

S3 Text: Clear candidates whose mutants were not consistently auxotrophic

Most of the clear candidates (that is, genes that were identified by at least two out of three annotation resources as filling gaps in the IMG predictions) were either important for fitness in minimal media, or were likely essential in rich media, or the lack of an auxotrophic phenotype could be explained by genetic redundancy. Here we describe the two exceptions. First, Psest_1986 from *P. stutzeri* contains a prephenate dehydrogenase domain (PF02153) fused to a 3-phosphoshikimate 1-carboxyvinyltransferase (TIGR01356). Although this gene is expected to be required for tyrosine biosynthesis, it was not important for fitness in most defined media conditions. The gene's mutant fitness was estimated from just two mutant strains, both of which lie in the linker region between the two domains. The location of these strains is striking because the linker region contains just 40 of the 747 amino acids of Psest_1986. Our interpretation is that both enzymatic activities are essential, and that these insertions were recovered because they allowed both domains to be expressed as separate proteins. Second, DvMF_1902 from *D. vulgaris* contains a D-isomer specific 2-hydroxyacid dehydrogenase catalytic domain (PF00389) and is annotated as D-3-phosphoglycerate dehydrogenase by both KEGG and SEED. It is 39% identical to a characterized D-3-phosphoglycerate dehydrogenase (MMP1588; [59]). This activity is expected to be required for serine biosynthesis, as a related strain of *Desulfovibrio* appears to use the standard pathway of serine synthesis via phosphorylated intermediates [60]. Also, the isotope labeling of serine in another strain of *D. vulgaris* is consistent with the standard pathway [61]. Nevertheless, DvMF_1902 was not important for fitness in defined media (all fitness values were -0.5 or higher). Although we did not identify any other strong candidate genes for this activity, DvMF_1902 could be redundant with another dehydrogenase.

Additional References

53. Dos Santos PC, Fang Z, Mason SW, Setubal JC, Dixon R. Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics*. 2012 May 3;13:162.
54. Hagino K, Onuma R, Kawachi M, Horiguchi T. Discovery of an endosymbiotic nitrogen-fixing cyanobacterium UCYN-A in *Braarudosphaera bigelowii* (Prymnesiophyceae). *PLoS ONE*. 2013 Dec 4;8(12):e81749.
55. Chen CK, Blaschek HP. Effect of acetate on molecular and physiological aspects of *Clostridium beijerinckii* NCIMB 8052 solvent production and strain degeneration. *Appl Environ Microbiol*. 1999 Feb;65(2):499–505.
56. Kanamori K, Weiss RL, Roberts JD. Ammonia assimilation pathways in nitrogen-fixing *Clostridium kluyverii* and *Clostridium butyricum*. *J Bacteriol*. 1989 Apr;171(4):2148–54.
57. Theisen AR, Ali MH, Radajewski S, Dumont MG, Dunfield PF, McDonald IR, et al. Regulation of methane oxidation in the facultative methanotroph *Methylocella silvestris* BL2. *Mol Microbiol*. 2005 Nov;58(3):682–92.
58. Shi T, Sun Y, Falkowski PG. Effects of iron limitation on the expression of metabolic genes in the marine cyanobacterium *Trichodesmium erythraeum* IMS101. *Environ Microbiol*. 2007 Dec;9(12):2945–56.

59. Helgadóttir S, Rosas-Sandoval G, Söll D, Graham DE. Biosynthesis of phosphoserine in the Methanococcales. *J Bacteriol.* 2007 Jan;189(2):575–82.
60. Germano GJ, Anderson KE. Serine biosynthesis in *Desulfovibrio desulfuricans*. *J Bacteriol.* 1969 Sep;99(3):893–4.
61. Tang Y, Pingitore F, Mukhopadhyay A, Phan R, Hazen TC, Keasling JD. Pathway confirmation and flux analysis of central metabolic pathways in *Desulfovibrio vulgaris hildenborough* using gas chromatography-mass spectrometry and Fourier transform-ion cyclotron resonance mass spectrometry. *J Bacteriol.* 2007 Feb;189(3):940–9.