# List of scripts and associated datasets

## DFEAnalysis

* **sfs**
  + SFS files for the four focal species, including bootstrap SFS
* **output\_polydfe**
  + Output files of polyDFE
* **scripts**
  + LaunchPolyDFE.sh: script to launch polyDFE
  + init\_model\_BandC.txt: initial parameters for polyDFE
  + range\_model\_BandC.txt: range of parameters for polyDFE
  + generateSFS.R: to generate SFS and bootstrap SFS from polymorphism data
  + processpolyDFE.R: to process polyDFE results and plot corresponding figures

## Orthology

* Files with names of the form “FocalSpecies\_Hordeum\_reciprocalbestblast.txt”: list of orthologs with blast information
* Files with name of the form “FocalSpecies\_Hordeum\_ds.txt”: list of orthologs with ds information
* HordeumGenes.txt: files with all Hordeum genes, position and some attributes
* **Scripts**
  + ScriptBlast.sh: reciprocal blast between focal species and hordeum
  + ScriptBlastSorting.R: sort and clean blast outputs and retrieve the best reciprocal best blast
  + ScriptBestBlastAlignDs.sh: align focal and Hordeum orthologs and compute the dS

## PolymorphismData

* One folder per species
* In each folder
  + Outputs of the dNdSpiNpiS program: one summary file, one file with all statistics per contig
  + For the four focal species: two additional files with information on each SNP: frequency and polarization (using Taeniatherum caput-medusae as outgroup), base composition polymorphism

## Recombination

* MareyMap\_Hordeum\_SNPJHutton2012\_On\_EnsemblPlantsVersionHv\_IBSC\_PGSB\_v21.txt: Marey map (for each markers: physical and genetic position)
* RecombinationRates\_AllHordeumGenes.txt: for each gene on the Hordeum genome: estimation of local recombination rate with three methods
* Files with name of the form “FocalSpeciesXXcM.txt: input dataset for the linked selection model (widows of xx cM)
* Files with name of the form “Results\_BSfit\_filter1000\_windowXXcM.txt”: results of the linked selection model
* **Scripts**
  + ScriptFigure3.R: to generate figure 3
  + ScriptFigureSupPiAlongChromosomes.R: to the generate the equivalent of figure 3 for all species (Figure S13)
  + LinkedSelectionSimple.R: Script to fit a simple, phenomenological, linked selection model (corresponds to values given on Figure 3)
  + GenerateFileForLinkedSelection\_bycM.R: script to generate datasets used as input for fighting linked selection model
  + EstimationLinkedSelection.R: to fit a mechanistic linked selection model
  + FigureLinkedSelectionResults.R: joint analysis of the files “Results\_BSfit\_filter1000\_windowXXcM.txt” to build figures
  + FigurePolymorphism\_vs\_recombination.R: draw piS, piN/piS and Ds as a function of categories of recombination.

## SpeciesRange

* Aegilops\_filtered.txt, Triticum\_filtered.txt: extracted data from Gbif after filtering out errors
* Grid\_1\_0.jpg: grid used for ArcGIS

## MorphologyData

* Trait distribution and imputation of missing data:

**Input file**: Supplementary\_table\_Sx\_Morpho\_traits\_original.csv (measures of reproductive organs per individual and species)

**Script**: Morpho\_traits1\_diversity\_and\_imputation.nb.html

* Characterization of the selfing syndrome with Principal Component Analysis:

**Input files**:

- Supplementary\_table\_Sx\_Morpho\_traits\_imputed.csv (measures of reproductive organs with missing data imputed);

- Summary\_polymorphism\_Triticeae\_extended.csv (genomic diversity estimates per species);

- MLtree\_OneCopyGenes\_ultrametric.tree (species tree for phylogenetic contrasts).

**Scripts**:

- Morpho\_traits2\_PCA.nb.html

- Morpho\_traits2\_PCA\_with\_3OUTGROUPS.nb.html

## PolymorphismMorphologyCorrelation

Analysis of the correlation between estimates of neutral polymorphism, purifying selection, species range and linked selection with the morphological selfing syndrome

**Input files**:

- Summary\_polymorphism\_Triticeae\_extended.csv (genomic diversity and range estimates per species)

- PCA\_morphological\_traits.Rdata (results of PCA analysis on morphology data)

- MLtree\_OneCopyGenes\_ultrametric.tree (species tree for phylogenetic contrasts).

**Script**: Morpho\_traits3\_correlation\_with\_diversity.nb.html