Gene Cart

IMG users can save genes of interest to **Gene Cart** for further studies. There are many ways to save genes into Gene Cart; for example,

- Use **Gene Search** to find genes of interests to add to **Gene Cart**;
- Find genes of interest during browsing;
- Find genes of interest in some functional analysis results;
- Load previously genes of interest saved in a **Workspace Gene Set**.

After a user adds some genes into Gene Cart, there will be many analysis options as shown in Figure 1:

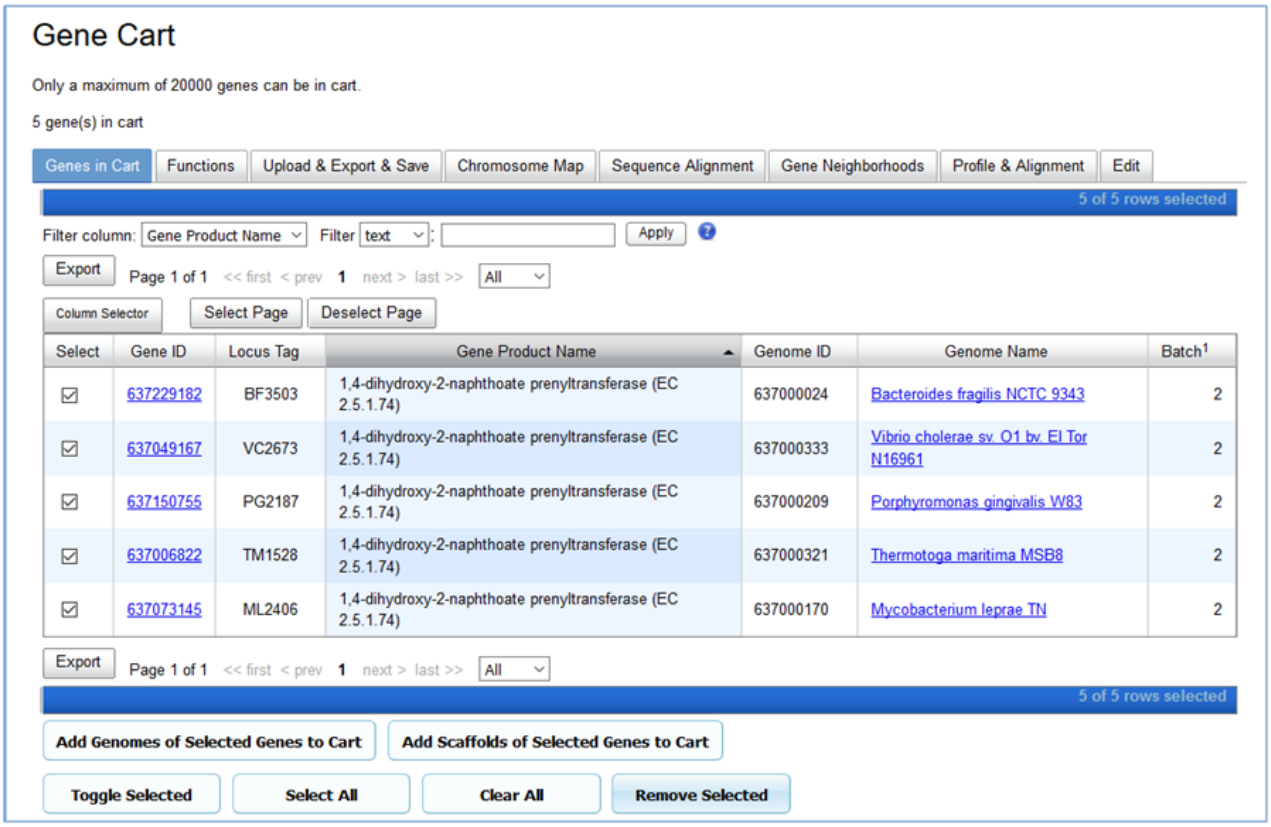

Figure 1. Gene Cart.

Genes in Cart

The first tab **Genes in Cart** shows all genes that have been added to the Gene Cart. Users can select up to 20,000 genes to add to the cart. The default list shows the following fields:

- Gene ID: IMG Gene OID;
- Locus Tag: locus tag of this gene;
- Gene Product Name: product name of this gene;

- Genome ID: The IMG Taxon OID of the genome which this gene belongs to;
- Genome Name: the name of the genome;
- Batch: An integer to distinguish genes added in different batches.

Users can select additional fields to be display in the **Table Configuration** section immediately below (see Figure 2).

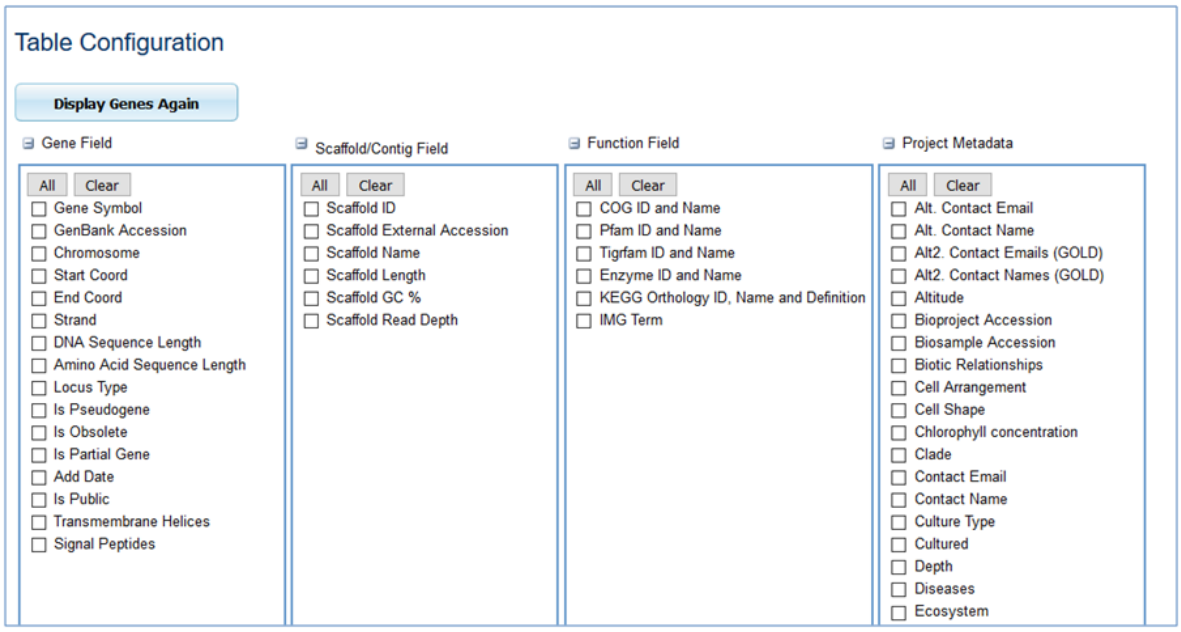

Figure 2. Gene Cart Table Configuration.

The fields are divided into four categories:

- Gene Field: Gene Symbol, GenBank Accession, Chromosome, Start Coord, End Coord, Strand, DNA Sequence Length, Amino Acid Sequence Length, Locus Type, Is Pseudogene, Is Obsolete, Is Partial Gene, Add Date, Us Public, Transmembrane Helices, Signal Peptides
- Scaffold/Config Field: Scaffold ID, Scaffold External Accession, Scaffold Name, Scaffold Length, Scaffold GC %, Scaffold Read Depth
- Function Field: COG ID and Name, Pfam ID and Name, Tigrfam ID and Name, Enzyme ID and Name, KEGG Orthology ID, Name and Definition, IMG Term
- Project Metadata: This includes all metadata fields from GOLD such as Depth, Disease, Culture Type, Ecosystem, etc.

Users can select fields of interest and then click the **Display Genes Again** button to show the additional fields.

Click the **Add Genomes of Selected Genes to Cart** button in Figure 1 will add genomes of all selected genes into the **Genome Cart**.

Click the **Add Scaffolds of Selected Genes to Cart** button in Figure 1 will add scaffolds/contigs of all selected genes into the **Scaffold Cart**.

There are 4 additional buttons under the table display:

- **Toggle Selected**: Un-select any preselected genes and select all genes that have not been selected previously.
- **Select All**: Select all genes in the cart.
- **Clear All**: Un-select all genes in the cart.
- **Remove Selected**: Remove all selected genes from the cart.

Functions

The **Functions** tab provides a couple of functional analysis operations for selected genes.

Users can add functions associated with selected genes to the Function Cart. To do so, first select a function category (COG, Pfam, TIGRfam, EC, InterPro, KO, MetaCyc, IMG Terms, IMG Pathways, IMG Parts List), and then click the **Add to Function Cart** button (see Figure 3(i)).

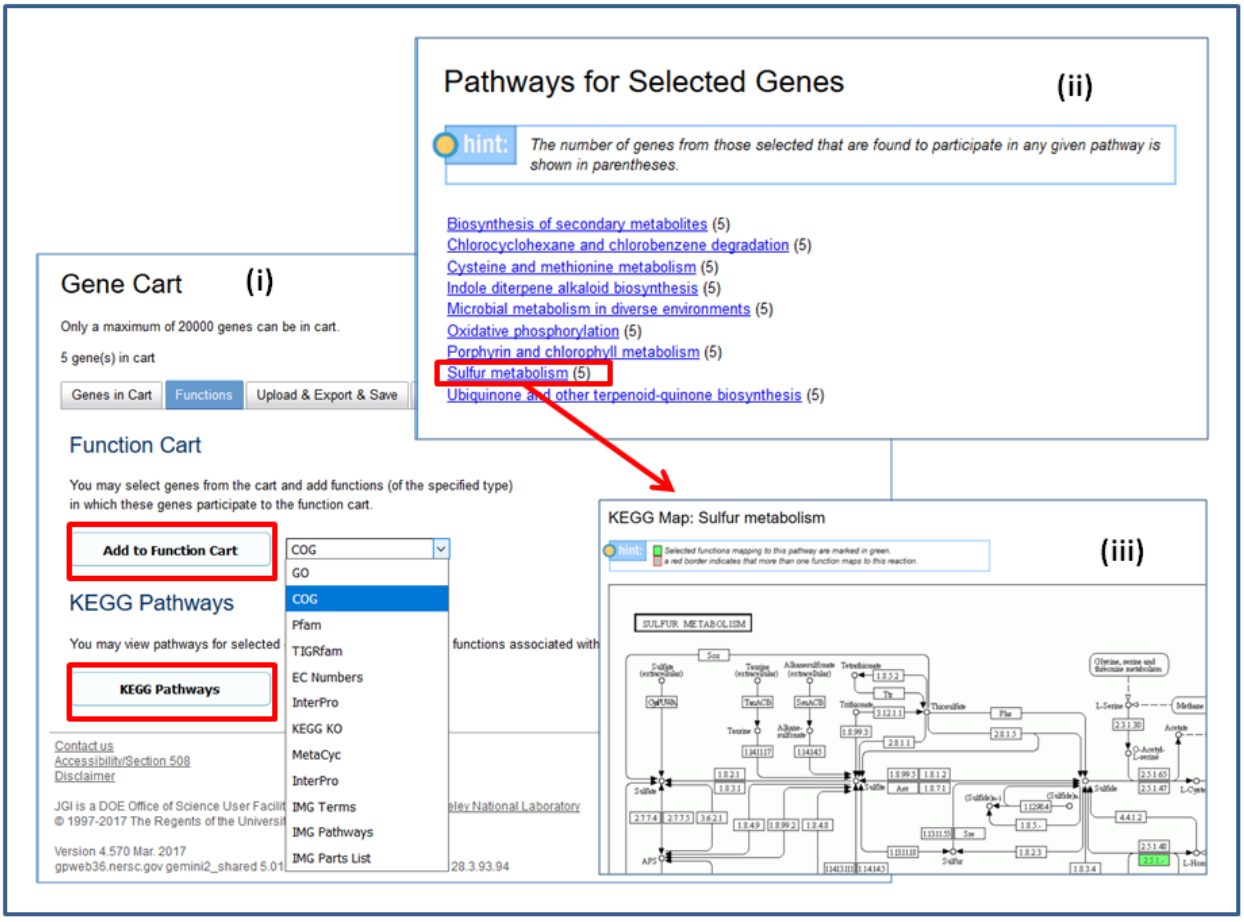

Figure 3. Gene Cart: Functions.

Clicking the **KEGG Pathways** button will lead to a new page listing all KEGG Pathways the selected genes participate (Figure 3(i) and (ii)). Users can click on a KEGG Pathway link to view the pathway map with highlights showing enzymes associated with the selected genes (see Figure 3(iii)).

Upload & Export & Save

This tab includes 3 functions: **Upload Gene Cart**, **Export Genes** and **Save Genes to My Workspace**.

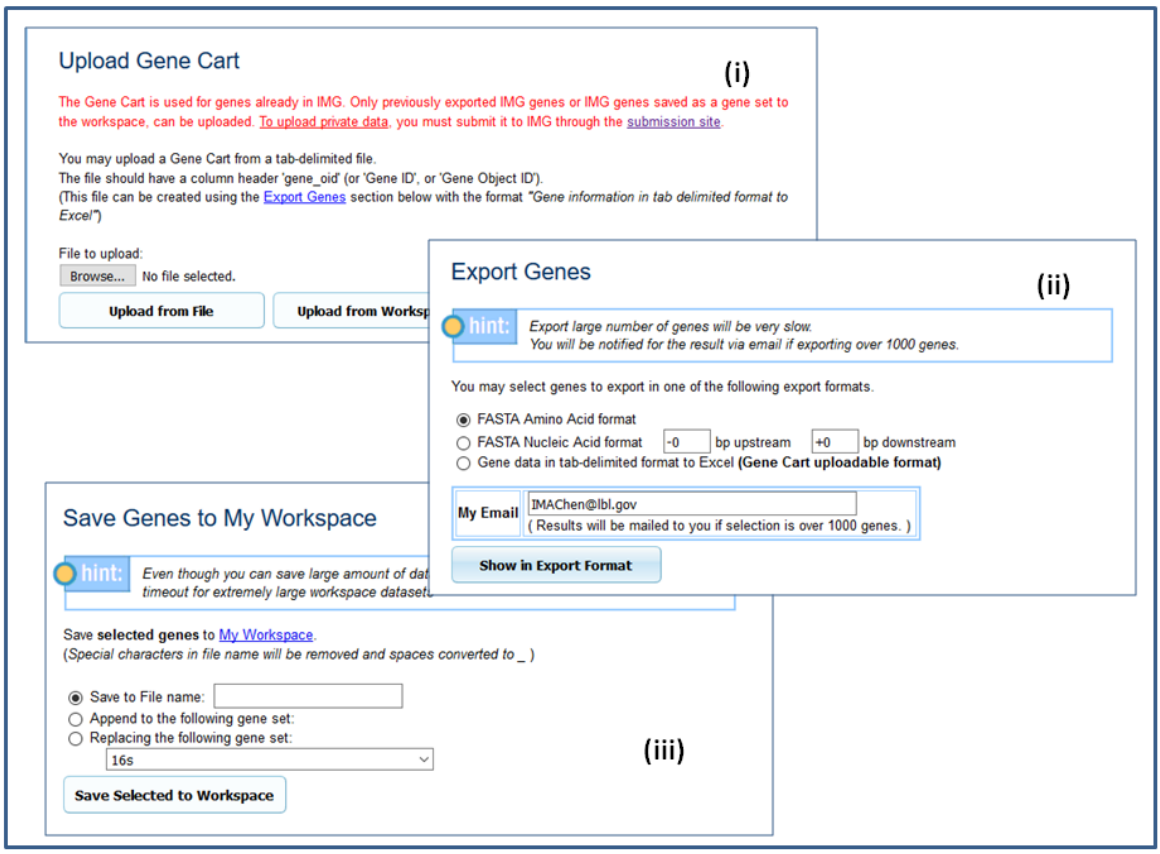

Figure 4. Gene Cart: Upload & Export & Save.

Upload Gene Cart

Users can either upload genes from a tab-delimited file or from Workspace Gene Set (Figure 4(i)). If a user wishes to select the **Upload from File** option, then the selected file must have a column header 'gene_oid', 'Gene ID', or 'Gene Object ID'. If the user is not sure what the file format should be, he/she can use the **Export Genes** function described below to obtain an example file.

Clicking the **Upload from Workspace** button will lead to the Workspace Gene Set page for gene set selection.

Export Genes

Users can export amino acid FASTA, nucleic acid FASTA or basic gene information in tab-delimited format (see Figure 4(ii)). Gene information in tab-delimited format can be saved for future uploading. If

the number of genes is very large (e.g., over 1000 genes), then the export process will run in the background, and the user will receive email notification of the result.

Save Genes to My Workspace

Users can also save selected genes in the cart to a Workspace Gene Set. The result can be saved to a new gene set, appended to an existing gene set, or replacing an existing gene set (see Figure 4(iii)).

Chromosome Map

The **Chromosome Map** tab provides a function for users to view chromosome maps of selected genes in the cart. Click the **Chromosome Map** button will lead to a table display of scaffolds of selected genes. Select proper band assignment, and then click the **Draw Map** button to view chromosome map(s) (see Figure 5).

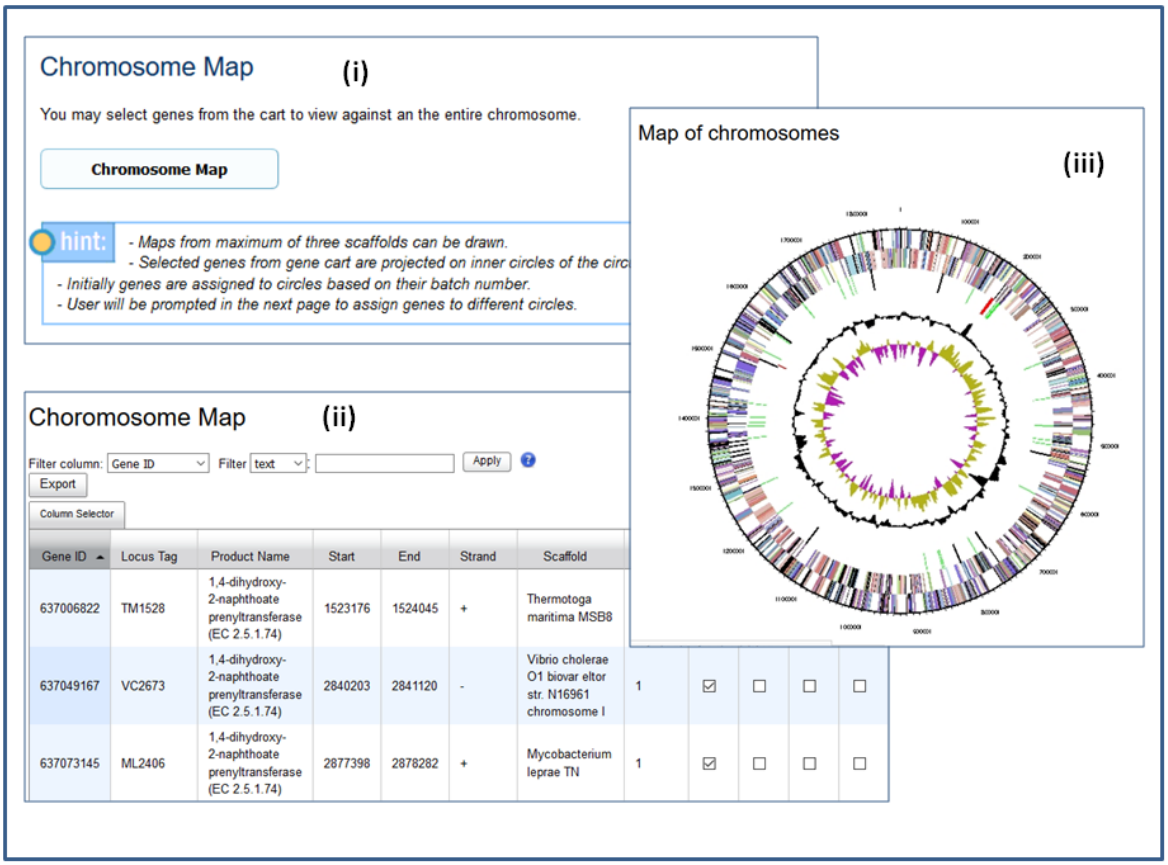

Figure 5. Gene Cart: Chromosome Map.

Sequence Alignment

The **Sequence Alignment** tab allows users to view protein or DNA sequence alignment of selected genes in the cart.

Figure 6. Gene Cart: Sequence Alignment.

Select either *Protein* or *DNA* option, and then click the **Do Alignment** button (see figure 6(i)) to view the protein alignment or DNA alignment of selected genes as shown in Figure 6(ii) and 6(iii), respectively. *Clustal Omega* (http://www.clustal.org/omega/) sequence alignments are displayed using *BioJS MSA* Viewer (http://msa.biojs.net/). Users can click on any gene name to view the gene detail.

The **Analyzed Genes** tab lists all selected genes (Figure 7(i)), and the **Rectangular Phylogram** tab shows gene clustering (Figure 7(ii)).

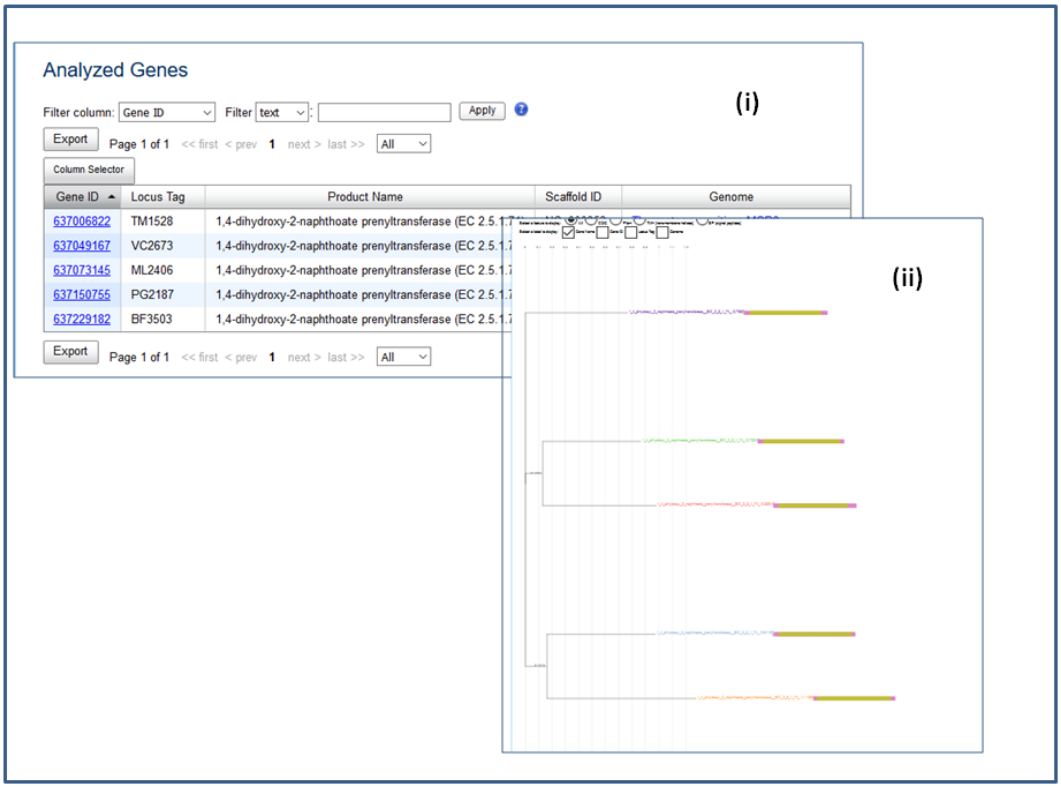

Figure 7. Gene Cart: Sequence Alignment -- Analyzed Genes and Rectangular Phylogram.

Gene Neighborhoods

Gene neighborhoods function allows users to view neighborhoods of selected genes in the Gene Cart. There are two display options:

- 5'-3' direction of each selected gene is left to right
- 5'-3' direction of plus strand is always left to right, on top

Select either option and then click the **Show Neighborhoods** button (in Figure 8(i)) to view the gene neighborhoods display (see Figure 8(ii)). Users can mouse over to view gene description or click to view the gene detail page.

Figure 8. Gene Cart -- Gene Neighborhoods

Profile & Alignment

The **Profile & Alignment** tab provides three functions: **Gene Profile**, **Occurrence Profile**, and **Function Alignment**.

Gene Profile

Gene Profile function allows users to view selected protein coding genes against selected genomes using unidirectional sequence similarities. Users will have to select 1 to 1000 genomes first, and then click either **View Genes vs. Genomes** or **View Genomes vs. Genes** button (see Figure 9(i)). The result table (in Figure 9(ii)) shows unidirectional sequence similarities for selected genes with BLAST cutoffs at specified minimum % *identity*, and maximum *E-value*. Users can click on any non-zero count to view gene list or gene detail (Figure 9(iii)). It is also possible to perform the analysis again using different minimum % *identity* and/or maximum *E-value*.

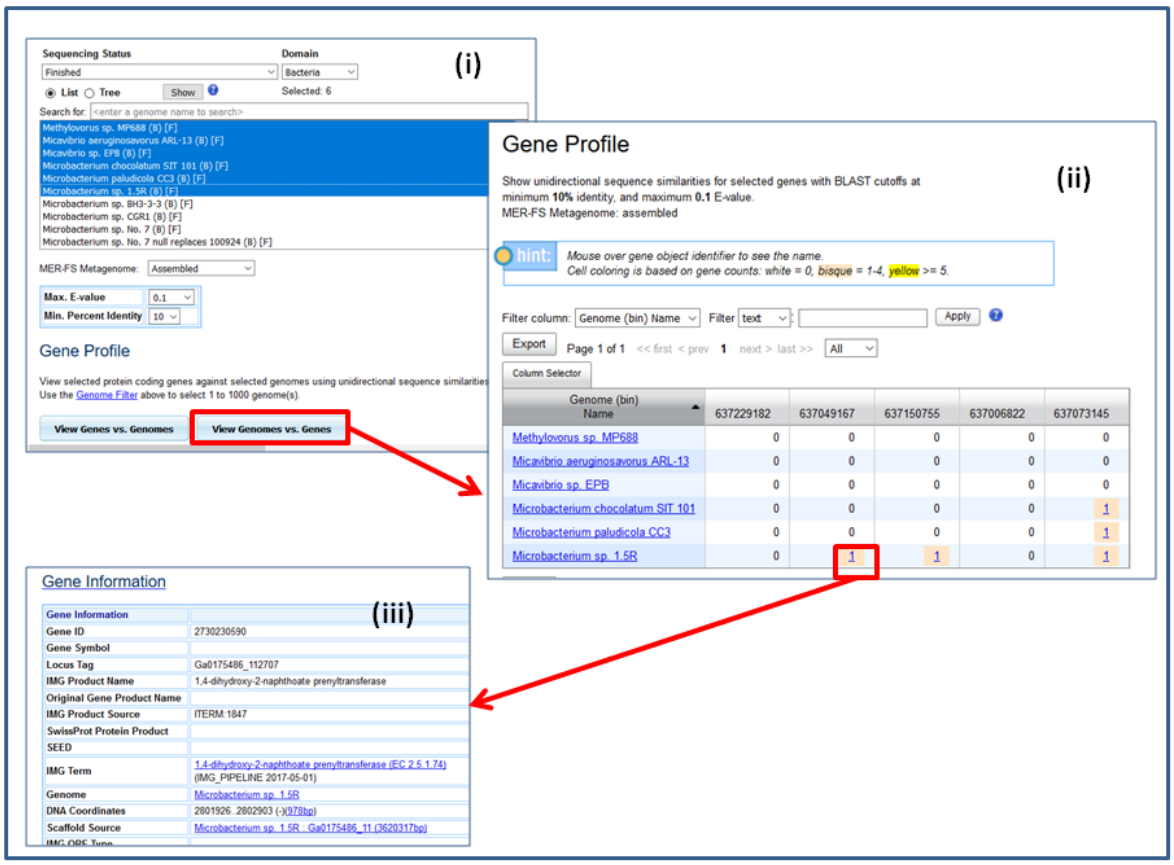

Figure 9. Gene Cart -- Gene Profile.

Occurrence Profile

This function shows phylogenetic occurrence profile for selected genes. There is a limitation of no more than 100 genes. Users can change the default *E-value* and *percent identity* cutoff. Click the **View Phylogenetic Occurrence Profiles** button to view the analysis result as shown in Figure 10. Users can mouse over the result to view genome hits.

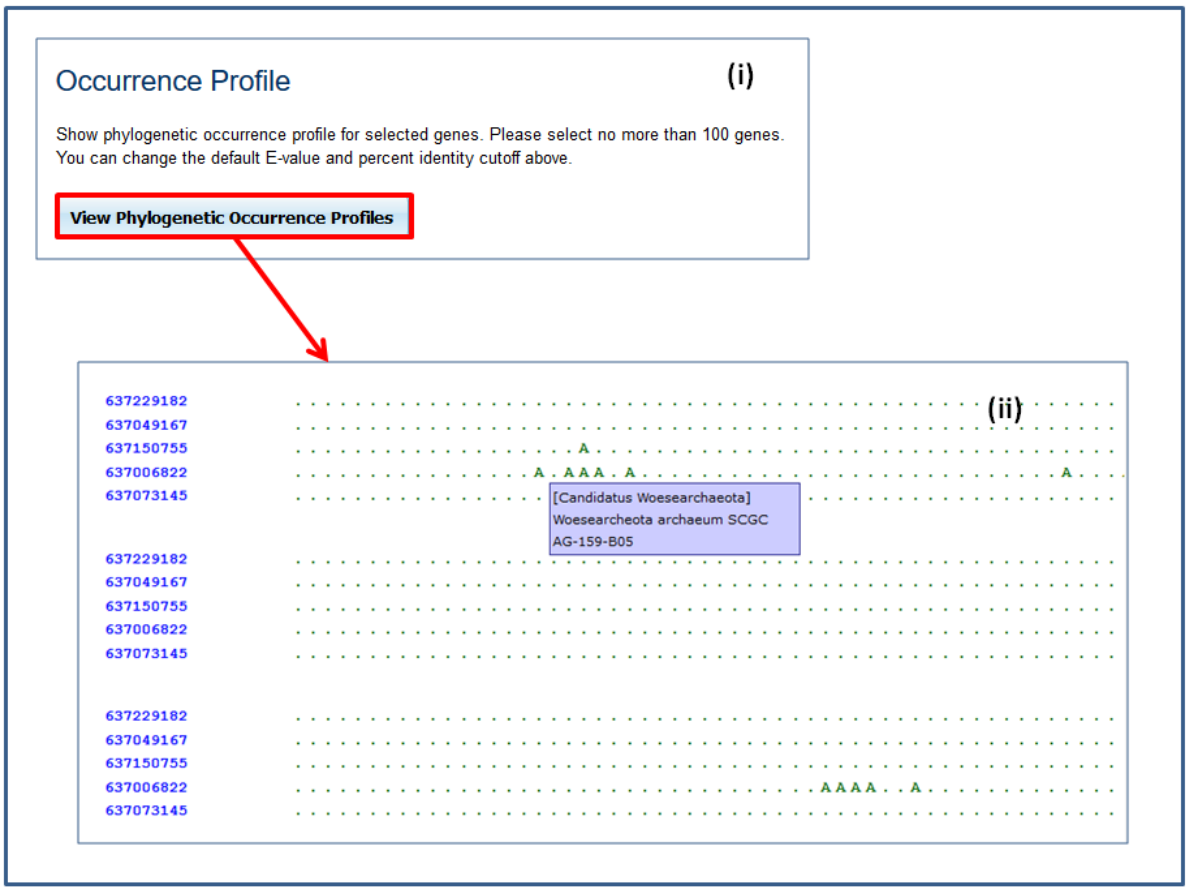

Figure 10. Gene Cart -- Occurrence Profile.

Function Alignment

This function lists alignments of function prediction for selected genes (limit to COG, KOG and pfam). Metagenome Genes are not supported. In order to perform this analysis, first select genes in the Gene Cart and then click the **Function Alignment** button (see Figure 11(i)).

The result shows hits against COG (Figure 11(ii)), KOG, and Pfam (Figure 11(iii)).

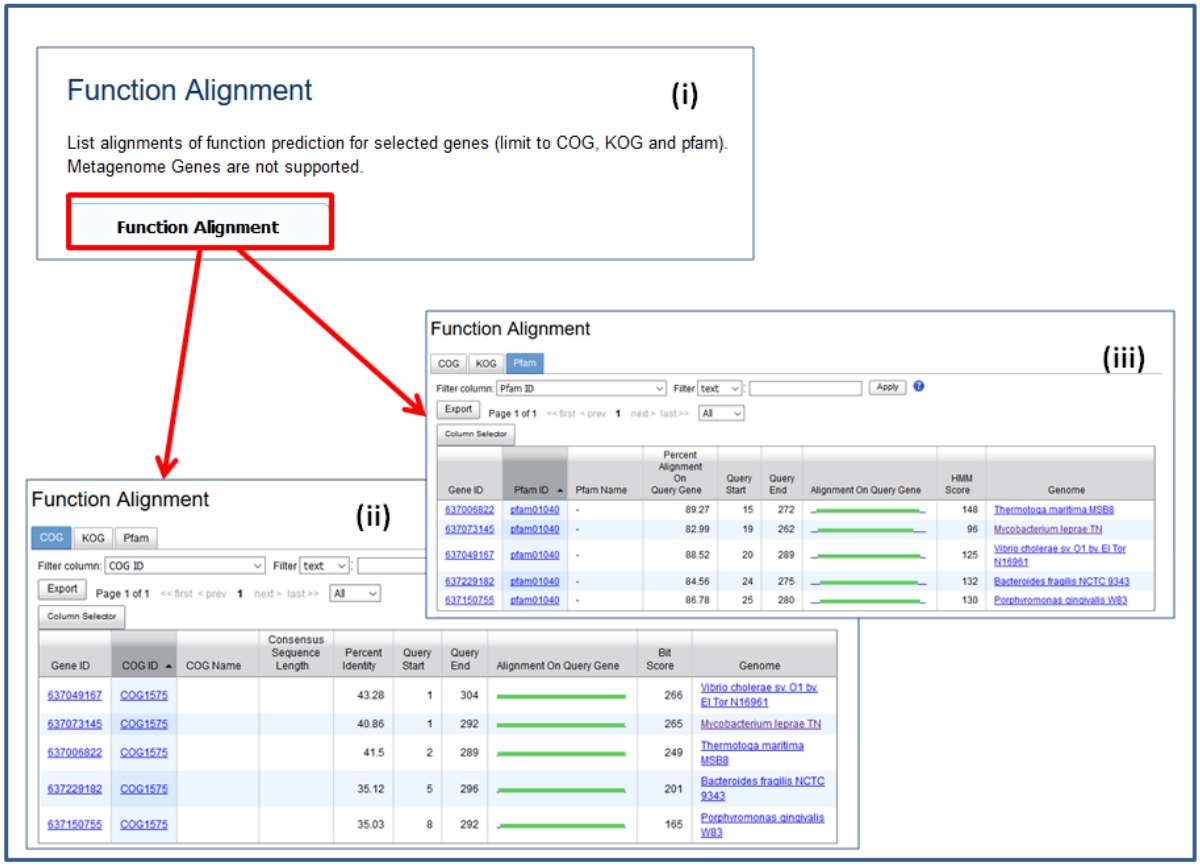

Figure 11. Gene Cart -- Function Alignment.

Edit

The **Edit** tab provides a function for users to enter MyIMG annotations of selected genes. Click the **Annotate Selected Genes** button (see Figure 12(i)) to view and to edit MyIMG annotations (see Figure 12(ii)).

Please refer to the MyIMG User Guide for more detail re. entering and updating MyIMG gene annotations.

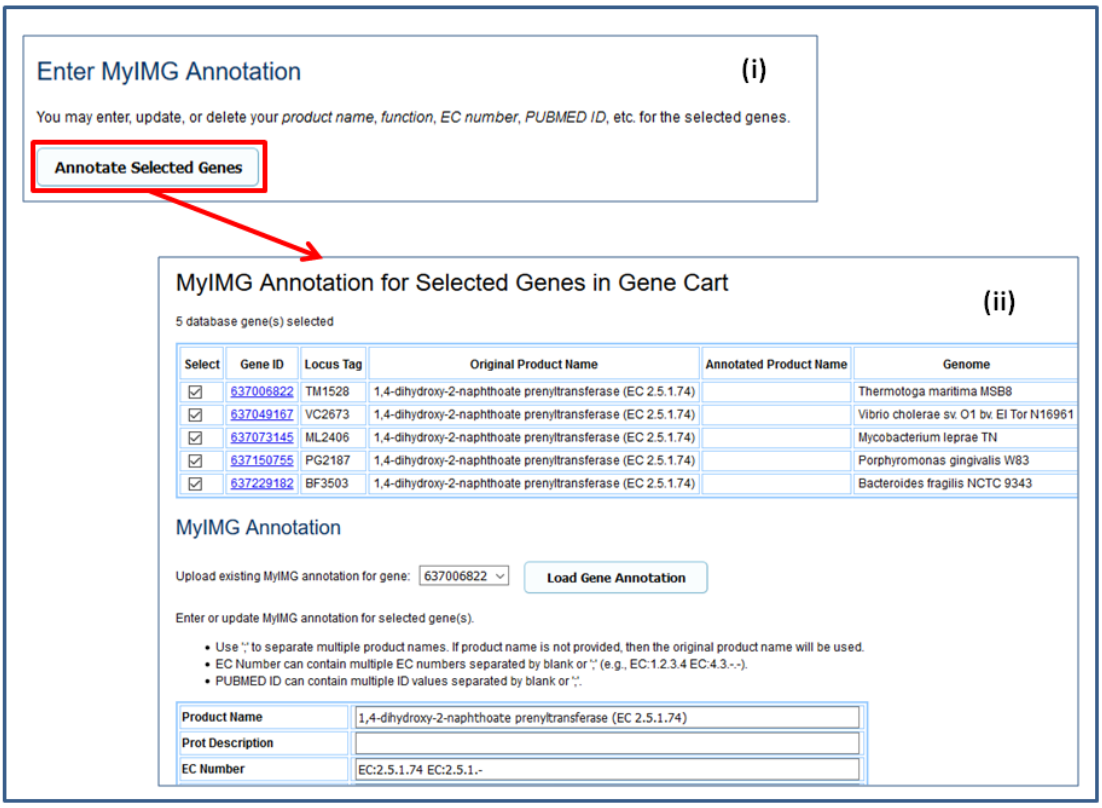

Figure 12. Gene Cart -- Enter MyIMG Annotation.