Annotation of the pathway for the biosynthesis of chorismate from erythrose-4-phosphate in *Kytococcus Sedentarius* using the IMG-ACT database

Ashlesha P. Odak, Patricia Masso-Welch, Rama Dey-Rao, Stephen Koury

Department of Biotechnical and Clinical Laboratory Sciences, School of Medicine and Biomedical Sciences, State University of New York at Buffalo, 26 Cary Hall, 3435 Main Street, Buffalo, NY 14214

### Abstract

*Kytococcus sedentarius* is an aerobic, gram positive, non-motile, non-photosynthetic bacterium isolated from a marine environment. The genome of *Kytococcus sedentarius* is of interest to ascertaining its metabolic potential in the biosynthesis of chorismate, due to the importance of this pathway for many aromatic compounds. This study utilized omics databases, such as KEGG, COG, TIGRFam, PFam, Phylogeny, ExPasy and MetaCyc, to elucidate the pathways for the biosynthesis of chorismate from erythrose-4-phosphate in *Kytococcus sedentarius*. The metabolic pathway was annotated using Swiss-Prot and non-redundant databases, T-COFEE, WebLogo, TMHMM, PSORT-B, and genome-wide tools. The results from this study were then used to generate a multiple sequence alignment using the Tree-Based Consonancy Objective Function For Alignment Evaluation (TCoFE) software (http://www.expasy.ch/enzyme/enzyme-search-ec.html). This software program was used to confirm the data obtained from the other tools. The genome of *Kytococcus sedentarius* was found to be part of a marine environment, and the pathway for the biosynthesis of chorismate was confirmed to be present in *Kytococcus sedentarius*.

### Materials and Methods

The sequences of the top 10 orthologs were obtained through the IMG database. Immunological analyses were performed to generate a multiple sequence alignment using the Tree-Based Consonancy Objective Function For Alignment Evaluation (TCoFE) software (http://www.expasy.ch/enzyme/enzyme-search-ec.html). This software program was used to confirm the data obtained from the other tools. The genome of *Kytococcus sedentarius* was found to be part of a marine environment, and the pathway for the biosynthesis of chorismate was confirmed to be present in *Kytococcus sedentarius*.

### Results

1. **Figure 1**: KEGG Pathway map for biosynthesis of chorismate from erythrose-4-phosphate. Genes highlighted in green are the genes used by *K. sedentarius*.

2. **Figure 2**: Genes encoding the enzymes dehydroshikimate kinase, 3-dehydroquinate synthase and 3-dehydroshikimate 5-phosphate synthesis for the biosynthesis of chorismate are illustrated in green. The red highlighted area indicates the orthologs for the gene. These were used to generate a multiple sequence alignment using the Tree-Based Consonancy Objective Function For Alignment Evaluation (TCoFE) software (http://www.expasy.ch/enzyme/enzyme-search-ec.html). This software program was used to confirm the data obtained from the other tools.

3. **Figure 3**: Gene neighborhood map for gene OID 644991533 illustrates 1) Gene encoding 3-dehydroquinate synthase and 3-dehydroquinate synthase showing absence of a peak indicating that the protein lacks a signal peptide. Cut-off value of probability for the alignment was determined to be 0.5.

4. **Figure 4**: Gene neighborhood map for gene OID 644991533 illustrates 1) Gene encoding dehydroquinate synthase with the red highlighted area showing the statistics for the alignment.

5. **Figure 5**: Gene neighborhood map for gene OID 644991533 illustrates 1) Gene encoding dehydroquinate synthase with the red highlighted area showing the statistics for the alignment.

6. **Figure 6**: Gene neighborhood map for gene OID 644991533 illustrates 1) Gene encoding dehydroquinate synthase with the red highlighted area showing the statistics for the alignment.

7. **Figure 7**: Sequence viewer for ORF search. The arrow points to the first nucleotide of the ORF. No red highlighted nucleotides indicating no start codon has been called by the gene caller.

### Conclusions

The results from this study were used to generate a multiple sequence alignment using the Tree-Based Consonancy Objective Function For Alignment Evaluation (TCoFE) software (http://www.expasy.ch/enzyme/enzyme-search-ec.html). This software program was used to confirm the data obtained from the other tools. The genome of *Kytococcus sedentarius* was found to be part of a marine environment, and the pathway for the biosynthesis of chorismate was confirmed to be present in *Kytococcus sedentarius*. The results from this study were then used to generate a multiple sequence alignment using the Tree-Based Consonancy Objective Function For Alignment Evaluation (TCoFE) software (http://www.expasy.ch/enzyme/enzyme-search-ec.html). This software program was used to confirm the data obtained from the other tools. The genome of *Kytococcus sedentarius* was found to be part of a marine environment, and the pathway for the biosynthesis of chorismate was confirmed to be present in *Kytococcus sedentarius*.