Open Journal of Genetics, 2012, 2, 1-4 **OJGen**
Published Online December 2012 (http://www.SciRP.org/journal/ojgen/)

Bioinformatic screening of the binding transcription sites in the regulatory regions of genes up-regulated in response to oxidative stress

Shkurat TP, Ponomareva NS, Aleksandrova AA, Shkurat MA, Butenko AI, Panich AE

Southern Federal University, Rostov-on-Don, Russia
Email: tshkurat@yandex.ru

Received 2012

ABSTRACT

This study focuses on bioinformatics search for new regulatory structures in the non-coding DNA, located around the patterns of gene expression levels changed significantly in response to oxidative stress. Hypothesized that all of the genes increase the expression in response to oxidative stress may have the same motifs in non-coding DNA. To search for motifs created an integrated collection database of transcription binding sites - JASPAR, TRANSFAC, Hocomoco TF Homo sapiens, Uniprobe TF Mus musculus. Two types of regulatory regions: the promoter region and the sequence with the capture of potential cis-regulatory modules. In the regulatory regions of genes increase the expression in response to oxidative stress, in contrast to the gene expression level did not change, families of transcription factors identified SOX (1-30) and HX (A, B, C, D).

Keywords: Gene Expression; DNA Microarrays; Noncoding DNA; Oxidative Stress; Transcription Factor; Sites of Transcription Factor Binding; DNA Motif

1. INTRODUCTION

Non-coding DNA sequences are considered to be a new challenge for molecular genetics, genomics, transcriptomics and proteomics. Studying them is exceptionally important for full understanding of biological processes under normal and pathological conditions. According to the report published by the international project ENCODE in September, 2012, more than 80% of the non-coding DNA is represented by different types RNA, regions responsible for histone modifications, open chromatin regions and transcription factor binding sites [1]. However, it is not yet fully understood how the regulatory information is encoded in DNA in order to define the positions of enhancers, silencers and other distantly operating regulatory elements (Noonan J.P., McCallion A.S., 2010; Lindblad-Toh K, e.a., 2011). Systematic analysis of the transcription profile provides us with the data on regulatory mechanisms and screening for the linked scenario of coding and non-coding DNA interactions that are basically important for understanding of system biology and molecular pathogenesis of a variety of human diseases.

We chose oxidative stress as a model of non-specific pathogenesis. It was previously shown that a single-time treatment of newborn rats with high oxygen pressure modifies intracellular metabolism and results in formation of basically new ratio of pro- and antioxidant activities in the organism. Moreover, this newly formed ratio seems to be stable and may be observed in the first generation progeny [2, 3]. Cells of newborn animals appear to be better used to variable physiological conditions compared to mature cells. It is suggested that treatment of newborn animals with low dose hyperbaric oxygenation will facilitate the formation of a wide spectrum oxygen sensitivity of a neural network that, in turn, will lead to increase in resistance of animals to oxidative stress in further ontogenesis [4]. Full genomic analysis of the transcribed sequences in the brain tissue of these animals revealed that genes, which form at least one expression genetic pattern and regulate six key processes including oxidative metabolism, synaptic transmission, intracellular transport, apoptosis and proliferation, membrane permeability membrane potentials and intercellular contact formation, are involved into the early genetic mechanisms of preadaptation to oxidative stress [7].

Objectives: The present study was aimed at bioinformatics screening of new regulatory structures in the promoter and *cis*-regulatory regions of the non-coding DNA, which participate in gene expression in response to oxidative stress.

2. MATERIALS AND METHODS

The object of the study were white outbreed rats *Rattus norvegicus* treated with high oxygen pressure (0.2 MPa, 1 h). Animals were sacrificed in 3 h after the treatment and gene expression was studied in the frontal lobe. The results of full genomic transcriptome study were previously described in [5].

To perform bioinformatic analysis two groups of genes were formed. The first group involved genes up-regulated in response to oxidative stress. The second group consisted of genes with unchanged expression. The first grouped was formed by genes NM_017138:Lamr1, NM_053440:Stmn2, NM_057207:Sv2b, NM_053339:Acox3, Scn7a, Trpc3, Nid2:NM_213627, Herc1, Ssc1:BC085795, Golph2:NM_023977, Actr1a, Crebzf, Pdk4:NM_053551, Mrpl3, Api5, Zfhx1b, Snrpb:BC083694, Snrpb. The second group included all genes of the chromosome 20 of *R. norvegicus*.

Regulatory regions of 264 genes available in KnownGene with exception of the first and the last gene were studied.

Two types of regulatory elements were studied in both groups. These included the promoter regions (1000, 800 b. p. upstream the start point and 200 b. p. downstream the termination point) and the sequences, overlapping the potential cys-regulatory elements (5000, 4000 b. p. upstream the start point and 1000 b. p. downstream the termination point)

The bioinformatic screening was seeking DNA motifs located near the first group genes. Here, the "recognition motif" implies the way to describe a set of similar oligonucleotides, which can be specifically recognized and bound by a certain regulatory protein.

To search the motifs we created an integrative collection of the binding transcription sites on the basis of JASPAR, TRANSFAC, UCSC ENCODE, Hocomoco TF *Homo sapiens* and Uniprobe TF *Mus musculus* databases. The integrative collection Uniprobe included 272 motifs for transcription factors of *M. musculus*. The collection Hocomoco TF *H. sapiens*, which was obtained by means of integration of the data from different sources, included 332 motives for 321 transcription factor. To perform bioinformatic analysis we chose the release of *R. norvegicus* genome rn4 (Baylor 3.4/rn4)

[http://genome.ucsc.edu/cgi-bin/hgGateway].

The conception of homotypic transcription factor binding site clusters, which were represented by a group of transcription factor binding site, was used as a model of a binding region. To assess the statistical confidence of a binding site cluster the "r-scan" model with fixed motive acceptance threshold was used (Papatsenko D., 2007). The statistical confidence of the "r-scan" was assessed as a probability to find at least the expected number of transcription factor binding site clusters in a sequence of certain length on account that at least 1 transcription factor binding site was already found, i.e. one cluster contains at least one binding site.

In order to perform reliable assessment of the presence of homotypic clusters they were counted in the first and second group. We assessed the chance (p) to choose such a gene, the regulatory region of which would contain at least one cluster (for promoters) and a cluster with a confidence of at least 0.0005 for *cis*-regulatory elements.

The p value can be assessed using the binomial distribution as a confidence of the fact that in the studied sample the transcription factor binding site clusters contain at least k number of regulatory regions, where k corresponds to experimentally found number of sequences which contain the transcription factor binding site clusters (for promoters) / at least 0.0005 confident transcription factor binding site clusters for *cis*-regulatory elements.

3. RESULTS AND DISCUSSION

It was shown that transcription factor binding sites of the SOX subfamily more often preceded the first group of genes (Table 1). The SOX subfamily involves 30 transcription factors, which contain heptamerous sequence (A/T) (A/T) CAA(A/T)G (Fig. 1) [5]. It is considered to be one of the most important subfamilies, which regulate development of both vertebrate and invertebrate animals. Biological functions of these proteins were investigated in a variety of mammalian tissues and cells during embryogenesis and postembryonic development. It is suggested that differentiation of the proteins for groups is typically due to specificity of their functioning in different tissues. They initiate differentiation program and activate tissue specific expression. The SOX subfamily is represented by multifunctional proteins. Among multiple roles that SOX proteins play in cells the transcription factors can either induce or suppress a variety of cellular processes. Moreover, it was shown that SOX proteins are involved in DNA reparation processes.

HXA3, hepoxilin A3; in the absence of cytosol glutathione peroxidase is transformed into 12S-hydroperoxyeicosatetraenoic acid on biologically active epoxides. The process is catalyzed by lipoxygenase.

HLX is expressed in mesodermal tissues in embryogenesis. Especially high expression was observed in

visceral mesenchyme, skeletal myoblasts and limb mesenchyme[6].

Gata 3 regulates the expression of genes, which are involved in cell growth control, immune functions and adipogenesis inhibition.

Table 1. Confident differences of transcription factor localization prior to the first group genes.

Collection	P <0.05	P <0.05,r-skan >9	P <0.01
Promoter_Hocomoco (s1000)	PAX6, ALX1, PITX2, OTX2, CRX, CDC5L, PO3F2	FOXA1, CDX2	SOX5, FOXO1,EVI1, FOXO3
Promoter_Uniprobe (s1000)	GATA6, BARX2, HXB6, MEOX1, SH SOX8, HXD3, HXA6, DLX1, ARI5A, PITX2, HXA10, HXA7, DLX3, HXA9, BARH1, SOX15, HXB7, HXC8, SOX13, HXA2, HDX, PDX1, SP100, SOX30, GATA3, PO6F1, HXA1	TBP, TF7L1, LMX1A	DLX2, ARH2, DBX1
CRM_Hocomoco (s5000)	SOX9, ARI3A, NFYA, ATF6A, CDX1, ZEP1, SOX5, NFYB, SOX2	EVI1, MBD2, MEF2C, MEF2A	-
CRM_Uniprobe (s5000)	SOX15, SOX30, ZN187, HXD13, PROP1, SOX18, RX, SRY, SOX7, SOX14, HXD1,DLX4,PHX2B, PAX6,HXD10,HBP1,OTP	PO3F2, PO2F1, SOX12	SOX5, SOX13, SOX17, DBX1, HLX-

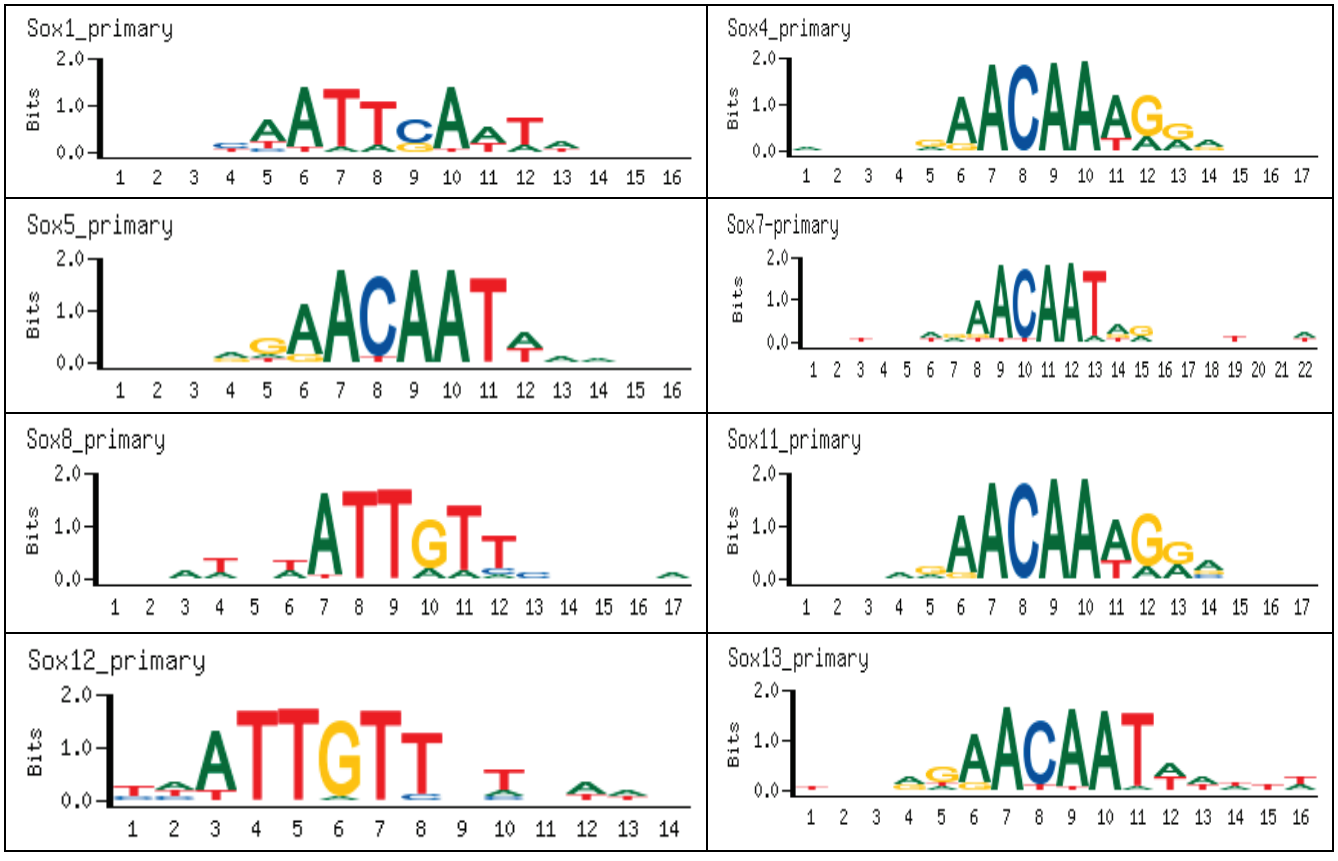

Figure 1. Consensus sequences of the SOX transcription factors SOX [(A/T) (A/T) CAA(A/T)G[5].

SP100 is a multifunctional protein factor often found in brain

Therefore, we have found out that primary treatment with hyperbaric oxygenation (0.2 MPa, 1 h) results in

up-regulation of a set of transcription factors, which are associated with the studied genes and their promoters.

The obtained data support the idea of specific and possibly cooperative interaction of transcription factors in both promoters and *cis*-regulatory elements. The mechanism of simultaneous gene expression induced by oxidative stress is not yet clear. It is suggested that the neighbor sequences of non-coding DNA may be involved in these processes. Our data allow us to conclude that non-coding DNA somehow participates in gene-gene interactions. The detailed study of all functions of non-coding DNA is considered to be a challenging task for biological science.

4. ACKNOWLEDGEMENTS

This work was supported by the Federal Targeted Program by The Ministry of education and science of Russian Federation, state contract no. 14.740.11.0006

REFERENCES

[1] The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome.- Nature (2012) , Vol. 489: p. 57-74.

[2] EP Guskov, TP Shkurat, EI Shimanskaja, SI Guskova Genetic effects of hyperbaric oxygen therapy// Mutation Research/Genetic ToxicologyVolume 241, Issue 4, August 1990, Pages 341–347.

[3] EP. Guskov, EV Mashkina, NI Belichenko, TV Varduni, G I Volosovtsova, IO Pokudina, GE Guskov and TP Shkurat Mutation processes in oxidative stress preadapted animals\\ Russian Journal of Genetics: Applied Research Volume 1, Number 2 (2011), 112-118.

[4] EP. Guskov, EV Mashkina, NI Belichenko, TV Varduni, G I Volosovtsova, IO Pokudina, GE Guskov and TP Shkurat Mutation processes in oxidative stress preadapted animals // Ecological Genetics. —2009. — V.VII, № 1. —pp.40-47.

[5] Guimont P, Grondin F, Dubois C: Sox9-dependent transcriptional regulation of the proprotein convertase furin. American journal of physiology. Cell physiology 2007, 293:C172-183

[6] Allen, J.D., T. Lints, N.A. Jenkins, N.G. Copeland, A. Strasser,R.P. Harvey, and J.A. Adams. 1991. Novel murine homeo box gene on chromosome 1 expressed in specific hematopoietic lineages and during embryogenesis. Genes & Dev.5: 509-520.

[7] T. P. Shkurat, A. V. Shestopalov, G. E. Guskov, V. N. Prokofiev, A. I. Butenko, T. V. Belik, M. Yu. Bibov, O. V. Lyangasova, M. A. Shkurat, E. V. Mashkina, S. S. Mandzhieva, E. V. Butenko, I. N. Sevastianova, P. Ryzhkov, E. M. Vechkanov Study of novel mechanisms of brain preadaptation to oxidative stress using DNA microarray // Problems of contemporary science and education. – 2011. – № 5; URL: www.science-education.ru/99-4888 .

