SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Characterization of labelled regulatory elements in embryonic stem cells and macrophages using quantitative and qualitative methods

by **Attila Horváth**

Supervisor: Prof. Dr. László Nagy, MD, PhD, DSc, MHAS

Co-Supervisor: Dr. Benedek Nagy, PhD

UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF MOLECULAR CELL AND IMMUNE BIOLOGY

DEBRECEN, 2019

Characterization of labelled regulatory elements in embryonic stem cells and macrophages using quantitative and qualitative methods

by **Attila Horváth**, MSc

Supervisor: Prof. Dr. László Nagy, MD, PhD, DSc, MHAS

Co-Supervisor: Dr. Benedek Nagy, PhD

Doctoral School of Molecular Cell and Immune Biology, University of Debrecen

Head of the **Examination Committee**:	Prof. Dr. László Fésüs, MD, PhD, DSc, MHAS
Members of the **Examination Committee**:	Prof. Dr. Imre Boros, DSc
	Dr. István Nagy, PhD

The examination took place at the Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen, Debrecen;
at 12:00; 23rd of May, 2017

Head of the **Defense Committee**:	Prof. Dr. László Fésüs, MD, PhD, DSc, MHAS
Reviewers:	Prof. Dariusz Plewczyński, DSc
	Dr. András Penyige, PhD
Members of the Defense Committee:	Prof. Dr. Imre Boros, DSc
	Dr. Magdolna Szántó, PhD

The PhD Defense takes place at the Lecture Hall of Bldg. A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen, Debrecen;
at 13:00; 16th of May, 2019

1. INTRODUCTION

Cells are the basic structural and functional units of life. The human body consists of trillions of cells, which can be categorized into around 200 different cell types based on their morphological and/or functional characteristics. With a few notable exceptions, the genomic DNA sequence is identical in the nuclei of all diploid cells within an organism. Although genomic DNA contains all the necessary information for creating any of the organisms' cell types, the fate of an individual cell is defined by the local microenvironment during differentiation. The resulting unique gene expression pattern will define the mature cell's function. The genomic determinants of this unique gene expression profile and its functional role in the cell's response to external stimuli are two of the most exciting topics of the post-genomic era. The cell-type-specific gene expression profile is the result of a series of transcriptional regulatory processes, including transcriptional initiation, elongation, transcript processing and degradation.

Transcription regulation in Eukaryotes

Transcription is the process by which the genetic information stored in the DNA is copied into RNAs. Unlike prokaryotes, eukaryotic cells encode for multiple RNA polymerases, including Pol I, Pol II and Pol III, which synthesize different types of RNAs. Pol I transcribes ribosomal RNAs (rRNAs), Pol III carry out the synthesis of transfer RNAs (tRNAs), small RNAs, the 5S rRNA and most long non-coding RNAs (lncRNAs), while Pol II transcribes mRNAs, small nuclear RNAs (snRNAs) and micro RNA (miRNA) precursors. The correct positioning of Pol II and other proteins, forming the so-called initiation complex, is essential to initiate transcription at the start of the gene body.

A promoter region is defined as the minimal DNA sequence which is sufficient for transcription initiation by an RNA polymerase. One of the best example for promoter-like elements is the TATA box, located ~25 base pairs upstream of coding regions for certain protein- or non-protein-coding (e.g. lncRNA) genes. Other promoter elements have also been discovered: the initiator (Inr), the center of which is located at +1 (one nucleotide downstream of the transcription start site), and the downstream promoter element (DPE), which is centered at +30. These DNA elements provide binding sites for the transcription initiation complex and will direct RNA polymerase where to begin transcription.

Tissue-specific gene expression variability in multicellular organisms rely on distal regulatory regions. These genomics regions, namely enhancers play a major role in the transcriptional initiation process, modulating the timing and rate of transcription of the associated genes.

The concept of enhancer

Enhancers are short DNA sequences that can induce the expression level of a gene, upon binding one or more regulatory protein(s) called transcription factors (TFs) and looping to the promoter of the regulated gene. An enhancer can regulate multiple genes, and one gene can have multiple enhancers. It has also been described that the promoters of certain genes can act as enhancers of distal genes. In contrast to promoters, enhancers can exert their gene regulatory function in a distance- and direction-independent manner; they do not need to be located near to the transcription start site of the gene.

Typically, tens of thousands of enhancer regions are engaged in transcriptional regulation in a given cell type, at a given time point. Enhancers provide gene regulatory flexibility by being able to anchor different sets of transcriptional regulators, which may be present under only certain circumstances, such as after the exposure to a certain external stimulus. An enhancer can exert its gene regulatory function by making a spatial contact with the initiation complex at the promoter region of the regulated gene. In theory, almost any distance in the nucleus can be bypassed by this so-called promoter-enhancer looping. However, the spatial organization of the chromatin, such as the presence of topologically associated domains (TADs), puts a constraint on the possible promoter-enhancer interactions. Moreover, it has been shown that enhancers can regulate different stages of the transcription cycle, including RNAPII recruitment, release of promoter-proximal pause and transcription elongation. Therefore, the detection and characterization of enhancers is critical for understanding how these elements contribute to cell-type specific functions.

Identification of enhancer regions

Identification of the enhancers of a certain gene is challenging for multiple reasons. First, the regulatory elements can be located at great distances (up to 1 Mb) both upstream or downstream the transcription start site (TSS) of the regulated gene. Moreover, studies have shown that not only the non-coding genome, but intronic and exonic regions can also contain enhancer elements. Surprisingly, gene promoters can also act as distal regulatory elements.

Genes often have more than one enhancers and one enhancer can regulate multiple genes. Finally, unlike promoters, there is no general sequence code for enhancers that could be predicted by in silico methods. A widely used strategy for enhancer prediction is to map histone modifications and other epigenomic marks using high-throughput sequencing-based techniques such as chromatin immunoprecipitation sequencing (ChIP-seq).

Histone modifications

In the eukaryotic nucleus, DNA is packed and ordered by histones in structures called nucleosomes. There are five types of histones: H1, which is involved in higher-order structures of chromatin (known as linker histones), H2A, H2B, H3 and H4 (core histones constituting the so-called histone octamer). Nucleosomes comprise core histones and DNA wrapped around them. In human cells, there is about ~2 meters of DNA, but it is about 90 millimeters of chromatin in a condensed form.

Histones can carry various post-translational modifications referred to as histone marks, which may regulate gene expression by making DNA less or more accessible to transcription. H3 is one of the most extensively modified protein among the core histones. The modifications of H3 can be used to distinguish between hetero or euchromatin states or identify different types of regulatory elements. Trimethylated H3K4 (H3K4me3, three methyl groups to lysine 4 of histone H3) is enriched around transcription start sites (TSSs) of active genes. In contrast, gene bodies of actively transcribed genes are associated with trimethylated H3K36 (H3K36me3, three methyl groups to lysine 36 of histone H3). Monomethylated H3K4 (H3K4me1, one methyl group to lysine 4 of histone H3) has been shown to co-localize on regulatory regions with pioneer TFs, characterizing general enhancers (irrespectively to transcriptional activity), and it is also enriched at promoter regions, although to a lesser extent. Finally, the histone H3K27ac (acetylation at the lysine 27 of histone H3) can be found both at promoters and enhancers, and marks active transcription.

Classification of enhancer states based on epigenetic signatures

In macrophage biology, PU.1 and H3K4me1/2 double-positive regions were considered as enhancer-like regulatory regions and active histone marks such as H3K27ac were used to distinguish between poised and active enhancers. This classification revealed the existence of latent or de novo enhancers which are not marked either by H3K4me1/2 or PU.1 binding prior to stimulus but they get activated by signal-dependent transcription factors (SDTFs). Although

these studies did not characterize the contribution of chromatin openness which might provide a better mechanistic insight into the different stages of enhancer formation.

2. AIMS OF THIS STUDY

Aim 1. Evaluation of the contribution of key macrophage TFs to chromatin openness and enhancer formation in steady state and polarized mouse macrophages

- Map the genomic binding sites of key TFs including PU.1, IRF8, JUNB, RUNX1 and CEBPA in order to examine interrelation the hierarchy among them
- Apply machine-learning approaches such as Random Forest and Support Vector Regression to build predictive models evaluating the contribution of the binding pattern of the studied TFs to chromatin openness and enhancer activation
- Classify the cistromes of the studied TFs based on chromatin openness and perform loss/gain of function experiments to investigate the deterministic role of low accessible regions in regulating gene expression in the steady state and/or in response to external stimuli
- Reveal whether there exist distinct TF modules binding low accessible genomic regions that regulate specific gene expression programs triggered by various macrophage polarizing stimuli
- Build a formal model using Automata Theory that describe the possible states of enhancer formation and transitions among them

Aim 2. Examination of the role of OCT4 in the early steps of RA-induced neurogenesis

- Map the OCT4 cistrome and investigate its relation to chromatin openness in both naïve and ground state ESCs
- Characterize the interaction between the low accessible OCT4 binding sites and retinoic acid signaling pathway playing critical role in the early steps of neuronal differentiation
- Perform loss of function experiments for OCT4 to delineate its deterministic role in forming differentiation-related enhancer and regulate retinoic acid target genes

- Construct a composite network using network motifs to describe the dual role of OCT4 in pluripotency state and in the context of early steps of RA-induced neurogenesis.

3. MATERIALS AND METHODS

Differentiation of bone marrow derived macrophages

Bone-marrow was flushed from the femur of wild-type C57BI6/J male animals. Cells were purified through a Ficoll-Paque gradient (Amersham Biosciences, Arlington Heights, IL) and cultured in DMEM containing 20% endotoxin-reduced fetal bovine serum and 30% L929 conditioned medium for 5 days. Isolated bone marrow-derived cells were differentiated for 6 days in the presence of L929 supernatant. At the 6^{th} day of differentiation, cells were exposed to IL-4 (20ng/ml) and LPS (100 ng/ml) for the indicated period of time.

Embryonic stem cell culture

Mouse E14 embryonic stem cell cultures were cultured on primary mouse embryonic fibroblast (PMEF) feeder cells (5% CO2 at 37 °C). ESC medium was prepared by supplementing DMEM Glutamax (Gibco) with 15% FBS (Hyclone), 1,000 U of LIF, penicillin/streptomycin, non-essential amino acids, and 2-mercaptoethanol. 2i ESCs were adapted for a minimum of five passages to grow in serum-free N2B27-based medium supplemented with LIF, PD0326901 (1 mM), and CHIR99021 (3 mM).

Ligands and Treatment

ESCs were treated with vehicle (DMSO) or with all-trans RA (Sigma, 1 mM stock in DMSO, 1/1,000 dilution).

siRNA knockdown

OCT4 and control siRNAs were obtained from Thermo Fisher. Mouse E14 cells were plated on gelatinized plates 12 hr before transfection. siRNA transfection was carried out with Lipofectamine 3000 (Invitrogen). OCT4 stealth RNAi oligo sequences: siOCT4_A_Fw: 5`-AUG CUA GUU CGC UUU CUC UUC CGG G-3`, 5`-CCCGGAAGAGAAAGCGAACUAGCAU-3`, siOCT4_A_Rev: 5`-CCC GGA AGA GAA AGC GAA CUA GCA U-3`, siOCT4_B_Fw: 5`-ACC UUC UCC AAC UUC ACG GCA UUG G-3`, siOCT4_B_Rev: 5`- CCAAUGCCGUGAAGUUGGAGAAGGU-3`. Transfected cells were cultured in embryonic stem (ES) medium for 24 hr prior to experiments. Cells were

ligand treated for 3-24 hrs before harvesting for eRNA and mRNA experiments, meaning that cells were harvested 24+3 or 24+24 hrs following the transfections. Similarly, ChIP experiments were performed 24 hrs following the transfection or the latest 48 hrs if ligand treatment was applied.

Microarray analysis

Control, Nanog RNAi and OCT4 RNAi .CEL files of GSE4189 microarray series were downloaded from NCBI/GEO database and imported into GeneSpring (version 13.0). Gene Level Experiment was carried out using the following parameters: Threshold raw signals to 1.0, Normalization algorithm: quantile, Baseline to median of control samples (GSM94856-GSM94860), Average over replicates in conditions. Entity lists containing pre-selected RA target genes or components of the retinoic acid signaling pathway were used for heatmap analysis.

RT-qPCR

RNA was isolated with Trizol reagent (Ambion). RNA was reverse transcribed with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to manufacturer's protocol. Transcript quantification was performed by qPCR reactions using SYBR green master mix (BioRad). Transcript levels were normalized to *Ppia*.

ChIP-seq

ChIP was performed as previously described (Barish et al., 2010; Daniel et al., 2014b; Siersbaek et al., 2012), with minor modifications. The following antibodies were used: OCT4 (sc-8628), PU.1 (sc-352), IRF8 (sc-6058x), JUNB (sc-46x), CEBPA (sc-61x), RUNX1 (sc-8563x), STAT6 (sc-981x), p65 (sc-372), H4ac (millipore 06-866), H3K4me1 (ab8895), H3K27ac (ab4729), IgG (Millipore,12-370), OCT4 (Santa Cruz, sc-8628), RXR (Santa Cruz, sc-774), RAR (Santa Cruz, sc-773), P300 (Santa Cruz, sc-585), H3K27me3 (Millipore, 07-449) and RNAPII-pS2 (ab5095). Libraries were prepared by Ovation Ultralow Library Systems (Nugen) from two biological replicates according to the manufacturer's instructions.

Western blot analysis

20 µg protein whole cell or nuclear extracts were separated by electrophoresis in 10 or 12.5% polyacrylamide gels and then transferred to Immobilon-P Transfer Membrane (Millipore Crp.,Billerica, Massachusetts). Membranes were probed with anti-Oct3/4 (Santa

Cruz;sc-5278), anti-RAR (Santa Cruz, sc-773), or anti-GAPDH (Santa Cruz, sc-32233) antibodies according to the manufacturer's recommendations.

GRO-seq

Global Run-On sequencing and library preparation was performed as described earlier (Core, Waterfall, & Lis, 2008). Libraries were made from two biological replicates of BMDMs. Libraries were sequenced with Illumina HiScanSQ sequencer.

ChIP-seq, GRO-seq and ATAC-seq analyses

Primary analysis of the ChIP-seq, GRO-seq, and ATAC-seq raw reads was carried out using a ChIP-seq analyze command line pipeline (Barta, 2011). Briefly, Burrows-Wheeler Alignment Tool (H. Li & Durbin, 2009) was used to align the reads to mm10 genome assembly (GRCm38) with default parameters. MACS2 2.0.10 (Zhang et al., 2008) was used for predicting TF peaks (q-value <= 0.01) and findPeaks.pl (with '-size 1000' and '-minDist 2500' options) for histone regions with the option '-style histone'. Artifacts were removed using the ENCODE blacklist (Consortium, 2012). 'Intergenic' and 'Intron' regions were considered as distal elements from HOMER (v4.2) annotation. Reads per kilobase per million mapped reads (RPKM) values of the predicted peaks was calculated using BedTools coverageBed and bash scripts. DiffBind v2.8.0 was used to infer differential binding sites from duplicates of STAT6 and p65 from CTR, 1h IL-4 and 1h LPS treated cells (p-value < 0.05), respectively, and from RNAPII-pS2 ChIP-seq time course experiments (p-value < 0.05 & FC>2) measured on distal regions (normalized DiffBind occupancy > 30) and on gene bodies (normalized DiffBind occupancy > 50), using untreated samples as controls. K-means clustering of RNAPII-pS2 regions both on distal elements and gene bodies (mm10 RefSeq) was performed using kmeans() function from the R package stats. GO analyses were performed using the clusterProfiler R package. Intersections, subtractions, and merging of the predicted peaks were done with BedTools (v2.23.0). Regions were considered overlapping if there was at least one common nucleotide. Consensus sets were defined by merging overlapping regions (in at least 2 samples). Proportional Venn diagrams were generated with VennMaster. Genome coverage files (BedGraphs) for visualization purposes were generated by makeUCSCfile.pl, and then converted into tdf files using igvtools (IGV2.3, Broad Institute) with the 'toTDF' option. Genomic distribution was analyzed using HOMER categories provided by annotatePeaks.pl (UTR regions were merged). De novo motif discovery was performed in the 150 bp vicinity

of the peak summits using findMotifsGenome.pl with options '-length –len '12,14,16,18,20,22' and '-size 200' on the repeat-masked mouse genome (mm10r) from HOMER. Integrative Genomics Viewer (IGV2.3, Broad Institute) was used for data browsing and creating representative snapshots. Normalized tag counts for Meta histograms and read distribution heatmaps (RD plots) were generated by annotatePeaks.pl with '-ghist' and '-hist 25' options from HOMER on one representative example of duplicates and then visualized by R (ggplot2) or Java TreeView. Motif matrices were remapped using annotatePeaks,pl with the '-mscore' option. Summits used for centering RD plots and motif remapping were identical to the summit of the peak with the highest MACS score from those used for deriving the consensus region.

Machine learning

Machine learning analyses were performed in R using the packages randomForest, e1071 and custom scripts. For training sets, 1,000 sites were randomly chosen from both labelled and 'HighAcc' categories for Random Forest and Support Vector Machine models (repeat-masked mouse genome mm10r). To avoid the well-known issue of 'overfitting' in data mining, all models were built using a k-fold (k=10) cross validation. In total, 31 Random Forest models were generated from all possible TF combination (one TF only, n=5; two TFs, n=10; three TFs, n=10; four TFs, n=5; all five TFs - 'full model', n=1). For validation sets, another 1,000 sites were randomly chosen from both labelled and 'HighAcc' categories that were not used for learning processes. Contribution scores (MeanDecreaseGini) were calculated using randomForest function with the 'importance = T' option. Sensitivity, Specificity and ROC values were calculated with the caret and ROCR packages. Boxplot of the 31 models and the ROC plot were generated using ggplot2 R package. For SVM model, Pearson correlation coefficient was calculated on an independent validation set using stats packages in R.

4. RESULTS

Random Forest classification hints the existence of low accessible PU.1 binding sites in macrophages

We systematically assessed all combinations of the studied TFs (PU.1, IRF8, JUNB, CEBPA and RUNX1) by using Random Forest method and measured their prediction accuracy. Interestingly, the model where only PU.1 binding pattern was used showed weak

prediction accuracy of openness (64.4%), only marginally higher than expected by chance. All the other models which accommodated only one of the other key TFs resulted in similarly weak predictive powers, however, slightly better than PU.1. While the model in which all the TFs were used as input ('full model'), resulted in a high accuracy prediction on an independent validation set (Accuracy = 0.82, Sensitivity = 0.88, Specificity = 0.77, AUC = 0.90. The Random Forest classification also revealed that all the other TFs had higher contribution than PU.1 in defining chromatin openness indicating that PU.1 binds low accessible genomics regions as well.

Based on these analyses we classified the PU.1-bound sites into three classes: (1) binding sites that were not associated with predicted NFRs, termed 'PU.1-labelled' regulatory elements (PU.1-LREs, 46,280 sites); (2) regions where PU.1 binding overlapped with open chromatin regions, termed 'PU.1 pos. HighAcc' sites (18,448 sites); and (3) PU.1-negative highly accessible regions, termed 'PU.1 neg. HighAcc' sites (18,967 sites). Strikingly, only one-third of the PU.1-bound genomic regions overlapped with NFRs.

Taken together, these results show that two-thirds of PU.1-bound regions are associated with low accessible chromatin and the lack of transcriptional activity; we termed these sites LREs. Moreover, both the characterization of 'PU.1 neg. LowAcc' sites and the ChromHMM-based analysis suggest that IRF8, RUNX1 and CEBPA also have labelled fractions and there might be unique co-LRE TF modules.

Key transcriptional regulators of macrophage form labelled regulatory elements

Our study revealed that LREs are widespread in cistromes of the studied TFs, and PU.1-labelled sites are the most prevalent LREs because PU.1 has the largest cistrome and its ability to bind alone. Thus, the labelled portion of IRF8, CEBPA and RUNX1 often overlap with PU.1 and they form co-LREs. Finally, among the possible combinations of co-LREs, PU.1+IRF8 co-LREs are the most abundant, having around three times as many binding sites than the second most abundant combination, which also contain the TFs PU.1 and IRF8.

The role of PU.1 and IRF8 co-LREs in cellular response to IL-4

Our findings highlight that PU.1+IRF8 co-LREs can be transformed into active enhancers activated by IL-4 stimulus. Using gain and loss of function experimental systems we set out to reveal the possible role of PU.1+IRF8 co-LREs in gene regulation. We measured the mRNA

level of three well-known IL-4 regulated genes (Retnla, Hbegf, and Arg1) in PUER cells (Pu.1-/- myeloid progenitor cells containing a PU.1-estrogen receptor ligand binding domain fusion protein), whose transcriptional activity can be turned on by adding tamoxifen. Our results in the PUER system demonstrated that transcriptionally active PU.1 up-regulated the expression of the three IL-4-target genes in the unstimulated state providing the context for efficient IL-4-mediated induction as well. To test the role of IRF8 in IL-4 regulated expression of the same three genes, we used wild type (WT) and Irf8-/- BMDMs. Our result showed that in the lack of Irf8 the inducibility by IL-4 treatment was partially (Hbegf, and Arg1) or completely (Retnla) impaired.

IRF8 maintains low accessible chromatin structure at a subset of labelled regulatory elements

37% of the sites that gained openesss in the absence of *Irf8* overlapped with IRF8-LREs, suggesting that IRF8 binding is indispensable to prevent chromatin opening at these sites. This analysis indicates that IRF8 might have a chromatin compacting effect on a subset of loci, potentially through either indirect binding or binding to non-canonical motifs as we neither identified PU.1-IRF composite motif nor any other ISRE-like elements. These findings highlight the importance of IRF8 as a regulator of chromatin and hints the existence of PU.1+IRF8-co-LREs and IRF8-LREs, where IRF8 may maintain low accessible chromatin, stabilizing the LRE state.

Labelled regulatory elements are dynamically utilized by macrophage polarization signals

Comparison of STAT6 (1h IL-4) and p65 (1h LPS) cistromes revealed 10,619 STAT6-specific, 8,466 p65-specific regions and 3,781 co-bound regulatory elements that are utilized by both TFs upon the certain activating signals. Both cistromes of TF had similar fractions of de novo, LRE- and highly accessible regions. Co-bound genomic regions showed high overlap with highly accessible sites (2,720), but only 892 labelled and a very small set of de novo sites (169) were also identified.

Motif analysis of NF-κB (represented by p65), STAT6 and PU.1 highlighted that: 1) the motif of PU.1 was weaker in almost all the cases at de novo sites and LREs compared to highly accessible sites, with no regard as to whether it overlapped by STAT6, p65 or both; 2) as expected, the p-value of the NF-κB motif was lower at p65-specific sites, and the de novo

regions had the strongest motifs. Moreover, co-bound sites possessed a much weaker NF-κB motif, while this motif was virtually not detected at STAT6-only sites; 3) the STAT6 motif had very similar characteristics to the p65 motif, showing specificity to STAT6-bound regions with the highest p-values for this motif at the de novo sites.

The cistromes of both p65 and STAT6 highly overlapped with PU.1-, PU.1+IRF8- and PU.1+IRF8+CEPBA+RUNX1-LREs ("All" group), although a higher portion of the STAT6 cistrome overlapped with these groups compared to p65's. Interestingly, p65 showed higher enrichment for IRF8-, RUNX1- and CEBPA-LREs than STAT6, suggesting that in certain cases, SDTFs preferentially use certain LRE groups (without regard to the number of the co-bound TFs). Collectively, these findings highlight that the SDTFs of the two main polarization programs bind LREs, where the SDTFs do not necessarily prefer co-LREs, and show preferential binding to certain TF modules (TF combinations) at LREs.

Our results highlight that LREs have a contribution to short-term polarization programs. To address the question whether co-LREs also have any role in intermediate or long-term transcriptional responses, we investigated the activated enhancer network of classical and alternative macrophage polarization programs using fine resolution, time course RNAPII-pS2 ChIP-seq experiments.

At the global scale, we identified 6,356 LREs that were up-regulated upon IL-4 or LPS treatment at least at one time point (FC>2 & p-value < 0.05, compared to the steady state BMDMs). Next, we clustered the up-regulated LREs into eight enhancer clusters (ECs), all of them with different transcriptional kinetics: EC1, EC2 and EC3 were IL4-specific, while clusters EC4-EC8 were specific for LPS. Having analyzed the cluster sizes, we found that LREs nearly equally contribute to early, intermediate and late induced TF modules.

To reveal whether these co-LRE-associated ECs regulate gene expression with similar kinetics, we filtered the genes based on two conditions: 1) on the gene body, Pol II S2 signal was induced upon IL-4 or LPS compared to the resting BMDM state (1,735, FC>2 & p-value < 0.05) and 2) their TSSs was in the 100-kb vicinity of at least one regulated EC. Next, we clustered these genes into eight clusters (GBs), and they showed similar transcriptional kinetics.

Our study confirms that LREs are dynamically utilized upon SDTF binding and activation both in the classical and alternative polarization programs of macrophages and have significant

role in the early stage of RA-induced neurogenesis. In summary, LREs are widespread both in ESCs and macrophages, extending the working model of enhancer formation and providing a novel class of the regulatory elements which contribute to dynamic gene expression regulation.

Modelling enhancer states using Nondeterministic Finite State Automata

To provide a formal model to describe the states of the chromatin and the transitions among them, we defined a Nondeterministic Finite Automaton. In our case, Q is the set of the possible states of chromatin which are determined by three components: The first one is Binding representing whether or not the given region is bound by any transcription factors (’bound’, ’unbound’), the second component is Openness (‘low’ or ‘high’ accessibility) determined by ATAC-seq signal and the third one is Activity that is represented by the change in active histone mark signal (’down-regulated’, ’inactive’, ’up-regulated’). This automaton can describe the interactions between the LDTF(s) and SDTF(s) in macrophages, and the results could serve as a new-paradigm which can be tested in other mammalian model systems.

OCT4-LREs in the context of RA-induced neurogenesis

Our previous finding raised the inquiring question whether LREs exist in a different cellular context as well and if so, these regulatory elements can also be activated in a signal specific manner. We identified ~21,000 OCT4 binding sites in naïve (“serum”) and ground state (“2i”) ESCs and classified the OCT4 cistrome based on chromatin openness predicted by DNase-seq. Notably, 37% of the OCT4 cistrome in ground state was associated to low accessible chromatin (OCT4+/DNase high and OCT4+/DNase low). The DNase signal showed a good correlation with ATAC-seq signals and this classification was in a good agreement in the naïve state as well.

Pathway analysis on the gene set associated with the OCT4+/DNase low genomic regions and found that these sites were enriched for WNT/b-catenin, Axonal Guidance Signaling, Epithelial Adherens Junction Signaling, PTEN Signaling, Signaling by Rho Family GTPases and RAR signaling pathways including WNT/b-catenin target genes such as Wnt3 and T/Brachyury and RAR:RXR including Hoxa1, Prmt8 and Cdx2 were enriched for OCT4.

Interestingly, upon the knockdown of OCT4 RA target genes showed higher basal mRNA expression. In contrast, the depletion of NANOG did not substantially increase the mRNA level of these genes. This result confirms that there exists a OCT4-related mechanism by which RA-regulated genes such as Hoxa1 are suppressed and this regulation mechanism is

independent from the canonical pluripotency gene expression network. qPCR of measurements confirmed that OCT4 represses the mRNA level of Hoxa1 gene and the eRNA of a putative enhancer 5 kb upstream of the TSS of Hoxa1 bound by OCT4 in the pluripotency state.

Next, we aimed to characterize the interaction between OCT4 and the RA pathway components. Only Rarα and Rarγ were down-regulated at the mRNA level upon the depletion of OCT4, however the latter with much higher induction. qPCR measurements of OCT4 binding and the eRNA level at the putative regulatory region of Rarγ after RNAi transfection revealed that OCT4 regulates Rarγ in the pluripotent state via direct binding to its putative enhancer region.

Modeling OCT4-related transcriptional circuit using network motifs.

This accumulated data led us to conceptualize our knowledge on the role of OCT4 in the early steps of RA-induced neurogenesis. We built a composite motif network consisting of two fundamental modules: (1) A C1-FFL in which OCT4 directly up-regulates the expression level of Rarγ in the ESC state and indirectly up-regulates Hoxa1 up RA treatment via collaborative binding of the RXR:RAR heterodimer (2) OCT4 and Hoxa1 implement a mutual repression motif where either OCT4 is turned on and Hoxa1 is turned off (ESC state) or vice versa (long term RA signaling).

5. DISCUSSION

In the post-genomic era, one of the biggest challenge is to integrate, interpret, and ultimately create meaningful models based on accumulating Next-Generation Sequencing data which require an increasing intellectual contribution from the computer science community. Therefore, interdisciplinary teams need to be set up to understand complex molecular biology processes by building predictive models, which seem to be essential for deciphering the processes of cell functions, replication, transcription, and translation.

Machine learning methods such as Random Forest and Support Vector Regression serve as novel tools to validate hypotheses by building classifiers and regressive models that can not only make predictions, but also determine the relative contribution of the input variables. We have demonstrated that chromatin openness profiled by ATAC-seq can be predicted from the binding pattern of key TFs using Random Forest classifier and Support Vector Regressor models. This computational approach revealed that (1) despite the fact that PU.1 is the most

prominent LDTF of macrophages, neither its solo binding nor the solo binding of other key TFs are not sufficient to predict NFRs, and (2) the whole TF panel we used in the machine learning approach predicted open chromatin with a very high (82%) accuracy, suggesting that the most relevant TFs were included in the model.

This analysis led us to the observation that more than half of the PU.1 binding sites do not overlap with NFRs. Therefore, we examined the potential function of this type of PU.1 binding sites in classical and alternative polarization programs induced by LPS and IL-4, respectively. Our study revealed that a certain part of the 'labelled' sites is opened and activated by the binding of STAT6 (induced by IL-4 – alternative polarization) or p65 (induced by LPS - classical polarization) in a signal-specific manner. A likely scenario is that in resting macrophages, the LDTF(s) attempt to establish open chromatin but no proper TF module is available to maintain open chromatin structure permanently. Further extension of our model will be possible when binding events can be examined at the single cell level and within the time-frame of milliseconds rather that within several minutes and by this way, the biases caused by population-based (bulk) techniques such as ChIP-seq can be avoided. Our results add another layer to the working model of enhancer activation and underline the importance of chromatin openness in shaping cell-type specific enhancer repertoires and gene regulation. Collectively, our results lend support for a spectrum model according to which the transcriptional responses are tailored by the trade-off of the SDTFs and the available enhancer repertoire determined by the LDTFs.

First, on one end of the spectrum there are *de novo* enhancers strongly dependent on the corresponding SDTF in a sequence-specific manner and the LDTFs can neither open nor bind these sites in the resting state. Therefore, at these sites certain SDTFs seem to be mandatory factors to initiate and maintain any regulatory events. Second, the newly identified, labeled enhancers are bound by LDTF(s) in the unstimulated state, but the LDTF(s) do not have the ability to maintain open chromatin permanently. The third class of enhancers is the group of poised enhancers, which are already open before the stimuli meaning that they are strongly supported by the available TFs in the unstimulated state, but they need the SDTF(s) to become activated. In former studies, the binding of PU.1 and H3K4me1/2 was used to define enhancers and H3K27ac (active histone marks) to discriminate between poised and active enhancers (5, 50). Fourth, the up-regulated, constitutively active enhancers can be driven without any stimuli

and in some cases the binding of the SDTF(s) can induce further up-regulation. Our results show that this group of enhancers requires a lower level of sequence-specificity compared to the *de novo* or labelled enhancers, and this raises the possibility that open chromatin has a distinct role in recruiting SDTFs via less-specific DNA binding and/or protein-protein interactions.

The need for formal models capturing the essence of a certain biological process appeared in parallel with the emergence of experimental molecular biology, even before the NGS Revolution. During the modelling process, the most critical step is abstraction, by which we eliminate the unnecessary components and highlight the critical features of the phenomena to be modeled. The Operon-model was the first well-characterized regulatory process representing a very effective mechanism of prokaryotic gene regulation by which the organism can rapidly adapt to external environmental stimuli. Another example, which applied Automata Theory approach, is the Chemoton model proposing an automaton to formalize the essential properties of life (self-replication, metabolism and a bilayer membrane). In this study, we constructed a Nondeterministic Finite Automaton termed Regulation Automaton capturing the essential steps of enhancer formation using only the attributes Binding, Openness and Activity.

We found a similar mechanism in embryonic stem cells. The pluripotency factor OCT4 occupies genomics regions associated to the genes of signaling pathways such as RA signaling that are inactive in the ESC state. These regions are typically low accessible and often positive for the repressive histone mark H3K27me3. Upon certain stimuli, however, these regions will be bound by SDTFs which subsequently activate them by recruiting co-activators, similarly as co-LREs collaborate with STAT6 or p65 upon certain polarization signals in macrophages. Notably, OCT4 maintains the gene expression of Rarγ in the unstimulated state while at low accessible regions it serves as a repressor of RA-target genes such as Hoxa1. In this regard, OCT4 is similar to IRF8 which also seems to have such a function, however, delineation of the repressive role of IRF8 at low accessible regions needs further studies. To formally describe this complex OCT4-related regulatory circuit, we build a composite network from the combination of two characteristic network motifs; (1) an C1-FFL depicting the direct and positive regulation of Rarγ in the unstimulated state and the indirect up-regulation of Hoxa1 through by RXR:RAR heterodimer, (2) the mutual repression of OCT4 and Hoxa1 genes providing a regulatory switch between one of the most prominent pluripotency factor and the

key regulator of early steps of neurogenesis required for the proper patterning of the early mouse hindbrain and the associated neural crest. By using such modeling approaches, these studies can serve as building blocks for creating a firmly established theory of enhancer formation.

6. SUMMARY

Summary 1. We have determined the contribution of key macrophage TFs to chromatin openness and enhancer activation in steady state and polarized mouse macrophages.

Our findings include

- Chromatin openness can be accurately predicted from the binding pattern of key TFs using machine learning methods such Random Forest and Support Vector Regressor both qualitatively and quantitively.
- The machine learning-based classification has also revealed that more than half of the PU.1 cistrome is associated to low accessible chromatin regions termed PU-labelled regulatory elements (LREs). Moreover our results show that having a remarkable fraction of LREs is a general phenomenon among the studied TFs (IRF8, CEBPA and RUNX1).
- Loss/gain of function experiments for PU.1 (PUER system) and IRF8 (Irf8-/-) has shown that PU.1 and IRF8 have an indispensable role in regulating gene expression in the steady state and/or in response to IL-4.
- There are indeed distinct TF modules collaboratively binding labelled regulatory elements (co-LREs) that regulate specific gene expression programs with different dynamics initiated by various macrophage polarizing stimuli including LPS (classical activation) and IL-4 (alternative activation).
- Identification of ~2,300 IRF8-LREs that gained openness in $Irf8^{-/-}$ cells suggesting a repressive role for IRF8 at these LREs.
- A formal model termed Regulation Automaton describing the possible states of enhancer formation and transitions among them.

Summary 2. We have characterized a novel role of OCT4 in the early steps of RA-induced neurogenesis.

Our finding include:

- Characterization of the OCT4 cistrome and its relation to chromatin openness has revealed that there is a remarkably fraction of OCT4 binding that are not associated to open chromatin (OCT4-labelled regulatory elements – OCT4-LREs) both in naïve and ground state.
- Examination of the interaction between OCT4-LREs and retinoic acid signaling pathway has uncovered that OCT4 plays a critical role in not only maintaining pluripotency state but also in the early steps of neuronal differentiation.
- siRNA mediated knockdown of OCT4 has shown that OCT4 is essential to activate differentiation-related enhancers of retinoic acid target genes such as Hoxa1.
- Construction of a composite network built from stereotypical network motifs describing the dual role of OCT4; maintaining the expression of Rxrγ and repressing Hoxa1 in pluripotency state and being essential role in mediating the up-regulation of Hoxa1 via RXR:RAR response elements upon 24h RA treatment.

7. LIST OF PUBLICATIONS

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK//2019.PL
Subject: PhD Publikációs Lista

Candidate: Attila Horváth
Neptun ID: BC1RA5
Doctoral School: Doctoral School of Molecular Cellular and Immune Biology

List of publications related to the dissertation

1. **Horváth, A.**, Dániel, B., Széles, L., Cuaranta-Monroy, I., Czimmerer, Z., Ozgyin, L., Steiner, L., Kiss, M., Simándi, Z., Póliska, S., Giannakis, N., Raineri, E., Gut, I. G., Nagy, B., Nagy, L.: Labelled regulatory elements are pervasive features of the macrophage genome and are dynamically utilized by classical and alternative polarization signals.
Nucleic Acids Res. 47 (6), 2778-2792, 2019.
DOI: http://dx.doi.org/10.1093/nar/gkz118
IF: 11.561 (2017)

2. Simándi, Z., **Horváth, A.**, Wright, L. C., Cuaranta-Monroy, I., De, L. I., Károlyi, K., Sauer, S., Deleuze, J. F., Gudas, L. J., Cowley, S. M., Nagy, L.: OCT4 Acts as an Integrator of Pluripotency and Signal-Induced Differentiation.
Mol. Cell. 63 (4), 647-661, 2016.
DOI: http://dx.doi.org/10.1016/j.molcel.2016.06.039
IF: 14.714

List of other publications

3. Ozgyin, L., **Horváth, A.**, Hevessy, Z., Bálint, B. L.: Extensive epigenetic and transcriptomic variability between genetically identical human B-lymphoblastoid cells with implications in pharmacogenomics research.
Sci Rep. 9, 1-16, 2019.
DOI: http://dx.doi.org/10.1038/s41598-019-40897-9
IF: 4.122 (2017)

4. Ivády, G., Madar, L., Dzsudzsák, E., Koczok, K., Kappelmayer, J., Krulisova, V., Macek, J. M., **Horváth, A.**, Balogh, I.: Analytical parameters and validation of homopolymer detection in a pyrosequencing-based next generation sequencing system.
BMC Genomics. 19, 1-8, 2018.
DOI: http://dx.doi.org/10.1186/s12864-018-4544-x
IF: 3.73 (2017)

5. Simándi, Z., Pájer, K., Károlyi, K., Sieler, T., Jiang, L. L., Kolostyák, Z., Sári, Z., Fekecs, Z., Pap, A., Patsalos, A., Contreras, G. A., Rehó, B., Papp, Z., Guo, X., **Horváth, A.**, Kiss, G., Keresztessy, Z., Vámosi, G., Hickman, J., Xu, H., Dormann, D., Hortobágyi, T., Antal, M., Nógrádi, A., Nagy, L.: Arginine Methyltransferase PRMT8 Provides Cellular Stress Tolerance in Aging Motoneurons.
J. Neurosci. 38 (35), 7683-7700, 2018.
DOI: http://dx.doi.org/10.1523/JNEUROSCI.3389-17.2018
IF: 5.97 (2017)

6. Czimmerer, Z., **Horváth, A.**, Dániel, B., Nagy, G., Cuaranta-Monroy, I., Kiss, M., Kolostyák, Z., Póliska, S., Steiner, L., Giannakis, N., Varga, T., Nagy, L.: Dynamic transcriptional control of macrophage miRNA signature via inflammation responsive enhancers revealed using a combination of next generation sequencing-based approaches.
Biochim. Biophys. Acta. Gene Regul. Mech. 1861 (1), 14-28, 2018.
DOI: http://dx.doi.org/10.1016/j.bbagrm.2017.11.003
IF: 5.179 (2017)

7. Fejes, Z., Czimmerer, Z., Szűk, T., Póliska, S., **Horváth, A.**, Balogh, E., Jeney, V., Váradi, J., Fenyvesi, F., Balla, G., Édes, I., Balla, J., Kappelmayer, J., Nagy, B. J.: Endothelial cell activation is attenuated by everolimus via transcriptional and post-transcriptional regulatory mechanisms after drug-eluting coronary stenting.
PLoS One. 13 (6), 1-20, 2018.
DOI: http://dx.doi.org/10.1371/journal.pone.0197890
IF: 2.766 (2017)

8. Czimmerer, Z., Nagy, Z. S., Nagy, G., **Horváth, A.**, Silye-Cseh, T., Kriston, Á., Jonás, D., Sauer, S., Steiner, L., Dániel, B., Deleuze, J. F., Nagy, L.: Extensive and functional overlap of the STAT6 and RXR cistromes in the active enhancer repertoire of human CD14+ monocyte derived differentiating macrophages.
Mol. Cell. Endocrinol. 471, 63-74, 2018.
DOI: http://dx.doi.org/10.1016/j.mce.2017.07.034
IF: 3.563 (2017)

9. Ozgyin, L., **Horváth, A.**, Bálint, B. L.: Lyophilized human cells stored at room temperature preserve multiple RNA species at excellent quality for RNA sequencing.
Oncotarget. 9 (59), 31312-31329, 2018.
DOI: http://dx.doi.org/10.18632/oncotarget.25764

10. Simándi, Z., **Horváth, A.**, Cuaranta-Monroy, I., Sauer, S., Deleuze, J. F., Nagy, L.: RXR heterodimers orchestrate transcriptional control of neurogenesis and cell fate specification.
Mol. Cell. Endocrinol. 471, 51-62, 2018.
DOI: http://dx.doi.org/10.1016/j.mce.2017.07.033
IF: 3.563 (2017)

11. Dániel, B., Nagy, G., **Horváth, A.**, Czimmerer, Z., Cuaranta-Monroy, I., Póliska, S., Hays, T. T., Sauer, S., Francois-Deleuze, J., Nagy, L.: The IL-4/STAT6/PPAR[gamma] signaling axis is driving the expansion of the RXR heterodimer cistrome, providing complex ligand responsiveness in macrophages.
Nucleic Acids Res. 46 (9), 4425-4439, 2018.
DOI: http://dx.doi.org/10.1093/nar/gky157
IF: 11.561 (2017)

12. Dániel, B., Nagy, G., Czimmerer, Z., **Horváth, A.**, Hammers, D. W., Cuaranta-Monroy, I., Póliska, S., Tzerpos, P., Kolostyák, Z., Hays, T. T., Patsalos, A., Houtman, R., Sauer, S., Francois-Deleuze, J., Rastinejad, F., Bálint, B. L., Sweeney, H. L., Nagy, L.: The Nuclear Receptor PPAR[gamma] Controls Progressive Macrophage Polarization as a Ligand-Insensitive Epigenomic Ratchet of Transcriptional Memory.
Immunity. 49 (4), 615-626, 2018.
DOI: http://dx.doi.org/10.1016/j.immuni.2018.09.005
IF: 19.734 (2017)

13. Czimmerer, Z., Dániel, B., **Horváth, A.**, Rückerl, D., Nagy, G., Kiss, M., Peloquin, M., Budai, M. M., Cuaranta-Monroy, I., Simándi, Z., Steiner, L., Nagy, B. J., Póliska, S., Bankó, C., Bacsó, Z., Schulman, I. G., Sauer, S., Deleuze, J. F., Allen, J. E., Benkő, S., Nagy, L.: The Transcription Factor STAT6 Mediates Direct Repression of Inflammatory Enhancers and Limits Activation of Alternatively Polarized Macrophages.
Immunity. 48 (1), 75-90, 2018.
DOI: http://dx.doi.org/10.1016/j.immuni.2017.12.010
IF: 19.734 (2017)

14. Gál, A., Balicza, P., Weaver, D. A., Naghdi, S., Joseph, S. K., Várnai, P., Gyuris, T., **Horváth, A.**, Nagy, L., Seifert, E. L., Molnár, M. J., Hajnóczky, G.: MSTO1 is a cytoplasmic pro-mitochondrial fusion protein, whose mutation induces myopathy and ataxia in humans.
EMBO Mol Med. 9 (7), 967-984, 2017.
DOI: http://dx.doi.org/10.15252/emmm.201607058
IF: 10.293

15. Imre, L., Simándi, Z., **Horváth, A.**, Fenyőfalvi, G., Ifj., N. P. P., Niaki, E. F., Hegedűs, É., Bacsó, Z., Weyemi, U., Mauser, R., Ausio, J., Jeltsch, A., Bonner, W., Nagy, L., Kimura, H., Szabó, G.: Nucleosome stability measured in situ by automated quantitative imaging.
Sci Rep. 7 (1), 1-15, 2017.
DOI: http://dx.doi.org/10.1038/s41598-017-12608-9
IF: 4.122

16. Kiss, M., Czimmerer, Z., Nagy, G., Bieniasz-Krzywiec, P., Ehling, M., Pap, A., Póliska, S., Botó, P., Tzerpos, P., **Horváth, A.**, Kolostyák, Z., Dániel, B., Szatmári, I., Mazzone, M., Nagy, L.: Retinoid X receptor suppresses a metastasis-promoting transcriptional program in myeloid cells via a ligand-insensitive mechanism.
Proc. Natl. Acad. Sci. U. S. A. 114 (40), 10725-10730, 2017.
DOI: http://dx.doi.org/10.1073/pnas.1700785114
IF: 9.504

17. Varga, T., Mounier, R., **Horváth, A.**, Cuvellier, S., Dumont, F., Póliska, S., Ardjoune, H., Juban, G., Nagy, L., Chazaud, B.: Highly Dynamic Transcriptional Signature of Distinct Macrophage Subsets during Sterile Inflammation, Resolution, and Tissue Repair.
J. Immunol. 196 (11), 4771-4782, 2016.
DOI: http://dx.doi.org/10.4049/jimmunol.1502490
IF: 4.856

18. Varga, T., Mounier, R., Patsalos, A., Gogolák, P., Peloquin, M., **Horváth, A.**, Pap, A., Dániel, B., Nagy, G., Pintye, É., Póliska, S., Cuvellier, S., Ben Larbi, S., Sansbury, B. E., Spite, M., Brown, C. W., Chazaud, B., Nagy, L.: Macrophage PPAR[gamma], a Lipid Activated Transcription Factor Controls the Growth Factor GDF3 and Skeletal Muscle Regeneration.
Immunity. 45 (5), 1038-1051, 2016.
DOI: http://dx.doi.org/10.1016/j.immuni.2016.10.016
IF: 22.845

19. Simándi, Z., **Horváth, A.**, Nagy, P., Nagy, L.: Prediction and Validation of Gene Regulatory Elements Activated During Retinoic Acid Induced Embryonic Stem Cell Differentiation.
JoVE. 2016 (112), e53978, 2016.
DOI: http://dx.doi.org/10.3791/53978
IF: 1.232

20. Czimmerer, Z., Varga, T., Kiss, M., Vázquez, C. O., Doan-Xuan, Q. M., Rückerl, D., Tattikota, S. G., Yan, X., Nagy, Z. S., Dániel, B., Póliska, S., **Horváth, A.**, Nagy, G., Varallyay, É., Poy, M. N., Allen, J. E., Bacsó, Z., Abreu-Goodger, C., Nagy, L.: The IL-4/STAT6 signaling axis establishes a conserved microRNA signature in human and mouse macrophages regulating cell survival via miR-342-3p.
Genome Med. 8 (1), 1-22, 2016.
DOI: http://dx.doi.org/10.1186/s13073-016-0315-y
IF: 7.071

21. Rácz, R., Bereczki, J., Sramkó, G., Kosztolányi, A., Tóth, J. P., Póliska, S., **Horváth, A.**, Barta, E., Barta, Z.: Isolation and Characterisation of 15 Microsatellite Loci from Lethrus apterus (Coleoptera: Geotrupidae).
Ann. Zool. Fenn. 52 (1-2), 45-50, 2015.
DOI: http://dx.doi.org/10.5735/086.052.0204
IF: 0.753

22. Simándi, Z., Czipa, E., **Horváth, A.**, Kőszeghy, Á., Bordás, C., Póliska, S., Juhász, I., Imre, L., Szabó, G., Dezső, B., Barta, E., Sauer, S., Károlyi, K., Kovács, I., Hutóczki, G., Bognár, L., Klekner, Á., Szűcs, P., Bálint, B. L., Nagy, L.: PRMT1 and PRMT8 regulate retinoic acid-dependent neuronal differentiation with implications to neuropathology.
Stem Cells. 33 (3), 726-741, 2015.
DOI: http://dx.doi.org/10.1002/stem.1894
IF: 5.902

23. Cuaranta-Monroy, I., Simándi, Z., Kolostyák, Z., Doan-Xuan, Q. M., Póliska, S., **Horváth, A.**, Nagy, G., Bacsó, Z., Nagy, L.: Highly efficient differentiation of embryonic stem cells into adipocytes by ascorbic acid.
Stem Cell Res. 13 (1), 88-97, 2014.
DOI: http://dx.doi.org/10.1016/j.scr.2014.04.015
IF: 3.693

24. Gyöngyösi, A., Dócs, O., Czimmerer, Z., Orosz, L., **Horváth, A.**, Török, O., Méhes, G., Nagy, L., Bálint, B. L.: Measuring expression levels of small regulatory RNA molecules from body fluids and formalin-fixed, paraffin-embedded samples.
Methods Mol. Biol. 1182, 105-119, 2014.
DOI: http://dx.doi.org/10.1007/978-1-4939-1062-5_10

25. Dániel, B., Nagy, G., Hah, N., **Horváth, A.**, Czimmerer, Z., Póliska, S., Gyuris, T., Keirsse, J., Gysemans, C., Van Ginderachter, J. A., Bálint, B. L., Evans, R. M., Barta, E., Nagy, L.: The active enhancer network operated by liganded RXR supports angiogenic activity in macrophages.
Genes Dev. 28 (14), 1562-1577, 2014.
DOI: http://dx.doi.org/10.1101/gad.242685.114
IF: 10.798

26. Czimmerer, Z., Hulvely, J., Simándi, Z., Varallyay, É., Havelda, Z., Szabó, E., Varga, A., Dezső, B., Balogh, M., **Horváth, A.**, Domokos, B., Török, Z., Nagy, L., Bálint, B. L.: A versatile method to design stem-loop primer-based quantitative PCR assays for detecting small regulatory RNA molecules.
PLoS One. 8 (1), 1-10, 2013.
DOI: http://dx.doi.org/10.1371/journal.pone.0055168
IF: 3.534

27. Meskó, B., Póliska, S., Váncsa, A., Szekanecz, Z., Palatka, K., Holló, Z., **Horváth, A.**, Steiner, L., Zahuczky, G., Podani, J., Nagy, L.: Peripheral blood derived gene panels predict response to infliximab in rheumatoid arthritis and Crohn's disease.
Genome Med. 5 (6), 59-69, 2013.
DOI: http://dx.doi.org/10.1186/gm463
IF: 4.942

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

28. Laczik, M., Tukacs, E., Uzonyi, B., Domokos, B., Doma, Z., Kiss, M., **Horváth, A.**, Batta, Z., Maros-Szabó, Z., Török, Z.: Geno viewer, a SAM/BAM viewer tool. *Bioinformation. 8* (2), 107-109, 2012.

Total IF of journals (all publications): 195,742
Total IF of journals (publications related to the dissertation): 26,275

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

03 May, 2019

DEBRECENI EGYETEM * Egyetemi és Nemzeti Könyvtár *