Seurat analysis with user defined highly variable genes

Strategy 2 ::

- step 1 - Intersection of highly variable genes between replicates of a species

- step 2 - Union of highly variable genes between species

Filtering criteria ::

1. No genes have been removed :: min.cells filtering criteria has not been used

2. Remove low-quality cells & cells with high total read count ::

- If expression of at least 200 genes have not been detected a cell, the cell has been removed

- Cells with a total read count more than 110000 have been removed

- Cells with read count from mitochondrial genes equal or more than 5% and read count from chloroplast genes more than 10% have been removed

Normalization method :: Log-Normalization

Selection of highly variable genes ::

- 3000 highly variable genes have been selected for each replicate of *Cardamine hirsuta*

- 2000 highly variable genes have been selected for the Apex of *Arabidopsis thaliana* (only one replicate)

Integration feature : Strategy 2 (step 1 + step 2)

Dimension reduction :: 50 PCs have been used to compute UMAP and TSNE

Cell clusters identification :: 50 PCs have been used to identify cell clusters with a range of resolution parameters, 0.1 – 1.2 with 0.1 increment