**Analysis plan EPPN project**

1. **Level: individual parent-progeny sets**

* Aim: compare parents to progeny (individual sets)
* Questions:
  + Are progeny phenotypes intermediate between parent phenotypes, more similar to Parent 1, more similar to Parent 2, or do they show transgressive segregation (above or below both parent values) for each trait?
  + Are progeny sets significantly different to each other? (comparison 2: novel allohexaploids)
* Planned output: Table for each of the four comparison groups
* Statistical analyses:
  + T-Test for comparing progenys set to P1, P2 and MPV in comparisons 1, 3 and 4
    - Shapiro-Wilk-Test
    - Levenes Test
* Coding path

Once this is all sorted, we can pick which traits it makes sense to show as examples in figure etc.

1. **Level: within the comparison groups**

* Aims:
  + To see if resynthesised *B. napus* traits are more similar to natural *B. napus* or to the parent *B. rapa* and *B. oleracea* parent lines (comp. 1)
  + To determine if the NCJ and JO allohexaploids have more similarity to each other than to their parent genotypes (comp. 2)
  + To determine if the JC hybrids have more similarity to each other than to their parent genotypes (comp. 3)
  + See if the progeny show hybrid vigour relative to the parents (all at the allohexaploid level). Look at Comparisons 2 and 4 together (parent diploid and tetraploid genotypes, comp 4)
* Questions:
  + Do hybrids have more similarity to each other than to their parent genotypes?
  + Does progeny show hybrid vigour relative to hexaploid parents and to diploid and tetraploid progenitors?
  + Are there species-specific traits?
* Output:

1. **Level: across groups (if we have similar trends in the individual and within-group comparisons)**

* Aims:
  + To see if there are similarities across groups
* Questions:
* Output

**Ideas on further analyses**

* calculate mid parent heterosis and better parent heterosis according to formula.
* Heat map showing differences between parent mean and progeny
  + transgressive segregation most interesting phenomenon
* heat map with maternal genotypes as references => maternal effects / differences mother vs progeny
* Check correlation between traits
  + Trade-off between growth and fertility?
* PCA for traits
* Distribution of progeny trait values regarding MPH and BPH? Normal distribution? Frequency of MPH and BPH?

**Sets of comparisons: ploidy level, number of genomes and heterozygosity**

**comparison 1: *B. rapa*, *B. oleracea*, synthetic *B. napus* from known parent genotype combinations**

* AA (diplod, homozygous) x CC (diploid, homozygous) = AACC (tetraploid, 2 genomes, homozygous)
* Effects of ploidy level and number of genomes (=allopolyploids)

**comparison 2: *B. juncea* J1, *B. carinata* C1 and C2, and the *B. juncea* x *B. carinata* tetraploids**

* AABB (tetraploid, 2 genomes, homozygous**?**) x BBCC (tetraploid, 2 genomes, homozygous**?**) = BBAC**?** (tetraploid, **3** genomes, homozygous)
* Effects of number of genomes. Same ploidy level.

**comparison 3: *B. oleracea* TO1000, *B. juncea* J3, and their allohexaploid O1J3**

* CC (diploid, homozygous) x AABB (tetraploid, 2 genomes, homozygous**?**) = AABBCC (hexaploid, 3 genomes, homozygous**?**)
* Effects of ploidy level (= number of different genomes, in this case)

**comparison 4: the 7 allohexaploid lines from *B. napus* x *B. carinata* x *B. juncea* combinations, with their *B. napus*, *B. juncea* and *B. carinata* parents**

* Order of crossing?
* AACC x BBCC x AABB = AABBCC
* Effects of ploidy level (= number of different genomes, in this case)

**a complex set of comparisons: each of the heterozygous hexaploids to their two parent hexaploid lines (and maybe all of them to the species parents, but let's see first how reasonable this is)**

* AABBCC (homozygous) x AABBCC (homozygous) = AABBCC (heterozygous)
* Effects of heterozygosity. Same ploidy level / number of different genomes.

**Consider this additional information for comparison group 4**

PH32 Fully heterozygous parent combination

PH33 Fully heterozygous parent combination

PH34 Fully heterozygous parent combination

PH37 Fully heterozygous parent combination

PH38 Shared B genome genotype between parents

PH41 Shared A genome genotype between parents

PH43 Shared C genome genotype between parents

PH46 Parent lines both extremely stable

PH49 Best between-species cross