NUCLEOSOME FORMATION POTENTIAL OF THE GENE REGULATORY REGIONS

Levitsky V.G.*, Proscura A.P., Podkolodnaya O.A., Ignatieva E.V., Ananko E.A.

Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
* Corresponding author: e-mail: levitsky@bionet.nsc.ru

Keywords: *transcription regulation, nucleosome positioning*

Summary

Motivation: Nucleosome organization of DNA plays the key role in transcription initiation. Exhaustive computer analysis of nucleosome formation potential of the regulatory gene regions appears to be of interest.

Results: Using the RECON program for constructing nucleosome formation potential (NFP) profiles, we analyzed 5 samples of promoters specific in gene expression pattern and 57 samples of transcription factor binding sites (TFBS) regions with flanks. It was shown that the NFP values for promoters are related to tissue- and stage-specificity of gene expression. The NFP pattern of the TFBS regions may also serve as a characteristic of the open chromatin structure.

Availability: http://wwwmgs.bionet.nsc.ru/mgs/programs/recon/

Introduction

An essential feature of eukaryotic DNA is its packaging in chromatin structure. Today, nucleosomes are recognized as highly dynamic units through which the eukaryotic genome can be regulated (Khorasanizadeh, 2004). Hence, the regulated processes of condensation-decondensation of chromatin during the preferential activation of genes are of undeniable importance to gene function. Strong experimental evidence is accumulating for nucleosome positioning in the promoter regions of the eukaryotic genes being of great importance in their regulation of transcription (Goriely *et al*., 2003). Currently, there is a growing number of reports about the interaction of different transcription factors (TF) with nucleosomal DNA (Hsiao *et al.*, 2002). For example, precise positioning of the DNA double helix on the surface of the histone octamer precludes binding of NFI and Oct-1/OTF-1 to their cognate sequences (Truss *et al.*, 1995). The interaction of TF with DNA is a dynamic process. Thus, it was demonstrated that the equilibrium constants describing this interaction decrease progressively from either side of the nucleosomal DNA toward its pseudodyad (Anderson, Widom, 2000). In this way, the distributed nucleosome positioning signals may be crucial in determining chromatin architecture *in vivo* in the regulatory regions. The context code of nucleosome positioning (Kiyama, Trifonov, 2002) controls both the nucleosome packaging in discrete chromatin regions and the accessibility of TFBS.

Based on the different TF-nucleosome interactions, the TFs are subdivided into those capable of binding to the DNA sites within nucleosome (type 1) and those incapable to do so (type 2). Type 1 include factors GR (Li, Wrange, 1995), SP1, GAL4, USF (Chen *et al.*, 1994), type 2 includes TBP, HSF (Taylor *et al.*, 1991) and NF1 (Blomquist *et al.*, 1996). To provide the interaction of type 2 factors with DNA, nucleosome should be displaced, or, at least, its conformational state should be remodeled occasionally by type 1 factors. Our task is to study of the NFP of gene promoters with different expression patterns and genomic sequences containing TFBS.

Methods and Algorithms

To build the NFP, a method based on discriminant analysis and on calculation of dinucleotide frequencies in the local regions of nucleosomal sites was applied (Levitsky *et al.*, 2001; Levitsky,

2004). The NFP positive values agree with significant predictions of nucleosome formation. We have used the confidence level $\alpha = 0.95$ in analysis.

A total number 211 promoters 300 bp long ([–300; +1] relative to the transcription start) were analysed. Classification of promoters according to the gene expression patterns allowed us to distinguish 5 promoter classes (Table). The patterns were identified on the basis of the information stored in the TRRD and literature sources. The nucleosomal organization of the TFBS DNAs of 57 types was studied in the fragments of 160 bp long regions, including an experimentally characterized site occupying the central position. The sequences were extracted from the TRRD database. The TFBS samples contained from 10 to 280 sequences.

Table. Samples of the promoter gene regions

Description	Number of sequences
Genes controlling biosynthesis of steroid hormones	32
Interferon regulated genes	44
Genes regulating intracellular cholesterol level	20
Cell cycle regulated genes	75
Erythroid specific regulated genes	40

Implementation and Results

To further clarify the relation between the nucleosome positioning in promoters and gene expression, we analyzed the NFP profiles for genes of 5 classes showing different expression patterns (Fig. 1).

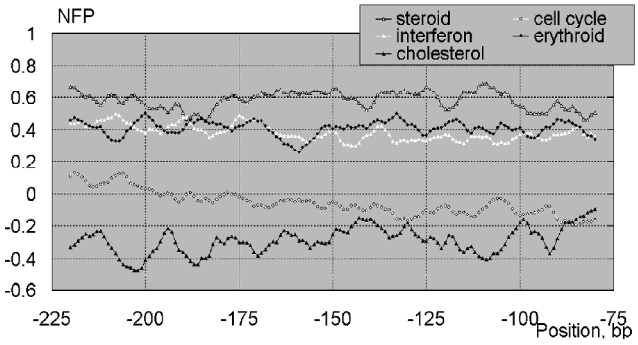

Fig. 1. Nucleosome formation potential profiles for the gene promoters showing different expression patterns: genes controlling biosynthesis of steroid hormones, interferon regulated genes, genes regulating intracellular cholesterol level, cell cycle regulated genes, erythroid specific regulated genes.

Figure 2 gives the NFP average values and confidence intervals for the sample of DNA fragments containing TFBS. It was found that most (46 of 57) of sites are located in the region with the positive NFP. Thus, all the sites were be subdivided into 3 groups: (i) sites with high positive NFP ($\varphi(X) > 0.5$, the total number of sites was 16); (ii) sites with low positive NFP ($0 < \varphi(X) < 0.5$, their number was 30); (iii) sites with negative NFP ($\varphi(X) < 0$, the total number of sites was 11).

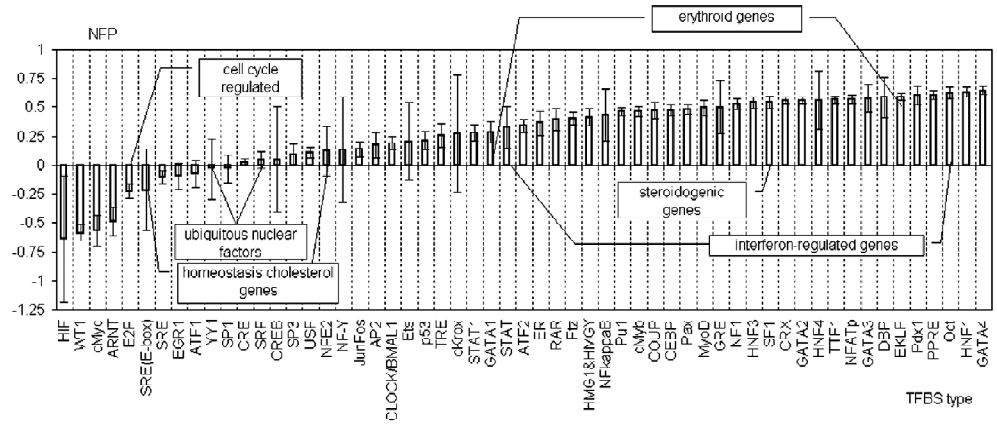

Fig. 2. Average values for nucleosome formation potential and respective confidence intervals for the sample of DNA fragments containing TFBS at the central position.

Discussion

Let us examine the NFP profile values for the promoters showing different expression patterns. The low values for the NFP of the genes regulating intracellular cholesterol level and cell cycle regulated genes presumably reflect the possible easy access of TF to the proximal promoters of these genes, which are regulated for rapid transcription initiation and providing high expression level. Let us now turn to the promoters of the erythroid specific regulated genes, genes controlling biosynthesis of steroid hormones and interferon regulated genes for which the NFP values are high. It should be noted that these genes are tissue-specific and inducible, in contrast to genes regulating intracellular cholesterol level and cell cycle regulated genes. For such genes nucleosome packaging in promoter DNA may be an important element in the regulatory mechanism of transcription. It may be assumed that the state of tightly packed chromatin is normal in this case. This state may be overcome by the appearance of appropriate TFs in cell nuclei. They interact with chromatin and alter or abolish DNA nucleosome packaging. Therefore, tissue-specific and inducible genes possess finer mechanisms for transcription initiation. Change in the nucleosome packaging pattern in promoter DNA and it transition from repressed to active state are indispensable for the transcription initiation mechanism.

Let us examine how the average NFP values for the sample of DNA fragments containing TFBS are distributed. It is noticeable that the distribution (Fig. 2) is consistent the reference of these site to a particular expression patterns of genes (Fig. 1). Thus, for example, the lowest NFP values were for the TFBS occurring in the cell cycle regulated genes (-0.23 E2F, -0.57 cMyc), the genes regulating lipid metabolism (-0.10 SRE), and also ubiquitous TFBS (-0.04 YY1, -0.03 SP1). The highest NFP values were for the TFBS most frequently occurring in the tissue-specific and inducible genes (0.62 Oct, 0.54 SF1).

Acknowledgements

The work was supported by the Russian Foundation for Basic Research (grants Nos. 02-04-48802, 03-04-48555, 03-04-48829, 01-07-90376, 02-07-90355, 03-07-96833, 03-07-90181); Ministry of Industry, Science, and Technologies of the Russian Federation (grant No. 43.073.1.1.1501); Russian Federal Research Development Program Research and Development in Priority Directions of Science and Technology (contract No. 38/2004); NATO (grants Nos. LST.CLG.979816 and LST.CLG 979815), Project No. 10.4 of the RAS Presidium Program “Molecular and Cellular Biology”. The authors are

grateful to Drs. Khlebodarova T.M., Lokhova I.V., Mischenko E.L., Suslov V.V., Stepanenko I.L., Turnaev I.I., Orlov Yu.L. for supporting the work and to Prof. N.A. Kolchanov for valuable discussion.

References

Anderson J.D., Widom J. Sequence and position-dependence of the equilibrium accessibility of nucleosomal DNA target sites // J. Mol. Biol. 2000. V. 296(4). P. 979–987.

Blomquist P., Li Q., Wrange O. The affinity of nuclear factor 1 for its DNA site is drastically reduced by nucleosome organization irrespective of its rotational or translational position // J. Biol. Chem. 1996. V. 271. P. 153–159.

Chen H., Li B., Workman J.L. A histone-binding protein, nucleoplasmin, stimulates transcription factor binding to nucleosomes and factor-induced nucleosome disassembly // EMBO J. 1994. V. 13. P. 380–390.

Goriely S., Demonte D., Nizet S., De Wit D., Willems F., Goldman M., Van Lint C. Human IL-12(p35) gene activation involves selective remodeling of a single nucleosome within a region of the promoter containing critical Sp1-binding sites // Blood. 2003. V. 101(12). P. 4894–4902.

Hsiao P.W., Deroo B.J., Archer T.K. Chromatin remodeling and tissue-selective responses of nuclear hormone receptors // Biochem Cell Biol. 2002. V. 80(3). P. 343–51.

Khorasanizadeh S. The nucleosome: from genomic organization to genomic regulation // Cell. 2004. V. 116(2). P. 259–272.

Kiyama R., Trifonov E.N. What positions nucleosomes? A model // FEBS Lett. 2002. V. 523. P. 7–11.

Levitsky V.G. RECON: a program for prediction of nucleosome formation potential // Nucleic Acids Res. 2004, (in press).

Levitsky V.G., Podkolodnaya O.A., Kolchanov N.A., Podkolodny N.L. Nucleosome formation potential of eukaryotic DNA: tools for calculation and promoters analysis // Bioinformatics. 2001. V. 17. P. 998–1010.

Li Q., Wrange O. Accessibility of a glucocorticoid response element in a nucleosome depends on its rotational positioning // Mol. Cell. Biol. 1995. V. 15. P. 4375–4384.

Truss M., Candau R., Chavez S., Beato M. Transcriptional control by steroid hormones: the role of chromatin. 1995. V. 191. P. 7–23.