論文の内容の要旨

論文題目　オープンクロマチン解析による脂肪細胞分化制御因子の同定

氏名　　中村　正裕

Identification of regulatory elements within the genome is crucial for understanding the mechanisms that govern cell-type-specific gene expression. Recent technological advances—such as chromatin immunoprecipitation followed by either microarray analysis or deep sequencing (ChIP-chip and ChIP-seq)—have led to integrated understanding of multilayered mechanisms of transcriptional regulation that range in complexity from DNA methylation to binding of transcription factors and chromatin remodelers to histone modification to regulation of higher-order chromatin structures. Genome-wide profiling of regulatory elements in the genome by different methods is currently an area of intense research activity. Traditionally, regulatory elements have been identified by DNase I hypersensitivity assay combined with Southern blot analysis, which detects accessible open chromatin DNA regions that are prone to nuclease digestion. DNase I hypersensitivity assay coupled with microarray or high-throughput sequencing (DNase-Chip or DNase-seq) were shown to be effectively applied in genome-wide identification of open chromatin regions.

The adipocyte is central in controlling energy balance and whole-body glucose and lipid homeostasis. Peroxisome proliferator-activated receptor gamma (PPARγ; NR1C3) is both necessary and sufficient for adipocyte differentiation. Necessary for both development and maintenance of mature adipocytes, PPARγ is crucial in systemic glucose and lipid

homeostasis, and, importantly, is the molecular target of thiazolidinediones, widely prescribed for obese diabetics. CCAAT enhancer binding proteins (C/EBPs; C/EBPα-β-δ) act with PPARγ, forming the adipogenic transcription cascade. C/EBPβ and δ are induced by adipogenic stimulus, inducing PPARγ, which activates expression of C/EBPα, which binds and further activates expression of PPARγ, providing a positive regulatory loop.

Genome-wide approaches now dissect the transcriptional mechanisms of adipocyte differentiation. ChIP-chip or ChIP-seq studies of adipogenic regulators have provided valuable mechanistic insights into adipogenic transcription never before gained by conventional experiments: New concepts include co-localization of PPARγ and cell-type-specific transcription factors, low conservation rate of PPARγ binding sites between murine and human adipocytes and the role of C/EBPβ as a pioneer factor that establishes ―hot spots where multiple adipogenic regulators cooperatively work in the very early stage of differentiation.

We generated genome-wide maps of open chromatin sites in 3T3-L1 adipocytes (on day 0 and day 8 of differentiation) and NIH-3T3 fibroblasts using formaldehyde-assisted isolation of regulatory elements coupled with high-throughput sequencing (FAIRE-seq). FAIRE detects open chromatin structure much the way the DNase I hypersensitivity assay does, but with advantages, like obviating the need for clean nuclei preparation and laborious enzyme titrations. Coupled with high-throughput sequencing (FAIRE-seq), FAIRE allows unbiased identification of potential regulatory elements without requiring prior knowledge of binding factors. This approach identified in the genome 37,781 FAIRE peaks in 3T3-L1 on day 0 and 26,611 on day 8, plus 36,111 in NIH-3T3 cells. Consistent with previous observations, 28% of the FAIRE peaks were detected near the transcription start sites (TSSs ± 500 bp) of RefSeq genes and are referred to as promoter FAIRE peaks, while 72% were located outside known TSSs, and are referred to as non-promoter FAIRE peaks. FAIRE peaks at the promoter were associated with active transcription and histone modifications of H3K4me3 and H3K27ac. Non-promoter FAIRE peaks were characterized by H3K4me1+/me3-, the signature of enhancers, and were largely located in distal regions. The non-promoter FAIRE peaks showed dynamic change during differentiation while the promoter FAIRE peaks were relatively constant.

By comparing open chromatin regions in preadipocytes and adipocytes, we identified the adipocyte- and preadipocyte-specific FAIRE peaks in the genome. Functionally, we demonstrated that the adipocyte-specific FAIRE peaks were associated with genes up-regulated by adipogenesis while the preadipocyte-specific FAIRE peaks were associated with genes down-regulated by adipogenesis.

We also conducted ChIP-seq analyses using antibodies specific for either PPARγ or RXRα in 3T3-L1 adipocytes at 36 hours and day 8 after induction of differentiation. The number of PPARγ binding sites increased during differentiation while that of RXRα

binding sites remained virtually constant. Significant overlap between the PPARγ and RXRα binding sites was consistent with the heterodimer formation of PPARγ and RXRα. Microarray and GO analysis revealed that the PPARγ binding sites were enriched in the vicinity of genes up-regulated by adipocyte differentiation and the bound genes were associated with adipocyte differentiation and lipid metabolism. We performed de novo motif analysis of genomic regions bound by PPARγ, and found that the AGGTCA-n-AGGTCA (called DR-1) shown was the most over-represented one (E-value 1.3×10^{-55}) . An extension AGT 5' outside of DR-1 appeared to correspond to the direct interaction between the DNA and the hinge region between the DNA-binding domain and the ligand-binding domain. A significant proportion of adipocyte-specific non-promoter FAIRE peaks overlapped the PPARγ/RXRα binding sites. Both PPARγ and C/EBPα binding sites were enriched in the fractions of adipocyte-specific FAIRE peaks, and they respectively accounted for 45.3 % and 11.7 % of the adipocyte-specific FAIRE peaks. These data support the role of PPARγ and C/EBPα as primary transcription factors that drive adipocyte-specific gene expression.

The total number of adipocyte-specific FAIRE peak clusters for different window sizes were calculated and compared with a random data set comprised of the same number of sites. The adipocyte-specific FAIRE peaks had a significantly higher number of clusters in a window size raging from 800bp to ~30kb. Similar results were obtained for the PPARγ binding sites.

Computational motif analyses of the adipocyte-specific FAIRE peaks revealed enrichment of a binding motif for nuclear family I (NFI) transcription factors. The motifs for PPARγ (and other DR1 motifs) and C/EBPs were among the list, consistent with their critical roles in adipogenic transcription. The murine NFI family consists of NFIA, NFIB, NFIC and NFIX, and was identified as a site-specific DNA-binding protein that bound to the adenovirus origin of replication . It forms a dimer to bind to the symmetric consensus sequence TTGGC(N5)GCCAA . The expression change of these factors in in vitro adipocyte differentiation and found that the expression of NFIA and NFIB were significantly induced during differentiation of 3T3-L1 and of another adipogenic cell line, 3T3-F442A . Consistent with this pattern, both NFIA and NFIB were highly expressed in a variety of adipose tissue depots in addition to the brain. ChIP assay showed that NFI occupy the adipocyte-specific FAIRE peaks and/or the PPARγ binding sites near PPARγ, C/EBPα and aP2 genes. Overexpression of NFIA in 3T3-L1 cells resulted in robust induction of these genes and lipid droplet formation without differentiation stimulus. Overexpression of dominant-negative NFIA or siRNA-mediated knockdown of NFIA or NFIB significantly suppressed both induction of genes and lipid accumulation during differentiation, suggesting a physiological function of these factors in the adipogenic program.

Lastly, we compared FAIRE peaks between 'undifferentiated' 3T3-L1 and NIH-3T3 cells.

The 3T3-L1- or NIH-3T3-specific FAIRE peaks were enriched in the vicinity of genes whose expression levels were higher in 3T3-L1 or NIH-3T3, respectively. Motif analysis of the 3T3-L1-specific FAIRE peaks showed that the binding motif for EBF had the highest enrichment ratio (1.81) and a statistically significant p-value of $3.9E^{-3}$. Although the p-value of the motif for PPARγ/RXR did not reach statistical significance, that motif had an enrichment ratio of 1.84. These two factors were among the handful that were proven to transform NIH-3T3 cells into adipocytes when ectopically introduced.

Together, our study demonstrates the utility of FAIRE-seq in providing a global view of cell-type-specific regulatory elements in the genome and in identifying transcriptional regulators of adipocyte differentiation.