DESeq2 differential gene expression analysis results legend

Tabular results:

*raw_gene_counts_GeneSymbols_GeneTypes.tsv   File containing raw counts per gene from ReadsPerGene.out.tab generated by STAR aligner. This file is suitable as input for various tools that calculate differential expression (DESeq2, edgeR, EBSeq...).
Columns:
1   GeneType
2   GeneSymbol
3   Ensembl Gene ID
4-   sampleIDs with raw counts (all integer numbers)

*vst_transf_counts_GeneSymbols_GeneTypes.tsv or
*reg_log_transf_counts_GeneSymbols_GeneTypes.tsv   F
ile containing regularized log2-transformed count data, generated by DESeq2 function `vst` or `rlog`. These transformed data are recommended for the generation of sample distance plots, PCA plots, heatmaps etc. (rather than the raw counts), as described here:
http://bioconductor.org/packages/devel/bioc/vignettes/DESeq2/inst/doc/
DESeq2.html#count-data-transformations
Columns:
1   GeneType
2   GeneSymbol
3   Ensembl Gene ID
4-   sampleIDs with transformed counts

*[samplegroup]_DESeq2results_GeneSymbols_GeneTypes.tsv File containing the results of the differential gene expression test performed by DESeq2 functions `DESeq` and `results`. One or several of these files may be generated for different group comparisons (samplegroup).
Columns:
1   GeneType
2   GeneSymbol
3   Ensembl Gene ID
4   baseMean   intermediate mean of normalized counts for all samples
5   log2FoldChange   log2 fold change (maximum likelihood estimate) for group 2 vs group 1
6   lfcSE   standard error for group 2 vs group 1
7   stat   Wald statistic: group 2 vs group 1
8   pvalue   Wald test p-value: group 2 vs group 1
9   padj   Benjamini-Hochberg adjusted p-values
10-   sampleIDs with normalized counts (normalized by size factors)

Plots:

These plots are based on the transformed count data and thus give a global overview of the samples in the experiment. The plots help to inspect sample groups and potential outliers.

*EuclDist.png   Plot of Euclidean Distance between samples. Sample groups are expected to cluster together (dark blue), unrelated samples from different groups are expected to be more distant from each other (light blue). The dendrogram additionally highlights the sample relationships.

* **PCA_biplot_with_grids_labels_samplegroup.svg** Plot of the first 2 dimensions of a principal component analysis, based on the transformed count data, samples may be colored by individual or by sample group, a. o. .