

E. coli b4226 (*ppa*) and
Mrub_0258 are orthologs; *E.*
coli b2501 (*ppk*) and
Mrub_1198 are orthologs

Brandon Wills

ppa gene – inorganic pyrophosphatase

Structure/Function:

- 175 amino acids
- 1 single domain
- Cytoplasm

Cellular Roles:

- Drive non-spontaneous reactions
- Macromolecule synthesis
- Glycolysis
- **Energy Metabolism**

Figure 1. Crystal structure of inorganic pyrophosphatase showing the active site and residues that surround it. Image taken from Yang et al., 2009

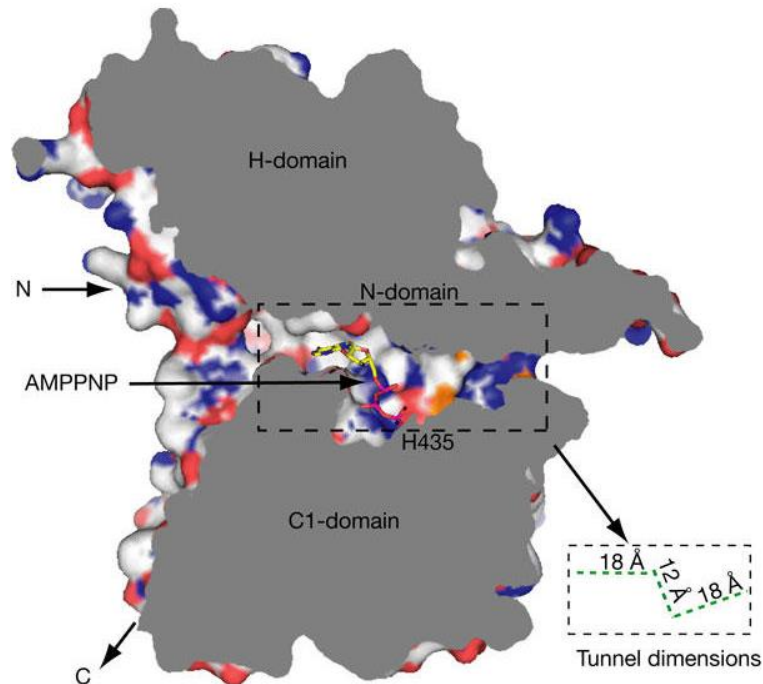
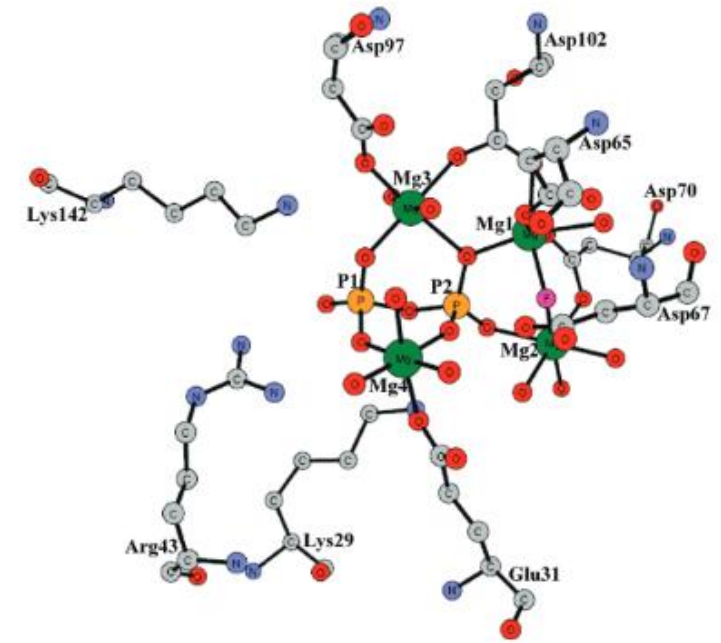


Figure 2. Side view of polyphosphate kinase showing the domains and the active site tunnel. Image taken from Zhu et al., 2005

ppk gene – polyphosphate kinase

Structure/Function:

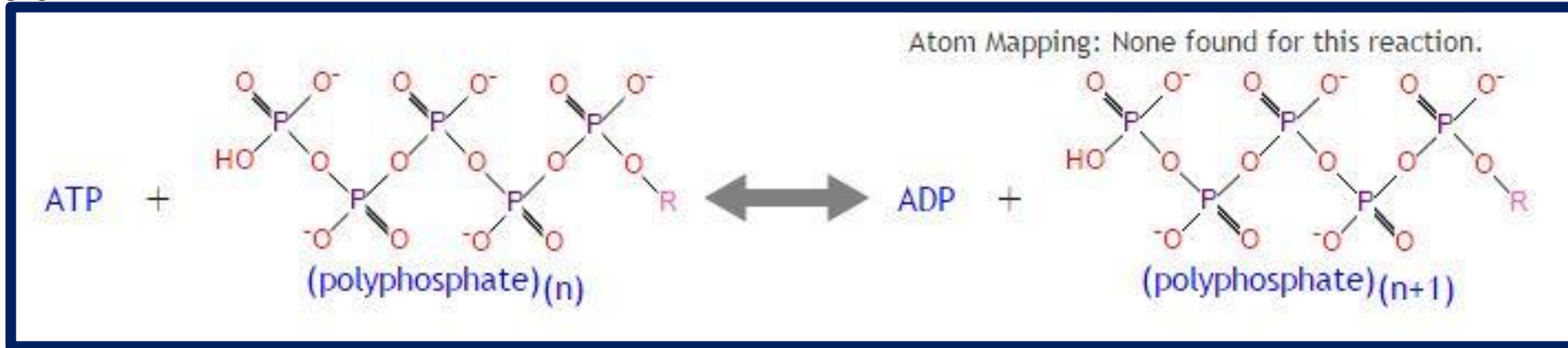
- 678 amino acids
- 4 separate domains
- Active site tunnel
- Outer Membrane

Cellular Roles:

- RNA Degradation
- Biofilm production
- Cell motility
- **Energy Metabolism**

Reactions Catalyzed

ppk



ppa

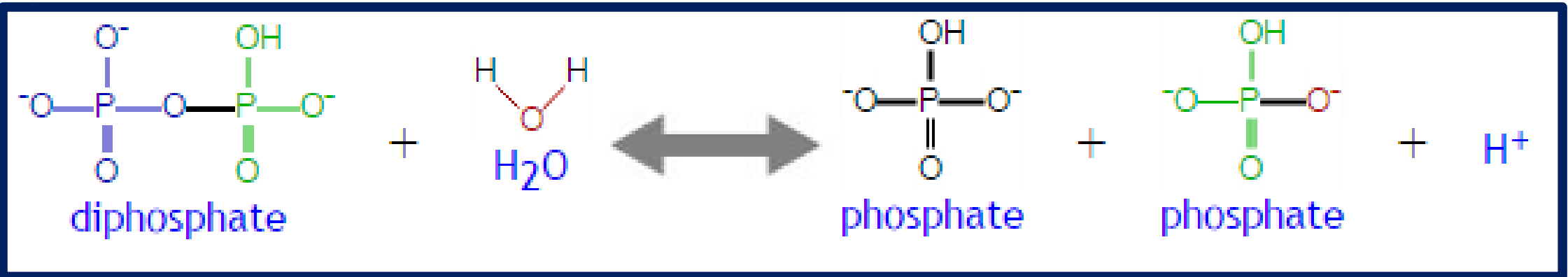


Figure 3. Reactions catalyzed by the enzymes encoded by *ppk* (top) and *ppa* (bottom). Images were taken from the MetaCyc website at <http://metacyc.org/>.

KEGG Pathway Map

Oxidative Phosphorylation

- Not directly involved
- Provide phosphates for ATP production
- Also provides phosphate for other cellular processes

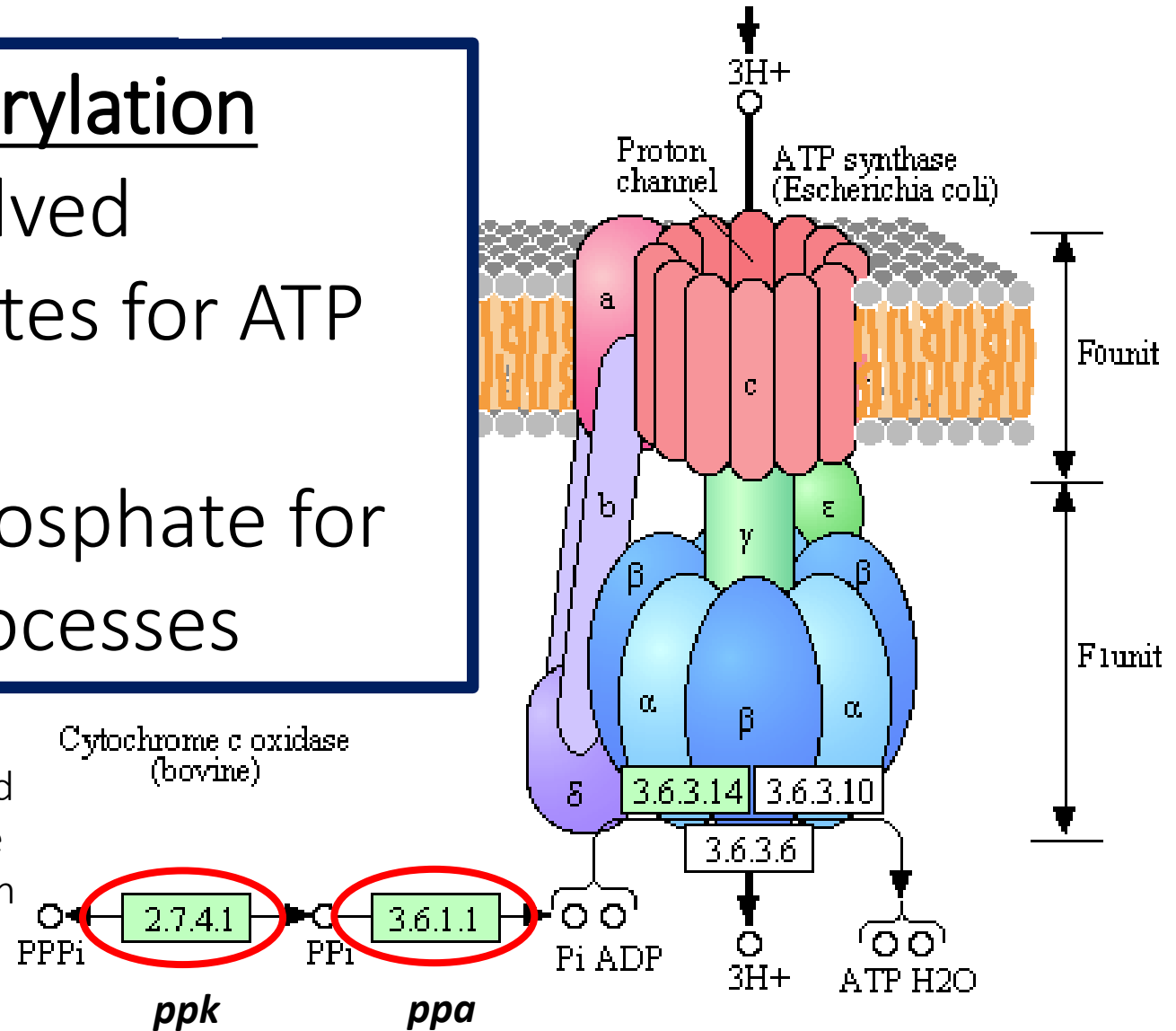


Figure 4. The enzymes encoded by *ppk* and *ppa* are indirectly linked with the oxidative phosphorylation pathway. Image was taken from the KEGG database at <http://www.genome.jp/kegg/>.

Functional Evidence

ppa

- Lahti R, Pitkaranta T, Valve E, Ilta I, Kukko-Kalske E, Heinonen J. 1988. Cloning and characterization of the gene encoding inorganic pyrophosphatase of *Escherichia coli* K-12. J Bacteriol 170(12):5901-5907.

ppk

- Zhu Y, Huang W, Lee S, Xu W. 2005. Crystal structure of a polyphosphate kinase and its implications for polyphosphate synthesis. EMBO reports 6(7):681-687.

Almost all the bioinformatics programs were consistent with each other for both the *ppa* gene and the Mrub_0258 gene

Crystal structures from different bacteria – still consistent enzyme name

These data suggest that these two genes might be orthologous to one another

Bioinformatics tool used	<i>E. coli ppa</i> gene	Mrub_0258 gene
BLAST <i>E. coli</i> against <i>M. ruber</i>	Score: 148 bits	
	E-value: 2e-49	
CDD Data (COG category)	COG Number: COG0221	
	Inorganic pyrophosphatase	
	E-value: 3.64e-85	E-value: 1.27e-68
Cellular Localization	Cytoplasm of the cell	
TIGRfam – protein family	N/A	
Pfam – protein family	Pf00719 (inorganic pyrophosphatase)	
	E-value: 1.1e-55	E-value: 4.6e-52
Protein Database	4UM4 and 2PRD	
	E-value: 2.27548e-100	E-value: 3.01152e-71
Enzyme commission number	E.C. 3.6.1.1 – Inorganic pyrophosphatase	
KEGG pathway map	Oxidative phosphorylation	
	KEGG number: 00190	

Table 1. *E. coli ppa* and Mrub_0258 are orthologs

Almost all the bioinformatics programs were consistent with each other for both the *ppk* gene and the Mrub_1198 gene

Cellular location modules showed some deviations (Not shown)

Deviations in results were assessed using primary literature published on the cellular location of this protein

Bioinformatics tool used	<i>E. coli ppk</i> gene	Mrub_1198 gene	
BLAST <i>E. coli</i> against <i>M. ruber</i>	Score: 262 bits E-value: 2e-81		
CDD Data (COG category)	COG Number: COG0855 Polyphosphate kinase E-value: 0.00 E-value: 3.97e-152		
Cellular Localization	Cytoplasmic membrane		
TIGRfam – protein family	TIGR03705 (polyphosphate kinase) E-value: 0.00 E-value: 4.20e-160		
Pfam – protein family	PF13089 (N-terminus Domain), PF02503 (Middle Domain), PF13090 (C-terminus Domain)		
	E-value	E value	E value
	5.8e-143 4.3e-88	5.8e-41 1.5e-11	3.1e-31 1.6e-19
Protein Database	1XDO E-value: 0.00 E-value: 8.28761e-70		
Enzyme commission number	E.C 2.7.4.1 – polyphosphate kinase		
KEGG pathway map	Oxidative phosphorylation KEGG number: 00190		

Table 2. *E. coli ppk* and Mrub_1198 are orthologs

Other Consistencies


- THMHH consistent for both gene pairs
- Same highly conserved amino acids
- None of the genes in operons

Cellular Location Deviations


Panel A

PSORT-B

go to PSORT-B at <http://www.psort.org/psortb>

Cytoplasmic score 

0.00


CytoplasmicMembrane score 

10.0


Panel B

PSORT-B

go to PSORT-B at <http://www.psort.org/psortb>

Cytoplasmic score 

2.11

CytoplasmicMembrane score 

7.88

Figure 5. PSORT-B results are inconsistent between the *ppk* gene and the Mrub_1198 gene. Panel A shows the PSORT-B results for *ppk*; Panel B shows the PSORT-B results for Mrub_1198. PSORT-B for *ppk* predicts that the protein is located in the cytoplasmic membrane. On the other hand, the PSORT-B results for Mrub_1198 are split between a cytoplasmic or cytoplasmic membrane location.

Is the protein found in the cytosol or the cytoplasmic membrane?

Cellular Location Deviations

NEITHER!

- A study by Akiyama et al., 1992 shows that the protein is actually found on the outer membrane
- It is a specific type of protein called a peripheral membrane protein
- Therefore, PSORT-B predicted the wrong cellular location

Horizontal Gene Transfer in Mrub_0258

Tree Rendering results

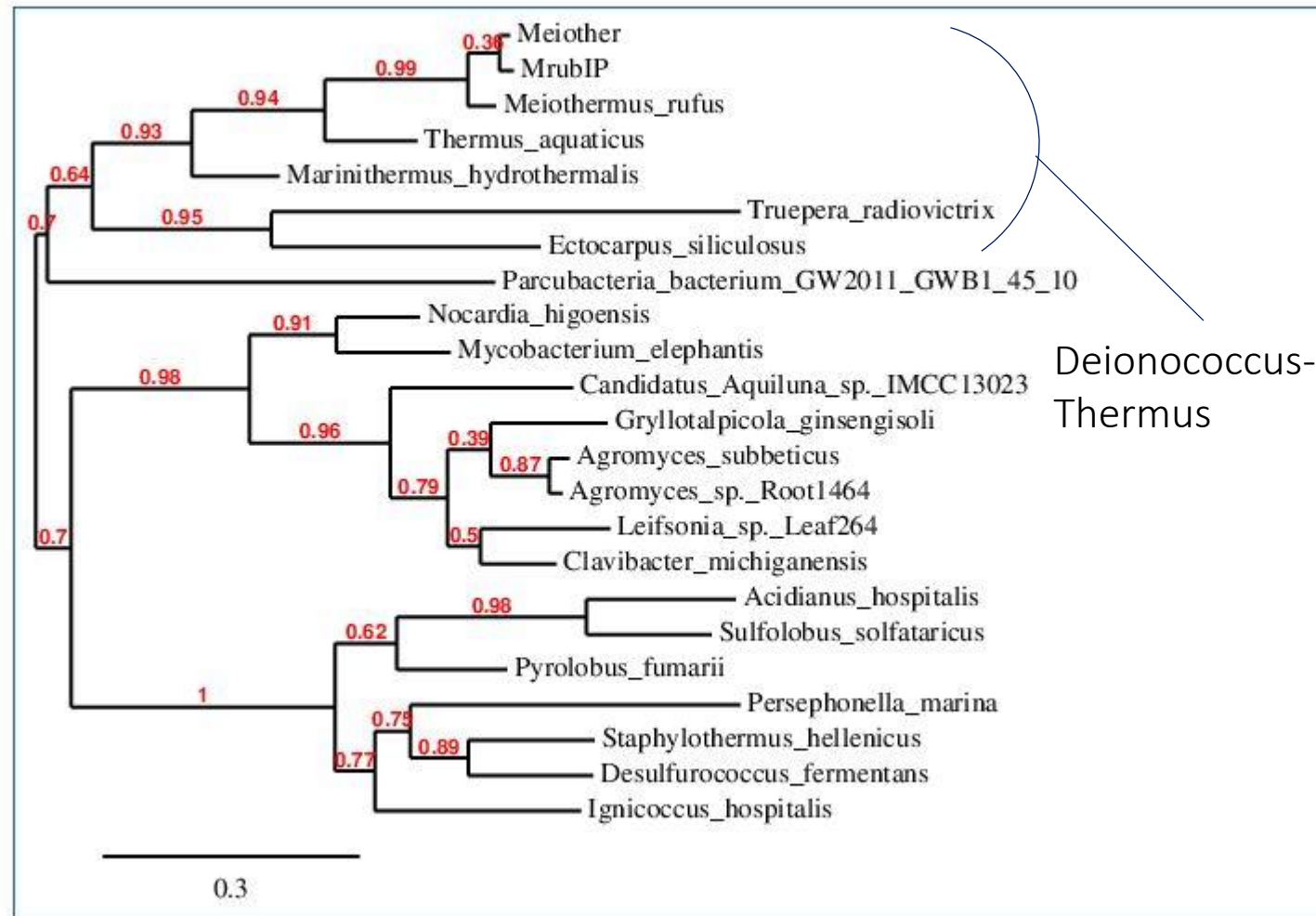


Figure 6. Mrub_0258 does not show evidence of recent HGT. The organisms near it on the phylogenetic tree are closely related. The BLAST program (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used to obtain sequences for the phylogenetic tree, which was created using phylogeny.fr.

Horizontal Gene Transfer in Mrub_1198

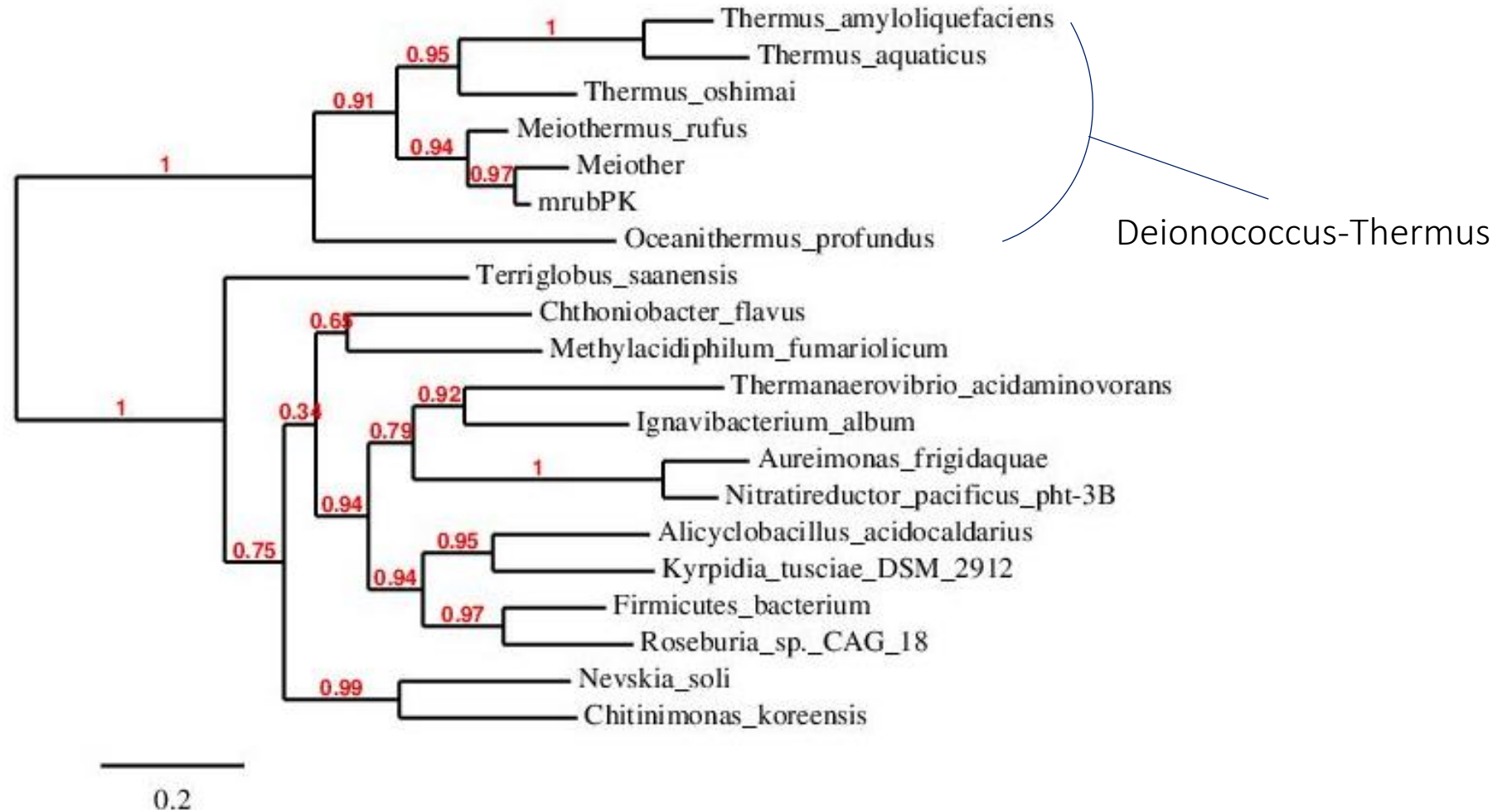


Figure 7. Mrub_1198 does not show evidence of recent HGT. The organisms near it on the phylogenetic tree are closely related. The BLAST program (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used to obtain sequences for the phylogenetic tree, which was created using phylogeny.fr.

Conclusion

Based on the frequency with which the bioinformatics matched between the gene pairs, the data suggest that:

- *ppa* and Mrub_0258 are orthologs
- *ppk* and Mrub_1198 are orthologs

References

- Akiyama M, Crooke E, Kornberg A. 1992. The polyphosphate kinase gene of *Escherichia coli*. Isolation and sequence of the ppk gene and membrane location of the protein. *J Biol Chem* 267(31):22556-22561.
- Blum E, Py B, Carpousis A, Higgins CF. 1997. Polyphosphate kinase is a component of the *Escherichia coli* RNA degradosome. *Molec Biol* 26(2):387-398.
- Chen W, Palmer RJ, Kuramitsu HK. 2002. Role of polyphosphate kinase in biofilm formation by *Porphyromonas gingivalis*. *Infection and Immunity* 70(8):4708-4715.
- Chen J, Brevet A, Fromant M, Leveque F, Schmitter J, Blanquet S, Plateau P. 1990. Pyrophosphatase is essential for growth in *Escherichia coli*. *Bacteriol* 172(10):5686-5689.
- Lahti R, Pitkaranta T, Valve E, Ilta I, Kukko-Kalske E, Heinonen J. 1988. Cloning and characterization of the gene encoding inorganic pyrophosphatase of *Escherichia coli* K-12. *J Bacteriol* 170(12):5901-5907.
- Sun J, Yang F, Zhao D, Yang X. 2014. Highly sensitive real-time assay of inorganic pyrophosphatase activity based on the fluorescent gold nanoclusters. *Anal Chem* 86(15):7883- 7889.
- Teplyakov A, Obmolova G, Wilson KS, et al. 1994. Crystal structure of inorganic pyrophosphatase from *Thermus thermophilus*. *Protein Sci* 3(7) 1098-1107. Terkeltaub RA. 2001. Inorganic pyrophosphate generate and disposition in pathophysiology. *Am J Physiol, Cell Physiol* 281(1) C1-C11.
- Yang L, Lioa R, Yu J, Liu, R. 2009. DFT study on the mechanism of *Escherichia coli* inorganic pyrophosphatase. *J Phys Chem B* 113(18):6505-6510.
- Zhu Y, Huang W, Lee S, Xu W. 2005. Crystal structure of a polyphosphate kinase and its implications for polyphosphate synthesis. *EMBO reports* 6(7):681-687.