

Our research team aims to probe the developmental effects of various transcription factors in embryonic stem cell derived dendritic cells (ES-DCs) and their precursors. We have previously demonstrated that overexpression of the RUNX3 transcription factor enhanced the maturation and the immunogenicity of ES-DCs. In this project we analyzed the genomic impact of the RUNX3 and the ZBTB46 transcription factors during the mouse ES-DC development. We also planned to investigate the effect of RUNX3 in human ES-DCs. However, we realized that it would be very challenging to select transgenic clones using human pluripotent stem cells. Therefore, we focused on the murine embryonic stem cell (ESC) derived dendritic cells and their progenitors. We employed genome-scale gene expression profiling upon the forced expression of RUNX3 or ZBTB46 during the various stage of ES-DC differentiation using bulk and single cell RNA sequencing. In this summary I will explain in detail our unpublished transcriptomic data about the RUNX3 regulatory network.

1. ZBTB46 and EGR2 dependent gene expression and differentiation in ESC derived blood cells

Our goal is to investigate the developmental effects of the DC specific transcription factors during the differentiation of ESC derived DCs. Among the tested DC affiliated transcription factors, we found that RUNX3 and ZBTB46 elicited the strongest phenotypic changes during blood cell development. Therefore, in parallel with the RUNX3, we also assessed the phenotypic and transcriptomic effects of ZBTB46 in ES-DC precursors. We documented that ectopic expression of ZBTB46 interfered with the ESC derived mesoderm formation and reduced number of myeloid blood cells were produced in the presence of this transcription factor. In parallel with the myeloid repression, however, *Zbtb46* overexpression was associated with enhanced erythroid blood cell development and increased adult hemoglobin expression at the early phase of ESC differentiation. Moreover, our RNA sequencing analysis revealed that numerous myeloid and immune response related genes, including *Irf8*, exhibited lower expression in the ZBTB46 primed cells. On the other hand, multiple genes were upregulated in the ZBTB46-instructed cells, including *Cyp26b1*. This study has been already published, further details can be found in the following papers.

Boto P, Gerzsenyi TB, Lengyel A, Szunyog B, Szatmari I. *Zbtb46*-dependent altered developmental program in embryonic stem cell-derived blood cell progenitors. *Stem Cells*. 2021; 39(10):1322-1334.

We also generated *Egr2*-inducible mouse ESC lines and myeloid cells. In collaboration with Laszlo Nagy's research group we analyzed the effects of the EGR2 transcription factor in murine ESC-derived myeloid cells. We probed the impact of the IL-4/EGR2 axis on macrophage alternative polarization. We observed that forced expression of *Egr2* led to elevated expression of two alternatively polarized macrophage marker genes (*Retnla* and *Chil3*) confirming the instructive role of EGR2 on the alternative macrophage developmental program. The results of this study have been already published:

Daniel B, Czimmerer Z, Halasz L, **Boto P**, Kolostyak Z, Poliska S, Berger WK, Tzerpos P, Nagy G, Horvath A, Hajas G, Cseh T, Nagy A, Sauer S, Francois-Deleuze J, **Szatmari I**, Bacsı A, Nagy L. The transcription factor EGR2 is the molecular linchpin connecting STAT6 activation to the late, stable epigenomic program of alternative macrophage polarization. *Genes Dev*. 2020; 34(21-22):1474-1492.

2. Modified ES-DC differentiation protocol

Of note, we have learned a lot from the RNA-seq experiments carried out with the ZBTB46 primed cells. This knowledge was later used for the RUNX3 dependent gene expression analyses. For the RNA-seq experiments, it was necessary to extract RNA from OP9 free cell populations, therefore, Flk1+ cells were purified with magnetic separation and this purified cell population was further cultured without OP9 feeders. It is worth mentioning that we had to repeat the cell differentiation experiments several times to obtain good quality of RNA. As a matter of fact, 3 out of 8 experiments were successful regarding the quality of RNA (RNA integrity number was better than 8). The possible reason for the poor quality of RNA was that the viability of the differentiated cells was lower in the absence of OP9 feeders, especially if the cell density was low. This observation inspired us to apply a modified ES-DC differentiation protocol.

In the original proposal we planned to use an OP9 cell based co-culture method to differentiate the transgenic ESCs into myeloid cells. Unfortunately, with this protocol, after 11 days of differentiation, the percentage of the CD11b/CD45 double positive myeloid cells was variable and often rather low (2-15%). In addition, we intended to assess the global gene expression pattern during this early phase of differentiation, therefore, it was crucial to get cells without OP9 feeders. Therefore, we tested an alternative differentiation procedure to generate myeloid blood cells. With this novel method, ESCs were first converted to embryoid bodies (EBs); after this step the EB derived disaggregated cells were cultured on monolayer in high density without OP9 feeders in the presence of GM-CSF. We obtained a much better hematopoietic differentiation: around 30-70% of CD11b/CD45+ blood cells were detected after 12 days. Importantly, these EB derived myeloid cells and ES-DCs exhibited a similar phenotype after *Runx3* induction compared with the OP9 co-cultured cells.

3. Transcriptome analysis of the RUNX3 instructed cells

The primary goal of this project was to identify RUNX3 regulated genes and pathways in ES-DCs and their precursors. In the original proposal we planned to carry out both transcriptomic and ChIP-seq (chromatin precipitation coupled DNA sequencing) experiments simultaneously using the very same samples. However, our pilot experiments revealed that it was very challenging to harvest enough uniformly differentiated cells for ChIP-seq analysis especially during the early phase of the differentiation because of the cellular heterogeneity. Moreover, we also realized that the doxycycline regulated transgenes are poorly expressed in the late phase of the differentiation, see Boto et al Stem Cells.; 39:1322-1334. Similarly, our intracellular RUNX3 staining also revealed that only a subset of the cells was positive in the 11-day differentiated cells (Figure 1).

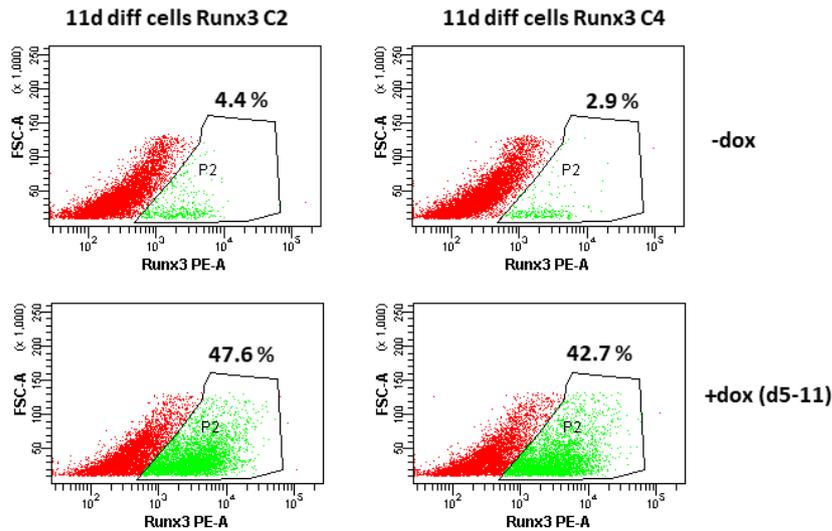


Figure 1. Representative flow plots show quantified protein expression of RUNX3 in 11-day differentiated *Runx3* inducible cell clones (C2 and C4) with intracellular flow cytometry. To induce the *Runx3* transgene, cells were treated with doxycycline (+dox).

In parallel with the native RUNX3 protein investigation, we have also engineered chemically inducible N-terminal tagged RUNX3 mouse ESC lines with site specific recombination. In these modified cell lines, the AviTEV-RUNX3 fusion sequence can be co-induced with a biotin ligase (BirA) with doxycycline treatment. After induction of these transgenes, the BirA enzyme will recognize the AviTEV peptide and ligate biotin to it, consequently the final product is N-terminally biotinylated (Bio-RUNX3). Numerous *Bio-Runx3* expressing mouse ESC clones were generated and characterized. The doxycycline dependent transgene expression was demonstrated with quantitative PCR, moreover, the biotinylated protein was identified with Western blot using SA-HRP (streptavidin-horseradish peroxidase) in partially differentiated cells (Figure 2).

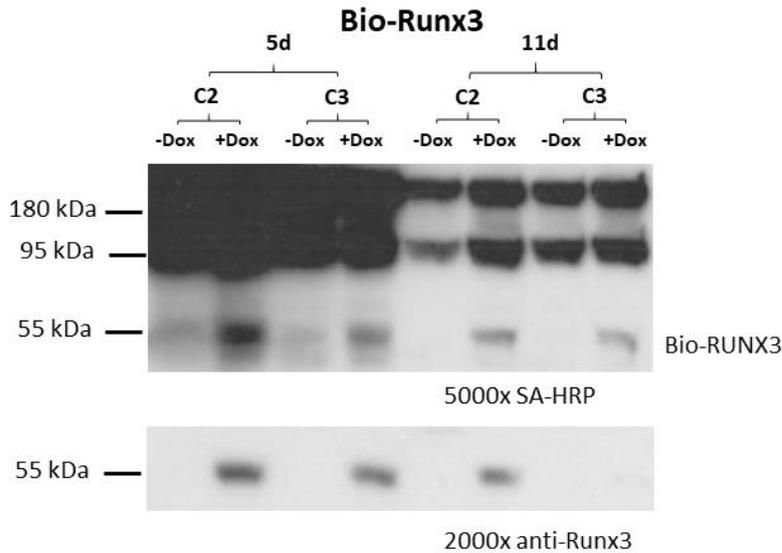


Figure 2. Bio-RUNX3 protein detection with Western blot testing 5- or 11-day differentiated, Bio-RUNX3-inducible cell clones (C1 and C3). SA-HRP (streptavidin-horseradish peroxidase) was used to detect the biotinylated protein. Cells were treated with doxycycline (+Dox) to induce the *Bio-Runx3* transgene. The high molecular weight bands represent the endogenous biotin containing enzymes.

We also analyzed the differentiation capacity of the Bio-RUNX3 expressing cells comparing with the native RUNX3 inducible cell lines. We found that the Bio-RUNX3 elicited a similar phenotype compared with the native RUNX3 protein during the ES-DC differentiation, although the maturation capacity of the Bio-RUNX3 primed ES-DCs were a little bit inferior. These findings suggest that the N-terminal biotinylated peptide sequence only slightly modified the effects of the RUNX3 on ES-DC differentiation. Therefore, the *Bio-Runx3* expressing cells could be applied for genomic mapping experiments using the high affinity biotin-streptavidin interaction. We also analyzed the protein expression of the Bio-RUNX3 using intracellular staining with flow cytometry. Similarly to the RUNX3 protein expression, relatively low percent of Bio-RUNX3 positive fraction was detected in the 11-day differentiated cells (Figure 3).

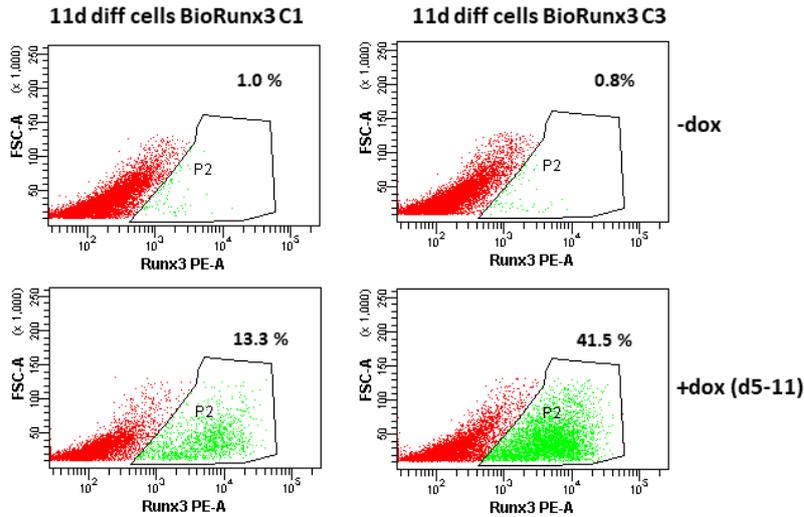


Figure 3. Representative flow plots show quantified protein expression of BIO-RUNX3 in 11-day differentiated *Bio-Runx3* inducible cell clones (C1 and C3) with intracellular flow cytometry. To induce the *Bio-Runx3* transgene, cells were treated with doxycycline (+dox).

This result suggests that the ChIP-seq experiments should be carried out in early progenitors instead of 11-day differentiated cells. However, at the early phase of differentiation, the cells are very heterogeneous, thus it would be very challenging to select a cell population with uniform gene expression profile. Therefore, we focused on other advanced genomics technologies (including single cell transcriptomic analysis) to characterize the impact of the RUNX3 transcription factors, in addition, we also studied the effects of RUNX3 in undifferentiated cells.

First, we have determined the global transcript profiles of undifferentiated ESCs upon the enforced expression of *Runx3* or *Zbtb46* (four parallel sample sets were tested, together 16 RNA samples). These 16 samples were run in parallel with the Illumina NextSeq 500 and the MGI DNBSEQ G400 next generation sequencing (NGS) platforms. Of note, both DNA sequencing platforms are available in our Genomics Core Facility. First, we determined the global gene expression signature with principle component analysis (PCA) including all samples and RNA transcripts. This analysis revealed that the RUNX3 and the ZBTB46 induced samples formed distinct gene expression signatures. In addition, it was observed that both NGS platforms (Illumina versus MGI) exhibited a similar global gene expression pattern (Figure 4).

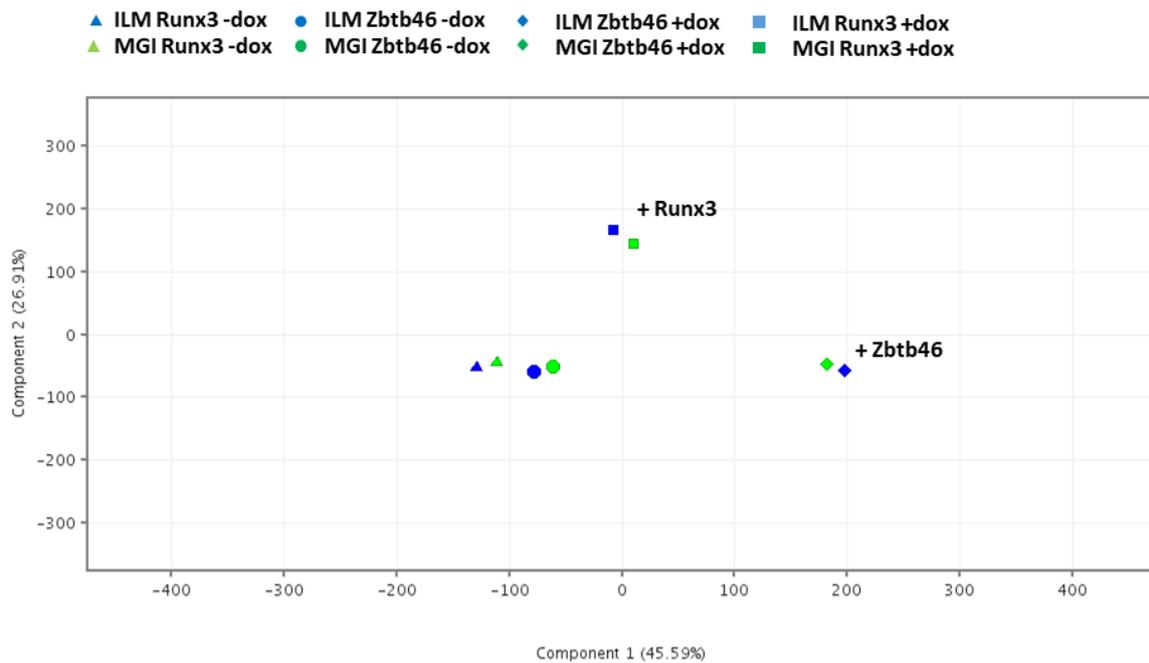


Figure 4. Principle component analysis was employed to visualize the alterations of sample type and transgene dependent gene expression patterns of the ESC samples. To induce the *Runx3* or *Zbtb46* transgenes, cells were treated with doxycycline (+dox). The RNA samples were run in parallel with the Illumina NextSeq 500 (ILM) and the DNBSEQ G400 (MGI) next generation sequencing platforms.

Next, those genes were selected which were up- or downregulated upon the RUNX3 or ZBTB46 induction. We used strict selection criteria: twofold change or greater, P-value cutoff was 0.05 using T test combined with multiple test correction. Unexpectedly, huge number of genes exhibited altered gene expression after three days of *Runx3* induction: 602 (Illumina) or 590 (MGI) genes were upregulated; in contrast, 349 (Illumina), or 328 (MGI) genes showed lower expression. If we combined the up- and downregulated genes 951 RUNX3 regulated genes were detected with the Illumina NGS platform and 928 genes with the MGI technology. Obviously most of the regulated genes were detected with both technologies: together 1098 RUNX3 regulated genes were identified (Figure 5). Similar analysis was performed testing the ZBTB46 inducible cells: 1392 ZBTB46 modulated genes were detected with the Illumina and/or the MGI platforms.

1098 RUNX3 regulated genes were detected with Illumina and/or MGI platforms in undifferentiated ESCs.

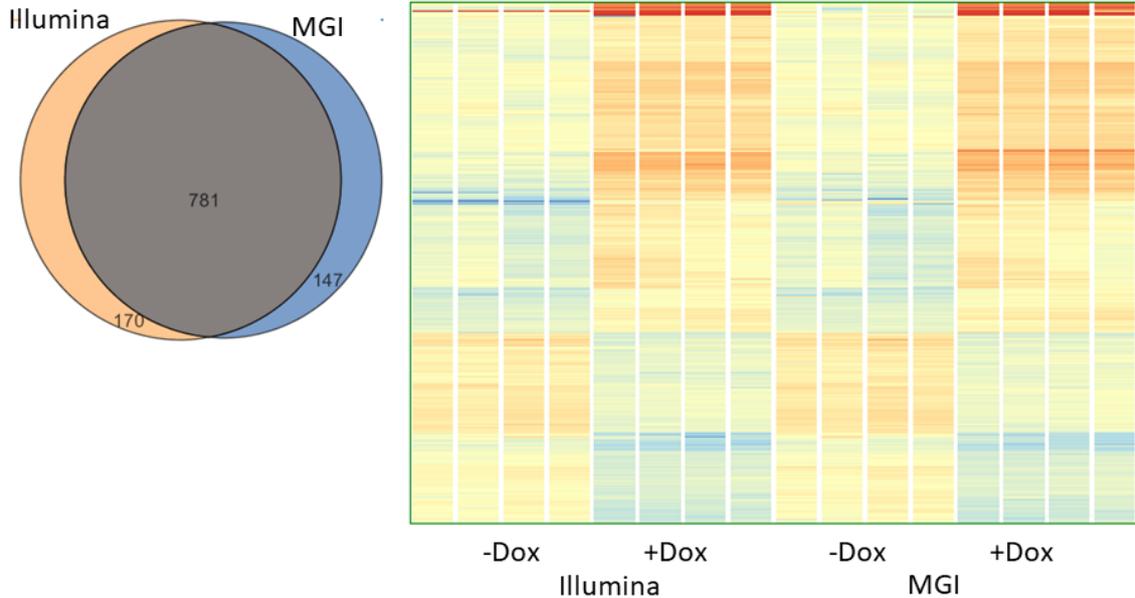


Figure 5. Hierarchical clustering of 1098 transcripts with altered expression upon RUNX3 induction detected with the Illumina and/or the MGI technologies. To induce the *Runx3* transgene, ESCs were treated with doxycycline (+dox).

In this project we focused on the putative RUNX3 target genes, therefore we investigated the RUNX3 upregulated genes. We observed that several members of the granzyme B gene family was upregulated, in addition elevated expression of the *Gdf6* and *Ccr7* was detected (Figure 6). We have validated some of these gene expression changes with quantitative PCR in undifferentiated ESCs (Figure 6).

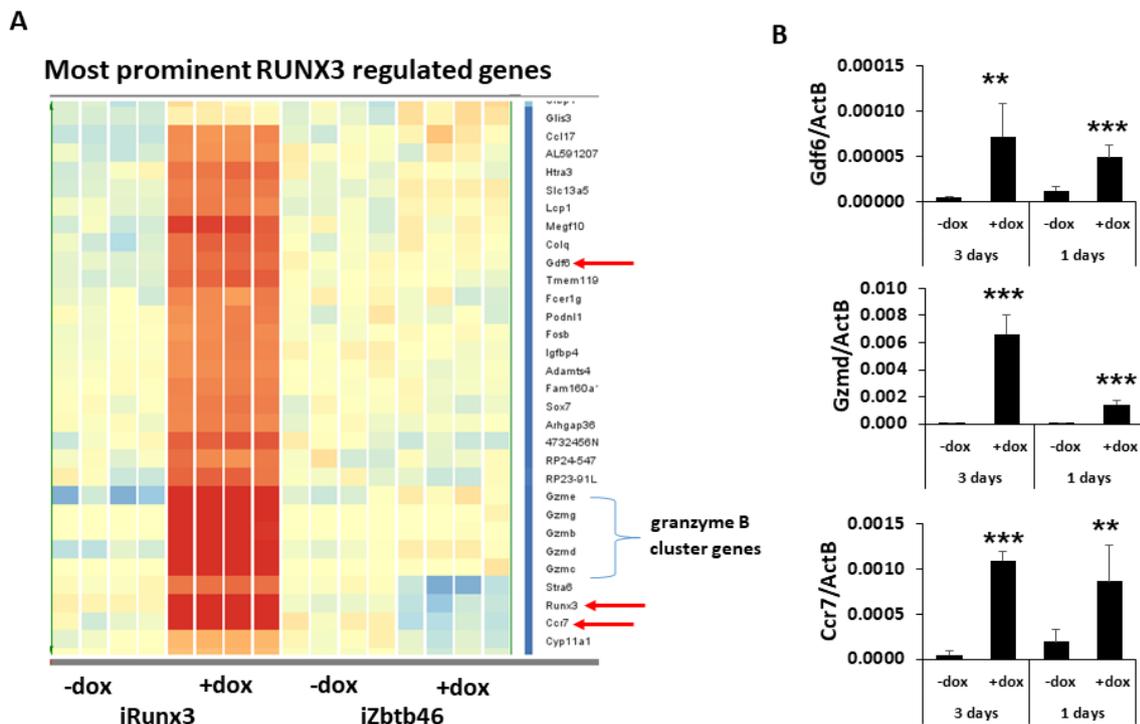


Figure 6. (A) Hierarchical clustering of the most prominent RUNX3 induced genes detected with the Illumina technology. To induce the *Runx3* or the *Zbtb46* transgenes, ESCs were treated with doxycycline (+dox). (B) Quantitative RT-PCR assay to validate the mRNA level of *Gdf6*, *Gzmd* and *Ccr7* testing undifferentiated ESCs.

Based on these data, next we analyzed the transcriptomic effect of the ectopically expressed RUNX3 in differentiated ES-DCs and they precursors using the MGI DNBSEQ G400 NGS platform for RNA-seq. Here we used EB derived progenitors for differentiation without OP9 feeders. Four parallel sample sets were tested: 9, 12 and 19-day differentiated cells plus 19-day differentiated cells with LPS activation; together 32 RNA samples were used. Comparing the non-differentiated ESCs fewer genes were regulated by RUNX3, especially at day 12: only 16 upregulated genes were detected (Figure 7).

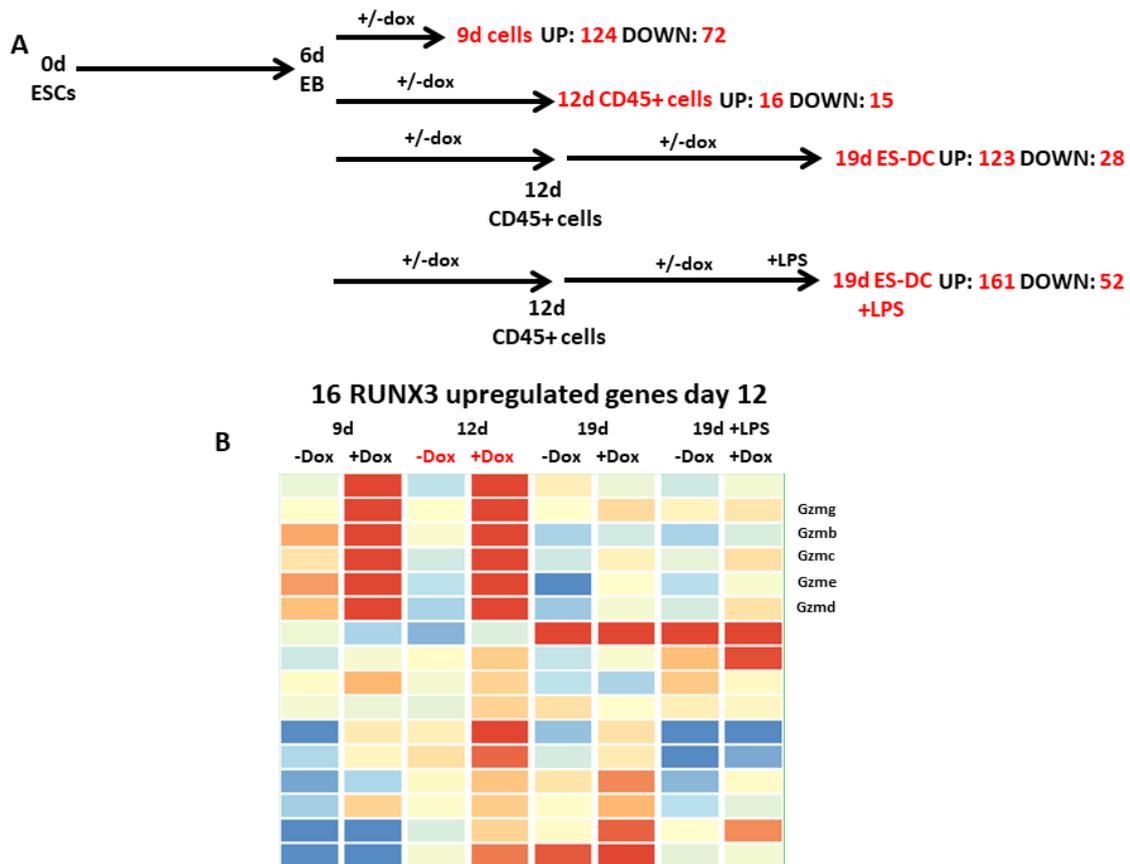


Figure 7. (A) Experimental scheme of the RUNX3-inducible ES-DC differentiation via EB formation. The indicated numbers represent the up- and downregulated genes upon doxycycline treatment. (B) Hierarchical clustering of 16 transcripts with altered expression upon RUNX3 induction at day 12.

This analysis also revealed that the 9-day and 12-day differentiated cells exhibited a partially overlapped upregulated gene set. For example, members of the granzyme B gene family (*Gzmb*, *c*, *d*, *e* and *g*) were similarly upregulated at day 9 and 12 by RUNX3 (Figure 7). In contrast, at late stage (19-day differentiated cells), a completely different set of genes were modulated by doxycycline treatment. We also realized that in later phase of the differentiation the RUNX3 level was very low due to transgene silencing (data not shown). Probably at day 19, the altered gene expression profile is the consequence of the distinct cellular composition of the RUNX3 instructed ES-DCs. We also observed that the differentiated cells represent a heterogeneous cell population, hence the putative RUNX3 target genes might differently expressed in the various subpopulations. The bulk RNA-seq technology, which was used for this study, cannot capture

the cell specific gene expression profiles. Therefore, we also aimed to employ single cell transcriptomic analysis. We have recently installed the 10x Genomics Chromium Controller system for single cell RNA sequencing. Using this technology, 8- and 12-day differentiated cells (with or without RUNX3 induction) were analyzed. Consistent with the flow cytometric RUNX3 protein data, a bimodal Runx3 expression profile was detected upon doxycycline treatment in the 8-day differentiated cells. It means that a subset of the cells was RUNX3 negative. More importantly, a distinct cell population was detected in the doxycycline treated cells and the RUNX3 expression was high in this population. In addition, the granzyme B genes were selectively expressed in this cell population. Moreover, several additional genes (for example, *Npnt* and *Rdh10*) were highly expressed in this RUNX3 dependent cell fraction (Figure 8).

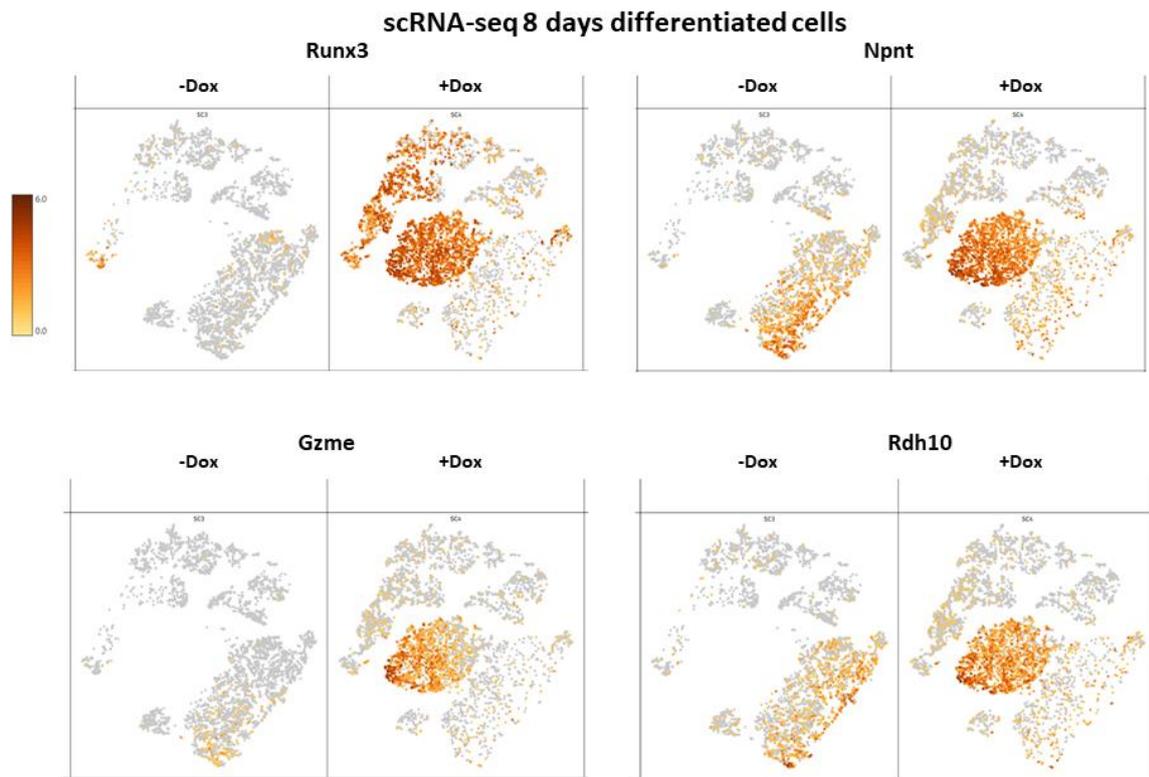


Figure 8. Single cell gene expression profiles of *Runx3*, *Gzme*, *Npnt* and *Rdh10* using t-SNE projection testing the 8-day differentiated cells. To induce the *Runx3* transgene, cells were treated with doxycycline (+Dox). The color scale represents the number of the indicated transcripts per cell.

Strikingly, the 12-day differentiated, doxycycline treated cells expressed less RUNX3, only a minor subpopulation was RUNX3 positive. These RUNX3+ cells also expressed the granzyme B genes (Figure 9 and data not shown). This result suggests that the RUNX3 level strongly correlates with the expression of the granzyme b genes. Granzymes are important for the killing activity of the cytotoxic T cells and NK cells. It is worth mentioning that RUNX3 has a prominent role in the development and activity of the cytotoxic T and NK cells. However, granzymes can also participate to elicit inflammatory response in myeloid cells, therefore, these enzymes can contribute for the RUNX3 dependent ES-DC development and activation. Our flow cytometric data indicated that at this stage more than 50% of the cells are CD45+, this was confirmed with single cell analysis (Figure 9). Interestingly, among the CD45+ cells unique subpopulation can be detected in the RUNX3 primed cells. We observed that this subpopulation

is selectively express *Ly6c2* marker suggesting that RUNX3 can facilitate the monocytic cell differentiation. Together our single cell analysis revealed that some genes are probably directly regulated by RUNX3, however, novel cell subpopulations can be also emerged during the ES-DC differentiation which are indirectly governed by RUNX3. Some of these changes can contribute for the enhanced immunogenicity of the RUNX3-programmed ES-DCs.

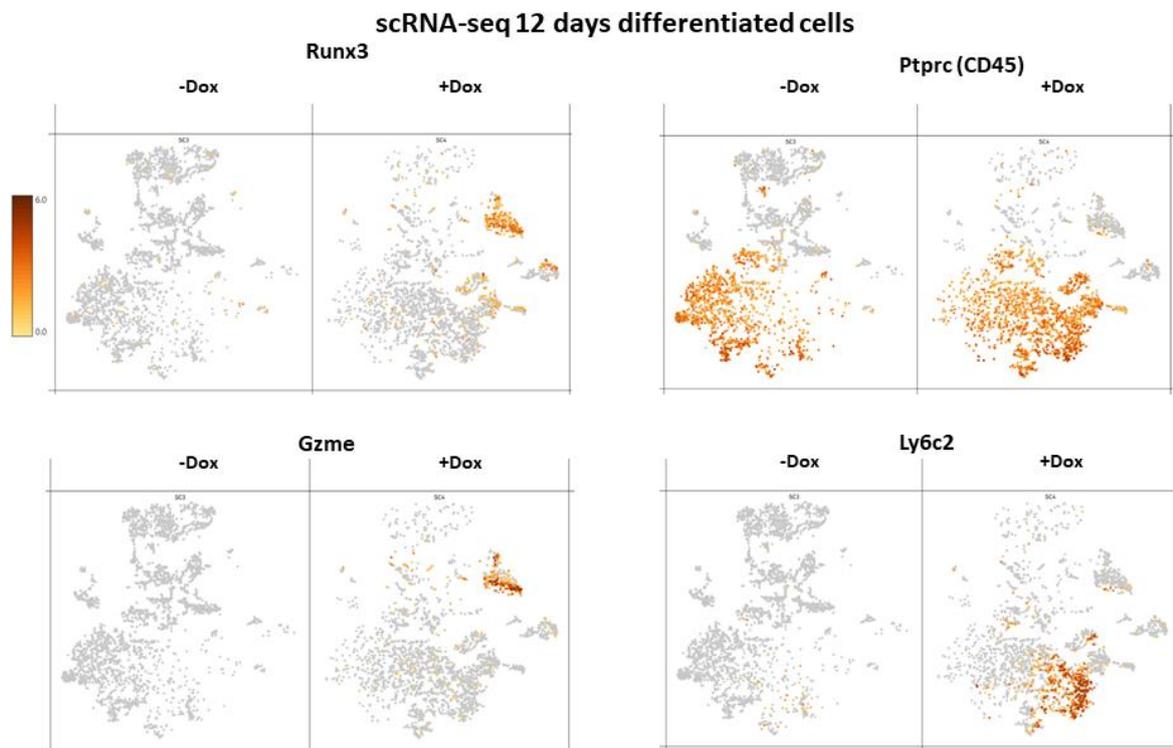


Figure 9. Single cell gene expression profiles of *Runx3*, *Gzme*, *Ptpnc* and *Ly6c2* using t-SNE projection testing the 12-day differentiated cells. To induce the *Runx3* transgene, cells were treated with doxycycline (+Dox). The color scale represents the number of the indicated transcripts per cell.

4. ES-DC differentiation from *Irf8* deficient cells with RUNX3 induction.

In this project we planned to examine the immune phenotype of the RUNX3-inducible ESC derived progenitors carrying inactivated myeloid master genes (*Irf8* or *Spi-1*). In collaboration with János Kádás and Attila Brunyánszki (UD-Genomed Ltd.) we generated *Irf8* KO cell lines with gene editing using the CRISPR Cas9 system. We also tried to generate *Spi-1* (PU.1) KO cells but in this case only heterozygous clones were obtained. Two *Irf8*-null ESC clones have been characterized: we confirmed in one of the tested clones (1/F10) with Western blot that the IRF8 protein was missing during the myeloid differentiation. Unexpectedly, despite of the *Irf8* deficiency we managed to generate myeloid cells from these modified ESCs. Next we reintroduce the RUNX3 in the *Irf8* KO background and assessed the RUNX3 dependent differentiation capacity of these cells. Without IRF8, forced expression of RUNX3 provoked a distinct phenotype: high percent of CD34+ cells were obtained upon the induction of RUNX3 (Figure 10). This result suggests that in the absence of *Irf8* RUNX3 skews the cell differentiation into the CD34+ endothelial or blood cell progenitors. We will further characterize this special CD34+ cell population.

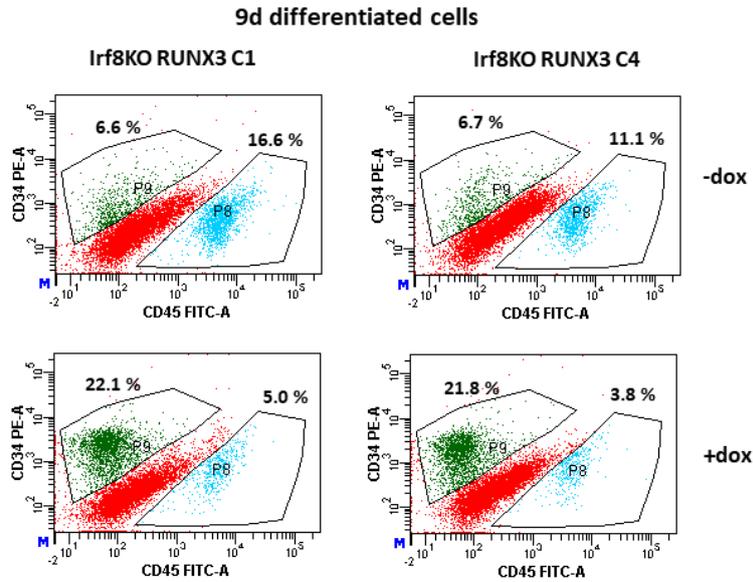


Figure 10. Representative flow plots and statistics show the percentage of the CD45 and CD34 positive populations of 9-day differentiated, *Runx3*-inducible, *Irf8* deficient cell clones (C1 and C4). To induce the *Runx3* transgene cells were treated with doxycycline (+dox).

Publication strategy

Despite of the huge amount of genomics results, we have moved less rapidly in the final period of the project, especially we delayed to publish the RUNX3 related transcriptomic results. It is worth mentioning that we published a review paper about the RUNX3 transcription factor (Boto et al. Crit Rev Immunol. 2018, 38:63-78). In addition, we have already prepared a manuscript in which we monitored the RUNX3 and ZBTB46 primed ESC gene expression profiles. In this report we have also compared the Illumina NextSeq 500 and the MGI DNBSEQ G400 (NGS) platforms. We also prepared a manuscript about the RUNX3 genomic effects during differentiation, in this report we intend to publish both the bulk and the single cell RNA sequencing results. Finally, we will prepare a separate manuscript about the effects of RUNX3 in the absence of *Irf8*.