科学研究費助成事業　　研究成果報告書

平成　２６　年　　５　月　２９　日現在

機関番号：　１４４０１
研究種目：　若手研究(B)
研究期間：　2011～2013
課題番号：　２３７１０２３４

研究課題名（和文）免疫応答の制御に関与するプロモータの網羅的な構造特徴の抽出とモデル化

研究課題名（英文）Comprehensive feature extraction and modeling of promoters controlling the immune response

研究代表者
ＶＡＮＤＥＮＢＯＮ　ＡＬＥＸＩＳ（VANDENBON, ALEXIS）

大阪大学・免疫学フロンティア研究センター・特任助教

研究者番号：６０５７０１４０

交付決定額（研究期間全体）：（直接経費）　　3,500,000円、（間接経費）　　1,050,000円

研究成果の概要（和文）：免疫システムの遺伝子発現制御の大規模解析を転写階層とエピジェネティクス階層について行った。転写因子ペアによるコンビナトリアル制御および転写因子の結合し易い領域を解析する手法を開発した。また、樹状細胞での転写因子のエンハンサーへのコンビナトリアル結合を解析した。さらに樹状細胞でのヒストン修飾およびRNAポリメラーゼII結合を実験的にゲノムワイドで検出し、リポ多糖刺激後のそれらの変化を解析した。

研究成果の概要（英文）：We conducted a large-scale analysis of gene expression control in the immune system on the transcriptional and epigenetic level. First, we developed methods for analysing combinatorial regulation by pairs of transcription factors (TFs), and positional preferences of binding of TFs to regulatory regions. Second, we analysed combinatorial binding of TFs to enhancers in dendritic cells. Finally, we performed experiments for genome-wide detection of several histone modifications and RNA polymerase II binding in dendritic cells, and analysed their changes after stimulation with lipopolysaccharide.

研究分野：　複合新領域

科研費の分科・細目：　ゲノム学・システムゲノム学

キーワード：　自然免疫　バイオインフォマティクス　配列解析　転写制御　エピジェネティクス　転写因子　マクロファージ　樹状細胞

１．研究開始当初の背景

Our bodies are protected from pathogens by a complex system of biological structures and processes, referred to as the immune system. When encountering a pathogen, signaling pathways are activated in specialized immune cells, resulting in the activation of regulators and subsequent changes in gene expression. These regulators include transcription factors (TFs) such as NF-κB and IRF family members [Takeuchi and Akira, Cell, 2010].

TFs play a critical role in the control of the immune response and in the regulation of transcription in general. They bind to specific DNA motifs within regulatory regions of genes, and facilitate the recruitment of RNA Polymerase II (Pol II) to the promoters of genes.

Recently, increasing attention is being paid to epigenetic regulation of gene expression. Local chromatin structure influences the accessibility of DNA to binding by TFs and Pol II, and depends on modifications such as DNA methylation and methylation or acetylation of specific amino acid residues in histone proteins. The importance of regulation of gene expression by TFs as well as by epigenetics is now widely accepted.

A number of studies have attempted to dissect the structure of regulatory regions controlling gene expression. In one study, we proposed a model that takes into account the presence of regulatory motifs, and their positioning with regard to each other and to the transcription start site [Vandenbon and Nakai, Nucleic Acids Research, 2010]. A limited number of studies have identified interactions between TFs and chromatin modifiers in the immune system [Gilchrist *et al.*, Nature, 2008].

However, a lot remains unclear. For example, by what mechanisms are epigenetic markers for enhancers and promoters defined? How do lineage-restricted, cell type-specific, and immune stimulus-activated TFs together regulate expression changes after stimulation? How does stimulus-induced activation of TFs influence changes in epigenetic markers, and vice versa? Particularly, very little is known about how TFs and chromatin modifiers work in concert to regulate gene expression changes.

２．研究の目的

The goal of this study was to increase our understanding of the regulation of the immune response, on a genome-wide level. We were interested in the structural features of regulatory regions, combinatorial regulation of transcription by pairs or sets of TFs, dynamics in chromatin structure over time after immune stimulation, and mutual influences between transcriptional and epigenetic regulation of gene expression.

３．研究の方法

(1) Combinatorial regulation of transcription.

We developed a new measure for predicting combinatorial regulation by pairs of TFs, based on the tendency of their binding sites to co-occur in regulatory regions. In brief, our method is based on the hypothesis that TFs whose binding sites are often present in the same regulatory regions together are likely to be involved in combinatorial regulation of their target genes. We applied our methodology on a large number of sets of co-expressed genes, including genes that are induced upon stimulation with various immune stimuli [Amit *et al.*, Science, 2009]. In the promoters of a set of induced genes, we found significant co-occurrence of binding sites for the TFs NF-κB and C/EBPα. For a number of genes, we experimentally confirmed the combinatorial regulation by these 2 TFs, using luciferase assays in which gene promoters were inserted in a plasmid upstream of a luciferase encoding regions. Following overexpression of each of the two TFs separately, and both TFs together, luciferase activity was measured in order to evaluate the presence or absence of combinatorial regulation between the two TFs.

(2) Positional preferences of TF binding.

We developed a new methodology for analysing preferential positioning of TF binding sites in regulatory regions with regard to a landmark, such as the transcription start site of genes. In brief, our method is based on the Parzen window method for estimating sample densities. In this study, the local density (or frequency) of regulatory motifs in sets of co-expressed genes is measured at specific distances from the genes' transcription start sites. Regions with significant enrichment of DNA motifs might reflect preferential positioning of TF binding in the input promoters with regard to

the transcription start site of chromatin related features. We applied our method on a large number of co-expressed genes, including immune-related genes.

(3) Sets of TFs binding to enhancers.
Using public ChIP-seq data for histone modifications [Ghisletti *et al.*, Immunity, 2010] and TSS-seq data, we defined enhancer regions in mouse macrophages. Subsequently, we merged this data with an additional set of ChIP-seq data for 25 TFs in closely related dendritic cells (DCs) [Garber *et al.*, Molecular Cell, 2012], and classified enhancers according to the combination of TFs binding them. Finally, we analysed for each such class of enhancers their properties and the properties of nearby genes, such as a) changes in expression levels of nearby genes after immune stimulation, b) binding by stimulus-activated TFs, and c) correlations in changes in TF binding to adjacently located enhancers.

(4) Dynamics of epigenetics after immune stimulation.
We conducted ChIP-seq experiments and obtained data for mouse DCs, in a time series of 10 time points after LPS stimulation (0h, 0.5h, 1h, 2h, 3h, 4h, 6h, 8h, 16h, and 24h), for RNA Polymerase II and a set of histone modifications (H3K4me1, H3K4me3, H3K9me3, H3K9K14Ac, H3K27ac, H3K27me3, H3K36me3). This large-scale time series data was processed and normalized, and combined with other data our collaborators had generated previously, for the same cell type and stimulation: gene expression data (RNA-seq), and transcription initiation event data (TSS-seq). Finally, we merged our data with publicly available ChIP-seq data for TF binding events in DCs stimulated with the same stimulus [Garber *et al.*, Molecular Cell, 2012]. Together, this data represents a unique resource for studying the dynamics of TF binding and epigenetics after immune stimulation of DCs. We designed a Hidden Markov Model that can capture changes in Pol-II binding and histone modifications following LPS stimulation. We analyzed the general tendencies of changes in enhancers and promoters, with special attention for LPS-inducible genes. We also analyzed potential correlations between TF binding events and epigenetic changes.

４．研究成果
(1) Combinatorial regulation of transcription.
Several studies have shown the importance of cooperativity between pairs or sets of regulators in controlling gene expression. Moreover, some studies suggest that the cooperative binding of TFs is important in the displacement of nucleosomes and thus plays an important role in changing chromatin structure. In this study we developed a method for the prediction of combinatorial regulation of sets of co-expressed genes by co-occurrences of transcription factor binding sites in their regulatory regions. The key result of this study was the prediction of combinatorial regulation by the TFs NF-κB and C/EBPα of a set of genes that are induced upon TLR stimulation in DCs. For a number of these genes, we experimentally confirmed this combinatorial regulation. One example, Lcn2, is shown in Fig. 1. The Lcn2 promoter was cloned into a plasmid upstream of a luciferase encoding region. Luciferase activity was induced >30-fold compared to control samples when cells

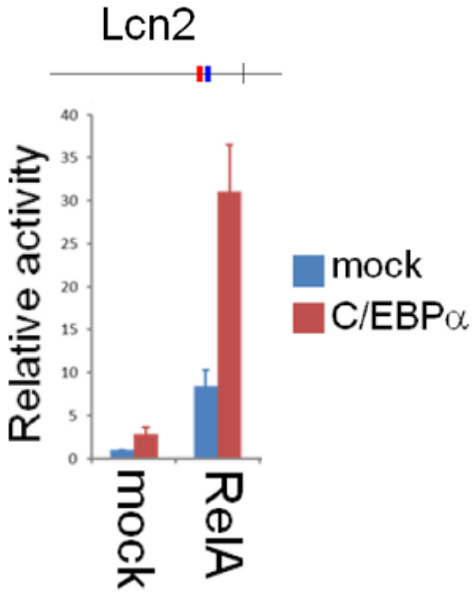

Fig. 1: Luciferase assay of the Lcn2 promoter. At the top, the Lcn2 promoter is shown with the predicted binding sites for RelA (red box) and C/EBPα (blue box). Below, the relative activity is shown for cells transfected with the Lcn2-luciferase plasmid only (mock), or in combination with the RelA plasmid with or without the C/EBPα plasmid. A >30-fold induction of luciferase activity was observed in cells transfected with both Rela and C/EBPα plasmids.

were transfected with plasmids encoding Rela (a subunit of NF-κB) and C/EBPα. This is a considerably higher induction than observed in cells transfected with the plasmid for Rela only, or C/EBPα only, suggesting co-operative regulation of the Lcn2 promoter by these 2 TFs.
An online tool has been constructed for this method (http://sysimm.ifrec.osaka-u.ac.jp/tfbs/remoled/).

(2) Positional preferences of TF binding.
In a second study, we developed a methodology for predicting preferential positioning with regard to some landmark of TF binding within regulatory region. Using this methodology, we can a) predict enrichment of regulatory motifs that are often missed by standard approaches, and b) study the positioning of TF binding relative to the transcription start site or centers of enhancer regions. Importantly, in this study we found that the majority of preferential enrichment of TF binding sites in promoter regions occurs in the region roughly between -300 and +300, with a peak around position -100 (Fig. 2). On the other hand, very little local enrichment was found in the region -700 to -500. These regions are likely to reflect epigenetic features of promoter regions. Strong positional preferences of TF binding might reflect interaction between TFs and nucleosome positioning. We applied this method to a large number of co-expressed genes, including genes that are induced upon TLR stimulation in DCs, and found position-specific enrichment of binding sites for several TFs.
An online tool has been constructed for this method (http://sysimm.ifrec.osaka-u.ac.jp/tfbs/locamo/).

(3) Sets of TFs binding to enhancers.
Using publicly available ChIP-seq data for enhancer-associated epigenetic markers and transcription initiation events we defined a genome-wide set of enhancers in macrophages. Combining these enhancers with publicly available genome-wide data for TF binding events for 25 TFs, we studied the sets of TFs binding to enhancers in DCs before and after stimulation with LPS.

In brief, we confirmed that enhancers located proximally to induced genes tend to be bound by key regulators of the immune response such as NF-κB, IRFs, and members of the STAT family of TFs. In addition, we found that induced genes tend to be associated with enhancers that are bound by a specific set of TFs already before stimulation occurs. We obtained results suggesting that pre-bound and lineage-specific TFs influence the binding of induced TFs, and vice versa. Finally, our results suggest that TF binding is influenced by the presence of nearby enhancers and the TFs binding them.

(4) Dynamics of epigenetics after immune stimulation.
Using the present grant we generated our own ChIP-seq data for a selection of histone modifications and Pol-II, and we combined this data with gene expression data and publicly available ChIP-seq data for a selection of TFs.
Using the above data we defined promoter regions and enhancer regions, and designed a hidden Markov model (HMM) that captures the states and dynamics in RNA Pol2 binding

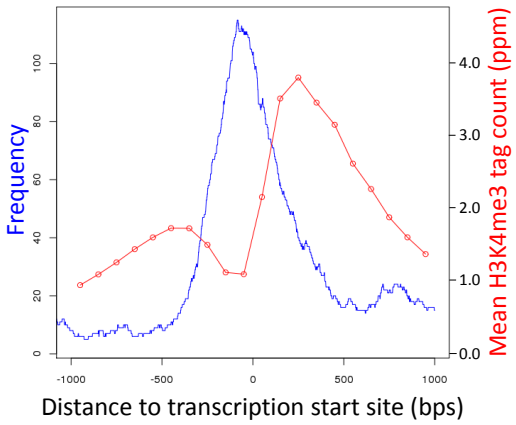

Fig. 2: Frequency of local TF binding site enrichment vs H3K4me3 histone modification around mouse transcription start sites. This graph shows the number of times we observed any position within promoters (X axis) to be located within a region of significant local enrichment in regulatory motifs (blue; left Y axis), and the average amount of H3K4me3 mapped reads (ppm) in bins of 100 bps around the genome-wide set of mouse transcription start sites (blue; right Y axis). The peak in local enrichment of regulatory motifs is located between the two peaks of H3K4me3 modifications.

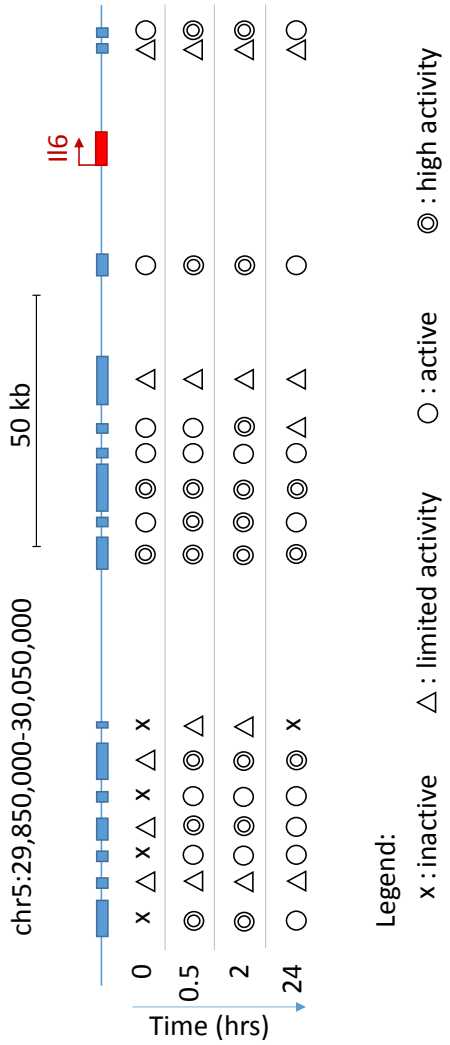

Fig. 3: Changes over time in activity of enhancers surrounding Il6 after LPS stimulation. The genomic region chr5:29,850,000-30,050,000 is shown, for time points 0, 0.5, 2, and 24 hours. Il6 is indicated in red, and blue boxes indicate H3K4me1-marked regions. Symbols represent enhancer activities: x: no activity; △: limited activity; ○: active; ◎: high activity.

and histone modifications at enhancers. Our model includes states with properties of active enhancers, repressed enhancers, and poised enhancers, as well as the transitions between these states over time. We found that after stimulation a total of 2,273 enhancers move from a poised state to an active state within 4 hours after stimulation, 1,113 of which are activated within the first 30 minutes. On the other hand, 1,343 enhancers become inactive, 533 of which within the first 30 minutes after stimulation. These results suggest considerable dynamics in epigenetic markers following immune stimulation. We also found significant associations between TF binding and changes in epigenetic markers.

As an illustration, we here describe the changes observed in the H3K4me1-marked regions surrounding the Il6 promoter (see Fig. 3). Il6 expression is rapidly induced upon LPS stimulation. A number of enhancers surrounding Il6 show epigenetic changes reflecting an increase in activity after stimulation, while another subset is already in an active state even before stimulation. One subset is rapidly activated after stimulation (0.5 h), coinciding with the rapid induction of Il6 expression. After 2 hours, the expression of Il6 reaches its peak, and with it the activation level of the enhancers. After 24 hours, the expression of Il6 has substantially dropped, and its enhancers have partly returned to their pre-stimulation states. Importantly, however, some retain a higher level of activation.

We have constructed an online database where our data can be visually inspected (http://sysimm.ifrec.osaka-u.ac.jp/genomebrowser/).

５．主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 2 件）

① Vandenbon A, Kumagai Y, Teraguchi S, Amada KM, Akira S, Standley DM, “A Parzen window-based approach for the detection of locally enriched transcription factor binding sites”, *BMC Bioinformatics*, Vol. 14, pp. 26, 2013, DOI: 10.1186/1471-2105-14-26. (査読有)

② Vandenbon A, Kumagai Y, Akira S, Standley DM, “A novel unbiased measure for motif co-occurrence predicts combinatorial regulation of transcription”, *BMC Genomics*, Vol. 13 Suppl 7, pp. S11, 2012, DOI: 10.1186/1471-2164-13-S7-S11. (査読有)

〔学会発表〕（計 8 件）

① Vandenbon A., “Sets of principal transcription factors define properties and dynamics of enhancers following LPS stimulation”, EMBO Workshop

Complex Systems in Immunology, Biopolis (Singapore), December 2-4, 2013.

② Vandenbon A., “Properties and dynamics of enhancers are defined by small sets of principal regulators”, 次世代シークエンサ現場の会　第３回研究会 (NGS 現場の会), Kobe International Conference Center, Kobe, (Japan), September 4-5, 2013.

③ Vandenbon A., “Towards Comprehensive Understanding of the Dynamic Control of the Immune Response on the Transcriptional and Epigenetic Level”, The 24th CDB Meeting: Genomics and Epigenomics with Deep Sequencing, RIKEN CDB, Kobe (Japan), June 13-14, 2013.

④ Vandenbon A., “Sequence Analysis Reveals Several Subtypes of Macrophage Enhancers”, TCUID2012, Osaka University, Osaka (Japan), October 29-31, 2012.

⑤ Vandenbon A., “A novel unbiased measure for motif co-occurrence predicts combinatorial regulation of transcription”, 11th International Conference on Bioinformatics (InCoB2012), Centara Grand at Central Plaza Ladprao, Bangkok (Thailand), 5 October 2012.

⑥ Vandenbon A., “Predicting combinatorial regulation of transcription from motif co-occurrence”, InCoB2012 (11th International Conference on Bioinformatics), Centara Grand at Central Plaza Ladprao, Bangkok (Thailand), October 3-5, 2012.

⑦ Vandenbon A., “Sequence analysis of mouse macrophage promoter and enhancer regions reveals several subclasses of enhancers”, ICSB (the 13th International Conference of Systems Biology), University of Toronto, Toronto (Canada), August 19-23, 2012.

⑧ Vandenbon A., “Sequence analysis of promoter and enhancer regions in mouse macrophages”, 次世代シークエンサ現場の会　第２回研究会 (NGS 現場の会), Hotel Hankyu Expo Park, Osaka (Japan), May 24-25, 2012.

〔その他〕
ホームページ等
http://sysimm.ifrec.osaka-u.ac.jp/genomebrowser/
http://sysimm.ifrec.osaka-u.ac.jp/tfbs/remoled/
http://sysimm.ifrec.osaka-u.ac.jp/tfbs/locamo/

６．研究組織
(1)研究代表者
Alexis Vandenbon (バンデンボン　アレクシス)
大阪大学・免疫学フロンティア研究センター・特任助教
研究者番号：60570140

(2)研究分担者
なし

(3)連携研究者
なし