

Ce projet nous a amené à étudier la séquence d'ADN TARA_7S 3634040 récoltée en mer Méditerranée (station 7), à 5 mètres de profondeur près de Tanger. Le fragment d'ADN récolté a été initialement séquencé au Génoscope à Evry. Avec l'aide de Pascal Hingamp du laboratoire de Génomique structurale (CNRS-Aix-Marseille-Université), nous avons utilisé les techniques de bioinformatique pour analyser et classer cette séquence d'ADN inconnue de 1010 pb.

TARA7S_3634040 ADN génomique (Tanger-Alger, station 7 [SRF])

AGTCACAACATCGCCACAACATCAGCGAGATTGTTTTCATGCAACATCGGGTGGTATGGGGCTTACTGGAATAATCATAAGTGCACACTATTTCAGCT
GCTTCCCATAAATCTTCGTTAAATTTCTCAGAAAACAATTAAAGGGGAGCTGTATAGAAGAGGATGTGAAGTCTTTGAGATTAATAGTGATGCTAC
TTACAGCGTTGCATGGATAGATTGCCTCAGCAAAGGAAAGGCTTGGGCGTAGTGTGATAATGCTTGGTGAGCATGCATCGCAAGGTGGATTA
GATATTGATACGACAGAAAGTCTCTGTTCCGTTTCTACTCCATCTGCGTTACTTAACAGTCTAACAGTGAAGGCCCTTAACGCCGCTTATTGGC
ACAAATCAAGGCAAAACCTAAGTCAGACAGTTGGGCTTATGCCTTACTTTATCCGTTAGATACTGTAGGTGAGTGGAAATAGGTTGTACGGTAAA
AAGGGATTTTGCATTTTCAGTGTGTCGTGCGTAAAGCTGAGTGGCGTGCAGAAATATGCGTAAGCTCTTAACCTGAAATCTCCACAGCGCGGAAGG
TTGCTTTCTAGCAGTACTTAAGCAGTTTGAATTTGCTAATGAGAACCCTCTCTCATTTCCAACAGAAGGTTACACATTTGGCACTAGACTTTAAGGCT
TCAGAAACAGCAATAAAGACGTTAAAGACTTGAAGATATGGTGGTAGGTATGGGTGGAAGGCTGTACCTAACGAAAGACGAGTAATGCAAG
AAAGCACCTTTAAAGCGACATATCCAACTGGGAAAAGTTTGAGGTAGTGGCGAGAACAATATGTTGCGGATTGGAAGATTAGTTCCACTCAGTCA
AAACGCTTAGGATTAGCATGAAGAGAATTTTAATTATTTGGGGCGAGCTCAGCTATTGCGGAGCACTGTGCGCGGATTGGGCTGCTAGAGGTGA
TGATTTGCACCTAGTGGCTAGAAATGAACGACAGCTTCAAAACATCG CAGCTGACTT

Transcription et traduction

Sequence Manipulation Site ORF Finder

Séquence de 292 acides aminés (chaque lettre correspond à un acide aminé)

VTTSPQHHADLFHATCGGMGLTGHISATIQLLPKSSLSISQKTIKADCEEACEVFEINS
DATYSVAWIDCLSKGKSLGRSVMILGEHASQGGLDIDIRQKVSVPFSTPSALLNSLTVK
AFNAAYWHKSRQNLSTQTVGLMPYFYPLDVTGGEWNRLYGKKGFLOFQCVPVKLDGV
ANMRKLLTEISNSGEGSFLAVLKQFGIANENLLSFPTEGYTLALDFKASETAIKTVKR
LEDMMVGMGGRLYLTKDAVMQESTFKATYPNWEKFEVVREQYGAIGKFSSTQSKRL
GLA*



Recherche de protéines homologues à la séquence trouvée à partir des banques de données internationales : BLAST = Basic Local Alignment Search Tool

Sequences producing significant alignments:

Select: All None Selected:0

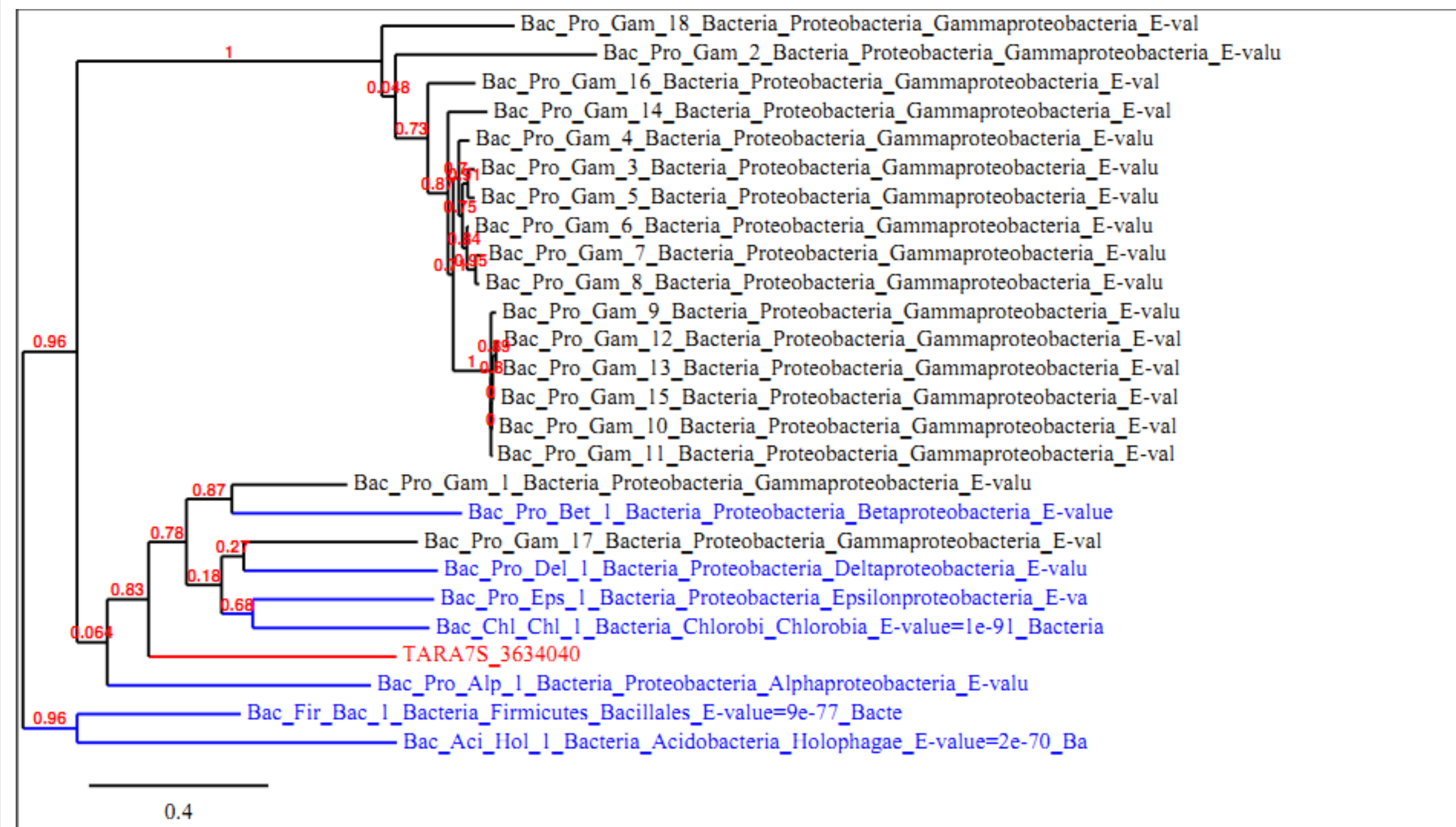
Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	E value	Max ident
FAD linked oxidase domain-containing protein [Pseudomonas aeruginosa VRFP01] >gb E12322.1 FAD linked oxidase domain-containing protein [Pseudomonas aeruginosa] >gb ABR82404.1 FAD linked oxidase domain protein [Pseudomonas aeruginosa]	3e-111	54%
FAD linked oxidase domain-containing protein [Pseudomonas aeruginosa PA7] >gb ABR82404.1 FAD linked oxidase domain protein [Pseudomonas aeruginosa]	4e-111	55%
hypothetical protein CT0170 [Chlorobium tepidum TLS] >gb AA71418.1 conserved hypothetical protein [Chlorobium tepidum TLS]	1e-106	55%
FAD linked oxidase domain protein [Vibrio parahaemolyticus AN-5034] >ref ZP_12107637.1 oxidoreductase, FAD-binding, putative [Vibrio parahaemolyticus 103]	2e-99	51%
oxidoreductase, FAD-binding [Methylococcus capsulatus str. Bath] >gb AAU91361.1 putative oxidoreductase, FAD-binding [Methylococcus capsulatus str. Bath]	3e-93	49%
FAD linked oxidase domain-containing protein [Hydrogenophaga sp. PBC1] >gb E1K92279.1 FAD linked oxidase domain-containing protein [Hydrogenophaga sp.]	2e-92	48%
FAD linked oxidase domain-containing protein [Chlorohelpton thalassium ATCC 35110] >gb ACF14054.1 FAD linked oxidase domain protein [Chlorohelpton thalassium]	1e-91	49%
FAD linked oxidase-like protein [Sulfurimonas denitrificans DSM 1251] >gb AB844496.1 FAD linked oxidase-like protein [Sulfurimonas denitrificans DSM 1251]	2e-89	48%
FAD linked oxidase domain-containing protein [Syntrophobacter fumaroxidans MPOB] >gb ABK15844.1 FAD linked oxidase domain protein [Syntrophobacter fumaroxidans]	4e-89	48%
oxidoreductase [Sulfovulum sp. NBC37-1] >gb BAF72471.1 oxidoreductase [Sulfovulum sp. NBC37-1]	3e-88	48%
FAD linked oxidase domain protein [Desulfurivibrio alkaliphilus AHT2] >gb ADH86452.1 FAD linked oxidase domain protein [Desulfurivibrio alkaliphilus AHT2]	6e-87	47%
hypothetical protein C900_04196 [Fulvivirga imtechensis AK7] >gb ELR73344.1 hypothetical protein C900_04196 [Fulvivirga imtechensis AK7]	4e-85	46%
oxidoreductase [Sulfovulum sp. NBC37-1] >gb BAF71292.1 oxidoreductase [Sulfovulum sp. NBC37-1]	1e-83	45%
FAD linked oxidase domain-containing protein [Sphingobium chlorophenolicum L-1] >gb AEG50395.1 FAD linked oxidase domain protein [Sphingobium chlorophenolicum]	2e-82	45%
putative oxidoreductase, FAD-binding [Algoriphagus sp. PR1] >gb EAZ82410.1 putative oxidoreductase, FAD-binding [Algoriphagus sp. PR1]	2e-82	47%
FAD linked oxidase domain protein [Burkholderia graminis C4D1M] >gb EDT11557.1 FAD linked oxidase domain protein [Burkholderia graminis C4D1M]	1e-80	47%

Un **BLAST** est une méthode de recherche informatique consistant à rechercher les séquences protéiques homologues au fragment TARA déterminé ci-dessus.

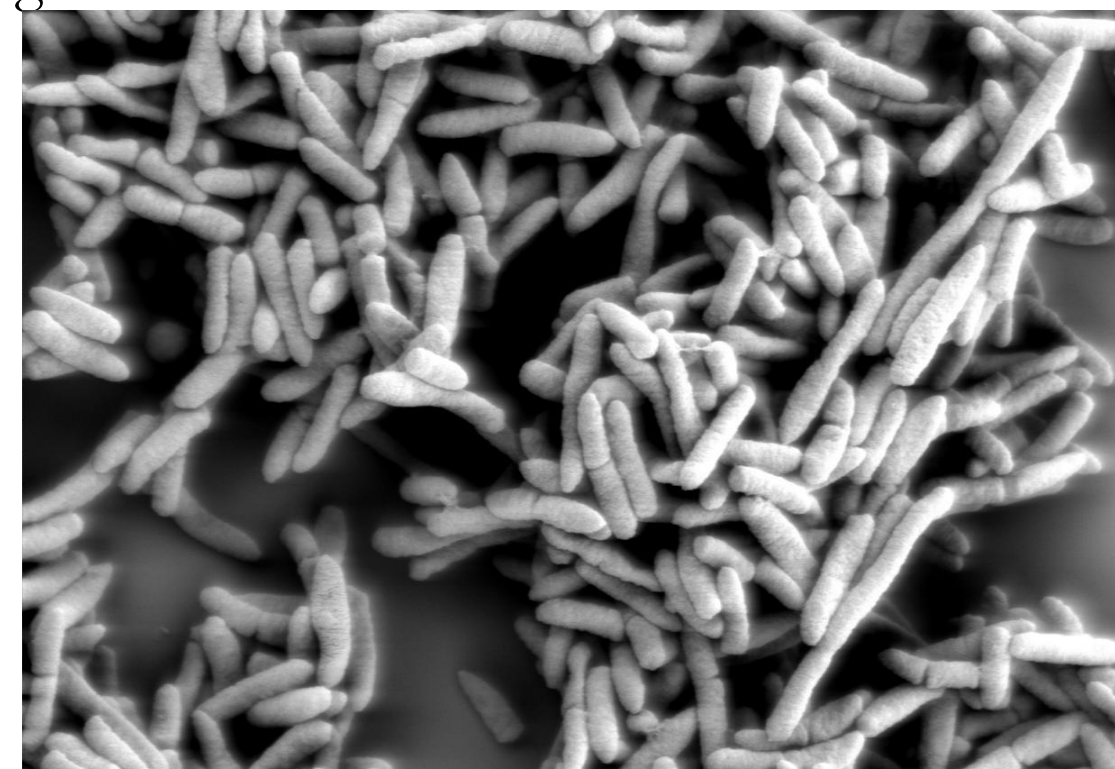
La séquence étudiée semble donc apparentée à celle d'un domaine protéique contenu dans une protéine enzymatique type « *FAD-linked oxidase* » présente chez une **gammaprotéobactérie**.

Établissement d'une phylogénie des Gammaprotéobactéries (www.phylogeny.fr). En rouge, apparaît la séquence TARA que nous avons étudiée et en bleu les groupes extérieurs.



La lecture de notre arbre phylogénétique ne semble pas confirmer qu'il s'agit d'une gammaprotéobactérie.

Par contre, nous pouvons raisonnablement affirmer que la séquence TARA étudiée doit appartenir à une **bactérie océanique type Protéobactérie**. Ce sont des bactéries à paroi Gram négatif.



Protéobactéries océaniques