scanpy.pp.recipe_zheng17

scanpy.pp.recipe_zheng17(*adata, n_top_genes=1000, log=True, plot=False, copy=False*)

Normalization and filtering as of [Zheng17].

Reproduces the preprocessing of [Zheng17] – the Cell Ranger R Kit of 10x Genomics.

Expects non-logarithmized data. If using logarithmized data, pass `log=False`.

The recipe runs the following steps

```
sc.pp.filter_genes(adata, min_counts=1)          # only consider genes with more than 1 count
sc.pp.normalize_per_cell(                          # normalize with total UMI count per cell
     adata, key_n_counts='n_counts_all'
)
filter_result = sc.pp.filter_genes_dispersion(   # select highly-variable genes
    adata.X, flavor='cell_ranger', n_top_genes=n_top_genes, log=False
)
adata = adata[:, filter_result.gene_subset]      # subset the genes
sc.pp.normalize_per_cell(adata)                    # renormalize after filtering
if log: sc.pp.log1p(adata)                         # log transform: adata.X = log(adata.X + 1)
sc.pp.scale(adata)                                 # scale to unit variance and shift to zero mean
```

Parameters:	**adata** : `AnnData`
	Annotated data matrix.

n_top_genes : `int` (default: `1000`)

Number of genes to keep.

log : `bool` (default: `True`)

Take logarithm.

plot : `bool` (default: `False`)

Show a plot of the gene dispersion vs. mean relation.

copy : `bool` (default: `False`)

Return a copy of `adata` instead of updating it.

Return type: `Optional` [`AnnData`]

Returns: : Returns or updates `adata` depending on `copy` .