Alena van Bömmel[1], *Mike Love*[1], *Ho-Ryun Chung*[2] *and Martin Vingron*[1]

(1) Computational Molecular Biology, Max Planck Institute for Molecular Genetics, Berlin, Germany
(2) Otto-Warburg Laboratory, Computational Epigenomics Group, Max Planck Institute for Molecular Genetics, Berlin, Germany

Detection of co-regulating transcription factors in 34 human cell types using predicted DNA-binding affinity on DNase hypersensitive sites.

BACKGROUND: Cell-type-specific gene expression is regulated by combinatorial interactions among transcription factors (TFs) binding to the DNA. Information about TFs' binding affinity to distal and proximal regulatory sequences can help determine which combinations of factors work together to regulate their target genes in cell-type-specific manner.

RESULTS: In this study, we provide detection of co-regulating TF pairs in 34 healthy human cell types which is based on statistical analysis of estimated ranked lists of TFs' target regions. Specifically, we first scanned all cell-type-specific DNase hypersensitive sites (DHSs) for single TF-DNA binding affinities using known motifs for 160 TFs and ranked the DHSs by their predicted binding affinity separately for each TF. We then studied the similarity of pairs of the ranked lists stratified by cell type by applying a statistical test for multiway contingency tables. Our significant TF pairs defined by the test in each cell type were validated by known protein-protein interactions (PPIs) and by detected co-binding of TFs in ChIP-seq data. We found that the known PPIs are significantly enriched (up to 12 fold) in the groups of our predicted co-regulating TFs and that we can recover a majority (56%) of predicted co-binding TF pairs from the ChIP-seq analysis. Furthermore, the predicted co-regulating TFs are supported in literature to be active regulators in the corresponding cell types.

CONCLUSION: Our findings show that the cell-type-specific gene expression is regulated by a large number of combinatorial TF interactions with dominating central regulators. However, the TF interaction networks substantially differ even for related cell lines.