Example .bed file format for multi-genome gene model input

| chromosome ID | start coordinate | stop coordinate | gene model ID | genome ID | strand |
|------------------|---------------------|--------------------|-----------------|--------------|--------|
| chr1 | 34617 | 40204 | Zm00001eb000010 | B73 | + |
| chr1 | 41214 | 46762 | Zm00001eb000020 | B73 | - |
| chr1 | 47662 | 53272 | Zm00030ab000010 | Ki11 | + |
| chr1 | 54156 | 59780 | Zm00030ab000020 | Ki11 | - |
| chr1 | 99852 | 104723 | Zm00040ab000020 | P39 | - |
| chr1 | 104584 | 115154 | Zm00040ab000030 | P39 | - |

This .bed file format is the required input for the multi-genome database build. Chromosomal coordinates make up the first three columns, followed by gene model ID, the genome ID (this genome ID will correspond to the genome IDs manually entered in the multi-genome "Genes in an Interval" genome drop-down menu; see Supplemental Figure S1), and strand. Genome ID can be either a genome's official ID or an alias.

Example metadata format

| File_handle | Data_id | Source | Туре | Link | Description |
|--|---------|---|------------|--|--|
| eightDAS_root | , . | NAM Consortium | Expression | | Plants were either field- or greenhouse-grown and roots were collected eight days after sowing. |
| | , | NAM Consortium | Expression | 0 | Plants were either field- or greenhouse-grown and embryos were collected 16 days after pollination. |
| whole_seed_2 4_days_after_ pollination | | Diepenbroc k 2017 [DellaPenna Lab] | | https://www. ncbi.nlm.nih. gov/pmc/arti cles/PMC577 | All plants in each plot were self-pollinated, and pollination dates were recorded. A single ear from a given plot was harvested at (each of) 12, 24, and 36 DAP. Immediately after harvest, whole ears were frozen in liquid nitrogen. The ears were stored at _80 degrees C until kernels could be removed from the still-frozen ears, placed in test tubes, and maintained at _80 degrees C. |

| whole_seed_1 2_days_after_ | whole seed 12 days after | - | | https://www. ncbi.nlm.nih. gov/pmc/arti cles/PMC577 | All plants in each plot were self-pollinated, and pollination dates were recorded. A single ear from a given plot was harvested at (each of) 12, 24, and 36 DAP. Immediately after harvest, whole ears were frozen in liquid nitrogen. The ears were stored at _80 degrees C until kernels could be removed from the still-frozen ears, |
|-------------------------------|-----------------------------|------|------------|--|---|
| pollination | pollination | Lab] | Expression | 4569/ | placed in test tubes, and maintained at _80 degrees C. |

This is the required metadata .csv format for the database build, including the column headers File_handle, Data_id, Source, Type, Link, and Description. File_handle must have the exact same name as the abundance file, minus the file extension. For instance, the file_handle "eightDAS_root" corresponds to an abundance .txt file named "eightDAS_root.txt". Note that file_handle IDs and their associated .txt or .fpkm_tracking files cannot begin with a numeral, and cannot contain any non-alphanumeric character except an underscore ("_"). File handles can be descriptive, such as eightDAS_root, or have SRA numbers from NCBI (i.e. SRR12345 corresponding to a file called SRR12345.fpkm_tracking). "Source" is the name of the consortium or project that originally sequenced the data; "Link" contains the URL that links the Source to the project publication or the project website. "Description" contains general methods used to generate the library.

Example abundance .txt file format

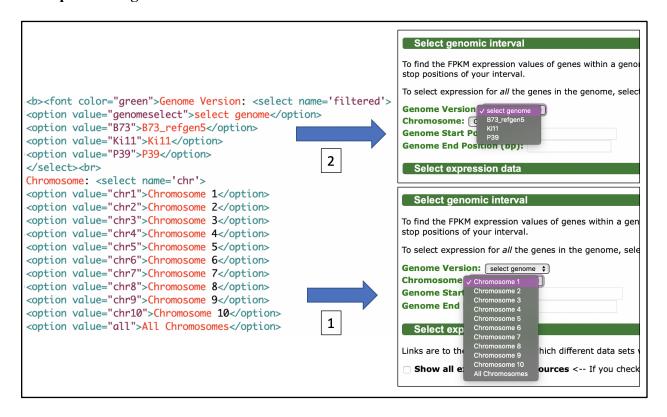
Cana madal ID

| Gene model ID | abundance |
|-----------------|------------|
| Zm00001eb000010 | 1.700325 |
| Zm00001eb000020 | 52.9438 |
| Zm00001eb000030 | 0.080943 |
| Zm00030ab061810 | 0.0817445 |
| Zm00030ab061820 | 16.4987 |
| Zm00030ab061830 | 0 |
| Zm00040ab201460 | 0.6890385 |
| Zm00040ab201470 | 0.428948 |
| Zm00040ab201490 | 0.29024503 |

Above is a subset of the "eightDAS_root.txt" input for the multi-genome database build. Gene model IDs for all genomes (distinguished by the prefixes Zm00001eb = B73, Zm00030ab =

Ki11, and Zm00040 = P39) are in Column A, and abundances (in this case, RNA-Seq fpkm values calculated by Cufflinks) is in Column B.

Drop-down organization for PHP files



Drop-down menu changes for chromosome ID (1) that must be manually edited in index_singlegenome.php, index_multigenome.php, and Protein_index.php files. Above is an example of the drop-down configurations for index_multigenome.php file on the left. At the bottom left (1) is the chromosome ID configuration for the "Genes in an Interval" Chromosome ID drop-down (lower right); this chromosome ID script is in all three different index_*.php files and must be edited in each to reflect the chromosome IDs of the target genome(s). For index_multigenome.php only (2), the Genome Version drop-down menu must be edited to reflect the genomes from the multi-genome bed file (described in Supplemental Table S1). Note that <option value= for the Genome Version dropdown menu in the php file corresponds to the genome ID listed in Column 5 of the bed file described in Supplemental Table S1.