**1. Raw input files**

Association files, imported from the PO subversion repository revision 1772 (<http://palea.cgrb.oregonstate.edu/viewsvn/Poc/trunk/associations/?pathrev=1772>):

po\_temporal\_gene\_arabidopsis\_tair.assoc

po\_growth\_genemodel\_zea\_MaizeGDB.assoc

These files are too large for github and can be viewed at the URL above.

**2. Edited input files**

At = Arabidopsis thaliana

Zm = Zea mays

Edited the association files so that they only contain rows for the classes of interest (subclasses of plant pro-ebryo stage) and removed one association from proembryo stage that wasn’t IEP (inferred by experimental procedure) or IDA (inferred by direct assay):

At\_concatenate\_files\_all\_columns.txt

Zm\_concatenate\_files\_all\_columns.txt

Extracted only column with gene names (col. 10 for Arabidopsis and col. 2 for maize) then removed any duplicate rows. This left two files with just the lists of genes for all plant embryo development stages:

At\_embryo\_genes\_uniq.txt

Zm\_embryo\_genes\_uniq.txt

The two files above were used as input for the InParanoid analysis (to determine homolog clusters). The output file contains the orthologs from the supercluster algorithm:

At\_Zmays\_inpara\_Zea\_mays.txt

The first column is a cluster id and will tell you if more than one cluster pair are orthologs. All genes with the cluster ID is the same are considered orthologs of each other.  This is the same as TableS2.txt, in the supplementary files.

The file above contains many genes that are not present in our association files. They were removed by deleting all rows for maize genes didn’t start with GRMZ (using grep -v -e EF -e AY -e AF -e AC At\_Zmays\_inpara\_Zea\_mays.txt )

This left 23185 rows (without header):

At\_Zmays\_inpara\_Zea\_mays\_GRZM.txt

Then removed ortho scores less than 0.5. Leaves 20096 rows (without header):

At\_Zmays\_inpara\_Zea\_mays\_50.txt

From this file, “.1” was removed from all the At gene ids and duplicate rows were removed. Extra columns were cut, leaving just two columns (id and gene).

At\_homolog\_ids.txt

A similar process was used for Zm genes, leaving two columns (id and gene):

Zm\_homolog \_ids.txt

In summary:

**At\_homolog\_ids.txt** – contains unique rows, with only columns for id and gene, with only those rows that had scores >= 0.50 for At genes. (13630 rows)

**Zm\_homolog \_ids.txt** – contains unique rows, with only columns for id and gene, with only those rows that had scores >= 0.50 for Zm genes. (10557 rows)

**3. Create species and stage specific sets of genes /homolog**

Sets of overlapping homologs were created from the two edited files – that is, all homologs that are present in both At and Zm. There were no genes that aren’t present in both species, because this list only contains those genes that passed through the Inparanoid analysis and thus includes only genes present in shared gene families.

The next step was to separate out homologs by each species/stage combinationband compare within and across stages.

To do this, we went back to the lists of genes for each stage and created new files that just have homolog IDs plus gene, by matching the lists of genes in each stage to Xx\_homolog\_ids.txt, where Xx is At or Zm. This resulted in the following files:

Atbilat.txt – 13823 rows

Atcotyl.txt – 13914 rows

Atglob.txt - 13763 rows

Atmature.txt – 13322 rows

Atproemb.txt – 10 rows

Zmcoleop.txt – 25065 rows

Zmproemb.txt – 27117 rows

Zmtruelf.txt – 28663 rows

Homolog IDs were added to the above files using the Excel match function:

Atbilat\_homolog.txt – 9781 rows

Atcotyl\_homolog.txt – 9831 rows

Atglob\_homolog.txt – 9741 rows

Atmature\_homolog.txt – 9519 rows

Atproemb\_homolog.txt – 8 rows

Zmcoleop\_homolog.txt – 8483 rows

Zmproemb\_homolog.txt – 8705 rows

Zmtruelf\_homolog.txt – 8765 rows

**4. Compare sets of genes across stages and species**

The tool at <http://nemates.org/MA/progs/Compare.html> was used to compare lists of gene or ortholog IDs to determine overlap among pairs of stages within and across species. All pairwise comparison were performed and listed below:

Within species, gene IDs:

Atbilat.txt vs. Atcotyl.txt

Atbilat.txt vs. Atglob.txt

Atbilat.txt vs. Atmature.txt

Atcotyl.txt vs. Atglob.txt

Atcotyl.txt vs. Atmature.txt

Atglob.txt vs. Atmature.txt

Zmcoleop.txt vs. Zmproemb.txt

Zmcoleop.txt vs. Zmtruelf.txt

Zmproemb.txt vs. Zmtruelf.txt

Within species, ortholog IDs:

Atbilat\_orthos.txt vs. Atcotyl\_orthos.txt

Atbilat\_orthos.txt vs. Atglob\_orthos.txt

Atbilat\_orthos.txt vs. Atmature\_orthos.txt

Atbilat\_orthos.txt vs. Atproemb\_orthos.txt

Atcotyl\_orthos.txt vs. Atglob\_orthos.txt

Atcotyl\_orthos.txt vs. Atmature\_orthos.txt

Atcotyl\_orthos.txt vs. Atproemb\_orthos.txt

Atglob\_orthos.txt vs. Atmature\_orthos.txt

Atglob\_orthos.txt vs. Atproemb\_orthos.txt

Atmature\_orthos.txt vs. Atproemb\_orthos.txt

Zmcoleop\_orthos.txt vs. Zmproemb\_orthos.txt

Zmcoleop\_orthos.txt vs. Zmtruelf\_orthos.txt

Zmproemb\_orthos.txt vs. Zmtruelf\_orthos.txt

Between species, ortholog IDs:

Atbilat\_orthos.txt vs Zmcoleop\_orthos.txt

Atbilat\_orthos.txt vs Zmproemb\_orthos.txt

Atbilat\_orthos.txt vs Zmtruelf\_orthos.txt

Atcotyl\_orthos.txt vs Zmcoleop\_orthos.txt

Atcotyl\_orthos.txt vs Zmproemb\_orthos.txt

Atcotyl\_orthos.txt vs Zmtruelf\_orthos.txt

Atglob\_orthos.txt vs Zmcoleop\_orthos.txt

Atglob\_orthos.txt vs Zmproemb\_orthos.txt

Atglob\_orthos.txt vs Zmtruelf\_orthos.txt

Atmature\_orthos.txt vs Zmcoleop\_orthos.txt

Atmature\_orthos.txt vs Zmproemb\_orthos.txt

Atmature\_orthos.txt vs Zmtruelf\_orthos.txt

Atproemb\_orthos.txt vs Zmcoleop\_orthos.txt

Atproemb\_orthos.txt vs Zmproemb\_orthos.txt

Atproemb\_orthos.txt vs Zmtruelf\_orthos.txt

The results are summarized in

Overlap\_summary.xlsx

**5. Create non-overlapping sets of genes between stages**

To facilitate comparisons among sets of genes, we performed all pairwise comparisons using <http://nemates.org/MA/progs/Compare.html> among species/stage specific gene and ortholog ID sets and created files containing list of genes that were in one stage but not the other. These files are stored in the folder “5No\_overlap\_genes” using names of the form:

Atbilat\_not\_Atcotyl.txt

...

Ortholog IDs cannot be used directly in Gene Ontology (GO) enrichment, so they were converted back into gene IDs using the match function in Excel.

As an example:

Match Atbilat\_not\_ Zmcoleop.txt to Atbilat\_orthos.txt using =IF(ISERROR(MATCH(A1,$E$1:$E$1249,0)),"",B1)

Sort only At gene names that match orthos.

Save working file as match\_Atbilt\_not\_Zmcoleop.xlsx in (in no overlap genes)

Saved list of At genes only match\_At\_orthos\_to\_bilatnotcoleop.txt

These files are stored in the folder “5no\_overlap\_genes” using names of the form:

match\_At\_orthos\_to\_bilatnotcoleop.txt

...

**6. Unique genes for each stage**

By species, we generated lists of genes unique to each stage by comparing lists of genes annotated to each stage to the complete list of genes for every stage but the one of interest, <http://nemates.org/MA/progs/Compare.html>. We first generated the lists for comparison:

At\_minus\_glob.txt

At\_minus\_bilat.txt

At\_minus\_cotyl.txt

At\_minus\_mature.txt

Zm\_minus\_proemb.txt

Zm\_minus\_coleop.txt

Zm\_minus\_truelf.txt

We then compared these to the lists of genes annotated to a stage (e.g., At\_minus\_glob.txt to At\_glob.txt) to generate lists of unique genes in each stage:

At\_glob\_u.txt 781 genes of 13739

At\_bilat\_u.txt 134 genes 0f 13798

At\_cotyl\_u.txt 246 genes of 13898

At\_mature\_u.txt 48 genes of 13319

Zm\_proemb\_u.txt 855 genes of 27117

Zm\_coleop\_u.txt 47 genes of 25065

Zm\_truelf\_u.txt 2169 genes of 28663

These lists are summarized in table S5, under supplementary files.

**7. Compare GO enrichment profiles**

Two sets of comparisons were made:

1. Intraspecific: Genes that were annotated to the globular stage but not the mature stage in Arabidopsis, versus genes that were annotated to the mature stage but not the globular stage in Arabidopsis.

2. Interspecific: Genes that were annotated to the coleoptilar stage in maize but not the extended cotyledonary stage in Arabidopsis, versus all genes that were annotated to the extended cotyledonary stage in Arabidopsis.

**7.1 Intraspecific comparison**

Using the non-overlapping sets of genes generated in set 5 above, we used AgriGO to compare the GO enrichment profile of Atglob\_not\_ATmature.txt (genes annotated to the globular stage but not the mature stage in Arabidopsis) to Atmature\_not.\_Atglob.txt (genes that were annotated to the mature stage but not the globular stage in Arabidopsis).

**Settings:**

Analysis tool: SEA

Supported species: Arabidopsis thaliana TAIR 10

Reference from selected background: Arabidopsis genome locus (TAIR 10)

Advanced options: PO slim

**Results:**

## Analysis Brief Summary

**Job ID:** 363277779 [Useful within 7 days]

**Job Name:** AtglobNOTmature

**Species:**

**GO type:** Plant GO slim

**Background/Reference:** Customized

**Annotated number in query list:** 1156

**Annotated number in background/reference:** 28397

**Significant GO terms:** 46

## Analysis Brief Summary

**Job ID:** 688842085 [Useful within 7 days]

**Job Name:** AtmatureNOTglob

**Species:**

**GO type:** Plant GO slim

**Background/Reference:** Customized

**Annotated number in query list:** 722

**Annotated number in background/reference:** 28397

**Significant GO terms:** 30

The results above were then used in the AgriGO tool SEACOMPARE, and the comparison results were saved as:

seacompare\_glob\_mature.xlsx

**7.2 Intrerspecific comparison**

Using the non-overlapping sets of genes generated in set 5 above, we used AgriGO to compare the GO enrichment profile of match\_Atorthos\_to\_coleopnotcotyl.txt (genes that were annotated to that were annotated to the coleoptilar stage in maize but not the extended cotyledonary stage in Arabidopsis) to Atcotyl.txt (genes that were annotated to the extended cotyledonary stage in Arabidopsis).

**Settings:**

Analysis tool: SEA

Supported species: Arabidopsis thaliana TAIR 10

Reference from selected background: Arabidopsis genome locus (TAIR 10)

Advanced options: PO slim

**Results:**

## Analysis Brief Summary

**Job ID:** 250031567 [Useful within 7 days]

**Job Name:** coleopnotcotyl

**Species:**

**GO type:** Plant GO slim

**Background/Reference:** Customized

**Annotated number in query list:** 267

**Annotated number in background/reference:** 28397

**Significant GO terms:** 14

## Analysis Brief Summary

**Job ID:** 900997977 [Useful within 7 days]

**Job Name:** Atcotyl

**Species:**

**GO type:** Plant GO slim

**Background/Reference:** Customized

**Annotated number in query list:** 13817

**Annotated number in background/reference:** 28397

**Significant GO terms:** 134

The results above were then used in the AgriGO tool SEACOMPARE, and the comparison results were saved as:

Seacompare\_coleopnotcotyl\_cotyl\_all.xlsx

File that includes just the GO terms from coleopnotcotyl:

Seacompare\_coleopnotcotyl\_cotyl.xlsx

**8. Supplementary tables from manuscript**

The folder “8Supplementary\_tables” contains the following supplementary files from Walls, Cooper et al. (DOI will be supplied upon publication):

Ortholog clusters as generated by InParanoid:

TableS2.txt

List of genes or gene models annotated to only one Plant Ontology plant embryo development stage:

TableS5.csv