

Table S2

Duf1628 (PilA) Analysis

Species	Accession Code	Gene	FlaFind 1.2 motif	Co-localized w/ putative pilins	TFP gene/additional notes
(Crenarchaeaota)					
<i>Thermoproteus neutrophilus</i> (<i>Pyrobaculum neutrophilum</i> V24Sta)	YP_001795233.1	Tneu_1872		Y	Tneu_1871 type II secretion system protein downstream
<i>Aeropyrum pernix</i> K1	NP_147667.2	APE_1013.1	RGIS	N	
(Euryarchaeaota)					
<i>Archaeoglobus fulgidus</i> DSM 4304	NP_069925.1	AF1096	KGVS	N	
<i>Ferroplasma placidus</i> DSM 10642	YP_003434907.1	Ferp_0458	RGIS	N	
	YP_003435552.1	Ferp_1118	KALT	N	
	YP_003435980.1	Ferp_1554	KAVS	N	
<i>Haloarcula marismortui</i> ATCC 43049	YP_134014.1	pNG2015		N	
	YP_134116.1	pNG4035		N	
	YP_135052.1	rmAC0290		N	
	YP_135368.1	rmAC0658	SAVS	N	
	YP_136350.1	rmAC1743		N	
	YP_136880.1	rmAC2355	RAVS	N	
	YP_137899.1	rmAC3526	DAVS	Y	rmAC3525 hypothetical protein (FlaFind1.2 positive) downstream same direction
	YP_137901.1	rmAC3528	DAVS	N	rmAC3526 hypothetical protein (FlaFind1.2 positive) upstream same direction, outside 2 gene limit

<i>Haloferax volcanii</i> DS2	YP_003534069.1	HVO_A0632	RAVS	Y	HVO_A0633 hypothetical protein (FlaFind1.2 positive) downstream same direction
	YP_003534070.1	HVO_A0633	RAVS	Y	HVO_A0632 hypothetical protein (FlaFind1.2 positive) upstream same direction
	YP_003534385.1	HVO_0308	RGSS	N	
	YP_003534935.1	HVO_0875		N	
	YP_003535029.1	HVO_0972	DAVS	N	
	YP_003536087.1	HVO_2062	SAVS	N	
	YP_003536471.1	HVO_2450	DAVS	Y	HVO_2451 hypothetical protein (FlaFind1.2 positive) upstream same direction
	YP_003536472.1	HVO_2451	RAVS	Y	HVO_2450 hypothetical protein (FlaFind1.2 positive) downstream same direction
	YP_003536720.1	HVO_2704	RGIS	N	
<i>Natronomonas pharaonis</i> DSM 2160	YP_325753.1	NP0180A		N	
	YP_326594.1	NP1876A	RGAV	Y	NP1880A hypothetical protein (FlaFind1.2 positive) downstream opposite direction
	YP_326632.1	NP1956A	RAVS	N	
	YP_326707.1	NP2110A	KALA	N	NP2108A flagella cluster protein upstream opposite direction
	YP_331148.1	NP4846A	RGQS	Y	NP4842A hypothetical protein (FlaFind1.2 positive) downstream same direction
	YP_331239.1	NP5032A	RAVA	Y	NP5030A hypothetical protein (FlaFind1.2 positive) upstream same direction
	YP_331401.1	NP5358A	RAVS	N	
<i>Methanosarcina mazei</i> Go1	NP_632501.1	MM_0477	KGLA	N	
	NP_633533.1	MM_1509		Y	MM_1510 hypothetical protein (FlaFind1.2 positive) upstream same direction
	NP_633534.1	MM_1510	KAVS	Y	MM_1511 hypothetical protein (FlaFind1.2 positive) upstream opposite direction
	NP_633535.1	MM_1511	KAVS	Y	MM_1510 hypothetical protein (FlaFind1.2 positive) upstream opposite direction
	NP_633537.1	MM_1513	SAVS	Y	MM_1514 hypothetical protein (FlaFind1.2 positive) upstream same direction
	NP_633538.1	MM_1514	EAAV	Y	MM_1516 Type II secretion system (FlaI) protein upstream same direction
	NP_633584.1	MM_1560	EGLS	N	

<i>Methanospirillum hungatei</i> JF-1	YP_501742.1	Mhun_0251		Y	Mhun_0252 hypothetical protein (non Fla-Find positive duf1628) downstream same direction
	YP_501743.1	Mhun_0252	SGVS	Y	Mhun_0251 hypothetical protein (non Fla-Find positive duf1628) upstream same direction
	YP_501757.1	Mhun_0267	EAVS	Y	Mhun_0268 hypothetical protein (non FlaFind positive duf1628) upstream same direction
	YP_501758.1	Mhun_0268		Y	Mhun_0267 hypothetical protein (FlaFind1.2 positive) downstream same direction
	YP_501786.1	Mhun_0296	SAVS	Y	Mhun_0297 hypothetical protein (FlaFind1.2 positive) upstream same direction
	YP_501787.1	Mhun_0297	DAVS	Y	direction; Mhun_0298 hypothetical protein (non FlaFind positive duf1628) upstream same direction
	YP_501788.1	Mhun_0298	SAVS	Y	Mhun_0297 hypothetical protein (FlaFind1.2 positive) downstream same direction
	YP_501789.1	Mhun_0299		Y	Mhun_0297 hypothetical protein (FlaFind1.2 positive) downstream same direction
	YP_501799.1	Mhun_0310		Y	Mhun_0311 hypothetical protein (FlaFind1.2 positive) downstream same direction
	YP_501800.1	Mhun_0311	EAVS	Y	Mhun_0310 hypothetical protein (non FlaFind positive duf1628) upstream same direction
	YP_502683.1	Mhun_1218	DAVS	Y	Mhun_1219 hypothetical protein (FlaFind1.2 positive) downstream same direction
	YP_502684.1	Mhun_1219	DAVS	Y	Mhun_1218 hypothetical protein (FlaFind1.2 positive) upstream same direction
	YP_503837.1	Mhun_2416	KAVS	N	
	YP_503846.1	Mhun_2426	DALS	Y	Mhun_2427 hypothetical protein (FlaFind1.2 positive), Mhun_2428 hypothetical protein (FlaFind1.2 positive) upstream same direction
	YP_503847.1	Mhun_2427	RAVS	Y	Mhun_2428 hypothetical protein (FlaFind1.2 positive) upstream same direction; Mhun_2426 hypothetical protein (FlaFind1.2 positive) downstream same direction
	YP_504089.1	Mhun_2674	SAVS	N	
	YP_504094.1	Mhun_2680		N	
	YP_504114.1	Mhun_2700		Y	Mhun_2701 hypothetical protein (non FlaFind1.2 positive duf1628) downstream same direction
YP_504115.1	Mhun_2701		Y	Mhun_2700 hypothetical protein (non FlaFind1.2 positive duf1628) upstream same direction	
<i>Thermoplasma acidophilum</i> DSM 1728	NP_393973.1	Ta0496	KAVS	N	
	NP_394310.1	Ta0851	KAVS	N	
(Thaumarchaeota)					

<i>Caldiarchaeum subterraneum</i>	BAJ51169	CSUB_C1318	RGLS	Y	CSUB_C1317 hypothetical protein (FlaFind1.2 positive) upstream same direction
-----------------------------------	-----------------	------------	------	---	---

FlaFind 1.2 positive [D/E] at -2

Not predicted by FlaFind 1.2

FlaFind positive [S] allowed at -2

Co-localized with putative pilin/pilus-biosynthesis gene

Co-localized with non-FF positive Duf1628