ATCC BAA-159 / DSM 3647 / Goe1 / Go1 / JCM 11833 / OCM 88) (Methanosarcina frisia)
1129 M1PD91 Methanosarcina mazei Tuc01
1130 A0A0E3PYT3 Methanosarcina mazei WWM610
1131 A0A0F8P197 Methanosarcina sp. 1.H.A.2.2
1132 A0A0F8SR54 Methanosarcina sp. 1.H.T.1A.1
1133 A0A0F8C747 Methanosarcina sp. 2.H.A.1B.4
1134 A0A0F8EQ81 Methanosarcina sp. 2.H.T.1A.15
1135 A0A0F8ESV2 Methanosarcina sp. 2.H.T.1A.3
1136 A0A0F8FHF3 Methanosarcina sp. 2.H.T.1A.6
1137 A0A0F8G5G5 Methanosarcina sp. 2.H.T.1A.8
1138 A0A0E3NWM Methanosarcina sp. MTP4
1139 A0A154NUR2 Methylobacterium radiotolerans
1140 A0A1E3H3I1 Methylobrevia pamukkalensis
1141 A0A1Z8KTN9 Micavibrio sp. TMED2
1142 A0A0F0LDS8 Microbacterium azadirachtae
1143 A0A0M2HBF1 Microbacterium ketosireducens
1144 A0A0B2A2N2 Microbacterium mangrovi
1145 A0A0F0L4Z2 Microbacterium oxydans
1146 A0A246E6P5 Microbacterium sp. AISO3
1147 A0A171DW03 Microbacterium sp. HM58-2
1148 A0A0Q5L8W0 Microbacterium sp. Leaf159
1149 A0A0Q5FPP4 Microbacterium sp. Leaf320
1150 A0A0Q8LJU1 Microbacterium sp. Root180
1151 A0A0F2C737 Microbacterium sp. SA39
1152 A0A161SK65 Microbacterium sp. TNHR37B
1153 U2XM45 Microbacterium sp. TS-1
1154 A0A0M2HA45 Microbacterium trichothecenolyticum
1155 A0A1C9W5R Microbulbifer aggregans
1156 A0A0X3U0N1 Microbulbifer sp. ZGT114
1157 F5XN73 Microlunatus phosphovorius (strain ATCC 700054 / DSM 10555 / JCM 9379 / NBRC 101784 / NCIMB 13414 / VKM Ac-1990 / NM-1)
1158 A0A109IJD6 Micromonospora rifamycinica

1159 A0A1S8Y220 Micromonospora sp. Rc5
1160 A0A246RH91 Micromonospora wenchangensis
1161 A1ZRN3 Microscilla marina ATCC 23134
1162 A0A1B2EJ22 Microvirga ossetica
1163 A0A0H1RG87 Microvirga vignae
1164 E6Q0K6 mine drainage metagenome
1165 A0A066UEGC Moraxella bovoculi 237
1166 A0A167ZV10 Moraxella ovis
1167 A0A0X8X008 Mucilaginibacter gotjawali
1168 A0A0B2BMY2 Mumia flava
1169 A0A0U1AWE Mycobacterium abscessus
1170 A0A1N0QLG Mycobacterium abscessus subsp. abscessus
1171 A0A1N6AFE1 Mycobacterium abscessus subsp. bolletii
1172 R4URJ3 Mycobacterium abscessus subsp. bolletii 50594
1173 A0A1U4ADK Mycobacterium abscessus subsp. massiliense
1174 A0A0H2ZT69 Mycobacterium avium (strain 104)
1175 T2GNH3 Mycobacterium avium subsp. hominissuis (strain TH135)
1176 A0A100VW Mycobacterium brisbanense
1177 A0A220Y6G5 Mycobacterium chimaera
1178 A0A0J6WN57 Mycobacterium chlorophenicum
1179 A0A0J6ZER6 Mycobacterium chubuense
1180 A0A024LTN1 Mycobacterium farcinogenes
1181 A0A0N9Y2F7 Mycobacterium fortuitum
1182 K0V4A6 Mycobacterium fortuitum subsp. fortuitum DSM 46621 = ATCC 6841
1183 A0A1V3X482 Mycobacterium kansasii
1184 X7XUS9 Mycobacterium kansasii 824
1185 X5L4C9 Mycobacterium mageritense DSM 44476 = CIP 104973
1186 A0A0H5RKE Mycobacterium neworleansense
1187 A0A0J6VE36 Mycobacterium obuense
1188 D5PGB3 Mycobacterium parascrofulaceum ATCC BAA-614
1189 A0A202FU52 Mycobacterium paratuberculosis
1190 A0A100IEI0 Mycobacterium pseudoshottsii
1191 A0A0D6GM4 Mycobacterium smegmatis
1192 A0QY21 Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155)
1193 L8FC86 Mycobacterium smegmatis (strain MKD8)
1194 A0A1A2BEH Mycobacterium sp. 852002-51759_SCH5129042
1195 A0A1A0X9U3 Mycobacterium sp. 852013-50091_SCH5140682
1196 A0A145SW7 Mycobacterium sp. QIA-37
1197 A0A2C9SVL6 Mycobacterium sp. shizuoka-1
1198 A1UEG1 Mycobacterium sp. (strain KMS)
1199 A0A0A1FQN Mycobacterium sp. VKM Ac-1817D
1200 A0A240AD36 Mycobacterium terrae
1201 A0A0U0UQT Mycobacterium tuberculosis
1202 L7N542 Mycobacterium tuberculosis (strain CDC 1551 / Oshkosh)
1203 X8FKH9 Mycobacterium ulcerans str. Harvey

1204 K0UTN0 Mycobacterium vaccae ATCC 25954
1205 X5LRH7 Mycobacterium vulneris
1206 X8AXV3 Mycobacterium xenopi 3993
1207 X8BH17 Mycobacterium xenopi 4042
1208 A0A0F7E671 Myxococcus fulvus 124B02
1209 F8CEF7 Myxococcus fulvus (strain ATCC BAA-855 / HW-1)
1210 A0A0H4XJ21 Myxococcus hansupus
1211 L7UL65 Myxococcus stipitatus (strain DSM 14675 / JCM 12634 / Mx s8)
1212 Q1CYM7 Myxococcus xanthus (strain DK 1622)
1213 C8XA27 Nakamurella multipartita (strain ATCC 700099 / DSM 44233 / CIP 104796 / JCM 9543 / NBRC 105858 / Y-104) (Microsphaera multipartita)
1214 A0A0H5D2L7 Nautella italica
1215 A0A0J1C2T6 Neisseria arctica
1216 A0A0B5CIZ5 Neisseria elongata subsp. glycolytica ATCC 29315
1217 A0A154TM64 Neisseria flavescens
1218 C5TNJ7 Neisseria flavescens SK114
1219 F9EZJ5 Neisseria macacae ATCC 33926
1220 E5UL28 Neisseria mucosa C102
1221 G4CK47 Neisseria shayeganii 871
1222 A0A1F1UJC7 Neisseria sp. HMSC03D10
1223 A0A1F1C597 Neisseria sp. HMSC055H02
1224 A0A1F0DW5 Neisseria sp. HMSC056A03
1225 A0A1F0AMB Neisseria sp. HMSC064E01
1226 A0A1F1FS36 Neisseria sp. HMSC067G11
1227 A0A1E9MTX Neisseria sp. HMSC068C04
1228 A0A1F1GKP1 Neisseria sp. HMSC069H12
1229 A0A1E9M5I0 Neisseria sp. HMSC070A01
1230 A0A1F1HPY2 Neisseria sp. HMSC073G10
1231 A0A1E9VWU Neisseria sp. HMSC077D05
1232 A0A1F1SLT0 Neisseria sp. HMSC31F04
1233 A0A1S1G8RC Neisseria sp. HMSC70E02
1234 A0A0T7FTX7 Neorhizobium galegae bv. officinalis
1235 A0A068TGW Neorhizobium galegae bv. officinalis bv. officinalis str. HAMB1 1141
1236 A0A0T7HMP Neorhizobium galegae bv. orientalis
1237 A0A068STZ0 Neorhizobium galegae bv. orientalis str. HAMB1 540
1238 A0A0U1NN44 Nereida ignava
1239 A0A1X3PB87 Nesterenkonia sp. PF2B19
1240 A0A084UCB4 Nitratireductor basaltis
1241 W9V8H0 Nitrincola nitratireducens
1242 Q0AH54 Nitrosomonas eutropha (strain C91)
1243 A0A0S8APY1 Nitrospira bacterium SG8_35_4
1244 A0A248K4P8 Nitrospirillum amazonense CBAmc
1245 U5EL89 Nocardia asteroides NBRC 15531
1246 A0A231GYE2 Nocardia cerradoensis
1247 A0A2B4QKU Nocardia farcinica
1248 A0A0B8ND93 Nocardia seriolae

1249 A0A1A0L1S4 *Nocardia* sp. 852002-20019_SCH5090214
1250 E9UX93 *Nocardiodaceae* bacterium Broad-1
1251 A0A1A9GMI1 *Nocardioides dokdonensis* FR1436
1252 A0A0Q7PV51 *Nocardioides* sp. Root140
1253 A0A0Q7ZNK7 *Nocardioides* sp. Root151
1254 A0A0Q8Q364 *Nocardioides* sp. Root190
1255 A0A0Q8WQC *Nocardioides* sp. Root240
1256 A0A0Q8DQG *Nocardioides* sp. Root614
1257 A0A0Q8HK2C *Nocardioides* sp. Root682
1258 A0A0Q8V1T5 *Nocardioides* sp. Root79
1259 A0A0Q9RUJ4 *Nocardioides* sp. Soil796
1260 A0A0Q9SCRf *Nocardioides* sp. Soil797
1261 A0A1M4E4Qf *Nonomuraea gerenzanensis*
1262 A0A1V0AKQ7 *Nonomuraea* sp. ATCC 55076
1263 A0A235I1V5 *Nostoc* sp. 'Peltigera membranacea cyanobiont' 210A
1264 A0A235IRH3 *Nostoc* sp. 'Peltigera membranacea cyanobiont' 232
1265 Q2G7E0 *Novosphingobium aromaticivorans* (strain ATCC 700278 / DSM 12444 / CIP 105152 / NBRC 16084 / F199)
1266 A0A117UY79 *Novosphingobium fuchskuhlense*
1267 A0A1U6HY74 *Novosphingobium mathurense*
1268 F1Z9M7 *Novosphingobium nitrogenifigens* DSM 19370
1269 A0A258XNZ9 *Novosphingobium* sp. 16-62-11
1270 A0A259E054 *Novosphingobium* sp. 17-62-19
1271 A0A258V6P5 *Novosphingobium* sp. 28-62-57
1272 A0A0N1AP13 *Novosphingobium* sp. AAP1
1273 A0A0N1LJ13 *Novosphingobium* sp. AAP93
1274 A0A102DFP0 *Novosphingobium* sp. FSW06-99
1275 A0A102DAI9 *Novosphingobium* sp. Fuku2-ISO-50
1276 A0A257GFJ3 *Novosphingobium* sp. PASSN1
1277 A0A1E4N4I9 *Novosphingobium* sp. SCN 63-17
1278 A0A1E4MI48 *Novosphingobium* sp. SCN 66-18
1279 A0A0B9AG37 *Novosphingobium subterraneum*
1280 U3A1A2 *Novosphingobium tardaugens* NBRC 16725
1281 A0A1Y5THU2 *Oceanibacterium hippocampi*
1282 A0A154W8E6 *Oceanibaculum pacificum*
1283 A9DQY8 *Oceanibulbus indolifex* HEL-45
1284 A0A0P7XVY2 *Oceanicaulis* sp. HLUCCA04
1285 A0A1M7RRW *Oceanicella actignis*
1286 A0A1N6GAYf *Oceanicola litoreus*
1287 A0A254R1P2 *Oceanicola* sp. 22II-s10i
1288 A0A235CKE2 *Oceanimonas baumannii*
1289 A0A233REU1 *Oceanimonas doudoroffii*
1290 H2FXE8 *Oceanimonas* sp. (strain GK1)
1291 A0A1Y0D711 *Oceanisphaera profunda*
1292 A0A1Y0D0L7 *Oceanisphaera* sp. AMac2203
1293 Q8ESC9 *Oceanobacillus iheyensis* (strain DSM 14371 / CIP 107618 / JCM 11309 / KCTC 3954 / HTE831)

1294 A0A0A1MHT8 Oceanobacillus oncorhynchi
1295 W9AFE3 Oceanobacillus picturae
1296 A0A137XSC8 Ochrobactrum anthropi
1297 A6WYP4 Ochrobactrum anthropi (strain ATCC 49188 / DSM 6882 / JCM 21032 / NBRC 15819 / NCTC 12168)
1298 A0A1J6HWP2 Ochrobactrum cytisi
1299 U4VF86 Ochrobactrum intermedium 229E
1300 C4WGF5 Ochrobactrum intermedium LMG 3301
1301 M5K1T6 Ochrobactrum intermedium M86
1302 A0A1A9FKX2 Ochrobactrum pseudogrignonense
1303 U1XHP2 Ochrobactrum sp. EGD-AQ16
1304 A0A1V3YKH5 Ochrobactrum sp. P6BS-III
1305 A0A238KNR1 Octadecabacter ascidiaceicola
1306 A0A0K0Y7L4 Octadecabacter temperatus
1307 A0A165SFE7 Oleiphilus sp. HI0065
1308 R4YTJ3 Oleispira antarctica RB-8
1309 A0A1Z8WET7 Opitutae bacterium TMED67
1310 A1RBL5 Paenarthrobacter aurescens (strain TC1)
1311 K5AE03 Paenibacillus alvei DSM 29
1312 A0A168JB90 Paenibacillus antarcticus
1313 A0A1R0XZV7 Paenibacillus borealis
1314 A0A089HPB3 Paenibacillus durus (Paenibacillus azotofixans)
1315 A0A161SCT3 Paenibacillus elgii
1316 A0A0W1AX88 Paenibacillus etheri
1317 A0A1V4HFG4 Paenibacillus ferrarius
1318 A0A168HN33 Paenibacillus glacialis
1319 A0A163LZB8 Paenibacillus glucanolyticus
1320 A0A2A5LKM7 Paenibacillus lautus (Bacillus lautus)
1321 A0A1R1HJD6 Paenibacillus macerans (Bacillus macerans)
1322 A0A168GS77 Paenibacillus macquariensis (Bacillus macquariensis)
1323 A0A167XGX2 Paenibacillus macquariensis subsp. defensor
1324 A0A264EMM6 Paenibacillus odorifer
1325 A0A1T2X1D1 Paenibacillus selenitireducens
1326 A0A0P9DRM1 Paenibacillus sp. A3
1327 A0A0D7KQ86 Paenibacillus sp. E194
1328 A0A0M1P4C6 Paenibacillus sp. FJAT-22460
1329 A0A1R0ZLN72 Paenibacillus sp. FSL A5-0031
1330 A0A089I662 Paenibacillus sp. FSL H7-0357
1331 A0A089JUA0 Paenibacillus sp. FSL H7-0737
1332 A0A0D3V785 Paenibacillus sp. IHBB 10380
1333 A0A1B8UWV Paenibacillus sp. KS1
1334 A0A1S2F3U7 Paenibacillus sp. LC231
1335 A0A0B0I7C6 Paenibacillus sp. P1XP2
1336 A0A1E3L0A7 Paenibacillus sp. TI45-13ar
1337 A0A264E2N8 Paenibacillus sp. VTT E-133280
1338 A0A081PA59 Paenibacillus tyrfis

1339 A0A0A8W0V4 Paeniclostridium sordellii (Clostridium sordellii)
1340 M7NKQ6 Paeniglutamicibacter gangotriensis Lz1y
1341 A0A086XQY2 Paenirhodobacter enshiensis
1342 A0A1U7CSL2 Paludisphaera borealis
1343 A0A0U2XVV8 Pandoraea norimbergensis
1344 R7X1Q8 Pandoraea sp. SD6-2
1345 A0A239SAD1 Pandoraea sputorum
1346 A0A0U3P646 Pannonibacter phragmitetus
1347 A0A0F5XVQ9 Pantoea sp. 3.5.1
1348 A0A174I3J8 Parabacteroides distasonis
1349 A6LC41 Parabacteroides distasonis (strain ATCC 8503 / DSM 20701 / CIP 104284 / JCM 5825 / NCTC 11152)
1350 A0A1H7HJL6 Paraburkholderia caballeronis
1351 A0A142YPP2 Paraburkholderia caribensis
1352 A0A1H0Z455 Paraburkholderia fungorum
1353 A0A1A9NCM Paraburkholderia ginsengiterrae
1354 A0A1H6I231 Paraburkholderia hospita
1355 A0A1I3UYM3 Paraburkholderia megapolitana
1356 A0A149PB03 Paraburkholderia monticola
1357 A0A1N7RJA0 Paraburkholderia ribeironis
1358 A0A0A6QGRi Paraburkholderia sacchari
1359 A0A1H3XWD Paraburkholderia sartisoli
1360 A0A248VEP8 Paraburkholderia sp. BN5
1361 A0A1L6HS45 Paraburkholderia sp. SOS3
1362 A0A1I9YIB6 Paraburkholderia sprentiae WSM5005
1363 A0A1X7K361 Paraburkholderia susongensis
1364 A0A244DCT3 Paraburkholderia terrae
1365 A0A1M6XT47 Paraburkholderia terricola
1366 A0A1H6ZJ89 Paraburkholderia tropica
1367 A0A1H1DQT9 Paraburkholderia tuberum
1368 Q146S0 Paraburkholderia xenovorans (strain LB400)
1369 S5XML4 Paracoccus aminophilus JCM 7686
1370 A0A099F5D0 Paracoccus halophilus
1371 A0A099GG23 Paracoccus sanguinis
1372 A0A0V0PZU1 Paracoccus sp. MKU1
1373 A0A1L9BQV3 Paracoccus sp. SM22M-07
1374 A0A2A2GLU1 Paracoccus sp. WN007
1375 A0A099F717 Paracoccus versutus (Thiobacillus versutus)
1376 A0A212BJ66 Paracoccus yeei
1377 A0A0U1KJ51 Paraliobacillus sp. PM-2
1378 A0A135HQQ9 Paramesorhizobium deserti
1379 A0A1C7P2K3 Pararhizobium polonicum
1380 H0E165 Patulibacter medicamentivorans
1381 C6XX09 Pedobacter heparinus (strain ATCC 13125 / DSM 2366 / CIP 104194 / JCM 7457 / NBRC 12017 / NCIMB 9290 / NRRL B-14731 / HIM 762-3)
1382 A0A2A2RVW Pedosphaera sp. Tous-C6FEB
1383 A0A1E4FC10 Pelagibacterium sp. SCN 63-23

1384 A0A1Y5SE37 *Pelagicola litorisediminis*
1385 A0A238L5C6 *Pelagimonas varians*
1386 A0A069RH13 *Peptoclostridium litorale* DSM 5388
1387 A0A1B0ZTMC *Phaeobacter gallaeciensis*
1388 A0A135INQ2 *Phaeobacter inhibens*
1389 A0A0P6RWA *Phaeobacter* sp. 11ANDIMAR09
1390 A0A254QXJ5 *Phaeobacter* sp. 22II1-1F12B
1391 A0A0P1H3R0 *Phaeobacter* sp. CECT 5382
1392 A0A0P1INL0 *Phaeobacter* sp. CECT 7735
1393 A0A0C2HNB8 *Phaeobacter* sp. S26
1394 A0A1G2WBX *Phenylobacterium* sp. RIFCSPHIGHO2_01_FULL_69_31
1395 A0A1G2X5F9 *Phenylobacterium* sp. RIFCSPHIGHO2_01_FULL_70_10
1396 A0A0Q7DHY3 *Phenylobacterium* sp. Root1277
1397 A0A0Q7FJH9 *Phenylobacterium* sp. Root1290
1398 A0A0Q8N5J1 *Phenylobacterium* sp. Root700
1399 A0A0Q8UIE9 *Phenylobacterium* sp. Root77
1400 A0A1E4GVH7 *Phenylobacterium* sp. SCN 70-31
1401 A0A227JMM1 *Phenylobacterium zucineum*
1402 B4RGK0 *Phenylobacterium zucineum* (strain HLK1)
1403 A0A0B9G044 *Photobacterium gaetbulicola*
1404 A0A0C5W5H *Photobacterium gaetbulicola* Gung47
1405 A0A1B8IG12 *Photobacterium phosphoreum*
1406 A0A1C0U4M3 *Photorhabdus asymbiotica* subsp. *australis*
1407 A0A208YD14 *Pigmentiphaga* sp. NML030171
1408 A0A177R0P0 *Planctomyces* sp. SCGC AG-212-M04
1409 A0A142WX71 *Planctomyces* sp. SH-PL14
1410 A0A142YP87 *Planctomyces* sp. SH-PL62
1411 A0A1V6DF56 *Planctomycetes bacterium* ADurb.Bin126
1412 A0A1L9P0F0 *Planktotalea frisia*
1413 A4C151 *Polaribacter irgensii* 23-P
1414 A2TYS8 *Polaribacter* sp. MED152
1415 D3B2W5 *Polysphondylium pallidum* (strain ATCC 26659 / Pp 5 / PN500)
1416 A0A225PDU6 *Ponticoccus lacteus*
1417 A0A0W7WL3 *Ponticoccus marisrubri*
1418 A0A0B2BYW1 *Porphyrobacter mercurialis*
1419 A0A099WYM *Porphyromonas cangingivalis*
1420 A0A293NHD9 *Porticoccaceae bacterium*
1421 A0A167RP12 *Powai lake megavirus*
1422 W4TMT4 *Propionibacterium acnes* JCM 18909
1423 W4TS23 *Propionibacterium acnes* JCM 18916
1424 W4U2X3 *Propionibacterium acnes* JCM 18918
1425 W4UGU7 *Propionibacterium acnes* JCM 18920
1426 Q6A7K7 *Propionibacterium acnes* (strain KPA171202 / DSM 16379)
1427 F3NYU6 [*Propionibacterium*] *humerusii* P08
1428 A0A1W9U178 *Propionibacterium* sp. 4572_24

1429 A0A1J4X7W6 Propionibacterium sp. CG1_02_60_36
1430 A0A1W9I9G9 Proteobacteria bacterium HN_bin10
1431 A0A1W9I5B6 Proteobacteria bacterium SG_bin9
1432 K8WC88 Providencia burhodogranariea DSM 19968
1433 A0A1V8RV42 Pseudaminobacter manganicus
1434 A0A0S2JZG0 Pseudoalteromonas phenolica
1435 A0A0W1LGY Pseudoalteromonas sp. H105
1436 A0A095VUC7 Pseudohalaea rubra DSM 19751
1437 A0A1G3D8W Pseudomonadales bacterium GWC2_63_15
1438 A0A2A4EXY0 Pseudomonas acidophila
1439 A0A0G5TIY0 Pseudomonas aeruginosa
1440 A6UYI4 Pseudomonas aeruginosa (strain PA7)
1441 A0A1H6BVG Pseudomonas aestusnigri
1442 A0A0P9J0D4 Pseudomonas amygdali
1443 A0A0P9U7U3 Pseudomonas amygdali pv. hibisci
1444 A0A0P9TG04 Pseudomonas amygdali pv. lachrymans (Pseudomonas syringae pv. lachrymans)
1445 F3IMV6 Pseudomonas amygdali pv. lachrymans str. M302278
1446 F3ETR8 Pseudomonas amygdali pv. mori str. 301020
1447 F3E346 Pseudomonas amygdali pv. morsprunorum str. M302280
1448 A0A0Q0BNX2 Pseudomonas amygdali pv. tabaci (Pseudomonas syringae pv. tabaci)
1449 F3K2X6 Pseudomonas amygdali pv. tabaci str. ATCC 11528
1450 K2SL24 Pseudomonas avellanae BPIC 631
1451 A0A031M8M Pseudomonas bauzanensis
1452 W8PKE1 Pseudomonas brassicacearum
1453 F2KF85 Pseudomonas brassicacearum (strain NFM421)
1454 A0A0P9MU5 Pseudomonas caricapapayae
1455 V4QC34 Pseudomonas chloritidismutans AW-1
1456 A0A0G3G5P Pseudomonas chlororaphis
1457 A0A127MTZ1 Pseudomonas citronellolis
1458 A0A1S2V3L1 Pseudomonas constantinii
1459 A0A0Q0YYC Pseudomonas endophytica
1460 A0A0P9TXJ2 Pseudomonas ficuserectae
1461 A0A0B3BZN1 Pseudomonas flexibilis
1462 A0A0F4TBS8 Pseudomonas fluorescens
1463 U1S6N4 Pseudomonas fluorescens EGD-AQ6
1464 G8Q395 Pseudomonas fluorescens F113
1465 E2XQ03 Pseudomonas fluorescens WH6
1466 A0A267AKD6 Pseudomonas fragi
1467 A0A0D0L8A0 Pseudomonas fulva
1468 F6ADS4 Pseudomonas fulva (strain 12-X)
1469 A0A0J6HVVH9 Pseudomonas helleri
1470 A0A0F4XW7 Pseudomonas kilonensis
1471 A0A0R2Y5N5 Pseudomonas libanensis
1472 A0A2C8F1D9 Pseudomonas lundensis
1473 A0A098SVW Pseudomonas lutea

1474 A0A059KY26 *Pseudomonas mandelii* PD30
1475 A0A1I7DRP8 *Pseudomonas marincola*
1476 A0A2A4FBW! [*Pseudomonas*] *mesoacidophila*
1477 A0A225DUDε *Pseudomonas oleovorans* subsp. *oleovorans*
1478 A0A1S8DH43 *Pseudomonas pachastrellae*
1479 A0A2A4AVP9 *Pseudomonas pelagia*
1480 S2KU52 *Pseudomonas plecoglossicida* NB2011
1481 A0A219A3U6 *Pseudomonas poae*
1482 A0A0J6KLU8 *Pseudomonas psychrophila*
1483 A0A1B2F4H7 *Pseudomonas putida* (*Arthrobacter siderocapsulatus*)
1484 A0A0C5RST4 *Pseudomonas putida* S13.1.2
1485 B1JA60 *Pseudomonas putida* (strain W619)
1486 A0A078M5A8 *Pseudomonas saudimassiliensis*
1487 A0A0P9WZT! *Pseudomonas savastanoi* pv. *nerii*
1488 A0A0P9ZTB9 *Pseudomonas savastanoi* pv. *savastanoi*
1489 A0A191YRK0 *Pseudomonas silesiensis*
1490 A0A1N7U2M7 *Pseudomonas simiae*
1491 Q59719 *Pseudomonas* sp.
1492 A0A1S1UYUε *Pseudomonas* sp. 06C 126
1493 A0A1C3GPS1 *Pseudomonas* sp. 1 R 17
1494 A0A0E9ZQPε *Pseudomonas* sp. 10-1B
1495 A0A1B5F5J3 *Pseudomonas* sp. 22 E 5
1496 A0A0F0F9L1 *Pseudomonas* sp. 2(2015)
1497 A0A1B5DAM1 *Pseudomonas* sp. 24 E 1
1498 A0A1B5DZGε *Pseudomonas* sp. 37 R 15
1499 A0A1B5DJF5 *Pseudomonas* sp. 44 R 15
1500 A0A1B5EAS6 *Pseudomonas* sp. 58 R 3
1501 A0A0N0BTRε *Pseudomonas* sp. 655
1502 A0A1V3SCY7 *Pseudomonas* sp. A25(2017)
1503 A0A212BBH3 *Pseudomonas* sp. A46
1504 J0Y6B6 *Pseudomonas* sp. Ag1
1505 A0A0I9SH83 *Pseudomonas* sp. BICA1-14
1506 A0A0F2QHZε *Pseudomonas* sp. BRH_c35
1507 A0A1D9ISN5 *Pseudomonas* sp. BS-2016
1508 A0A1V3MS54 *Pseudomonas* sp. C9
1509 S6HMS4 *Pseudomonas* sp. CF149
1510 A0A1V4FGV7 *Pseudomonas* sp. Ea RS28
1511 A0A0D9A9Mε *Pseudomonas* sp. ES3-33
1512 A0A0D6SUR7 *Pseudomonas* sp. FeS53a
1513 W2DGM3 *Pseudomonas* sp. FH4
1514 S2FK83 *Pseudomonas* sp. G5(2012)
1515 W6UZS9 *Pseudomonas* sp. GM41(2012)
1516 A0A263NQ63 *Pseudomonas* sp. IB20
1517 A0A0P7MPZε *Pseudomonas* sp. In5
1518 A0A1E5XA91 *Pseudomonas* sp. J237

1519 A0A246FWX(Pseudomonas sp. K2115
1520 A0A1C2K796 Pseudomonas sp. K35
1521 M5QQ88 Pseudomonas sp. Lz4W
1522 A0A1Y0KLY3 Pseudomonas sp. M30-35
1523 A0A1T1HQE(Pseudomonas sp. MF4836
1524 A0A0A1GEX(Pseudomonas sp. MT-1
1525 A0A2A3MFH(Pseudomonas sp. MT5
1526 A0A257E5Z5 Pseudomonas sp. PGPPP1
1527 A0A0N1DC31 Pseudomonas sp. RIT-PI-q
1528 A0A0Q6QAD(Pseudomonas sp. Root329
1529 A0A0Q8CQ3(Pseudomonas sp. Root569
1530 A0A0Q8IRR4 Pseudomonas sp. Root68
1531 A0A0Q8PJZ1 Pseudomonas sp. Root71
1532 A0A1C1W69(Pseudomonas sp. S1E40
1533 P13397 Pseudomonas sp. (strain NK87)
1534 A0A105T8Z6 Pseudomonas sp. TAD18
1535 A0A0T6VMV(Pseudomonas sp. TTU2014-105ASC
1536 A0A0S4HRC(Pseudomonas sp. URMO17WK12:I11
1537 A0A1V4LKD3 Pseudomonas sp. VI4.1
1538 A0A2A2FW1(Pseudomonas sp. WN033
1539 I4CSF9 Pseudomonas stutzeri CCUG 29243
1540 W8R682 Pseudomonas stutzeri (Pseudomonas perfectomarina)
1541 A0A0R2YNI1 Pseudomonas synxantha
1542 A0A085UV41 Pseudomonas syringae
1543 A0A0K8M1N(Pseudomonas syringae pv. actinidiae
1544 S6SFV9 Pseudomonas syringae pv. actinidiae ICMP 18807
1545 S6SKG4 Pseudomonas syringae pv. actinidiae ICMP 19096
1546 A0A0P9J079 Pseudomonas syringae pv. alisalensis
1547 A0A0P9JNW(Pseudomonas syringae pv. antirrhini
1548 A0A0P9NE39 Pseudomonas syringae pv. castaneae
1549 A0A0N0X8S9 Pseudomonas syringae pv. cilantro
1550 A0A0P9LJZ6 Pseudomonas syringae pv. coriandricola
1551 A0A0P9PAM(Pseudomonas syringae pv. daphniphylli
1552 A0A0P9UP84 Pseudomonas syringae pv. delphinii
1553 A0A0P9RAL4 Pseudomonas syringae pv. helianthi
1554 A0A0N0WUZ Pseudomonas syringae pv. maculicola
1555 A0A0P9XSH(Pseudomonas syringae pv. primulae
1556 A0A0P9YT75 Pseudomonas syringae pv. ribicola
1557 A0A0Q0CPS(Pseudomonas syringae pv. solidagae
1558 A0A0P9ZMA(Pseudomonas syringae pv. spinaceae
1559 S3NNZ0 Pseudomonas syringae pv. syringae SM
1560 A0A0Q0E9B3 Pseudomonas syringae pv. tagetis
1561 A0A099SNY9 Pseudomonas syringae pv. tomato
1562 Q87WM0 Pseudomonas syringae pv. tomato (strain ATCC BAA-871 / DC3000)
1563 A0A0Q0ELZ0 Pseudomonas syringae pv. viburni

1564 A0A0J6GS90 *Pseudomonas taetrolens*
1565 A0A176NE46 *Pseudomonas thivervalensis*
1566 A0A0M4R7T3 *Pseudomonas versuta*
1567 A0A1Y6JJ77 *Pseudomonas viridiflava*
1568 A0A0J6IH54 *Pseudomonas weihenstephanensis*
1569 A0A098FMG6 *Pseudomonas xanthomarina*
1570 A0A1Y2MRT7 *Pseudonocardia autotrophica* (*Amycolata autotrophica*) (*Nocardia autotrophica*)
1571 A0A292YZU5 *Pseudonocardia* sp. N23
1572 A0A0A0EBG4 *Pseudoceanicola atlanticus*
1573 A0A1Y5SNP5 *Pseudooctadecabacter jejudonensis*
1574 A0A081MJ35 *Pseudorhizobium pelagicum*
1575 A0A257GP62 *Pseudorhodobacter* sp. PARRP1
1576 A0A0S9KV91 *Pseudorhodoferax* sp. Leaf265
1577 A0A1Y5RVI3 *Pseudoruegeria aquimaris*
1578 A0A165T3A5 *Pseudovibrio axinellae*
1579 A0A165V164 *Pseudovibrio* sp. Ad26
1580 B6QYG6 *Pseudovibrio* sp. JE062
1581 G8PHS5 *Pseudovibrio* sp. (strain FO-BEG1)
1582 A0A165Q4T3 *Pseudovibrio* sp. W74
1583 A0A161XKM1 *Pseudovibrio* sp. WM33
1584 A0A144NFV2 *Psychrobacter alimentarius*
1585 U4T618 *Psychrobacter aquaticus* CMS 56
1586 Q1Q8U8 *Psychrobacter cryohalolentis* (strain K5)
1587 A0A1R4EGM *Psychrobacter pasteurii*
1588 A0A1R4GUP5 *Psychrobacter piechaudii*
1589 F5SNM6 *Psychrobacter* sp. 1501(2011)
1590 A0A229GP87 *Psychrobacter* sp. DAB_AL32B
1591 A0A1U6GQ15 *Psychrobacter* sp. DAB_AL43B
1592 S4YWH3 *Psychrobacter* sp. G
1593 A0A2A2B406 *Psychrobacter* sp. JB193
1594 A0A1R4I736 *Psychrobacter* sp. JB385
1595 X0QQK5 *Psychrobacter* sp. JCM 18900
1596 X0QIF0 *Psychrobacter* sp. JCM 18901
1597 X0QSL3 *Psychrobacter* sp. JCM 18902
1598 X0RIH1 *Psychrobacter* sp. JCM 18903
1599 A0A0Q9ZRI8 *Psychrobacter* sp. P11F6
1600 A0A0M4U492 *Psychrobacter urativorans*
1601 D5BRY7 *Puniceispirillum marinum* (strain IMCC1322)
1602 A0A191ZW16 *Ralstonia insidiosa*
1603 A0A0D5AQW *Ralstonia mannitolilytica*
1604 A0A291ELC1 *Ralstonia pickettii* (*Burkholderia pickettii*)
1605 U3QQX9 *Ralstonia pickettii* DTP0602
1606 C6BID3 *Ralstonia pickettii* (strain 12D)
1607 B2UC09 *Ralstonia pickettii* (strain 12J)
1608 A0A223H6L1 *Ralstonia pseudosolanacearum*

1609 D8NJ41 *Ralstonia solanacearum* CFBP2957
 1610 D8NBY7 *Ralstonia solanacearum* CMR15
 1611 A0A1L3DIY2 *Ralstonia solanacearum* FJAT-1458
 1612 A0A177S2X0 *Ralstonia solanacearum* (*Pseudomonas solanacearum*)
 1613 D8NTL3 *Ralstonia solanacearum* PSI07
 1614 Q8XYJ3 *Ralstonia solanacearum* (strain GMI1000) (*Pseudomonas solanacearum*)
 1615 F6G7H0 *Ralstonia solanacearum* (strain Po82)
 1616 A3RX73 *Ralstonia solanacearum* UW551
 1617 A0A0B1YN14 *Ralstonia* sp. A12
 1618 S9S1M4 *Ralstonia* sp. AU12-08
 1619 A0A0J9DWC *Ralstonia* sp. MD27
 1620 A0A100HXG4 *Ralstonia* sp. NT80
 1621 G3A899 *Ralstonia syzygii* R24
 1622 A0A127JSC9 *Ramlibacter tataouinensis*
 1623 A0A1Y6GHY5 *Raoultella ornithinolytica* (*Klebsiella ornithinolytica*)
 1624 A0A166H147 *Rathayibacter tanacetii*
 1625 A0A0Q7AMF5 *Rhizobacter* sp. Root404
 1626 A0A257XGZ6 *Rhizobiales* bacterium 12-66-7
 1627 A0A1Q4BVI8 *Rhizobiales* bacterium 62-47
 1628 A0A1M3AV42 *Rhizobiales* bacterium 63-7
 1629 A0A1M2Z684 *Rhizobiales* bacterium 65-79
 1630 A0A0N0LFA6 *Rhizobium acidisoli*
 1631 A0A120FWW4 *Rhizobium altiplani*
 1632 A0A2A6GKC4 *Rhizobium anhuiense*
 1633 A0A2D0ABR4 *Rhizobium esperanzae*
 1634 F2AAZ5 *Rhizobium etli* CNPAF512
 1635 Q2K6A3 *Rhizobium etli* (strain CFN 42 / ATCC 51251)
 1636 B3PTX9 *Rhizobium etli* (strain CIAT 652)
 1637 W6RDB9 *Rhizobium favelukesii*
 1638 A0A2A6M534 *Rhizobium fredii* (*Sinorhizobium fredii*)
 1639 G9A8D7 *Rhizobium fredii* (strain HH103) (*Sinorhizobium fredii*)
 1640 N6V9P4 *Rhizobium freirei* PRF 81
 1641 A0A1S9GKK4 *Rhizobium laguerreae*
 1642 A0A1B1CN08 *Rhizobium leguminosarum*
 1643 A0A1B8R1QC *Rhizobium leguminosarum* bv. trifolii
 1644 W0IMH0 *Rhizobium leguminosarum* bv. trifolii WSM1689
 1645 A0A0U3ATX7 *Rhizobium leguminosarum* bv. viciae
 1646 Q1ME15 *Rhizobium leguminosarum* bv. viciae (strain 3841)
 1647 A0A1S9H0G4 *Rhizobium leguminosarum* bv. viciae USDA 2370
 1648 A0A1B4YK92 *Rhizobium loti* (*Mesorhizobium loti*)
 1649 Q98HG2 *Rhizobium loti* (strain MAFF303099) (*Mesorhizobium loti*)
 1650 A0A222HER5 *Rhizobium meliloti* (*Ensifer meliloti*) (*Sinorhizobium meliloti*)
 1651 Q92UN4 *Rhizobium meliloti* (strain 1021) (*Ensifer meliloti*) (*Sinorhizobium meliloti*)
 1652 A0A1Q9B080 *Rhizobium oryzae*
 1653 A0A1Q8ZN64 *Rhizobium oryzae*

1654 A0A2A6H568 Rhizobium phaseoli
1655 A0A0M3GIZ2 Rhizobium phaseoli Ch24-10
1656 A0A1L9CVU4 Rhizobium pusense
1657 A0A0F4G225 Rhizobium radiobacter (Agrobacterium tumefaciens) (Agrobacterium radiobacter)
1658 A0A061MJ24 Rhizobium rhizogenes NBRC 13257
1659 A0A1Q9AM88 Rhizobium rhizosphaerae
1660 A0A2A5KR35 Rhizobium sophoriradicis
1661 A0A248W6S6 Rhizobium sp. 10195
1662 A0A1L8PNC4 Rhizobium sp. 58
1663 A0A1M3QF16 Rhizobium sp. 60-20
1664 A0A1M2Z2Z4 Rhizobium sp. 63-7
1665 A0A0N0JP16 Rhizobium sp. AAP116
1666 A0A0N0J5Q8 Rhizobium sp. AAP43
1667 A0A1B9RWR Rhizobium sp. AC27/96
1668 A0A1B9SHI5 Rhizobium sp. AC44/96
1669 A0A292EBQ9 Rhizobium sp. ACO-34A
1670 A0A2A6JHA7 Rhizobium sp. C5
1671 W6WR90 Rhizobium sp. CF080
1672 A0A1W6KTY1 Rhizobium sp. CIAT894
1673 A0A2A6KE93 Rhizobium sp. FH14
1674 A0A2A6R6K9 Rhizobium sp. H4
1675 A0A098RR64 Rhizobium sp. H41
1676 A0A2A6JRW1 Rhizobium sp. J15
1677 A0A2A6I3L1 Rhizobium sp. L18
1678 A0A2A6HJS4 Rhizobium sp. L43
1679 A0A2A6L2I5 Rhizobium sp. L9
1680 A0A0M3BF13 Rhizobium sp. LC145
1681 A0A1S1TF47 Rhizobium sp. LCM 4573
1682 A0A0Q5D449 Rhizobium sp. Leaf306
1683 A0A0Q5VUY1 Rhizobium sp. Leaf371
1684 A0A0Q5X570 Rhizobium sp. Leaf383
1685 A0A0Q6FS00 Rhizobium sp. Leaf453
1686 A0A2A6LLB4 Rhizobium sp. M10
1687 L0NIB8 Rhizobium sp. NT-26
1688 A0A1V3FBZ9 Rhizobium sp. P44RR-XXIV
1689 A0A246S9G7 Rhizobium sp. R634
1690 A0A2D0B191 Rhizobium sp. R635
1691 A0A246TYN5 Rhizobium sp. R693
1692 A0A246U7E8 Rhizobium sp. R72
1693 A0A0Q6RH82 Rhizobium sp. Root1203
1694 A0A0Q6RIJ7 Rhizobium sp. Root1204
1695 A0A0Q6SC92 Rhizobium sp. Root1212
1696 A0A0Q6VVH4 Rhizobium sp. Root1220
1697 A0A0Q6ZP87 Rhizobium sp. Root1240
1698 A0A0Q7Y0T3 Rhizobium sp. Root149

1699 A0A0T1XQP6 Rhizobium sp. Root482
1700 A0A0Q7PNF3 Rhizobium sp. Root483D2
1701 A0A0Q8NV26 Rhizobium sp. Root708
1702 A0A1S1PKL4 Rhizobium sp. RSm-3
1703 A0A0D6SJW4 Rhizobium sp. UR51a
1704 A0A198YT22 Rhizobium sp. WYCCWR10014
1705 A0A1E4YB23 Rhizobium sp. YK2
1706 A0A095VI55 Rhizobium sp. YS-1r
1707 A0A1Q9A3U3 Rhizobium taibaishanense
1708 A0A1H8N0J7 Rhizobium tibeticum
1709 L0LI23 Rhizobium tropici CIAT 899
1710 A0A177QCR3 Rhizomicrobium sp. SCGC AG-212-E05
1711 A0A1V3P819 Rhodanobacter sp. C05
1712 A0A154QJV9 Rhodanobacter thiooxydans
1713 D5ANB4 Rhodobacter capsulatus (strain ATCC BAA-309 / NBRC 16581 / SB1003)
1714 A0A0E2PBB7 Rhodobacter capsulatus Y262
1715 A0A1E3D2B6 Rhodobacter johrii
1716 A0A239D7T4 Rhodobacter megalophilus
1717 A0A1H4KA45 Rhodobacter sp. 24-YEA-8
1718 A0A0R2WMN Rhodobacter sp. BACL10 MAG-120419-bin15
1719 A0A0R2TNQ6 Rhodobacter sp. BACL10 MAG-120910-bin24
1720 A0A0R2TJK4 Rhodobacter sp. BACL10 MAG-121220-bin24
1721 V7EFI8 Rhodobacter sp. CACIA14H1
1722 A0A1D9MGL7 Rhodobacter sp. LPB0142
1723 Q3J3S2 Rhodobacter sphaeroides (strain ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158)
1724 A0A1N7JM58 Rhodobacter vinaykumarii
1725 A0A2A4NTN5 Rhodobacteraceae bacterium
1726 A0A1B6YIN3 Rhodobacteraceae bacterium EhC02
1727 A0A1G3FSR3 Rhodobacteraceae bacterium GWE1_64_9
1728 A0A1G3FZY3 Rhodobacteraceae bacterium GWF1_65_7
1729 A0A257EWN4 Rhodobacteraceae bacterium PARR1
1730 A0A257Z7B2 Rhodobacterales bacterium 12-65-15
1731 A0A258F267 Rhodobacterales bacterium 32-66-7
1732 A0A259M970 Rhodobacterales bacterium 34-62-10
1733 A0A1Y5HN86 Rhodobacterales bacterium 56_14_T64
1734 A0A1Q4CU76 Rhodobacterales bacterium 65-51
1735 A0A1G3GE61 Rhodobacterales bacterium RIFCSPHIGHO2_02_FULL_62_130
1736 A0A2A5ETK7 Rhodobiaceae bacterium
1737 A0A1F2Q4G2 Rhodococcus erythropolis (Arthrobacter picolinophilus)
1738 T5IF99 Rhodococcus erythropolis DN1
1739 A0A0U3JJP0 Rhodococcus erythropolis R138
1740 C3JU66 Rhodococcus erythropolis SK121
1741 C0ZQE0 Rhodococcus erythropolis (strain PR4 / NBRC 100887)
1742 A0A165LS60 Rhodococcus fascians D188
1743 I0W6I4 Rhodococcus imtechensis RKJ300 = JCM 13270

1744 Q0SC99 Rhodococcus jostii (strain RHA1)
1745 A0A1B1KET7 Rhodococcus opacus (Nocardia opaca)
1746 W8GQ03 Rhodococcus opacus PD630
1747 A0A069J9P1 Rhodococcus qingshengii
1748 A0A2A3GIH4 Rhodococcus sp. ACPA1
1749 A0A2A3EYS1 Rhodococcus sp. ACS1
1750 A0A0D8HSX4 Rhodococcus sp. AD45
1751 A0A1W7QA5 Rhodococcus sp. Br-6
1752 N1LYT5 Rhodococcus sp. EsD8
1753 J1R5I6 Rhodococcus sp. JVH1
1754 A0A150A6S1 Rhodococcus sp. LB1
1755 U0FDH0 Rhodococcus sp. P27
1756 A0A143Q7E3 Rhodococcus sp. PBTS 1
1757 A0A143QHT7 Rhodococcus sp. PBTS 2
1758 A0A135GTE9 Rhodococcus sp. SC4
1759 L2TBG3 Rhodococcus wratislaviensis IFP 2016
1760 X0PQA1 Rhodococcus wratislaviensis NBRC 100605
1761 A0A037UQ60 Rhodomicrobium udaipurense JA643
1762 A0A255R6Q8 Rhodopirellula sp. MGV
1763 A0A0D7EKS1 Rhodopseudomonas palustris
1764 Q6N5K0 Rhodopseudomonas palustris (strain ATCC BAA-98 / CGA009)
1765 Q07HP2 Rhodopseudomonas palustris (strain BisA53)
1766 Q13CT2 Rhodopseudomonas palustris (strain BisB5)
1767 E6VER5 Rhodopseudomonas palustris (strain DX-1)
1768 B3QAR9 Rhodopseudomonas palustris (strain TIE-1)
1769 A0A1H8TEL4 Rhodopseudomonas pseudopalustris
1770 A0A0N1C912 Rhodopseudomonas sp. AAP120
1771 A0A095CR32 Rhodovulum sp. NI22
1772 A0A1Z9TEMC Rickettsiales bacterium TMED269
1773 A0A173SRD8 Roseburia intestinalis
1774 A0A1Y5RXJ7 Roseisalinus antarcticus
1775 W4HPV7 Roseivivax atlanticus
1776 X7EKU0 Roseivivax halodurans JCM 10272
1777 X7F2C3 Roseivivax isopora LMG 25204
1778 A0A1X6ZHX4 Roseivivax jejudonensis
1779 Q161I2 Roseobacter denitrificans (strain ATCC 33942 / OCh 114) (Erythrobacter sp. (strain OCh 114)) (Roseobacter denitrificans)
1780 A3X3D7 Roseobacter sp. MED193
1781 A0A1J5LNQ1 Roseobacter sp. MedPE-SWchi
1782 A4EWR1 Roseobacter sp. SK209-2-6
1783 A0A1L7AF21 Roseomonas gilardii
1784 A0A1S8D874 Roseomonas mucosa
1785 A0A2C7AGT4 Roseomonas rhizosphaerae
1786 A0A1V2H3V2 Roseomonas sp. M3
1787 A0A1Q2YXL0 Roseomonas sp. TAS13
1788 A0A1X7BXA6 Roseovarius aestuarii

1789 A0A0T5NZX0 *Roseovarius atlanticus*
1790 A0A1X6ZMI7 *Roseovarius gaetbuli*
1791 A0A1X6YYU5 *Roseovarius halotolerans*
1792 A0A0T5P652 *Roseovarius indicus*
1793 A0A1V0RJX9 *Roseovarius mucosus*
1794 A0A1V1UE16 *Roseovarius* sp. A-2
1795 A0A176FG89 *Roseovarius* sp. HI0049
1796 A6E3Q6 *Roseovarius* sp. TM1035
1797 A0A0L6CVV1 *Roseovarius tolerans*
1798 A0A017HPX6 *Rubellimicrobium mesophilum* DSM 19309
1799 S9QZ20 *Rubellimicrobium thermophilum* DSM 16684
1800 A0A1H4A5D9 *Rubrimonas cliftonensis*
1801 F3LLE8 *Rubrivivax benzoatilyticus* JA2 = ATCC BAA-35
1802 A0A238KXG2 *Ruegeria arenilitoris*
1803 A0A0P1E894 *Ruegeria atlantica*
1804 A0A0X3TXF1 *Ruegeria marisrubri*
1805 A0A1X7A821 *Ruegeria meonggei*
1806 Q5LV92 *Ruegeria pomeroyi* (strain ATCC 700808 / DSM 15171 / DSS-3) (*Silicibacter pomeroyi*)
1807 A0A0X3TSF5 *Ruegeria profundus*
1808 A0A0C1J220 *Ruegeria* sp. ANG-R
1809 A0A0C1J930 *Ruegeria* sp. ANG-S4
1810 A0A0P1IH18 *Ruegeria* sp. CECT 5091
1811 A0A1E3DEM1 *Ruegeria* sp. PBVC088
1812 Q1GCS7 *Ruegeria* sp. (strain TM1040) (*Silicibacter* sp.)
1813 A0A174R2F8 [*Ruminococcus*] *torques*
1814 W7Y3I7 *Saccharicrinis fermentans* DSM 9555 = JCM 21142
1815 A3K8B3 *Sagittula stellata* E-37
1816 A0A1Q8SWX *Salinicola socius*
1817 A0A1Q8T807 *Salinicola* sp. MH3R3-1
1818 A0A1S2DGM1 *Salinicola* sp. MIT1003
1819 A0A0S2I4P7 *Salinivirga cyanobacteriivorans*
1820 A0A2A4XAM4 SAR86 cluster bacterium
1821 A0A0R2W5V4 SAR92 bacterium BAACL16 MAG-120619-bin48
1822 A0A0R2TUR5 SAR92 bacterium BAACL26 MAG-121220-bin70
1823 A0A139WUT1 *Scytonema hofmannii* PCC 7110
1824 A0A1I7EDC6 *Sedimentitalea nanhaiensis*
1825 A0A0X2PD81 *Serratia*
1826 A0A084YPF9 *Serratia grimesii*
1827 A0A0X8SJ20 *Serratia liquefaciens*
1828 A0A069CQX3 *Serratia liquefaciens* FK01
1829 A0A1Q4P008 *Serratia marcescens*
1830 Q07YF3 *Shewanella frigidimarina* (strain NCIMB 400)
1831 B8CPP0 *Shewanella piezotolerans* (strain WP3 / JCM 13877)
1832 A0A0C3QTN1 *Shewanella* sp. cp20
1833 A0A0P1FES5 *Shimia marina*

1834 A0A0M9EELC *Shimia* sp. SK013
1835 A0A021XH97 *Shinella* sp. DD12
1836 A0A1A9GB88 *Shinella* sp. HZN7
1837 A0A1L3LNT1 *Sinorhizobium americanum*
1838 A0A1L3L4T4 *Sinorhizobium americanum* CCGM7
1839 C3K6K6 *Sinorhizobium fredii* (strain NBRC 101917 / NGR234)
1840 A0A0T6ZNB2 *Sinorhizobium fredii* USDA 205
1841 I3X279 *Sinorhizobium fredii* USDA 257
1842 H0G851 *Sinorhizobium meliloti* CCNWSX0020
1843 F7XC84 *Sinorhizobium meliloti* (strain SM11)
1844 A0A178XHH4 *Sinorhizobium saheli*
1845 A0A1V3R4M6 *Sinorhizobium* sp. A49
1846 A0A2A6P936 *Sinorhizobium* sp. BJ1
1847 A0A0T6ZSL8 *Sinorhizobium* sp. GL28
1848 A0A226CPW4 *Sinorhizobium* sp. LM21
1849 A0A2A6MJ16 *Sinorhizobium* sp. NG07B
1850 A0A0T6XVS1 *Sinorhizobium* sp. Sb3
1851 A0A0D2VRC6 *Skermanella aerolata* KACC 11604
1852 W9GZH9 *Skermanella stibioresistens* SB22
1853 A0A235E3I4 *Sphingobacterium cellulitidis*
1854 A0A1R4KN13 *Sphingobacterium faecium* PCAi_F2.5
1855 A0A239XWS6 *Sphingobacterium mizutaii*
1856 A0A1V3MYC6 *Sphingobacterium* sp. CZ-UAM
1857 A0A077XW28 *Sphingobacterium* sp. PM2-P1-29
1858 D4Z350 *Sphingobium japonicum* (strain NBRC 101211 / UT26S)
1859 A0A013WVX6 *Sphingobium* sp. Ant17
1860 A0A173KVC3 *Sphingobium* sp. EP60837
1861 A0A0Q4I429 *Sphingobium* sp. Leaf26
1862 A0A257DKE9 *Sphingomonadaceae* bacterium PASS1
1863 A0A259IAD6 *Sphingomonadales* bacterium 39-62-4
1864 A0A1M3I6E9 *Sphingomonadales* bacterium 63-6
1865 A0A1G3K345 *Sphingomonadales* bacterium GWF1_63_6
1866 A0A1G3JW06 *Sphingomonadales* bacterium RIFCSPLOWO2_12_FULL_63_15
1867 A0A245ZHX8 *Sphingomonas dokdonensis*
1868 A0A0C5L545 *Sphingomonas hengshuiensis*
1869 A0A239GD87 *Sphingomonas laterariae*
1870 A0A245ZS72 *Sphingomonas mucosissima*
1871 A0A031JEB0 *Sphingomonas paucimobilis* (*Pseudomonas paucimobilis*)
1872 A0A147IZK6 *Sphingomonas sanguinis*
1873 A0A1M2ZJJ3 *Sphingomonas* sp. 66-10
1874 A0A0M2LVJ2 *Sphingomonas* sp. Ag1
1875 A0A1X9YCS5 *Sphingomonas* sp. KC8
1876 A0A0Q4DKY6 *Sphingomonas* sp. Leaf10
1877 A0A0Q4G6K6 *Sphingomonas* sp. Leaf20
1878 A0A0Q4EXT7 *Sphingomonas* sp. Leaf208

1879 A0A0Q4HXJ3 *Sphingomonas* sp. Leaf22
1880 A0A0S9CIS6 *Sphingomonas* sp. Leaf231
1881 A0A0Q4L8Y2 *Sphingomonas* sp. Leaf34
1882 A0A0Q5PA79 *Sphingomonas* sp. Leaf343
1883 A0A0Q4L2Z1 *Sphingomonas* sp. Leaf38
1884 A0A0T0PUE0 *Sphingomonas* sp. Leaf412
1885 A0A0Q4QA4C *Sphingomonas* sp. Leaf67
1886 A0A196MU22 *Sphingomonas* sp. TDK1
1887 A0A0W1QLG *Sphingomonas* sp. WG
1888 A0A2A4B2I1 *Sphingomonas* spermidinifaciens
1889 A0A0X8R772 *Sphingopyxis* granuli
1890 A0A246K3M4 *Sphingopyxis* witflariensis
1891 A0A220W7L8 *Sphingorhabdus* flavimaris
1892 A0A1V5WIK4 *Spirochaetes* bacterium ADurb.Bin215
1893 A0A1V5HU64 *Spirochaetes* bacterium ADurb.BinA120
1894 A0A0U1KTE0 *Sporomusa* sp. An4
1895 A0A246HND7 *Stenotrophomonas maltophilia* (*Pseudomonas maltophilia*) (*Xanthomonas maltophilia*)
1896 A0A023Y3E2 *Stenotrophomonas* rhizophila
1897 A0A1D8YF00 *Stenotrophomonas* sp. LM091
1898 B8L7U8 *Stenotrophomonas* sp. SKA14
1899 Q092U4 *Stigmatella aurantiaca* (strain DW4/3-1)
1900 A0A239X4J1 *Streptococcus acidominimus*
1901 A0A0H1LVR0 *Streptococcus agalactiae*
1902 A0A0G6P5S0 *Streptococcus equi* subsp. *equi*
1903 B4U3D8 *Streptococcus equi* subsp. *zooepidemicus* (strain MGCS10565)
1904 A0A239ST89 *Streptococcus merionis*
1905 Q8DTV5 *Streptococcus mutans* serotype c (strain ATCC 700610 / UA159)
1906 A0A2A4KXG5 *Streptococcus parauberis*
1907 M3IZ37 *Streptococcus parauberis* KRS-02083
1908 A0A0Y4GBH0 *Streptococcus pneumoniae*
1909 F8LIY3 *Streptococcus salivarius* (strain CCHSS3)
1910 F8LQ12 *Streptococcus salivarius* (strain JIM8777)
1911 A0A139P670 *Streptococcus* sp. DD11
1912 A0A139NLY0 *Streptococcus* sp. DD12
1913 A0A139NRA7 *Streptococcus* sp. DD13
1914 T0UHV3 *Streptococcus* sp. HSISS1
1915 T0V0E1 *Streptococcus* sp. HSISS2
1916 T0V5Y5 *Streptococcus* sp. HSISS3
1917 A0A116LKQ3 *Streptococcus suis*
1918 R5ZW35 *Streptococcus thermophilus* CAG:236
1919 A0A117E7E7 *Streptomyces acidiscabies*
1920 A0A1Z1WPQ *Streptomyces alboflavus*
1921 X0P2D6 *Streptomyces albulus* PD-1
1922 A0A0X7JDU9 *Streptomyces albus* subsp. *albus*
1923 S3ZE18 *Streptomyces aurantiacus* JA 4570

1924 A0A0K9XKY2 Streptomyces caatingaensis
1925 A0A0C5G7M Streptomyces cyaneogriseus subsp. noncyanogenus
1926 N0D5B0 Streptomyces fulvissimus DSM 40593
1927 A0A140DJY8 Streptomyces gandocaensis
1928 A0A1B1B067 Streptomyces griseochromogenes
1929 A0A0L8MPD1 Streptomyces griseoflavus
1930 D9XU43 Streptomyces griseoflavus Tu4000
1931 A0A0X3X681 Streptomyces griseus subsp. griseus
1932 D9WQJ9 Streptomyces himastatinicus ATCC 53653
1933 A0A0F7VM75 Streptomyces leeuwenhoekii
1934 A0A086N889 Streptomyces luteus
1935 A0A1D7VM3 Streptomyces lydicus
1936 A0A1E7M696 Streptomyces nanshensis
1937 V6JFK2 Streptomyces niveus NCIMB 11891
1938 A0A1E7KKR Streptomyces oceani
1939 A0A250V3K6 Streptomyces olivochromogenes
1940 A0A191USC4 Streptomyces parvulus
1941 A0A1Y2NKS5 Streptomyces platensis
1942 A0A1E7K6U3 Streptomyces qinglanensis
1943 A0A0A0NFP7 Streptomyces rapamycinicus NRRL 5491
1944 A0A0U5HFC4 Streptomyces reticuli
1945 A0A0M8R3U Streptomyces rimosus subsp. pseudoverticillatus
1946 A0A0N0S725 Streptomyces rimosus subsp. rimosus
1947 A0A124C331 Streptomyces scabiei
1948 A0A1P8TQR Streptomyces silaceus
1949 A0A0W7WSC Streptomyces silvensis
1950 A0A1D2IJJ3 Streptomyces sp. AVP053U2
1951 A0A1Q5DVU Streptomyces sp. CB01249
1952 A0A1Q5CL17 Streptomyces sp. CB02414
1953 A0A1Q5LQM Streptomyces sp. CB02488
1954 A0A1Q4WE4 Streptomyces sp. CB02923
1955 A0A1Q4YR1C Streptomyces sp. CB03578
1956 A0A151DH83 Streptomyces sp. CC71
1957 A0A0N6ZT93 Streptomyces sp. CCM_MD2014
1958 A0A1K2FMI7 Streptomyces sp. F-1
1959 A0A0S1UGE Streptomyces sp. FR-008
1960 A0A0N0SSD Streptomyces sp. H036
1961 B4UYT5 Streptomyces sp. Mg1
1962 A0A1V2RPTC Streptomyces sp. MP131-18
1963 A0A0M8WZ2 Streptomyces sp. NRRL B-3648
1964 A0A0X3V258 Streptomyces sp. NRRL F-4489
1965 A0A0N0SA61 Streptomyces sp. NRRL F-5755
1966 A0A0N0YWV Streptomyces sp. NRRL F-6602
1967 A0A0N0B1J8 Streptomyces sp. NRRL F-7442
1968 A0A0X3X4J0 Streptomyces sp. NRRL S-1521

1969 A0A0N1K9U5 *Streptomyces* sp. NRRL WC-3753
1970 A0A069K0H7 *Streptomyces* sp. NTK 937
1971 A0A1B9EIG0 *Streptomyces* sp. PTY08712
1972 A0A1A9IUG9 *Streptomyces* sp. SAT1
1973 A0A022MG25 *Streptomyces* sp. Tu 6176
1974 A0A0M8SV66 *Streptomyces* sp. WM4235
1975 A0A0M8S759 *Streptomyces* sp. WM6368
1976 A0A0N0SCU6 *Streptomyces* sp. WM6372
1977 A0A0L8QS22 *Streptomyces varsoviensis*
1978 A0A0L8MWR *Streptomyces virginiae*
1979 A0A101NPK3 *Streptomyces yokosukanensis*
1980 A0A073IF04 *Sulfitobacter donghicola* DSW-25 = KCTC 12864 = JCM 14565
1981 A0A196R0Z5 *Sulfitobacter geojensis*
1982 A0A061ST62 *Sulfitobacter mediterraneus*
1983 A0A221K285 *Sulfitobacter pseudonitzschiae*
1984 A0A1J0WJ09 *Sulfitobacter* sp. AM1-D1
1985 A0A085BRZ7 *Sulfitobacter* sp. CB2047
1986 A0A196NMT8 *Sulfitobacter* sp. EhC04
1987 A0A166T1Z3 *Sulfitobacter* sp. HI0023
1988 A0A167HP95 *Sulfitobacter* sp. HI0054
1989 A0A176HXM2 *Sulfitobacter* sp. HI0076
1990 A0A176IEM4 *Sulfitobacter* sp. HI0082
1991 A3SVT9 *Sulfitobacter* sp. (strain NAS-14.1)
1992 A0A0L8V7R6 *Sunxiuqinia dokdonensis*
1993 A0A1V4XGE4 Syntrophaceae bacterium PtaB.Bin095
1994 A0A1D3URL4 *Tannerella forsythia* (*Bacteroides forsythus*)
1995 A0A109ZY28 *Tardiphaga robiniae*
1996 A0A1H6MY42 *Tardiphaga* sp. OK245
1997 A0A239JW73 *Tardiphaga* sp. OK246
1998 A0A1P8MS08 *Tateyamaria omphalii*
1999 A0A0B4CAC3 *Tateyamaria* sp. ANG-S1
2000 A0A098GAL2 *Tatlockia micdadei* (*Legionella micdadei*)
2001 A0A238UE61 *Tenacibaculum jejuense*
2002 A0A0J6IW53 *Testudinibacter aquarius*
2003 W6K428 *Tetrasphaera australiensis* Ben110
2004 A0A077M3T4 *Tetrasphaera japonica* T1-X7
2005 A0A0P1FIN3 *Thalassobius gelatinovor* (*Ruegeria gelatinovora*)
2006 A0A0P1FSF4 *Thalassobius mediterraneus*
2007 A0A154L810 *Thalassospira lucentensis*
2008 A0A074JQ98 *Thalassospira permensis* NBRC 106175
2009 A0A199YG65 *Thalassospira profundimaris*
2010 A0A0D8CM43 *Thalassospira* sp. HJ
2011 A0A154LCG1 *Thalassospira* sp. MCCC 1A01148
2012 A0A1X4N4U6 *Thalassospira* sp. MCCC 1A01428
2013 A0A154VF20 *Thalassospira* sp. MCCC 1A02898

2014 A0A1X4MZA6 Thalassospira sp. MCCC 1A03138
2015 A0A136HUG4 Thalassospira sp. Nap_22
2016 A0A1Q4UMP1 Thalassospira sp. TSL5-1
2017 A0A199Y2W6 Thalassospira tepidiphila MCCC 1A03514
2018 A0A154KSG4 Thalassospira xiamenensis
2019 R7RUU8 Thermobrachium celere DSM 8682
2020 E8PLX2 Thermus scotoductus (strain ATCC 700910 / SA-01)
2021 A0A0B0SF43 Thermus sp. 2.9
2022 G8N8V5 Thermus sp. CCB_US3_UF1
2023 W2U5E2 Thermus sp. NMX2.A1
2024 Q72KS9 Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039)
2025 B8GT11 Thioalkalivibrio sulfidiphilus (strain HL-EbGR7)
2026 A0A2A4CN89 Thioclava arenosa
2027 A0A1T2AMX5 Thioclava marina
2028 A0A1S6IHX6 Thioclava nitratireducens
2029 A0A074K0M8 Thioclava pacifica DSM 10166
2030 A0A1T2B248 Thioclava sp. DLFJ4-1
2031 A0A1T2BAI7 Thioclava sp. DLFJ5-1
2032 A0A1W9XLR1 Thioclava sp. EIOx9
2033 A0A1T2A5H2 Thioclava sp. F28-4
2034 A0A254UQA8 Thioclava sp. JM3
2035 A0A1T2BXU1 Thioclava sp. L04-15
2036 D6CNE5 Thiomonas arsenitoxydans (strain DSM 22701 / CIP 110005 / 3As)
2037 A0A238D0E3 Thiomonas delicata (Thiomonas cuprina)
2038 A0A078BIV3 Thiomonas sp. CB2
2039 A0A160SPN4 Thiomonas sp. CB3
2040 A0A238DI14 Thiomonas sp. X19
2041 A0A1Z9U3J3 Thiotrichales bacterium TMED285
2042 A0A162KYF5 Tistrella mobilis
2043 I3TJC5 Tistrella mobilis (strain KA081020-065)
2044 F2NW26 Treponema succinifaciens (strain ATCC 33096 / DSM 2489 / 6091)
2045 A0A1Y5RU04 Tropicibacter litoreus R37
2046 A0A0N7M0Q4 Tropicibacter multivorans
2047 A0A238JFC7 Tropicibacter phthalicus
2048 A0A239K2A8 Tropicimonas sediminicola
2049 A0A1C6JX26 uncultured Anaerotruncus sp.
2050 A0A1C6HKF7 uncultured Bacteroides sp.
2051 A0A1C6G057 uncultured Blautia sp.
2052 A0A1C5WZV1 uncultured Clostridium sp.
2053 A0A1C5SG29 uncultured Coprococcus sp.
2054 A0A1C5Q3J1 uncultured Dorea sp.
2055 A0A1C6HH82 uncultured Eubacterium sp.
2056 A0A1C6IY35 uncultured Lachnospira sp.
2057 A0A075GLM4 uncultured marine group II/III euryarchaeote KM3_177_C07
2058 A0A075GZT7 uncultured marine group II/III euryarchaeote KM3_35_H09

2059 A0A1C5PT61 uncultured Ruminococcus sp.
2060 A0A1Y5PUT5 uncultured Sphingopyxis sp.
2061 A0A1Y5Q6P0 uncultured Stenotrophomonas sp.
2062 A0A1X6WK7 Vagococcus fluvialis bH819
2063 A0A0S3PSDC Variibacter gotjawalensis
2064 A0A250DM92 Variovorax boronicumulans
2065 A0A0N8PAS1 Variovorax paradoxus
2066 T1XH04 Variovorax paradoxus B4
2067 C5CP86 Variovorax paradoxus (strain S110)
2068 A0A126ZJR0 Variovorax sp. PAMC 28711
2069 A0A1I5W9H0 Variovorax sp. PDC80
2070 A0A0Q6NAU Variovorax sp. Root318D1
2071 A0A0Q7CE9 Variovorax sp. Root411
2072 A0A0Q7H2W Variovorax sp. Root434
2073 A0A0Q7LA23 Variovorax sp. Root473
2074 A0A1E4INS6 Variovorax sp. SCN 67-85
2075 A0A109BPV2 Variovorax sp. WDL1
2076 A0A1V6JU22 Verrucomicrobia bacterium ADurb.Bin006
2077 A0A1V6JCM4 Verrucomicrobia bacterium ADurb.Bin018
2078 A0A1V6GRFC Verrucomicrobia bacterium ADurb.Bin070
2079 A0A1W6VLV Vibrio alginolyticus
2080 U3AN41 Vibrio azureus NBRC 104587
2081 A0A193KB01 Vibrio breoganii
2082 A0A1C3JAB8 Vibrio celticus
2083 A0A1E5D9B3 Vibrio cyclitrophicus FF160
2084 A0A066UX83 Vibrio fortis
2085 A0A0B8NQX Vibrio ishigakensis
2086 A0A0L1L557 Vibrio lentus
2087 A0A090TR17 Vibrio maritimus
2088 U4KI36 Vibrio nigripulchritudo
2089 A0A090P688 Vibrio ponticus
2090 A0A0C2NV18 Vibrio renipiscarius
2091 A0A1B1NP12 Vibrio scophthalmi
2092 A6CVK8 Vibrio shilonii AK1
2093 A0A1R4K7U3 Vibrio sp. JB196
2094 A0A0B8PKL5 Vibrio sp. JCM 19236
2095 A0A2C6ML20 Vibrio sp. PID17_43
2096 A0A1C3II24 Vibrio splendidus
2097 A0A1E5FVQ4 Vibrio splendidus 12E03
2098 A0A1E5ESR7 Vibrio splendidus 1F-157
2099 A0A1E5AXQ5 Vibrio splendidus ZS-139
2100 A0A109D629 Vibrio toranzoniae
2101 F9T9N2 Vibrio tubiashii ATCC 19109
2102 A0A090SJM3 Vibrio variabilis
2103 A0A1W6MBX Vibrio vulnificus

- 2104 A0A024QGQ: Virgibacillus massiliensis
- 2105 W4F018 Viridibacillus arenosi FSL R5-213
- 2106 A0A0K1PGS: Vulgatibacter incomptus
- 2107 A0A1L6R8U4 Weissella jogaejeotgali
- 2108 A0A0K0XV16 Wenzhouxiangella marina
- 2109 A0A2A3LF05 Xanthomonadaceae bacterium NML93-0399
- 2110 A0A2A3VLL2 Xanthomonadaceae bacterium NML93-0792
- 2111 A0A222WGX: Xanthomonas citri pv. mangiferaeindicae (agent of mango bacterial black spot)
- 2112 A0A0N8HXI0 Xanthomonas sp. Mitacek01
- 2113 A0A1Y2SKJ5 Xenorhabdus vietnamensis
- 2114 A0A1B1PSM: Yangia sp. CCB-MM3
- 2115 A0A0F4GRK: Zymoseptoria brevis

SCREEN SAMPLE OF UNIPROT PAGE FOR 6-AMINOHEXANOATE HYDROLASE:

The screenshot shows the UniProtKB search results for the query "6-aminohexanoate hydrolase". The search results are displayed in a table with columns for Entry, Entry name, Organism, Protein names, Gene names, and Length. The results are filtered to show 25 entries out of 5,019 total. The table includes entries from various organisms such as Flavobacterium sp. (strain K172), Pseudomonas sp. (strain NK87), Bacteroides cellulosilyticus, and Listeria monocytogenes. The protein names are listed as "6-aminohexanoate-cyclic-dimer hydrolase" and "6-aminohexanoate-dimer hydrolase". The gene names are listed as nylA, nylB, nylB', nylB'_1 BcellWH2_02191, nylB'_1 AA416_00538, nylA_2 ERS409707_02665, and FM121_01050. The lengths of the proteins range from 392 to 723 amino acids.

Entry	Entry name	Organism	Protein names	Gene names	Length
P13398	NYLA_FLASK	Flavobacterium sp. (strain K172)	6-aminohexanoate-cyclic-dimer hydrolase	nylA	493
P07061	NYLB_FLASK	Flavobacterium sp. (strain K172)	6-aminohexanoate-dimer hydrolase	nylB	392
P13397	NYLA_PSES8	Pseudomonas sp. (strain NK87)	6-aminohexanoate-cyclic-dimer hydrolase	nylA	493
P07062	NYLC_FLASK	Flavobacterium sp. (strain K172)	6-aminohexanoate-dimer hydrolase	nylB'	392
A0A0P0GQ77	A0A0P0GQ77_9BACE	Bacteroides cellulosilyticus	6-aminohexanoate-dimer hydrolase	nylB'_1 BcellWH2_02191	576
A0A108TGJ3	A0A108TGJ3_9BACE	Bacteroides cellulosilyticus	6-aminohexanoate-dimer hydrolase	nylB'_1 AA416_00538	576
A0A0Y7KNZ2	A0A0Y7KNZ2_LISMN	Listeria monocytogenes	6-aminohexanoate-cyclic-dimer hydrolase	nylA_2 ERS409707_02665	723
A0A1X6WK75	A0A1X6WK75_9ENTE	Vagococcus fluvialis	6-aminohexanoate-	FM121_01050	638

At the bottom of the page, there is a notice: "We'd like to inform you that we have updated our Privacy Notice to comply with Europe's new General Data Protection Regulation (GDPR) that applies since 25 May 2018." There is also a "Do not show this banner again" option.

Testing the Hypothesis that the Nylonase NylB Protein
Arose *de novo* via a Frameshift Mutation

Salvador Cordova and John Sanford

SupplementaryTable S6

List of proteins that were listed as both a beta lactamase and a 6-aminohexanoate hydrolase nylonase
in GenBank that were determined by searching for NylB homologs using psi-BLAST

Organism with Nylonase/Beta Lactamase	Accession Number As Beta Lactamase	Accession Number as 6-Aminohexanoate Hydrolase Nylonase	Comment
Bradyrhizobium sp. ORS 375	CCD97170.1	CCD97170.1	listed as both 6AH Nylonase and Beta Lactamase simultaneously
Candidatus Puniceispirillum marinum IMC ADE39034.1	ADE39034.1	ADE39034.1	listed as both 6AH Nylonase and Beta Lactamase simultaneously
E. Coli	SCQ13749.1	SCQ13749.1	protein listed as a Beta Lactamase, but gene listed as nylB nylonase
Delftia acidovorans	KFJ13307.1	WP_034393572.1	6AH Nylonase 100% identical with Beta Lactamase
Comamonas thiooxydans	CUA91648.1	WP_034404398.1	6AH Nylonase 100% identical with Beta Lactamase
Comamonas testosteroni	EFI60192.1	WP_034362870.1	6AH Nylonase 100% identical with Beta Lactamase
Sinorhizobium sp. RAC02	AOF94051.1	WP_069063755.1	6AH Nylonase 100% contained inside Beta Lactamase
Burkholderia singularis	SMG00640.1	WP_089340811.1	6AH Nylonase 100% contained inside Beta Lactamase
Burkholderia sp. yr281	SDJ12107.1	WP_091788686.1	6AH Nylonase 100% contained inside Beta Lactamase
Bradyrhizobium arachidis	SFV17448.1	WP_092218251.1	6AH Nylonase 100% contained inside Beta Lactamase

Testing the Hypothesis that the Nylonase NylB Protein
Arose *de novo* via a Frameshift Mutation

Salvador Cordova and John Sanford

Supplementary Table S7

List of proteins that were listed as both an amidase and a 6-aminohexanoate hydrolase nylonase
in GenBank that were determined by searching for NylA homologs using psi-BLAST

Organism with Nylonase	Accession Number as Amidase	Accession Number as 6-Aminohexanoate Hydrolase Nylonase	Comment
Rhodococcus erythropolis	WP_031330824.1	EQM35287.1	100% identical
Sphingomonas sp. KC8	WP_010125486.1	ARS27527.1	100% identical
Paenibacillus amylolyticus	WP_100527413.1	PJN62497.1	100% identical
Pseudomonas syringae	WP_020342474.1	EPN22118.1	100% identical
Marinobacter hydrocarbonoclasticus ATCC 49840]	WP_041654449.1	CCG93690.1	Amidase 100% contained inside 6AH Nylonase, covers 99% of 6AH Nylonase
Thiomonas delicata	WP_094159176.1	SBP866 99.1	Amidase 100% contained inside 6AH Nylonase, covers 99% of 6AH Nylonase
Leptospira interrogans	WP_001136113.1	KWV26884.1	Amidase 100% contained inside 6AH Nylonase, covers 99% of 6AH Nylonase

Testing the Hypothesis that the Nylonase NylB Protein
Arose *de novo* via a Frameshift Mutation

Salvador Cordova and John Sanford

Supplementary Table S8

Screen captures of MUSCLE 6.0 alignment of 6 NylB and 4 NylB' proteins.

Proteins 1-8 have experimental support for nylonase activity.

Protein 9 (Paenibacillus sp AD87) had the lowest e-value in psi-Blast search for NylB in Flavobacterium KI72

Protein 10 (E. Coli) is from a well known organism

Highlighted columns have 80% or higher conservation across selected proteins.

Asterisks "*" above a column indicate 100% conservation

Distance Matrix at the bottom

Positions 1-51

Species/Abbrv	
1. NylB_Flavobacterium_KI72	-----M N ARSTGQH P ARY P GAAAGEPTLDSWQEAPHNRWAFAR L G E LL P
2. NylB_Prime_Flavobacterium_KI72	-----M N T P TTGSH P ARY P SAAAGEPTLDSWQE P PHNRWAF A HL G EM V P
3. NylB_Agromyces_KY5R	-----M N ARSTGQH P ARY P GAAAGEPTLDSWQEAPHNRWAFAR L G E LL P
4. NylB_Prime_Agromyces_KY5R	-----M N ARSTGQH P ARY P GAAAGEPTLDSWQEAPHNRWAFAR L G E LL P
5. NylB_Bacillus_cereus	-----M N -----FKQ-----
6. NylB_Prime_Bacillus_cereus	-----M D -----FKQ-----
7. NylB_Pseudomonas_aeruginosa_PA01	-----M L G L P L DRREG--PPMPA-LPLSRRALAALALVFALPALAEDWP
8. NylB_Pseudomonas_NK87	-----M N T V PP F RDPT-V P GNSH-I P RQD-WDRAPWNRWTFQHVRELLP
9. NylB_Prime_Paenibacillus_sp._AD87	-----M N -----PSSIPSMLH
10. NylB/beta-lactamase_E._coli	MIRKPLALALILAALPAAAMAQHCGSLTLD-VCPTPYDQ-TLPAAKDMLS

Positions 52-131

Spe	
1.	-----TAAVSR-RDPATPAEPVVR-LDALA-----TRL P D L EQ R LEETCTDAFL V L RGSEVLAE E Y R AGE
2.	-----SAAVSR-RPVNAPGHALAR-LGAIA-----AQL P D L EQ R LEQTYTDAFL V L RGTEVVAE E Y R AGE
3.	-----TAAVSR-RDPATPAEPVVR-LDALA-----TRL P D L EQ R LEETCTDAFL V L RGSEVLAE E Y R AGE
4.	-----TAAVSR-RDPATPAEPVVR-LDALA-----TRL P D L EQ R LEETCTDAFL V L RGTEVVAE E Y R AGE

5. -----LENKFEKKKVNTFLVYQKGKLTTEYYKTPE

6. -----LENKFEKKKVNTFLVYQKGELTTEYYKTPE

7. -----DTDWPRAQQPAGPA-----LEAFERYAFGRRDDV--RRKGI RTDAVVVIRDGRLVYERYAGPS

8. -----TTKVWRGTGPASPLPVDLRDIDAVSFAAEGQSHTIAGFLETSYADGFVLV LHGGKIVAERYLNGM

9. -----QIIPPLDLR---SC---LVSMRGEIIEHYRNQE

10. WDQTSRVIGFRNDYRNYAGDVFR-HGASTPLERA EKQLTDARYTLNGHTWNLQDYLKRENVSGMLVLKDGKVAWKYLAEGN

Positions 132-212

Spe	*	**	*	*				
1.	APDDRHLLMSVSKSLCGTVV	GALIDE	GRI-DPAQPV	TEYVPEL---	AGSVYDGPSVLQVLD	MQISIDYNE	EDY--	VDPASEV
2.	APDDRHLLMSVSKSLCGTVV	GALVDE	GRI-DPAQPV	TEYVPEL---	AGSVYDGPSVLQVLD	MQISIDYNE	EDY--	VDPASEV
3.	APDDRHLLMSVSKSLCGTVV	GALIDE	GRI-DPAQPV	TEYVPEL---	AGSVYDGPSVLQVLD	MQISIDYNE	EDY--	VDPASEV
4.	APDDRHLLMSVSKSLCGTVV	GALVDE	GRI-DPAQPV	TEYVPEL---	AGSVYDGPSVLQVLD	MQISIDYNE	EDY--	VDPASEV
5.	CTNNLYKINSITKSIVSLLIGIAID	KGYINDIHTPI	TEWIENV---	PEGKHD-LTLYHLL	TMTTGEEWKE	-F--	GNGVVF	P
6.	CANNLYKINSITKSIVSLLIGIAID	KGYINDIHTSI	TEWIENV---	PGEKHD-LTLYHLL	TMTTGEDWKE	-F--	GNGVVF	P
7.	RAETPHLTSVSKSLLATLLGVAEGE	GRF-QLDDP	VARYYP---	PFARHPEVTLRHL	LNWSSGLDWQE	DYEFAPLK	SSV	
8.	APHTQHLSQSVAKSVVGTVAGILID	RGVV-NPAALL	THYLPEL---	EATAYRGATVQHVL	DMTSGVVFDE	TY--	TALDSHM	
9.	AGTHIAKINSCTKSVLSALICIAMDH	GWLPEASTPI	STFFPQLVSDPD	PRKSEITLEQL	LTMTAGFNWDE	-F--	GGQNSFP	
10.	TDTTLWTSRSVGSVVSTLVGIAIQQ	GKIHS	LDDLITVYEPEL---	KGTAWDGVTLKQLIQHTS	GVEWNE	EDY--	TDPQSHF	

Positions 213-293

Spe	*	*	*	*		
1.	QTHDRSAGWRTRRDGD-PADTYEFLTTL	RG-DGGTG-EFQYCSANTDV	LAWIVERVTGL	RYVEALSTYLWAKLDADR	DATI	
2.	QTHGRSAGWRTRATGD-PADTYEFLTTL	RG-DGSTG-EFQYCSANTDV	LAWIVERVTGL	RYVEALSTYLWAKLDADR	DATI	
3.	QTHDRSAGWRTRRDGD-PADTYEFLTTL	RG-DGGTG-EFQYCSANTDV	LAWIVERVTGL	RYVEALSTYLWANVDADR	DATI	
4.	QTHGRSAGWRTRRHGD-PADTYEFLTTL	RG-DGSTG-EFQYCSANTDV	LAWIVERVTGL	RYVEALSTYLWAKLDADR	DATI	
5.	NDFVESANW-----	VQYILEKPIIEEPAT-KMNYNS	GSSHLLSYIIQKAT	GMPT	EQFAKRYL	FNPLEI-TEYEW

6.	NDFVESEN	-----	VQYILEKPIIEEPAT	-	KMNYNS	GS	SHLLSYIIQEAT	GM	STERFAKKY	LEDPLKI	-	NEYEW
7.	VAMLYTRG	-----	RDDMAGFVAETPQ	-	ARPPGRRFRYS	SG	DSNVLAAALHGML	GADYAEYPWRAL	FEPLGI	-	RS	AVW
8.	AQLDVASG	K	DSPNDWPTHVWDLILSLKDLECPHGASFRYS	SI	ETDVLAFVLQRAAAAPLAELVSREL	WAPMGAEEDAYF						
9.	R-MTRTDHW	-----	VNFALEQRLSHVPGT	-	YMEYNS	GV	SQILSAIILMQST	GISVAEFAERY	FGPLGI	-	QQ	YEW
10.	ARLTQCEA	----	HPGVYACVRKIVTGLAR	-	QHPAGEQWS	YS	GGAWLLGDILERAT	MSLAAWLEQAL	WQPAGMAHDGVW			

Positions 294-374

Spe	*	*	*	*	*	*	*	*	*	*	*					
1.	TVDQTGFG	-	FANGGV	SCTARD	LARVGR	MMLDGGVA	-	PGGRVV	SQGW	VESVLAGGS	REAMTDEGFTSAFPE	-	GSYTRQ	WW	--	
2.	TVDTTGFG	-	FAHGGV	SCTARD	LARVGR	MMLDGGVA	-	PGGRVV	SE	DVRRVLAGGS	SHEAMTDKGFTNTFPD	-	GSYTRQ	WW	--	
3.	TVDQTGFG	-	FANGGV	SCTARD	LARVGR	MMLDGGVA	-	PGGRVV	SQGW	VESVLAGGS	REAMTDEGFTSAFPE	-	GSYTRQ	WW	--	
4.	TVDTTGFG	-	FAHGGV	SCTARD	LARVGR	MMLDGGVA	-	PGGRVV	SE	DVRRVLAGGS	SHEAMTDKGFTNTFPD	-	GSYTRQ	WW	--	
5.	QQDSQGIY	-	VGGFG	MKMKSKD	LLKLGI	LCLQNGIW	-	QNEIV	SSKW	LEESS	-----	TPLFETYEHV	-	GAYGYH	WW	--
6.	QQDPQGIY	-	VGGFG	MKMKSTD	LLKLGI	LCLQNGYW	-	NEKEIV	SAKW	LEESS	-----	RAQFETYEHV	-	GAYGYH	WW	--
7.	ERDAAGTF	-	VGSSYV	YMSARD	LARVGL	LLMQRHGRW	-	RERQLL	PLAW	VEFNLT	PFARYRPSAEEEGEAVPG	-	GQ	---	WW	--
8.	TVDTAGYA	-	LGDGG	FNATLRD	YARFALLH	LRGGEI	-	DGRRIV	SPGW	IAATR	-	FGADPALFGDIYREALPA	-	GAYHNQ	FW	--
9.	ESDPQGVH	-	TGGFG	LKLLPED	LLKFG	QLFLQQGMW	-	NGQSMI	SSDHV	SRST	-----	KPAITVTPPNHGS	YAWH	WW	--	
10.	HAYQQGKH	D	VGAHGF	NATLED	WGRFGE	FVAR	DGRLSNGKQLV	PAGW	FDQAA	---	SWTKALNSVSAAHPE	-	GIYGYQ	WW	NN	

Positions 375-455

Spe	*	**										
1.	-C-----	TGNERG	-----	NVSGIGI	HGQNLWLD	PRTDSVIVKLS	SSWPD	--	PDTRHWHGL	QSGILLDVS	RALDAV	
2.	-C-----	TGNERG	-----	NVSGIGI	HGQNLWLD	PLTDSVIVKLS	SSWPD	--	PDTEHWHRL	QNGILLDVS	RALDAV	
3.	-C-----	TGNERG	-----	NVSGIGI	HGQNLWLD	PRTDSVIVKLS	SSWPD	--	PDTRHWHGL	QSGILLDVS	RALDAV	
4.	-C-----	TGNERG	-----	NVSGIGI	HGQNLWLD	PLTDSVIVKLS	SSWPD	--	PDTEHWHRL	QNGILLDVS	RALDAV	
5.	-V-----	LNNERF	-----	HIPYCMY	FAMGYGGQYIVVI	PQLEVVAVIS	SHMPK	---	RGLVPLK	LFI	-EHVQENYK	FV
6.	-V-----	LHNERF	-----	HIPYCIY	FAMGYGGQYIVVI	PQLEVVISIIS	SHMPK	---	RGLVPLN	LFI	-EHVQGNSNYI	

7. -L-----NRAVKGAPAPWPDVAETAFAALGHWGQALYVLP EEKLVIVRYADDRDGRFRHNEFLRLAQA AAFAGEARP-----
8. -I-----EDTARG-----AYMARGVFGQLIYIDPEADFAAVILSSWPE--FVSTTRLRTALAAVRAVREALSA--
9. -VGAYTNETAEAKF--TEQPTSILHYYARGFGGQFVYIVPALELVVVLTNDKRRK-----KEKPPLDVFPKQIAPQLAKLL
10. AI-----PANAQNVQPTPQEG LKGS LWALGIYGVVIMVNR AEHLVIVQWSTWPEQAEPSFNAQPLEAALMYS AIARELR--

MUSCLE Parameters:

Gap Open -2.9, Gap Extend 0, Hydrophobicity Multiplier 1.2, Max Memory in MB 487, Max Iterations, 8,
Clustering Method (Iteration 1,2) UPGMB, Clustering Method (Other Iterations) UPGMB, Min Diag Length (lambda).

MUSCLE Citation: Edgar, Robert C. (2004), MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Research 32(5), 1792-1797.

M6: Pairwise Distances (D:\muscle\supplementary_s9.meg)

	1	2	3	4	5	6	7	8	9	10
1. NylB CAA24927.1 Flavobacterium KI72										
2. NylB Prime CAA26616.1 Flavobacterium KI72	21.000									
3. BAE97621.1 NylB Agromyces KY5R	2.000	23.000								
4. BAE97630.1 NylB Prime Agromyces KY5R	19.000	2.000	21.000							
5. NylB SME41718.1 Bacillus cereus	209.000	212.000	210.000	212.000						
6. SME45458.1 NylB Prime Bacillus cereus	211.000	214.000	212.000	214.000	33.000					
7. NylB-like AAG07735.1 Pseudomonas aeruginosa PAO1	197.000	198.000	198.000	198.000	211.000	212.000				
8. NylB BAA01524.1 Pseudomonas NK87	175.000	174.000	175.000	176.000	206.000	213.000	196.000			
9. OAX47860.1 NylB Paenibacillus sp. AD87 7e-5	215.000	213.000	216.000	212.000	165.000	167.000	199.000	213.000		
10. SCQ13749.1 nylb gene/beta-lactamase protein Escherichia coli	203.000	208.000	203.000	208.000	207.000	206.000	194.000	190.000	208.000	

Table. Estimates of Evolutionary Divergence between Sequences

The number of amino acid differences per sequence from between sequences are shown. The analysis involved 10 amino acid sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 277 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 [1].

1. Tamura K., Stecher G., Peterson D., Filipski A., and Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution 30: 2725-2729.

>Testing the Hypothesis that the Nylonase NylB Protein
>Arose de novo via a Frameshift Mutation
>Salvador Cordova and John Sanford
>Supplementary Table S9

>NylB CAA24927.1 *Flavobacterium* KI72
MNARSTGQHPARYPGAAAGEPTLDSWQEAPHNRWAFARLGELLPTAAVSRRDPATPAEPV
VRLDALATRLPDLEQRLEETCTDAFLVLRGSEVLAEYYRAGFAPDDRHLLMSVSKSLCGT
VVGALIDEGRIDPAQPVTEYVPELAGSVYDGPSVLQVLDMQISIDYNEDYVDPASEVQTH
DRSAGWRTRRDGDPADTYEFLTTLRGDGGTGEFQYCSANTDVLAWIVERVTGLRYVEALS
TYLWAKLDADRDATITVDQTGFQFANGGVSTARDLARVGRMMLDGGVAPGGRVVSQGWV
ESVLAGGSREAMTDEGFTSAFPEGSYTRQWWCTGNERGNVSGIGIHGQNLWLDPRTDSVI
VKLSSWPDPDTRHWHGLQSGILLDVSRALDAV

>NylB Prime CAA26616.1 *Flavobacterium* KI72
MNTPTTGSHPARYPSSAAAGEPTLDSWQEPPHNRWAF AHLGEMVPSAAVSRRPVNPAGHAL
ARLGAIAAQLPDLEQRLEQTYTDAFLVLRGTEVVAEYYRAGFAPDDRHLLMSVSKSLCGT
VVGALVDEGRIDPAQPVTEYVPELAGSVYDGPSVLQVLDMQISIDYNEDYVDPASEVQTH
GRSAGWRTRATGDPADTYEFLTTLRGDGGTGEFQYCSANTDVLAWIVERVTGLRYVEALS
TYLWAKLDADRDATITVDTTGFGFAHGGVSTARDLARVGRMMLDGGVAPGGRVVSQGWV
RRVLAGGSHEAMTDKGFTNTFPDGSYTRQWWCTGNERGNVSGIGIHGQNLWLDPLTDSVI
VKLSSWPDPDTEHWHRLQNGILLDVSRALDAV

>BAE97621.1 NylB *Agromyces* KY5R
MNARSTGQHPARYPGAAAGEPTLDSWQEAPHNRWAFARLGELLPTAAVSRRDPATPAEPV
VRLDALATRLPDLEQRLEETCTDAFLVLRGSEVLAEYYRAGFAPDDRHLLMSVSKSLCGT
VVGALIDEGRIDPAQPVTEYVPELAGSVYDGPSVLQVLDMQISIDYNEDYVDPASEVQTH
DRSAGWRTRRDGDPADTYEFLTTLRGDGGTGEFQYCSANTDVLAWIVERVTGLRYVEALS
STYLVANVDADRDATITVDQTGFQFANGGVSTARDLARVGRMMLDGGVAPGGRVVSQGWV
VESVLAGGSREAMTDEGFTSAFPEGSYTRQWWCTGNERGNVSGIGIHGQNLWLDPRTDSVI
VKLSSWPDPDTRHWHGLQSGILLDVSRALDAV

>BAE97630.1 NylB Prime *Agromyces* KY5R
MNARSTGQHPARYPGAAAGEPTLDSWQEAPHNRWAFARLGELLPTAAVSRRDPATPAEPV
VRLDALATRLPDLEQRLEETCTDAFLVLRGTEVVAEYYRAGFAPDDRHLLMSVSKSLCGT
VVGALVDEGRIDPAQPVTEYVPELAGSVYDGPSVLQVLDMQISIDYNEDYVDPASEVQTH
GRSAGWRTRRHGDPADTYEFLTTLRGDGGTGEFQYCSANTDVLAWIVERVTGLRYVEALS
STYLVANVDADRDATITVDTTGFGFAHGGVSTARDLARVGRMMLDGGVAPGGRVVSQGWV
VESVLAGGSHEAMTDKGFTNTFPDGSYTRQWWCTGNERGNVSGIGIHGQNLWLDPLTDSVI
VKLSSWPDPDTEHWHRLQNGILLDVSRALDAV

>NylB SME41718.1 *Bacillus cereus*
MNFKQLENKFEKKKVNTFLVYQKGLTTEYYKTPECTNLYKINSITKSIVSLLIGIAIDKGYINDIHTP
ITEWIENVPEGKHDLTLYHLLMTTGEEWKEFGNGVVPNDVFESANWVQYILEKPIIEEPATKMNYNSG
SSHLLSYIIQKATGMPTEQFAKRYLFNPLEITEYEWQQDSQGIYVGGFGMKMKSKDLLKLGILCLQNGIW
QGNEIVSSKWLEESSTPLFETYEYHVGAYGYHWWVLNNERFHIPYCMYFAMGYGGQYIVVIPQLEVVAVIS
SHMPKRGLVPLKLFIEHVQENYKFV

>SME45458.1 NylB Prime *Bacillus cereus*
MDFKQLENKFEKKKVNTFLVYQKGLTTEYYKTPECANLYKINSITKSIVSLLIGIAIDKGYINDIHTS
ITEWIENVPEGKHDLTLYHLLMTTGEDWKEFGNGVVPNDVFESANWVQYILEKPIIEEPATKMNYNSG
SSHLLSYIIQKATGMPTEQFAKRYLFNPLEITEYEWQQDPQGIYVGGFGMKMKSTDLLKLGKLCQNGYW
NEKEIVSAKWLEESSRAQFETYEYHVGAYGYHWWVLNNERFHIPYCIYFAMGYGGQYIVVIPQLEVVAIS
SHMPKRGLVPLNLFIEHVQGNNSNYI

>NylB-like AAG07735.1 *Pseudomonas aeruginosa* PAO1
MLGLPLDRREGPPMPALPLSRRALALVFPALPAEDWPDTDWPRAPQPPAGPALEAFE
RYAFGRRDDVRRKGIRTDVVIRDGRLVYERYAGPSRAETPHLTWSVSKSLLATLLGVA
EGEGRFQLDDPVARYYPPFARHPEVTLRHLNWSGLDWQEDYEFAPLKSSVVAMLYTRG
RDDMAGFVAETPQARPPGRRFRYSSGDSNVLAAALHGMLGADYAEYPWRALFEPLGIRSA
VWERDAAGTFVGGSSVYMSARDLARVGLLMQRHGRWRERQLLPLAWVEFNLTFFARYRPS
AEEEGEAVPGGQWWLNRAVKGAPAPWPDVAETAFAALGHWGQALYVLPEEKLVIVRYADD
RDGRFRHNEFLRLAQAFAFAGEARP

>NylB BAA01524.1 *Pseudomonas* NK87
MNTVPPFRDPTVPGNSHIPRQDWDWRAPWNRWTFQHVRELLPTTKVWRGTGPASPLPVDLRDIDAVSFAAE
GQSHTIAGFLETSYADGFLVLHGKIVAERYLNGMAPHTQHLSQSVAKS VVGT VAGILIDRGVVNPAALL
THYLPELEATAYRGATVQHVLDMTSGVVFDETYTALDSHMAQLDVASGWKDSPNPDWPTHVWDLILSLKD
LECPHGASFRYRSIETDVLAFVLQRAAAAPLAELVSRELWAPMGAEEDAYFTVDTAGYALGDGGFNATLR
DYARFALLHLRGGEIDGRRIVSPGWIAATRFADPALFGDIYREALPAGAYHNQFWIEDTARGAYMARGV
FGQLIYIDPEADFAAVILSSWPEFVSTTRLRTALAAVRVREALSA

>OAX47860.1 NylB' *Paenibacillus* sp. AD87 7e⁻⁵
MNPSSIPSMHLHQIIPPLDLRSLVSMRGEIIEHYRNQEAGTHIAKINSCTKSIVLSALICIAMDHGWLPE
ASTPISTFFPQLVSDPDRKSEITLQLLTMTAGFNWDEFGGQNSFPRMTRTDHWNFALEQRLSHVPGT
YMEYNSGVSQILSAILMQSTGISVAEFAERYLFGPLGIQQYEWESDPQGVHTGGFGLKLLPEDLLKFGQL
FLQQGMWNGQSMISSDHVSRSTKPAITVTPPNHGSYAWHWWVGAYTNETAEAKFTEQPTSILHYYYARGF

GGQFVYIVPALELVVLTNDKRKKEKPPLDVFPKQIAPQLAKLL

>SCQ13749.1 nylb gene/beta-lactamase protein Escherichia coli

MIRKPLALALILAALPAAAMAQHCGSLTDVCPTPYDQTLPAAKDMLSWDQTSRVIGFRNDYRNYAGDVF
RHGASTPLERAEKQLTDARYTLNGHTWNLQDYLKRENVSGMLVLKDGKVAWKYLAEGNTDRTLWTSRSVG
KSVVSTLVGIAIQQGKIHSLLDITVYEPELKGTAWDGVTLKQLIQHTSGVEWNEDYTDQSHFARLTQC
EAHPGVYACVRKIVTGLARQHPAGEQWSYSSGGAWLLGDILERATGMSLAAWLEQALWQPAGMAHDGVWH
AYQQGKHDVGAHGFNATLEDWGRFGEFVARDGRLSNGKQLVPAGWFDQAASWTKALNSVSAAHPEGIYGY
QWWNNAIPANAQNVQPTPQEGLKGSWALGIYGQVIMVNR AEHLVIVQWSTWPQAEPFNAQPLEAALMY
SAIARELR

nylonase_supplemental_preprints_v1.pdf (4.55 MiB)

[view on ChemRxiv](#) • [download file](#)
