

Supplementary Material:

MBSP1: a biosurfactant protein derived from a metagenomic library with activity in oil degradation

Sinara Carla da Silva Araújo¹; Rita C. B. Silva-Portela¹; Daniel Chaves de Lima¹; Marbella Maria Bernardes da Fonsêca¹; Wydemberg J. Araújo¹; Uaska Bezerra da Silva¹; Amanda P. Napp²; Evandro Pereira²; Marilene H. Vainstein²; Lucymara Fassarella Agnez-Lima^{*}.

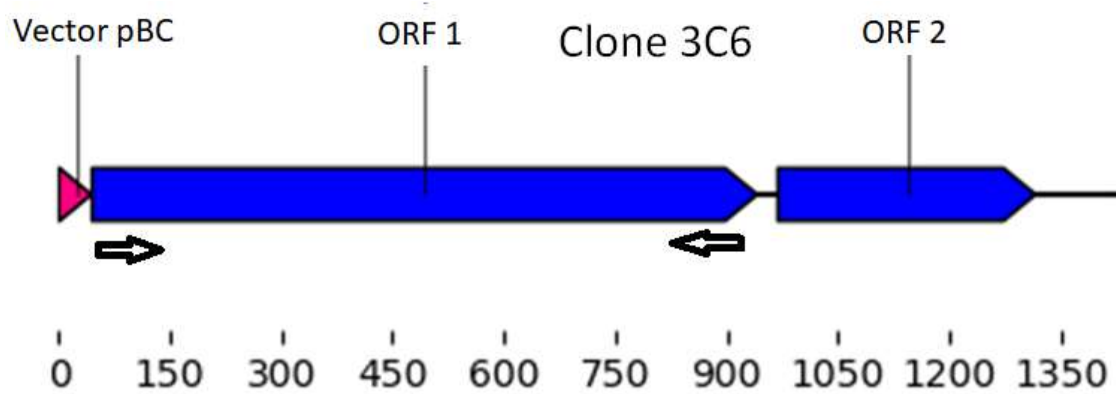
*Corresponding Author:

Dr. Lucymara Fassarella Agnez-Lima
Departamento de Biologia Celular e Genética, Centro de Biociências,
Universidade Federal do Rio Grande do Norte,
Campus Universitário, Lagoa Nova, Natal-RN,
59078-900, Brazil.
Tel. # 55.84.3211-9209, Fax # 55.84.3215-3346
E-mail: lfagnez@ufrnet.br

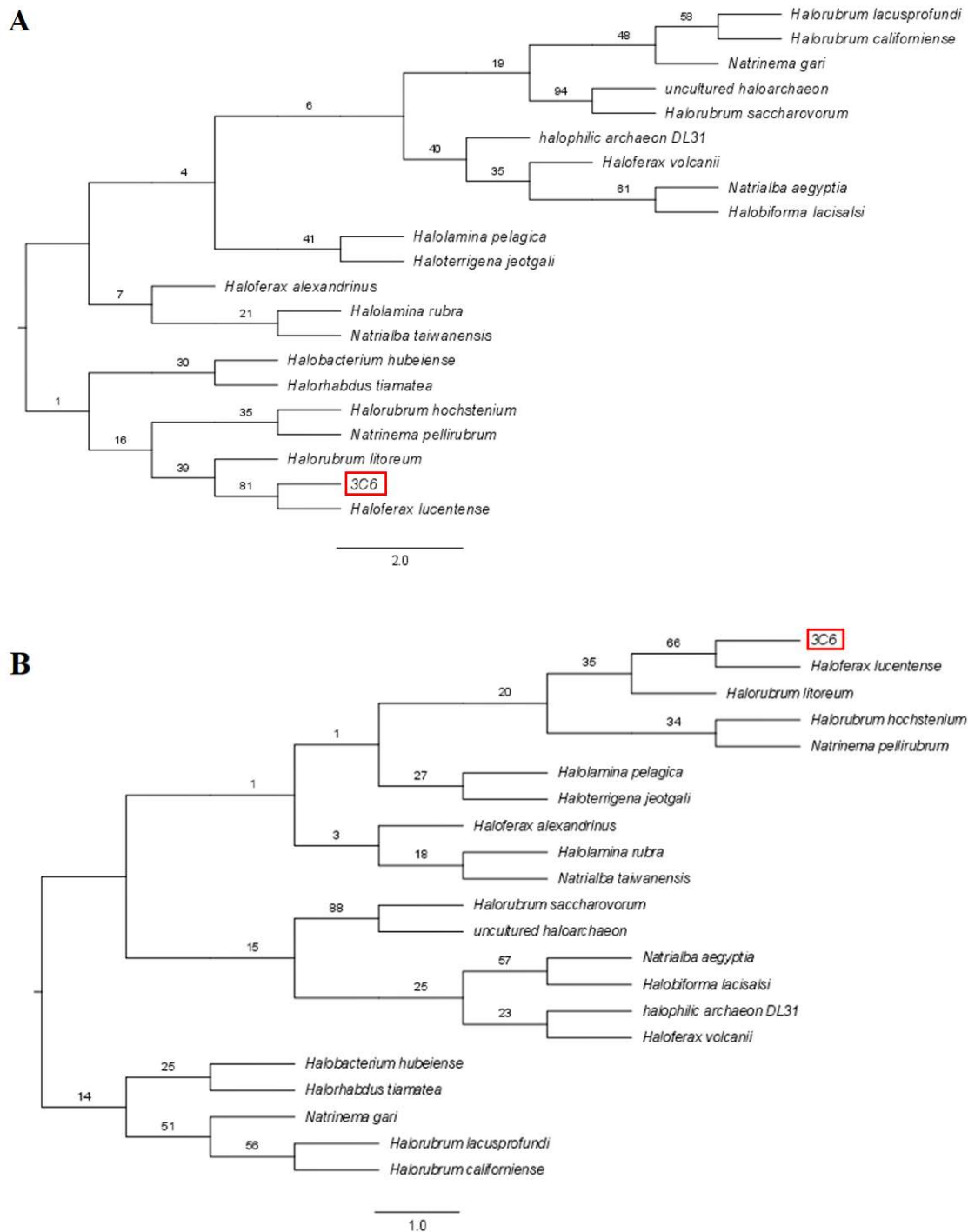
This file contains:

Supplementary Figures S1 to S6

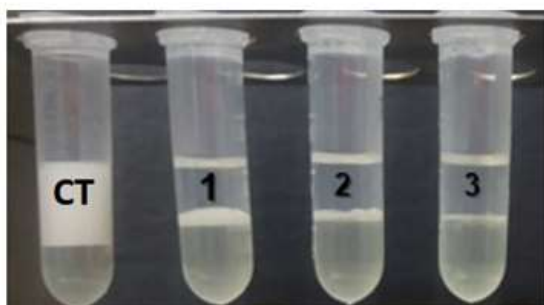
Supplementary Tables S1 to S2



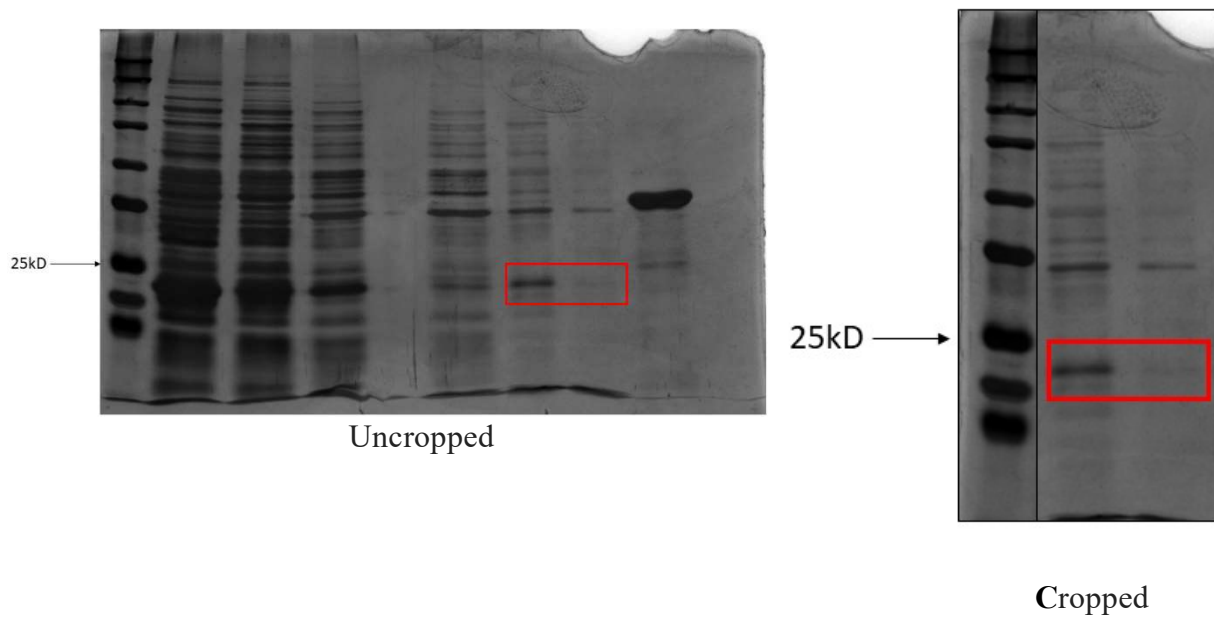
Supplementary Figure S1: ORFs identified in clone 3C6. The ORF1 contains 897 bp and was subcloned into a pHis-parallel1 expression vector in this work (↔primers used for sequence amplification). Both ORFs show high identity with conserved hypothetical ORFs from *Halobacteriaceae* family.



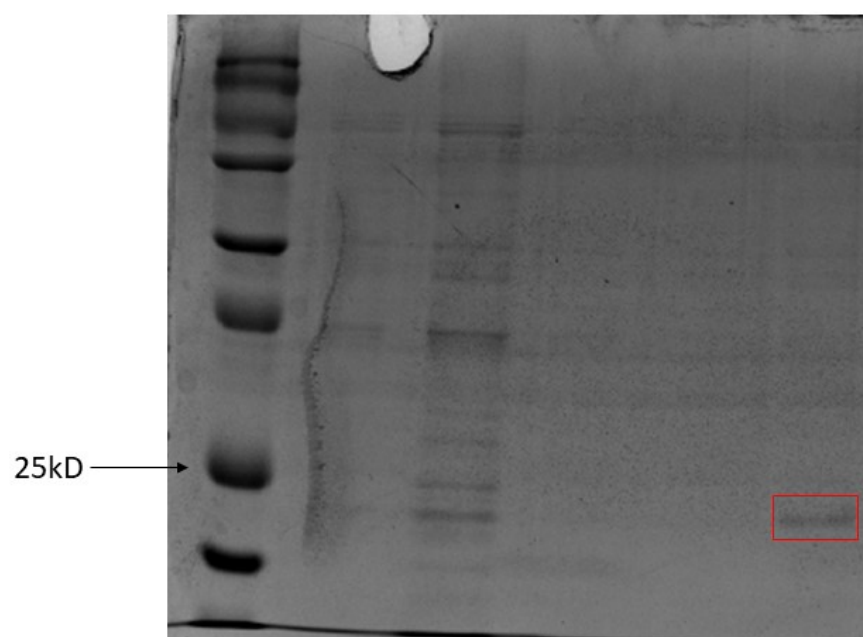
Supplementary Figure S2: Phylogenetic trees evidencing highest similarity with *Haloferax lucentense* proteins. A) Phylogenetic tree obtained from alignment by maximum parsimony; B) Phylogenetic tree obtained from alignment by maximum likelihood.

A**B**

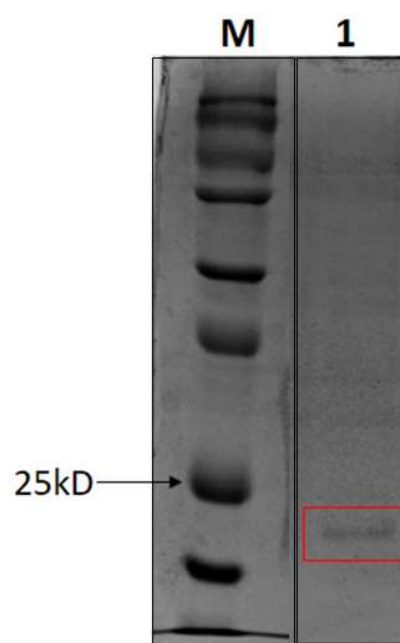
Supplementary Figure S3: Stability of MBSP1 obtained by acid precipitation measured by emulsification test using kerosene as a substrate A) Addition of proteinase-k: CT- control without addition of proteinase K; 1- 0.1 mg/ml of proteinase K; 2- 0.2 mg/ml of proteinase K; 3- 0.3 mg/ml of proteinase K. B) Emulsion with MBPS1 heated to 100 °C.



Supplementary Figure S4: Uncropped and cropped **figure 3C**. The highlighted samples are part of the study. Additional samples are not part of this study. The black line delineates where the gel was cut.

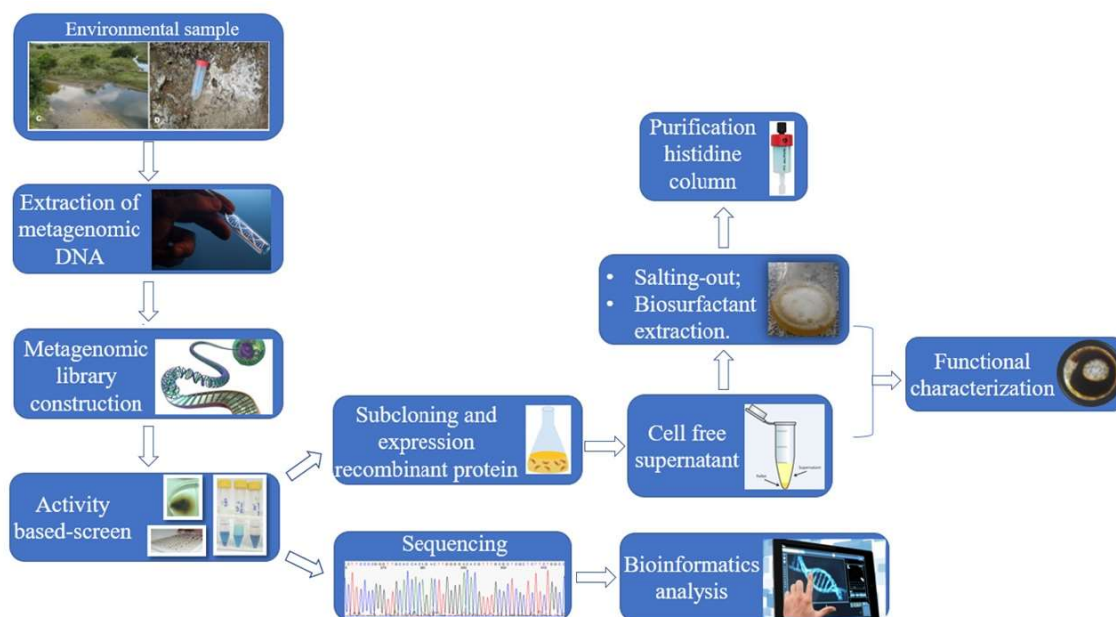


Uncropped



Cropped

Supplementary Figure S5: Uncropped and cropped **Figure 3D**. The highlighted sample is part of the study. Additional samples are not part of this study. The black line delineates where the gel was cut.



Supplementary Figure S6: Visualization of the work flow for prospecting and obtaining a clone with biosurfactant activity obtained from metagenomic library.

The construction of the metagenomic library, functional screening, expression of the recombinant protein and its characterization.

Supplementary Table S1: Soil characteristics and geographical references of the area of sample collection.

Geographic reference	Conductivity (dS.m⁻¹)	pH	Granulometry (g.Kg⁻¹)			Sodium (mg.dm⁻³)
S 05° 51’ 926’’	4,3	6,5	Sand	Clay	Silt	90
W 35° 21° 310’’			955	20	25	

Supplementary Table S2: BlastP results of the 3C6 ORF.

Description of BlastP hit (ORF3C6)	E-value	Identity	Accession
hypothetical protein [<i>Haloferax lucentense</i>]	0.0	93%	WP_004062196.1
hypothetical protein [<i>Halorubrum trapanicum</i>]	0.0	91%	WP_096396384.1
hypothetical protein [<i>Natrialba taiwanensis</i>]	0.0	90%	WP_083867309.1
hypothetical protein [<i>Haloterrigena jeotgali</i>]	0.0	92%	WP_084158367.1
hypothetical protein [<i>Halorubrum</i> sp. SD683]	0.0	94%	WP_086215034.1
hypothetical protein [<i>Halobacterium hubeiense</i>]	0.0	89%	WP_059058917.1
hypothetical protein [<i>Halobellus clavatus</i>]	0.0	88%	WP_089767349.1
hypothetical protein [<i>Haloarchaeobius iranensis</i>]	6e-180	88%	WP_089736699.1
hypothetical protein [<i>Halolamina pelagica</i>]	1e-179	88%	WP_074879642.1
Conserved hypothetical protein [<i>Halorhabdus tiamatea</i> SARL4B]	2e-179	91%	CCQ35110.1