Dear Participants
In order to provide you with ECTS credits, we need to have a final control step, in our case – a report (approx. 3 pages). Let's set the deadline for the submission – the 20th of January. I am ready to answer your questions, if needed.

As we discussed, you have 2 opportunities:
(A) to write a report related to your own data;
(B) to work on the proposed topic (see case problem below).

A report for case (A) should contain:

1. A brief description of the area.
2. Goal of the study.
3. Description of the data.
4. List of methods used for the analysis.
5. Results and conclusions.

A report for case (B) should contain:

1. Data presentation.
2. Answer to the question.
3. Results and conclusions.
4. your original R source code.

Case problem: mRNA / miR data for 8 cell lines

Data files

http://edu.sablab.net/data/txt/mRNA_8CL_pub.txt
http://edu.sablab.net/data/txt/miRNA_8CL_pub.txt

contain expression values of mRNA and miR measured for 8 cell lines. mRNA and miR arrays were performed for the same samples.

a) Present the data (distributions, boxplots, description)

b) Investigate data using PCA. Draw conclusions about mRNA and miR data reproducibility. Which cell lines are similar?

c) Cell line A375 is characterized by overexpression of MAGEA4 gene. Can you identify this cell line among other skin cancer cell lines?

d) Perform differential expression analysis (DEA) b/w all given skin cell lines and all other cell lines. Report findings: significant miRs and mRNAs and sample clustering based on to significant genes.

e) MITF is known to be a target of mir-182. Compare expression profiles for this gene and miR.

f) *(optional)* Filter out miR and mRNAs with low expression and small variance. For the rest, build correlations b/w mRNA and miR profiles. Compare negatively correlated miR/mRNA pairs with TargetScan predictions (see predicted pairs in the file http://edu.sablab.net/data/TargetScan_Human_2012.zip)

The following custom online functions may help you

plotDataPDF – plot distribution of your data:	source("http://sablab.net/scripts/plotDataPDF.r")
limmaEBS2Class – differential expression analysis:	source("http://sablab.net/scripts/limmaEBS2Class.r")