Microarray-based mRNA expression analysis - Gene regulation by mRNA export protein, THOC5

1) Experimental study design

In each of two biological experiments six different cytoplasmic RNA samples were generated from bone marrow derived macrophages of one ROSA26-CreER^{T2} control mouse and one CreER^{T2} Thoc5 (flox/flox) mouse, respectively. These samples correspond to the following six conditions:

1) Cells from ROSA26-CreER^{T2} control mouse at day 0 (0d differentiation)
2) Cells from ROSA26-CreER^{T2} control mouse 3 days after CSF-1 treatment (3d differentiation)
3) Cells from ROSA26-CreER^{T2} control mouse 3 days after CSF-1 treatment in the presence of tamoxifen (3d differentiation + tamoxifen)
4) Cells from CreER^{T2} THOC5 (flox/flox) mouse at day 0 (0d differentiation)
5) Cells from CreER^{T2} THOC5 (flox/flox) mouse 3 days after CSF-1 treatment (3d differentiation)
6) Cells from CreER^{T2} THOC5 (flox/flox) mouse 3 days after CSF-1 treatment in the presence of tamoxifen (3d differentiation + tamoxifen)

"Day 0 samples" of a given genotype were always utilized as common reference in two microarrays and were accordingly co-hybridized against both "3-days differentiated samples" (treated or none with tamoxifen) of the same genotype and biological experiment. Thus, four dual-color microarrays (Microarrays #1-4) from each of the two biological replicate series (Experiments #1 and #2, respectively) were performed, giving rise to 8 dual-color microarrays in total. The complete set of microarrays from the second biological experiment was performed with inverted labeling directions (dye-swap). An overview of the experimental study design is given in Table 1.

Experiment #	Microarray #	sample description	condition #	RNA	cRNA	Ratio	Ratio #
1	1	BoneMarrowMacrophages(ERT2_Cre_control_mouse#35) 0d differentiation	1	6855	6855.1.aaU(1+11)1.(555)1	red / green	1
		BoneMarrowMacrophages(ERT2_Cre_control_mouse#35) 3d differentiation	2	6859	6859.1.aaU(1+11)1.(647)1		
	2	BoneMarrowMacrophages(ERT2_Cre_control_mouse#35) 0d differentiation	1	6855	6855.1.aaU(1+11)1.(555)1	red / green	2
		BoneMarrowMacrophages(ERT2_Cre_control_mouse#35) 3d differentiation+tamoxifen	3	6860	6860.1.aaU(1+11)1.(647)1		
	3	BoneMarrowMacrophages(ERT2_Cre_Thoc5_flox/flox_mouse#398) 0d differentiation	4	6857	6857.1.aaU(1+11)1.(555)1	red / green	3
		BoneMarrowMacrophages(ERT2_Cre_Thoc5_flox/flox_mouse#398) 3d differentiation	5	6865	6865.1.aaU(1+11)1.(647)1		
	4	BoneMarrowMacrophages(ERT2_Cre_Thoc5_flox/flox_mouse#398) 0d differentiation	4	6857	6857.1.aaU(1+11)1.(555)1	red / green	4
		BoneMarrowMacrophages(ERT2_Cre_Thoc5_flox/flox_mouse#398) 3d differentiation+tamoxifen	6	6866	6866.1.aaU(1+11)1.(647)1		
2	1	BoneMarrowMacrophages(ERT2_Cre_control_mouse#36) 0d differentiation	1	6856	6856.1.aaU(1+11)1.(647)1	green / red	1
		BoneMarrowMacrophages(ERT2_Cre_control_mouse#36) 3d differentiation	2	6862	6862.1.aaU(1+11)1.(555)1		
	2	BoneMarrowMacrophages(ERT2_Cre_control_mouse#36) 0d differentiation	1	6856	6856.1.aaU(1+11)1.(647)1	green / red	2
		BoneMarrowMacrophages(ERT2_Cre_control_mouse#36) 3d differentiation+tamoxifen	3	6863	6863.1.aaU(1+11)1.(555)1		
	3	BoneMarrowMacrophages(ERT2_Cre_Thoc5_flox/flox_mouse#400) 0d differentiation	4	6858	6858.1.aaU(1+11)1.(647)1	green / red	3
		BoneMarrowMacrophages(ERT2_Cre_Thoc5_flox/flox_mouse#400) 3d differentiation	5	6868	6868.1.aaU(1+11)1.(555)1		
	4	BoneMarrowMacrophages(ERT2_Cre_Thoc5_flox/flox_mouse#400) 0d differentiation	4	6858	6858.1.aaU(1+11)1.(647)1	green / red	4
		BoneMarrowMacrophages(ERT2_Cre_Thoc5_flox/flox_mouse#400) 3d differentiation+tamoxifen	6	6869	6869.1.aaU(1+11)1.(555)1		

Table 1: The experimental study design.

2) Experimental procedure and data processing

The “Whole Mouse Genome Oligo Microarray V2” (G4846A, ID 026655, Agilent Technologies) used in this study contains 44397 oligonucleotide probes covering roughly 32000 murine transcripts. Synthesis of Alexa555- or Alexa647-labeled cRNA was performed

with the “Amino Allyl MessageAmp™ II aRNA Amplification Kit” (#5190-0444, Life Technologies) according to the manufacturer’s recommendations, except that the molar proportion of used aminoallyl-UTP and UTP was adjusted to 1+11 (instead of 1+1). cRNA fragmentation, hybridization and washing steps were carried-out exactly as recommended in the “Two-Color Microarray-Based Gene Expression Analysis Protocol V5.7” (Agilent Technologies). Slides were scanned on the Agilent Micro Array Scanner G2565CA (pixel resolution 5 µm, bit depth 20). Data extraction, processing and intra-array normalization of raw fluorescence intensity values were performed with the “Feature Extraction Software V10.7.3.1” by using the recommended default extraction protocol file: GE2_107_Sep09.xml. For inter-array normalization, processed fluorescence intensity values of the green channel, (gProcessedSignal or gPS) or the red channel (rPS) of all microarrays were subjected to global linear scaling, except for the first microarray of the whole series which served as the reference. All gPS and rPS values were multiplied by an array-specific scaling factor which was calculated by dividing the gPS 75th percentile of the first microarray (Array #1) by the gPS 75th percentile value of the particular microarray to be normalized (Array i in the formula shown below). Accordingly, inter-array normalized PS values (nPS) for all samples (microarray data sets), all probe measurements from both channels (red and green) were calculated by the following formula:

$$nPS_{Array\ i} = PS_{Array\ i} \times (75^{th}\ Percentile\ gPS_{Array\ \#1} / 75^{th}\ Percentile\ gPS_{Array\ i})$$

3) Filtering procedure to identify THOC5-dependent transcripts

In order to identify THOC5-dependent transcripts, whole data were subjected to a multistep filtering procedure. All of the following criteria had to be satisfied:

a) Arithmetic mean intensity of nPS values calculated from both channels > 50 in each of the eight dual-color microarrays.

b) Resulting ratio value calculated from: “(Ratio 4 / Ratio 3) / (Ratio 2 / Ratio 1)” (see Table 1, columns 7 and 8) >2 (THOC5-dependent upregulation) or < 0.5 (THOC5-dependent downregulation) in each of the two biological experiments.
In other words:
Fold change for “tamoxifen-effect” in ERT2 Cre Thoc5 (flox/flox) cells over “tamoxifen-effect” in ERT2 Cre control cells >2 (THOC5-dependent upregulation) or <-2 (THOC5-dependent downregulation) in each of the two biological experiments.

c) Log2 ratio of microarray #4 > average log2 ratio calculated from microarrays #1-3 plus 3 standard deviations (THOC5-dependent upregulation) or < average log2 ratio of microarrays #1-3 minus 3 standard deviations (THOC5-dependent downregulation) in each of the two biological experiments.

d) The respective transcript has to be classified as being functionally characterized and reasonably annotated (for details concerning our probe classification system, visit: www.mh-hannover.de/Transcriptomics.html and consult our manual: “Crude probe characterization_RCUT_2012-08-22.pdf”).

e) Filtering was performed at the level of gene probe measurements rather than at the gene level. In cases when measurements of more than one probe, directed against a particular transcript, match the applied filter criteria, the value from that probe with the lowest Feature Number (provided in the Feature Extraction result file) was selected for final visualization.

4) Heatmap visualization and functional classification of filtering results

Heatmaps were generated by use of Microsoft Excel 2003. Color-codes for fold change values were allocated by executing visual basic macros, programmed in our lab. Functional classification of transcripts, matching applied filtering criteria was performed with annotations provided by Ingenuity Pathway Analysis, Content Version 11904312 (2012-01-17).

5) Rationale of overall experimental design and applied filtering strategy

In essence, the microarray-study had been designed to identify transcripts that were both, regulated during 3-days of CSF-1 induced macrophage differentiation and dependent on Thoc5 in proper expression during this tested period.

A consistently two-fold deviation in cytoplasmic mRNA abundance in THOC5 knockout samples compared to control conditions (normal differentiation) was fixed as final criterion. All annotated transcripts that matched this criterion were depicted in the heatmap of figure 2a. The quantitative fold deviation is directly represented in the color-code in columns 13 and 14, since all of the depicted transcripts show at least the lightest blue- or grey color, representing (at least) 2-fold elevated or diminished mRNA levels in THOC5 knockout samples, respectively. However, in order to clearly distinguish causal THOC5 dependence from experimental noise and methodical side-effects, it was necessary to integrate a number of control samples into the overall microarray design and to develop a robust multistep filtering strategy, in order to apply the twofold regulation threshold on accordingly corrected results. Fold-change vales as depicted in columns 13 and 14 can be regarded as corrected or, more precisely, two times normalized ratio values. Initial ratios, representing expression changes during 3 days of differentiation (columns 1-8, redish-greenish color scale) were first corrected for Tamoxifen side-effects giving rise to columns 9-12 (ratios of ratios, blue-grey color scale). Even though, of course not yet present in the depicted results in figure 2A, this normalization efficiently eliminated transcripts, responding to Tamoxifen apart from the CRE-mediated THOC5 knockdown situation.

The next normalization step additionally demands for specificity of the THOC5 knockdown effects against samples from control mice without floxed alleles. It was realized by calculating

ratios of “ratios of ratios” and finally gave rise to those very values, that were finally addressed for 2-fold deviations as filter criteria (see filter criterion 3b).

In order to exclude less-reliable measurements, resulting from either experimental bias at the low intensity end or from transcripts with extraordinary high variability in expression in the examined situation, appropriate filter criteria were additionally introduced (see filter criteria 3a and 3c, respectively).

6) Reevaluation of filtering results

In order to finally verify the results as being causally related to Thoc5 knock-out rather than reflecting intrinsic variability among samples (accidentally selected for by inappropriate filter criteria) we applied our filtering procedure after a virtual exchange (permutation) of the genotype status among samples and datasets (ERT2 Cre control samples were exchanged by ERT2 Cre Thoc5 (flox/flox) samples and vice versa). Applying the complete filtering procedure this way revealed no single up or downregulated transcript (compared to 65 and 100 transcripts revealed as up or downregulated with correct genotype allocation, respectively).

Additionally, hierarchical clustering on log ratios from all 8 microarrays revealed the highest degree of overall similarity among those two data sets that reflect expression signatures under actual Thoc5 knockout conditions from the two biological experiments (clustering applied on 19606 detectable entities with the following settings: Euclidean Similarity Measure and Average Linkage Rule, GeneSpring V12.0 program). Theses results show that Thoc5 knockout caused a specific and reproducible disturbance in the overall expression signature that could be distinguished from background variability among all control conditions analyzed in parallel (Figure 1).

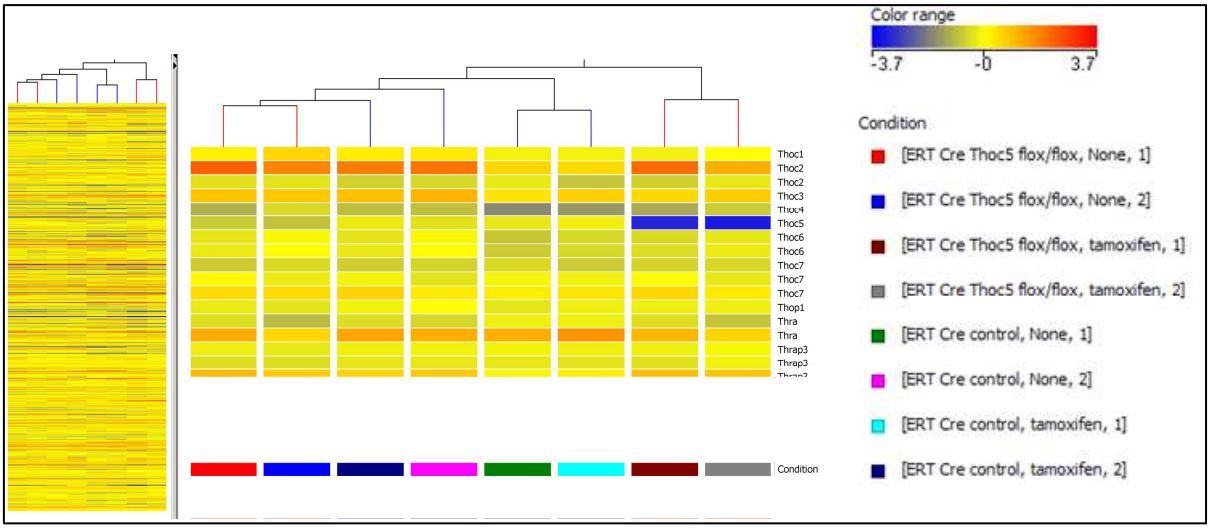

Figure 1: Clustering on log-ratios of all eight dual-color microarray datasets. Clustering was applied on log2 ratios of 19606 detectable entities in GeneSpring V12.0 program with the following settings: Euclidean Similarity Measure and Average Linkage Rule.