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

Ashleasha 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

The biosynthesis of chorismate occurs via the shikimate pathway in seven enzyme-mediated steps. This pathway is a central pathway for the synthesis of various aromatic compounds in bacteria, fungi, yeasts, algae, plants and certain animals. The genome of *Kytococcus sedentarius* was isolated from a marine environment. The genome of this organism was analyzed and it was found that the gene encoding 3-dehydroquinate dehydratase responsible for conversion of dehydroquinate to 3-dehydroshikimate may have undergone a horizontal gene transfer from other organisms. The possible occurrence of horizontal gene transfer was confirmed by the IMG-ACT tools. The functions of all the genes in the pathway were confirmed by the KEGG and MetaCyc results.

**Materials and Methods**

The following modules were looked at to annotate the genes using the IMG-ACT website:

1. Basic Information
2. Gene Product
3. Cellular Localization Data
4. Predictive Open-Reading Frames
5. Structure-Based Enzyme
6. Enzymatic Function
7. Duplication and Degradation
8. Homology Gene Transfer

**Sequence based similarity:**

This model was used to determine the homology of the protein sequence to KEGG substrate to similar sequences in other organisms. The protein sequence for the gene was extracted as the gene sequence in the NCBI Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The software gives the gene product name, organism, sequence length (in amino acids), and E value. The protein sequence was then compared from other organisms and arranged them according to the best match of search sequence. The sequence with the lowest E value and highest bit score was selected.

**Domain Database (COG) accessed through the topmost BLAST results pga shows how the gene product is related to other orthologous proteins. The figure shows the clusters of orthologous groups (COG) based on identity to query sequence.

**Results**

**Figure 1:** KEGG Pathway map for biosynthesis of chorismate from erythrose-4-phosphate. Gene highlighted in green are the gene of *K. Sedentarius*.

**Figure 2:** Pathway diagram for biosynthesis of chorismate from erythrose-4-phosphate. Depiction of proteins showing the statistics for the alignment.

**Figure 3:** Domain neighbor map for gene OID 644991533 illustrating absence of transmembrane helices. Cut-off value of probability for prediction of location is 0.5.

**Figure 4:** Gene neighbor map for gene OID 644991533 illustrating 1) Gene encoding chorismate synthase 2) Gene encoding dihydrodiketone 3) Gene encoding dihydro-hydroquinone synthase.

**Figure 5:** Gene ortholog neighbor map for gene OID 644990396 encoding 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase illustrating absence of a peak indicating that the protein lacks a signal peptide.

**Figure 6:** Graphical representation of ORF search. The arrow points to the first nucleotide of the sequence. No red highlights indicates no start codon has been called by the gene caller.

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

**Figure 8:** 3-dehydroquinate dehydratase is a key enzyme in the shikimate pathway. The figure shows the enzyme's domain orientation, domain boundaries, and the consensus sequence aligned with the homologs.

**Figure 9:** Alignment of the top BLAST hit for gene OID 644991532 encoding chorismate synthase with the red highlighted area showing the statistics for the alignment.

**Conclusions**

Annotating all of the 7 genes involved in the biosynthesis of chorismate from erythrose-4-phosphate revealed that the selection of the enzymes involved in the synthesis of chorismate was found to be involved in 2 steps in the pathway: 1) Conversion of 2-keto-3-deoxy-D-glycero-d-gluco-heptulosonate 7-phosphate (KDG) to shikimate-3-phosphate 2) Conversion of chorismate to shikimate. The gene OID 644990396 was found to be involved in 2 steps in the pathway: 1) Conversion of 2-keto-3-deoxy-D-glycero-d-gluco-heptulosonate 7-phosphate (KDG) to shikimate-3-phosphate 2) Conversion of chorismate to shikimate.

**References**


**Figure 10:** A) Graphical representation of ORF search. The arrow points to the first nucleotide of the sequence. No red highlights indicates no start codon has been called by the gene caller.