What I would like to have in the end:

1. A frequency distribution analyses of all the TEs found in my C. grandiflora population.

This would include a discrimination between:

* siRNAs+ TEs vs siRNAs- TEs
* Centromere vs chromosome arm
* TEs in CDS vs TEs close to genes vs TEs far away from genes
* Retrotransposons vs DNA transposons

1. A frequency distribution analyses of the TEs found in the C. rubella reference genome and in the C. grandiflora population.

This would include a discrimination between:

* TEs targeted by uniquely mapping small RNAs vs TEs targeted by multi mapping small RNAs vs TEs not targeted by small RNAs
* Centromere vs chromosome arm
* TEs in CDS vs TEs close to genes vs TEs far away from genes

1. A frequency distribution analyses of the TEs found in the C. rubella reference genome and in the C. grandiflora population linked to the expression level of the genes next to them.

This would include a discrimination between:

* Genes with high expression vs genes with low expression
* Additionally, the discriminations in point 2

1. An age of allele analyses of the TEs found in the C. rubella reference genome
2. A 4-fold SFS of the SNPs found in the C. grandiflora population.