

Numéros de CDPS	Informations disponibles dans les	
	Organisme	Locus du gène
1	<i>Kutzneria</i> sp. 744	KUTG_00871
2	<i>Streptomyces</i> sp. AA4	SSMG_06201
3	<i>Austwickia chelonae</i> NBRC 105200	AUCHE_08_05160
4	<i>Saccharothrix espanaensis</i> DSM 44229	BN6_23810
5	<i>Amycolatopsis azurea</i> DSM 43854	C791_7200
6	<i>Amycolatopsis decaplanina</i> DSM 44594	H074_03090
7	<i>Streptomyces somaliensis</i> DSM 40738	SsomD4_010100015357
8	<i>Candidatus Odysseella thessalonicensis</i> L13 HMO	COdytL_10466
9	LPE509 plasmid	LPE509_p00072
10	<i>Sphingobium japonicum</i> UT26S	SJA_C1-32620
11	<i>Kutzneria</i> sp. 744	KUTG_09682
12	<i>Saccharopolyspora spinosa</i> NRRL 18395	SspiN1_010100037249
13	<i>Marinomonas mediterranea</i> MMB-1	Marme_0276
14	<i>Streptomyces cattleya</i> NRRL 8057	SCAT_0901
15	<i>Streptomyces cattleya</i> NRRL 8057	SCAT_4268
16	<i>Streptomyces sviveus</i> ATCC 29083	SSEG_10527
17	<i>Parabacteroides</i> sp. 20_3	HMPREF9008_03529
18	<i>Moorea producens</i>	LYNGBM3L_69800
19	<i>Pseudomonas aeruginosa</i> PA21_ST175	H123_18841
20	<i>Diplorickettsia massiliensis</i> 20B	Dmas2_010100010809
21	<i>Diplorickettsia massiliensis</i> 20B	Dmas2_010100011124
22	<i>Diplorickettsia massiliensis</i> 20B	Dmas2_010100005805
23	WGS	CAGGBEG34_30028
24	<i>Rickettsiella grylli</i>	RICGR_1344
25	<i>Parachlamydia acanthamoebae</i> str. Hall's coccus	pah_c026o110
26	<i>Candidatus Protochlamydia amoebophila</i> UWE25	pc1814
27	<i>Micavibrio aeruginosavoru</i> s EPB	A11S_1546
28	<i>Legionella longbeachae</i> D-4968	LLB_2968
29	<i>Fluoribacter dumoffii</i> Tex-KL	FdumT_010100010802
30	<i>Rickettsiella grylli</i>	RICGR_0139
31	<i>Rickettsiella grylli</i>	RICGR_1349
32	Uncultured bacterium ACD_69C00020	ACD_69C00020G0002
33	<i>Fluoribacter dumoffii</i> Tex-KL	FdumT_010100011357
34	<i>Diplorickettsia massiliensis</i> 20B	Dmas2_010100005975
35	<i>Micavibrio aeruginosavorus</i> ARL-13	MICA_581
36	<i>Vibrio brasiliensis</i> LMG 20546	VIBR0546_16848
37	<i>Burkholderia lata</i>	Bcep18194_C7320
38	<i>Gordonia effusa</i> NBRC 100432	GOEFS_047_00070
39	<i>Nocardia brasiliensis</i> ATCC 700358	O3I_025450
40	<i>Streptomyces albulus</i> CCRC 11814	K530_11825
41	<i>Streptomyces albulu</i> s CCRC 11814	K530_11977
42	<i>Streptomyces albulus</i> CCRC 11814	K530_34281
43	<i>Pseudomonas</i> sp. M47T1	PMM47T1_20098
44	<i>Pseudomonas protegens</i> Pf-5	PFL_1389
45	<i>Gibberella zeae</i> PH-1	FG04615.1
46	<i>Desulfovibrio aespoensis</i> Aspo-2	Daes_1150
47	<i>Streptomyces</i> sp. AA4	SSMG_00951
48	<i>Streptomyces cattleya</i> NRRL 8057	SCAT_0198
49	<i>Streptomyces venezuelae</i> ATCC 10712	SVEN_1864

Annexe électronique 1 : Informations disponibles d

bases de données pour chaque CDPS potentielle	
Identifiant de la protéine	Commentaire
EWM10567	
WP_009082649	
WP_006503025	
WP_015099810	
WP_005165333	obablement partielle à cause d'une annotation de codon initiateur incorrecte
WP_007028561	
WP_010473184	obablement partielle à cause d'une annotation de codon initiateur incorrecte
WP_010302596	
WP_015445053	
WP_013041345	obablement partielle à cause d'une annotation de codon initiateur incorrecte
EWM19378	
WP_010314882	
WP_013659485	
WP_014141671	ablement trop long à cause d'une annotation de codon initiateur incorrecte
WP_014144981	
WP_007379823	
WP_005863245	
WP_008190827	
WP_003158562	
WP_010598945	obablement partielle à cause d'une annotation de codon initiateur incorrecte
WP_010599006	
WP_010598044	
WP_006683143	
WP_006034660	
WP_006341088	
WP_011176359	obablement partielle à cause d'une annotation de codon initiateur incorrecte
WP_015467885	
WP_003635969	
WP_010654300	
WP_006034819	
WP_006034794	
EKD46099	
WP_010654411	
WP_010598076	
WP_014102141	
WP_006880660	
WP_011350007	
WP_007317485	
WP_014985896	
WP_016572541	
WP_016572571	obablement partielle à cause d'une annotation de codon initiateur incorrecte
WP_016576960	
WP_008372311	
WP_011059731	
XP_384791	
WP_013514096	
WP_009073312	
WP_014140974	
WP_015033069	

ans les bases de données et séquences relatives à chaque CDPS p

Données de séque	
Séquence codante de référence	Gène synthétique ou amplicon PCR
KUTG_00871	CCATGGCCTTTACCGAAGAACCGCTGA
SSMG_06201	CCATGGCCTTTGAAGCAAATCCGCTGA
AUCHE_08_05160	CCATGGCAACCGTTACCGAAGTTCTGC
BN6_23810	CCATGGCACTGCTGCCGACACCGCTG
NZ_ANMG01000073	CCATGGCACTGCGTCCGCTGCCGCTG
H074_03090	CCATGGCACCGCGTCCGCTGCCGCTG
NZ_AJJM01000075	CCATGGCCTTTGAAATTGAACCGCTGA
COdytL_10466	CCATGGCAATCCGCAGCAGAAAATTG
LPE509_p00072	CCATGGATAACCGCAATGTTATTGTTA
complément des nucléotides 3246475-3247194 of NC_014006	CCATGGATGTTACCGTTCTGAGCCATC
KUTG_09682	CCATGGCAAGTCATGGTCTGGTCTGT
SspiN1_010100037249	CCATGGCAAATGCAGGTCGTGAGCCG
Marme_0276	CCATGGCACTGAGCATTGAAGCACAG
nucléotides 154-887 of SCAT_0901	CCATGGCACGTAGCGGTGATTTAACT
SCAT_4268	CCATGGCAGATGCAGCAGGTCCGGAT
SSEG_10527	CCATGGCACATGATTGGGCAATGCCT
HMPREF9008_03529	CCATGGCAAGCCTGCTGTATGAACGTC
LYNGBM3L_69800	CCATGGCAAGCAGCCATCGTGATGAA
H123_18841	CCATGGCAAATCTGATCGCAGAAGATT
complément des nucléotides 3383-4258 of NZ_AJGC01000004	CCATGGAAGTGCATGCCAAAATTAACG
Dmas2_010100011124	CCATGGCACCGGAAAACTATACCAATA
Dmas2_010100005805	CCATGGCATTCCGACCATTAACTGA
CAGGBEG34_30028	CCATGGCCAAAATCGCCTATTTCAAAG
RICGR_1344	CCATGGCACGCTACAAAAAATGAAAA
pah_c026o110	CCATGGCCAACAGCTATATTTGGCGTC
nucléotides 2168543-2169298 of NC_005861	CCATGGCCAACAAATATACCTGGCGCA
A11S_1546	CCATGGCAAAAGTTGCACCGAAAAGC
LLB_2968	CCATGGCACCGAAAAGTTCGTGGTCCG
FdumT_010100010802	CCATGGCACCGAAAAGTTCGTGGTAGC
RICGR_0139	CCATGGCCCATAGCATGAAACATCATA
RICGR_1349	CCATGGCCCTGAAAAATAACAATAATA
ACD_69C00020G0002	CCATGGCCCAGATTAACAAACCGCGT
FdumT_010100011357	CCATGGCCAGCATGATTGATAACCGTA
Dmas2_05975	CCATGGCACCGAAAACCAGCCGTCTG
MICA_581	CCATGGCCAACATAAAAGTGAAAATTC
VIBR0546_16848	CCATGGCAACCAAAGCCCGTTTTAAAC
Bcep18194_C7320	CCATGGCCTATAAAGTTAGCCTGCAGG
GOEFS_047_00070	CCATGGCAACCCTGACCCTGGCAGCA
O3I_025450	CCATGGCAACCACCACCACCCTGCTGA
K530_11825	CCATGGCCAATAGCTATCTGGTTGATG
nucléotides 81705-82424 of NZ_AROY02000068	CCATGGCACTGAGCGGTCTGGTTCCG
K530_34281	CCATGGCAGATACCAGCATGGGTGCC
PMM47T1_20098	CCATGGCAAGCGCAACCTATAAAGCA
PFL_1389	CCATGGCGAGTAAGGTACAGGAGATC
FG04615.1	CCATGGCGTATCATCTCGATCTGGAAC
Daes_1150	CCATGGCGATCATCAAAGCTACTACA
SSMG_00951	GTAGAACCCTGACCGCCGCTTGC
SCAT_0198	CCATGGCGCTGCATCGAACGTCCTTCT
SVEN_1864	CCTGCAGAGACGCTCCGACTCCAG

otentielle

ences relatives à chaque CDPS poten

Séquence en acides aminés

MAFTEEPLTSHCAGVRDSSAHACVGVSPF
MAFEANPLSSDCAALYPGAEHACVGVSLF
MATVTEVLLRDVVFRREPLTESCAQIVAEA
MALLPTLTEHCAAPLRSGRHVCGVSPF
MALRPLLTEHCADPLARAEHVCIGISPFN
MAPRPLLTEHCADPLALAEHVCIGVSPFN
MAFEIEPLTERCRGLLPEAVHVCIGVSPFN
MAIPQQKIDPQDFEYNHLVDSNEHALIGIS
MDNRNVIVTPHLITSLSATLYQQKEHVLLG
MDVTVLSHRALTVTSIRPGDRMNIGDYGH
MASHGRGRGSIVFTLDWLAETARHRDHV
MANAGRQPMRHEPATEESARIFQRRRH
MALSIEAQSGEAKSILKELKANTDYHAFSI
MARSGDFNCRPLTPNCARVVDAGEHAVV
MADAAGPDTAVGFGFFPVTHTCRLIWEEG
MAHDWAMPAAWAPSTEARSGDGFLATPF
MASLLYERRLESCYIERIYPYSESNAFLGV
MASSHRDEETIENFDWNQLSLEQLKLLQ
MANLIAEDFSTRSLLAGERYKAKIAFVSPH
MELHAKINGLEKHLKQHPTKRDFSIVNVIS
MAPENYTNKVKFHSKLEPGSILIYGVSKDN
MAFPTIKLKKHLHSADGKTASVLMIIISIDQ
MAKIAYFKGLEQFILENSLIKNKPIKIVIGIS
MARYKKMKKAKFQGNVDIIKNAQCLLMGI
MANSYIWRQNFTATFTDQEKTVRGKLAFK
MANKYTWRKNFTAILSNOKEIIKGKLAFKT
MAKVAPKSLSIWKEYKTACVGISMHSENH
MAPKVRGPRDGKLITASFDGKEKSLFPSL
MAPKVRGSKTGKYVTATFDSKDNSKSRFP
MAHSMKHHKISGLNTIKKDFKTPFVQEAS
MALKNNNNNSIKAPSYRIAHCGNFPNYLNLN
MAQINKPRAVFKCSEEFKLRFAHSDCLLTI
MASMIDNRIVKASFRENPEERKLFQSS
MAPKTSRLSYSGFYKSHNREKLFQTT
MANYKVKILGPNATHSHPVGRDLISVGQP
MATKARFKHDGEKQELVGKCCILAISVGQ
MAYKVSLQAQGFQAHEYVGKTALLLISV
MATLTLAAPVGSVSHKAAFDLKGQDYDIA
MATTTTTLLSASHKAAYDLRSDGITTDRS
MANSYLVLDGRTKANFGGAAKLLTAGTRA
MALSGLVPAHDHSMREEILGNRGRKIRQR
MADTSMGAGGSGATDFGADARAGTVAG
MASATYKAAFETDLSDYRGKHALLLISV
MASKVQEINLKAYFNKGEEHEFDGKRVV
MAYHLDLEPFAANGHCAVAICLGETPLDD
MAIISYYKTTQAQIEARRWNPYLGISINN
MAISNGFRVEPLTAACQEVWESRRHAVFG
MALHRTSFSVEPLTETCRVVVHNRDHVVL
AHVVIGVSPGNSYFTAQRLRELTRWGM

